Factorization and Jet Clustering Algorithms for Deep Inelastic Scattering

B.R. Webber
Theory Division, CERN, CH-1211 Geneva 23, Switzerland

Abstract

Jet clustering algorithms for final states in deep inelastic lepton scattering are discussed from a theoretical viewpoint. The importance of factorization and the special rôle of the Breit frame of reference are stressed. Predictions of jet cross sections in factorizable JADE- and $k_{\perp}$-type algorithms are presented.
1. Factorization

Jet physics at HERA will be an important testing-ground for QCD, in which processes involving both incoming and outgoing hadrons can be probed in a cleaner environment than that of hadron-hadron collisions. A new feature of deep inelastic lepton scattering (DIS) compared with $e^+e^-$ annihilation is the presence of a 'beam jet' containing the remnant of the incoming hadron together with QCD radiation from its struck constituent. The algorithms used to define jets should distinguish as clearly as possible between this beam jet and the 'hard jets' produced in the short-distance DIS process. This leads one naturally to consider the properties of jet clustering algorithms with respect to factorization between the beam fragmentation and the hard process.

The importance of factorization of the DIS structure functions has long been recognized [1]. For simplicity, let us consider here the case of purely electromagnetic neutral-current scattering. Recall that the differential cross section with respect to the variables $x = Q^2/2p \cdot q$ and $y = p \cdot q/p \cdot l$ can be written in the general form

$$\frac{d^2\sigma}{dx dy} = \frac{2\pi \alpha^2 S}{Q^4} \left\{ [1 + (1 - y)^2] F_T(x, Q^2) + 2(1 - y) F_L(x, Q^2) \right\}$$

where $F_T$ and $F_L$ are structure functions that correspond respectively to transverse and longitudinal polarization of the virtual photon (in an older notation, $F_T = 2x F_1$ and $F_L = F_2 - 2x F_1$). Then factorization means that for either polarization $P = T, L$ the structure function is expressible in the form

$$F_P(x, Q^2) = \sum_i \int_0^1 dy \frac{x}{\eta} C_{P,i} \left( \frac{x}{\eta}, \alpha_s(\mu^2), Q^2, \mu^2 \right) D_i(\eta, \mu^2)$$

where $D_i(\eta, \mu^2)$ represents the distribution of longitudinal momentum fraction $\eta$ for partons of type $i$ in the target hadron, probed at momentum transfer scale $\mu^2$.

At scales $Q^2$ and $\mu^2$ sufficiently large for perturbation theory to make sense, the coefficient functions $C_{P,i}$ are perturbatively computable. One finds that

$$C_{P,i}(z, \alpha_s, Q^2, \mu^2) = \epsilon_i^2 \delta_{PT} \delta(1 - z) + \mathcal{O}(\alpha_s)$$

where $\epsilon_i$ is the parton charge, so that in zeroth-order QCD (the parton model) the scattering is purely transverse and samples the quark and antiquark densities at momentum fraction $\eta = x$. In first order, longitudinal scattering also occurs, both it and the transverse contribution are sensitive to the gluon distribution, and the parton momentum fraction $\eta$ is generally larger than $x$ ($0 < z = x/\eta \leq 1$).

In contrast to the coefficient functions, the parton distributions $D_i$ are intrinsically non-perturbative. However, once measured experimentally at some scale $\mu^2$, they are universal, that is, they can be convoluted with the appropriate coefficient functions to calculate any other observable having the factorizable form (2). Furthermore they can be evolved perturbatively from the scale of measurement $\mu^2$ to any other scale (e.g., $Q^2$) by means of the Altarelli-Parisi equations.

Factorization is a deep property of QCD, without which the calculation of structure functions and many other observables would be impossible. This is because one actually encounters divergences in the perturbative calculation of the coefficient functions $C_{P,i}$,
arising from collinear parton emission from the incoming parton $i$. These divergences are associated with long-distance physics and are not cured by renormalization. It proves possible to extract all such divergences from the coefficient functions, to all orders in perturbation theory, as universal factors which can be absorbed into the parton distributions $D_i$. In this way all sensitivity to long-distance, non-perturbative QCD, which must in reality regularize the collinear divergences, is isolated in the parton distributions, which are taken from experiment.

Let us now consider the application of a jet algorithm to classify each final state uniquely according to its number of jets at some jet resolution scale $d_{\text{cut}}$, to be defined later. We shall adopt the usual DIS nomenclature and write $\sigma^{(n)}$ for the cross section for an $(n+1)$-jet final state, where the 1 refers to the beam (remnant) jet and the $n$ to hard jets resolved from the remnant. In order to be able to compute such a cross section perturbatively from the universal parton distributions $D_i$, we need to be able to write it in the factorized form (2). Otherwise we may find that some other parton distributions, possibly different for each $n$, need to be defined and measured, or even that no finite prediction can be made if the collinear divergences cannot be factorized.

An important feature of factorization is that the coefficient functions depend on the Bjorken variable $x$ and the momentum fraction $\eta$ of the struck parton only via the ratio $z = x/\eta$. As a consequence, they can be calculated at the parton level, without any explicit reference to the incoming hadron momentum $p$. To preserve this feature, the jet algorithm should not introduce any other form of dependence on $x$ or $\eta$ through the criteria used to separate the total cross section into its components with different values of $n$. In addition, one would prefer not to introduce any dependence on $y$ or other lepton variables, so that each structure function separately can be split into $(n+1)$-jet contributions. In summary, a jet algorithm that is well-formulated from the theoretical viewpoint should provide jet cross sections of the form

$$
\frac{d^2\sigma^{(n)}}{dx\,dy} = \frac{2\pi\alpha_s^2 S}{Q^4} \left\{ [1 + (1 - y)^2] F_T^{(n)}(x, Q^2, d_{\text{cut}}) + 2(1 - y) F_L^{(n)}(x, Q^2, d_{\text{cut}}) \right\}
$$

(4)

where

$$
F_P^{(n)}(x, Q^2, d_{\text{cut}}) = \sum_i \int_x^1 d_{\eta} \eta C_{P,i}^{(n)} \left( \frac{x}{\eta}, \alpha_s(\mu^2), Q^2, \mu^2, d_{\text{cut}} \right) D_i(\eta, \mu^2).
$$

(5)

Let us review how the dependence of the coefficient functions on the ratio $x/\eta$ comes about for the full structure functions $F_P$. In this case the coefficient function $C_{P,i}$ is computed [2] from the total cross section for interaction between an (unpolarized) incoming parton of type $i$ and momentum $\eta p$ and a virtual photon of polarization $P$ and momentum $q$. This can only depend on $i$, $P$ and the incoming momenta $\eta p$ and $q$, and must be Lorentz-invariant. Hence schematically, neglecting $p^2$ in the Bjorken limit,

$$
C_{P,i} = C_{P,i}[\eta p \cdot q, q^2] = C_{P,i}[Q^2/(2x/\eta), -Q^2],
$$

(6)

which is indeed a function of $x/\eta$ (and $Q^2$), rather than of $x$ and $\eta$ separately.

Now let us extend this argument to the $(n+1)$-jet coefficient function $C_{P,i}^{(n)}$ [3,4]. We shall see below that in addition to the scalar resolution variable $d_{\text{cut}}$ and the incoming and
outgoing momenta of the hard parton-photon subprocess, one generally needs to introduce at least one auxiliary vector $\vec{p}$ in order to define a jet clustering algorithm fully. After integration over the outgoing momenta, one is left with a scalar function of $\eta p$, $q$, $\vec{p}$ and $d_{\text{cut}}$, i.e.

$$C^{(n)}_{P;i} = C^{(n)}_{P;i}[\eta p \cdot q, q^2, \eta p \cdot \vec{p}, q \cdot \vec{p}, \vec{p}^2, d_{\text{cut}}] . \quad (7)$$

In order to preserve the $\eta$ dependence in Eq. (4), the auxiliary vector should be expressible as a linear combination of $p$ and $q$, say $\vec{p} = \alpha p + \beta q$, with coefficients $\alpha$ and $\beta$ that can in principle be any functions of $x$ and $Q^2$. Substituting into Eq. (7), we find

$$C^{(n)}_{P;i} = C^{(n)}_{P;i}[Q^2/(2x/\eta)\,\eta, -Q^2, \beta Q^2/(2x/\eta), \alpha Q^2/2x, \beta Q^2(\alpha/x - \beta), d_{\text{cut}}] . \quad (8)$$

For this to be a function of $x/\eta$, $Q^2$ and $d_{\text{cut}}$ only, we require $\alpha = x f(Q^2)$ and $\beta = g(Q^2)$. Thus the jet algorithm will only be factorizable if the auxiliary vector has the form

$$\vec{p} = x f(Q^2)p + g(Q^2)q . \quad (9)$$

If several auxiliary vectors are required to define the algorithm, they must all be of the form (9).

2. **Factorizable JADE-type algorithm**

The most natural way to perform jet clustering in DIS final states is to proceed as far as possible in the same way as in $e^+e^-$ annihilation. In particular we can start with a JADE [5] type of algorithm, based on a resolution variable of the form

$$d_{ij} = 2p_i \cdot p_j \simeq (p_i + p_j)^2 . \quad (10)$$

One first computes $d_{ij}$ for all pairs of final-state particles. If the smallest value is less than $d_{\text{cut}}$, then one combines the corresponding pair into a cluster $(ij)$ with momentum (in the simplest scheme) $p_{(ij)} = p_i + p_j$. The relevant $d_{ij}$ values are recomputed using the cluster in place of its constituent particles. The procedure is iterated until all remaining $d_{ij}$ values are larger than $d_{\text{cut}}$, at which point the clusters are defined as jets.

We have chosen not to normalize the resolution (10) to any particular energy scale yet, in order to emphasize that the clustering procedure is scale-independent. It is $d_{\text{cut}}$ that sets the scale, by specifying when we stop clustering and define the clusters as jets. In $e^+e^-$ annihilation, the natural choice of scale for $d_{\text{cut}}$ is the total centre-of-mass energy-squared $S$. Thus we write $d_{\text{cut}} = y_{\text{cut}} S$ in that case, where $y_{\text{cut}}$ is a dimensionless constant. In DIS the analogous variable to $S$ would appear to be the hadronic centre-of-mass energy-squared, $W^2 = (p + q)^2 = Q^2(1 - x)/x$ [6-8]. We see however that this would introduce a non-factorizing $x$ dependence through the factor $(1 - x)/x$. In other words, the corresponding auxiliary vector $p + q$ is not of the required form (9). The only factorizable choice of auxiliary vector is that in Eq. (9), which gives

$$\vec{p}^2 = g(f - g)Q^2 = F(Q^2) . \quad (11)$$

Thus the natural choice is

$$d_{\text{cut}} = f_{\text{cut}} Q^2 , \quad (12)$$
where the dimensionless constant \( f_{\text{cut}} \) is analogous to \( y_{\text{cut}} \), but we denote it differently to remind ourselves that the DIS algorithm is different from \( e^+e^- \), and to avoid confusion with the DIS variable \( y \).

The simplest example of a suitable auxiliary vector is
\[
\vec{p} = 2xp + q .
\] (13)

Then \( \vec{p}^2 = Q^2 \) and \( \vec{p} \cdot q = 0 \), i.e. the virtual photon momentum is purely spacelike in the rest frame of \( \vec{p} \), which is called the Breit frame. While Eq. (9) allows other choices, every frame with a timelike rest vector of the form (9) can be transformed into the Breit frame by an \( x \)-independent boost along the direction of \( q \). Thus the Breit frame, or at least this family of related frames, assumes special significance in the theory of DIS jet algorithms. Note that unfortunately the hadronic centre-of-mass and HERA laboratory frames do not belong to the preferred family.

Even with \( d_{\text{cut}} \propto Q^2 \), the algorithm defined above is not factorizable. This is because the final state involves not only the products of the hard parton-photon scattering but also the hadron remnant, with momentum \( p_r = (1 - \eta)p \). Since this vector is not of the form (9), treating the remnant as an ordinary final-state particle will lead to non-factorizing jet cross sections. The physical reason for this is clear: including the remnant in the clustering procedure causes the products of the hard parton scattering and those of the soft remnant fragmentation to become inextricably entangled. On the other hand, the remnant itself can only be identified by applying some form of jet algorithm, so we cannot exclude it from clustering altogether.

These difficulties can be resolved by introducing a further auxiliary vector \( p_b = xp \) and modifying the JADE clustering procedure as follows. In addition to the quantities \( d_{ij} \) defined by Eq. (10), we now compute also
\[
d_{ib} = 2p_i \cdot p_b = 2xp_i \cdot p
\] (14)
for every final state particle or cluster. If the smallest of all the current \( d_{ij} \) and \( d_{ib} \) values is a \( d_{ij} \), then that pair is clustered as before. But if it is a \( d_{ib} \), then the particle or cluster \( i \) is classified as part of the beam (remnant) jet and is not available for further clustering. Since \( p_b \) is of the canonical form (9), no violation of factorization will result from this way of separating the beam jet from the hard scattering products.

Notice that we must not replace \( p_b = xp \) by \( p'_b = xp + p_i \) after assigning particle \( i \) to the beam jet, because the auxiliary vector would then no longer be of the form (9). In particular, after absorbing the hadron remnant with momentum \( p_r = (1 - \eta)p \), which has \( d_{rb} = 0 \) and hence is always assigned to the beam jet, we would have \( p'_b = (1 - \eta + x)p \). Instead, \( p_b \) remains fixed and the remnant has no effect on clustering.\(^4\)

The operation of this “factorizable JADE” algorithm can be illustrated using the order-\( \alpha_s \) subprocesses \( \gamma^*q \to gq, \gamma^*\bar{q} \to g\bar{q} \) and \( \gamma^*g \to q\bar{q} \). We denote the two outgoing parton momenta by \( p_{1,2} \) and define the momentum fractions
\[
z_i = p_i \cdot p/p \cdot q \quad (i = 1, 2) .
\] (15)

\(^4\)Apart from this feature, the algorithm is similar to the ‘mixed’ algorithm of Körner et al. [7].
Then $0 \leq z_i \leq 1$ and $z_1 + z_2 = 1$. The jet resolution variables are
\[ d_{12} = 2p_1 \cdot p_2 = (\eta / x - 1)Q^2, \quad d_{ib} = 2xp_i \cdot p = z_i Q^2, \quad d_{(12)b} = 2x(p_1 + p_2) \cdot p = Q^2. \] (16)

For a fixed value of $x / \eta$, there is only one relevant free kinematic variable, which we take to be $z_1$. Then the regions with different jet multiplicities can be depicted on a $(z_1, f_{\text{cut}})$ plot as shown in Fig. 1. For $x / \eta < 2/3$, the value of $d_{12}$ is irrelevant since either $d_{1b}$ or $d_{2b}$ is smaller and one of the hard partons disappears first into the beam jet. When $x / \eta > 2/3$ and $z_1, z_2 > \eta / x - 1$, the two hard partons are clustered before merging with the beam. In this region, at least one hard jet is resolved for any $f_{\text{cut}} < 1$.

![Figure 1: Regions in the $(z_1, f_{\text{cut}})$-plane, labelled by $n$ for an $(n+1)$-jet final state, according to the factorizable JADE algorithm: (a) $x / \eta < 2/3$; (b) $x / \eta > 2/3$.](image)

In $e^+e^-$ annihilation, some advantages have been found [5] in defining the cluster resolution variable as
\[ d_{ij} = 2E_i E_j (1 - \cos \theta_{ij}), \] (17)
which coincides with the covariant definition (10) only in the massless limit. The corresponding modification of Eq. (14) would be
\[ d_{ib} = 2x E_i E_p (1 - \cos \theta_i). \] (18)

Since these are no longer covariant expressions, we have to specify some frame of reference in which they are to be evaluated. Again, this corresponds to introducing an auxiliary vector, whose rest frame is the preferred one. To preserve factorization, the vector must be of the form (9), which excludes the hadronic centre-of-mass and HERA laboratory frames, but not the Breit frame, specified by Eq. (13).

For the $O(a_s)$ hard subprocesses discussed above, the only change when using the non-covariant algorithm occurs when the two hard partons are clustered to form a jet of finite mass. In the Breit frame we find that
\[ E_{(11)} = Q \eta / 2x, \quad E_p = Q / 2x, \quad \cos \theta_{(12)} = \text{sign}(1 - 2x / \eta) \] (19)
and so for the case of interest \((x/\eta > 2/3)\) Eq. (18) gives

\[
d_{(12)d} = (\eta/x)Q^2,
\]

which means that one hard jet may now be resolved up to \(f_{\text{cut}} = \eta/x < 3/2\), as indicated in Fig. 1. Notice that Eq. (20) again involves only the factorizable ratio \(x/\eta\), as a consequence of our use of the Breit frame. Notice also the difference from the \(e^+e^-\) algorithm, in which no jets can be resolved when \(y_{\text{cut}} > 1\).

To predict the corresponding \((n+1)\)-jet cross sections for a given value of \(f_{\text{cut}}\), we have to integrate the subprocess cross sections [9] over the relevant ranges of \(z_1\) for each value of \(x/\eta\) to obtain the coefficient functions \(C_{F,i}^{(n)}\) for insertion in Eqs. (4,5). For example, the \((2+1)\)-jet coefficient functions are of the form

\[
C_{F,i}^{(2)}(x/\eta = z, f_{\text{cut}} = f) = g_\alpha \frac{\alpha_s}{2\pi} \Theta(1 - 2f) \Theta(1 - z - fz) K_{F,i}^{(2)}(z, f)
\]

where \(g_\alpha = e_\alpha^2\), \(g_\beta = \sum_{q=1}^{N_f} e_q^2\) and

\[
K_{T,q}^{(2)}(z, f) = \frac{4}{3} \left[ z + \frac{z^2}{1 - z} \ln \left( \frac{1 - f}{f} \right) + \left( 2 + z - \frac{3/2}{1 - z} \right) (1 - 2f) \right]
\]

\[
K_{T,q}^{(2)}(z, f) = [z^2 + (1 - z)^2] \left( z \ln \left( \frac{1 - f}{f} \right) - (1 - 2f) \right)
\]

\[
K_{L,q}^{(2)}(z, f) = \frac{8}{3} z(1 - 2f), \quad K_{L,q}^{(2)}(z, f) = 4z(1 - z)(1 - 2f).
\]

Convoluting these functions with the parton distributions according to Eq. (5), one obtains the structure function contributions shown, together with those for \(n = 0,1\), in Fig. 2(a). For these calculations the MRS D' parton distributions [10] have been used. We see that, as expected, the \((2+1)\)-jet fraction falls with increasing \(f_{\text{cut}}\) while the fraction with no resolved hard jets rises. Consequently the \((1+1)\)-jet contribution, which gives the whole cross section in lowest order for \(f_{\text{cut}} < 1\), now exhibits first a rise and then a fall with increasing \(f_{\text{cut}}\). The factorization property ensures that the form of the curves does not depend strongly on \(x\) or \(Q^2\). Notice that in fixed-order calculations there are logarithmic divergences at \(f_{\text{cut}} = 0\) and 1, which are expected to be regularized by form-factor effects when such logarithms are resummed to all orders [3,4].

3. Factorizable \(k_T\)-algorithm

The JADE-type jet algorithm outlined in the previous Section is satisfactory from the formal viewpoint of perturbative computability and factorization. However, in \(e^+e^-\) annihilation such algorithms have been found to have some undesirable phenomenological features: large hadronization corrections, non-intuitive assignment of particles to jets, and generation of ‘phantom jets’ [11,12]. These features arise from the use of invariant mass as a resolution criterion, rather than directed energy flow: the algorithm counts ‘lumps’ instead of jets. In addition to the phenomenological problems, the invariant mass criterion leads to theoretical difficulties in the calculation of multijet rates at the parton level. For example, the exponentiation of multiple soft gluon contributions, although present in the
Figure 2: Multijet contributions to DIS structure functions to order $\alpha_S$, according to the factorizable (a) JADE-type and (b) $k_\perp$-type algorithms.

matrix elements, is not seen in the cross sections, owing to the algorithm's tendency to cluster soft gluons together, even when they are far apart in angle [13,14].

These deficiencies have been largely overcome through the introduction, at the last Durham phenomenology workshop [15], of $k_\perp$-type algorithms [16-18] for $e^+e^-$ annihilation, in which Eq. (17) is replaced by

$$d_{ij} = 2\min\{E_i^2, E_j^2\}(1 - \cos\theta_{ij}). \quad (23)$$

This means that a soft particle is clustered with the particle or jet that is nearest in angle, regardless of the energy of the latter, which prevents the generation of 'phantom jets' and ensures soft gluon exponentiation.

The extension of $k_\perp$-type algorithms to DIS, as proposed in Ref. [3], is straightforward in the light of the discussion in the previous Section. To ensure factorization, the energy scale for $d_{\text{cut}}$ should be $Q^2$, and the particle energies and angles should be evaluated in the Breit frame. The resolution variable (18) for the beam jet becomes

$$d_{ip} = 2E_i^2(1 - \cos\theta_i). \quad (24)$$

Note that we do not need to introduce $p_b = xp$ here because the value of $d_{ip}$ is now invariant under rescaling of $p$. However, we still implicitly introduce the auxiliary vector $\vec{p} = 2xp + q$ by specifying the use of the Breit frame. In terms of this vector, in the massless approximation, we may write

$$d_{ij} = 2p_i \cdot p_j \min\left\{\frac{p_i \cdot \vec{p}}{p_j \cdot \vec{p}}, \frac{p_j \cdot \vec{p}}{p_i \cdot \vec{p}}\right\}, \quad d_{ip} = \frac{2p_i \cdot p \cdot p_i \cdot \vec{p}}{p \cdot \vec{p}}. \quad (25)$$
For the $\mathcal{O}(\alpha_S)$ subprocesses, using the notation of Eq. (15) and $z = x/\eta$, we have explicitly

\[
d_{12} = \frac{(1-z)}{z}Q^2\min\left\{\frac{z + z_1 - 2zz_1}{1 - z - z_1 + 2zz_1}, \frac{1 - z - z_1 + 2zz_1}{z + z_1 - 2zz_1}\right\}
\]
\[
d_{ip} = z_i(1 - z - z_i + 2zz_i)Q^2/z.
\]

(26)

If the hard partons are clustered first, we also require

\[
d_{(12)p} = \frac{Q^2}{z^2}\Theta(2z - 1),
\]

(27)

where we have used the Breit frame results (19) in Eq. (24), rather than the massless approximation (25).

The regions of different jet multiplicity in the $(z_1, f_{cut})$ plot defined by Eqs. (26,27) are more complicated than those in Fig. 1, but they still factorize and the numerical evaluation of the jet cross sections according to Eqs. (4,5) remains straightforward. Results on the structure function contributions are shown in Fig. 2(b). The qualitative features are similar to those for the JADE-type algorithm, although the cross sections for higher jet multiplicities at a given value of $f_{cut}$ are generally smaller for the $k_\perp$-algorithm, as found also in $e^+e^-$ annihilation.

Acknowledgements

The work described here was done in collaboration with Yuri Dokshitzer and Stefano Catani; further details will eventually appear in Ref. [4]. I would like to acknowledge discussions on related topics with Steve Ellis, Mike Seymour and Zoltan Kunszt, and to thank the organizers of this meeting for making it so pleasant and stimulating.

References


