Suggestions for a General Energy Reconstruction Scheme for the ATLAS Calorimeters
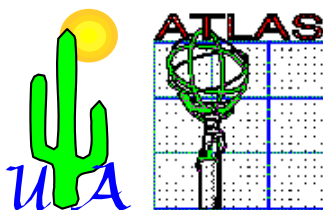
Peter Loch

Dept. of Physics, University of Arizona

Tucson, Arizona 85721

USA

March 4, 1997

Abstract

A possible general scheme for energy reconstruction for the ATLAS calorimeters is described. A major ingredient is the concept of an intrinsic electromagnetic energy scale, as defined by the true or ideal electron calibration constant, on which experimental and simulated signals can be compared at every step of the reconstruction chain. The other important part of the suggested reconstruction model is local calibration, a strategy based on the idea to calibrate clusters of calorimeter cells with correlated signals in such a way that the final energy reconstructed from these objects corresponds to the true energy deposit at the cluster location. The nature of the energy deposit, which within some restrictions can be measured from the spatial signal distribution in a cluster, can be included in the calibration. This may allow to improve energy resolution in the non-compensating ATLAS calorimeters and, more generally, is a way to apply calibration functions determined with single particle response in the more complex signal environment of particle jets.

Suggestions for a General Energy Reconstruction Scheme for the ATLAS Calorimeters

Peter Loch

Dept. of Physics, University of Arizona

Tucson, Arizona 85721

USA

March 4, 1997

Abstract

A possible general scheme for energy reconstruction for the ATLAS calorimeters is described. A major ingredient is the concept of an intrinsic electromagnetic energy scale, as defined by the true or ideal electron calibration constant, on which experimental and simulated signals can be compared at every step of the reconstruction chain. The other important part of the suggested reconstruction model is local calibration, a strategy based on the idea to calibrate clusters of calorimeter cells with correlated signals in such a way that the final energy reconstructed from these objects corresponds to the true energy deposit at the cluster location. The nature of the energy deposit, which within some restrictions can be measured from the spatial signal distribution in a cluster, can be included in the calibration. This may allow to improve energy resolution in the non-compensating ATLAS calorimeters and, more generally, is a way to apply calibration functions determined with single particle response in the more complex signal environment of particle jets.

1 Introduction

Fundamental calibration biases and problems

The quality of energy reconstruction in any high energy collider experiment is naturally affected by intrinsic calorimeter performance (linearity, resolution etc.) and acceptance (hermetic coverage, including inactive material, energy leakage in cracks or thinner regions etc.), both of which can be regarded as hardware acceptance. Signal processing in the reconstruction software can of course correct some of the hardware deficiencies, but can also introduce additional changes in the detector acceptance: signal and physics object[1] definition require selections which may lead to signal losses or gains by applying more or less efficient methods within a limited detector granularity. These effects have to be understood from testbeam data and simulations, and full physics simulations. They are not always fully correctable on event by event level, and may therefore add to detector resolution for a specific physics variable.

The calibration procedure for a calorimeter must then take into account the relation between energy reconstruction and full detector acceptance, including whatever signal definition

---

[1]identified single particles (mostly e, μ, γ), jets, or missing transverse energy carried by neutrinos.

is applied to the raw data. At least for this reason calibration is not easily obtainable from testbeams with typically different hardware environments and possibly different reconstruction and analysis cuts – after all, the best signal definition may only be known once the physics experiment is running. It is also especially problematic in non-compensating calorimeters like the ones in ATLAS, where even the intrinsic response depends on the incoming particle type, and jet calibration can only be extrapolated from the single particle testbeams, in addition to *in situ* calibration using transverse momentum balance. It is obvious, though, that the later approach fully integrates hardware (dead materials, dead channels etc.) and software (jet finding algorithms etc.) acceptance into the calibration. This makes in situ calibration only really acceptable if at least the hardware acceptance is stable in time. Also, some detector regions may hardly be accessible, like the forward region, where the total average transverse energy is typically too small compared to the detector resolution to allow for a determination of calibration parameters within reasonable errors.

An approach to calibration and energy reconstruction: intrinsic energy scales and local calibration

One concept to overcome some of these problems includes an energy reconstruction scheme based on a *local calibration ansatz* with well-defined energy scales. The most important one is the *intrinsic electromagnetic energy scale*, introduced at the lowest level in the reconstruction chain, and therefore common basis for all further signal treatment of testbeam and physics data and simulations. It is characterized by a *true* or *ideal electron calibration constant*, which has to be determined independent of any signal definition – a necessary condition to transfer any (!) calibration from testbeam to final experiment, and one of the major features in this reconstruction model.

The response on what we call the *final local energy scale* always depends on a signal definition, which usually provides input to calibration parameters. Among these inputs can be estimators of *signal nature* of response objects, i.e. the electromagnetic or hadronic origin of (typically) clusters of calorimeter cells with correlated signals. Providing the calorimeter has a fine enough read out granularity, meaning there is a certain sensitivity in some measurable variables related to these clusters with respect to the different shower characteristics, compensation can be achieved by applying calibration procedures according to the cluster nature. In this picture clusters of hadronic nature would receive another kind of calibration, like signal weighting, than electromagnetic clusters.

This idea, together with a signal definition based on cell clusters and including the already mentioned cluster classification prior to energy reconstruction, are the other major features of this model. It is referred to as local calibration, as it requires calibration procedures to be normalized to energy deposit at the cluster location. Also, calibration of clusters instead of full particle calibration can be a concept to extract corresponding parameters and functions from single particle testbeams and apply them to physics signals, under the assumption that an identical "look" of these objects, measured in a given space of variables like signal density and dispersion or compactness, means the same relation between signal and deposited

energies in testbeam and physics experiment. This is especially interesting for the hadronic calibration.

The ultimate goal is of course full energy reconstruction, providing the the best estimate of the true energy of physics objects like isolated particles and jets. It can be based on the reconstructed clusters and their energies on the final local scale, where each cluster can be seen as a (possibly massive) *pseudoparticle* with a 4-vector assigned to it. This step depends even more on selection criteria applied to identify particles and jets, and is a good part of physics analysis. A reconstruction program typically includes only one given default algorithm for each physics object. All data needed to re-define physics objects should be provided in addition.

Optimal energy reconstruction of physics objects may also require to include tracks, for example for photon/electron separation. All other steps preceding this very last one should be performed using calorimeter signals only, though, because reliable energy measurement in a given calorimeter must not be influenced by tracker performance, or even by neighbouring calorimeters. Improvement of the energy measurement by combining signals (generalizing cell clustering) from adjacent calorimeters is of course foreseen, but local calibration as a concept also allows to estimate true energy deposits in individual calorimeters.

The most obvious conceptional problem with this model is that it relies heavily on response simulations, mainly for testbeams. This is less critical for the determination of the electromagnetic energy scale, as electron response can typically be simulated at a 1% level with respect to experimental data [1]. More problematic is the hadron response simulation needed for calibration normalization at cluster level, especially because this model is somewhat sensitive to details of shower development. Better than 5% normalization error has been achieved for the H1 parallel plate liquid argon calorimeter with lead and iron absorber [2], with present day shower models.

Basics of this energy reconstruction model have been explicitly introduced by the H1 collaboration [1, 2, 3], but many other high energy physics experiments are implicitly using the same approaches. The model discussed in this note is also extended compared to the original H1 suggestions cited above, where the local calibration and reconstruction scheme has not been completely implemented for jets.

Layout of this note

In the following section 2 we describe the overall design of the suggested energy reconstruction scheme and introduce its basic ingredients. We do not have any (quantitative) expectations for physics performance for a given calorimeter or region in ATLAS at this time – this is a much larger project which has to be done to decide on a strategy and justify *any* chosen energy reconstruction approach. Naturally though, the model discussed here is designed to include all features that allow to monitor biases introduced at every given step in the reconstruction chain, and maybe even keep these biases small.

Section 3 focuses on the discussion of the first step in the reconstruction chain, the determination and application of the ideal electromagnetic energy scale. In the following section

4 we introduce a general signal definition based on cell clusters, while the final energy reconstruction based on these objects is discussed in section 5. This includes a brief discussion of particle identification, jet finding and missing transverse energy calculation, and the (software) interface between reconstruction and analysis.

The scope of this note, and the intention behind it, is certainly limited to the introduction of a *complete* concept for a *general purpose*, least physics biased energy reconstruction based on some features observed in the response of operating non-compensating calorimeters. No effort to actually adopt any of the described schemes for actual ATLAS energy reconstruction needs has been made so far, not even basic feasibility checks within the designed calorimeter read out granularity[2]. It may well be that physics performance requirements can be met with somewhat simpler approaches, especially in certain calorimeter regions. It has been shown, though, that the concept of an ideal energy scale together with a cluster based signal definition, including electromagnetic cluster filtering, can bring hadronic energy scale errors down to better than 5% [2, 4], at least in a low noise environment.

Local calibration has not yet been fully exploited in a physics environment, at least not to our knowledge. Specifically the effect of the large pile up fluctuation at LHC on hadronic energy scales are certainly not yet understood and have to be studied for this model as well as for any other solution.

2   General energy reconstruction scheme

In this section we discuss the overall energy reconstruction scheme together with its basic ingredients. Before we can do this we actually have to define some of the expressions often used in the course of this discussion – not all are commonly understood in the same way (section 2.1).

Most of the motivations and requirements to energy reconstruction have been presented in the introduction of this paper. They are collected in more systematic way in section 2.2, just before we finally introduce the overall design of the model in section 2.3.

2.1   Some definitions

The first major concept implemented into the suggested model is the intrinsic electromagnetic energy scale EMS. It is basically defined by the true or ideal experimental electron calibration constant c_e.

In this context we need to introduce our somewhat idealistic understanding of *intrinsic response* in general. We define this response as the "true" signal corresponding a given *deposited energy* – not the incoming energy ! – in a calorimeter. This signal is purely defined by the sampling fraction and only affected by sampling and intrinsic shower fluctuations. It is ideal with the respect that it is not influenced by charge or light collection efficiencies, specific

---

[2]a first round on the electron calibration following the concepts in section 3 has been done for the forward calorimeter [5]

behaviour of electronic chains, or any specific acceptance limitations due to dead material in front, the limited size of a given calorimeter module, or analysis cuts. The intrinsic response is obviously not easily accessible, even in a testbeam setup.

Based on this understanding, c_e is defined as the average ratio of the intrinsic response of a given calorimeter to the energy E_{dep} deposited by electrons at a certain beam energy E_{beam}. Naturally we assume intrinsic linearity of the electron signal, so c_e is supposed to be independent of the deposited (!) energy. The relation between E_{beam} and E_{dep} is determined by the calorimeter acceptance[3], and may well depend on E_{beam}, for example.

The EMS signal is then defined by the product of a "real" experimental response Q, measured in the same units as the ideal signal, and the ideal calibration constant c_e. There is no specific restriction on the analysis methods applied to determine Q, which in return means that the EMS signal is depending on these analysis cuts, and other acceptance limitations due to dead channels, for example. It also means that the reconstructed energy on this scale is generally not equal to E_{dep} or, even less likely, E_{beam}!

The analysis cuts mentioned above are part of what we call a *signal definition*. This step is typically motivated by signal significance enhancement, but also to generate objects for local calibration. Principal contributions to this signal definition are cell level noise cuts, topological noise suppression and cluster building. At the very end of the reconstruction chain, and to a part maybe already at analysis level, particle identification and jet finding are also included.

2.2   Fundamental requirements and motivations

The basic requirements and motivations for the energy reconstruction model suggested in this note are generally driven by ideas of lowest possible biases to all imaginable physics channels in addition to well controlled systematic errors on energy scales, most of which are introduced in the calibration transfer from testbeam to the final calorimeters. Naturally one also requires linearity in the jet and isolated particle response, together with an optimal energy resolution.

These very general requirements can be translated into more specific items and summarized together with some important consequences for the energy reconstruction in the following list:

(1) Minimization of systematics requires that experimental data can be compared to corresponding simulations after every step of the reconstruction chain. It also requires that it must be possible to apply the full reconstruction to testbeam data as well as physics data.

(2) Linear response for hadrons in non-compensating calorimeters can be achieved by signal weighting techniques [2]. The corresponding functions have to be extracted from testbeam data or testbeam and physics simulations. Universal application to experimental

---

[3]a rather high acceptance has to be assumed, still, for the reasons given on page 12 in section 3 $\Rightarrow$ dead material in front should of course be minimized, and calorimeter modules should not be too small with respect to electromagnetic shower sizes!

data and simulations – a consequence of the requirement stated in point (1) of this list – requires a well defined fundamental EMS.

(3) Biases towards certain physics or physics models should be avoided in the general purpose energy reconstruction. This introduces a problem for the hadronic sector, as calibration functions for jets are usually not accessible from single hadron testbeam data. If determined by jet response simulations, biases towards the used jets, specifically the fragmentation model, can be avoided by using more localized signal objects within a signal definition including cluster formation and classification. This strategy also allows to apply the calibration functions to testbeam data by concept, if not even extract the corresponding parameters from this single hadron response.

(4) The EMS signal mentioned in point (2) needs to be defined in a way that its relation to the deposited energy is, at a given signal definition, the same at the testbeam and the final calorimeter for the same kind of particles at identical energies. It is only in this case that the universal application requirement mentioned in the previous points can be met. This again can be realized by the introduction of the ideal electromagnetic calibration constant defined in section 2.1, together with a charge scale for experimental data. Signals on this charge scale need to be fully corrected for any differences in the charge collection efficiency or electronics between testbeam and physics experiment.

(5) More a consequence of the previously stated items than an actual requirement itself, is the "correct" use of testbeam data for calibration purposes. The fundamental EMS must certainly be determined experimentally from electron response. Cluster builder parameters can also be extracted from this source, especially for an "electromagnetic oriented" cluster formation. Hadron calibration functions and parameters determined with single pion testbeams may not be applicable to jet response, even within a cluster based approach. Single pion testbeam data is still important, though, for monitoring hadronic shower package performance and determination of the systematic error of the final local energy scale (FES) by testing calibrations extracted from jet response simulations.

From this list we extract two major tasks for calibration, the first being the determination of the fundamental ideal electron calibration constant, to define the EMS of a given calorimeter or calorimeter region. It is rather obvious that the EMS signal does not provide an estimate for the "true" energy of a physics object (single particle or jet). This can easily be understood from the following example: the reconstructed energy for an electron depends not only on the electron calibration constant, but also on the number of cells included in its calculation. It therefore depends on a signal definition[4].

It is also obvious that the very first step in calibration, and therefore also in the energy reconstruction, must be data specific, as it must take into account the data type (experiment

---

[4]it also depends on corrections beyond the application of a simple calibration constant, as for energy losses in dead materials, for example.

or simulation) and the "local" environment. A part of this step is electronics calibration and charge corrections used to unfold deficiencies in charge or light collection and to minimize differences in electronics or read out between testbeam and final experiment setup. Formally a charge scale for experimental signals is introduced, which even allows to use different electronics and read out systems at testbeams, for example.

The second task is the final energy calibration, which is preceded by the signal definition. Input to this are EMS cell signals, representing one physics event or testbeam particle. Signal losses and artificial gains due to more or less efficient selections in the signal definition process must be corrected here by normalizing the applied calibration functions to the true energy deposit at the signal object location. Other corrections for dead material losses in front or in calorimeter cracks must be determined in addition.

2.3   Energy reconstruction scheme

The suggested energy reconstruction model is discussed step by step in this paragraph, with the global design shown in figure 1. More details on the discussed steps, including some remarks on the actual determination of calibration constants and parameters, and some suggestions on cluster finding and classification, can be found in the following sections of this note.

Step 1: electromagnetic energy scale

The first step in the energy reconstruction for the experiment is to convert all raw data cell by cell to a charge scale signal, with all necessary corrections for cross talk, electronic gain variations etc. done. These are usually included in the electronics calibration procedure. Other corrections on top of this are more global, like for charge collection inefficiencies introduced by liquid argon contamination, and must be applied here, too.

The conversion of the signal from this charge scale to EMS is then rather trivial, one just has to multiply the signal with the ideal electromagnetic calibration constant (see section 3 for further details).

For simulations the procedure is somewhat similar, except that the raw data, usually visible energies, is only multiplied by the inverse electron sampling fraction, and no other corrections are needed. The experimental noise has to be included, of course. This can be done by explicitly adding randomly triggered "empty" experimental events to the Monte Carlo event, on cell by cell bases (see section 3), or simulate the noise in each calorimeter cell by an appropriate model.

As already mentioned in the item (1) in section 2.2, the following reconstruction steps should be identical for testbeam and experiment data and simulations.

Step 2: signal definition

The next step in the energy reconstruction is the fundamental signal definition. In the experiment (testbeam or physics) the input to this step is basically a list of calorimeter cells,
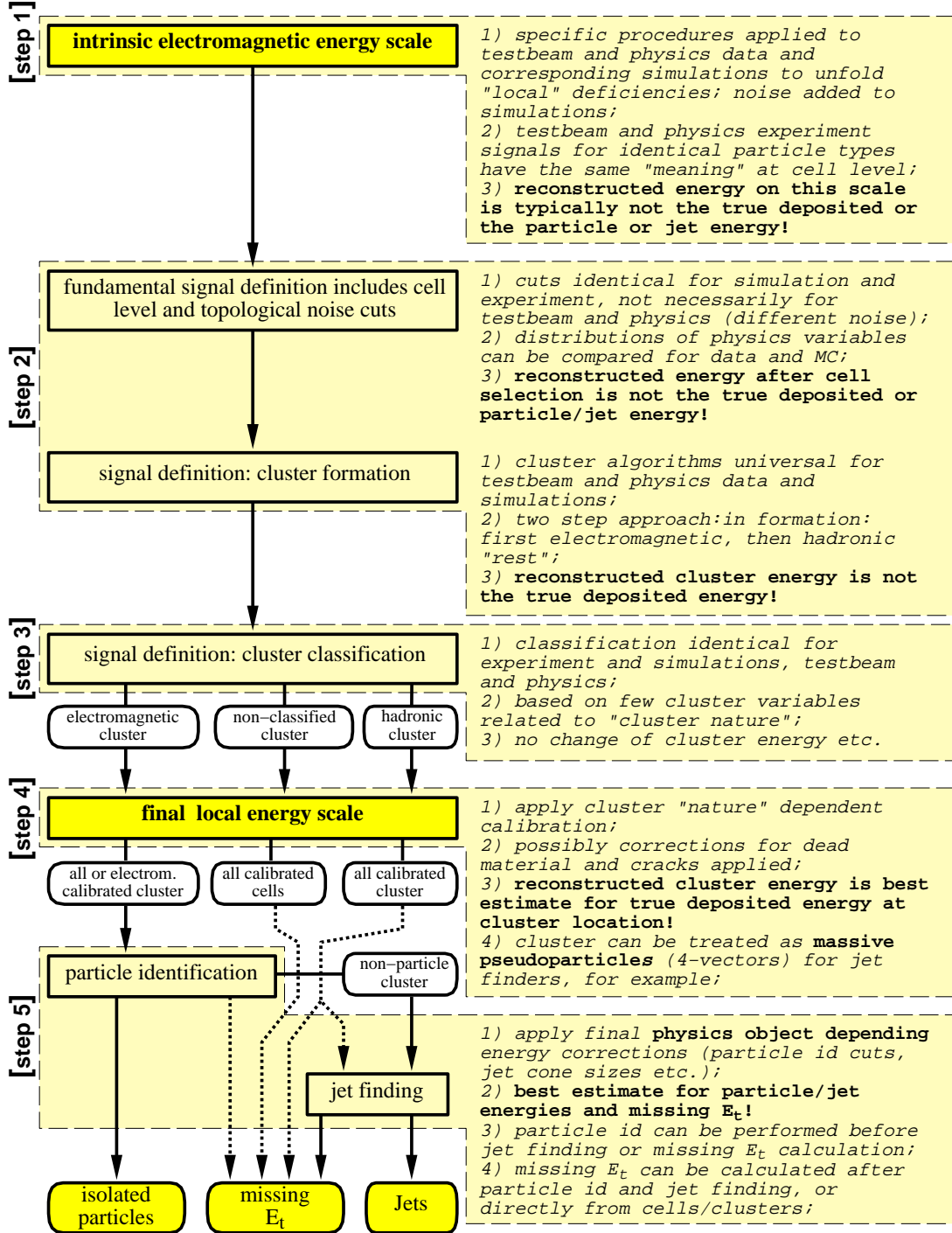
[step 1]

intrinsic electromagnetic energy scale

1) specific procedures applied to testbeam and physics data and corresponding simulations to unfold "local" deficiencies; noise added to simulations;
2) testbeam and physics experiment signals for identical particle types have the same "meaning" at cell level;
3) reconstructed energy on this scale is typically not the true deposited or the particle or jet energy!

[step 2]

fundamental signal definition includes cell level and topological noise cuts

1) cuts identical for simulation and experiment, not necessarily for testbeam and physics (different noise);
2) distributions of physics variables can be compared for data and MC;
3) reconstructed energy after cell selection is not the true deposited or particle/jet energy!

signal definition: cluster formation

1) cluster algorithms universal for testbeam and physics data and simulations;
2) two step approach:in formation: first electromagnetic, then hadronic "rest";
3) reconstructed cluster energy is not the true deposited energy!

[step 3]

signal definition: cluster classification

1) classification identical for experiment and simulations, testbeam and physics;
2) based on few cluster variables related to "cluster nature";
3) no change of cluster energy etc.

electromagnetic cluster

non–classified cluster

hadronic cluster

[step 4]

final local energy scale

1) apply cluster "nature" dependent calibration;
2) possibly corrections for dead material and cracks applied;
3) reconstructed cluster energy is best estimate for true deposited energy at cluster location!
4) cluster can be treated as **massive pseudoparticles** *(4-vectors) for jet finders, for example;*

all or electrom. calibrated cluster

all calibrated cells

all calibrated cluster

[step 5]

particle identification

non–particle cluster

1) apply final **physics object depending** *energy corrections (particle id cuts, jet cone sizes etc.);*
2) best estimate for particle/jet energies and missing E$_t$!
3) particle id can be performed before jet finding or missing E$_t$ calculation;
4) missing E$_t$ can be calculated after particle id and jet finding, or directly from cells/clusters;

jet finding

isolated particles

missing E$_t$

Jets

Figure 1: *General overview on the energy reconstruction sequence.*

which may be subjected to zero suppression, cell noise cuts, and topological noise finders. It is important to apply all selection criteria to the EMS signals. This allows to apply the same cuts to the simulated events (noise added). Note that some or all of these fundamental selections can even be done online using EMS, as present day's online data processors usually allow to perform the few simple calculations at appropriate speed.

Once the calorimeter cell list is reduced to significant cells, a general cluster finder is applied to collect cells into three dimensional objects. Signal related variables used to measure significance and other parameters used in this step should be on EMS, to allow for universal application in experiment and simulations. The cluster finder is usually tuned to find electromagnetic objects first, with the boundary condition that electron and photon response is kept in one cluster. It is imaginable, though, that once all electromagnetic clusters are found and the corresponding cells are flagged/removed from the cell list, another algorithm optimized for hadron response can be run on the remaining cells[5].

Step 3: signal classification

The next and important step is cluster classification, where each cluster is analyzed with respect to its electromagnetic or hadronic *nature*. Some of this classification can be done in the cluster finding step already, if the two step approach discussed above is chosen. In any case, three classes of clusters remain: electromagnetic, hadronic, and not classifiable. Details on these classes are discussed in section 5.

Step 4: final local energy scale

Next, calibration functions are applied to these clusters, depending on their nature, their location and other variables. The reconstructed energy assigned to a cluster is then the best estimate for the true energy deposited at this place, not yet the particle or jet energy. Additional corrections for dead material in front can be applied based on cluster location alone, or after particle identification and jet finding, see below. This depends on how general the dead material losses can be corrected, i.e. what is the optimal approach for resolution. Energy losses in cracks or due to dead channels etc. can maybe only be corrected at cluster level, but this depends strongly on specific calorimeter granularities and sampling characteristics.

Step 5: final physics calibration

The final step in the energy reconstruction includes particle identification and jet finding. Especially electromagnetic clusters are good candidates for single photons or electrons, even though the cuts applied to find them may be very relaxed compared to particle identification cuts. Once a cluster is accepted as isolated single photon or electron, another level of corrections may be necessary to determine the true particle energy, now certainly including dead material losses.

---

[5]thanks to J. Schwindling for pointing this out at the ATLAS LARG jet simulation meeting Jan. 29, 1997.

Jet finding is done on all clusters that are not clearly identified as single particles. This typically includes some electromagnetic clusters, certainly most of the hadronic and maybe some non-classified. A good approach to jet finding may be to treat all clusters as (possibly massive) pseudo-particles, i.e. assign 4-vectors to them.

3 Intrinsic Electromagnetic Energy Scale (EMS)

In section 2 we discussed the overall design of an energy reconstruction model for non-compensating calorimeters. One of the major ideas behind the model is the introduction of an intrinsic signal and a related energy scale, which is identical for testbeam and the final calorimeter, and allows to compare all experimental data with corresponding simulations. In this section we present a concept on how to determine the ideal electromagnetic calibration constant defining the intrinsic scale, and give some more details on the specific ideas and characteristics behind it.

The true or ideal electromagnetic calibration strategy has originally been suggested for the H1 liquid argon calorimeters [6, 7] and has been successfully implemented into the H1 energy calibration and reconstruction effort [1, 2].

3.1 Determination of the ideal electron calibration constant

The ideal electron calibration constant c_e can be defined by the average ratio of the ideal signal $\mathcal{Q}$, typically a charge, and the energy E_{dep} deposited by electrons inside a given calorimeter (following the definitions in section 2.1):

$$c_e = \langle\ \mathcal{Q}\ /\ E_{dep}\ \rangle \tag{1}$$

The real experimental signal Q can be rather close to $\mathcal{Q}$ on average, except for possible asymmetric noise, and not always correctable acceptance limitations like dead or hot channels. Upstream dead material and finite calorimeter dimensions usually affect E_{dep} in the same way as $\mathcal{Q}$, and have therefore no big effect on c_e.

The major influence on $\mathcal{Q}$, though, are analysis methods and cuts mainly applied to calculate Q within optimal resolution. These introduce a signal definition dependence of c_e, as they are only affecting Q, but not E_{dep}. In the sense of the universality requirement to c_e, these dependencies are not acceptable.

The normalization E_{dep} is often not known very well experimentally, even for electrons in a testbeam. Especially upstream dead material can introduce severe relative energy losses at lower energies, introducing an effective non-linearity and therefore a large error on c_e and the intrinsic energy scale. Limited geometrical acceptance of calibration modules can add to this effect.

Unfolding of this uncertainty is possible, though, within remaining systematics of about 1%, with electron response simulations featuring a detailed description of the calorimeter and the materials around it. A strategy to achieve this is described in the following.
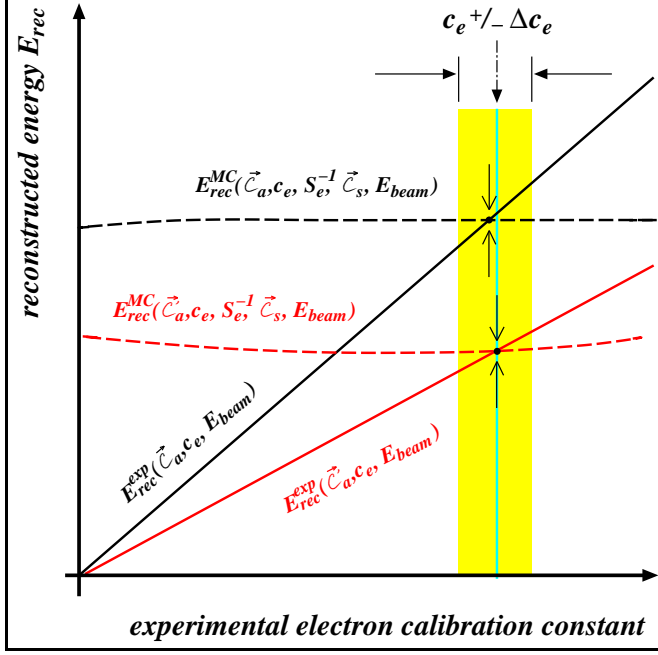
Figure 2: *Schematic view on the determination of the ideal experimental electron calibration constant c_e from the energies reconstructed with different analysis methods and cuts, denoted by $\vec{\mathcal{C}}_a$ and $\vec{\mathcal{C}}'_a$, at different electron energies E_{beam}, for experiment (E_{rec}^{exp}) and simulation (E_{rec}^{sim}).*

Equalizing data and simulation response

One way to extract c_e within all deficiencies described above is based on the rather trivial assumption that simulated electron response should describe the experimental response with all respects. This includes the average response, energy resolution, shower profiles and space resolution. The only parameter explicitly useful for the determination of any calibration is the average response, at a given E_{beam} measured as average reconstructed EMS energy $\langle E_{rec}^{exp} \rangle$ in the experiment, and as average reconstructed EMS energy $\langle E_{rec}^{sim} \rangle$ in the simulation.

For each event E_{rec}^{exp} can be related to the experimental signal Q using the yet unknown ideal calibration constant c_e:

$$E_{rec}^{exp} = c_e \cdot Q, \qquad \text{with} \qquad Q = \underbrace{\sum_i^N q_i}_{\text{selected cells}} \tag{2}$$

where Q has to be calculated from fully corrected charge signals q_i in calorimeter cells[6]. These charges have to be reconstructed using the electronics calibration, thus removing pre-amplifier gain fluctuations and pedestal drifts. Typical additional corrections are for charge collection inefficiencies and cross talk. It is only after these corrections that q_i has the meaning of a charge scale, independent of the testbeam setup and therefore the same for the final experiment. Note that Q can be the sum on cells selected by noise cuts and/or clustering; E_{rec}^{exp} naturally depends on these selections in the same way Q does.

---

[6]we assume each calorimeter cell corresponds to one electronic channel, for simplicity.

The simulated response E_{rec}^{sim} is also calculated event by event using

$$E_{rec}^{sim} = \underbrace{\sum_i^N E_{rec,i}^{sim}}_{\text{selected cells}} , \qquad \text{with} \qquad E_{rec,i}^{sim} = S_e^{-1} \cdot E_{vis,i} + c_e \cdot q_{noise,i} \qquad (3)$$

$E_{vis,i}$ is the simulated cell signal, while $q_{noise,i}$ is the noise equivalent charge from the experiment, added to each cell in the simulation. This charge has to be corrected in exactly the same way as the experimental cell signal q_i. Note that the experimental cell selection criteria have to be applied to $E_{rec,i}^{sim}$ for correct modeling.

The inverse electron sampling fraction S_e^{-1} is derived from the visible E_{vis} and deposited energies E_{dep} for electrons at a given E_{beam}:

$$S_e^{-1} = \langle\, E_{vis} \,/\, E_{dep}\, \rangle^{-1} \approx (\langle E_{vis} \rangle \,/\, \langle E_{dep}\rangle)^{-1} \qquad (4)$$

in a full acceptance simulation (no cell selection, for example).

One important point should be mentioned here. Even though E_{vis} corresponds to E_{dep} and not to E_{beam}, and has therefore the feature of being correctly normalized intrinsically, one still has to make sure that the amount of dead material in the testbeam is at minimum. S_e^{-1} is only an energy independent constant if the major part of the electromagnetic shower can be sampled inside the active calorimeter (especially the shower maximum) [7].

The ideal experimental calibration constant c_e is then constraint by the most obvious requirement that for identical E_{beam} the average response for experiment and simulations have to be identical *for all possible signal definitions and analysis cuts* $\vec{\mathcal{C}}_a$, applied in the same way to both kind of data:

$$\left\langle\, E_{rec}^{exp}(\vec{\mathcal{C}}_a, c_e, E_{beam})\, \right\rangle \;\overset{!}{=}\; \left\langle\, E_{rec}^{sim}(\vec{\mathcal{C}}_a, c_e, S_e^{-1}, \vec{\mathcal{C}}_s, E_{beam})\, \right\rangle \qquad (5)$$

$\vec{\mathcal{C}}_s$ represents the set of particle tracking energy cutoffs and other possible simulation conditions. If these are the same as the ones used in the determination of S_e^{-1} (see eq. (4)), one expects that the dependence of $E_{rec,i}^{sim}$ on $\vec{\mathcal{C}}_s$ basically factorizes out, see eq. (3). This has been observed for cutoffs within reasonable ranges.

The actual determination of c_e at a given beam energy can be performed iteratively, with given analysis conditions $\vec{\mathcal{C}}_a$. Note that even the qualitative dependence of E_{rec}^{sim} on c_e is non-trivial, especially if noise cuts and other signal definitions based on significance are applied to data and Monte Carlo, see figure 2. The same figure shows a graphical solution for the determination of c_e: a fixed charge scale signal $Q = Q(\vec{\mathcal{C}}_a)$ leads to a linear dependence of E_{rec}^{exp} on c_e, where the slope is just Q. The intersection $\langle E_{rec}^{exp}(c_e, \ldots)\rangle = \langle E_{rec}^{sim}(c_e, \ldots)\rangle$ obviously defines c_e. Changing the analysis cuts $\vec{\mathcal{C}}_a$, or even going to another beam energy, certainly changes E_{rec}^{exp} and E_{rec}^{sim}, but ideally not their intersection in c_e.

Whatever residual dependencies of c_e on E_{beam} or the analysis cuts remain, reflects the deviation from (intrinsic) signal linearity for electrons, defined by

$$\frac{\left\langle E_{rec}^{sim}\left(c_e, E_{beam}, \ldots\right)\right\rangle - \left\langle E_{rec}^{exp}\left(c_e, E_{beam}, \ldots\right)\right\rangle}{\left\langle E_{rec}^{exp}\left(c_e, E_{beam}, \ldots\right)\right\rangle}$$

It therefore defines the basic systematic error Δc_e on the whole calibration and energy reconstruction, as EMS is our fundamental signal scale. To estimate this error safely requires several rounds of analysis at various E_{beam} and with changing noise cuts and cluster definitions. Also, most signal significance based cell selection algorithms implicitly introduce sensitivities to (especially radial) shower development. Monte Carlo has therefore to describe this feature as well as the average response, which is usually the case for the electromagnetic shower model implemented in GEANT 3 [8], see electron shower profiles in [1] (also for the determination of non-linearities and details on systematic errors).

3.2 Reconstruction on the electromagnetic energy scale

The reconstruction of EMS calorimeter signals happens at cell level, prior to any data reduction and noise suppression depending on cell signal significances and/or correlations. Practically this means it has to be done in the online data processing for the physics experiment, and at a very early stage in the reconstruction of the corresponding physics simulation, or testbeam data and simulation.

As already discussed in the previous section, and shown in figure 3, the very first step towards the EMS signal is the conversion of the digitized experimental signal to a physical quantity like charge. This step is performed by the appropriate electronics calibration, including corrections for imperfections in the analog or digital signal handling. These are often parameterized as functions of the digitized signal itself.

Corrections for charge collection efficiencies and cross talk could be performed at the same level, but they are often better controlled if the base for parameterization is the charge conversion described before is already done. In any case, the signal q_i in a given calorimeter cell i is on an absolute and fully corrected charge scale after these procedures have been applied.

What is left for the experimental signal in this cell is just applying the ideal electron calibration constant c_e to convert it to an energy $E_{rec,i}^{exp}$ reconstructed on EMS (similar to eq. 2):

$$E_{rec,i}^{exp} = c_e \cdot q_i$$

The basic conversion of the simulated signal to EMS is already described in eq. (3). We suggest to follow the strategy of adding experimental noise instead of modelling it. The reason is that especially deviations from Gaussian behaviour in experimental noise can often not be modelled adequately (descriptions of tails). Adding is more straight forward, but requires experimental access to noise, i.e. a set of randomly triggered "empty" events with non-zero suppressed read out. These are often easily available in testbeams, where noise is
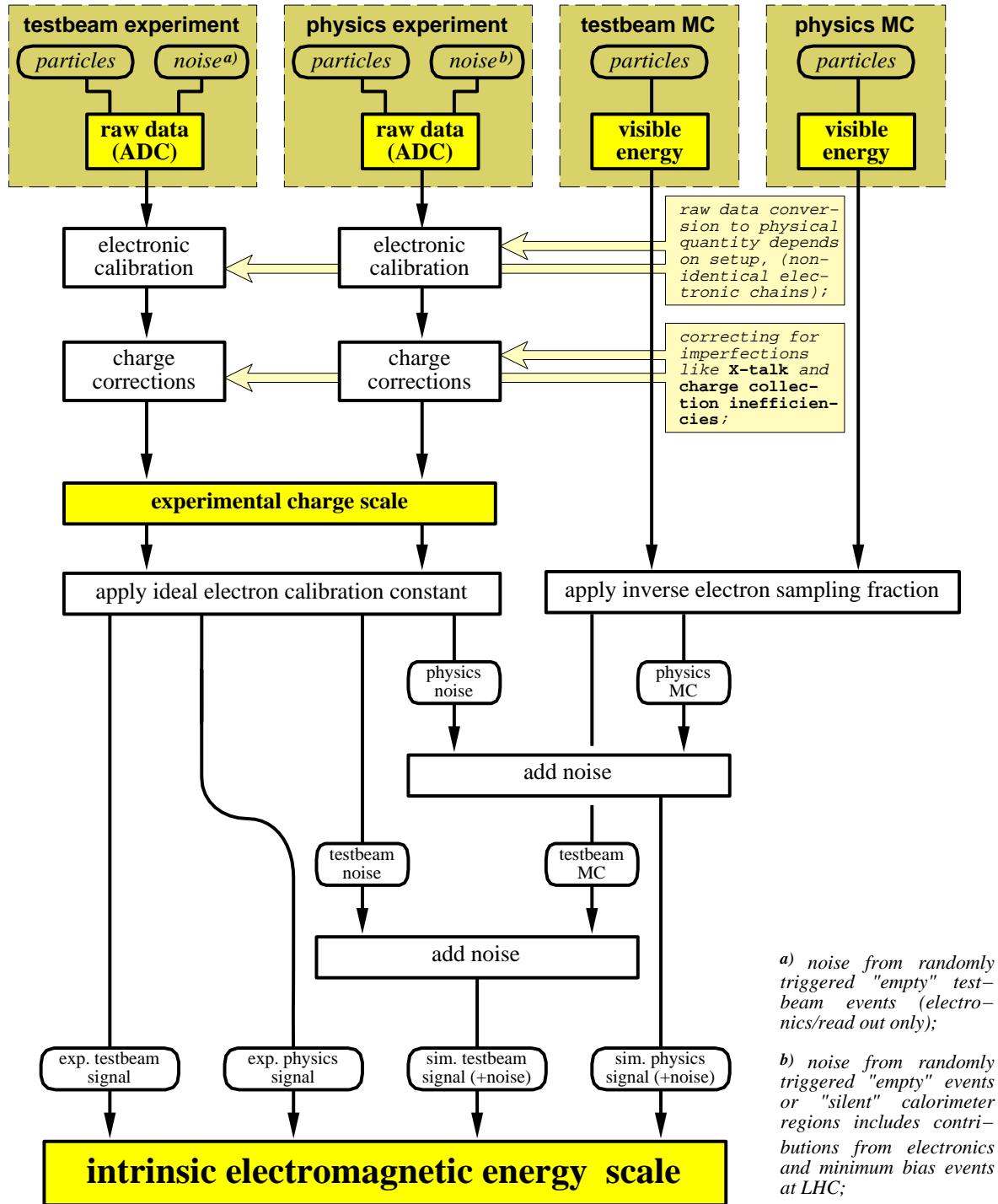
Figure 3: *Cell level energy reconstruction on the electromagnetic energy scale.*

typically generated by the electronics. Extracting the experimental noise at the LHC is not that straight forward, as it includes minimum bias physics. One possible strategy is maybe to look into "quiet" regions of triggered physics events, but this requires non-zero suppressed calorimeter read out.

4   Signal Definition

Our understanding of "signal definition" is not really related to specific demands from reconstruction of *certain* physics, more so to general calorimeter performance optimization for *all* possible physics signatures. This can also be understood as optimal use of calorimeter signals in a given read out granularity, allowing for best possible noise suppression and consequently, signal significance enhancement (see also remarks in section 2.1). In this section we present a possible concept along this line. The general sequence in this approach is given in figure 4.

As for most models collected in this paper, some ideas have been realized for the H1 liquid argon calorimeter. Some details on the topological noise suppression discussed in section 4.1 have been published [3, 4]. The same references contain some description of the H1 clustering algorithms, which is the major input for the model suggested in section 4.2. The most complete description of a possible cluster strategy and details on possible selection criteria and cuts are, at least to our knowledge, only available as an H1 internal note [9].

4.1   Fundamental cell selection

The first step in the signal definition process is usually performed online during data taking, by applying a zero suppression algorithm based on a signal-to-noise cut. This obviously requires knowledge of the 1σ noise[7] in each calorimeter read out cell. Another possible algorithm and maybe more appropriate for LHC applications, is baseline subtraction, where the background noise is determined by pre and post sampling around the signal. A cell selection based on the signal-to-background ratio or the signal-to-background difference can then be applied. Note that in this and the following discussions any signal fluctuations from electronic noise or physics pile-up are generally referred to as "noise". Also, all variables related to cell signals and used in this or the following cell selection algorithm should be measured on EMS, to allow to apply the same selections to simulated data with noise added. This can often be achieved even in the online signal processing, as all variables in noise look-up tables can be stored on this energy scale, and the signal can quickly be converted.

In addition to this basic noise suppression technique it is often necessary to apply a topological signal finder. Such an algorithm can be a very simple pattern recognition consisting of two steps: first, cells with highly significant signals, measured by a signal-to-noise estimator or an absolute energy or transverse energy cut, are selected. Then, neighbours around these cells which have less significant signals above a given lower threshold are also kept. This

---

[7]depending on the actual calorimeter *rms* noise may be better defined, and can be used in the same sense as 1σ noise in the following discussion.
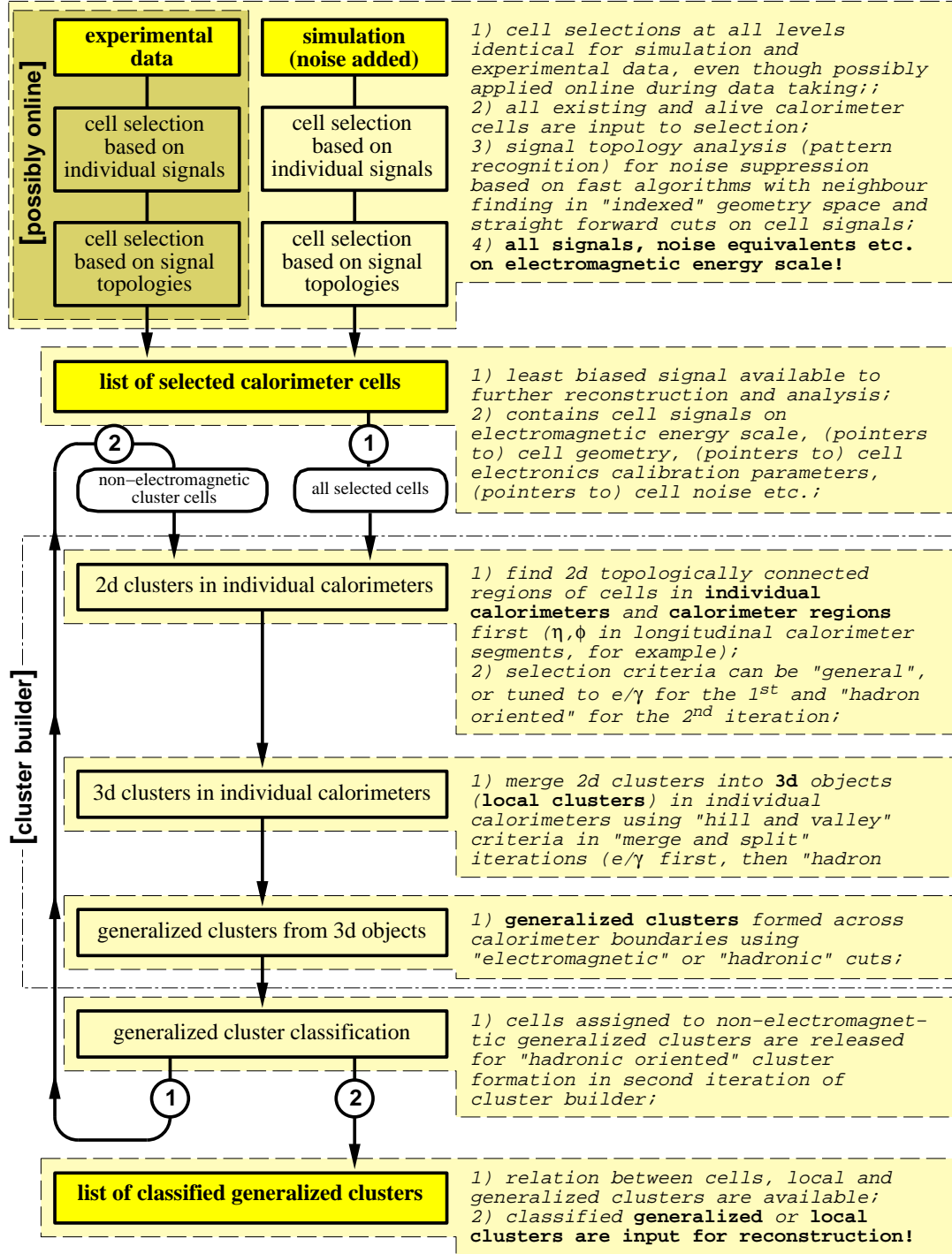
15

experimental data

simulation (noise added)

cell selection based on individual signals

cell selection based on individual signals

cell selection based on signal topologies

cell selection based on signal topologies

1) cell selections at all levels identical for simulation and experimental data, even though possibly applied online during data taking;;
2) all existing and alive calorimeter cells are input to selection;
3) signal topology analysis (pattern recognition) for noise suppression based on fast algorithms with neighbour finding in "indexed" geometry space and straight forward cuts on cell signals;
4) **all signals, noise equivalents etc. on electromagnetic energy scale!**

list of selected calorimeter cells

1) least biased signal available to further reconstruction and analysis;
2) contains cell signals on electromagnetic energy scale, (pointers to) cell geometry, (pointers to) cell electronics calibration parameters, (pointers to) cell noise etc.;

② non−electromagnetic cluster cells

① all selected cells

[cluster builder]

2d clusters in individual calorimeters

1) find 2d topologically connected regions of cells in **individual calorimeters** *and* **calorimeter regions** *first (η,ϕ in longitudinal calorimeter segments, for example);*
2) selection criteria can be "general", or tuned to e/γ for the 1st and "hadron oriented" for the 2nd iteration;

3d clusters in individual calorimeters

1) merge 2d clusters into **3d** *objects (***local clusters***) in individual calorimeters using "hill and valley" criteria in "merge and split" iterations (e/γ first, then "hadron*

generalized clusters from 3d objects

1) **generalized clusters** *formed across calorimeter boundaries using "electromagnetic" or "hadronic" cuts;*

generalized cluster classification

1) cells assigned to non−electromagnet-tic generalized clusters are released for "hadronic oriented" cluster formation in second iteration of cluster builder;

① ②

list of classified generalized clusters

1) relation between cells, local and generalized clusters are available;
2) classified **generalized** *or* **local clusters are input for reconstruction!**

Figure 4: *General overview on the signal definition sequence.*

selection removes the majority of all cells surviving the strictly cell signal based cut discussed above by requiring a rather large (positive!) signal for isolated cells. On the other hand it does not cut into individual particle showers too much by specifically taking into account cell signal correlations, and keeping cells with less significant positive or negative signals.

A typical example for this strategy is to find cluster seed cells with signals E_i on electromagnetic energy scale, and adjacent cells with signals E_j on the same scale by:

$$\frac{E_i}{E_{noise,i}} > a, \quad \text{with } a = 3 \dots 4, \quad \text{and} \quad \left| \frac{E_j}{E_{noise,j}} \right| > b, \quad \text{with } b = 2 \dots 3$$

where $E_{noise,i}$, $E_{noise,j}$ are the 1σ EMS noise equivalents. Note that in this example negative signals are kept to "compensate" for a possible bias towards positive signal contributions from noise - but only if they are close to cells with significant positive signals. Isolated cells with an even highly significant negative signal should be removed as pure noise contribution.

Online application of the topological noise suppression, as well as implementation in offline reconstruction, requires a fast neighbour finder. Whatever solution is adapted for this task (database with list of neighbours for each cell, analytical determination of this list using geometry indices, etc.), topological noise suppression should therefore rather use an indexed than a real coordinate space. This usually restricts neighbouring cell finders to the geometrical boundaries of a given calorimeter, or calorimeter region with a certain read out geometry.

It is certainly necessary to understand the cell signal based and the event topology dependent noise suppression discussed here in the LHC environment at ATLAS. Especially the acceptance for negative isolated or non-isolated signals has to be tuned to optimum performance with least possible biases to any given physics channel. Another important point is the actual choice of variables to define signal significance. From point of shower development and electronics noise, energy based cuts are favoured. Noise from physics pile-up at LHC scales more with transverse energy [10], though, and may require this variable to be used for significance definition.

4.2   Cell clustering

The cells surviving the fundamental noise suppression discussed above are then subjected to a cluster finder. Several strategies are possible at this level. The design of the optimal algorithm is certainly depending on the calorimeter granularity. Ultimately there should only be one strategy, and it should use the EMS cell signal. This again allows universal application to data and simulations.

Clustering itself, as a possible signal definition, is mainly motivated by the following points:

- enhance signal over background from pile-up or other noise sources by using cell signal correlations from the shower development;

- separate electromagnetic from hadronic energy deposit as efficient as possible within a given calorimeter granularity, for energy reconstruction within optimal resolution;

- locate and identify isolated particle response in the calorimeter, especially for electrons and photons, and maybe muons;

- and in the same sense, the optimal use of the calorimeter for two particle separation, especially interesting for photons.

It is important to note that in contrary to the topological noise suppression algorithm, which can work in an indexed geometry space (see previous section 4.1), clustering usually requires real space coordinates. This is a natural consequence out of the idea that clusters have a physical meaning, i.e. are following electromagnetic or hadronic shower development. It is also needed for generalization, meaning the combination of cell clusters found in individual calorimeters or calorimeter regions with different read out geometries or technologies, to larger objects. This is most obvious in setups where the reference system for a given calorimeter granularity and/or its (quasi) pointing behaviour changes from rapidity and azimuthal coordinates to some linear space coordinates like radius and azimuth, for example.

Global cluster builder strategy

Clustering in our model consists of several layers, see figure 4. As already mentioned, the input to the first level is a list of cells with geometries and EMS signals, which passed the noise suppression algorithms discussed above:

1. *local 2d clustering*

    Topologically connected cells in each longitudinal segment of each calorimeter are clustered to form two dimensional regions around cluster seeds. Those are cells with highly significant signals.

2. *local 3d clustering*

    These regions are connected longitudinally to larger three dimensional objects by applying some overlap criteria. This process is typically still limited to a certain calorimeter or calorimeter region. Some of the originally found 2d clusters may be broken up again in this process, to optimize the 3d cluster building.

3. *generalized 3d clusters*

    The locally found 3d objects are combined to even larger clusters, now crossing the boundaries of a certain calorimeter or region. The generalized clusters are often confined within a given calorimeter, and therefore identical to the locally found 3d objects.

We suggest two iterations of these three step approach, where in the first iteration cell clusters are built using criteria optimized to electromagnetic shower development (figure 4).

This means that the corresponding algorithms should be tuned in such a way that cells from isolated electron or photon response are collected into exactly one cluster at high efficiency.

After this "electromagnetic cluster building", the found objects are classified with respect to their nature. Again, the efficiency of this classification depends strongly on the calorimeter granularity, or, to be more exact, on its granularity with respect to the electromagnetic and hadronic (lateral) shower sizes. We expect three different classes of clusters:

electromagnetic clusters are usually characterized by their compactness and the early shower starting point. They are typically located in the electromagnetic calorimeter, but may extend into the first longitudinal segment of the hadron calorimeter. These characteristics can be tested by three signal ratios, for example [12], each calculated from cluster quantities:

- ratio of signal summed over cluster cells located in the electromagnetic calorimeter to total cluster signal;
- ratio of signal summed over cluster cells located in the first longitudinal calorimeter segment to total cluster signal;
- ratio of signal summed over n cells with highest signals or signal significances (*hot spot*) to total cluster signal.

Actually definitions of these ratios certainly depend on the actual calorimeter, especially concerning the number n of cells contributing to the hot spot in the cluster. Cuts on these variables are tuned to minimize the number of clusters for electrons and photons, as mentioned above. This strategy only aims at clusters originating from primary photons or electrons in jets, and isolated particles. Finding intrinsic electromagnetic subshowers in hadronic showers is often not possible in realistic calorimeters, as it requires a very fine and often impractical read out granularity.

hadronic clusters can not be characterized as straight forward as the electromagnetic ones. Usually one can define all non-electromagnetic clusters as being hadronic. If a signal weighting scheme is applied for final hadron calibration, one should require a certain significance of the total cluster signal with respect to the noise in the corresponding cells. Also, weighting may be parameterized as function of different cluster variables (find a description of some possible parameters at the end of the paragraph on generalized clusters on page 22). Hadronic clusters should be large enough to allow calculation of these numbers. This may actually determine the cuts applied in the second, "hadronic" cluster building iteration: those can be tuned to construct clusters that meet optimal energy resolution requirements (optimize number of hadron clusters in pion response, for example).

non-classified clusters are those whose nature could not safely be determined in the previous two iterations of cluster building. This class typically contains clusters consisting of very few cells (typically one) with a significant signal, from which none or not all

of the variables used to determine the cluster nature could be safely evaluated. Other members of this class are clusters with insignificant total signals.

After classification cells collected into electromagnetic clusters are marked. Cells not assigned to electromagnetic clusters are subjected to the hadronic cluster builder in the second iteration. After this all cells should be assigned to one of the cluster classes, for final calibration purposes.

Two dimensional clustering: significance versus acceptance

Without presenting a fully detailed algorithm – which is far beyond the scope of this note – we would like to recollect some features of the individual steps to generalized clusters.

The very first step, the finding of topologically connected signal regions within a longitudinal calorimeter segment, looks at first very much like the pattern recognition applied in the topological noise suppression discussed earlier. Some differences are imaginable, though, especially concerning the seeds for these regions.

Like in jet finding, two scenarios with different biases on physics reconstruction can be explored. The first one is the approach of *constant acceptance*, where cluster seeds are defined as cells with energies E_i or transverse energies $E_{t,i}$ above a constant threshold E_{min} or $E_{t,min}$, respectively. The choice of variables depends on calorimeter location and on how global E_{min} or $E_{t,min}$ are supposed to be defined.

The thresholds in this approach are often defined to be safely above the expected noise. They will naturally not be changed very frequently, therefore allow for a constant acceptance. E_{min} or $E_{t,min}$ may certainly also depend on calorimeter location, cell sizes and other specific parameters depending on energy flow from physics and read out geometry.

The advantage of this approach is that the simulation of physics events is much more straight forward, especially if E_{min} ($E_{t,min}$) are very stable numbers. The disadvantage is maybe that the signal significance is not constant, and some sensitivity to physics may be lost if these thresholds are set too high. Generally the approach of constant acceptance is more suitable for a well understood calorimeter with rather constant noise conditions.

Alternatively to this approach one can think about an approach where initial cluster finding is based on signal significance rather than absolute magnitude. This is realized in the model of *constant significance*, where cluster seeds are those cells with E_i/E_i^{noise} ($E_{t,i}/E_{t,i}^{noise}$) above a certain threshold. E_i^{noise} ($E_{t,i}^{noise}$) is the noise equivalent (transverse) energy of this cell, measured on the electromagnetic energy scale, like the signal itself.

This approach basically allows to easily adjust the detector sensitivity to the present noise conditions within the frequency of updates of the cell noise database. Naturally the acceptance in (transverse) energy is changing with these updates. The advantage is, though, within this time base it is always at maximum, which in turn means continuously high physics sensitivity.

The disadvantage of a constant sensitivity approach is that it may introduce more complex treatment of simulations to keep up with more frequent background changes – even though

this is only important at the reconstruction level. The constant acceptance approach also requires adequate noise to be added to simulations, following "true" experimental conditions.

The constant sensitivity model seems to be more appropriate for early physics running, where background conditions may change more often.

In any case, topologically connected cells can be found with very general cuts, not even tuned to specific particle response, by finding seeds in two dimensions and keep all cells around them.

Three dimensional clustering

Collecting cells into three dimensional objects is more depending on cut optimizations with respect to certain shower shapes. As already lined out in the overview given on page 18, we suggest a two iteration approach[8]: first build clusters using parameters optimized for electron/photon response, and remove clearly identified electromagnetic objects. Second, apply a hadronic cluster builder on the cells not belonging to electromagnetic clusters.

It is basically impossible to discuss any specific solution for $3d$ clustering without referring to an actual calorimeter design. Some of the aspects generally included in the possible strategies are definition of overlaps between $2d$ objects along the direction of flight of the incoming particles, and (angular or radial) distances between this direction and cluster seeds. Also, the actual projective (angular or radial) distance between cluster seeds themselves can be used. Choices of angular or radial distances are depending on conical or cylindrical approaches. The first one is more oriented towards energy flow in jets, while the cylindrical approach tries to follow single particle shower development.

Two dimensional objects can certainly be split in the process of $3d$ cluster building, and the resulting or original areas can be merged, depending on chosen selection criteria. Usually we expect that this process depends on electromagnetic or hadronic "orientation" of the cluster builder, i.e. on the step in the iteration process discussed earlier. Split and merge parameters are often related to the spatial signal distribution, where the (relative or absolute) depth of "signal valleys" and the slope of "signal hills" can be used. The choice of energy or transverse energy – both measured on EMS – is depending on the orientation of the cluster builder, and maybe related to the projective read out geometry of the calorimeter.

Generally we expect that clustering at this level does not cross calorimeter boundaries. This is not really required, though, in case of adjacent calorimeters with about the same significance (!) of EMS signals for hadrons, practically meaning comparable e/π signal ratios.

Generalized clusters

As suggested above, $3d$ clusters are usually formed within the boundaries of a given calorimeter or calorimeter module first. Whether or not they are combined to larger generalized clusters depends on the specific design of boundaries between calorimeters or calorimeter modules, and the read out granularity in these regions. Generally we expect that if the amount

---

[8]it should be mentioned that the $2d$ selection may also be done with two differently tuned sets of cuts.

of material between two calorimeters is small, like between the electromagnetic and hadronic endcap modules in ATLAS [11], generalized clusters may extend beyond this boundary.

Combining two adjacent clusters in two different calorimeters to form a generalized cluster across such a boundary can even be reasonable in case of a relatively deep and massive crack, like (maybe ?) between the barrel electromagnetic and hadronic calorimeters in ATLAS [11], especially if each cluster provides enough information to correct for the energy loss in the inactive material. This requires a sufficient read out granularity of the calorimeter regions around this crack. As already pointed out in the introduction of this section, real space coordinates are certainly advantageous for any generalization attempt.

After signal definition, generalized as well as local $3d$ clusters are described by a set of variables, some of which are:

- the total EMS cluster signal E_{cl}, defined as

$$E_{cl} = \sum_{i=1}^{N_{cl}} E_{rec,i},$$

  where N_{cl} is the number of cells in this cluster, and $E_{rec,i}$ the EMS cell signal;

- the cluster location, defined by a set of indices for calorimeter region, segment etc., and/or the cluster centre of gravity $\vec{X}_c = (x_{cl,1}, x_{cl,2}, x_{cl,3})$ in some real space coordinates;

- the cluster size, defined by the volume V_{cl}, some area A_{cl} projected on a given direction (like the one defined by the event vertex and $\vec{X}_{cl}$), and the corresponding dimensions in Cartesian space $(\Delta x, \Delta y, \Delta z)$. Linear $(\Delta \theta \times \Delta \phi)$ or non-linear $(\Delta \eta \times \Delta \phi)$ angular dispersion may be used in addition. Cluster dimensions can also be expressed by

$$\sigma_{x_{k,cl}} = \frac{1}{E_{cl}} \cdot \sum_{i=1}^{N_{cl}} \left[E_{rec,i} \cdot (x_{k,i} - x_{k,cl})^2 \right], \quad \text{with } x_k = x, y, z, \theta, \phi, \eta, \ldots \text{ in}$$
$$\text{cells } (x_{k,i}) \text{ or clusters } (x_{k,cl});$$

- the global signal density $\rho_{cl} = E_{cl}/V_{cl}$, in addition to the cell signal densities $\vec{\varrho} = (E_{rec,1}/V_1, E_{rec,2}/V_2, \ldots, E_{rec,N_{cl}}/V_{N_{cl}})$, where V_i is the cell volume ;

- the cluster shape $\delta E_{cl}/\delta x_{cl,k}$ with respect to some of the coordinates discussed above;

- and the cluster nature, defined by signal fractions as discussed earlier for electromagnetic clusters (page 19).

As usual, all signal related variables should be calculated on EMS.

5 Final Energy Calibration and Reconstruction

Final energy reconstruction consists of two basic tasks. First, the reconstructed EMS energy must be converted to the best estimate of the true deposited energy at a given cluster location. This conversion is achieved by local calibration functions and parameters, sensitive to the nature of the energy deposit.

The second step is then to reconstruct particle and jet energies, and missing transverse momentum or other global event quantities. Input to this step are clusters or cells after final local calibration.

The final calibration depends for both tasks on the signal definition. This introduces similar restrictions like the ones already presented in the sections on clustering, on a general discussion of the final calibration and reconstruction. Especially the calibration functions are very specific not only to the calorimeter geometry and read out granularity, but also to some hardware performance aspects like e/π. Suggestions for the H1 calorimeter can be found in [3, 4] and, in more detail, in [13]; calibration functions inspired by this approach have been already been successfully applied to data from a combined ATLAS liquid argon and tile calorimeter testbeam [14] and studied for simulated hadron response in the ATLAS hadronic endcap liquid argon calorimeter [15].

5.1 Final local calibration

The relation between true deposited energy in a given cell cluster and the corresponding EMS signal is certainly depending on the nature of the deposit[9] (hadronic or electromagnetic), e/π signal ratio, the cell selection cuts, and the cluster finding and classification efficiency. The calibration tasks generated by these influences are discussed in the following paragraphs.

Electromagnetic cluster calibration

Once a cluster is classified as electromagnetic, its total EMS signal E_{cl} may already provide a good estimate of the "true" energy E_{dep} deposited by electrons or photons. This is especially true if the cluster is not located too close to a calorimeter crack.

Still, E_{cl} is of course affected by the signal definition cuts, as already discussed in sections 3 and 4. Electron response in testbeams, though, usually allows to understand and correct these effects at a level of 1% [1], often together with simulations. Constraints for the corresponding calibration functions are optimized signal linearity and energy resolution for a given signal definition.

Local calibration in its most limited understanding does not require any correction for energy losses outside the cell cluster, as its whole philosophy is to reconstruct the energy at the cluster location. But the correction for signal definition cuts already softens this picture

---

[9]muons are not included here or in previous discussions. They can be treated like isolated electrons or photons, i.e. the appropriate calibration function is applied after identification, and clusters belonging to muon response are not available for jet finding.

Figure 5: *Electromagnetic cluster calibration normalization. The cluster EMS signal E_{cl} is defined by the sum of signals in clustered cells and must be normalized for calibration using the sum of deposited energies in all cells hit by the shower.*
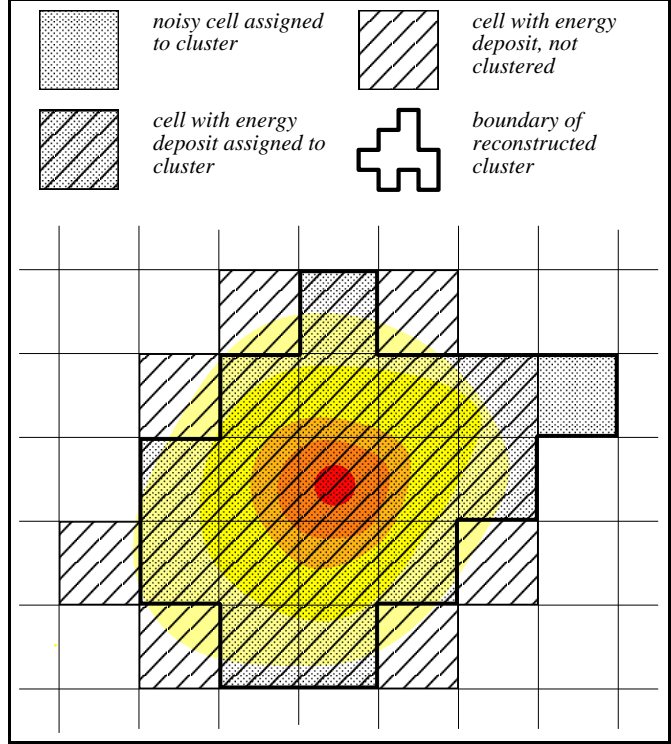
at least for electrons and photons, as the signals in calorimeter cells surviving the signal definition are used to reconstruct the energy deposited in a full electromagnetic shower, i.e. to correct for lost cells with energy deposit but small signal or large noise, and for added noisy cells without energy deposit. See also figure 5.

The same argument can be made for dead material losses in electromagnetic showers, with one additional observation: the effect of signal definition cuts on the electron and photon calibration is the same in testbeam and physics experiment, while especially the inactive material in front of a certain calorimeter may be different for both setups, resulting into different correction functions and parameters. The calibration functions correcting noise cuts and cluster finding should therefore be determined independently of any corrections for energy losses in dead material.

Another important difference can be that the corrections for the signal definition can be determined experimentally, at least for electrons. Simulations are only needed for monitoring the extracted calibration functions and estimation of the usually small (by comparison) energy losses in upstream material.

Hadronic calibration functions

Hadronic calibration functions often introduce complex non-linear signal weights for optimal energy resolution in non-compensating calorimeters. Their extraction from testbeam data is

24
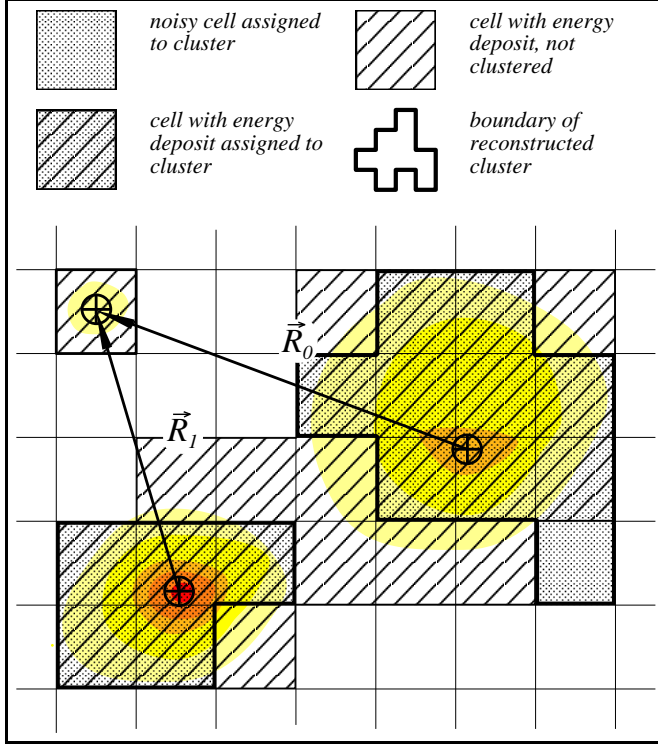
Figure 6: *Normalization for hadronic cluster calibration. Simulated energy deposit in cells with no or insignificant signals can be assigned to the nearest cluster in space, measured by the (energy weighted) distance $\left|\vec{R}_i\right|$ between cluster/cell barycenters.*

not at all straight forward, as they are much more sensitive to signal definition and limited acceptance of even large calorimeter calibration modules. And even if these functions are determined by testbeam pions, one still has to find a way to apply them to jets, including a verification of validity in this different environment. Differences in inactive upstream material are usually less severe for hadron calibration normalization than for electrons.

The energy reconstruction model discussed here is based on local calibration, i.e. the hadron calibration functions are applied only to hadronic clusters, and are normalized to the energy deposit at the cluster location. The non-classified clusters mentioned in section 4.2 on page 20 are usually part of the hadronic response, but also artificially generated by cluster finder inefficiencies. They are discussed in the next paragraph.

The normalization problem for hadronic energy deposit is dominated by signal definition cuts. Even hadronic showers from single pions in a testbeam environment are often split into several clusters with very different characteristics. This is mainly due to the fact that the correlations between signals even in adjacent calorimeter cells can be relatively weak. Also, hadronic clustering should be tuned to the diversity observed in hadronic showers, meaning that event by event variations in the shower development are actually measured and used for optimal reconstruction at highest energy resolution. In addition this is a possible way to carry hadronic cluster calibration obtained from single pion testbeam data to jets: localized clusters with the same characteristics, measured by adequate variables like signal densities, cluster shapes etc., are reconstructed by the same functions in both kind of data.

25

The weak signal correlation and the related wide shower spread increase the probability of fairly isolated small energy deposits in the development of the hadronic shower, which may not even be converted into any significant signal, resulting into the possible loss of the corresponding calorimeter cell at any of the signal definition steps. On the other hand, these energies can often not be neglected in the normalization of the calibration functions, mainly for signal linearity at lower energies. One way to include them is the use of simulations, where the deposited energy is available cell by cell, and can be assigned to the closest cluster, see figure 6.

Input to this approach can be single particle [7] or jet response simulations [2, 13]. Both are sensitive to hadronic shower models and certainly need systematic evaluation. Testbeam data usage is then limited to tuning of the shower Monte Carlos and to monitoring the performance of the calibration functions with respect to energy resolution and linearity, the later defined by comparisons to simulations [2].

Alternatively one may use a global normalization based on the beam energy, where deposited energy is assigned to each cluster based on its contribution to the total event signal (or other appropriate variables). This concepts often suffers from limited acceptance of calibration modules, which must be understood, for example from testbeam simulations.

The question of what kind of calibration functions are appropriate can not generally be answered. Experience with the reasonably segmented H1 liquid argon calorimeter shows that the signal definition including the filtering of electromagnetic clusters, which is part of a signal weighting scheme (!), is actually dominating linearity and resolution for lower energetic jets with (in linear space) rather wide opening angles [3]. This may change at higher energies, like at ATLAS, even though we expect to benefit from typically large signals requiring only relatively straight forward calibration functions. Concerning the effect of these functions on energy resolution, it is imaginable that the relatively large noise at LHC makes the signal definition itself dominating this performance parameter.

Non-classified clusters

The deposited energy in clusters consisting of an inadequate number of cells (mostly one) to calculate variables needed as input for hadronic calibration functions can not really be estimated safely by any means. Some limited guesses on the nature of the energy deposit are maybe possible from the cluster location: electromagnetic deposit is more probable close to the front of the calorimeter, while isolated signals deeper in the calorimeter can rather safely be assumed to be of hadronic origin. But even these coarse evaluations do not really help to make a safe guess on the relation between signal and deposited energy.

Calibration functions for these objects can be extracted, however, from simulations or global normalization attempts in testbeam or physics experiment. The later can be constraint by resolution optimization in transverse momentum balance, for example. In any case, these functions can only provide an estimate for the deposited energy within a fairly poor resolution; they are therefore typically limited to straight forward constant factors or slightly signal dependent corrections.
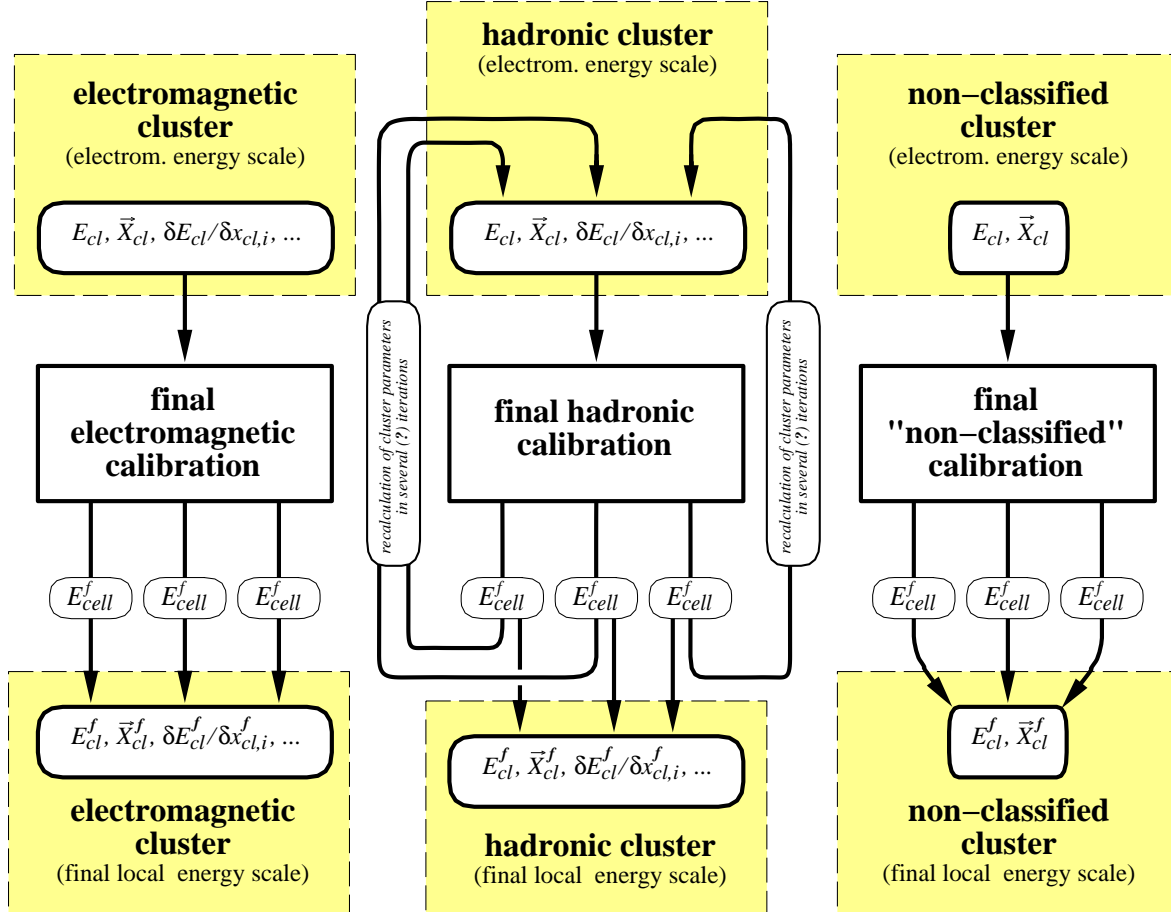
Figure 7: *Final local energy reconstruction scheme. Each cluster is characterized by a set of variables, like the cluster energy signal $E_{cl}^{(f)}$, the cluster location $\vec{X}_{cl}^{(f)} = \left(x_{cl,1}^{(f)}, x_{cl,2}^{(f)}, x_{cl,3}^{(f)}\right)$, the cluster signal shape $\delta E_{cl}^{(f)}/\delta x_{cl,i}^{(f)}$, and others. $E_{cell}^{(f)}$ is the cell energy signal. Superscript f denotes FES signals (usually the best estimate for the energy deposited at the cluster location), or quantities derived from this signal. No superscript means signals and related variables on EMS.*

5.2 Local energy reconstruction

Once the calibration functions and parameters for each cluster class have been determined, energy reconstruction follows the scheme lined out in figure 7. The EMS cluster signals are converted to the FES signals using the appropriate calibration procedure.

For electromagnetic objects, this typically includes a rather straight forward correction for signal definition cuts, possibly followed by dead material corrections, see earlier discussion on normalization for this case. The same is true for non-classified clusters, where the previously discussed simple calibration functions are applied to the signal.

Hadronic clusters are usually subjected to more sophisticated calibration procedures. Even different choices of algorithms depending on cluster variables can be realized, for example depending on the cluster signal [4]. Iterative procedures to optimize energy resolution may also be needed. In this case sensitive cluster variables are recalculated from cell energies updated in each iteration, until a stability requirement (typically on the total cluster energy) is fulfilled. Note that in case of software signal weighting, parameters for the weighting functions can be calculated from cluster variables. The weights are then calculated for each cell using these parameters and the cell signal or signal density.

Dead material corrections for hadronic objects may also be applied, following the discussion of this subject in case of electromagnetic clusters. Here it is also necessary to decouple the hadronic calibration functions from these corrections as much as possible, for universal application to testbeam and physics signals, see discussion on electromagnetic calibration above.

Generally all individual cell signals in a given cluster should be converted to FES, just like the total cluster signal. Even though this final cell signal is due to the typically enormous fluctuations only a poor estimate for the true deposited energy *in this cell* – even for electrons, see section 3 – it is still important if missing transverse energy is calculated using cells.

5.3 Full final state reconstruction

Locally calibrated cell clusters are the input for the last step in the reconstruction chain, the determination of the energy of physics objects. The sequence of tasks performed for this step in given in figure 8: first, electromagnetic clusters are subjected to electron and photon identification. Those which do not pass as isolated particles then define, together with hadronic and non-classified clusters, the input for jet finding. Each identified photon, electron and jet is subjected to a very last round of energy corrections for particle identification and jet finder cuts (if any!), and other possible acceptance corrections depending on the particle/jet type.

Particularly important corrections at this last step can be related to energy losses in upstream dead material, which usually depend on the particle type. High precision photon and electron energy reconstruction, for example, needs to take into account the differences between photon and electron showers at this point. Here is one possible violation of our "calorimeter only" attempt on energy reconstruction: reconstructed tracks in the inner de-
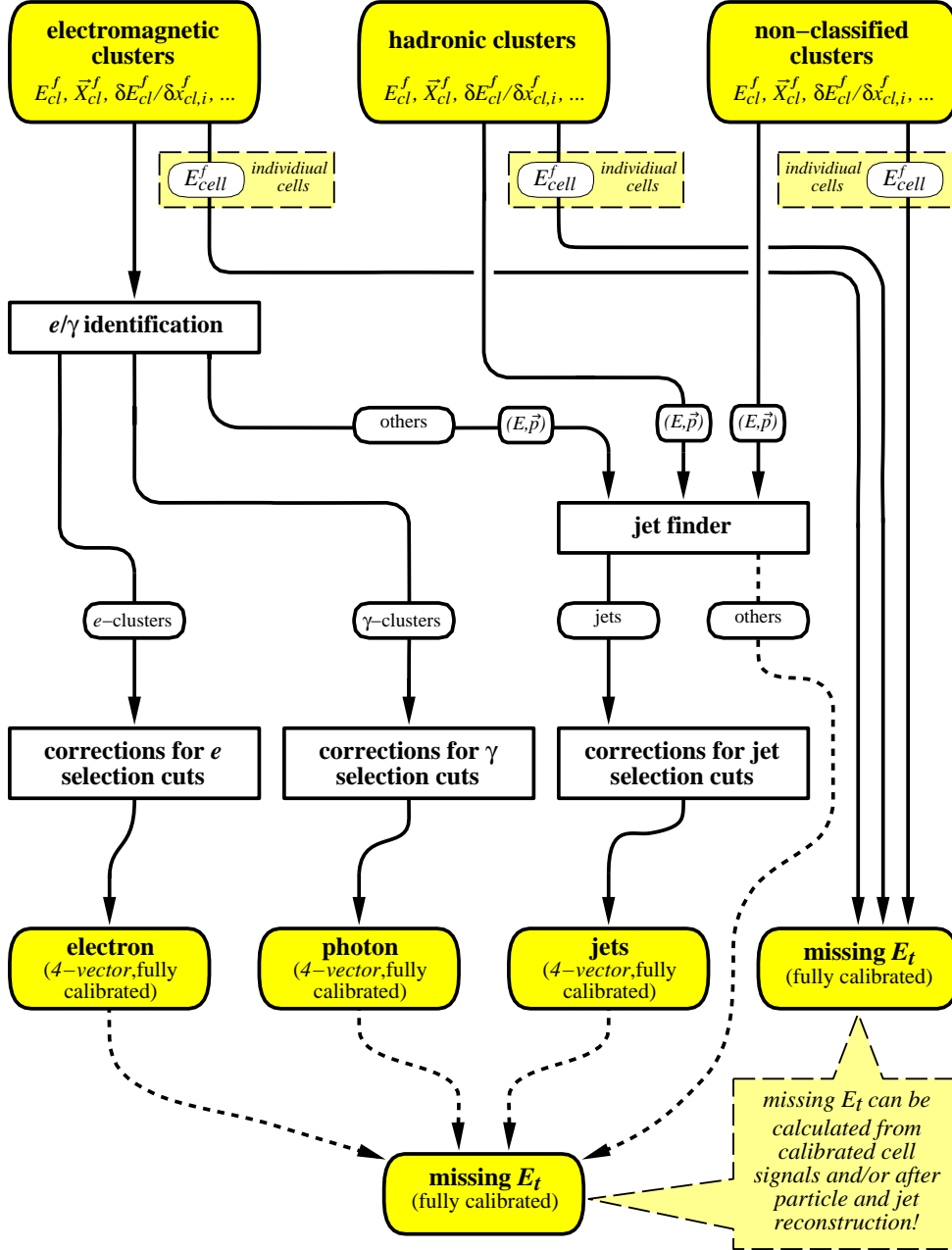
Figure 8: *Full calibration including particle identification and jet finding. The broken lines indicate an alternative reconstruction of the missing transverse momentum from fully reconstructed particles and jets and non-classified clusters, instead of the (resolution preferred ?) single cells approach. Superscript f indicates FES signals, or quantities derived from these signals (see also figure 7); $(E, \vec{p})$ denotes 4-vector input.*

29

tector pointing to an electromagnetic cluster in the calorimeter help to separate photons from electrons. Using the calorimeter alone for this task requires very fine read out of the first longitudinal segment and very little inactive material in front to be sensitive to the small differences in the electron and photon shower development.

The missing transverse energy $E_{t,miss}$ can be calculated from the fully reconstructed final state particles, jets (including all corrections) and whatever clusters are not assigned to any physics object. The cell signal should again be corrected, too, like after local energy reconstruction. This allows to use cell geometry and energies for optimal resolution in the $E_{t,miss}$ calculation.

Another approach is to use the FES cell energy signal for this calculation, with the disadvantage that not all acceptance corrections are applied. On the other hand, some of these corrections may be very model depending, and the corresponding biases of particle and jet identification in the $E_{t,miss}$ calculation should be avoided. Also, overall acceptance corrections obtained from simulations or in situ from events with perfect transverse momentum balance, are still possible in this calculation[10].

One eminent advantage of the local energy reconstruction model is that generalized clusters provide rather well calibrated and massive (if more than one cell) 4-vectors. They can therefore be treated as pseudoparticles in the following physics object definition, with quite realistic estimates of the true energy flow in their direction – a very interesting feature for comparisons of different jet finding schemes.

Reconstruction and physics analysis: some software aspects

Particle identification and jet finding, maybe not so much $E_{t,miss}$ calculation, are to a point already part of the physics analysis, even though one default algorithm for each of these tasks must be implemented into the reconstruction, mainly for event filter purposes. General purpose calibration and reconstruction can therefore only provide a stable signal with small and well understood biases, on a well known energy scale, as input for the physics object definition.

Consequently, all relevant data allowing the re-reconstruction of the physics objects must be provided as the minimum input to analysis. For systematic studies more information is needed, though: EMS and FES cell signals must both be available, in addition to the fully corrected energies after object formation. It is often useful to keep the visible and deposited energy in calorimeter cells in addition, which is of course only available for simulated data.

From point of software, cell clusters can be seen as logical concepts realized by lists of pointers to calorimeter cells. Generalized clusters are then just a list of pointers to these local clusters, see figure 9. These lists, together with access to the algorithms used to collect the "smaller" objects, makes cluster formation and classification fully transparent at analysis level.

The physics objects should be built in the same way as clusters, allowing to establish a hierarchical relation between each individual calorimeter cell surviving the noise suppression,

---

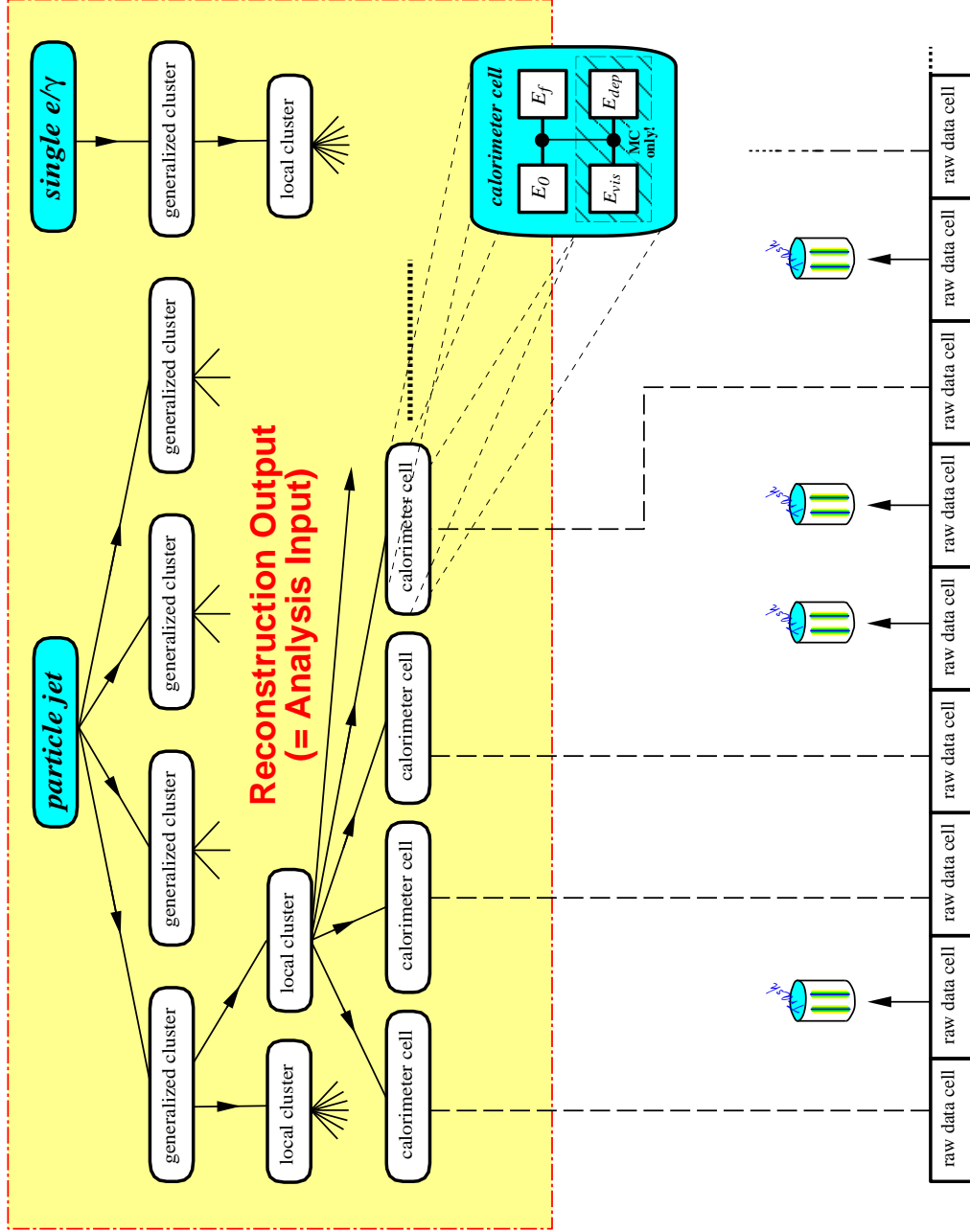[10]these corrections are not included in the scheme shown in figure 8!

Figure 9: *Basic data structure of the reconstruction output. E_0 is the EMS and E_f the FES cell signal, both available for data and Monte Carlo. The visible (E_{vis}) and deposited (E_{dep}) cell energies are only available for the simulation. The full raw data cell list is usually not available in the physics experiment, but should be available for testbeam data, to study the effect of the signal definition, see section 4.*

31

the final object *and all important steps in between!* It should be possible to follow these relations from the cells to the physics objects and vice versa, again to allow for transparency of the chain. Note that the raw data cell list can already be reduced by online noise suppression, as discussed in section 4. This step can therefore not be redone in analysis, especially not for physics data.

Conclusions and Outlook

The concept for a complete energy reconstruction model for isolated particles, jets and missing transverse energy, using calorimeter signals only and without bias towards certain LHC physics, has been presented. The fundamental ingredient of this model is the intrinsic electromagnetic energy scale featuring the ideal (unbiased) electron calibration constant, as extracted from testbeam response. This scale is important for comparisons of experimental data and the corresponding simulations, and is universal for testbeam and physics.

The other major feature of this model is local reconstruction, meaning calibration functions normalized to energy deposited locally in the calorimeter. This step must be preceded by a signal definition, which consists of noise suppression algorithms and localization of energy deposit by cell clustering. The following determination of the electromagnetic or hadronic nature of this energy deposit using cluster related quantities is the first step towards linearity and optimal energy resolution in non-compensating calorimeters.

Another advantage of localizing signals in calorimeters is the possible extraction of hadronic calibration functions for jets from single hadron testbeam data. Also, reconstructed and calibrated clusters can be interpreted as pseudoparticle 4-vectors for jet finders.

A potential weakness is the dependence of the calibration functions on (especially hadronic) shower simulation models, if those must be used for local normalization. These strategies certainly need further investigations for real calorimeter designs. The choice of calibration functions itself is also strongly related to the calorimeter geometry and read out granularity, and needs to be studied for each design in ATLAS. The suggested strategy for the determination of the electromagnetic calibration seems to be generally applicable, even though it depends on the more reliable electromagnetic shower Monte Carlo for acceptance unfolding.

We do not suggest to implement the described model without further studies. Too many (design depending) details are missing, especially concerning the sensitivity of the signal in a given calorimeter to shower nature. It is also imaginable that the rather large pile-up noise at LHC puts severe limitations to the somewhat crucial clustering and signal classification. These fluctuations, and the need to illuminate all calorimeter regions in ATLAS with sufficient statistics, requires a huge amount of response simulations.

We hope, though, that the presented model can be used as a starting point for software design, and that the other message of this note, the strong relation between the signal definition and the final calibration, will be studied more carefully for all ATLAS calorimeters in the LHC environment in the near future

Acknowledgements

References

[1] H1 Calorimeter Group (B. Andrieu et al.), Nucl. Instr. and Meth. A350 (1994) 57-72

[2] H1 Calorimeter Group (B. Andrieu et al.), Nucl. Instr. and Meth. A336 (1993) 499-509

[3] H1 Calorimeter Group (B. Andrieu et al.), Nucl. Instr. and Meth. A336 (1993) 460-498

[4] V. Shekelyan, *Simulation and Reconstruction in the H1 Liquid Argon Calorimetry*, in Proc. of Int. Conf. on Monte Carlo Simulation in High Energy and Nuclear Physics MC93, Tallahassee, Feb. 22-26, 1993, ed. P. Dragovitsch, S.L. Linn and M. Burbank (1994)

[5] M.I. Ferguson et al., *Electron testbeam results for the ATLAS liquid argon forward calorimeter prototype*, Nucl. Instr. and Meth. A 383 (1996) 399-408.

[6] J. Gayler, H. Küster and P. Loch, *Determination of the Electromagnetic Energy Scale for the IF Calorimeter*, H1 internal note H1-04/91-171 (1991)

[7] P. Loch, thesis (in German), DESY-FH1K-92-02 (1992)

[8] R. Brun and F. Carminati, *GEANT Detector Description and Simulation Tool*, CERN Programming Library Long Writeup W5013 (1993)

[9] L. Görlich and H.P. Wellisch, *Documentation of the LAr clustering*, H1 internal note H1-12/91-204 (also H1-MPI-162) (1991)

[10] R. Davis and P. Savard, *A Study of Pileup Noise in the Barrel and Endcap Calorimetry*, ATLAS internal note CAL-NO-84 (1996)

[11] ATLAS coll., *ATLAS Liquid Argon Calorimeter Technical Design Report*, CERN/LHCC/96-41 (1996);
ATLAS coll., *ATLAS Tile Calorimeter Technical Design Report*, CERN/LHCC/96-42 (1996)

[12] H1 Calorimeter Group (B. Andrieu et al.), Nucl. Instr. and Meth. A344 (1994) 492-506

[13] H.G. Oberlack, *Compensation by software: single particles and jets in the H1 calorimeter*, in Proc. of. Int. Conf. on High Energy Physics, Singapore 1990 (1991);
J.P. Kubenka et al., *Hadronic Calibration of the H1 LAr Calorimeter using Software Weighting Techniques*, MPI pre-print MPI-PhE/94-03 (1994)

[14] M.P. Casado et al., *H1-inspired analysis of the 1994 combined test of the Liquid Argon and Tilecal calorimeter prototypes*, ATLAS internal note TILE-NO-75 (1996)

[15] A. Kiryunin, *Study of the hadronic energy resolution of a LAr hadronic calorimeter by applying the H1 weighting technique*, ATLAS internal note CAL-NO-16 (1993)