Longitudinally-Invariant $k_{\perp}$ -Clustering Algorithms for Hadron-Hadron Collisions*

S. Catani[†]

Theory Division, CERN, CH-1211 Geneva 23, Switzerland

Yu.L. Dokshitzer

Nuclear Physics Institute, Gatchina, St. Petersburg 188350, Russia and
Department of Theoretical Physics, University of Lund,
Sölvegatan 14A, S-22362 Lund, Sweden

M.H. Seymour

Department of Theoretical Physics, University of Lund,
Sölvegatan 14A, S-22362 Lund, Sweden

and

B.R. Webber[‡]

Theory Division, CERN, CH-1211 Geneva 23, Switzerland

Abstract

We propose a version of the QCD-motivated ' $k_{\perp}$ ' jet-clustering algorithm for hadron-hadron collisions which is invariant under boosts along the beam directions. This leads to improved factorization properties and closer correspondence to experimental practice at hadron colliders. We examine alternative definitions of the resolution variables and cluster recombination scheme, and show that the algorithm can be implemented efficiently on a computer to provide a full clustering history of each event. Using simulated data at $\sqrt{s} = 1.8$ TeV, we study the effects of calorimeter segmentation, hadronization and the soft underlying event, and compare the results with those obtained using a conventional cone-type algorithm.

CERN-TH.6775/93

LU TP 93-2

February 1993

*Research supported in part by the Italian Ministero dell'Università e Ricerca Scientifica.

[†]On leave of absence from INFN, Sezione di Firenze, Italy.

[‡]On leave of absence from Cavendish Laboratory, University of Cambridge, UK.

1. Introduction

The structure of hadronic final states in high energy collisions may be described in a simple way in terms of jet characteristics. Nowadays, jet cross section data are used both for precise quantitative tests of QCD (measurement of the QCD coupling α_S and scale $\Lambda_{\overline{\text{MS}}}$, studies of QCD coherence, etc.), and for looking for breakdown of the standard model and new physics. Although the intuitive definition of a jet as ‘a large amount of hadronic energy in a small angular region’ is sufficient to account for many qualitative features of hadron production, any detailed quantitative analysis requires a precise jet definition. One needs a jet algorithm able to specify a jet configuration unambiguously, both in theoretical calculations and in experimental practice, starting from the hadrons detected in the final state.

At present there are essentially two classes of jet algorithms in use [1]: cone-type algorithms ‘à la Stermann-Weinberg’ [2] and clustering algorithms of the type firstly introduced by the JADE collaboration [3]. In cone-type algorithms jets are defined by maximizing the amount of energy which can be covered by cones of fixed size, whilst in clustering algorithms particles are assigned to jets iteratively according to whether a given energy-angle resolution variable y_{ij} exceeds a fixed resolution parameter y_{cut} . In our opinion, clustering algorithms are more precisely defined because they do not suffer from ambiguities related to dealing with overlapping cones in multijet events. It is surely true that clustering algorithms are more suitable for processes, like e^+e^- -annihilation, with no hadrons in the initial state.

Recently clustering algorithms for e^+e^- -annihilation have been reconsidered and a new jet resolution variable, namely the minimal relative transverse momentum between particles, has been introduced [4-7]. The corresponding clustering algorithm, known as the ‘ $k_\perp$ ’- or ‘Durham’ algorithm, has been widely investigated both from the theoretical [4-8] and experimental [9,10] side. The use of the $k_\perp$ -resolution variable is suggested by the present understanding [11] of the perturbative QCD evolution of partonic systems and offers several phenomenological and theoretical advantages over the invariant mass resolution of the original JADE algorithm. In fact, the use of a resolution variable of invariant mass type induces soft particle kinematic correlations [4-6] which can lead to a non-intuitive classification of events and to unnatural assignments of particles to jets. The $k_\perp$ -variable avoids such kinematic correlations, thereby reducing hadronization corrections and simplifying the calculation of higher-order logarithmic contributions in perturbative QCD.

Motivated by these considerations in e^+e^- jet physics, in a recent paper [12] we proposed a non-trivial generalization of the $k_\perp$ -algorithm to deep inelastic lepton-hadron scattering (DIS). By ‘non-trivial’ we mean that the $k_\perp$ -algorithm for DIS is not a simple transplantation of the e^+e^- -algorithm to a different environment involving also initial-state hadrons. The presence of colliding hadrons in the initial state poses the new problem of dealing with the hadron remnants and isolating the latter from high- $p_\perp$ jets produced by hard scattering of partons. From the viewpoint of perturbative QCD this means that the jet algorithm has to ensure the factorization of initial-state radiation into universal (process independent) structure functions of the hadrons. Therefore the $k_\perp$ -algorithm for DIS combines the use of the $k_\perp$ -resolution variable for e^+e^- with an improved clustering procedure able to deal with the underlying soft event accompanying the fragmentation of the initial-state hadron.

In spite of the fact that jets in $p\bar{p}$ -collisions [13-15] are commonly defined in terms of cone algorithms, and that an agreed cone definition was achieved at the 1990 Snowmass Workshop [16], in the same letter [12] we sought to provoke the $p\bar{p}$ -collider community by proposing the same $k_{\perp}$ type of clustering algorithm also for hadron-hadron collisions. In this paper we amplify that proposal, improving its theoretical and phenomenological properties and showing that it does indeed represent a practical method of analyzing the complex final states encountered at hadron colliders.

First of all, in Sect. 2, we review the properties required of a jet definition in hadronic collisions, in order for it to be useful for comparing perturbative calculations with experimental data. Next, in Sect. 3, we concentrate on clustering algorithms of the $k_{\perp}$ type. We start with a general discussion of the clustering procedure, which has two distinct phases in hadronic collisions: one to define the hard subprocess and one to resolve in more detail the jet structure of the hard final state. The clustering algorithm itself has two components: a jet resolution variable and a recombination scheme. Even within the class of algorithms of $k_{\perp}$ type, one has some freedom in the definition of both components. Only the soft and collinear limits are crucial for consistency with jet dynamics in QCD. We consider alternative extrapolations away from these limits for both the resolution variable and the recombination scheme. In particular, we introduce variables and schemes that are invariant under boosts along the beam directions. In Sect. 4 we show that such ‘longitudinally-invariant’ algorithms have improved properties with respect to factorization of initial-state radiation, which simplifies the perturbative calculation of multi-jet cross sections.

Readers who are more interested in the applications of longitudinally-invariant $k_{\perp}$ -type algorithms to hadron-hadron collisions than in their theoretical aspects may begin at Sect. 5, where we explain the computer implementation of the algorithm that we have adopted for numerical studies. Results based on simulated data at $\sqrt{s} = 1.8$ TeV from the Monte Carlo program HERWIG are described in Sect. 6. We examine in particular the difference between jet properties evaluated by applying the algorithm at the parton, hadron and calorimeter-cell levels. In Sect. 7 we compare these results with those obtained using the more conventional ‘cone’ type of algorithm, while in Sect. 8 we compare the jet structure of the hard final state with that seen in e^+e^- annihilation using the original $k_{\perp}$ -algorithm of Ref. [4]. Finally, in Sect. 9, we draw some conclusions and discuss other applications of the algorithm, some of which have already been reported in Ref. [17].

2. Jet definition in hadron-hadron collisions

In the case of hadron collisions the *jet definition* has to fulfil the requirements of being

- i) simple to use in experimental analyses
- ii) simple to use in theoretical calculations
- iii) infrared and collinear safe
- iv) subject to small hadronization corrections
- v) able to factorize initial-state collinear singularities into universal distributions

- vi) not strongly affected by contamination from hadron remnants and the underlying soft event.

Requirements i) and ii) are self-evident. The others follow from the motivation of comparing data with theory, namely QCD. So far, we are able to perform QCD calculations, using perturbation theory, only for the region of small distances in which high energy collisions produce partons (quarks and gluons). At larger distances these produced partons are confined by the colour force field and are forced to dress themselves up into colourless hadrons. Although the hadronization process is certainly not perturbative, the preconfinement (or local parton hadron duality) [18] and factorization [19] properties of QCD imply that the hadronic flows follows the partonic flow quite closely, with transfers of momentum and other quantum numbers that are local in phase space and involve scales of the order of the hadronic scale, around 1 GeV. Thus provided jet observables are constructed to be insensitive to long distance physics, in particular to small values of the parton masses, perturbative predictions are expected to be reliable to this sort of precision.

For e^+e^- collisions this is achieved by the property iii) above, i.e. jet cross sections at parton level must be finite order by order in perturbation theory in the limiting case of massless final-state partons. Obviously, hadronization corrections may still sizeably affect jet algorithms satisfying the property iii). Therefore one should try to estimate hadronization effects by using some model, and then to minimize them by choosing an appropriate jet definition.

In the case of collisions involving initial state hadrons there are new features not present in e^+e^- annihilation. Most importantly, the centre-of-mass energy no longer controls the hardness of the process: the high- $p_\perp$ scattering subprocess has to be properly defined and separated from the underlying soft event. Therefore the jet definition for hadron collisions has to fulfil the extra requirements v) and vi). In particular, the initial state collinear singularities present in the QCD calculation at parton level have to be removed in order for the hard scattering part to be computable in perturbation theory. This is achieved by factorizing these singular contributions in a universal (process independent) way and associating them with hadron structure functions measured in inclusive processes such as DIS. Once the factorization procedure has been carried out, jet cross sections are computable in perturbation theory up to higher-twist contributions, i.e. contributions which are suppressed by powers of some hadronic scale over a properly defined hard scattering (factorization) scale. These power corrections can in turn be minimized by optimizing the jet definition so as to fulfil the condition vi).

The requirements i)-vi) are implemented in a *jet clustering algorithm* by defining

j_1) a test variable (energy-angle resolution) d_{kl}

j_2) a recombination procedure

j_3) a hard scattering scale.

The test variable d_{kl} is needed in order to specify whether or not two hadrons h_k, h_l belong to the same cluster, whilst the recombination procedure tells us how cluster resolution variables are related to those of the hadrons belonging to them. The introduction of a hard

scattering scale is necessary only in the case of hadron collisions. For e^+e^- -annihilation, no factorization of hard and soft subprocesses has to be performed and the hard scattering scale coincides with the centre-of-mass energy.

3. $k_\perp$ -type clustering algorithms

In this section we set out the details of our proposed [12] jet algorithm for hadron-hadron collisions. It is based on two main points: the jet algorithm is a clustering algorithm and the jet resolution variable is of relative transverse momentum type.

The first point follows from our opinion that a clustering procedure is preferable since it allows one an unambiguous and exhaustive assignment of hadrons to jets using the *same* procedure both in theoretical calculations and in experimental practice. This avoids some practical disadvantages of cone algorithms related to the difficulties in finding the optimal configuration of cones and in standardizing the treatment to be adopted for overlapping jets [16,20].

The use of the $k_\perp$ -resolution variable follows from the theoretical prejudice (based on present understanding of perturbative QCD) that QCD jets are not sprays of hadrons confined in cones of fixed angle. Soft hadrons produced *coherently* by the fragmentation of hard partons should instead be naturally assigned to the jet of the hard parton nearest in angle independently of the actual value of its angular distance [11,21]. In other words, our jets have an effective radius depending on the hardness of the jet itself and on the colour flow of the hard subprocess.

The $k_\perp$ -algorithm for hadron-hadron collisions is defined in the following subsections. In Sect. 3.1 we describe the clustering procedure already introduced in Ref. [12] and discuss in detail the issue of defining the hardness scale. Alternative prescriptions for introducing resolution variables of transverse momentum type are considered in Sect. 3.2, in order to simplify the practical implementation of the algorithm and improve factorization properties. In Sect. 3.3 several recombination procedures are presented and their features are discussed. Finally, Sect. 3.4 deals with the relationship between event shape variables and jet cross sections in different formulations of the algorithm.

3.1 The clustering procedure

Let us consider the hadron-hadron scattering process

$$a(p) + b(\bar{p}) \rightarrow h_1(p_1) + \cdots + h_n(p_n) + X \quad . \quad (1)$$

Here $a(p), b(\bar{p})$ are the initial-state hadrons, $h_k(p_k)$ ($k = 1, \dots, n$) denotes any hadron detected in the final state and X stands for hadrons with three-momenta parallel to either $\mathbf{p}$ or $\bar{\mathbf{p}}$.

The $k_\perp$ -algorithm is defined in the *centre-of-mass frame* of the colliding hadrons

$$p = E(1, \mathbf{0}, 1) \ , \ \bar{p} = E(1, \mathbf{0}, -1) \ , \ S = 2p\bar{p} = 4E^2 \ . \quad (2)$$

We denote by E_k the energy of the final-state hadron h_k , by θ_{ij} the angle between the momenta of hadrons h_i and h_j , and by $\theta_{kB} = \min(\theta_{ka}, \theta_{kb})$ the angle of the hadron momentum p_k with respect to the nearer initial-state hadron direction, all in the centre-of-mass frame

(2). Jets are defined by a two-step clustering procedure which amounts to A) pre-clustering of hadrons into ‘beam jets’ and ‘hard final state jets’ and B) resolving the structure of the hard final state jets into sub-jets.

A) Pre-clustering of hadrons.

- $a_1)$ For every final state hadron h_k and for every pair h_k, h_l one computes the corresponding value of the resolution variables d_{kB} and d_{kl} . The precise definition of these variables is discussed in Sect. 3.2. For the moment the only property we need is that in the small angle limit they reduce respectively to the transverse momentum $k_{\perp kB}$ of the hadron with respect to the beam direction and to the minimal relative transverse momentum $k_{\perp kl}$ of one hadron relative to the other

$$\begin{aligned} d_{kB} &\simeq E_k^2 \theta_{kB}^2 \simeq k_{\perp kB}^2, & \text{for } \theta_{kB} \rightarrow 0 \\ d_{kl} &\simeq \min(E_k^2, E_l^2) \theta_{kl}^2 \simeq k_{\perp kl}^2, & \text{for } \theta_{kl} \rightarrow 0 \end{aligned} \quad (3)$$

- $a_2)$ One then considers the smallest value among $\{d_{kB}, d_{kl}\}$. If d_{ij} is the smallest value, h_i and h_j have to be combined into a single cluster (‘pseudoparticle’) with momentum $p_{(ij)}$ according to a recombination prescription to be specified in Sect. 3.3. If d_{iB} is the smallest value, the hadron h_i has to be included in the ‘beam jets’.
- $a_3)$ This procedure has to be repeated from step $a_1)$ for all particles and pseudoparticles not included in the beam jets, until all objects (particles and/or pseudoparticles) have d_{kl}, d_{kB} larger than some stopping parameter d_{cut} . One ends up with the beam jets[§] and objects which we call *hard final state jets*.

The stopping parameter of the clustering procedure, d_{cut} ($\Lambda^2 \ll d_{\text{cut}} \leq S$, where Λ is the QCD scale and $\sqrt{S}$ the centre-of-mass energy), defines the hard scale of the process. It can be introduced in two different ways. One can fix its value $d_{\text{cut}} = E_{\text{cut}}^2$ ‘a priori’ for all events[¶] or alternatively one can determine $d_{\text{cut}} = E_{tn}^2$ on an event-by-event basis as the minimum value of $\{d_{kl}, d_{kB}\}$ when the given event contains a fixed number n of hard final state jets. In other words, one can stop the clustering procedure when a given number n of clusters have been reconstructed and identify d_{cut} with the corresponding minimum value among $\{d_{kl}, d_{kB}\}$. In the former case E_{cut} defines the (arbitrary) boundary between the low- $p_{\perp}$ scattering and hard scattering subprocesses, much as the resolution parameter y_{cut} defines the resolution scale for jets in e^+e^- -collisions [3]. In the latter case the corresponding E_{tn} can be regarded as a shape variable for the n -jet inclusive cross section, analogous to the generalizations y_n of the shape variable y_3 [22] in e^+e^- -annihilation. In Sect. 3.4, we discuss in more detail the relationship between these two ways of regarding the stopping parameter.

The pre-clustering procedure described so far allows one to identify low- $p_{\perp}$ scattering fragments and include them in the beam jets, thus factorizing the hard scattering subprocess. Once the subprocess has been defined, we are in the same position as for e^+e^-

[§]Note that undetected hadrons, denoted by X in (1), have vanishing transverse momenta with respect to the beams. Therefore they are automatically and consistently assigned to the beam jets.

[¶]This definition is that originally introduced in Ref. [12].

annihilation and we can proceed to resolve jet structures in the same way, considering only the particles assigned to hard final state jets. Let us recall the procedure for the sake of completeness.

B) Resolving event structure into sub-jets.

b_1) Define a resolution parameter $y_{\text{cut}} = Q_0^2/d_{\text{cut}} < 1$.

b_2) For any hadron in a hard final state jet^{||} consider the rescaled resolution variable y_{kl}

$$y_{kl} = d_{kl}/d_{\text{cut}} \quad . \quad (4)$$

b_3) If y_{ij} is the smallest value of y_{kl} computed in b_2) and $y_{kl} < y_{\text{cut}}$, combine (p_i, p_j) into a single cluster (pseudoparticle) $p_{(ij)}$ according to the recombination prescription.

b_4) Repeat this procedure from step b_2) until all pairs of objects (particles and/or clusters) have $y_{kl} > y_{\text{cut}}$. Whatever objects remain at this stage are called *sub-jets*.

Note that, as for the pre-clustering procedure A), one can avoid the definition of a fixed resolution parameter y_{cut} by introducing the latter implicitly as an n -jet shape variable y_n [22]. We discuss this option further in Sect. 3.4.

The above definition provides a QCD-motivated implementation of the jet requirements i) - vi). In particular, the algorithm is infrared and collinear safe because according to Eq. (3) the resolution variable d_{kl} vanishes for vanishing energies and relative angles. The fact that $d_{kB} \rightarrow 0$ when $\theta_{kB} \rightarrow 0$ and that hadrons included in the beam jets do not influence the recombination procedure ensure the factorization of initial state collinear singularities. As discussed in Sect. 4, the precise form of the factorization depends on the extrapolation of the resolution variable (3) away from the small-angle limit and on the recombination scheme.

The two-step clustering procedure allows us to include the hadron remnants in the beam jets while associating with jets the coherent (small-relative-angle) soft radiation accompanying the hard parton scattering. The final results are sub-jets with relative transverse momenta $\Delta k_{\perp}$ where $d_{\text{cut}} > (\Delta k_{\perp})^2 \geq Q_0^2$, produced within hard final state jets with large transverse momenta $p_{\perp}^2 \geq d_{\text{cut}}$ relative to each other and the incoming hadrons. The parameters Q_0^2 and d_{cut} fix respectively the fragmentation and hardness scales of the jets. If we set $y_{\text{cut}} = 1$ then the two scales are identified and the clustering procedure B) is not carried out. In e^+e^- annihilation we have instead the opposite situation: the pre-clustering stage A) is suppressed by setting $d_{\text{cut}} = S$. Then every event has two hard final state jets and the resolution of multijet substructure by the $k_{\perp}$ -algorithm occurs entirely at stage B).

Even once we have defined the hard final state jets in stage A), it is important to realise that we cannot treat them as independent. Sub-jets resolved during stage B) must be attributed to coherent emission by the system of hard final state jets, and not to independent emission from any particular one of them.

^{||}Note that hadrons already assigned to the beam jets via pre-clustering do not participate in this clustering procedure.

We have presented the clustering procedure as a two-stage process in order to make clear that the jet definition involves two logically distinct steps. First the hard scattering event has to be defined, and then one can proceed to resolve the structure of the hard final state as for e^+e^- annihilation. As a matter of fact, one could in principle implement the step B) without making use of the definition (4) for y_{kl} , replacing it by any other resolution variable commonly used in e^+e^- algorithms. Following our definition, on the other hand, the whole clustering procedure can actually be carried out in one pass (cf. Sect. 5).

3.2 The jet resolution variable

The small-angle form (3) of the resolution variables d_{kB} and d_{kl} controls the main features of jet production in the soft and collinear limits, ensuring that the coherence properties of QCD [11,21] are properly taken into account by the jet algorithm. Away from the collinear limit, different extrapolations of Eq. (3) are allowed. In choosing an extrapolation, the main constraints we have to keep in mind are that d_{kB} and d_{kl} have to be increasing functions of the angles and that their functional dependence is responsible for the detailed form of the factorization of structure functions and hard cross section contributions.

In Ref. [12] we introduced the following definition of the jet resolution variables

$$\begin{aligned} d_{kB} &= 2E_k^2 (1 - \cos \theta_{kB}) , \\ d_{kl} &= 2 \min(E_k^2, E_l^2) (1 - \cos \theta_{kl}) . \end{aligned} \quad (5)$$

This definition was motivated by a desire to use exactly the same type of variables as those presently used in the $k_\perp$ -algorithm for e^+e^- annihilation. We still recommend the definition (5) for jets in DIS (where the $k_\perp$ -algorithm should be implemented in the Breit frame [12]). For the case of hadron-hadron collisions, however, we suggest an alternative form in order to improve factorization properties, using kinematic variables closer to those used by cone algorithms and in experimental practice at hadron colliders.

The basic observation is that in the small-angle limit one has

$$\begin{aligned} E_k^2 \theta_{kB}^2 &\simeq E_k^2 \sin^2 \theta_k = p_{tk}^2 , \\ E_k^2 \theta_{kl}^2 &\simeq E_k^2 [(\theta_k - \theta_l)^2 + (\phi_k - \phi_l)^2 \sin^2 \theta_k] \simeq p_{tk}^2 [(\eta_k - \eta_l)^2 + (\phi_k - \phi_l)^2] . \end{aligned} \quad (6)$$

where $(p_{tk}, \theta_k, \phi_k)$ are the cylindrical coordinates of the momentum p_k with respect to the incoming hadron momentum p , and $\eta_k = -\ln \tan(\theta_k/2)$ is the corresponding pseudorapidity. Note that we are treating final-state hadrons as massless particles so that transverse momentum and transverse energy are equivalent, as are rapidity and pseudorapidity, at the hadronic level. Thus, comparing Eqs. (6) and (3), we see that the resolution variables d_{kB} and d_{kl} admit a *longitudinal-boost-invariant* extrapolation to large angles in the form

$$\begin{aligned} d_{kB} &= p_{tk}^2 , \\ d_{kl} &= \min(p_{tk}^2, p_{tl}^2) R_{kl}^2 , \end{aligned} \quad (7)$$

where the *generalized radius* R_{kl}^2 is given by

$$R_{kl}^2 = f(\eta_k - \eta_l, \phi_k - \phi_l) , \quad (8)$$

f being any monotonic function with the small-angle behaviour

$$f(\eta_k - \eta_l, \phi_k - \phi_l) \simeq (\eta_k - \eta_l)^2 + (\phi_k - \phi_l)^2, \quad \text{for } |\eta_k - \eta_l|, |\phi_k - \phi_l| \rightarrow 0. \quad (9)$$

The simplest definition of the generalized radius R_{kl} which makes the $k_\perp$ -algorithm invariant under longitudinal boosts is

$$R_{kl}^2 = (\eta_k - \eta_l)^2 + (\phi_k - \phi_l)^2. \quad (10)$$

This definition corresponds exactly to that used by cone algorithms and it is very simple to implement experimentally.

The alternative definition

$$R_{kl}^2 = 2 [\cosh(\eta_k - \eta_l) - \cos(\phi_k - \phi_l)] \quad (11)$$

may instead look more attractive from a theoretical point of view because multiparton QCD matrix elements are built up by eikonal factors of the type

$$\frac{p_k \cdot p_l}{(p_k \cdot p_i)(p_l \cdot p_i)} = \frac{\cosh(\eta_k - \eta_l) - \cos(\phi_k - \phi_l)}{p_{ti}^2 [\cosh(\eta_k - \eta_i) - \cos(\phi_k - \phi_i)] [\cosh(\eta_l - \eta_i) - \cos(\phi_l - \phi_i)]}. \quad (12)$$

3.3 Recombination schemes

Introducing longitudinal-boost-invariant resolution variables is not enough to guarantee the same property for the full clustering procedure: the recombination scheme has to be considered as well.

For historical reasons several different recombination prescriptions have been introduced for clustering algorithms in e^+e^- -annihilation [3,23]. They differ among themselves in the way in which the resolution variables are defined for pseudoparticles, i.e. after two particles have been merged into a single cluster. The point is that it was noticed in the original JADE algorithm (in which the invariant mass is used as resolution variable) that the jet properties and in particular the size of hadronization corrections were strongly affected by the recombination procedure. In the case of the $k_\perp$ -algorithm for e^+e^- collisions that is no longer true and hence the simplest recombination prescription (called the E -scheme) is used at present. It amounts to computing the resolution variable for a pseudoparticle (p_i, p_j) as if it was a particle with four-momentum $p_{(ij)} = p_i + p_j$.

Coming back to the $k_\perp$ -algorithm for hadron-hadron collisions, we can therefore adopt the same E -recombination scheme. However, it is worth considering other recombination prescriptions. In fact, as discussed in the previous subsection, there are alternative and equally good (i.e. longitudinal-boost-invariant) definitions of $k_\perp$ -type resolution variables and in that respect there is no reason to think of the E -scheme as the most ‘natural’. Moreover, some recombination prescriptions may be more suitable to implement in computer codes for experimental analyses. In the following we present three different recombination schemes and discuss their main features.

s₁) The covariant E -scheme.

This scheme corresponds to considering the pseudoparticle (p_i, p_j) as a real particle with momentum $p_{(ij)} = p_i + p_j$. More precisely, once the hadrons h_i and h_j have been merged into a cluster, the corresponding resolution variables are given by

$$\begin{aligned} d_{(ij)B} &= (\mathbf{p}_{ti} + \mathbf{p}_{tj})^2, \\ d_{(ij)k} &= \min\{(\mathbf{p}_{ti} + \mathbf{p}_{tj})^2, \mathbf{p}_{tk}^2\} R_{(ij)k}^2, \end{aligned} \quad (13)$$

where the generalized radius is

$$R_{(ij)k}^2 = f(y_{(ij)} - \eta_k, \phi_{(ij)} - \phi_k), \quad (14)$$

f being any function with the properties discusses in Sect. 3.2, e.g. that given by Eq. (10) or (11). Note that transverse momenta are added vectorially in (13). Note also that the pseudoparticle is not massless ($p_{(ij)}^2 \neq 0$). Therefore in computing the generalized radius (14) one has to use the true rapidity $y_{(ij)}$ of the momentum $p_i + p_j$ instead of the pseudorapidity, if the recombination procedure is to be longitudinal-boost-invariant.

s₂) The p_t -weighted scheme.

Here the functional form (7) of the resolution variables is used for both particles and pseudoparticles, and the pseudoparticle kinematic variables are obtained by p_t -average of the corresponding particle variables. Thus the transverse momentum, pseudorapidity and azimuth of the pseudoparticle (p_i, p_j) are defined as

$$\begin{aligned} p_{t(ij)} &= p_{ti} + p_{tj}, \\ \eta_{(ij)} &= \frac{p_{ti}\eta_i + p_{tj}\eta_j}{p_{t(ij)}}, \\ \phi_{(ij)} &= \frac{p_{ti}\phi_i + p_{tj}\phi_j}{p_{t(ij)}}. \end{aligned} \quad (15)$$

Note that the particle transverse momenta enter Eq. (15) as *scalar* quantities.

s₃) The monotonic p_t^2 -weighted scheme.

Once p_i and p_j have been merged into a single cluster, its resolution variables are defined by

$$\begin{aligned} d_{(ij)B} &= (p_{ti} + p_{tj})^2, \\ d_{(ij)k} &= \min\{(p_{ti} + p_{tj})^2, p_{tk}^2\} R_{(ij)k}^2, \end{aligned} \quad (16)$$

and the radius $R_{(ij)k}$ is obtained as follows

$$R_{(ij)k}^2 = \frac{p_{ti}^2 R_{ik}^2 + p_{tj}^2 R_{jk}^2}{p_{ti}^2 + p_{tj}^2}. \quad (17)$$

Further recombinations are then defined iteratively. Unlike the E - and p_t -weighted schemes, in this case the recombination procedure is *not associative*: the resolution variables of a cluster depend on the ordering in which particles have been merged into a cluster (for instance, $R_{((ij)k)l} \neq R_{((ik)j)l}$).

The covariant E -scheme could be slightly less convenient for practical implementation because at each step of the algorithm the four-momentum kinematics have to be reconstructed. However this scheme may be preferable for theoretical calculations if the definition (11) of the generalized radius is used.

The simplest recombination procedure, the p_t -weighted scheme, coincides with that used in cone algorithms. It corresponds to the small-angle version of the E -scheme prescription and is quite convenient for use in association with the definition (10) of the generalized radius.

We refer to the p_t^2 -weighted scheme s_3) as *monotonic* because, although it may look quite cumbersome, it is actually the simplest one (for practical implementation) that leads to monotonic resolution variables. By this we mean that after each recombination the values of the resolution variables never decrease, i.e. $\min\{d_{kB}, d_{kl}\} \leq \min\{d'_{kB}, d'_{kl}\}$ where d and d' are respectively the computed resolution variables before and after the recombination. In fact, it is trivial to check that in this scheme after having merged (p_i, p_j) into a single cluster one has $d_{(ij)B}, d_{(ij)k} \geq d_{ij}$. We discuss the implications of this property in Sect. 3.4.

Note that in the small-angle limit Eqs. (16,17) are equivalent to

$$\begin{aligned} p_{t(ij)} &= p_{ti} + p_{tj} \\ \eta_{(ij)} &\simeq \frac{p_{ti}^2 \eta_i + p_{tj}^2 \eta_j}{p_{ti}^2 + p_{tj}^2} \\ \phi_{(ij)} &\simeq \frac{p_{ti}^2 \phi_i + p_{tj}^2 \phi_j}{p_{ti}^2 + p_{tj}^2}. \end{aligned} \tag{18}$$

It may seem unnatural to use a scheme which does not reduce to vector addition of momenta near the collinear limit, but one should note that one of the aims is to minimize hadronization corrections. After the effects of hadronization have been incorporated, particles tend to acquire additional ‘random’ transverse momenta, which perturb the directions of soft particles more than hard particles. This means that schemes like the monotonic p_t^2 -weighted one, which, as shown by Eqs. (18), gives additional weight to the directions of hard particles when calculating the direction of the jet, may be less perturbed by hadronization than those which give equal weight to all particles, as in vector addition.

We conclude this section by noting that the recombination issue addressed here concerns only the clustering procedure and should not be confused with the definition of jet characteristics. After particles have been assigned to jets, jet observables can be defined from any infrared and collinear safe combination of the momenta of the particles belonging to a given jet.

3.4 Monotonicity and event shape variables

Although the literature on jet algorithms is by now quite extensive, the issue of monotonicity of resolution variables does not seem to have been discussed before. The issue is important because in a monotonic algorithm the number of jets can only change by at most ± 1 when the stopping parameter is varied by a small amount, whereas in a non-monotonic one it can change by several units. Correspondingly, one would expect, as we shall illustrate in Sect. 8, that a monotonic procedure should lead to smaller, more stable

hadronization corrections, because in a non-monotonic scheme a small shift in kinematics due to hadronization can give rise to a large jump in the jet multiplicity.

The two definitions of the stopping parameter d_{cut} introduced in Sect. 3.1 are in fact closely connected in the case of a monotonic algorithm. If we denote by $\sigma^{n\text{-jet}}(E_{\text{cut}})$ the n -jet exclusive cross section obtained for a given a priori value of $d_{\text{cut}} = E_{\text{cut}}^2$, and by $d\sigma/dE_{tn}$ the inclusive cross section giving the distribution of the shape variable E_{tn} (E_{tn}^2 being the value of the smallest resolution variable when the event has n hard final state jets), then for a monotonic algorithm we find

$$\sigma^{n\text{-jet}}(E_{\text{cut}}) = \int_{E_{\text{cut}}}^{\infty} dE_{tn} \frac{d\sigma}{dE_{tn}} - \int_{E_{\text{cut}}}^{\infty} dE_{t,n+1} \frac{d\sigma}{dE_{t,n+1}} . \quad (19)$$

Thus the distributions of the shape variables E_{tn} are sufficient in this case to enable us to compute the n -jet exclusive cross sections for any value of E_{cut} .

Many commonly used algorithms, including $k_{\perp}$ -type algorithms with the E - or p_t -weighted recombination schemes discussed above, are not strictly monotonic. In such cases, Eq. (19) is not precisely true as stated, although it will be a good approximation in practice if the degree of non-monotonicity is small. We can make it precise by modifying the definition of E_{tn} or $\sigma^{n\text{-jet}}(E_{\text{cut}})$, or both. The minimal modification is to re-define the shape variable E_{tn} for non-monotonic algorithms as follows. Denoting by d_m the smallest resolution variable when the event has m hard final state jets, we define

$$E_{tn}^2 = \max_{m \geq n} \{d_m\} . \quad (20)$$

Clearly this coincides with the usual definition for a monotonic algorithm, and is a monotonic function of n . In this case Eq. (19) is correct for $\sigma^{n\text{-jet}}(E_{\text{cut}})$ as defined by the clustering procedure in Sect. 3.1.

An alternative way to introduce a monotonic shape variable for a non-monotonic algorithm is to define

$$E_{tn}^2 = \min_{m \leq n} \{d_m\} . \quad (21)$$

In this definition, the clustering procedure has to be carried through to the end (a single cluster), and then “backed off” to the desired jet multiplicity. Correspondingly, the exclusive cross sections $\sigma^{n\text{-jet}}(E_{\text{cut}})$ now refer to the jet multiplicity n after the *last* merging in which d_n became greater than E_{cut}^2 , rather than the first. This definition might have some theoretical advantages if real and virtual contributions are computed separately with a fixed infrared and collinear cut-off Q_{min} , as is the case in Monte Carlo simulations or in some procedures used to perform fixed-order numerical calculations. Here the $\{d_m\}$ with $m > n$ are likely to be more Q_{min} -dependent than those with $m < n$. Obviously this dependence should cancel for $E_{tn} \gg Q_{\text{min}}$, but at less asymptotic scales the cutoff-dependence of the event shape (21) may be smaller.

The first definition (20) looks better for experimental purposes and for conventional (analytic) fixed-order calculations, because for any given number of final-state particles one has to perform a smaller number of recombinations. Furthermore it represents the least departure from current practice, and can easily be implemented in any existing clustering program.

It is clear that similar arguments apply to the second stage of the clustering algorithm, in which sub-jet structure is resolved by considering the resolution variables $y_{kl} = d_{kl}/d_{\text{cut}}$. Even the well-established $k_{\perp}$ -algorithm for e^+e^- annihilation, using the E -recombination scheme, is not strictly monotonic, and hence a redefinition of the event shape variables y_n along the lines of Eq. (20), viz.

$$y_n \rightarrow y'_n = \max_{m \geq n} \{y_m\} , \quad (22)$$

should be considered.

For most of the numerical studies presented in Sects. 6-8, the E -recombination scheme was used and the shape variable E_{tn} and jet cross sections $\sigma^{n\text{-jet}}(E_{\text{cut}})$ were defined in the conventional way, and not modified according to either of the above proposals. This means that the relation (19) does not hold precisely, but the amount of non-monotonicity is small in practice and hence it should remain a good approximation.

We should mention that in numerical studies we have found that it is possible for an algorithm to be “too monotonic”, i.e. that hadronization corrections become large when recombination typically produces a large increase in the affected resolution variables. This is the case, for example, if one adopts a ‘monotonic p_t -scheme’ analogous to the monotonic p_t^2 -scheme defined above. The optimal situation seems to be one in which the effect of recombination on the resolution variables is as smooth as possible (see Sect. 6 and 8).

4. Factorization properties and boost invariance

In addition to being infrared and collinear safe, any jet algorithm for processes involving incoming hadrons must satisfy factorization of initial-state collinear singularities. When factorization holds, physical hard-scattering cross sections can be calculated perturbatively in terms of partonic subprocess cross sections and universal, scale-dependent parton distributions inside the colliding hadrons.

A general n -jet cross section defined according to the $k_{\perp}$ type of algorithm satisfies the factorization formula

$$\begin{aligned} \sigma_{ab}^{(n)}(S, d_{\text{cut}}, y_{\text{cut}}, \{p_J\}) &= \sum_{ij} \int_0^1 dz_1 \int_0^1 dz_2 F_{i/a}(z_1, \mu_F) F_{j/b}(z_2, \mu_F) \\ &\cdot \hat{\sigma}_{ij}^{(n)}(z_1 \sqrt{S}, z_2 \sqrt{S}, d_{\text{cut}}, \mu_F, \mu_R, \alpha_S(\mu_R), y_{\text{cut}}, \{p_J\}) \quad , \quad (23) \end{aligned}$$

where $F_{i/a}$ and $F_{j/b}$ are the parton distributions of the incoming hadrons a and b , $\hat{\sigma}_{ij}^{(n)}$ represents the n -jet subprocess cross section for colliding partons of type i and j and jet momenta $\{p_J\}$, and μ_F , μ_R are arbitrary factorization and renormalization scales.

In proving Eq. (23) we can rely on the factorizability [19] of the total ij -subprocess hard scattering cross section at scale d_{cut} . We have to show that this property carries over to each individual n -jet cross section after the separation into multijet fractions. This will be the case if the jet algorithm can be formulated entirely in terms of the kinematic variables of the subprocess, without reference to the overall kinematics of the hadron-hadron collision.

The subprocess variables are the incoming parton momenta $q_1 = z_1 p$, $q_2 = z_2 \bar{p}$, and the outgoing momenta $\{p_k\}$. The jet resolution variables for the original form of the hadronic

$k_\perp$ -algorithm [12], given by Eq. (5), can be written in terms of these quantities and their corresponding hadronic centre-of-mass energies $\omega_1, \omega_2, \{E_k\}$ as follows (neglecting masses):

$$\begin{aligned} d_{kB} &= 2E_k \min(p_k \cdot q_1/\omega_1, p_k \cdot q_2/\omega_2) , \\ d_{kl} &= 2p_k \cdot p_l \min(E_k/E_l, E_l/E_k) . \end{aligned} \quad (24)$$

Thus in order to apply the original $k_\perp$ -algorithm we need the subprocess invariants and also the incoming parton energies $\omega_i = \frac{1}{2}z_i\sqrt{S}$, which accounts for the explicit dependence on these quantities in Eq. (23). Given ω_1 and ω_2 , the energies $\{E_k\}$ are expressible in terms of invariants as

$$E_k = \frac{1}{2}(p_k \cdot q_1/\omega_1 + p_k \cdot q_2/\omega_2) , \quad (25)$$

and hence no other non-invariants are required.

The separate dependence of the n -jet subprocess cross section on z_1 and z_2 in Eq. (23), while not affecting factorization, implies a non-invariance of this cross section with respect to longitudinal boosts of the subprocess centre-of-mass. Thinking of the process as a collision between beams of partons with a range of longitudinal momenta, we would prefer to define a cross section which is invariant under such boosts. This will be the case if the dependence on z_1 and z_2 occurs only via the invariant $z_1 z_2 S = 2q_1 \cdot q_2 = \hat{s}$, the subprocess c.m. energy. A jet algorithm of the type defined by Eqs. (7,8), being explicitly formulated in terms of longitudinal-boost-invariant quantities, clearly satisfies this requirement. Thus in this case the factorization formula (23) becomes

$$\begin{aligned} \sigma_{ab}^{(n)}(S, d_{\text{cut}}, y_{\text{cut}}, \{p_J\}) &= \sum_{ij} \int_0^1 dz_1 \int_0^1 dz_2 F_{i/a}(z_1, \mu_F) F_{j/b}(z_2, \mu_F) \\ &\cdot \hat{\sigma}_{ij}^{(n)}(z_1 z_2 S, d_{\text{cut}}, \mu_F, \mu_R, \alpha_S(\mu_R), y_{\text{cut}}, \{p_J\}) . \end{aligned} \quad (26)$$

There are several computational advantages of using a longitudinal-boost-invariant algorithm of the type leading to Eq. (26). The Born cross section for two hard final-state jets corresponds to the usual leading-order two-jet prediction with $d_{\text{cut}} = p_{T\text{min}}^2$. Furthermore the tools exist for a full next-to-leading-order calculation [14,24], after minor modifications to convert from a cone to a cluster algorithm. In addition, the resummation of large logarithms in the region of small d_{cut} and/or y_{cut} , along the lines of Refs. [4,12], remains possible.

5. Computer implementation

The $k_\perp$ type of algorithm described in Sect. 3 can be easily implemented in a computer program. The simplest approach is to associate with each pseudoparticle constructed in the clustering procedure a set of state variables from which the resolution variables can be computed. As emphasised in Sect. 3.3, these state variables need not necessarily be taken to describe the jet characteristics at the end of the procedure.

In the covariant E -scheme we need all four components of the four-momentum of each pseudoparticle in order to compute the resolution variables, for any definition (8) of the generalized radius R . It is then natural to associate these four-momenta directly with the jets formed in the clustering procedure.

In the p_t -weighted scheme, on the other hand, we need only the three scalar state variables (p_t, η, ϕ) to compute the resolution variables, for any definition of R . In this case, however, the state variables of a jet may differ considerably from the resultant transverse momentum, rapidity and azimuth obtained, for example, by summing the four-momenta of the particles in the jet.

In the monotonic p_t^2 -weighted scheme, for any cluster X of particles obtained by merging the sub-clusters Y and Z , we may define p_t^2 -weighted cluster attributes of the form

$$\langle w \rangle_X = \frac{p_{tY}^2 \langle w \rangle_Y + p_{tZ}^2 \langle w \rangle_Z}{p_{tY}^2 + p_{tZ}^2} \quad (27)$$

to be used as state variables in the resolution computation. The number of variables required now depends on the definition of the generalized radius. For the simple definition (10), the four variables $\{p_t, \langle \eta \rangle, \langle \phi \rangle, \langle \eta^2 + \phi^2 \rangle\}$ suffice to write

$$R_{XY}^2 = \langle \eta^2 + \phi^2 \rangle_X + \langle \eta^2 + \phi^2 \rangle_Y - 2 (\langle \eta \rangle_X \langle \eta \rangle_Y + \langle \phi \rangle_X \langle \phi \rangle_Y) , \quad (28)$$

whereas for the definition (11) the five-variable set $\{p_t, \langle \cosh \eta \rangle, \langle \sinh \eta \rangle, \langle \cos \phi \rangle, \langle \sin \phi \rangle\}$ may be used. Note that, with the definition (27), $\langle \eta^2 + \phi^2 \rangle \neq \langle \eta \rangle^2 + \langle \phi \rangle^2$, $\langle \cosh \eta \rangle^2 - \langle \sinh \eta \rangle^2 \neq 1$ in general, and similarly for $\langle \cos \phi \rangle^2 + \langle \sin \phi \rangle^2$. Here again, the state variables need not correspond in any direct way to the attributes of a jet deduced from its constituent particles.

In fact, for the monotonic scheme a more general treatment is made possible by storing the generalized radius between each pair of particles, R_{ij}^2 , in an array, and literally implementing Eq. (17) at each merging. This means that any definition (8) of R_{ij}^2 can be used, even one that cannot be described in terms of a finite number of state variables. It also allows generalizations of (17) such as replacing p_t^2 by p_t , or R^2 by R . In fact, we have found that none of these schemes perform as well as the three defined in Sect. 3.3 (see Sect. 6). Although the array requires a considerable amount of memory when the multiplicity is large, this is not a problem on modern workstations, and actually results in an improvement in speed relative to the state variable approach. With the exception of the more sophisticated definition of R^2 in Eq. (11), all the schemes actually take the same amount of time to cluster events.

For all the results reported in the following, the simplest form (10) for R has been used, together with the covariant E recombination scheme except where stated otherwise.

We have implemented the algorithm in **FORTRAN**, and the code is available from the authors**. The approach taken is very similar to that of earlier $e^+e^- k_\perp$ -clustering algorithms except that each of the functional units has been put in a separate program block. It is hoped that this will make the program easier to use, and to modify if necessary.

Clearly the algorithm takes N steps to merge all the jets (where N is the initial number of particles). At each of these steps $\sim N^2$ combinations of resolution variables d_{kl} must be checked to find the smallest. Although these are also tabulated in an array, and only those affected by each merging are recalculated, it turns out that searching this table for

**email address: mike@thep.lu.se

its smallest member is the slowest part of the algorithm, for the multiplicities expected at hadron colliders. We have exploited this fact to simplify the use of the program, by separating the clustering and analysis stages. The first stores a ‘merging history’ and only returns the simplest output information: the d_m value at which each merging takes place. Various other routines can then be called to give the information required for more sophisticated analyses, such as the number of jets at each of a number of d_{cut} values, the jet momenta at a given d_{cut} , or lists of which jets at one scale d_{cut} belong to which jets at a higher scale d'_{cut} . Each of these takes an amount of time $\sim N$, which is negligible relative to the clustering stage. In fact we find that the program takes about $0.2N^3 \mu\text{s}$ on a 10 MIP workstation.

The program allows any combination of collision type (ee, ep, or pp), angular variable ($2(1 - \cos \theta)$, $\Delta\eta^2 + \Delta\phi^2$ or $f(\Delta\eta, \Delta\phi)$) and recombination scheme (E , p_t or p_t^2), with or without a monotonic sequence of $\{d_{kB}, d_{kl}\}$ guaranteed by the generalization of Eq. (17). Of course several of these combinations make little sense: for example, any of the boost-invariant schemes in e^+e^- collisions.

6. Monte Carlo results

Having established the theoretical properties of a jet definition, the next most important feature is the size of hadronization corrections. In this section we use the HERWIG Monte Carlo event generator [25] to study the corrections due to hadronization, the underlying soft event, and the effect of using calorimeter cells rather than particle tracks.

The HERWIG event generator models the production of hadrons in hadron-hadron collisions through the combination of the lowest-order matrix elements for QCD $2 \rightarrow 2$ scattering, a parton shower algorithm for the initial and final states, and closely related phenomenological cluster models for the non-perturbative processes of hadronization and underlying soft event generation. The underlying event is taken to be a soft collision between the clusters containing the spectator partons from the incoming hadrons. This soft collision gives rise to a longitudinal distribution of clusters parametrized in the same way as a minimum-bias soft event at the corresponding centre-of-mass energy [26]. The default parameter set (which is tuned to data from e^+e^- annihilation) has been shown to give reasonable agreement with data from hadron-hadron collisions [27]. We generated 50000 events, for $p\bar{p}$ collisions with a total centre-of-mass energy of 1.8 TeV. All the parameters were left at their default values, for which the underlying soft event contributes a transverse energy of roughly 1 GeV per unit area in (η, ϕ) . The minimum transverse momentum of the hard scattering was set to 50 GeV, so events with two or more hard final state jets at $E_{\text{cut}} > 50$ GeV should be well described.

We have used a simulation of an idealized calorimeter [28] to estimate how robust the algorithm is against mismeasurements of particle direction. The coverage and segmentation used ($|\eta| < 5$, with 150 cells in η and 100 in ϕ) correspond roughly to those planned for detectors at the LHC and SSC, with an assumed perfect energy resolution. Thus the output of each cell is simply the sum of the transverse energies of all particles passing through it.

Unlike the jet analyses in e^+e^- annihilation, where jets are produced by final state radiation (i.e. according to our definition in this paper only two hard final state jets are produced) and the y_{kl} value of each merging is the most relevant quantity, there is a wide

variety of variables we could measure. We can focus our interest on the hard jets, or their sub-jet structure, and can choose whether or not to make a distinction between merging of two hard jets and of one hard jet with the beam jets. Here we just consider hard jets, and come back to sub-jets in a later section. We define the n -jet shape variable E_{tn} as discussed in Sect. 3.1. Thus, when the n th from last merging is with the beam E_{tn} is the p_t of the n th hardest jet, and when it is with another jet, E_{tn} is the relative $k_\perp$ of the pair.

In Fig. 1 we show the shape variable cross-sections $d\sigma/dE_{tn}$ for $n = 2, 3, 4$. In each case the curves correspond to: ‘partons’, i.e. all partons after initial-state and final-state parton showering; ‘hard hadrons’, i.e. including all hadrons which result from hadronization of the partons; ‘all hadrons’, i.e. those from the underlying event as well; and ‘calorimeter cells’, i.e. using the calorimeter simulator described above. In each case it can be seen that the spectrum from hard hadrons is very similar to that from partons, indicating a small average hadronization correction. The spectrum from all hadrons is also similar, but shifted up by about 3 GeV, which corresponds to the underlying event contribution in an effective jet cone of radius $R \sim 1$ (see Sect. 7). The calorimeter results are similar to the all-hadron results. Since the hard process was generated with $p_t > 50$ GeV, we have an approximate kinematic limit $E_{t1} + E_{t2} \gtrsim 100$ GeV after initial-state radiation has shifted the p_t of the hard process. Thus the shape variable cross-sections are actually the conditional ones, $d\sigma/dE_{tn}|_{E_{t1}+E_{t2}>100 \text{ GeV}}$, and the E_{t2} distribution in particular falls rapidly for $E_{t2} < 50$ GeV.

We next consider the event-to-event fluctuations in hadronization. Figure 2 shows correlations between the E_{t3} value found from the partons and those from the various sets of hadrons. Defining the fractional mismeasurement of E_{t3} ,

$$\Delta E_{t3} = \frac{E_{t3}(\text{hadrons}) - E_{t3}(\text{partons})}{E_{t3}(\text{hadrons}) + E_{t3}(\text{partons})}, \quad (29)$$

we find that for the hard hadrons ΔE_{t3} has a mean value of -0.014 and an r.m.s. of 0.088 , and that 95% of the points lie within $\Delta E_{t3} < 0.20$. This is entirely due to the hadronization, and can be compared to the similar quantity Δy_3 for the $k_\perp$ algorithm in e^+e^- . In Ref. [7] this was shown to have a mean value of 0.109 and an r.m.s. of 0.202 . The equivalent plot for all hadrons is shifted upwards due to particles from the underlying event entering the jets, and slightly broadened reflecting the additional uncertainty in reconstructing them. The mean and r.m.s. values of ΔE_{t3} for all hadrons and for calorimeter cells are listed in Table 1, for each of the three recombination schemes defined in Sect. 3.3.

We have studied these distributions, as well as those described later, using a wide variety of the options which are available in our program. The results can be briefly summarized as showing that all the non-boost-invariant schemes (which use the angle definition (5) of the resolution variable) are similar, and perform much worse than the boost-invariant schemes, which with one exception are also similar to one another. The exception is the monotonic p_t -scheme defined by Eq. (17) with p_t^2 replaced by p_t , which is considerably worse than the others. As we mentioned in Sect. 3.4, this can be traced to the fact that this scheme is “too monotonic”, i.e. the correction which renders the sequence monotonic actually dominates in the resolution variables towards the end of clustering. Amongst the other boost-invariant schemes, the three defined in Sect. 3.3 perform somewhat better than the other combinations, with the covariant E -scheme and monotonic p_t^2 -scheme marginally

better than the p_t -scheme. Recall that the figures discussed above and all subsequent ones were generated using the covariant E -scheme.

We have also investigated the effect of using the more sophisticated definition (11) of the generalized radius, and find that in the majority of events it behaves identically to the simpler one (10). In particular, the mean and r.m.s. of the correlations between (11) at parton level and (10) at hadron level are almost identical to those purely within the latter scheme. This suggests that the difference between the definitions (10) and (11) is smaller than the effect of hadronization.

7. Comparison with cone algorithm

At present most jet studies in hadron collisions use cone algorithms to define jets, so it is useful to consider the similarities and differences between the two types of algorithm. As noted earlier, when the n th from last merging is with the beam jets, E_{tn} is equal to the transverse momentum of the n th hardest jet. Since this jet is being merged with the beam, it must be a distance of at least one unit in (η, ϕ) from the nearest jet. The $k_\perp$ algorithm is therefore similar to a cone with a radius of 1 when the sub-jets within a jet are not resolved, at least for hard emission. We further define $E_{tn}(B)$ to be the E_{cut} value at which the n th from last merging with the beam occurs, and $E_{tn}(C)$ to be the transverse energy of the n th hardest jet with a cone algorithm, with a radius of 1. Note that since the last few mergings are usually with the beam, we have $E_{t1,2}(B) = E_{t1,2}$ in all events and $E_{t3}(B) = E_{t3}$ in about 80% of events.

Although a standard cone algorithm was agreed upon between theorists and experimenters in the Snowmass Accord [16], it is still common practice to use various approximations to that algorithm. We use one such approximation, GETJET [28] from the ISAJET [29] event generator package. This forms jets from the hit cells of a calorimeter in the following steps.

1. Find the cell with the highest E_t , which has not already been used.
2. If this has $E_t < E_{t\text{stop}}$, then stop.
3. Sum all cells with $E_t > E_{t\text{cut}}$ within a circle of radius R_{jet} .
4. Calculate the direction of the jet from the vector sum of all contributing cells.
5. Go to 1.

We use the default values for $E_{t\text{stop}}$ and $E_{t\text{cut}}$, of 1.5 GeV and 0.5 GeV respectively, and consider various jet radii, R_{jet} . This differs from the Snowmass Accord jet definition in three ways: in the use of the vector sum to find the jet direction, rather than the E_t -weighted average of η and ϕ ; in the lack of an optimization stage to ensure that the jet direction coincides with the cone centre; and in the parameter $E_{t\text{cut}}$, which reduces the sensitivity to the underlying soft event. Because of steps 1 and 3, the jet definition would not be collinear and infrared safe if used at the level of massless partons instead of calorimeter cells. As long as the partons are assigned to cells, however, the calorimeter segmentation provides the necessary cutoff, and this and the above differences have little effect in our Monte Carlo

studies. The most important difference between algorithms seems to be the treatment of overlapping cones, which is not specified in the Snowmass Accord. In the definition just given, all radiation in the overlapping region is assigned to the jet with the higher-energy ‘seed’ cell.

In Fig. 3 we compare the $E_{tn}(B)$ and $E_{tn}(C)$ spectra for $n = 3$ and 4 from partons (assigned to cells in the case of the cone analysis) and from calorimeter cells. They are indeed very similar. In Fig. 4 the parton–calorimeter correlations for $E_{t3}(B)$ and $E_{t3}(C)$ are shown, where they are also seen to be very similar, with $E_{t3}(B)$ having a slightly higher fraction of badly mismeasured values, resulting in a slightly worse r.m.s. (Table 1). In fact the cone algorithm cannot construct three jets in all events, and Fig. 4(b) shows only those events with three or more reconstructed jets at both the parton and calorimeter levels. However, a considerable fraction (almost 5%) of events contain three or more reconstructed jets in one case but not the other. If these are given $\Delta E_{t3}(C)$ values of ± 1 , corresponding to $E_{t3}(C) = 0$ when less than three jets are constructed, then the r.m.s. is increased to 0.24 which is considerably worse than for $E_{t3}(B)$. The equivalent correlations for $E_{t2}(B)$ and $E_{t2}(C)$ both have a smaller spread, due to the reduced uncertainty in reconstructing the hardest jets, with $E_{t2}(B)$ slightly better than $E_{t2}(C)$.

Although we have said that the $k_{\perp}$ -algorithm reproduces a cone of unit radius, this is not strictly the case for soft emission. In the $k_{\perp}$ -algorithm, soft particles which are near each other tend to clump together before eventually being merged with a jet or with the beams. It is hoped that soft gluons, which are emitted coherently from hard partons, are slightly more likely to be pulled into the jets, whereas particles from the underlying event are just as likely to fall out as in. Therefore the $k_{\perp}$ -algorithm has an effective cone size larger than 1 with respect to soft emission. In Fig. 5(a) we show the total amount of transverse energy which is pulled into the two hardest jets from outside unit cones centred on their directions, ω , as a function of E_{t2} . One would expect this energy transfer to be dominated by large angle emission from hard partons, and so be proportional to the jet energies, and we see that this is indeed the case. The fact that the parton result has a small intercept at zero indicates that the effects of non-soft radiation are small. The additional contributions due to non-perturbative processes are seen to be almost entirely additive, with only a small change in gradient, indicating that they are produced incoherently with respect to the jets. In Fig. 5(b) the opposite quantity, ϕ , is shown, which is the amount of transverse energy inside the unit cones which was not associated with the jets. This is seen to be much more slowly increasing with E_{t2} with roughly equal proportions coming from partons and the underlying event, and less from hadronization. Note that for the all-hadron case, ϕ is typically around 2 GeV, while ω is around two percent of E_{t2} , plus 2 GeV. We conclude that although some soft particles are transferred across the cone boundary, their energies are not sufficient to affect greatly the phenomenological properties studied here.

8. Comparison with e^+e^- $k_{\perp}$ -algorithm

The $k_{\perp}$ algorithm is already well established in e^+e^- annihilation. One of the advantages of extending it to hadron-hadron studies is that it will allow more direct comparisons to be made between the jets in e^+e^- and those in hadron-hadron collisions. In this section we study the sub-jet multiplicities in events which have a reasonably large transverse energy, $75 \text{ GeV} < E_{t2} < 100 \text{ GeV}$. We define $E_{tn}^2(M)$ to be the d_{cut} value at which the n th from

last sub-jets were merged inside one of the two hard jets surviving to E_{t2} , and further define the shape variable $y_n = (E_{tn}(M)/E_{t2})^2$. This means that y_n is the value of y_{cut} where the event changes from having n to $n - 1$ sub-jets. Since we expect sub-jet production to be dominated by final-state gluon emission, it is very closely related to the y_n used in e^+e^- studies.

In Fig. 6 we show the sub-jet fractions as a function of y_{cut} . It can be seen that down to about $y_{\text{cut}} = 3 \times 10^{-3}$ the average hadronization correction is small, and below that value grows quite rapidly. However the average correction due to the underlying event becomes significant by about $y_{\text{cut}} = 10^{-2}$, and has the effect of shifting all the graphs upwards in y_{cut} due to the fact that each of $E_{tn}(M)$ and E_{t2} are separately shifted upwards by similar amounts. The size of this effect agrees with our results in Sect. 6, i.e. a transverse energy correction of a few GeV due to the underlying event. Relative to the e^+e^- equivalents, all the curves are shifted downwards in y_{cut} . This is because of the presence of the beam jets: a final-state emission at high y_{cut} is more likely to give a jet that is merged with the beam jets.

Figure 7 shows the correlations between the parton-level and various hadron-level values for y_3 , the value at the last merging. These are seen to be considerably worse than their E_{t3} equivalents, with a significant fraction quite badly mismeasured. The corresponding mean and r.m.s. values are shown in Table 1. It can be seen that although the average hadronization correction is smaller than in e^+e^- , the r.m.s. is somewhat larger. As can be seen from Table 1, the spread of y_3 values is recombination scheme-dependent. In Fig. 8 we show the equivalent correlations in the monotonic p_t^2 -scheme, which are seen to have a narrower central band of points, but still have a large scatter of badly mismeasured events.

The spread of Δy_3 values is considerably worse than the e^+e^- case because the hadron-hadron algorithm is sensitive to events where two jets are at a separation of about $R = 1$ from each other. In the e^+e^- algorithm, if the angle between two such jets is slightly shifted during hadronization, then one of the jets may be closer to a third jet, and those merged, but the resulting jet will usually end up being merged with the first jet anyway. In the hadron-hadron case, however, one jet could be merged with the beam or with the other jet, depending on their exact opening angle. Once it is merged with the beam, it is never available for further merging. Thus although E_{tn} will not be much changed, $E_{tn}(B)$ and $E_{tn}(M)$ will be completely switched to $\sim E_{t(n\pm 1)}$, and the subsequent merging history will be different. This is the cause of the stray events far from the centre line. It seems that problems of this kind are unavoidable at some level, since one must differentiate between emission which is factorized into the fragmentation functions, and that which is factorized into the distribution functions, and emission close to the boundary will always be sensitive to hadronization corrections. Their relative effect on jet rates, however, decreases as the transverse energies of the hard final state jets increase.

Using the cone algorithm it is also possible to study sub-jets, by varying the cone radius. We define the number of sub-jets by analogy with the $k_\perp$ case as follows. We run the jet algorithm with unit radius and find the hardest two jets. Then we rerun the algorithm with various radii R , classifying as sub-jets only those whose axes lie within one unit in (η, ϕ) of a hard jet axis. These sub-jet fractions are shown in Fig. 9, where it can be seen that the hadronization corrections are large, and furthermore are not under control even

for $R \rightarrow 1$. In Fig. 10 we show the correlation in R_3 , the jet radius at which the number of sub-jets changes from 3 to 2. There can be seen to be very little correlation, leading to a much higher reconstruction uncertainty than in the $k_\perp$ case.

We conclude that although the hadronization uncertainty is somewhat larger than that encountered in e^+e^- studies, it is certainly much smaller than from a cone algorithm used in a similar way.

9. Conclusions

We propose a new class of jet algorithms for hadron-hadron collisions, based on the clustering approach, using distance measures of $k_\perp$ type. The algorithms lead to an unambiguous and exhaustive classification of hadronic final states into multijet configurations at the exclusive level. A distinction is made between hard final state jets and the remnants of the incoming hadrons, or beam jets, which allows the factorization of initial-state singularities into universal distribution functions. The structure of this factorization and the practical implementation of the algorithms are improved by making them invariant under boosts along the beam directions. Within this procedure, there is still considerable flexibility to choose various recombination schemes and resolution variables. In this paper we have discussed several such schemes, and seen that their practical properties are similar.

A subset of the jets reconstructed by the cluster algorithm, those which are ‘just about’ to be merged with the beam jets, were shown to have similar properties to those reconstructed by an algorithm of cone type. However, much of the interest in jet physics involves the internal structure of such jets, and we have shown that the cluster algorithm is superior to the cone algorithm for such studies, since it suffers a much smaller correction in relating the measurable hadronic final state to the calculable partonic final state. This can be traced to the fact that variables of $k_\perp$ type do not give rise to strong kinematic correlations [4,5] and thus, assuming local parton hadron duality, would be expected to give non-perturbative corrections which only modify the results at the hadronization scale of around 1 GeV. Furthermore, ambiguities present in cone algorithms about the treatment of overlapping cones are avoided by the unique assignment of every particle to exactly one hard final state jet or beam jet.

One of the variants of the scheme we propose was studied in a preliminary way in Ref. [17], where it was shown to have advantages over conventional cone-type algorithms outside the realm of pure QCD and jet physics. The two algorithms were compared for the reconstruction of semi-leptonic decays of heavy Higgs bosons at super-collider energies. This situation is a particular problem for conventional cone algorithms, since the hadronically decaying W or Z boson typically has a rather large transverse boost. This almost guarantees that the two resulting jets will be reconstructed with overlapping cones. In contrast, the resolution parameter in the cluster algorithm can be adjusted from event to event, so that two jets are unambiguously reconstructed. Assuming that the momentum of the leptonically decaying W or Z is perfectly known, this was shown to give a spread in reconstructed masses which is four times smaller than that from the cone algorithm, for WW/ZZ pairs of mass 600 GeV. This is a direct reflection of the reduced hadronization uncertainty in our cluster algorithm relative to the cone algorithm. Further analysis along these lines is in progress [30].

The $k_{\perp}$ -algorithm discussed in this paper is infrared and collinear safe at parton level and guarantees the factorization of initial-state collinear singularities into universal parton distributions. Therefore, the corresponding jet cross sections can be computed in QCD perturbation theory using standard techniques. In particular, the tools already exist for a full next-to-leading order calculation [14,24]. Furthermore, the use of a jet resolution variable of $k_{\perp}$ type makes it possible to go beyond fixed-order perturbation theory and resum classes of large logarithmic corrections, which may spoil the convergence of the perturbative expansion in the kinematical regions $d_{\text{cut}}/S \rightarrow 0, 1$ and $y_{\text{cut}} \rightarrow 0$. Resummed calculations can indeed be carried out along the lines of Refs. [4,12].

Although throughout this paper we have always referred to hadron-hadron collisions, the $k_{\perp}$ type of algorithms proposed here can be used also for *photoproduction* processes or, in general, for any high energy collision initiated by particles having hadron-like behaviour.

The use of the $k_{\perp}$ -algorithm will permit more detailed comparisons between experimental results and perturbative QCD to be made. It will also allow, for the first time, detailed quantitative studies of similarities and differences between jets produced in hadron-hadron collisions and those produced in e^+e^- annihilation.

Acknowledgements

We have enjoyed conversations on this topic with S.D. Ellis, Z. Kunszt, G. Marchesini and D.E. Soper.

References

1. B. Flaughner and K. Meier, FERMILAB-Conf.-90/248-E, to be published in *Research Directions for the Decade*, Proc. Summer Study on High Energy Physics, Snowmass, Colorado, 1990.
2. G. Sterman and S. Weinberg, Phys. Rev. Lett. 39 (1977) 1436.
3. JADE Collaboration, W. Bartel et al., Zeit. Phys. C33 (1986) 23; S. Bethke et al., Phys. Lett. 213B (1988) 235.
4. S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock and B.R. Webber, Phys. Lett. 269B (1991) 432.
5. S. Catani, in *QCD at 200 TeV*, Proc. 17th INFN Eloisatron Project Workshop, Erice, Italy, June 1991, eds. L. Cifarelli and Yu.L. Dokshitzer (Plenum Press, New York, 1992); Yu.L. Dokshitzer, lecture given at the 29th Course of the Int. School of Sub-nuclear Physics, Erice, Italy, 1991, in *Physics at the Highest Energy and Luminosity*, ed. A. Zichichi (Plenum Press, New York, 1992); Yu.L. Dokshitzer and M. Olsson, Lund preprint LU TP 92-17; S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock and B.R. Webber, Cambridge preprint Cavendish-HEP-91/12, in preparation.
6. N. Brown and W.J. Stirling, Phys. Lett. 252B (1990) 657; Zeit. Phys. C53 (1992) 629.
7. S. Bethke, Z. Kunszt, D.E. Soper and W.J. Stirling, Nucl. Phys. B370 (1992) 310.

8. S. Catani, Yu.L. Dokshitzer, F. Fiorani and B.R. Webber, Nucl. Phys. B377 (1992) 445, Nucl. Phys. B383 (1992) 419.
9. OPAL Collaboration, P.D. Acton et al., Zeit. Phys. C55 (1992) 1; ALEPH Collaboration, D. Decamp et al., Phys. Lett. 284B (1992) 163; L3 Collaboration, O. Adriani et al., Phys. Lett. 284B (1992) 471.
10. ALEPH Collaboration, contributed paper at XXVI Int. Conf. on High Energy Physics, Dallas, August 1992; L3 Collaboration, contributed paper at XXVI Int. Conf. on High Energy Physics, Dallas, August 1992; OPAL Collaboration, P.D. Acton et al., preprint CERN-PPE/93-02.
11. A. Bassetto, M. Ciafaloni and G. Marchesini, Phys. Rep. 100 (1983) 202; Yu.L. Dokshitzer, V.A. Khoze, A.H. Mueller and S.I. Troyan, *Basics of Perturbative QCD*, Editions Frontières, Gif-sur-Yvette, 1991.
12. S. Catani, Yu.L. Dokshitzer and B.R. Webber, Phys. Lett. 285B (1992) 291.
13. UA1 Collaboration, G. Arnison et al., Phys. Lett. 132B (1983) 214, Phys. Lett. 172B (1986) 461; UA2 Collaboration, J. Alitti et al., Phys. Lett. 257B (1991) 232; CDF Collaboration, F. Abe et al., Phys. Rev. Lett. 62 (1988) 613, Phys. Rev. Lett. 68 (1992) 1104.
14. S.D. Ellis, Z. Kunszt and D.E. Soper, Phys. Rev. Lett. 62 (1989) 726, Phys. Rev. D 40 (1989) 2188, Phys. Rev. Lett. 64 (1990) 2121, Phys. Rev. Lett. 69 (1992) 1496; Z. Kunszt and D.E. Soper, Phys. Rev. D 46 (1992) 192.
15. F. Aversa, P. Chiappetta, M. Greco and J.Ph. Guillet, Nucl. Phys. B327 (1989) 105, Phys. Rev. Lett. 65 (1990) 401, Zeit. Phys. C46 (1990) 253; F. Aversa, P. Chiappetta, L. Gonzalez, M. Greco and J.Ph. Guillet, Zeit. Phys. C49 (1991) 459.
16. J.E. Huth et al., FERMILAB-Conf.-90/249-E, to be published in *Research Directions for the Decade*, Proc. Summer Study on High Energy Physics, Snowmass, Colorado, 1990.
17. M.H. Seymour, Cambridge PhD thesis, 1992.
18. D. Amati and G. Veneziano, Phys. Lett. 83B (1979) 87; A. Bassetto, M. Ciafaloni and G. Marchesini, Phys. Lett. 83B (1979) 207; G. Marchesini, L. Trentadue and G. Veneziano, Nucl. Phys. B181 (1980) 335; Ya.I. Azimov, Yu.L. Dokshitzer, V.A. Khoze and S.I. Troyan, Phys. Lett. 165B (1985) 147, Zeit. Phys. C27 (1985) 65.
19. D. Amati, R. Petronzio and G. Veneziano, Nucl. Phys. B140 (1978) 54, Nucl. Phys. B146 (1978) 29; R.K. Ellis, H. Georgi, M. Machacek and G.G. Ross, Phys. Lett. 78B (1978) 281, Nucl. Phys. B152 (1979) 285; C.T. Sachrajda, Phys. Lett. 73B (1978) 185, Phys. Lett. 76B (1978) 100.
20. S.D. Ellis, Z. Kunszt and D.E. Soper, Phys. Rev. Lett. 69 (1992) 3615.
21. R.K. Ellis, G. Marchesini and B.R. Webber, Nucl. Phys. B286 (1987) 643, E Nucl. Phys. B294 (1987) 1180.

- 22. ALEPH Coll., D. Decamp et al., Phys. Lett. 255B (1991) 623.
- 23. Z. Kunszt, P. Nason, G. Marchesini and B.R. Webber, in ‘Z Physics at LEP’, CERN 89-08, vol.1, p. 373.
- 24. W.T. Giele, E.W.N. Glover and D.A. Kosower, preprint CERN-TH.6750/92.
- 25. G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L. Stanco, Computer Phys. Comm. 67 (1992) 465.
- 26. UA5 Collaboration, G.J. Alner et al., Nucl. Phys. B291 (1987) 445.
- 27. G. Marchesini and B.R. Webber, Nucl. Phys. B310 (1988) 461, Phys. Rev. D 38 (1988) 3419; CDF Collaboration, F. Abe et al., Phys. Rev. D 44 (1991) 601 and contributions to XXVI Int. Conf. on High Energy Phys., Dallas, Texas, 1992.
- 28. F.E. Paige, CALSIM and GETJET, private communication.
- 29. F.E. Paige and S.D. Protopopescu, BNL report 38034 (1986).
- 30. M.H. Seymour, in preparation.

Tables

Recombination Scheme			Correlation between partons and		
			Hard Hadrons	All Hadrons	Calo. Cells
E -scheme	E_{t3}	mean	-0.014	0.025	0.028
		r.m.s.	0.088	0.099	0.100
	$E_{t3}(B)$	mean	-0.030	0.036	0.039
		r.m.s.	0.126	0.139	0.140
	y_3	mean	0.013	0.070	0.080
		r.m.s.	0.293	0.303	0.309
p_t -scheme	E_{t3}	mean	-0.008	0.034	0.040
		r.m.s.	0.091	0.103	0.105
	$E_{t3}(B)$	mean	-0.018	0.049	0.055
		r.m.s.	0.128	0.142	0.146
	y_3	mean	0.008	0.059	0.087
		r.m.s.	0.293	0.303	0.313
monotonic p_t^2 -scheme	E_{t3}	mean	-0.013	0.026	0.033
		r.m.s.	0.087	0.098	0.100
	$E_{t3}(B)$	mean	-0.020	0.043	0.049
		r.m.s.	0.125	0.136	0.139
	y_3	mean	-0.019	0.019	0.043
		r.m.s.	0.274	0.284	0.289
cone	$E_{t3}(C)$	mean			-0.026
		r.m.s.			0.132

Table 1: Mean and r.m.s. fractional mismeasurement of E_{t3} , $E_{t3}(B)$, $E_{t3}(C)$ and y_3 , between parton level, and the various hadron levels. For definitions, read the text.

Figure Captions

- Figure 1: Differential jet cross-sections in the longitudinally-invariant $k_{\perp}$ algorithm (E-scheme), $d\sigma/dE_{tn}$,
 (a) $n = 2$,
 (b) $n = 3$,
 (c) $n = 4$.
 Histograms are from partons (solid), hard hadrons (dashed), all hadrons (dot-dashed) and calorimeter cells (dotted).
- Figure 2: Correlations between the E_{t3} values obtained from partons (x -axes) and
 (a) hard hadrons,
 (b) all hadrons,
 (c) calorimeter cells.
- Figure 3: As Fig. 1 for $E_{tn}(B)$,
 (a) $n = 3$,
 (b) $n = 4$,
 and for $E_{tn}(C)$, from partons (solid) and calorimeter cells (dashed),
 (c) $n = 3$,
 (d) $n = 4$.
- Figure 4: Correlations between the E_{t3} values obtained from partons and from calorimeter cells,
 (a) $E_{t3}(B)$,
 (b) $E_{t3}(C)$.
- Figure 5: Average transverse energies crossing the boundaries of unit cones, centred on the hardest two jets, as a function of E_{t2} . Histograms are as in Fig. 1; bars show the errors due to Monte Carlo statistics.
 (a) ω , the average transverse energy outside the cones but in the jets,
 (b) ϕ , the average transverse energy inside the cones but not in the jets.
- Figure 6: The sub-jet fractions as a function of y_{cut} . Curves are as in Fig. 1.
- Figure 7: Correlations between the y_3 values obtained from partons (x -axes) and
 (a) hard hadrons,
 (b) all hadrons,
 (c) calorimeter cells,
 using the covariant E -scheme.
- Figure 8: As Fig. 7, but using the monotonic p_t^2 -scheme.
- Figure 9: The sub-jet fractions from the cone algorithm as a function of R . Curves are as in Fig. 3(c) and (d).
- Figure 10: Correlations between the R_3 values obtained from partons (x -axis) and calorimeter cells, using the cone algorithm.

Figure 1(a)

Figure 1(b)

Figure 1(c)

Figure 2(a)

Figure 2(b)

Figure 2(c)

Figure 3(a)

Figure 3(b)

Figure 3(c)

Figure 3(d)

Figure 4(a)

Figure 4(b)

Figure 5(a)

Figure 5(b)

Figure 6

Figure 7(a)

Figure 7(b)

Figure 7(c)

Figure 8(a)

Figure 8(b)

Figure 9

Figure 8(c)

Figure 10