H1-01/05-616

An energy flow algorithm for Hadronic Reconstruction in OO: Hadroo2

MATTI PEEZ[a], BENJAMIN PORTHEAULT[b], EMMANUEL SAUVAN[c]

a,c: CPPM Marseille, b: LAL Orsay

Abstract

This note describes in detail the algorithm performing the identification and measurement of the Hadronic Final State in the OO framework, and the corresponding high P_T jet calibration procedure.

[a]matti@cppm.in2p3.fr. Now at UAM, Madrid
[b]portheau@lal.in2p3.fr
[c]sauvan@cppm.in2p3.fr

1 Introduction

One of the tasks of the OO framework is to provide to the user reconstructed particles at the μODS level. To achieve this goal physics algorithm now benefit of all the expert knowledge integrated during the HERA I operating phase. Particle identification is made by a set of different finders running sequentially, namely the electron finder, the muon finder, and the hadronic final state (HFS) finder. Additional finders may create so-called composed particles, such as the D^* finder or the jet finder. The input of the jet finder for the "exclusive jets" is the output of the HFS finder.

An energy flow algorithm is characterised by the combination of information coming from different sub-detectors. Following the guideline of a general improvement in the measurement of physical quantities with the H1 detector, the HADROO (for Hadronic Reconstruction in OO) algorithm was developed by M. Peez and C. Vallée [1], introducing the idea of using either the track or the calorimetric information for the creation of a particle candidate, depending on the error of the track measurement. This was the first step toward an energy flow algorithm.

This implementation was then refined, including also a better calorimeter noise rejection and an absolute calibration of the hadronic final state, based on reconstructed jets and suited for high Q^2 analyses. This actual implementation, called Hadroo2, will be described in this note.

The minor conceptual difference between Hadroo2 and a so-called energy flow algorithm — such as for example the D0 one [2] or the ZEUS one [3] — is that a one-to-one attribution of a cluster to a track is not performed[1]. In this view, it is an inclusive oriented algorithm, however it suits also to exclusive analyses because of the detailed track treatment.

This note will be organised in the following way: first, a description of the basic inputs of the algorithm (tracks and clusters) will be done. Particularly, the noise treatment applied to calorimetric objects will be detailed. Then the algorithm itself will be described, and compared with other HFS algorithms developed in H1. Its application to the reconstruction of high P_T jets and a suited calibration procedure is developed in the last part of this document.

2 Selection of the input objects: Tracks and Clusters

2.1 Tracks

As the spirit is to benefit from expert knowledge, the tracks used are the standard "good quality" tracks as defined by the heavy flavour group, the so-called "Lee West" tracks [4]. These tracks, measured with the central and forward tracking detectors (see Fig. 1), are classified in three categories, Central, Combined and Forward, requiring the quality cuts detailed in Table 1. If a track satisfies several sets of cuts, the preference order is Central, Combined, Forward. Both primary and secondary vertex fitted tracks can be selected but preference is given to primary. For HERA II data and MC, pure forward and combined tracks are excluded because their kinematics as well as their error measurement are at the moment not well studied and described. The selected tracks build up the input of the Hadroo2 algorithm.

[1] This problem is not trivial because of combinatorial ambiguities and its difficulty depends crucially on the features of the clustering algorithm.

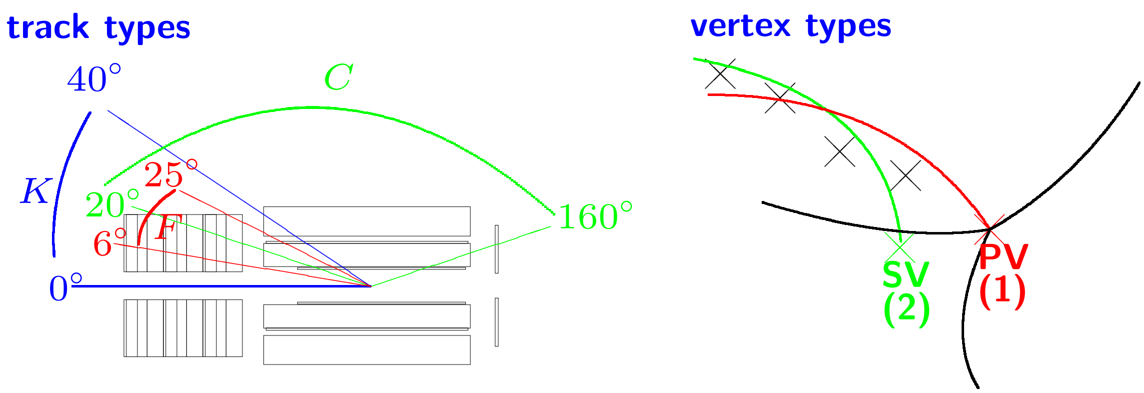


Figure 1: Different track types and their angular domain, and the two vertex hypothesis for a single track. Both primary and secondary vertex fitted tracks can be selected.

combined (K)	**central (C)**
$p_T > 120$ MeV $0° < \theta < 40°$ $\lvert dca' \rvert \leq 5$ cm $R_{start} \leq 50$ cm $R_{length} \geq 0$ cm $\Delta p/p \leq 99999.9$ $N_{CJC\ \ hits} \geq 0$ $\chi^2_{track-vertexfit} \leq 50$ $\chi^2_{cent.-fwd.tracker} \leq 50$	$p_T > 120$ MeV $20° < \theta < 160°$ $\lvert dca' \rvert \leq 2$ cm $R_{start} \leq 50$ cm $R_{length} \geq$ 10cm for $\theta \leq 150°$ $R_{length} \geq$ 5cm for $\theta > 150°$ $N_{CJC\ \ hits} \geq 0$
forward (F)	
$p_T > 1$ MeV $6° \leq \theta \leq 25°$ $R_0 \leq 10$ cm $\chi^2_{trackfit}/NDOF \leq 10$ $\chi^2_{track-vertexfit} \leq 25$ $nPrimary + nSecondary$ $PlanarSegments \geq 1$ $nPlanar + nRadialSegments \geq 2$ $\Delta p/p \leq 9999.9$ $p \geq 0.5$ GeV	

Table 1: Summary of the different cuts used in the track selection. If a track satisfies several set of cuts, the preference order is Central, Combined, Forward.The dca is the distance of closest approach of the track extrapolation to the vertex and dca' is the distance of closest approach in the x,y plane at $z = z_{vertex}$.

2.2 Clusters

The clusters are aligned and beam tilted in a proper way using run-dependent alignment factors. Calorimetric clusters are made only out of LAr or SpaCal. Iron or Plug cluster are not considered (beside the mediocre energy resolution of the Iron calorimeter, a significant fraction of Iron clusters are noise or background[2]). If a cluster in LAr have cells in Iron or Plug, these cells are removed from the cluster. Note that the negative energy cells in clusters are kept, as it must be to avoid a systematic positive bias in energy measurements.

The cluster energy momentum four-vector is made of the addition of massless cells four vectors (in this way clusters acquire a “mass”). The position of the center of gravity is determined with a linear energy weighting of the cells positions.

2.2.1 Weighting

As the LAr calorimeter has the well-known behaviour of being non-compensating, weighting algorithms are necessary to compensate the lower response to hadrons with respect to electron for a same energy [5]. Such a weighting procedure is already applied at the reconstruction level, in H1REC, identifying clusters as originating from electromagnetic particles or from hadrons.

But in the present algorithm this classification was modified. All clusters with at least 95 % of their energy in electromagnetic part and with also 50 % of it in the first two layers of the electromagnetic calorimeter are taken at the electromagnetic scale. All other clusters are considered as originating from hadrons and the hadronic energy scale, determined by the H1REC weighting algorithm, is considered. It was shown by S. Hellwig and K. Daum that this improves the energy resolution in D^* analysis and that the total reconstructed energy of the HFS was closer to the true level [6].

2.2.2 Noise suppression

The default situation All the measurement relying on the LAr calorimeter are affected by a relatively large amount of noise (few GeV per event). This noise is due to detector effects such as noise in the electronics or pile-up deposition of energy coming from non ep physics like halo or cosmic muons. The impact of this noise on physics analysis is clearly not negligible. For an inclusive analysis, the distribution

$$y_h = \frac{\sum_h E_h - Pz_h}{2E_0} \tag{1}$$

is specially affected. At low y_h (when $E_h \sim Pz_h$) most of the hadrons are produced in the forward direction. Any noisy cluster misidentified as part of the hadronic final state will count in the sum of Eq. (1) with a weight increasing with θ. So even relatively low energy noisy clusters in the barrel part of the LAr will strongly bias the y_h distribution. This situation is depicted in Fig. 2 where the different contribution to the distorsion of the measurement of y_h are depicted. Two first sources of bias in the measurement of y_h are the misidentification

[2]Note that the inclusion of tail catcher clusters with connected activity in the LAr could help to improve the determination of the energy of high P_T jets. This study is therefore planned for further developments of the HFS finder

of part of the energy of the scattered electron as hadrons[3] and the presence of photons due to QED initial state radiation. These two contributions are discussed in details in [11] and will be explicitely removed in the present study. After the removal of such events, the remaining bias introduced by noise can be observed in Fig. 2. On the right plot the difference between the reconstructed and the true value of y_h as a function of the true y_h shows that the events with $y_h \sim 10^{-2}$ have a systematic bias of the order of 60 %, even after the application of noise suppression at the reconstruction level.

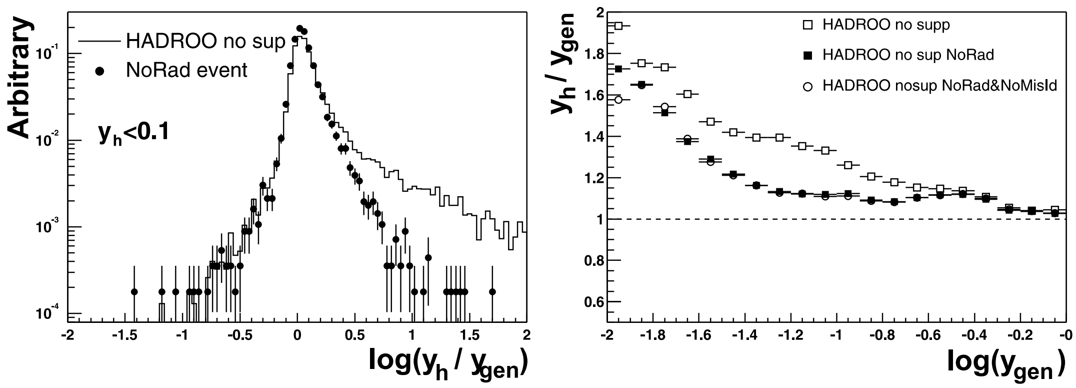


Figure 2: Comparions between the reconstructed and true values of y_h using a neutral current Monte Carlo event sample. The left figure shows the effect of the radiative NC events on the y_h/y_{gen} distribution (Note the logarithmic scales). The right figure shows the mean of y_h/y_{gen} as a function of y_{gen} and the effect of removing explicitely contributions due to radiative events (labeled "NoRad"), and misidentification of the scattered electron ("NoMisId").

Beside topological background finders [9, 10] dedicated to the rejection of an entire event which does not originate from an ep collision, the noise suppression algorithms described here are designed to remove the unphysical clusters while keeping the event. They are specially tuned to remove the previously described high θ background. Note that before all this there is already the so-called topological noise suppression (ETNS) (see Ref. [5]) which is applied at the reconstruction level (H1REC). In Monte Carlo noise is added on top of the simulated energy deposit. This noise come from real data taken during dedicated random trigger runs.

Noise suppression strategy First, all one-cell clusters are considered as not physical and removed, as well as clusters with energy $E_{clu} < 0.2$ GeV in LAr or 0.1 GeV in SpaCal. Then a set of background finders (as developed in [11]) are applied. Now these finders will be described and their performance studied.

[3]The imperfect cluster algorithm can give rise to multiple clusters for the scattered electron in particular when it hits Φ crack between octants

2.2.3 The FSCLUS algorithm

The principle of the FSCLUS algorithm, noise suppression inherited from a fortran algorithm, is the suppression of low energy isolated clusters. If the energy E_{clu} of a cluster is such that $E_{clu} < E_1$ the energy E_{sphere} in a sphere of radius R around the cluster is computed and if $E_{sphere} < E_2$ then the cluster is suppressed. This allows low energy cluster to survive if they are near more energetic ones *e.g.* if they are due to a shower fluctuation. The values for the different thresholds are $E_1 = E_2 = 0.4$ GeV and $R = 40$ cm for $\theta_{clu} > 15°$, $E_1 = E_2 = 0.8$ GeV and $R = 20$ cm for $\theta_{clu} < 15°$. Consistently with the first suppression the threshold E_1 is lowered to 0.2 GeV for clusters in the electromagnetic part of LAr. Clusters near the beam pipe in the SpaCal calorimeter are also suppressed if $\sqrt{x_{clu}^2 + y_{clu}^2} < 9.6$ cm. The performance of the FSCLUS algorithm is shown in Fig. 3: the bias is reduced by 20 % and is now at a level of 40 %. So this noise suppression is clearly not efficient enough and has to be combined with other algorithms.

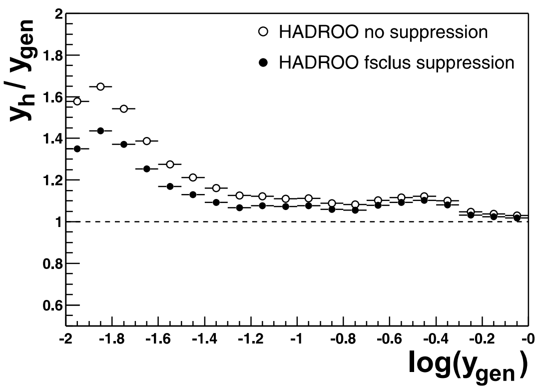


Figure 3: Mean of y_h/y_{gen} distribution as a function of y_{gen} (Note the logarithmic scales). The situation before and after the application of FSCLUS is depicted by open and solid circles, respectively. Neutral current events from a Monte Carlo sample have been used.

2.2.4 The HALOID algorithm

The HALOID algorithm is devoted to the suppression of energy deposit due to halo muons on top of real physics events. The signature is a narrow energy deposit parallel to the beam axis. To suppress such a pattern, for each cluster it is defined two cylinders of radius $R_1 = 25$ cm and $R_2 = 65$ cm. If there is energy deposit in these cylinders in at least 4 wheels including 2 CB wheels, and if at least two of the following criteria are true:

$$E_{cylinder\,1} \geq 0.5 E_{cylinder\,2} \tag{2}$$

$$N_{clusters\,in\,cylinder\,1} \geq 0.5 N_{clusters\,in\,cylinder\,2} \tag{3}$$

$$N_{cells\,in\,cylinder\,1} \geq 0.5 N_{cells\,in\,cylinder\,2} \tag{4}$$

the cluster is flagged as noise and suppressed. The improvement in the measurement of y_h before and after the suppression is shown in Fig. 4 based on a charged current MC sample.

There is a clear improvement in the y_h reconstruction of these very biased events.

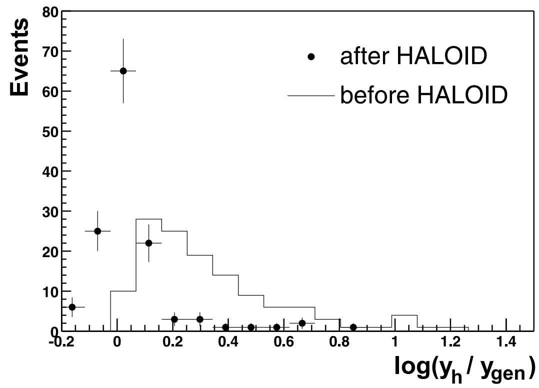


Figure 4: Improvement in y_h reconstruction after the HALOID algorithm for charged current MC events in which there is an overlap halo muon.

2.2.5 The HNOISE algorithm

Contrary to halo muons, cosmic muons or coherent noise do not have a characteristic pattern of energy deposit. However, on general ground, any deposit in the hadronic part of LAr should be connected to activity in the electromagnetic part or linked with tracks. The HNOISE algorithm look for clusters in the hadronic part and suppress them if the following conditions are all fulfilled:

- There is no energy deposit in the first hadronic layer or there is energy deposit in the first hadronic layer and there is no more energetic clusters at a distance less than 75 cm.
- There is no electromagnetic energy in a safety cylinder of 50 cm radius. The axis of this cylinder is defined by the interaction vertex and the barycenter of the considered cluster.
- There is no vertex fitted track with a dca of less than 50 cm.

This finder help again to remove a part of the noise, as shown in [11] . However there is still noise contribution at large angles leading to a bias in the y_h distribution. The NEWSUP algorithm is designed to remove this remaining background.

2.2.6 The NEWSUP algorithm

The NEWSUP algorithm is inspired form FSCLUS: it is designed to suppress low energy isolated clusters. However, to remove completely the noise a threshold higher than previously applied is needed, but only in the central region of LAr where the $E - Pz$ contribution of a false particle candidate biases the y_h of the event by a large value. Contrary to FSCLUS this algorithm care about track-cluster link and if there is a vertex fitted track with $dca \leq 25$ cm for an electromagnetic cluster or $dca \leq 50$ cm in the hadronic part the corresponding cluster is not suppressed. The same thresholds as in FSCLUS are applied, except that now $E_1 = E_2 = 1.5$ GeV for $\theta > \alpha_h$. The angle α_h is chosen to be the maximum between the angle of the most backward track and the inclusive hadronic angle[4] $\tan(\gamma_h/2) = (E_h - Pz_h)/P_T^h$. The algorithm is run iteratively until there is no cluster suppressed. The results are presented in Fig. 5 where situation for charged current events is depicted. The energy reconstructed at

[4]If $\gamma_{h\,LAr} \leq 50^o$ or $y_{h\,LAr} \leq 0.1$ the SpaCal clusters do not enter in the calculation of γ_h.

high angles is much greater than the generated one without any noise suppression. After the application of all suppression algorithms developed here the measurement is closer to the true level.

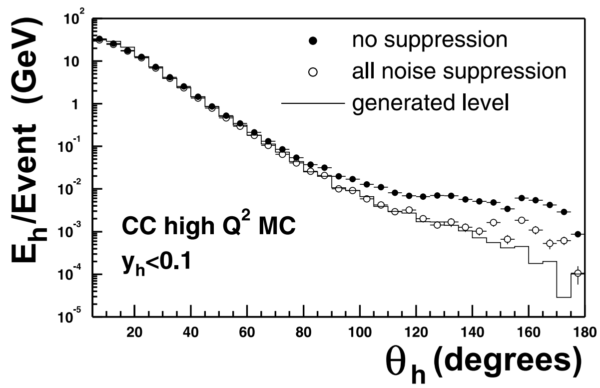


Figure 5: Comparison between the total reconstructed and generated energy distributions for low y_h events with all noise suppressions (open circles) and without (solid circles). The true level is represented by an histogram. Charged current events from a Monte Carlo sample have been used.

All the previous discussions were purely based on MC files. We have now to check that all the noise suppression is applicable to real data. This is done in Fig. 6 where the energy fraction suppressed from data and MC are compared as a function of y_h. A good agreement is observed and the amount of suppressed noise energy is comparable to the previous FORTRAN implementation of the algorithms, as presented in [11]. The conclusion is that the combination of these complementary noise finders allows a good reconstruction of the y_h kinematic variables. A good suppression is also very important in the views of a calibration procedure aiming at the knowledge of the true energy.

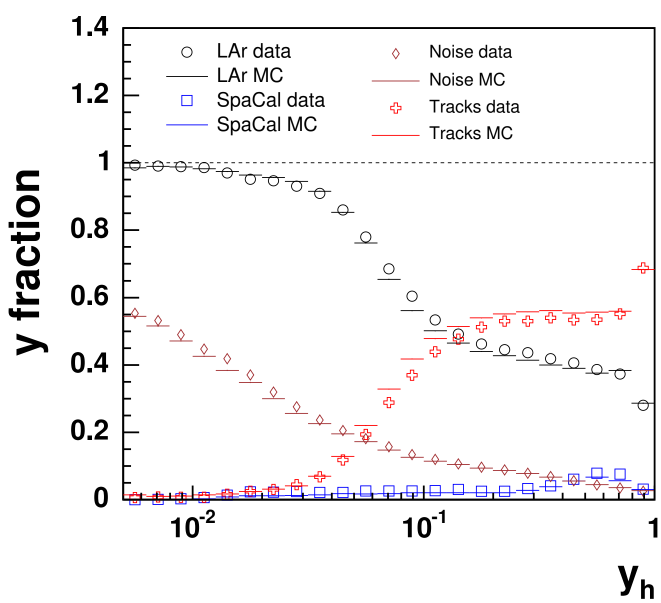


Figure 6: Comparison of the fraction of the suppressed noise energy in data (points) and in MC (histograms). The other components contributing to the final y_h are also displayed. 2003-2004 data and Django MC have been used here.

2.2.7 Safety tests of the noise finders

As these finders are applied by default, careful studies have been performed to see if signal relevant for exclusive studies were not suppressed. Tests were made on different MC samples, namely D^* events in photoproduction and in DIS, and diffractive J/Ψ events.

The principle of the tests was to look if these additional noise suppressions was killing genuine signal. The distance in the η, φ plane between a generated particle (with $P_T^{gen} > 180$ MeV, so that it reaches the calorimeter) and each particle candidate was computed:

$$d = \sqrt{(\eta_{gen} - \eta_{cand})^2 + (\varphi_{gen} - \varphi_{cand})^2}. \quad (5)$$

The minimal distance is supposed to give the corresponding candidate associated to the generated particle. By looking at this minimal distance before noise suppression d_{nosup} and after noise suppression d_{sup} we can see if signal has been suppressed.

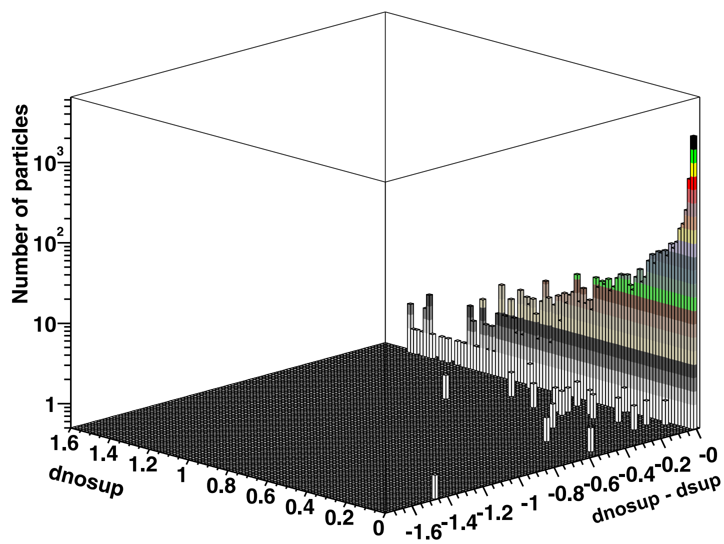


Figure 7: Numbers of entries regarding the minimal distance in η, φ between a generated particle and a particle candidate before d_{nosup} and after d_{sup} noise suppression, in the D^* photoproduction event sample.

The figures 7 and 8 are two dimensional histograms of d_{nosup} versus $d_{nosup} - d_{sup}$. It is straightforward to see that first, most entries are concentrated at $d_{nosup} - d_{sup} = 0$, so the noise suppression was safe, and in the $d_{nosup} - d_{sup} = 0$ plane the region $d_{nosup} \simeq 0$ dominate, so the generated particle was correctly matched to a candidate. The region to look for signal suppressed is the region of $d_{nosup} \simeq 0$ (the particle is well associated to a generated one) and $d_{nosup} - d_{sup} < 0$. We see two such entries on the histogram of Fig. 7, at $d_{nosup} - d_{sup} = -0.5$ and -1.5. The first one is a π^- killed by the NEWSUP algorithm, and the second one a n killed by the HNOISE algorithm. For the histogram of Fig. 8, five particles (two γ, two K_L^0 and a n) are found to be suppressed, this being mainly due to the special topology of photoproduction charm events with a large number of very low energy particles in a large η range. A conclusion can be drawn by looking at the Table 2 where one can see that the loss of genuine signal is at a very low and acceptable rate with respect to the signal suppressed.

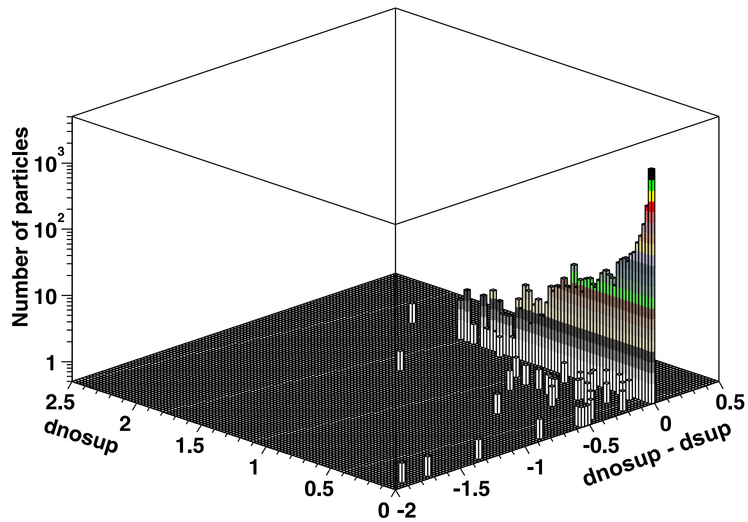


Figure 8: Numbers of entries regarding the minimal distance in η, φ between a generated particle and a particle candidate before d_{nosup} and after d_{sup} noise suppression, in the D^* low Q^2 DIS event sample.

D^* photoproduction sample		Inefficiency
2 signal killed	717 newsup clusters	0.2 %
3 signal killed	114 hnoise clusters	2.6 %
D^* low Q^2 DIS sample		Inefficiency
1 signal killed	562 newsup clusters	0.2 %
1 signal killed	140 hnoise clusters	0.7 %

Table 2: Summary of the signal suppression.

For sake of completeness a test was made on a diffractive J/Ψ sample. It was found that two μ with no tracks associated and not found by the muon finder were suppressed by HNOISE, this for 10^4 J/Ψ events. So the noise suppression is clearly safe for diffractive vector meson production.

To conclude the noise finding achieves a very good compromise between efficiency and safety. The output list of noise suppressed clusters is the input of the Hadroo2 algorithm which is now going to be described in detail. Every noise suppressed cluster is flagged and re-used to calculate the total hadronic LAr four-vector produced by noisy cells. This information is stored on HAT in the variables which start with the string 'HfsClusNoiseXXXX' and allows then to study the impact of the different noise finders.

3 The Hadroo2 algorithm

The Hadroo2 algorithm realises the creation of the HFS particles. Note that if there are identified electrons or muons which are not flagged as isolated[5], they are considered as being part of the Hadronic Final State but their four vector remains unchanged and their associated tracks and clusters are excluded from any additional treatment.

The algorithm starts with the previously described list of selected tracks and clusters. The cornerstone idea of the energy flow algorithm is the combination of the tracks and clusters. As we may have both for a charged particle, we want to keep the best measurement. To achieve this, we propose to compare relative resolutions of the tracker or of the calorimeter for the measurement of the same amount of energy.

3.1 Comparison of tracker and calorimeter resolutions

Each track is supposed to originate from a pion, with energy

$$E^2_{track} = P^2_{track} + m^2_\pi = P^2_{T,track}/\sin^2\theta + m^2_\pi. \tag{6}$$

The error on this energy is obtained by standard error propagation using some of the track fitting error information:

$$\frac{\sigma_{E_{track}}}{E_{track}} = \frac{1}{E_{track}}\sqrt{\frac{P^2_{T,track}}{\sin^4\theta}\cos^2\theta\sigma^2_\theta + \frac{\sigma^2_{P_T}}{\sin^2\theta}} \tag{7}$$

where σ_{P_T} and σ_θ are the corresponding error on P_T and θ and neglecting their correlations. It was checked that the use of the full covariance matrix gave similar results within 2 % at most.

Now we evaluate for each track what would be the corresponding error of this particle as measured with the calorimeter. This decision turns out to be only based on the track, but it is not possible to make any decision based on the calorimeter deposit as this one is a priori unknown due to possible contribution of neutral particles. We made the assumption that the corresponding error on the measurement of this particle in the LAr [5] would lead to the error $\sigma_{E\,LAr\,expect.}$

$$\left(\frac{\sigma_E}{E}\right)_{LAr\,expectation} = \frac{\sigma_{E\,LAr\,expect.}}{E_{track}} = \frac{0.5}{\sqrt{E_{track}}}. \tag{8}$$

The relative resolutions defined by Eqs. (7) and (8) are then compared to determine which of the tracker or the calorimeter provides the best measurement. The track is considered as a “good one” if

$$\frac{\sigma_{E_{track}}}{E_{track}} < \frac{\sigma_{E\,LAr\,expect.}}{E_{track}} \tag{9}$$

The Fig. 9 shows the relative resolutions of the track compared to the LAr expectation. We observe that the tracker measurement is better up to 12 GeV for forward tracks, 25 GeV for

[5] A muon is isolated if the calorimeter energy in a cylinder around the extrapolated muon track is < 5 GeV (cylinder radius of 35 cm in electromagnetic, 75 cm in hadronic LAr section) and if there is no other selected track in a cone of radius $R_{\eta-\varphi} = 0.5$.

An electron is isolated if the calorimeter energy not attributed to any other identified electron in a cone around the electron of radius $R_{\eta-\varphi} = 0.5$ is less than 3 % of the electron energy. All SpaCal electrons are considered as isolated.

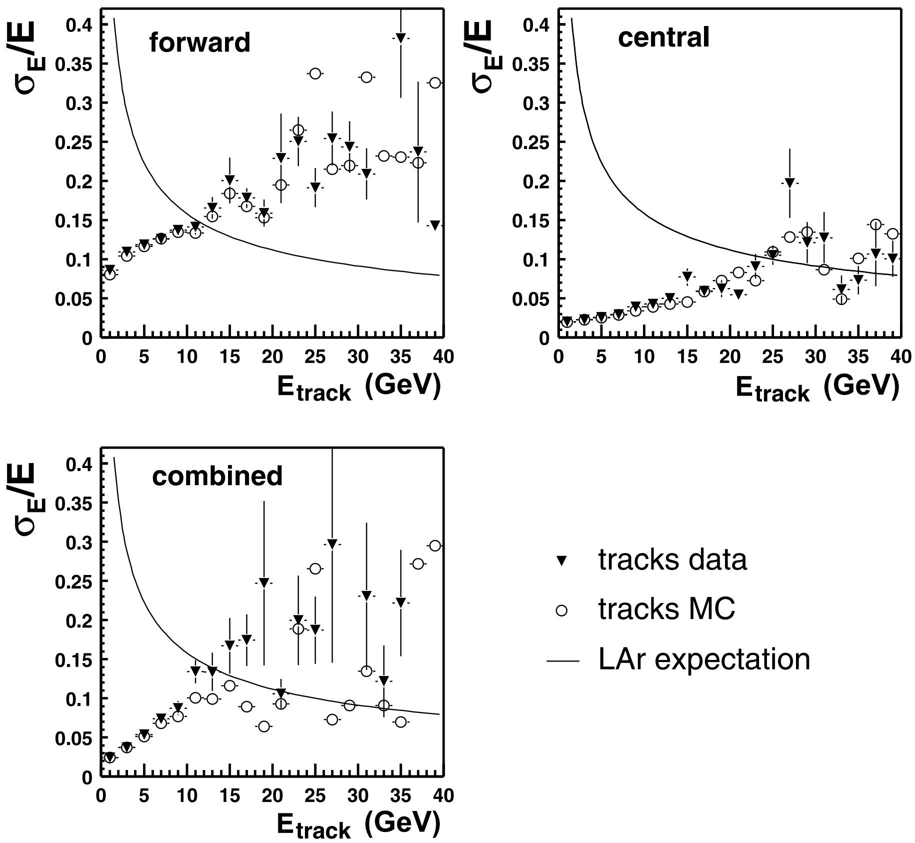


Figure 9: Relative resolution of the different types of tracks compared to the LAr expectation.

central tracks and about 13 GeV for combined tracks. We also observe that the error of the track measurement is reasonably well described by the MC, at least up to the turnaround energy.

To also optimise the global energy measurement, selected charged tracks are ordered by increasing P_T, in order to associate first the clusters to the well measured low P_T tracks. Then the algorithm do a loop over selected tracks and for each track test the Eq. (9) and try to associate calorimetric clusters to the track.

3.2 Track measurement preferred

If Eq. (9) is true, the track measurement is used to make a particle candidate. In this case the calorimetric energy has to be suppressed to avoid double counting. Each track is extrapolated up to the surface of the calorimeter as an helix, and inside LAr as a straight line. The calorimetric energy $E_{cylinder}$ is computed as the sum of all clusters in the overlapping volume of a 67.5° cone and two cylinders of radius 25 cm in the electromagnetic part of LAr and 50 cm in the hadronic part (see Fig. 10). This volume will be referred hereafter as the "cylinder". The numerical values are such that the cylinder reasonably contains the full hadronic shower. Small variations of these values do not lead to significant changes in the performance of the algorithm.

Then the track energy E_{track} is compared to the calorimetric energy inside the cylinder

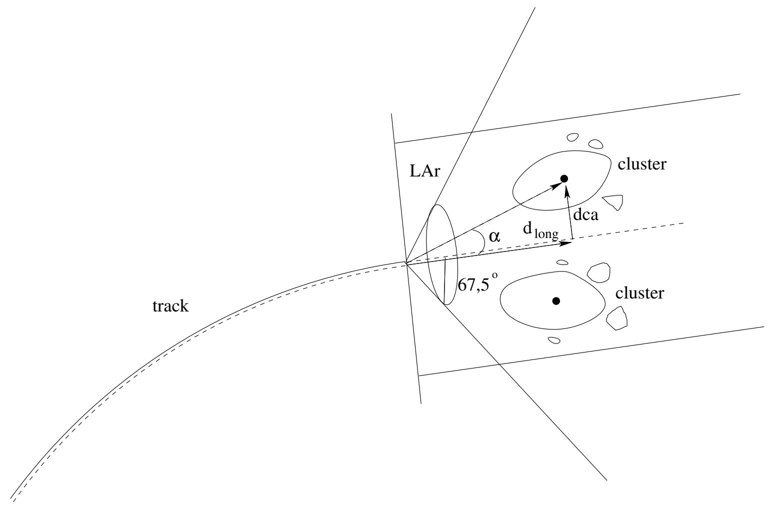


Figure 10: The axis of the cone and the cylinders is the straight line extrapolation of the particle trajectory into the calorimeter. The distance of closest approach (dca) of a cluster is defined with respect to this line. This drawing is courtesy of A. Perieanu.

$E_{cylinder}$, taking into account possible fluctuations of both measurement within their standard errors[6] and if

$$E_{cylinder} < E_{track} \times \left[1 + 1.96 \sqrt{ \left(\frac{\sigma_{E_{track}}}{E_{track}} \right)^2 + \left(\frac{\sigma_E}{E} \right)^2_{LAr\, expectation} } \right] \tag{10}$$

an amount of calorimetric energy $E_{suppressed}$ equal to $E_{cylinder}$ has to be suppressed completely. Otherwise only an amount of energy $E_{suppressed} = E_{track}$ is suppressed. Clusters are suppressed one after the other by increasing dca and up to the needed energy. To reach the exact $E_{suppressed}$ energy, some clusters may be only partially removed and their energy is adjusted.

The meaning of Eq. (10) is the following: the calorimeter measurement may have fluctuated, but the well measured track give a constraint on the amount of energy coming from charged particles; so we discard all the calorimeter measurement except if the observed fluctuation is above 95 % C.L. of the error. If Eq. (10) is false the energy difference $E_{cylinder} - E_{track}$ is assumed to originate from neutral particles or other charged tracks. So it is a way of deciding whether there is additional energy not belonging to the primary track or not without always believing the upward fluctuations of the LAr energy measurement.

3.3 Calorimetric measurement preferred

If Eq. (9) is false then the energies $E_{cylinder}$ and E_{track} are compared and if

$$E_{track} \in \left[E_{cylinder} - 1.96\, \sigma_{E_{cylinder}}, E_{cylinder} + 1.96\, \sigma_{E_{cylinder}} \right] \tag{11}$$

[6] This feature of the algorithm was suggested by K. Daum

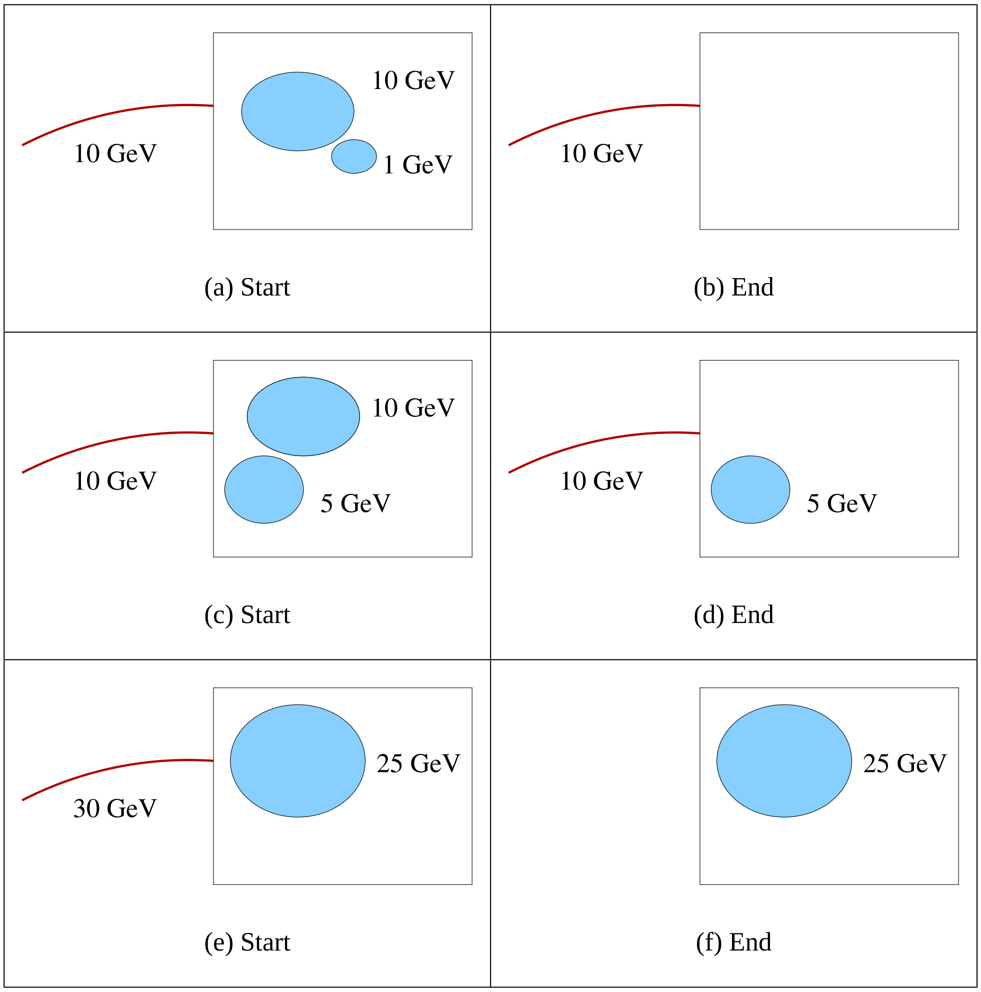


Figure 11: Example: behaviour of the Hadroo2 algorithm given three starting situations involving tracks and clusters. On the first line, a 10 GeV track measured with a 4 % accuracy is kept (Eq. (9)) and all the calorimetric information is removed (Eq. (10)). On the second line the track information is still kept, however the cylinder energy of 15 GeV is determined to contain a neutral component (following the Eq. (10)) and only the track energy is subtracted. On the third line the track is not well measured (15 % accuracy) and the calorimetric information is used.

(with $\sigma_{E_{cylinder}} = 0.5\sqrt{E_{cylinder}}$) the track energy is considered to be compatible with the calorimetric deposit and the calorimetric measurement is used to define a particle candidate. Otherwise, if

- $E_{track} < E_{cylinder} - 1.96\,\sigma_{E_{cylinder}}$, the track measurement is used and calorimetric energy is subtracted as in Sec. 3.2.

- $E_{track} > E_{cylinder} + 1.96\,\sigma_{E_{cylinder}}$, the track is suppressed and an hadron is defined using the calorimetric clusters[7].

Indeed, when the compared energies are compatible, the hadrons are well measured but most of the time the measurement of the calorimeter is more accurate. When the track energy is much larger than the calorimeter energy, it is most of the time due to a bad measurement of an high P_T track.

3.4 Treatment of residual clusters

Once all the tracks have been treated, particles candidates are made out of remaining clusters using the calorimetric energies. The momentum of these clusters is rescaled to obtained massless particles. Thes particles correspond to neutral hadrons with no associated track or to charged particles with a badly measured track.

4 Comparison with other HFS algorithms

The main kinematic variable used in the next sections are defined using the hadronic and double-angle methods. The total hadronic transverse momentum P_T^h is defined by

$$P_T^h = \sqrt{\left(\sum_h P_x^h\right)^2 + \left(\sum_h P_y^h\right)^2} \tag{12}$$

where the summation h extends over all reconstructed hadrons at the μODS level. The quantity θ_h and θ_h^e stand for the hadronic inclusive polar angle calculated respectively with the hadronic and positron variables using:

$$\tan(\theta_h/2) = \frac{\sum_h (E_h - Pz_h)}{P_T^h} \tag{13}$$

and

$$\tan(\theta_h^e/2) = \frac{2E_0^e - (E_e - Pz_e)}{P_T^e} \tag{14}$$

where E_0^e, E_e, Pz_e and P_T^e are respectively the energy of the incident positron, the energy, the longitudinal and transverse momenta of the scattered positron.

The total transverse momentum P_T^{da} is calculated using the double angle method from the angles of the positron and of the hadronic system:

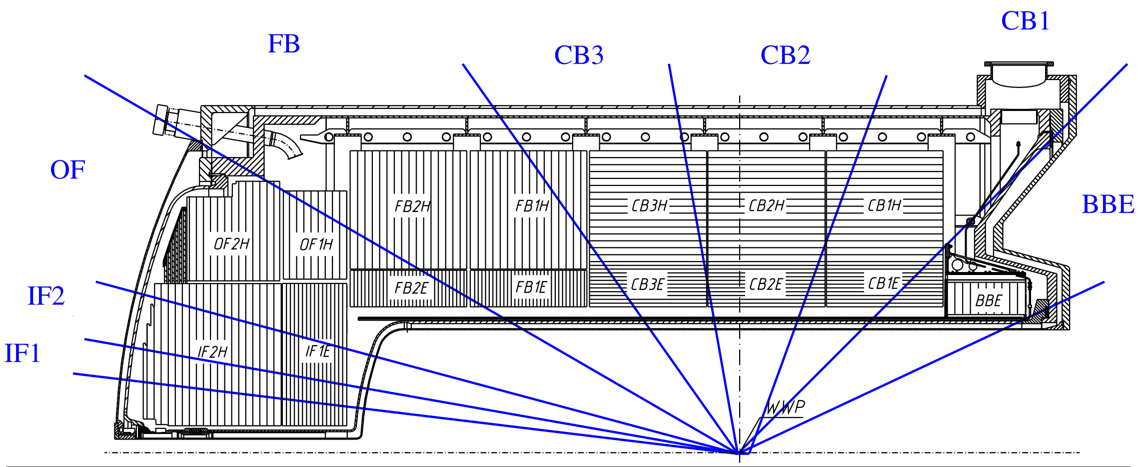
$$P_T^{da} = \frac{2E_0^e}{\tan\frac{\theta_e}{2} + \tan\frac{\theta_h}{2}}. \tag{15}$$

The P_T balance P_T^{bal} stands for the ratio of the hadronic transverse momentum and the double angle transverse momentum:

$$P_T^{bal} = \frac{P_T^h}{P_T^{da}}. \tag{16}$$

[7]Note that technically the four-vector of the particle candidate associated to the track is changed using the calorimeter informations. Only if there is no calorimetric energy behind the track, its particle candidate four-vector is set to zero.

In this note the following area depicted on Fig. 12 will be used. It is an angular division roughly named after the corresponding calorimeter wheels.



IF1	IF2	OF	FB	CB3	CB2	CB1	BBE
7°-10°	10°-15°	15°-30°	30°-55°	55°-80°	80°-110°	110°-135°	135°-155°

Figure 12: Definition of the different areas of the LAr calorimeter defined for the HFS calibration.

In the past, several other approaches for the reconstruction of the hadronic final state have been used. For inclusive high Q^2 analysis, algorithm using only calorimeter information has been used — this is referred here as the "clusters only" algorithm. This suffers from the drawback that low P_T tracks component is not included in the HFS reconstruction and is therefore missing.

The widely used FSCOMB algorithm has been one of the first attempts to combine tracks and clusters, but tracks were only considered up to[8] a P_T of 2 GeV. In FSCOMB the subtraction also is done in such a way that only E_{track} is suppressed, and never $E_{cylinder}$. So there is no equivalent to Eq. (10) and the energy measurement of LAr is always trusted.

4.1 Composition in tracks and clusters of the HFS particles

Figure 13 shows the relative contribution of clusters to P_T^h for neutral current (NC DIS) events with only one jet and for data and MC events. The details of the event selection used here are given in Sec. 5.2. The fraction of tracks is then the complement to one of the cluster fraction presented in Fig. 13.

A clear pattern of dependencies upon P_T^{da} and θ_{jet} appears. We see that first, the contribution of tracks is decreasing when the transverse energy of the jet is increasing. This is consistent with the fact that more clusters are chosen at high energies. The main dependency is the one with respect to θ_{jet}. For $\theta_{jet} < 15°$ the forward and combined track contributions are rather low, and the HFS particles are clearly cluster-dominated by about 80 %. In the OF region the central track contribution starts to play a role and the cluster contribution decreases. At the end, the cluster contribution in the central region is around 40 %. We can also observe that

[8]When the algorithm was developed, high P_T tracks had not been extensively studied.

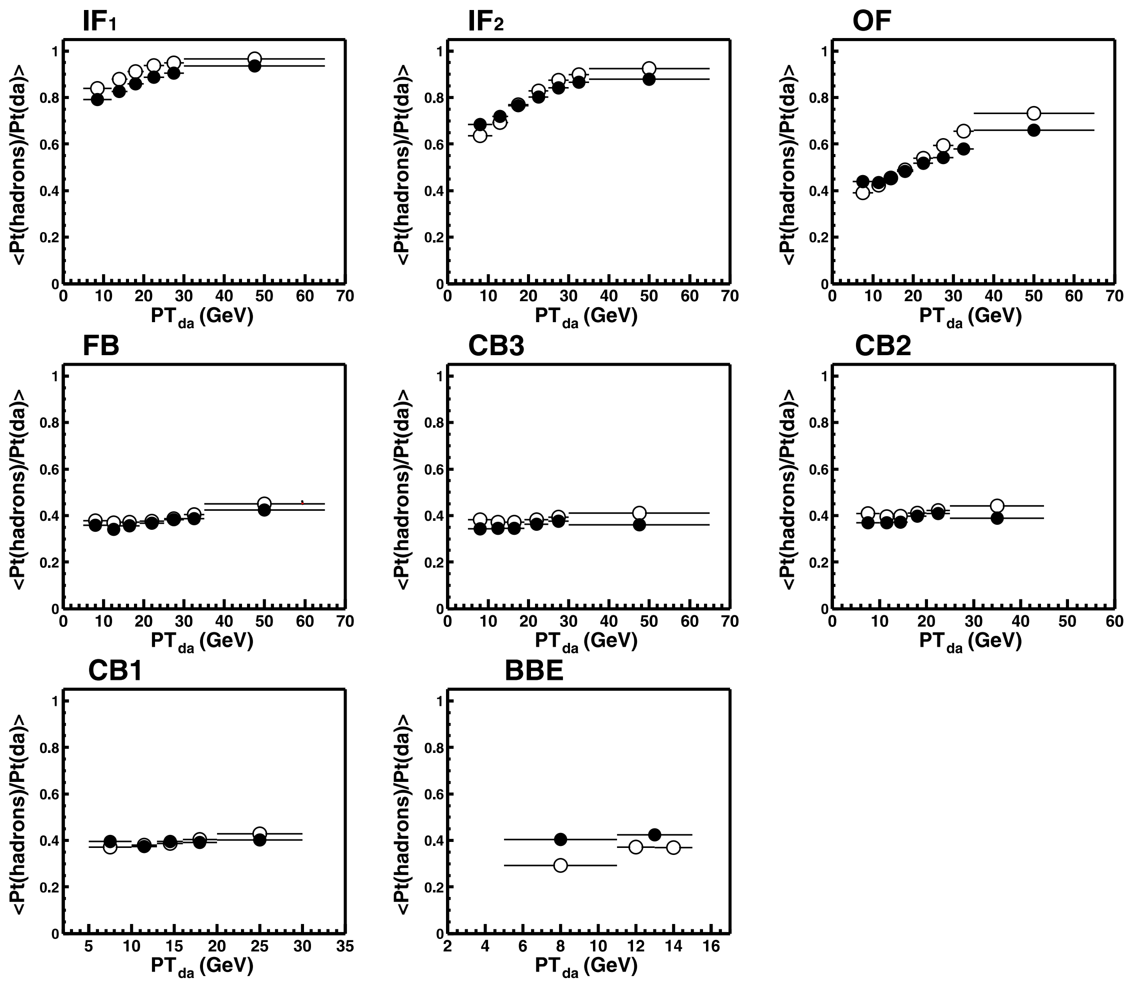


Figure 13: Fraction of the total hadronic P_T due to clusters for high Q^2 NC DIS one jet events, as a function of θ_{jet} and P_T^{da}. Data are presented using solid circles and MC with open circles.

the contributions of tracks and cluster in the data are reasonably well described by the MC, as expected from the good description of the track relative resolution as shown in Fig. 9.

Figure 14 is a comparison of the $P_T^{bal}(\theta_{jet}, P_T^{da}) = P_T^h / P_T^{da}$ for the same high Q^2 NC DIS one jet sample reconstructed with different algorithms. No cluster calibration is applied. In almost all calorimeter wheels, a 10 % shift of the P_T^{bal} mean values of the Hadroo2 algorithm is observed due to the different weighting scheme used (see Sec. 2.2.1). On this plot the contribution of tracks is clear, already for the OF region: the P_T^{bal} is much flatter with respect to P_T^{da}. In the IF1 and IF2 regions, where track contribution is negligible, all the P_T^{bal} have a similar shape.

4.2 Resolution

The evolution of P_T^{bal} distributions as a function of P_T^{da} and θ_{jet} for FSCOMB and Hadroo2 algorithms are compared in Fig. 15. The evolution of mean values of P_T^{bal} distributions are

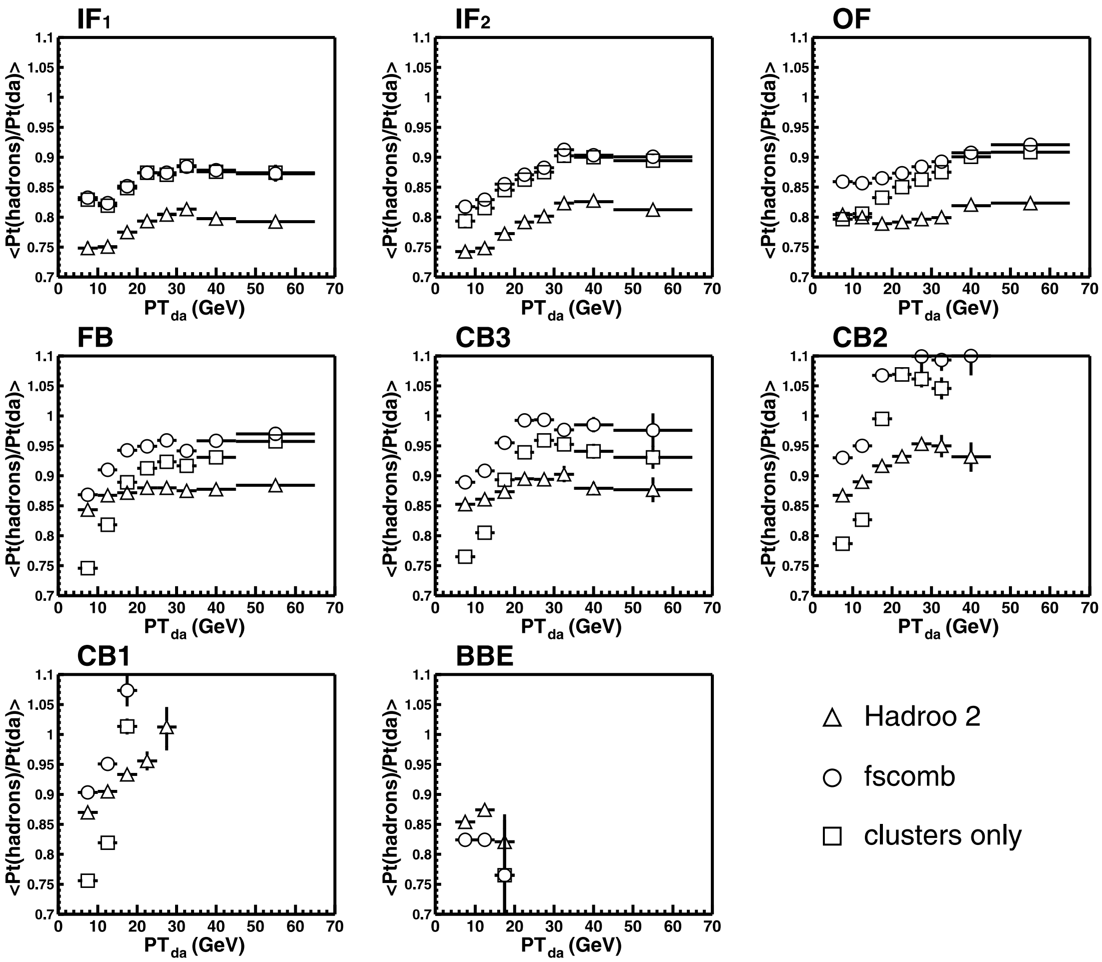


Figure 14: Comparison of the $P_T^{bal} = P_T^h / P_T^{da}$ dependency upon P_T^{da} for different HFS algorithms: FSCOMB, Hadroo2 and clusters only.

depicted on the left figures while the right figures present the evolution of the relative resolution, defined as $\sigma(P_T^{bal}) / < P_T^{bal} >$. NC DIS events with only one jets have again been used (see Sec. 5.2). The mean values corresponding to the Hadroo2 algorithm are shifted down by ~ 10 % due to the weigthing scheme used. Nevertheless the resolution obtained with the Hadroo2 algorithm is comparable to FSCOMB and even better in the backward region of the LAr calorimeter.

Further improvement of the resolution can be obtained by combining the Hadroo2 algorithm with the new energy weighting scheme for hadronic clusters proposed in [7, 8]. This is displayed by blue points on Fig. 15. We can observe that the resolution is improved in all regions of the calorimeter and especially in the central part of the barrel and for 10 GeV $< P_T^{da} < 20$ GeV. The evolution of the mean values of P_T^{bal} as a function of P_T^{da} is also flat using this new weighting. Therefore it is planned to include this new energy weighting in next developments of the HFS finder.

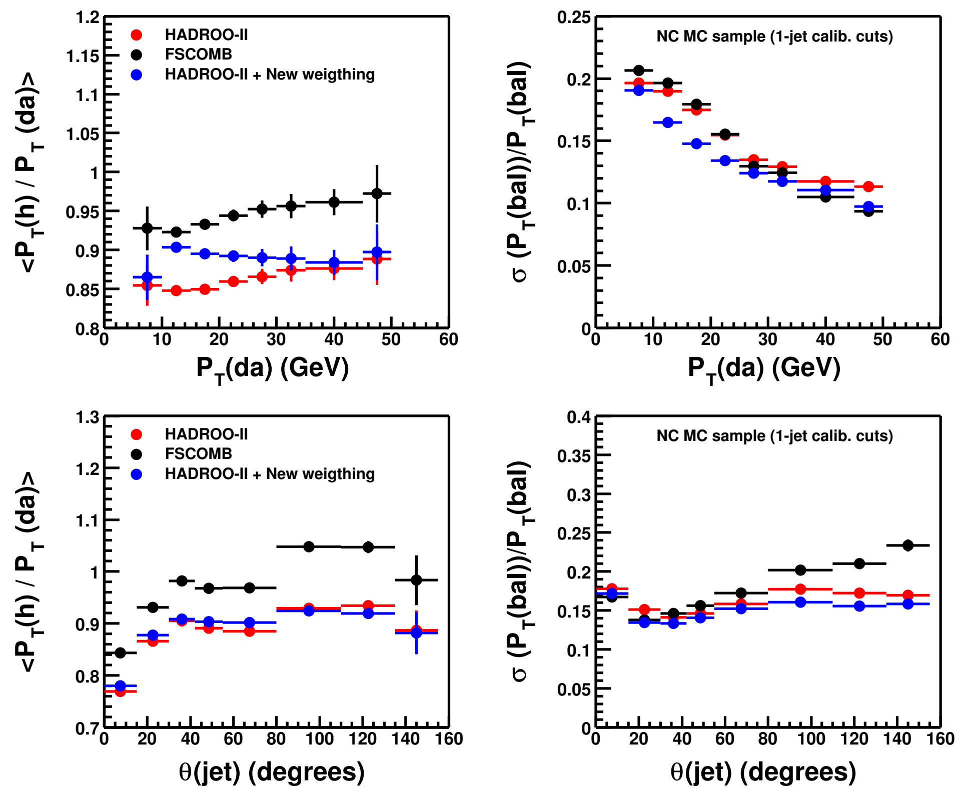


Figure 15: Evolution of mean values (left figures) and relative resolutions (right figures) of P_T^{bal} as a function of P_T^{bal} and θ_{jet}, for FSCOMB and Hadroo2 algorithms. The influence of the "new weighting" [7, 8] (blue dots) is also presented.

5 Jet Calibration procedure

This section is devoted to the calibration of jets for high Q^2 inclusive measurements. The knowledge of the absolute energy scale and its error is a key point for lots of analyses, ranging from searches and "exotic" analyses where we want to reconstruct an invariant mass, to jets and inclusive physics where the understanding of the error on the hadronic energy scale is crucial. This is especially true for the Charged Current analysis where all the kinematics variables are reconstructed using the HFS.

5.1 The principle of calibration

Once the hadron finding algorithm has been fully specified, a suited calibration procedure can be applied. The selected tracks are already calibrated and the calibration procedure must not change their energy. In figure 16 we observe that the K^0 mass peak obtained with the default *H1PartK0Finder* has an accuracy better than 1 %.

The aim is therefore to perform a jet calibration but only changing the energy of calorimeter clusters. The method of jet calibration used here is derived from [12]. The reference quantities used for the calibration are determined with the double angle kinematics. The hadronic trans-

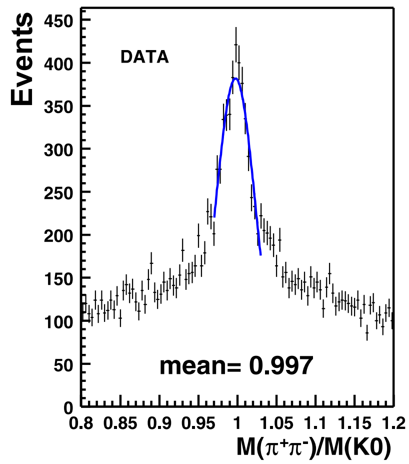


Figure 16: K0 mass spectrum (normalised to the K0 mass) obtained with the μODS standard *H1PartK0Finder*. The invariant mass of the two pions is computed using the good quality tracks. The mean of a Gaussian fit (blue line) is centered to one with less than 1 % deviation.

verse momentum determined with this method is independent of the LAr energy calibration to a good approximation. The calibration is said to be *absolute* if the measured P_T^h coïncides with the P_T^{da} (see Sec. 4 for definition of the variables). The use of the double angle method as a reference has several consequences: first, the calibration sample chosen to determine the calibration constants must be such that the P_T^{da} measurement is well under control. Secondly this method does not rely on MC which is separately calibrated and no relative calibration is needed. Finally the method is also independent of the electron calibration.

5.2 Determination of the calibration constants

The event sample used to determine the calibration coefficients is defined by the following selection:

- Good quality selection (High Voltage, Vertex, background finders, *etc*),
- $Q^2 \geq 100\, GeV^2$,
- 1 electron with $P_T^e \geq 10$ GeV,
- only one jet,
- good P_T^{da} measurement cuts:
  - Anti ISR cut $P_T^e / P_T^{da} > 0.88$
  - Anti leakage cuts : $E_{SpaCal}/E_{total} < 1$ %
  - $P_T^{SpaCal}/P_T^{total} < 1$ %
  - $E_{iron}/E_{total} < 1$ % or $P_T^{iron}/P_T^{total} < 1$ %
  - $d\theta = |\theta_{had} - \theta_{jet}| < 1.5$. This cut was shown to improve the double angle measurement at low P_T^{jet} and θ_{jet} (see [12]).

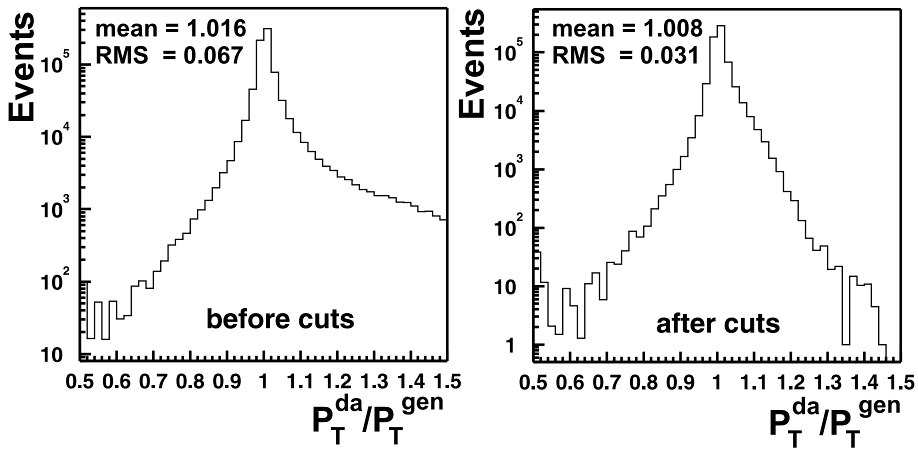


Figure 17: P_T^{da}/P_T^{gen} distributions before and after having applied the good double angle measurement cuts. The bias of P_T^{da} to higher values due to QED ISR is significantly reduced by these cuts.

The figure 17 shows the ratio P_T^{da}/P_T^{gen} before and after the good P_T^{da} measurement cuts for high Q^2 neutral current MC.

The improvement of the P_T^{da} measurement is clear, especially the bias of the P_T^{da} to larger values (due to initial QED radiation) is significantly reduced. Hence we can say that the double angle measurements are well under control. Note that no cut on the hadronic energy is used, because indeed such a cut would bias the distributions used to calibrate. In these one jet events the hadronic final state is entirely contained in a single material region of the LAr and we have an approximation of the difference between the true P_T of the jet (approximated as P_T^{da}) and the response (or lack of response) of the detector.

The evolution of mean values of P_T^{bal} distributions upon P_T^{da} — called F_{ptbal} — is fitted separately for several θ regions. The functional form used for the fit is

$$F^{\theta}_{ptbal}(P_T^{da}) = A_\theta(1 - \exp^{-B_\theta - C_\theta P_T^{da}}) \quad (17)$$

The P_T^{bal} distributions and the result of the fit are shown in Fig. 18.

During the calibration procedure described in the next section each jet will then be corrected by this factor F_{ptbal}. But, as these coefficients are determined using an high P_T (greater than 10 GeV) selection, the extrapolation of F_{ptbal} to low P_T jets cannot be reliably trusted. Therefore, only jets with $P_T^{jet} > 4$ GeV will be calibrated with this method. In very forward region, $\theta_{jet} < 7°$, affected by leakage in the beam-pipe no absolute calibration can be reasonably applied too. Jets reconstructed in the SpaCal calorimeter ($\theta_{jet} > 155°$) are also not calibrated.

In order to also calibrate remaining hadrons which are not part of a jet, or in jets not calibrated using F_{ptbal}, the dependence of the mis-calibration is also determined as a function of θ_h^e only, as presented in Fig. 19. This will be use to determine calibration coefficients D_θ for each θ bin as defined in Fig. 12. The coefficients D_θ will be applied to all remaining hadrons, separately for data and MC to perform an absolute calibration, except in the region $\theta_{hadron} < 7°$ where a relative calibration is applied. Here only data events will be calibrated, to bring the response of the LAr calorimeter to the one simulated in the MC.

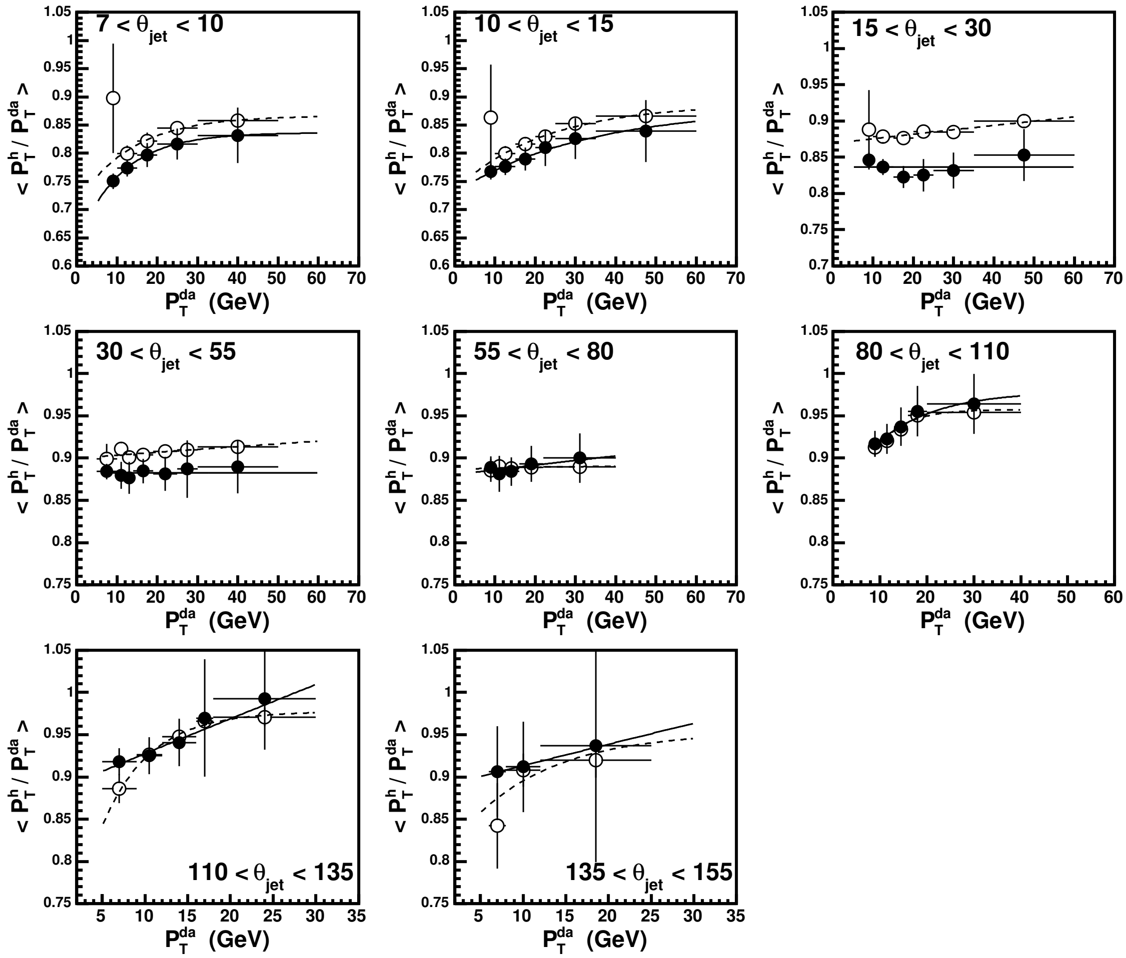


Figure 18: Evolution of the mean values of the P_T^{bal} distributions with P_T^{da} for the calibration sample in the different θ_{jet} regions. Solid and open circles stand for the data and MC, respectively. The plain and dashed lines are fits of the functional form of Eq. (17) to data and MC points, respectively. This example is for 1999p-2000 data and RAPGAP Monte Carlo.

5.3 Application of the calibration

In a first step all hadrons in jets will be calibrated, jet by jet. As the calibration should be applied only to clusters, we have to disentangle for each jet hadrons reconstructed from tracks and from clusters. For each jet we can define the fraction of P_T^{jet} carried by clusters before calibration C_{cls} as

$$C_{cls} = \frac{P_T^{uncalibrated\,clusters}}{P_T^{tracks} + P_T^{uncalibrated\,clusters}}. \tag{18}$$

The fraction of P_T^{jet} carried by tracks is the complement $(1 - C_{cls})$. Note that here the fraction C_{cls} is an approximation because it is determined before any calibration of the energy of clusters. If F_{ptbal} is the absolute correction defined in Sec. 5.2 it is easy to see that the correction

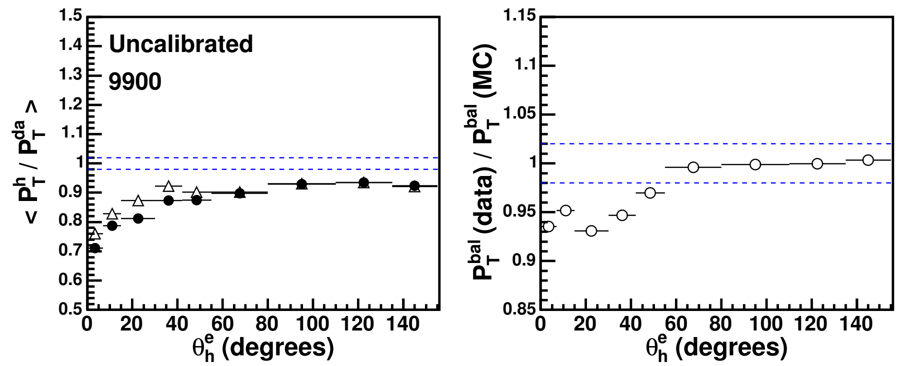


Figure 19: Mean values of P_T^{bal} distributions for each LAr wheel (left figure), for data (solid circles) and MC (open circles). The ratio between data and MC points is presented in the figure on the right. This example is for 1999p-2000 data and RAPGAP Monte Carlo.

factor f we need to apply only to all clusters in the jet is given by

$$f = \frac{1 - F_{ptbal} \times (1 - C_{cls})}{F_{ptbal} \times C_{cls}} \tag{19}$$

For each jet F_{ptbal} which was determined as a function of θ and P_T^{da} (Eq. (17)) will be calculated using instead the mean polar angle of the jet θ_{jet} and its transverse momentum P_T^{jet}. Indeed P_T^{da} can not be used now as, for a general selection, the double angle measurement may not be reliable and the total transverse momentum can be also shared between different jets. In order to have an approximation of the "true" transverse momentum an iterative procedure is used. The uncalibrated P_T^{jet} is used as the argument in Eq. (17) for a first approximation of f. The resulting $P_T^{jet'}$ is then used to compute the final f used to calibrate. For each jet the calibration is performed by multiplying the cluster energy by the f factor. Then in order to be consistent with the kt jet kinematics $\varphi_{jet}, \eta_{jet}, P_T^{jet}$ are properly recomputed and the final jet is massless.

The jets are not calibrated if $P_T^{jet} < 4$ GeV or $\theta^{jet} < 7°$ or $\theta^{jet} > 155°$. The total hadronic final state can be decomposed in hadrons belongings to calibrated jets and remaining hadrons:

$$P_{HFS,Uncalibrated} = \sum_i P_{jet\,i,Uncalibrated} + P_{HFS,not\,in\,jet} \tag{20}$$

where $P_{HFS,not\,in\,jet}$ the part of the HFS not in jets brings a negligible P_T contribution in high Q^2 events[9]. In a second step $P_{HFS,not\,in\,jet}$ will be calibrated using D_θ coefficients depending on the polar angle of each hadron θ_{hadron}. All cluster hadrons will be absolutely calibrated, except for $\theta_{hadron} < 7°$ where the calibration is applied only to data events using $D_{\theta<7°} = D^{data}_{\theta<7°} \,/\, D^{MC}_{\theta<7°}$. The total calibrated hadronic system is then obtained with:

$$P_{HFS,Calibrated} = \sum_i P_{jet\,i,Calibrated} + P_{HFS,not\,in\,jet,calibrated} \tag{21}$$

[9]In the theoretical prescription implemented in the jet finder [13] one end with nothing but only jets in the HFS. However as it is not really reliable to go down to arbitrary low P_T a cut of $P_T^{jet} \geq 2.5$ GeV is introduced for the writing of jets on μODS.

Note that all the calibration coefficients determined here are specific both to Hadroo2 and to the kt jet algorithm.

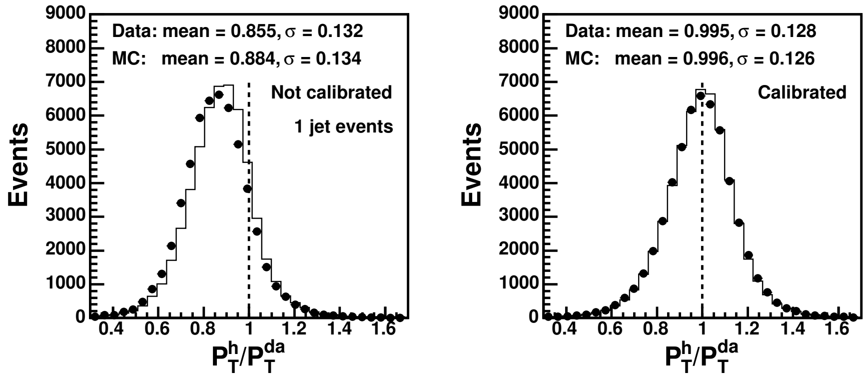


Figure 20: P_T^h/P_T^{da} distribution for one jet events, before (left) and after (right) the application of the jet calibration. The mean and σ values are obtained using a Gaussian fit to the central part of the distributions.

5.4 Tests of the Calibration

The tests are performed with a much larger event sample, requiring the following set of cuts:

- Good quality selection (High Voltage, Vertex, background finders, *etc*),
- $Q^2 \geq 100$ GeV^2,
- 1 electron with $P_T^e \geq 10$ GeV,
- $P_T^h/P_T^e > 0.35$,
- Anti ISR cut $\sum_{h,e}(E - Pz) > 42$ GeV,
- $\theta_{jet} > 7°$, this ensures that the jets are well contained in the calorimeter acceptance.

Now a total $E - Pz$ cut is allowed to reduce the effect of ISR. This different set of cuts will allow to check that the method does not depend on the selection used for the determination of the coefficients. Different NC DIS event sub-samples containing only one, two and three jets will be used for the tests. The two and three jets event samples are independent from the events used for the calibration and therefore provide good tests.

5.4.1 One jet check sample

First the tests with a one jet check sample are performed on 1999p-2000 data and an NC DIS Monte Carlo events generated using RAPGAP. The calibrated and uncalibrated P_T^{bal} distributions are presented in Fig. 20. The distribution is now centered at one and the width is reduced.

The evolution of the mean values of P_T^{bal} distributions as a function of P_T^{da} and θ_h^e is presented in Fig. 21. The ratio $P_T^{bal}(data)/P_T^{bal}(MC)$ is also shown. It is well described within 2 %.

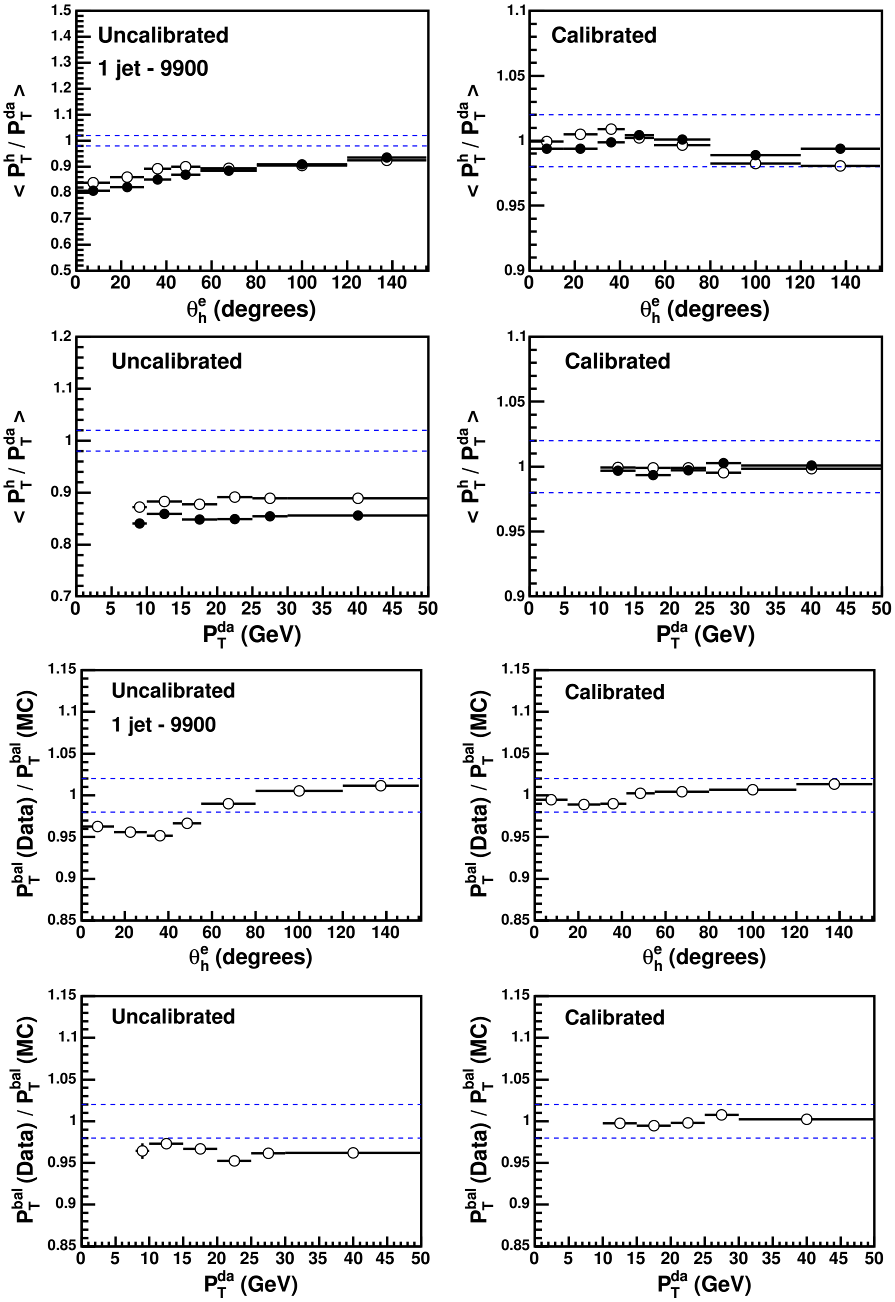


Figure 21: Mean of the P_T^h/P_T^{da} distributions (upper figures) and the ratio $P_T^{bal}(data)/P_T^{bal}(MC)$ (bottom figures) as a function of P_T^{da} and θ_h^e for one jet events, before and after calibrations.

5.4.2 Two and three jets check sample

Especially for cross section measurements of high P_T jets, the minimisation of the hadronic scale uncertainty, and therefore the optimisation of the jet calibration, is necessary. The uncorrected and corrected P_T^{bal} distributions are presented in Fig. 22. Again, after corrections the data and MC agreement is improved and the absolute momentum balances are centered around

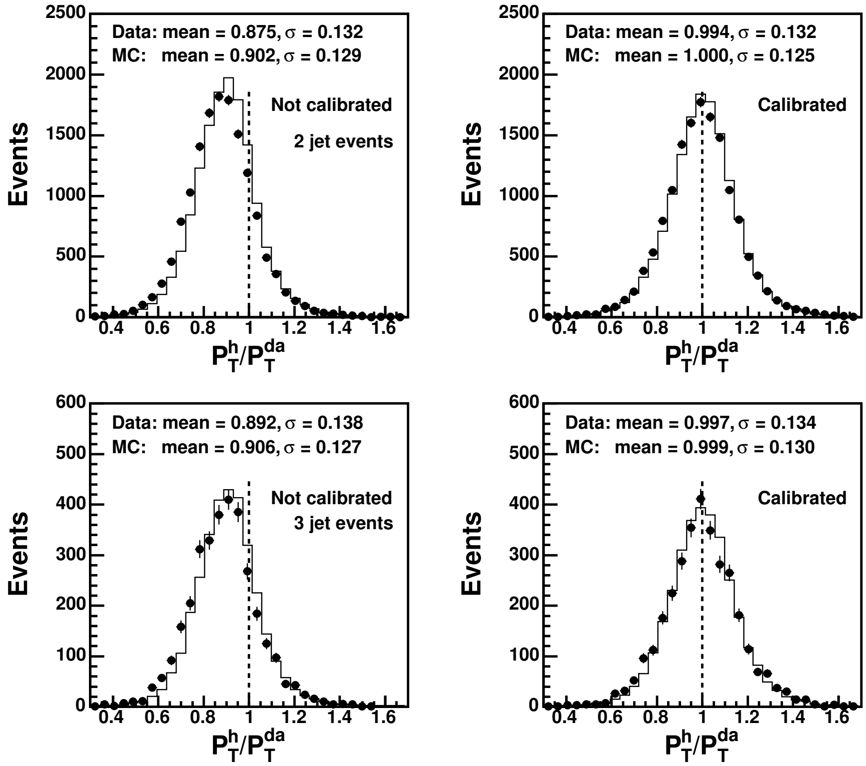


Figure 22: P_T^h/P_T^{da} distributions for two (upper figures) and three (bottom figures) jets events, before (left) and after (right) the application of the jet calibration. The mean and σ values are obtained using a Gaussian fit to the central part of the distributions.

one. The mean values of P_T^{bal} are displayed on Figs. 23 and 24 as a function of P_T^{da} and θ_h^e for two and three jet events.

The results obtained with a two jet sample show that the overall P_T balance is centered around 1.0 and that the systematic shift does not exceed 2 %. We can observe that the absolute hadronic scale is obtained within 2 % after coorection, for the data and the MC. The systematic uncertainties are also of the order of 2 % in all the P_T^{da} and θ_h^e ranges.

5.4.3 Inclusive check sample

With the inclusive check samples which has a larger statistics, the calibration can be tested as a function of P_T^{da} and θ at the same time. The mean values of corrected and uncorrected P_T^{bal} distributions as a function of P_T^{da} and θ_h^e are presented in Fig. 25. This example shows, on data event only, the effect of the jet calibrations. We can observe that after correction the absolute hadronic energy scale is well obtained within 2 %. The effect of the jet calibrations on the agreement between data and MC P_T^{bal} distributions is displayed in Fig. 26. One can see that again after having applied the calibration, the systematic error is well contained within 2 % in all P_T^{da} bins of each θ region.

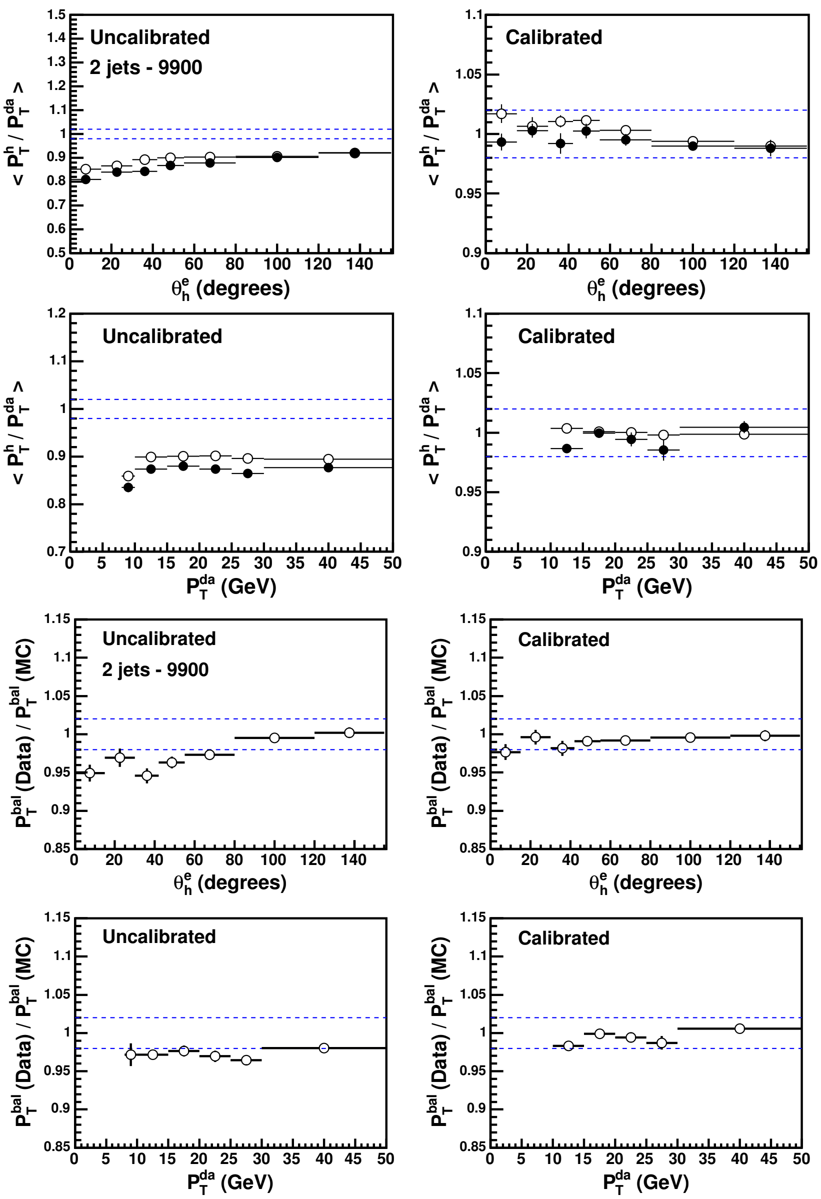


Figure 23: Mean of the P_T^h/P_T^{da} distributions (upper figures) and the ratio $P_T^{bal}(data)/P_T^{bal}(MC)$ (bottom figures) as a function of P_T^{da} and θ_h^e for two jets events.

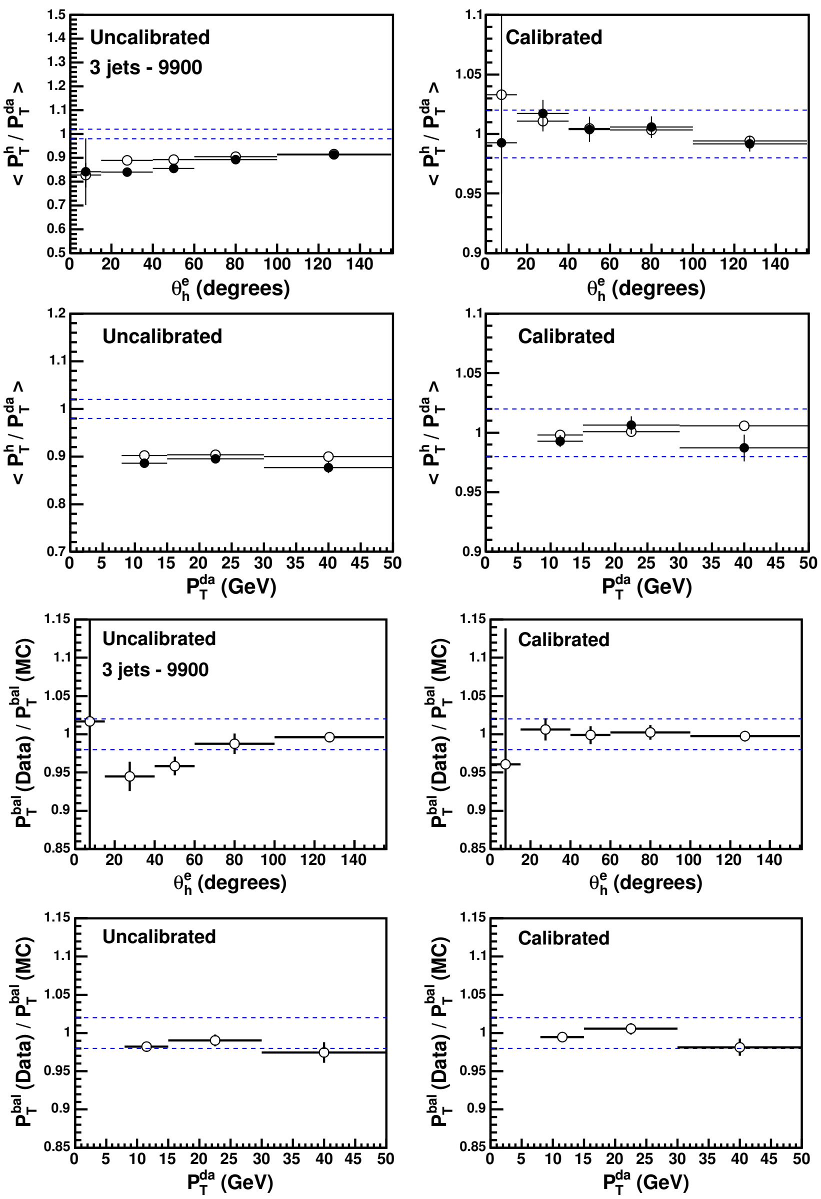


Figure 24: Mean of the P_T^h/P_T^{da} distributions (upper figures) and the ratio $P_T^{bal}(data)/P_T^{bal}(MC)$ (bottom figures) as a function of P_T^{da} and θ_h^e for three jets events.

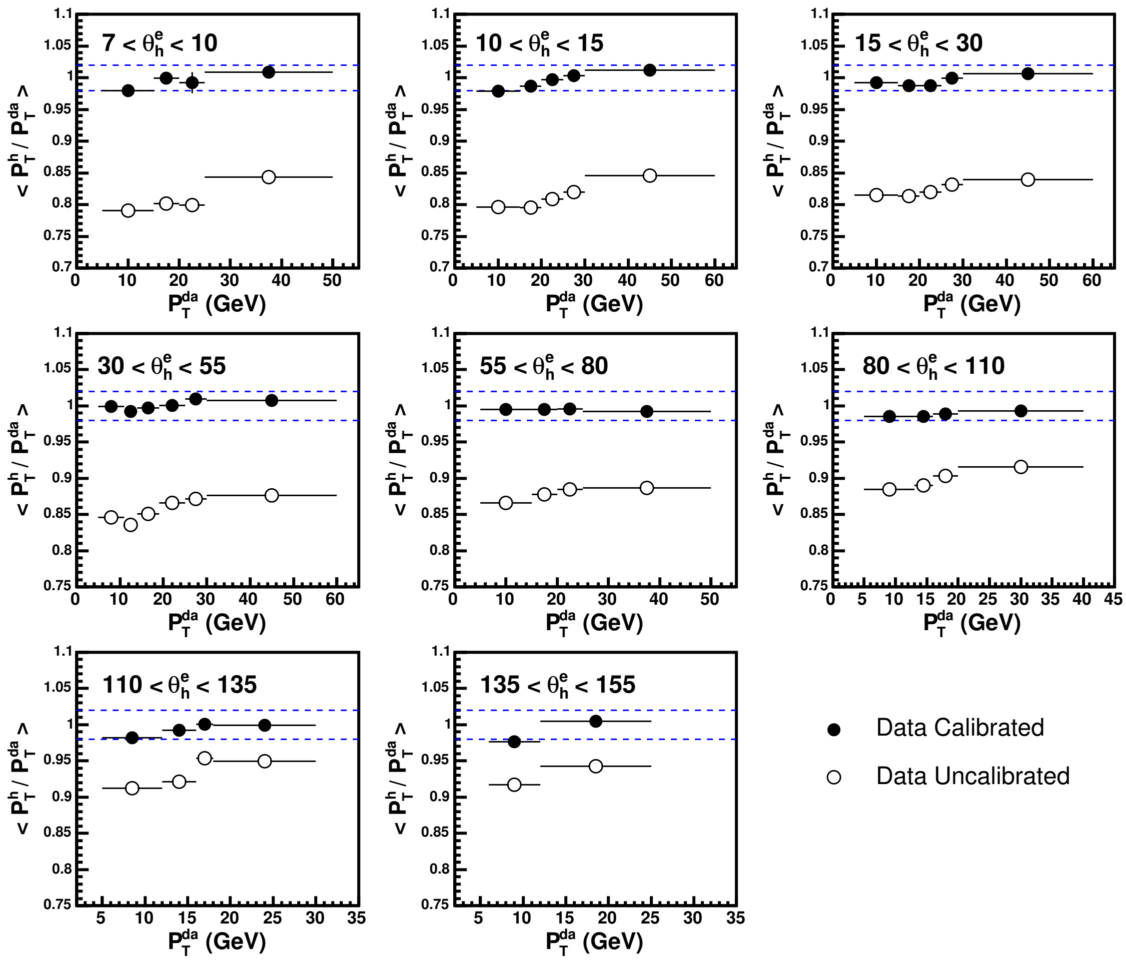


Figure 25: Evolution of the mean values of the P_T^{bal} distributions for data with P_T^{da} for the inclusive check sample in the different θ_h^e regions, before (open circles) and after (solid circles) application of the calibration. This example is for 1999p-2000 data.

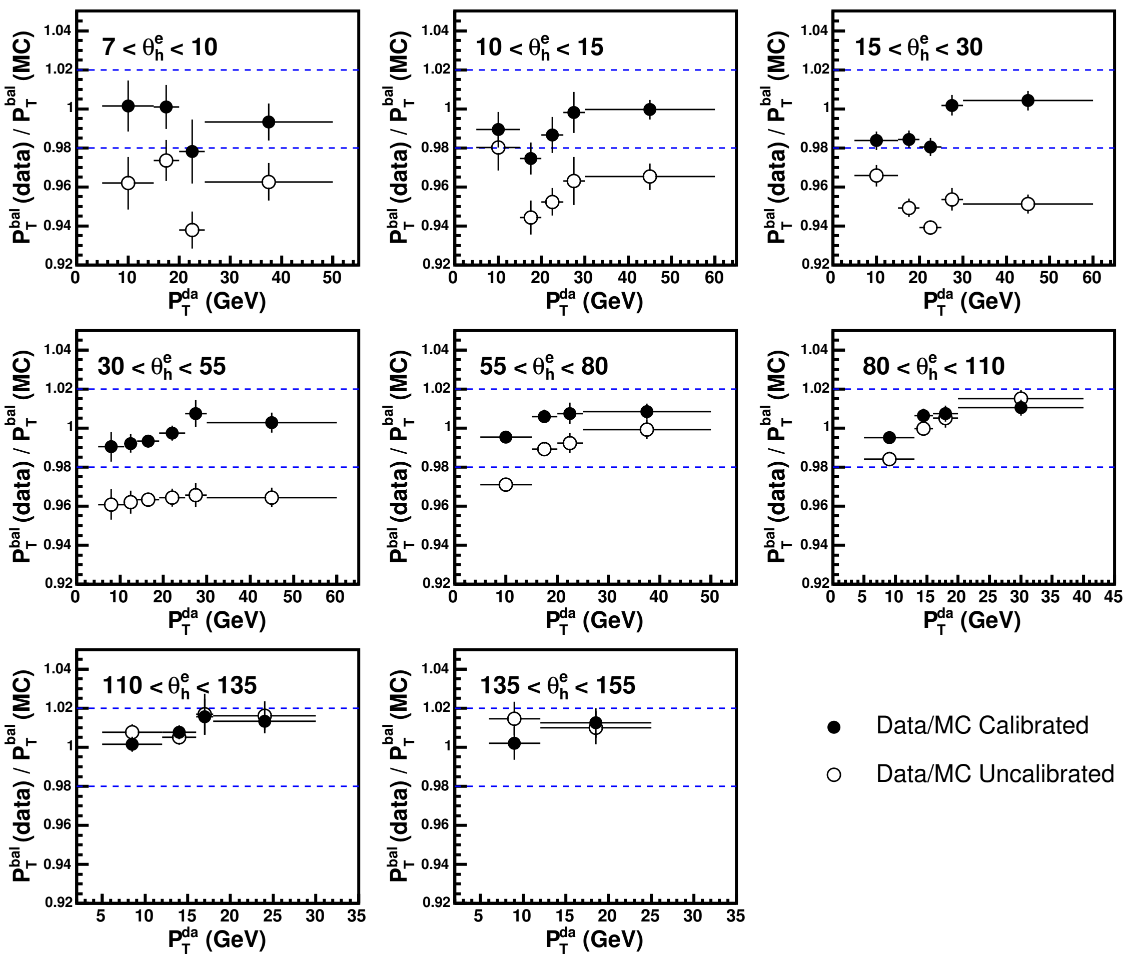


Figure 26: Evolution of the agreement between data and MC P_T^{bal} distributions with P_T^{da} for the inclusive check sample in the different θ_h^e regions. This example is for 1999p-2000 data and RAPGAP Monte Carlo.

5.5 Resolution improvements

Using the inclusive check sample, we can verify the effect of the hadronic calibrations on the resolution of the P_T^h measurement. The evolution of the relative resolutions $\sigma(P_T^{bal})/P_T^{bal}$ as a function of θ_h^e and P_T^{da} is presented in Fig. 27. The resolutions are calculated before and after the application of the jet calibration. In general, the relative resolutions are improved by the calibrations.

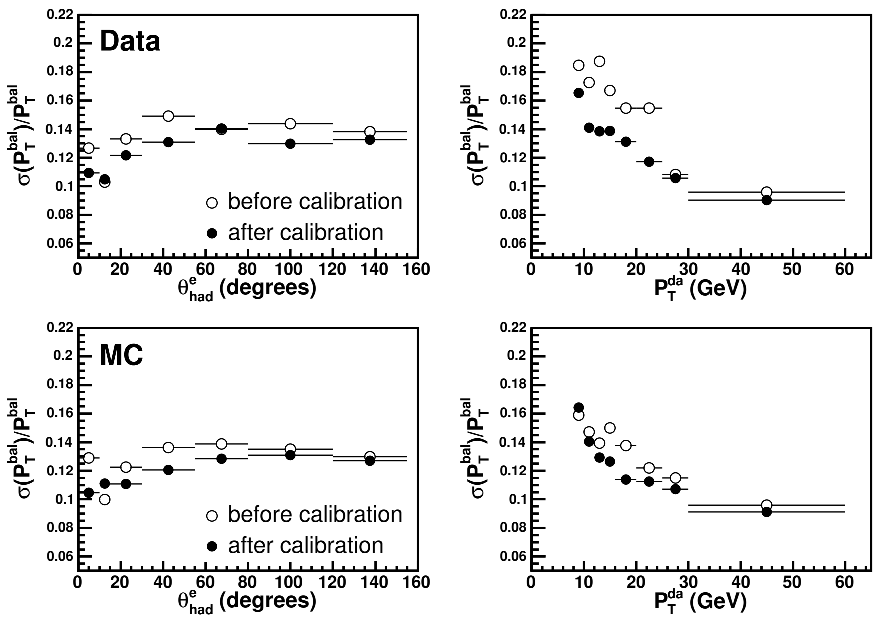


Figure 27: Evolution of the relative resolutions, calculated before and after applying the jet calibrations, as a function of θ_h^e (left) and P_T^{da} (right). The effect on data and MC is presented in upper and bottom figures, respectively.

5.6 Practical implementation in analysis: how to use it ?

This section will be devoted to give an example of the use of the actual jet calibration available in the *H1JetCalibration* package into an H1OO analysis. Indeed the reconstructed hadrons and jets available on μODS are not presently calibrated and each user has to apply by itself the jet calibration in his own analysis. We should here stress that this calibration was only developed for high P_T jets (greater than 10 GeV) and high Q^2 inclusive analyses ($Q^2 > 100$ GeV^2) and it can not be guaranteed that this calibration is working also on low Q^2 events.

The jet calibrations should first be initialised at the beginning of your job (before entering in the event loop) using:

```
H1JetCalibration* JetCalib = new H1JetCalibration;
JetCalib->InitHadroo2KtJetCalibration((int)RunYear,
                                      (int)RunType);
```

The RunType has to be 0 for data and 1 for MC. The available run periods for RunYear are defined as 96-97 = 3, 98-99 = 4 , 99-00 = 5 and 03-04 = 7.

Then inside the event loop, each event has to be calibrated by calling:

```
Int_t nJets=ModsPartJet.GetEntries();
Float_t fPthCalib=0;
Float_t fGammahCalib=0;
Float_t fEmpzhCalib=0;
TLorentzVector* jetscal= new TLorentzVector[nJets];
TLorentzVector HadCalTotVec=
               JetCalib->GetHadroo2KtJetCalibration(fPthCalib,
               fGammahCalib,fEmpzhCalib,jetscal);
delete [] jetscal; //--- after having used it
```

where HadCalTotVec is the four-vector of the calibrated hadronic final state, fPthCalib the calibrated P_T^h, fGammahCalib is θ_h after calibration and fEmpzhCalib the $E - Pz$ of the hadronic system. The calibrated four-vectors of all jets in the events are provided in the array jetscal (note that the user has to take care of deleting properly this array).

The total $E-Pz$ and missing transverse momentum P_T^{miss} of the event (Epz and Ptmiss, respectively) can be re-calculated by adding all isolated identified electrons and muons to the total HFS four-vector with:

```
TLorentzVector TotalVec=HadCalTotVec;
H1PartEmArrayPtr   ModsPartEm;
H1PartMuonArrayPtr ModsPartMuon;
for(Int_t i=0;i<ModsPartEm->GetEntries();i++){
 if(ModsPartEm[i]->IsIsolatedLepton())
     TotalVec+=ModsPartEm[i]->GetFourVector();
}
for(Int_t i=0;i<ModsPartMuon->GetEntries();i++){
 if(ModsPartMuon[i]->IsIsolatedLepton())
     TotalVec+=ModsPartMuon[i]->GetFourVector();
}
Float_t Epz=TotalVec.E()-TotalVec.Pz();
Float_t Ptmiss=TotalVec.Pt();
```

The hadronic kinematic variables can be also re-calculated:

```
Eh =  HadTotVec.E();
Pzh = HadTotVec.Pz();
Pth = HadTotVec.Pt();
Phh = HadTotVec.Phi();
if (Pth>0) Thh = 2*TMath::ATan((Eh-Pzh)/Pth);
Gammah = Thh;
yh  = (Eh-Pzh)/(2*GenPl);
Q2h = (Pth*Pth)/(1-yh);
if(yh!=0) xh = Q2h/(yh*GenS);
```

as well as the kinematic variables from the double angle method:

```
Q2da = 4*pow(GenPl,2)*sin(Gammah)*(1+TMath::Cos(The))
     /(TMath::Sin(Gammah)+TMath::Sin(The)-TMath::Sin(Gammah+The));
xda = (GenPl/GenPp)*(TMath::Sin(Gammah)+TMath::Sin(The)
    +TMath::Sin(Gammah+The))/(TMath::Sin(Gammah)+TMath::Sin(The)
    -TMath::Sin(Gammah+The));
yda = Q2da / (xda*GenS);
Ptda = 2*GenPl/(TMath::Tan(The/2)+((Eh-Pzh)/Pth));
Eda = 2*GenPl*sin(Gammah)/(sin(Gammah)+sin(The)-sin(Gammah+The));
```

where `The` is the polar angle of the scattered electron and `GenPl` = 27.598 and `GenPp` = 919.971 the energies of the incoming lepton and proton beams, respectively.

6 Conclusion

Along this note the motivations and details implemented in the actual Hadroo2 HFS finder have been described. Firstly the distortion in the measurement of the kinematic variables in the low y region was investigated. Inspired by previous FORTRAN implementations, dedicated noise suppression algorithm have been implemented into Hadroo2 to correct this distortion. Detailed checks have shown that these algorithms are working properly and that no signal important to exclusive analyses is suppressed. In a second part, details of the Hadroo2 algorithm and the way track and clusters measurements are chosen and combined were explained. Results of comparisons of Hadroo2 with other HFS reconstruction algorithms were presented. They show that Hadroo2 improves effectively the HFS reconstruction and the resolution, especially in the high P_T domain. Finally, the method of jet calibration available in *H1JetCalibration* was presented. This calibration is applicable for any event samples provided the transverse momentum of either the scattered electron or the hadronic system is larger than 10 GeV. It is available for all running periods, from 1994 to 2004 and checks have shown that the absolute hadronic scale is reached within 2 % and that the systematic errors are of the order of 2 %. Concerning future developments of the HFS reconstruction, it was shown that the application of the new weighting scheme presented in [7, 8] can help for additional improvement of the resolution. Work is also ongoing to extend the actual hadronic calibrations to all events, including low Q^2 ones and to apply it directly to each μODS hadron.

References

[1] M. Peez, “Recherche de déviations au Model Standard dans les processus de grande énergie transverse sur le collisionneur électron - proton HERA”, Ph.D. thesis, Université de Lyon (2003), DESY-THESIS-2003-023. (available at http://www-h1.desy.de/psfiles/theses/).

[2] M. Ridel, “Reconstruction of the energy flow and search for squarks and gluinos in the D0 experiment. (In French),” FERMILAB-THESIS-2002-08 (available at http://www.slac.stanford.edu/spires/find/hep/www?r=fermilab-thesis-2002-08).

[3] G.M. Briskin, “Diffractive Dissociation in ep Deep Inelastic Scattering”, DESY-THESIS-1998-036.

[4] QHQTRK manual, (Heavy Flavor Working Group track selection code by Lee West).

[5] I. Abt *et al.* [H1 Collaboration], “The H1 detector at HERA,” Nucl. Instrum. Meth. A **386** (1997) 310.

[6] S. Hellwig, “Untersuchung der $D^{\star} - \pi_{slow}$ Double Tagging Methode in Chamanalysen”, Diplomarbeit, Universität Hamburg (2004).
(available at http://www-h1.desy.de/psfiles/theses/).

[7] C. Issever, K. Borras, D. Wegener, “An Improved Weighting Algorithm to Achieve Software Compensation in a Fine Grained LAr Calorimeter”, H1-03/03-608.

[8] M. Ellerbrock, “An Improved Weighting Algorithm to Achieve Software Compensation in a Fine Grained LAr Calorimeter”, Ph.D. thesis, Phys. Inst. der Univ. Heidelberg, DESY-THESIS-2004-015. (available at http://www-h1.desy.de/psfiles/theses/).

[9] E. Chabert *et al.*, “QBGFMAR, an updated PHAN package for cosmic and halo muon topological rejection in high P_T physics analysis”, H1-11/98-556.

[10] C. Veelken, “H1NonepBgFinder - Rejection of cosmic muon and beam-halo events in the H1OO framework”, H1-09/02-603.

[11] J. Cao and Z. Zhang, “Towards an Unbiased Measurement of Kinematic Variables at Low y Region”, H1-12/99-580.

[12] M. Jacquet, Z. Zhang *et al.*, “Absolute hadronic jet calibration of the H1 Liquid Argon calorimeter”, H1-04/99-571.

[13] S. D. Ellis and D. E. Soper, “Successive Combination Jet Algorithm For Hadron Collisions,” Phys. Rev. D **48** (1993) 3160, [arXiv:hep-ph/9305266].