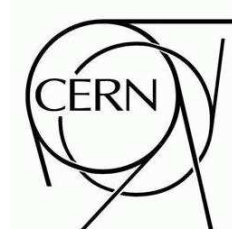




ATLAS NOTE

ATL-COM-CAL-2008-004

March 2, 2009



A Layer Correlation Technique for ATLAS Calorimetry Calibration at the 2004 ATLAS Combined Beam Test

Tancredi Carli

Peter Speckmayer

CERN – European Organization for Nuclear Research

CH-1211 Genève 23, Switzerland

Karl-Johan Grahn

Physics Department, Royal Institute of Technology (KTH)

AlbaNova University Center

SE-106 91 Stockholm, Sweden

Francesco Spanò

Columbia University

Nevis Laboratories

136 South Broadway, P. O. Box 137

Irvington, NY 10533, USA

ATL-CAL-PUB-2009-001
02 March 2009



Abstract

A method for calibrating the response of a segmented calorimeter to hadrons is developed. The ansatz is that information on longitudinal shower fluctuations gained from a principal component analysis of the layer energy depositions can improve energy resolution by correcting for hadronic invisible energy and dead material losses: projections along the eigenvectors of the correlation matrix are used as input for the calibration. The technique is used to reconstruct the energy of pions impinging on the ATLAS calorimeters during the 2004 Barrel Combined Beam Test at the CERN H8 area. Simulated Monte Carlo events are used to derive corrections for invisible energy lost in nuclear reactions and in dead material in front and in between the calorimeters. For pion beams with energies between 20 and 180 GeV, the particle energy is reconstructed within 3% and the resolution is improved by about 20%.

Introduction

In order to measure the energy of a hadron impinging on a calorimeter it is necessary to overcome the fact that part of the energy is deposited in a non-measurable (invisible) form, i.e. nuclear break-up and excitation, energy deposits arriving out of the sensitive time window (such as delayed photons), soft neutrons, etc. All of these effects are absent in the case of particles interacting only electromagnetically [1, 2]. All secondary particle showers generated by hadrons will have a component of such particles. As the average size of this component increases with the energy of the impinging hadron [3], the ratio of visible to invisible energy increases, resulting in a non-linear response of the calorimeter. In addition, the large event-by-event fluctuations of the size of the electromagnetic component ¹⁾ degrade the measured resolution.

The calorimeter material and size are not always aimed at (or capable of) equalizing the hadronic and electromagnetic energy response (compensation). In these cases offline calibration techniques have been developed that use the properties of the visible deposited energy to derive energy corrections aimed at restoring linearity in the response and improving the resolution. Average corrections can restore the linearity of the response, but they cannot compensate the calorimeter response on an event-by-event basis, so they can hardly improve the resolution. Techniques that are spatially sensitive to the difference between electromagnetic and hadronic energy deposits and their fluctuations are therefore more suited to produce event-by-event corrections that will also affect the resolution.

The correlations between the energy deposits in the longitudinal development of the shower have been shown [4] to have information on the electromagnetic and hadronic nature of the shower. An improvement in resolution is expected to result from capturing more of the properties of the different fluctuations. In this study a technique for energy compensation is developed that uses the properties of such correlations to derive a Monte Carlo-based event-by-event compensation technique. The application is quite specific to ATLAS, but the framework is rather general and it can be tested on any segmented calorimeter. Layer weighting has already been explored in ATLAS in a different fashion for jet energy calibration [5].

In section 1 the experimental setup of the ATLAS 2004 Barrel Combined Beam Test is presented. The data samples and the associated simulated event samples are described. Section 2 describes the basic selection and reconstruction of the events. The Layer Correlation technique is featured in section 3 in general terms to provide the basis for its application in the case of a segmented calorimeter. The various components of the calibration technique are also presented. Then its implementation in the present case and its validation on simulated samples are illustrated in sections 4 and 5, respectively. The results in terms of linearity and resolution are finally shown and discussed in section 6.

1 Experimental Setup, Data and Monte Carlo Samples

The 2004 ATLAS Combined Beam Test setup (CBT04) is fully described in [6]. For this analysis the fully combined data taking period has been used, where in addition to the Liquid Argon (LAr) and Tile (TileCal) calorimeters, some modules of the ATLAS inner detector were operational, i.e. the Pixel detector, the silicon strip SemiConductor Tracker (SCT) and the straw tube Transition Radiation Tracker (TRT). In addition, special beam-line detectors were installed to monitor the beam position and reject background events. For a beam impinging at a pseudorapidity of $\eta = 0.45$ the expected amount of material in front of the LAr is about $0.44 \lambda_I$ ²⁾. The LAr modules are longitudinally segmented in 3 layers that extend in total for $1.3 \lambda_I$. The dead material between TileCal and LAr spans about $0.6 \lambda_I$. Finally TileCal's three longitudinal segments stretch in total for about $8.2 \lambda_I$.

¹⁾Due to large phase space for hadronic interactions in matter.

²⁾The interaction length λ_I is defined in [2, 7]. The values refer to interactions caused by protons.

The data taken during this fully combined CBT04 period consist of single positive pions impinging on the calorimeters at azimuth angle $\phi = 0$ and pseudorapidity $\eta = 0.45$. The runs being used are shown in table 1. All data were collected without magnetic field around the pixel and silicon detectors. Data samples were reconstructed using Athena 12.0.7.

The positive pion beams are known to have a sizable proton contamination, f_{prot} , defined as the fraction of the events in a sample that result from protons impinging on the calorimeters. The contamination is also reported in table 1 and it is estimated ³⁾ by recent Transition Radiation Tracker measurements [8] exploiting the higher energy losses of protons with respect to pions. This estimate is in good agreement with measurements performed by a Cherenkov counter in the 2002 test beam [9], except at 180 GeV.

Table 1: Data samples taken in the CBT04 used in the present analysis. 2102225 is nominally an e^+ run where π^+ are selected.

RunNumber	$E_{\text{beam}}^{\text{nom}}$ (GeV)	E_{meas} (GeV)	N. ev. bef. cuts	N. ev. after cuts	Content	f_{prot}
2102396	20	20.19	49871	7812	π^+	0%
2102347	50	50.30	81103	15091	π^+	45%
2102355	100	100.54	55589	7250	π^+	61%
2102225	180	179.68	85967	4318	$\pi^+ (e^+)$	76%

Monte Carlo samples were produced simulating both pions and protons impinging on the CBT04 setup at the same coordinates as the data. The interaction of the particles in the detector was simulated with Geant4.7 [10] using both the QGSP [11] and the QGSP_BERT [12] physics lists. The QGSP physics list uses a phenomenological model describing the hadron–nucleus interaction by the formation and fragmentation of excited strings together with the de-excitation of an excited nucleus. The QGSP_BERT model adds the Bertini description of the intra-nuclear hadronic cascade of the secondary hadrons deriving from string decay: such a description is not present in QGSP where the secondaries are allowed to be produced only outside the nucleus.

The full reconstruction of the events used the CBT04 geometry and the full simulation of the detector read-out in the ATLAS Athena Software version 12.0.5.

Two sets of statistically independent samples were produced by dividing the available sample into two approximately equal parts: one set (“correction” samples in the following) was used to derive weights and dead material corrections (see sections 3.1 and 3.2), while the other set was used as simulated data to derive the performance expectation (“signal” samples in the following). The “correction” and “signal” samples are shown in table 2.

2 Event Selection and Energy Reconstruction

For each event, cell clusters are derived by adding up the energy in neighboring cells with a dynamical topological algorithm in three dimensions into “topological clusters”. The algorithm is described in [13]. The energy thresholds used for the seed, neighbor and boundary determination are set to respectively four, two and zero times the expected noise in the given cell. The reconstructed energy in layer L is obtained by building all the 3D topoclusters in the event and summing only the energy of the cells in layer L that belong to a topocluster. The total reconstructed energy is then derived by summing over the

³⁾With the instrumentation at hand, the proton contamination can be measured only on average, but not on an event-by-event basis.

Table 2: Fully simulated Monte Carlo samples used for the present analysis. The “correction” sample is used to derive the corrections, while the ”signal” sample is used to compare to the data.

E_{beam} (GeV)	“Correction sample”				“Signal sample”			
	π^+		p		π^+		p	
	Bef. cuts	Aft. cuts	Bef. cuts	Aft. cuts	Bef. cuts	Aft. cuts	Bef. cuts	Aft. cuts
15	50000	33098	50000	32315	50000	33075	50000	32711
17	45000	31352	50000	34991	45000	31452	50000	34975
20	50000	36387	50000	33119	50000	36327	50000	36862
22	45000	33440	50000	37532	45000	33367	50000	37339
25	45000	34053	50000	38160	45000	33899	50000	38293
35	50000	39420	50000	39027	50000	39339	50000	39010
40	50000	39640	45000	35389	50000	39535	50000	39277
45	50000	39947	50000	39426	50000	39871	50000	39481
50	50000	40171	45000	35510	50000	40248	50000	39574
55	45000	36431	45000	35547	50000	40514	50000	39495
60	45000	36509	45000	35668	50000	40647	50000	39537
65	50000	40670	50000	39606	50000	40570	50000	39685
70	35000	28494	25000	19853	35000	28549	30000	23855
80	25000	20448	35000	27919	25000	20337	35000	27791
90	35000	28771	31536	25083	35000	28680	35000	27805
100	25000	20510	20000	15859	30000	24717	25000	19839
110	15000	12362	15000	11948	15000	12313	20000	15736
120	35000	28774	30000	23957	35375	29036	35000	27579
130	25000	20626	44785	35461	30000	24798	45000	35526
150	50000	38556	50000	35979	50000	38538	50000	36248
170	50000	38569	50000	35804	50000	38606	50000	36026
180	50000	38807	40000	28804	50000	38625	45000	32126
190	50000	38701	45000	32308	50000	38740	50000	35847
210	40000	30965	45000	32236	40000	30915	45000	31994
230	30569	23670	21015	14913	40000	31027	30000	21444
Total	1040569	810371	1032336	776414	1075357	833725	1095000	828055

N_{lay} longitudinal layers in the calorimeter:

$$E_L^{\text{rec}} = \sum_{\text{cell.t3D}} E_{\text{cell.t3D},L}^{\text{rec}} \quad (1)$$

$$E_{\text{tot}}^{\text{rec}} = \sum_L E_L^{\text{rec}} \quad (2)$$

Events are selected, if they meet the following requirements.

A signal in the trigger scintillator and a measurement in adjacent beam chambers that is compatible with one particle passing close to the nominal beam line are required. In addition, at least one global track with a least 6 hits in the Pixel or the SCT and at least 20 hits in the TRT is asked for. The track in the TRT must be compatible with a pion track, i.e. no more than 2 higher-level hits must be present. Events with a second track in the TRT are rejected: this ensures that the pion does not interact strongly before the TRT.

In addition, there must be at least one cluster with at least 5 GeV in the calorimeter. This cut rejects muons contained in the beam. To reject some residual electron background, events with more than 99% of their energy in the LAr calorimeter are rejected.

The same selection is applied on simulated Monte Carlo events as on data, with the exception of cuts related to beam chambers and scintillators.

3 The Layer Correlation Method

The Layer Correlation calibration method (LC in the following) is based on a principal component analysis (PCA) [14] on simulated events of the calorimeter layer energy deposits. The ansatz is that different fluctuation properties characterize events with different invisible energy content: variables that can describe the various possible fluctuations should discriminate the different corrections to be applied to recover invisible losses due to hadronic interactions. In addition, dead material losses will vary depending on the shower development: fluctuation-sensitive variables should then also help to recover dead material losses while capturing their correlation with invisible energy fluctuations.

The aim of the LC technique is to derive an event-by-event correction to the energy deposited in each longitudinal calorimeter layer on the basis of Monte Carlo simulated events. A unified treatment for compensation and dead material correction is obtained by deriving both corrections from the same set of variables.

In every event the energy deposited in each layer is corrected by a weight to account for compensation. An additive correction to the total compensated energy accounts for dead material effects. In formulas:

$$E_L^{\text{weighted}} = w_L E_L^{\text{rec}} \quad (3)$$

$$E_{\text{tot}}^{\text{weighted}} = \sum_L E_L^{\text{weighted}} \quad (4)$$

$$E_{\text{tot}}^{\text{corr}} = E_{\text{tot}}^{\text{weighted}} + E_{\text{tot}}^{\text{DM}} \quad (5)$$

where w_L is the weight correcting for invisible energy in layer L and $E_{\text{tot}}^{\text{DM}}$ is the dead material correction.

The N_{lay} weights and the most important part of the dead material corrections are functions of linear combinations of the observed layer energy deposits. The idea is to use those linear combinations whose variance provides the largest contribution to the fluctuations of the total energy, i.e. are more sensitive to differences in invisible energy content. Such combinations are obtained from a principal component analysis of the of the N_{lay} layer energy deposits on simulated events. So, intuitively, the corrections are functions of the “directions of largest independent fluctuations”.

In order to derive the corrections, a fully simulated sample of single pions and protons impinging on the calorimeter is used. In such events the true deposited energy by the pions in each layer of the calorimeter is available and it is divided into four components: electromagnetic visible, hadronic visible, invisible and escaped. Visible energy results from ionization of the calorimeter material. Invisible energy is energy not directly measurable in the detector, such as break-up energy in nuclear interactions. The escaped energy represents the small contribution from neutrinos, high energy muons and, possibly, neutrons and low-energy photons escaping the calorimeters. All the terms are defined in [15].

The events in a fully simulated Monte Carlo sample are usually generated at a fixed beam energy in order to test the calorimeter response. The corrections derived from a fixed beam energy sample are, in principle, dependent on that information, i.e. they depend on the same quantity (pion energy) for the reconstruction of which they should be used. Sample mixing and iteration techniques are used to overcome such dependence and properly define a procedure for converging towards the appropriate corrections using only the visible energy in the calorimeter. The description of the weighting technique and the dead material corrections in the following sections (3.1 and 3.2) assumes that simulated events have a fixed beam energy. Section 3.3 deals with overcoming the beam energy dependence when deriving the corrections.

3.1 Compensation Weights

Each event is associated with a set of N_{lay} layer energy deposits ($E_1^{\text{rec}}, \dots, E_{N_{\text{lay}}}^{\text{rec}}$), one per calorimeter layer, representing a point in an N -dimensional vector space referred to below as the space of layer energy deposits. They are reconstructed energies as defined in equation 1.

For each event i , there is also an ideal set of N_{lay} coefficients that would re-weight each reconstructed energy deposit in layer L to the true deposited energy (known in Monte Carlo):

$$w_{L,i}^{\text{ideal}} = E_{L,i}^{\text{true}} / E_{L,i}^{\text{rec}}. \quad (6)$$

$E_{L,i}^{\text{rec}}$ ($E_{L,i}^{\text{true}}$) is the reconstructed (true) energy deposited in the L^{th} layer in the i^{th} event.

The task is to find a set of weights w_L that as closely as possible approximate the ideal weights. In general, for each layer L , the weight is an N_{lay} -dimensional function of the layer energy deposits.

It is desirable to reduce the number of dimensions that the weights depend on, while still capturing a large amount of event fluctuation information and maintaining a good separation between events with different content of invisible energy. This is achieved through the principal component analysis.

The N_{lay} -dimensional covariance matrix of the layer energy deposits is calculated as

$$\text{Cov}(M, L) = \langle E_M^{\text{rec}} E_L^{\text{rec}} \rangle - \langle E_M^{\text{rec}} \rangle \langle E_L^{\text{rec}} \rangle, \quad (7)$$

where

$$\langle E_M^{\text{rec}} E_L^{\text{rec}} \rangle = \frac{\sum_i E_{M,i}^{\text{rec}} E_{L,i}^{\text{rec}}}{N_{\text{ev}}} \quad \text{and} \quad \langle E_M^{\text{rec}} \rangle = \frac{\sum_i E_{M,i}^{\text{rec}}}{N_{\text{ev}}}. \quad (8)$$

The sums are performed over all the N_{ev} events in the sample.

A new orthogonal basis in the space of layer energy deposits is derived. It consists of the eigenvectors of the covariance matrix. The coordinates of the point in the N_{lay} -dimensional vector space corresponding to an event i can then be expressed in the new eigenvector basis as:

$$E_M^{\text{rec}} = \sum_{\text{eig}} \alpha_{M,\text{eig}}^{\text{rec}} E_{\text{eig}}^{\text{rec}}, \quad (9)$$

where $\alpha_{M,\text{eig}}$ is the matrix that performs the rotation to the new basis.

Projections of events along the covariance matrix eigenvectors represent independent fluctuations. The variances of those fluctuations are given by the corresponding eigenvalues. The eigenvectors are put in descending order according to their eigenvalues, meaning that the first eigenvectors determine the directions along which most of the event fluctuations take place.

Exploiting this property, weights are derived as a function of an N -dimensional subspace of the N_{lay} -dimensional space of layer energy deposits, spanned by the first N eigenvectors. In the absence of an analytic formulation, the layer weights w_L are estimated by Monte Carlo sampling: multi-dimensional cells are built, partitioning the N -dimensional vector space along the directions of the base eigenvectors. In general, these cells are multi-dimensional hyper-cubes. They are referred to as bins below.

For each bin k one defines the weight as the average of the ideal weights of equation 6:

$$w_{k,L} = \langle E_{L,i}^{\text{true}} / E_{L,i}^{\text{rec}} \rangle_k = \frac{1}{N_{\text{ev},k}} \sum_i E_{L,i}^{\text{true}} / E_{L,i}^{\text{rec}} \quad (10)$$

where the summation is performed for the $N_{\text{ev},k}$ events in the bin.

If each event has a weight p_i^4 , the average is modified accordingly:

$$w_{k,L} = \langle E_{L,i}^{\text{true}} / E_{L,i}^{\text{rec}} \rangle_k = \frac{\sum_i p_i E_{L,i}^{\text{true}} / E_{L,i}^{\text{rec}}}{\sum_i p_i}. \quad (11)$$

3.2 Dead Material Corrections

The energy loss in non-instrumented (dead) material between the LAr and the Tile calorimeters is expected to be the largest dead material effect [16] in the reconstruction of the pion energy. The LC technique is used also for correcting for this loss (see section 3.2.1).

Additional dead material losses (e.g. due to material in front of the calorimeter) give rise to smaller corrections. For these effects, a parametrization of the energy loss as a function of the best estimate of the pion energy is used (see section 3.2.2).

3.2.1 Dead Material between LAr and Tile

A similar technique to the calorimeter compensation is developed for the LAr–Tile energy loss. Each event is associated with a point in the layer energy deposit vector space as explained in the last section. It also has a true total energy lost in the dead material between LAr and Tile: $E_{\text{LArTile}}^{\text{DM,true}}(i)$.

The dead material correction, $E_{\text{LArTile}}^{\text{DM}}$, for each event i , is derived as an T -dimensional function of the layer energy deposits. In general, the subspace chosen for deriving the dead material correction and its dimension T can be different from the one chosen for compensation, both in content (spanned by different eigenvectors) and in dimension (T can be different from N). The value of $E_{\text{LArTile}}^{\text{DM}}$ is estimated by Monte Carlo sampling with the same bin definition as in section 3.1. For any T -dimensional bin k one defines:

$$E_{\text{LArTile},k}^{\text{DM}} = \langle E_{\text{LArTile},i}^{\text{DM,true}} \rangle_k \quad (12)$$

where the average is performed for the events in that bin.

3.2.2 Other Dead Material Corrections

While the energy losses between the calorimeters dominate, there are still other regions where dead material losses can occur. These are losses located in the material upstream of the LAr calorimeter, between the LAr presampler and the first LAr layer and energy leakage beyond the Tile Calorimeter.

⁴For instance, to equalize the number of events for all data sets.

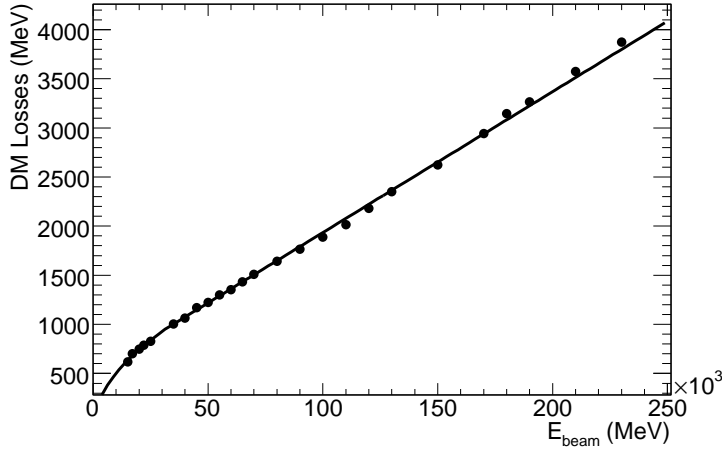


Figure 1: Mean Dead material losses other than those between the LAr and Tile calorimeters as a function of the beam energy. Statistical errors are smaller than the markers. Filled circles indicate the mean loss obtained from Monte Carlo simulation. The line indicates a parametrization to interpolate between the beam energies.

To compensate for these losses the mean energy loss was determined as a function of beam energy and the resulting data points were fitted using a suitable functional form:

$$E_{\text{other}}^{\text{DM}}(E_{\text{beam}}) = \begin{cases} C_1 + C_2 \sqrt{E_{\text{beam}}} & \text{if } E_{\text{beam}} < E_0 \\ C_3 + C_4 (E_{\text{beam}} - E_0) & \text{otherwise,} \end{cases} \quad (13)$$

where $E_0 = 30000$ MeV. The fit can be seen in figure 1. The resulting fitted parameters are

$$C_1 = (-75 \pm 31) \text{ MeV} \quad C_2 = (5.78 \pm 0.22) \sqrt{\text{MeV}} \quad C_3 = (931 \pm 5) \text{ MeV} \quad C_4 = 0.01435 \pm 0.0001 \quad (14)$$

3.2.3 Total Dead Material Correction

The total dead material correction needed in equation 5 is derived from summing the two contributions derived in sections 3.2.1 and 3.2.2

$$E_{\text{tot},k}^{\text{DM}}(E) = E_{\text{LArTile},k}^{\text{DM}} + E_{\text{other}}^{\text{DM}}(E) \quad (15)$$

where

- k is the bin in the appropriate T -dimensional space of layer energy deposits used to build the LC estimate for the energy loss in the dead material between LAr and Tile obtained from formula 12 and
- E is the best estimate for the total deposited pion energy used to estimate E_{beam} in formula 13 (also see section 3.3).

3.3 Beam Energy Dependence

If calculated from simulated pion samples generated at different beam energies, the average weight function defined in equation 10, the dead material corrections obtained in equation 12, and the associated eigenvectors are in principle different.

In order to overcome such a dependence, a single set of eigenvectors and a single set of look-up tables (one for each calorimeter layer for weighting and one for dead material corrections) are constructed using all available samples. These sets are used to derive the corrections to any data or Monte Carlo sample under study.

The procedure can be also be used iteratively: the weighted estimate of the energy is used to make a new choice of the correction tables until the returned value is stable. This can be coupled (or not) to the iteration on the dead material corrections (see below) so as to obtain the best energy estimate. The iteration cut-off is a tunable parameter.

For the LAr–Tile dead material correction the three dimensions of the look-up table are all shown to scale with beam energy, i.e. a table determined at a given beam energy can be turned into one at a different beam energy by scaling all the dimensions with the ratio of the two energies (see section 4.5). The look-up table for LAr–Tile dead material corrections can then be obtained as a function of the beam energy normalized components of the eigenvectors i.e. each event coordinates are expressed as

$$E_M^{\text{rec, norm}} = E_M^{\text{rec}} / E = \sum_{\text{eig}} \alpha_{M, \text{eig}}^{\text{rec}} E_{\text{eig}}^{\text{rec}} / E, \quad (16)$$

where the variables have the same meaning as in equation 9 and E is the best estimate of the beam energy of the simulated pion in that event (see below).

An iteration technique is used for determining all the dead material corrections. At each step the best estimate of the reconstructed energy, $E_{\text{tot}}^{\text{corr}}$, after all corrections, is used to set

- the scaling factor $1/E$ (equation 16) for the LAr–Tile correction and
- the best pion energy estimate in the parametrization for the other dead material corrections.

This allows one to re-calculate the dead material corrections and to get a new $E_{\text{tot}}^{\text{corr}}$ estimate by choosing a new two-dimensional bin in the look-up table and getting a new value from the parametrization. In the initial step $E_{\text{tot}}^{\text{corr}}$ is just the pion energy after compensation weights are applied.

3.4 Using the Corrections

The weights can be applied to any data set by taking the following steps:

- Associate each event to a bin in both the N and T -dimensional spaces defined in sections 3.1 and 3.2 by expressing its energy deposition vector in the new eigenvector basis derived from the simulated events.
- Extract the corrections for the energy of each given layer from the look-up tables defined for compensation weights and dead material estimates. Apply them according to formulas 5 and 15.
- Use the iteration for dead material corrections as discussed in section 3.3.

4 Implementation

Five steps are necessary to derive the compensation weights and dead material corrections from simulated Monte Carlo samples:

- Define the samples for weights generation and performance study (section 4.1).
- Provide the energy deposited in each layer at the electromagnetic scale (section 4.2).

- Calculate the covariance matrix (section 4.3).
- Extract the eigenvectors and eigenvalues (section 4.4).
- Build the compensation weights and dead material corrections look-up tables (section 4.5).

The steps for applying the corrections outlined in section 3.4 can be followed.

4.1 Samples Definition

All the available “correction” Monte Carlo samples are used to build a “mixed” sample, which is input to constructing the corresponding weight and dead material tables. In order to account for the beam content mentioned in section 1, at each energy of the interval, a pion and a proton sample generated at that energy, are considered. If the samples have different numbers of events, a sample-dependent weight is applied to equalize them. Equalization weights are determined initially so that all pion and proton samples in the interval are normalized to the same size before selection cuts. Given the proton contamination f_{prot} at a given energy, pion and proton events for each same-energy pair of samples are assigned a weight of f_{prot} and $1 - f_{\text{prot}}$, respectively. The resulting “mixed” samples constitute the global “mixed” sample.

A different choice of the pion–proton mix will result in a different “mixed” sample: weights can then be derived for each available data set taking into account the measured values reported in table 1.

The pion and proton “signal” samples of table 2 are also paired in energy, normalized and mixed with the same proportions as in the data as above to obtain a “mixed signal” sample on which to apply calibration corrections and to compare with data.

4.2 EM Scale Energy Determination

The estimate of the energy deposited in each calorimeter layer derived in equation 1 is expected to provide the correct reconstruction of the energy electromagnetically deposited in the calorimeters. The weights are defined in formulas 10 and 11 with this assumption.

4.3 Calculation of Covariance Matrix

The layer energy covariance matrix $\text{Cov}(M, L)$ (equations 7 and 8) is calculated using events from the “mixed” sample. The full seven-dimensional covariance matrix is used i.e. $N = N_{\text{lay}} = 7$ (see section 3.1.)

In any given event a symmetric energy cut is applied on each layer energy such that the energy for that layer is re-defined as E_M^{rec} , if $|E_M^{\text{rec}}| > E_M^{\text{thr}}$, zero otherwise.

In this way the contribution to the covariance from a layer only containing noise is set to zero. The energy threshold vector is

$$\vec{E}_{\text{thr}}(\text{GeV}) = (\sigma(1), \sigma(2), \sigma(3), \sigma(4), \sigma(5), \sigma(6), \sigma(7)) \quad (17)$$

where $\vec{\sigma}(\text{GeV}) = (0.032, 0.108, 0.03, 0.150, 0.039, 0.070, 0.042)$. The cuts were optimized to obtain the best expected compensation performance on fully simulated Monte Carlo samples at 50 GeV.

4.4 Extraction and Interpretation of Eigenvectors and Eigenvalues

The covariance matrix $\text{Cov}(M, L)$ is then diagonalized: the normalized eigenvectors and corresponding eigenvalues are calculated for all the samples.

The directions of the eigenvectors are observed to be stable with beam energy and choice of physics list.

A physical interpretation of the eigenvalues and normalized eigenvectors can be obtained by looking at figures 2 and 3 where the eigenvector components are calculated for a fully simulated pion sample impinging at $(\eta = 0.45, \phi = 0)$ in the CBT04 setup.

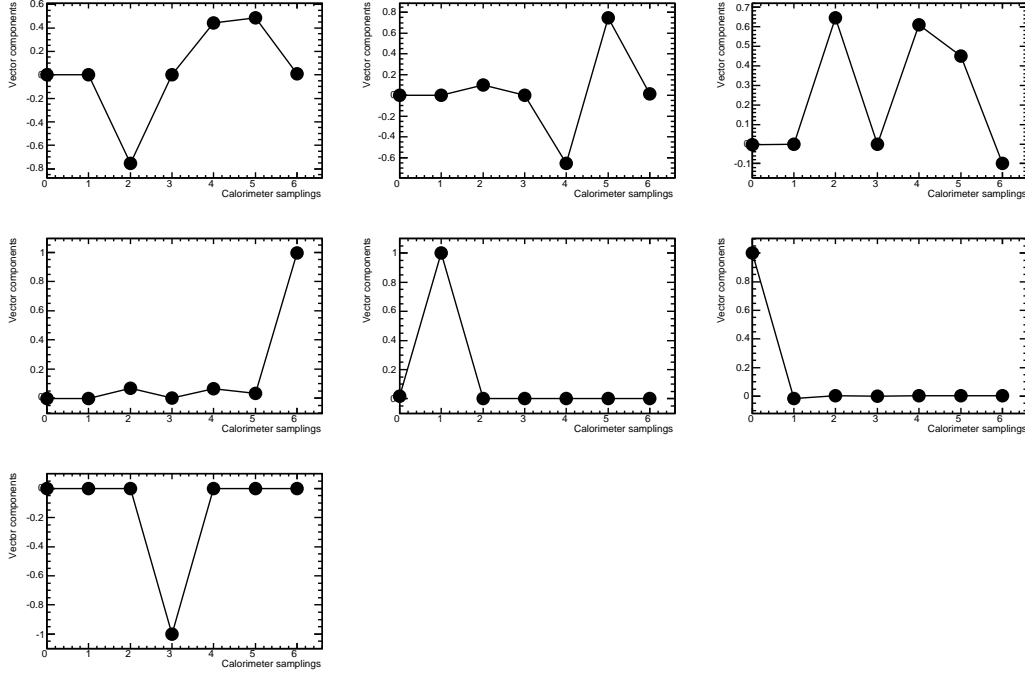


Figure 2: Eigenvector components for the seven layers of the ATLAS combined beam test for simulated pions.

We find that:

- $E_{\text{eig}0}^{\text{rec}} \approx \frac{1}{\sqrt{6}}(-2E_{\text{LAR,middle}} + E_{\text{Tile,A}} + E_{\text{Tile,BC}})$
- $E_{\text{eig}1}^{\text{rec}} \approx \frac{1}{\sqrt{2}}(-E_{\text{Tile,A}} + E_{\text{Tile,BC}})$
- $E_{\text{eig}2}^{\text{rec}} \approx \frac{1}{\sqrt{3}}(E_{\text{LAR,middle}} + E_{\text{Tile,A}} + E_{\text{Tile,BC}})$
- $E_{\text{eig}3}^{\text{rec}} \approx E_{\text{Tile,D}}$
- $E_{\text{eig}4}^{\text{rec}} \approx E_{\text{LAR,strips}}$
- $E_{\text{eig}5}^{\text{rec}} \approx E_{\text{LAR,presampler}}$
- $E_{\text{eig}6}^{\text{rec}} \approx -E_{\text{LAR,back}}$

So in more qualitative, but suggestive, way we can maintain that:

- $E_{\text{eig}0}^{\text{rec}} \approx$ “Difference between Tile and LAr”
- $E_{\text{eig}1}^{\text{rec}} \approx$ “Difference between Tile second (middle) layer and Tile first layer”
- $E_{\text{eig}2}^{\text{rec}} \approx$ “Total Energy”
- $E_{\text{eig}3}^{\text{rec}}$ to $E_{\text{eig}6}^{\text{rec}} \approx$ “Individual layers”

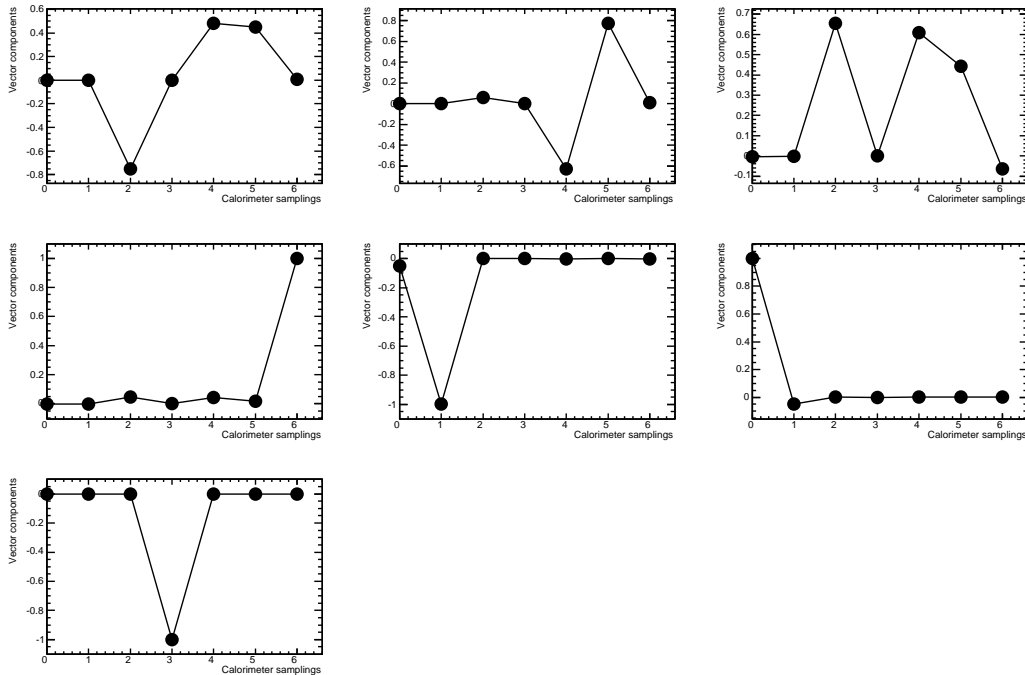


Figure 3: Eigenvector components for the seven layers of the ATLAS combined beam test for a simulated mix of protons and pions with 45% proton contamination (lower plot).

4.5 Preparation of Look-up Tables for Compensation Weights and Dead Material Corrections

Look-up tables for compensation weights and dead material correction are derived using the information from the covariance matrix. A set of eight tables is built: one for each layer of the calorimeter and one to derive the dead material correction between Tile and LAr. For compensation weights, the $w_{k,L}$ functions defined in equation 11 are estimated in bins of the two-dimensional space spanned by the eigenvectors corresponding to the two highest eigenvalues i.e. $N = 2$ (see section 3.1). Thus, each layer is associated with a two-dimensional look-up table. For a given layer the average weights in each two-dimensional bin are calculated using only the energy values that passed the cuts defined in section 4.3. The table has the same number of equally spaced bins along the two dimensions: 128×128 . Weights for the presampler layer of the LAr are not calculated, even if the presampler is kept in the covariance matrix. No weights are applied to the energy deposited in the presampler layer, and energy deposited in the presampler itself is taken as part up the upstream dead material losses.

For LAr–Tile dead material corrections, the average corrections defined in equation 12 are calculated in bins of the two-dimensional space spanned by the eigenvectors corresponding to the first and third eigenvalues i.e. $T = 2$ (see section 3.2.1). Correction tables are derived from a same-size 128×128 bins look-up table.

In addition the proton weights and corrections are corrected by the factor $\frac{E_{\text{beam}}}{E_{\text{beam}} - m_{\text{proton}}}$ for the fact that, for a proton, the sum of the total true deposited energy in the calorimeter is $E_{\text{beam}} - m_{\text{proton}}$.

For both look-up tables, the coordinates of the vector of energy deposits in the chosen eigenvector basis are obtained according to equation 9. Typical compensation weight tables for pion–proton mixed samples are shown in figures 4 and 5: they illustrate the look-up tables for the second (middle) layer of the LAr calorimeter and for the first and second layer of the Tile calorimeter for a pion–proton mixed

sample with 45% contamination.

A clear separation is achieved between regions dominated by electromagnetic events (average weights close to unity) and those where hadronic events are predominant (average weights significantly larger than unity). The functions depend on the layer being described.

The dead material look-up table is shown in figure 6 for a pion–proton mixed sample with 45% contamination. The plot shows the distribution of the rescaled dead material energy as a function of the rescaled event coordinates. The different dead material fractions are clearly separated. In addition, the samples at different energies behave very similarly as a function of the re-scaled variables.

4.6 Correcting the Data

The corrections to the data can be applied as described in section 3.4. In order to derive the total energy before corrections (corresponding to equation 2) at the electromagnetic scale, a series of additional inter-calibration factors are applied to the energy depositions in the data events.

The LAr calorimeter energy depositions are scaled by a factor 1.03 derived from studies of the response to electrons. A re-analysis of TileCal 2002 beam test data results in scaling each TileCal energy deposition by a factor 1.018 to correct the pC to GeV calibration constant. No additional inter-calibration factors are applied. In addition, Monte Carlo for energies up to 100 GeV were simulated with an erroneous overall Birks’ law factor, pushing down the response in the LAr calorimeter by a few per cent. To compensate for this, and additional overall factor of 1.0176 is applied.

5 Method Validation on Monte Carlo

The weighting technique is validated on Monte Carlo samples in separate steps:

- Reconstruct the true deposited energy in the calorimeters (compensation validation).
- Reconstruct the full energy of the incoming particles, including dead material corrections, and quantify the performance in terms of linearity and resolution.

The performance is evaluated in terms of bias and resolution. The weights and dead material corrections are derived from the “correction samples” and applied on the statistically independent “signal samples”. The results in this section are derived for pions only.

5.1 Compensation

The event-by-event difference $E_{\text{tot}}^{\text{weighted}} - E_{\text{tot}}^{\text{true}}(\text{calo})$ is considered, where $E_{\text{tot}}^{\text{true}}(\text{calo})$ is the true total energy deposited in the calorimeter. The bias is defined as the average value $\langle E_{\text{tot}}^{\text{weighted}} - E_{\text{tot}}^{\text{true}}(\text{calo}) \rangle$ and the resolution is obtained by calculating the standard deviation $\sigma(E_{\text{tot}}^{\text{weighted}} - E_{\text{tot}}^{\text{true}}(\text{calo}))$.

The performance of the LC technique is compared to a simple calibration scheme (called f_{comp} in the following) which uses beam energy information: each event in the sample is weighted with the same factor $f_{\text{comp}} = \langle E_{\text{tot}}^{\text{true}} \rangle / \langle E_{\text{tot}}^{\text{reco}} \rangle$ where $\langle E_{\text{tot}}^{\text{true}} \rangle$ ($\langle E_{\text{tot}}^{\text{reco}} \rangle$) is the average true total (reconstructed) energy deposited in the given sample in the whole calorimeter, but not in the dead material. The f_{comp} calibration scheme provides a reference/scale to which the improvement in resolution of the LC weighting can be compared.

The results of the validation procedure are shown in figure 7. The bias is zero by construction for the f_{comp} weighting. The LC weighting mostly gives a slight positive bias of about 0.6%. At the edges of the energy range studied, the bias instead turns slightly negative.

The resolution improvement increases with beam energy. It is 10% at 50 GeV and 20% at 230 GeV.

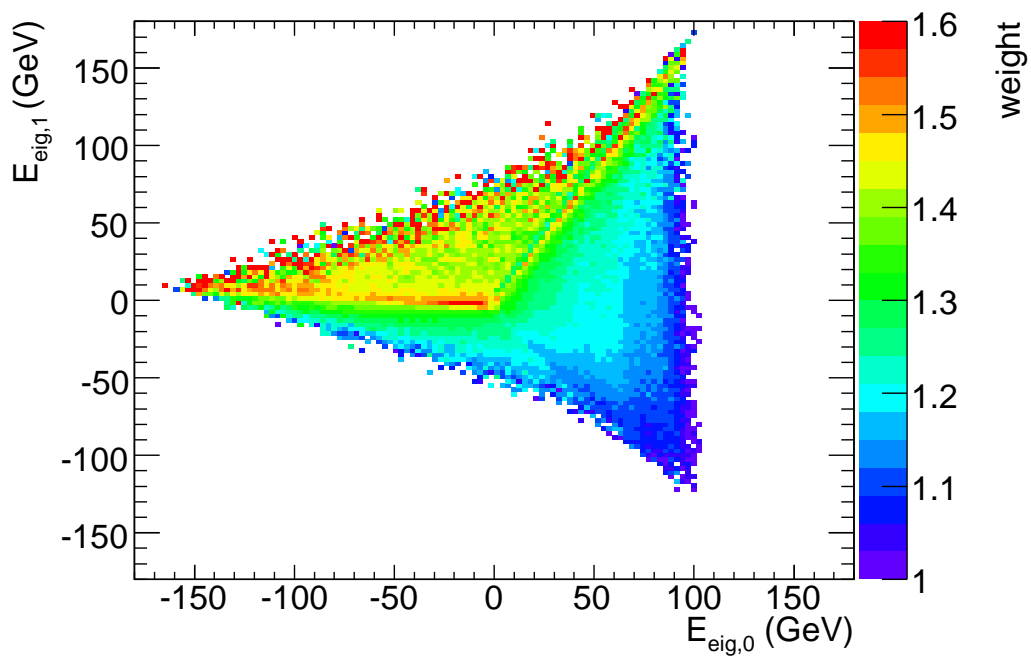
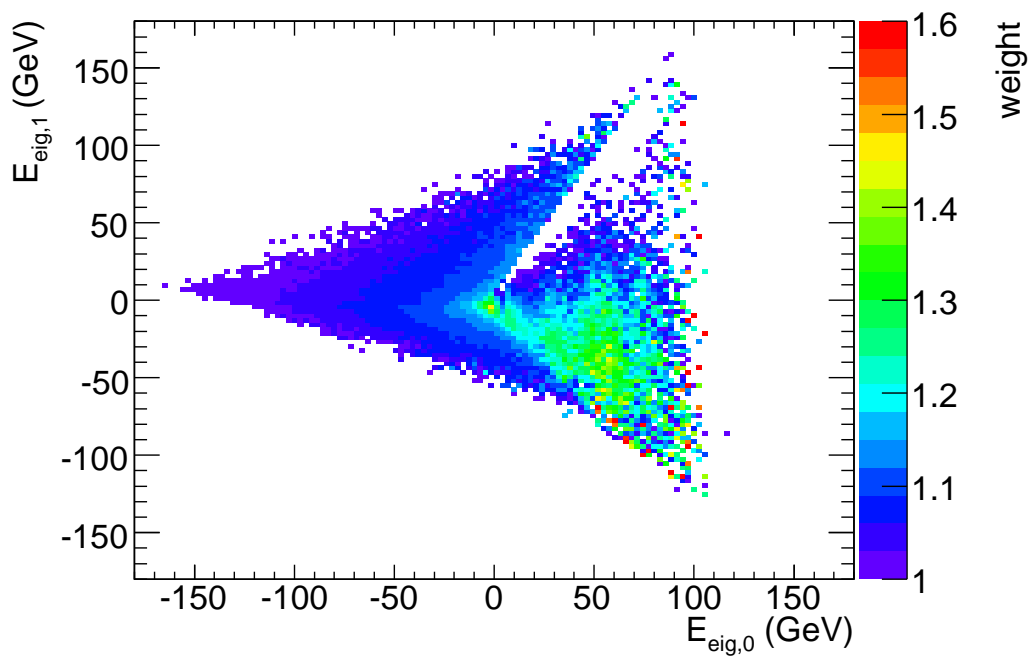


Figure 4: Weight table for simulated pion–proton mixed events (45% proton contamination) in LAr third (middle) layer and Tile first layer.

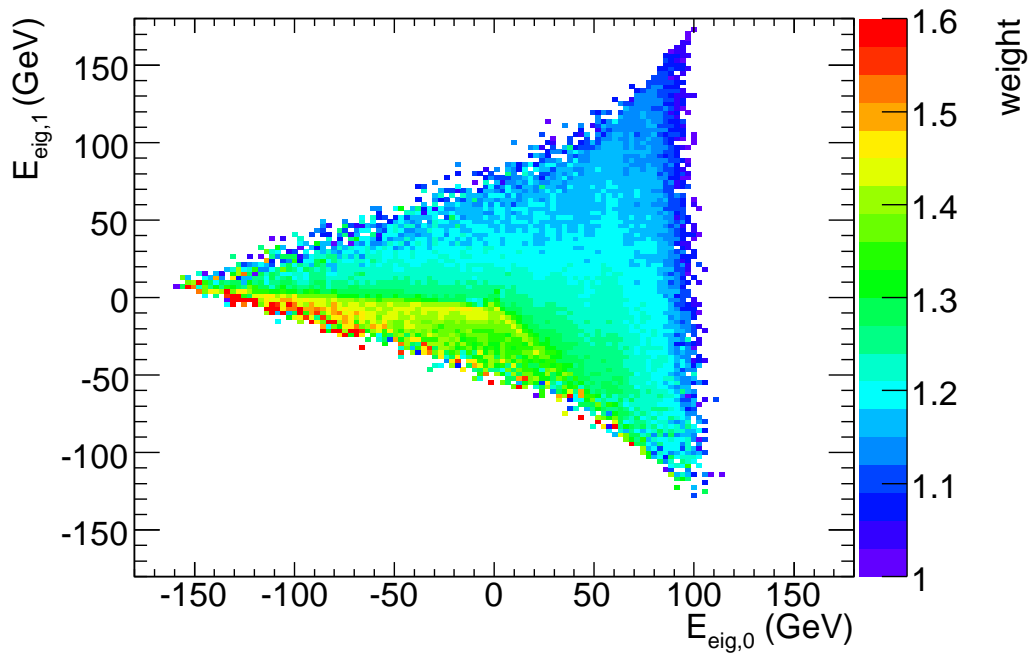


Figure 5: Weight table for simulated pion-proton mixed events (45% proton contamination) in Tile second layer.

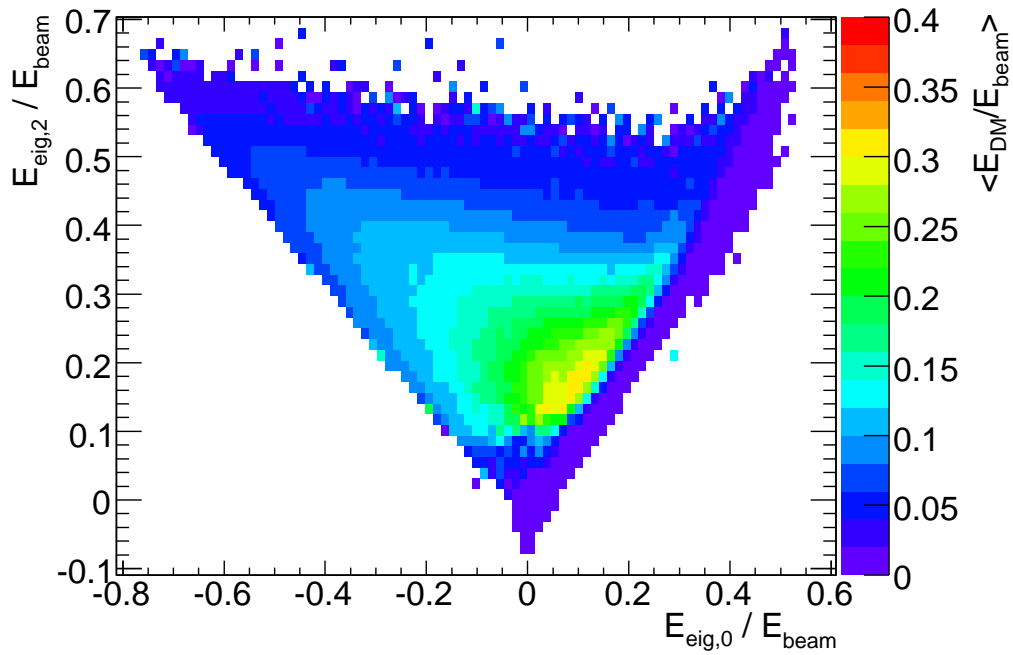


Figure 6: Look-up table for LAr-Tile dead material corrections for 45% proton contamination

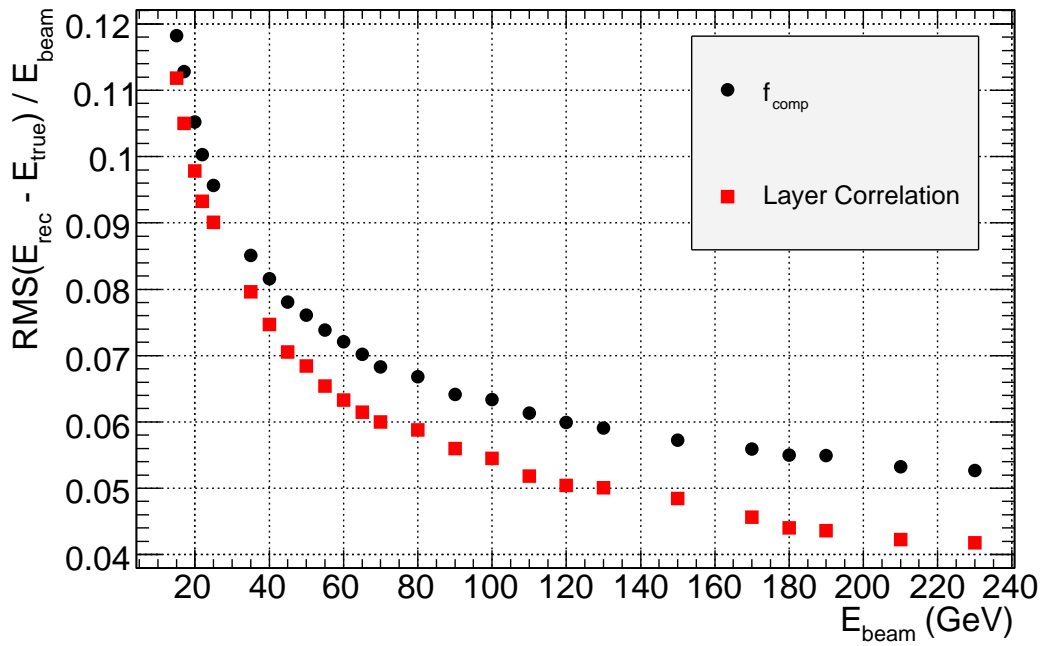
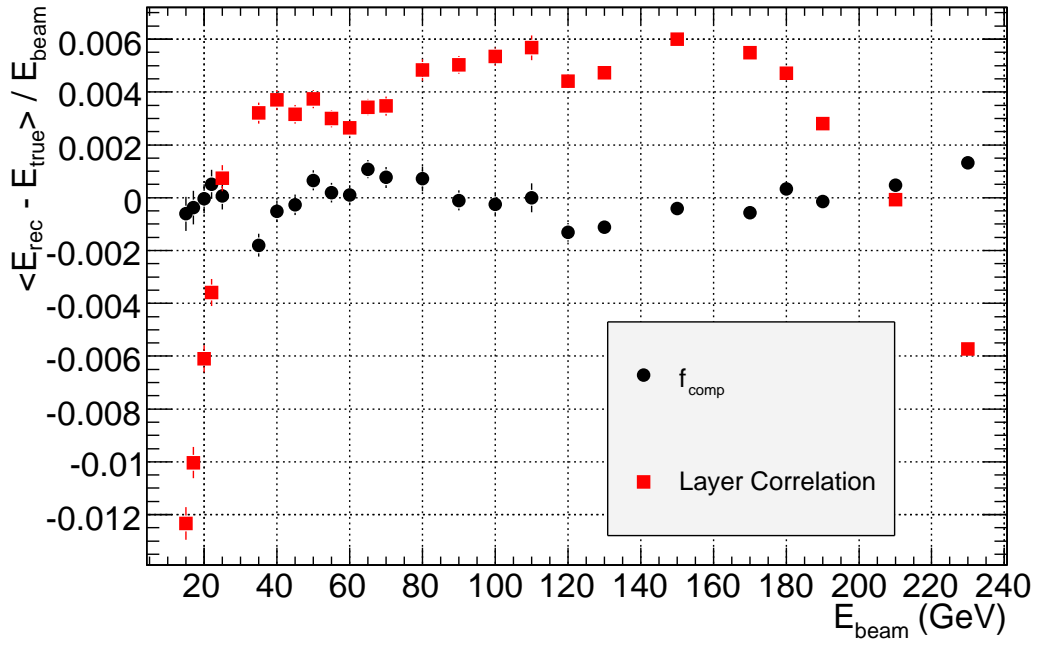


Figure 7: Bias and resolution of weighted energy minus true deposited energy for energy deposited in the calorimeters in simulated samples for f_{comp} weighting and LC weighting.

5.2 Dead Material Corrections

Figure 8 again shows the bias of the weighted energy, and also the bias of the dead material corrections. Contrary to the weighted energy, the LAr–Tile dead material correction mostly has a slight negative bias, with a positive bias at low energy. This cancels out most of the bias from the weighting.

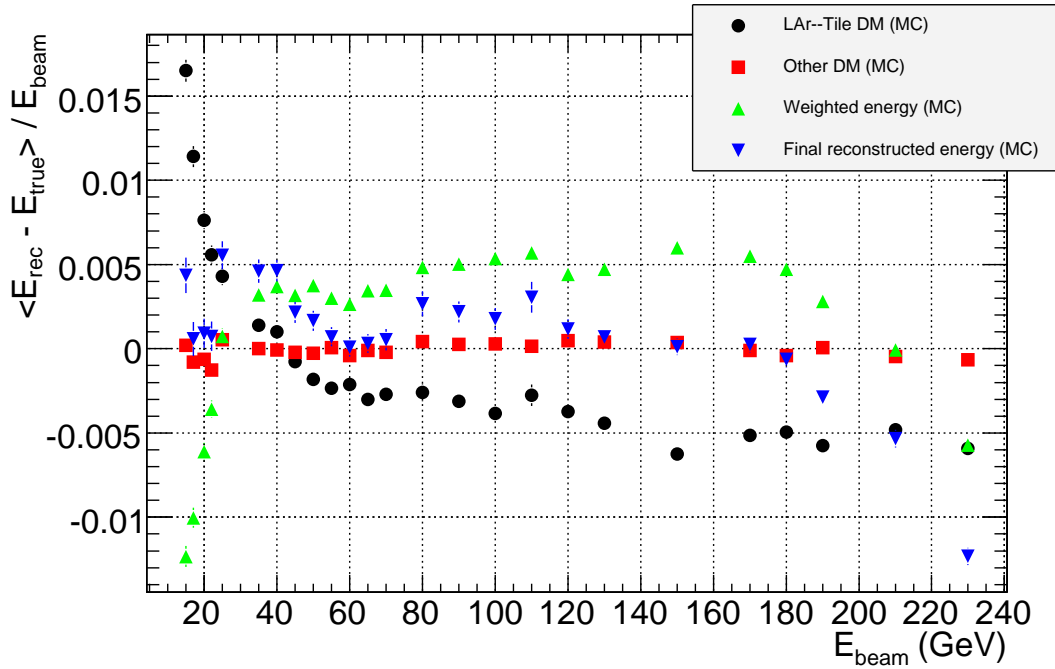


Figure 8: Bias (reconstructed energy minus true deposited energy, divided by beam energy) for the three individual corrections: Weighted calorimeter energy, correction for dead material between the liquid argon and Tile calorimeters, and other dead material corrections. Lastly, the bias of the final reconstructed energy, which is the sum of the three.

5.3 Linearity and Resolution

The performance for the fully corrected energy reconstruction is finally assessed in terms of linearity with respect to the beam energy and relative resolution.

The reconstructed energy distribution is fitted with a Gaussian distribution in the interval $(\mu - 2\sigma, \mu + 2\sigma)$, where μ and σ are the mean value and the standard deviation, respectively.

The mean value E_{fit} and the standard deviation σ_{fit} of the fitted Gaussian are used together with the beam energy E_{beam} to define the linearity and the relative resolution:

- the linearity is $E_{\text{fit}}/E_{\text{beam}}$ as a function of E_{beam} , while
- the relative resolution is $\sigma_{\text{fit}}/E_{\text{fit}}$ as a function of E_{beam} .

Both linearity and relative resolution are derived for the energy distribution at four stages:

- at the electromagnetic scale,
- at the compensated scale, i.e. after applying the compensation weights,

- after compensating weighting and application of dead material correction for losses between LAr and Tile, and
- after compensation weighting and all dead material corrections.

The evolution of the energy distributions after applying the subsequent corrections are shown in figures 9 and 10. Figure 11 shows linearity and relative resolution for the (20 GeV, 180 GeV) interval for fully simulated pions.

The linearity plot shows that the electromagnetic scale is reconstructing only two-thirds of the beam energy. The compensation weights push the recovery to about 90% of the beam energy. Finally the dead material corrections allow beam energy to be recovered within 1% for pion energies above 40 GeV and within 3% for lower energies ⁵⁾.

The relative resolution improves when applying the different correction steps. At high beam energies (above $E_{\text{beam}} = 100$ GeV) the contribution of the compensation weights to the improvement in resolution has the same magnitude as that of the LAr–Tile dead material corrections. At lower beam energies dead material corrections account for about 70% of the relative resolution improvement down to about $E_{\text{beam}} \simeq 30$ GeV. Below $E_{\text{beam}} \simeq 30$ GeV all the corrections account for a similar fraction of the improvement: other dead material corrections than those for LAr–Tile account for about 20% of the resolution improvement, they are marginal above that threshold.

6 Results

The distributions of the fundamental inputs to the calibration corrections are compared in data and Monte Carlo. The weights and dead material corrections derived from Monte Carlo “mixed correction” samples are finally applied on both data and “mixed signal” Monte Carlo samples.

6.1 Data to Monte Carlo Comparisons

The pion–proton “mixed signal” samples are used to compare data and Monte Carlo for the distribution for the first three components of the layer energy vector along the basis of covariance matrix eigenvectors as defined in section 4.4. Figure 12 shows such a comparison for a proton contamination of 45%. Very good agreement is obtained between data and simulation.

The normalized energy distributions (in unit bins of energy and events) for data and Monte Carlo are compared for all energies in figures 13 and 14. The four stages of corrections are shown. The quality of the initial description of data by Monte Carlo is not modified by the application of the compensation weights and dead material corrections (also see section 6.2).

⁵⁾The apparent discontinuity between the results at energies below 150 GeV and those above might be due to a geometry change in the description of the beam test setup: three cm of aluminum were included in the Inner Detector system for energies larger or equal than 150 GeV.

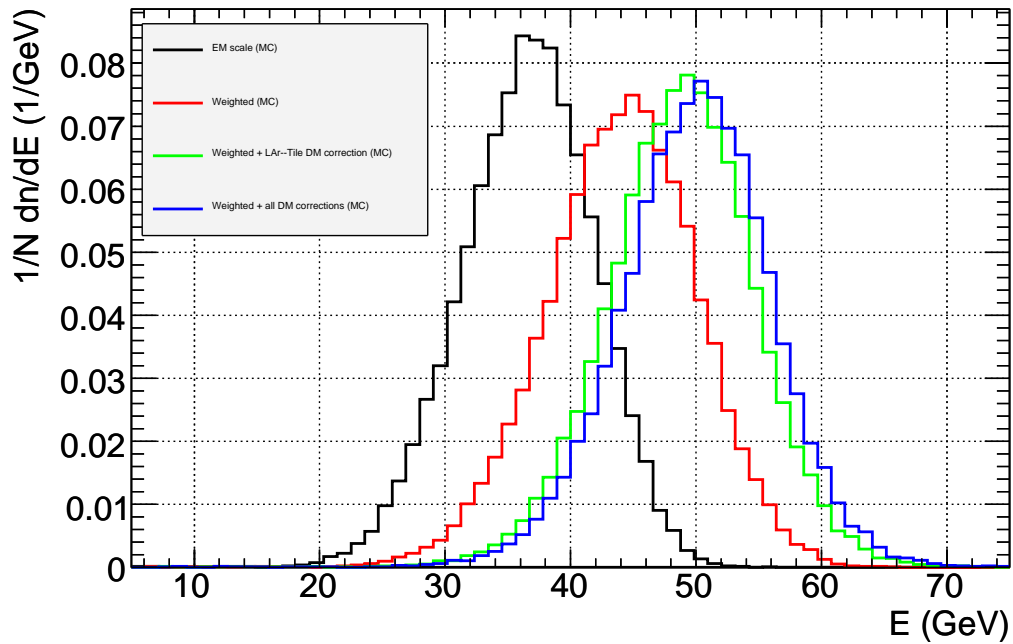
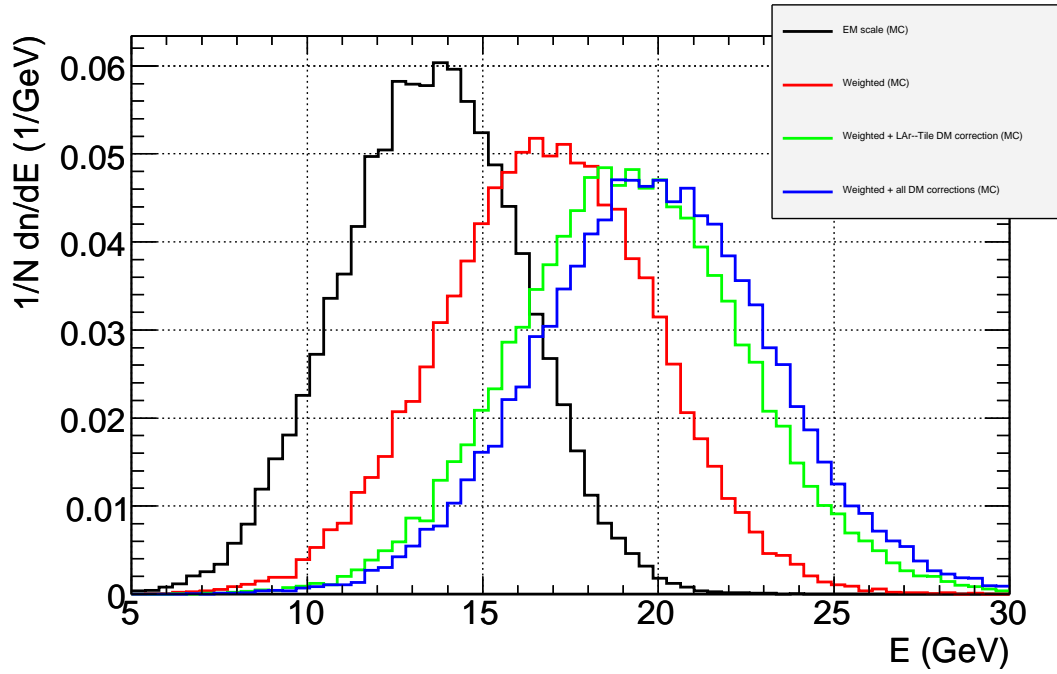


Figure 9: Evolution of reconstructed energy distribution at the different correction stages mentioned in the text: from electromagnetic scale to full dead material and compensation correction. The upper plot shows the simulation of 20 GeV pions, while the lower plot features 50 GeV pions.

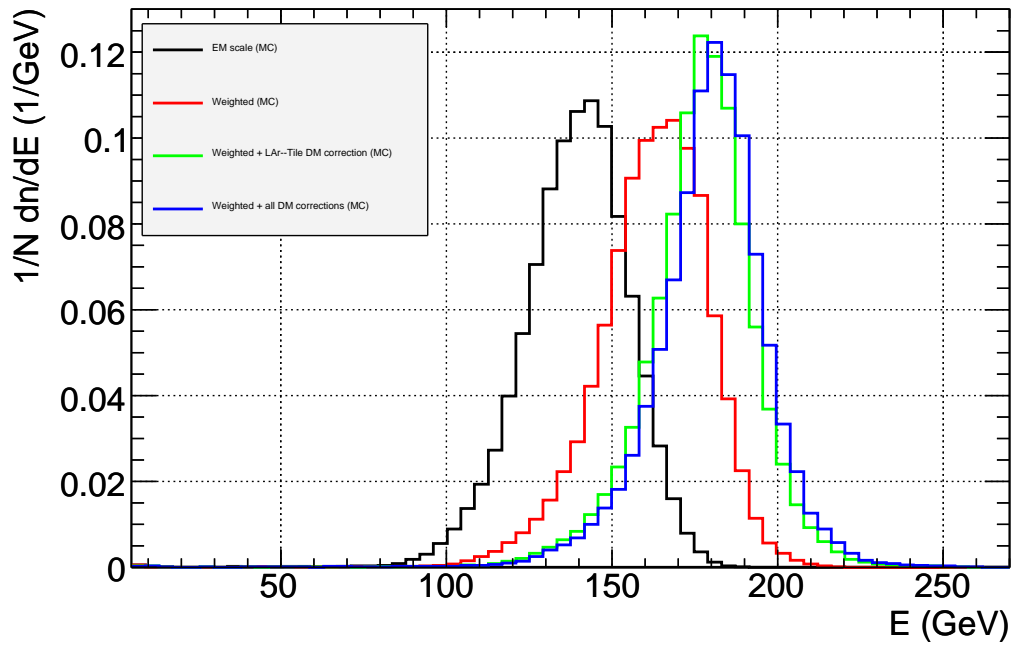
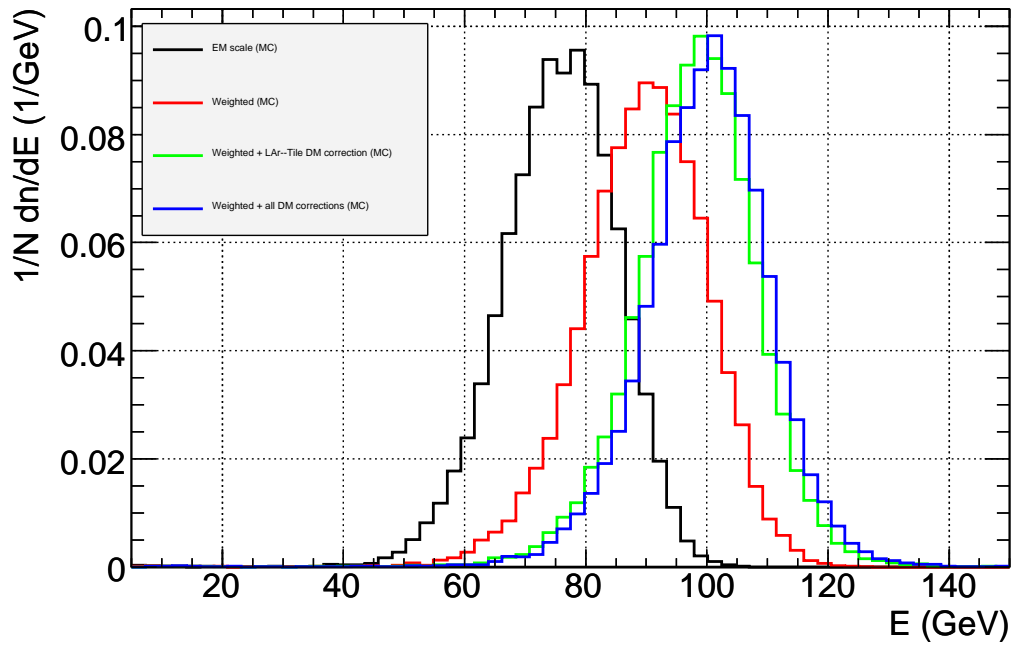


Figure 10: Evolution of reconstructed energy distribution at the different correction stages mentioned in the text: from electromagnetic scale to full dead material and compensation correction. The upper plot shows the simulation of 100 GeV pions, while the lower plot features 180 GeV pions.

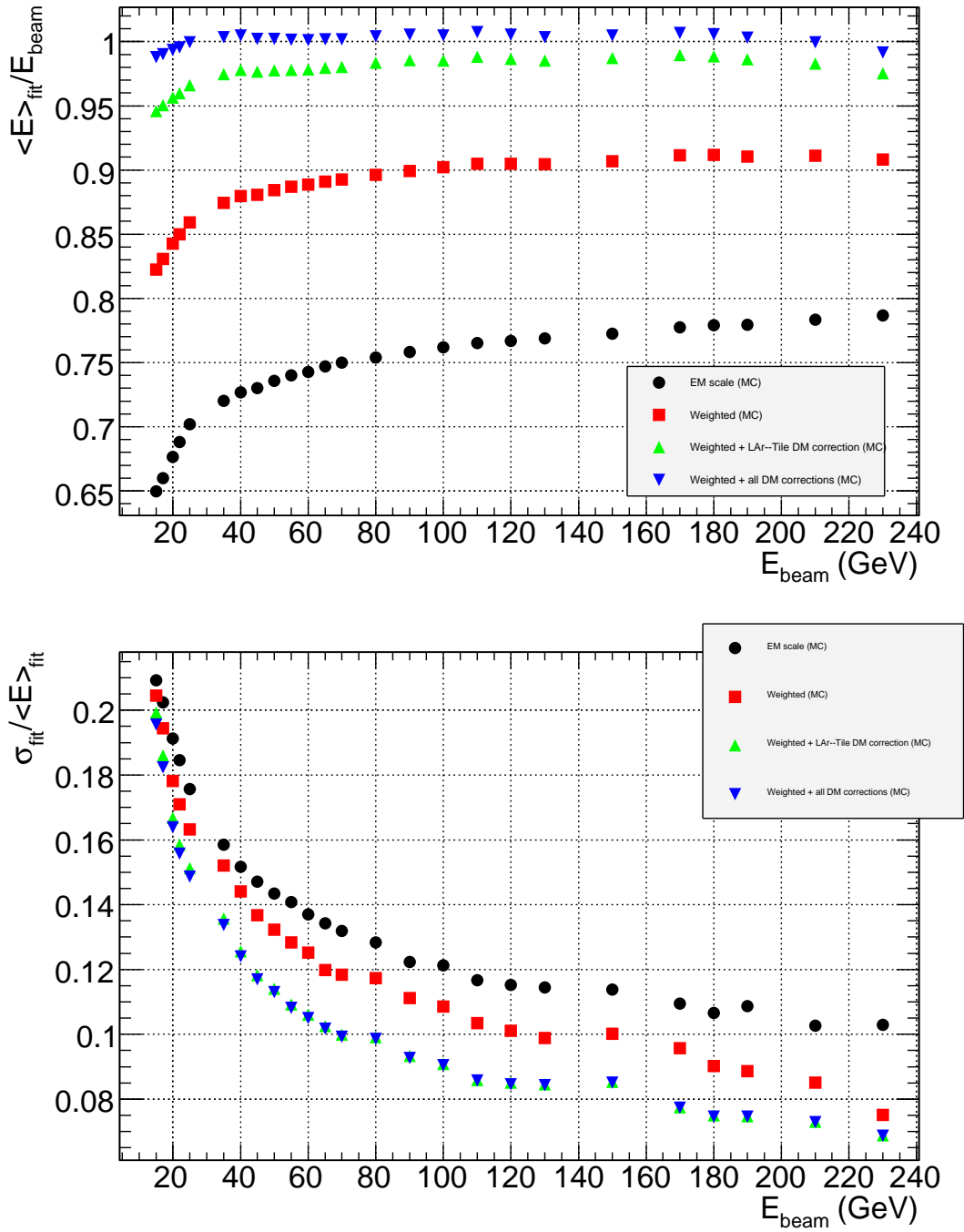


Figure 11: Linearity (upper plot) and relative resolution (lower plot) for fully simulated pure pion samples.

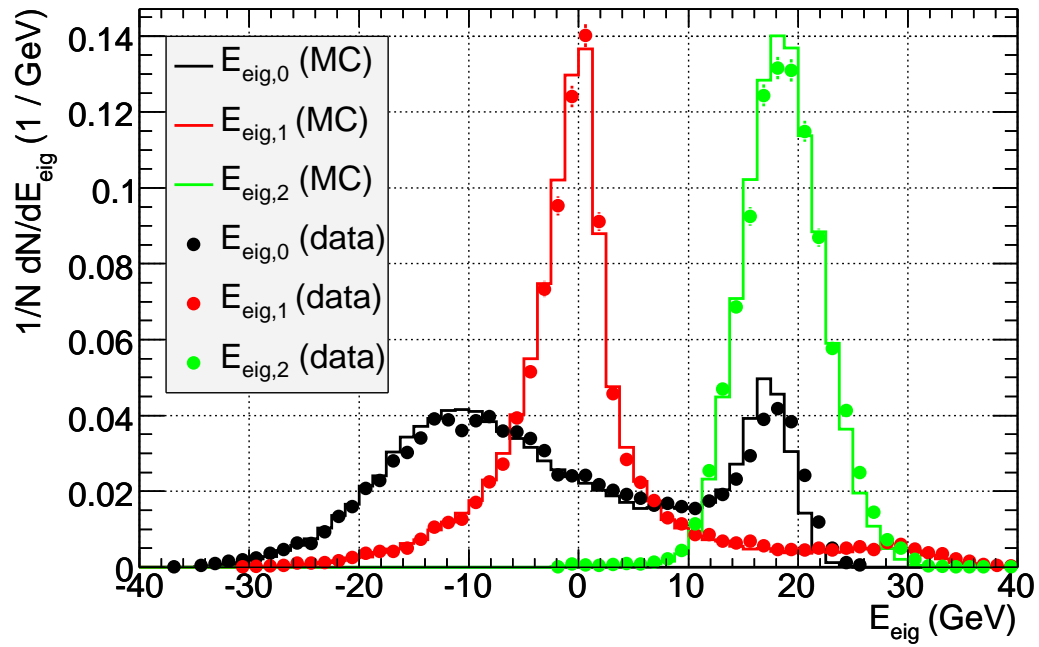


Figure 12: Distribution of the first three eigenvector components for data (filled circles) and Monte Carlo pion–proton “mixed signal” with a proton contamination of 45% (table 1).

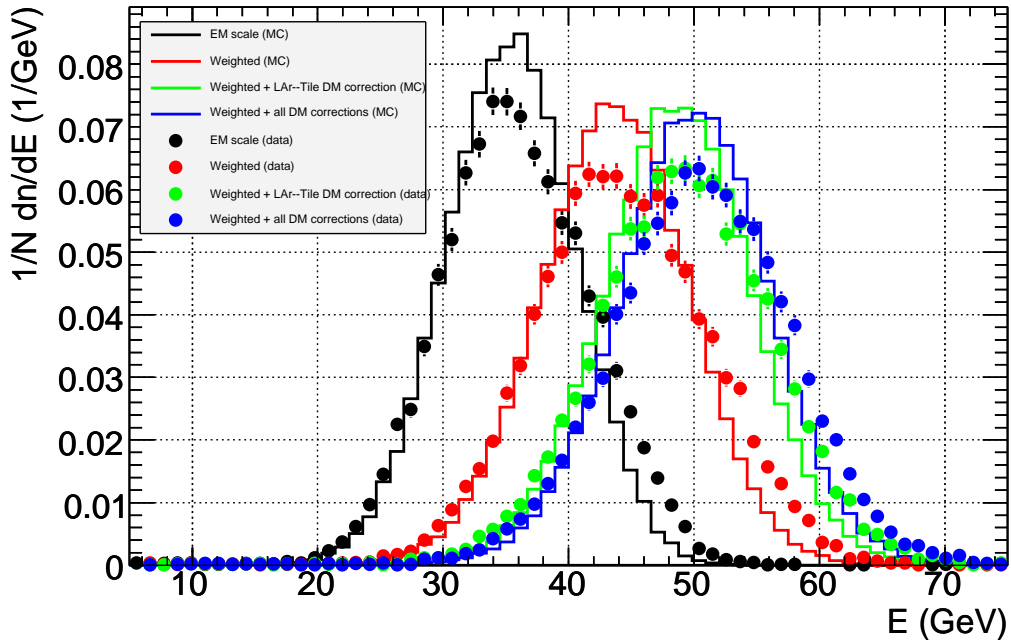
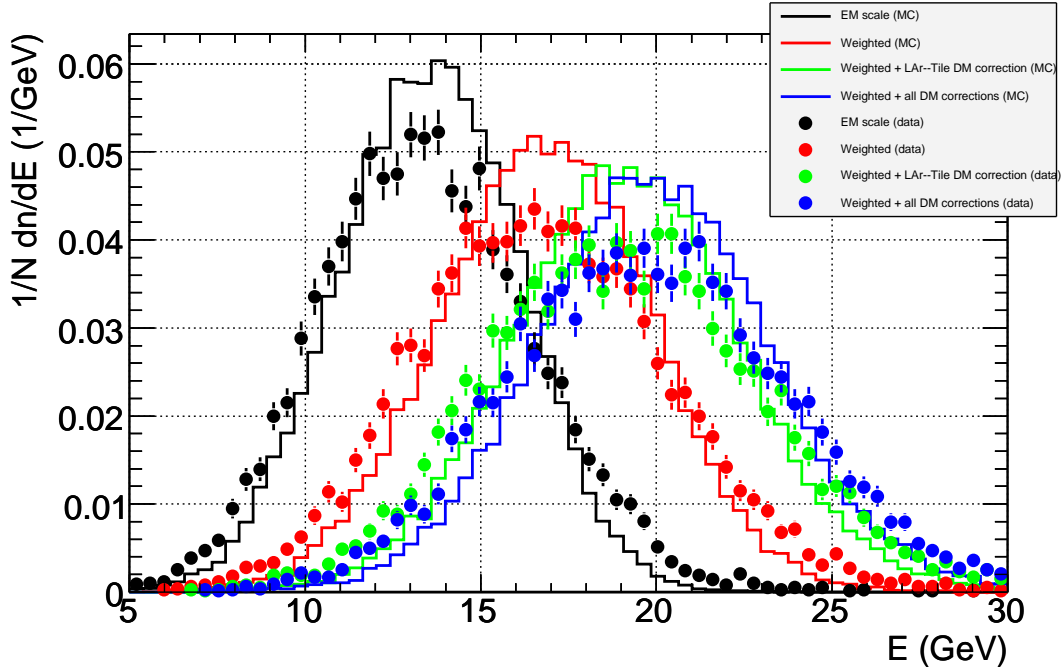


Figure 13: Normalized energy distribution for $E_{\text{beam}} = 20$ GeV (upper plot) and $E_{\text{beam}} = 50$ GeV (lower plot) after applying subsequent corrections for compensation and dead material effects. Data (filled circles) are compared with Monte Carlo simulation (solid lines).

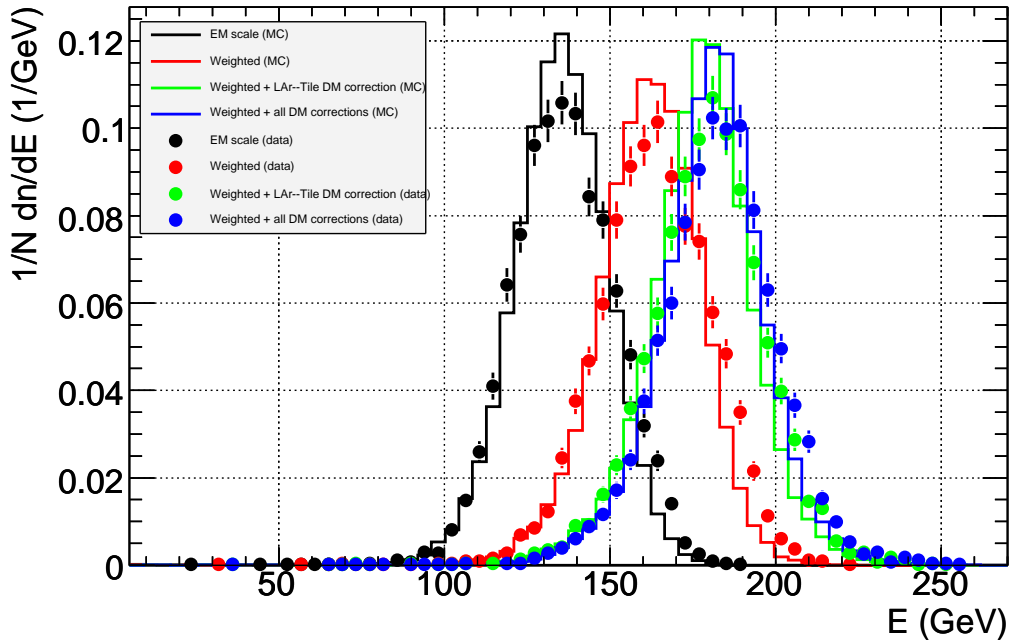
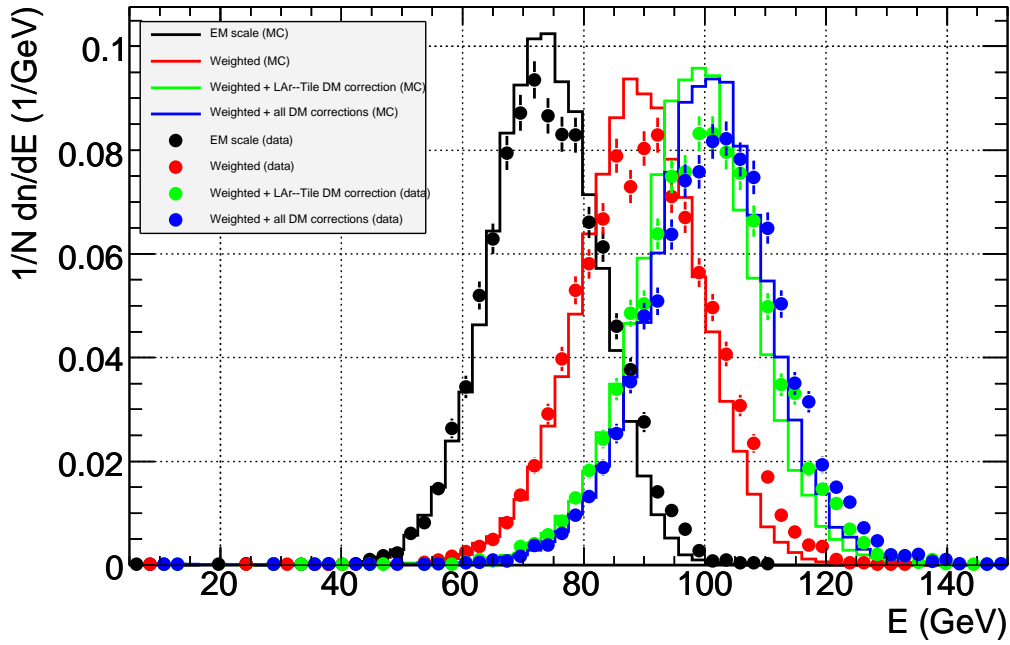


Figure 14: Normalized energy distribution for $E_{\text{beam}} = 100$ GeV (upper plot) and $E_{\text{beam}} = 180$ GeV (lower plot) after applying subsequent corrections for compensation and dead material effects. Data (filled circles) are compared with Monte Carlo simulation (solid lines) .

6.2 Linearity and Resolution

The linearity and relative resolution are extracted at all energies for both data and “mixed signal” Monte Carlo samples. Figure 15 shows that weighting recovers from 80% to 90% of the incoming beam energy. The dead material between LAr and Tile accounts for an additional 5% to 8%. The remaining dead material corrections allow linearity to be recovered within 3% (within 1% above 20 GeV). Weighting results are more important at high energies. Dead material effects play a more significant role at low energies particularly at 20 GeV where other corrections than LAr–Tile dead material are important to get to within 3% of the beam energy.

Figure 16 shows that the relative resolution is improved by about 17% to 22% in data when evolving from the electromagnetic scale to the fully corrected energy scale. A similar relative improvement is obtained in the Monte Carlo expectations: from 17% to 29%. The relative resolution is however smaller in Monte Carlo than in data: the discrepancies, at each correction stage, vary between 6% and 24% depending on the energy. The discrepancies in the shape of the total energy distribution are more pronounced at lower energies and they are already present at the electromagnetic scale (see figures 13 and 14).

The effect of the calibration technique on the energy reconstruction can be explored by considering the following double ratios:

$$\text{DRatio}_{\text{Lin}} = \left(\frac{E_{\text{fit,data}}}{E_{\text{fit,MC}}} \right)_{\text{rec}} / \left(\frac{E_{\text{fit,data}}}{E_{\text{fit,MC}}} \right)_{\text{corr}} \quad (18)$$

$$\text{DRatio}_{\text{Res}} = \left(\frac{(\sigma_{\text{fit}}/E_{\text{fit}})_{\text{data}}}{(\sigma_{\text{fit}}/E_{\text{fit}})_{\text{MC}}} \right)_{\text{rec}} / \left(\frac{(\sigma_{\text{fit}}/E_{\text{fit}})_{\text{data}}}{(\sigma_{\text{fit}}/E_{\text{fit}})_{\text{MC}}} \right)_{\text{corr}} \quad (19)$$

They represent the variation in the data-to-Monte Carlo ratio when corrections are applied to bring the reconstructed energy from the electromagnetic (rec) scale to the fully corrected hadronic (corr) scale. The ratio is considered both for the linearity and for the relative resolution as defined in section 5.3. The deviation of the double ratio from unity is a measure of the effect of the weighting technique on the description of the data by the simulation.

Figure 17 shows the evolution of the double ratio at three different stages of energy correction. The double ratio for linearity and resolution are consistent with unity within 0.7% and 5%, respectively.

For linearity such changes are of the same order of magnitude as the discrepancies between data and Monte Carlo at the electromagnetic scale: the agreement between data and Monte Carlo simulation is the same for all correction stages. This means that the Monte Carlo is able to predict the corrections that should be applied on the data. The ability of the Monte Carlo to reproduce the data at the electromagnetic scale (i.e. before any correction) seems to be the most critical limiting factor. The non-statistical nature of the fluctuations in the double ratios is affected by systematic effects such as the details of the peak determination in fitting and the different levels of proton contamination at different energies.

The sensitivity of the results to the assumed proton fraction at each beam energy was estimated by varying the fraction used to calculate the corrections. With the assumed fraction adjusted up or down one standard deviation of the TRT measurement, the relative variation in linearity and resolution in data and Monte Carlo was found to be of the order of 1% for $E_{\text{beam}} = 20$ GeV and 50 GeV and less than 0.5% for other energies.

For the relative resolution the changes are small, if compared with the discrepancies at the (rec) scale: the discrepancies do not get worse when the corrections are applied to the data. From preliminary studies the new Geant4 version (4.9) is able to provide a better description of the resolution in the data and it should be used for future analyses.

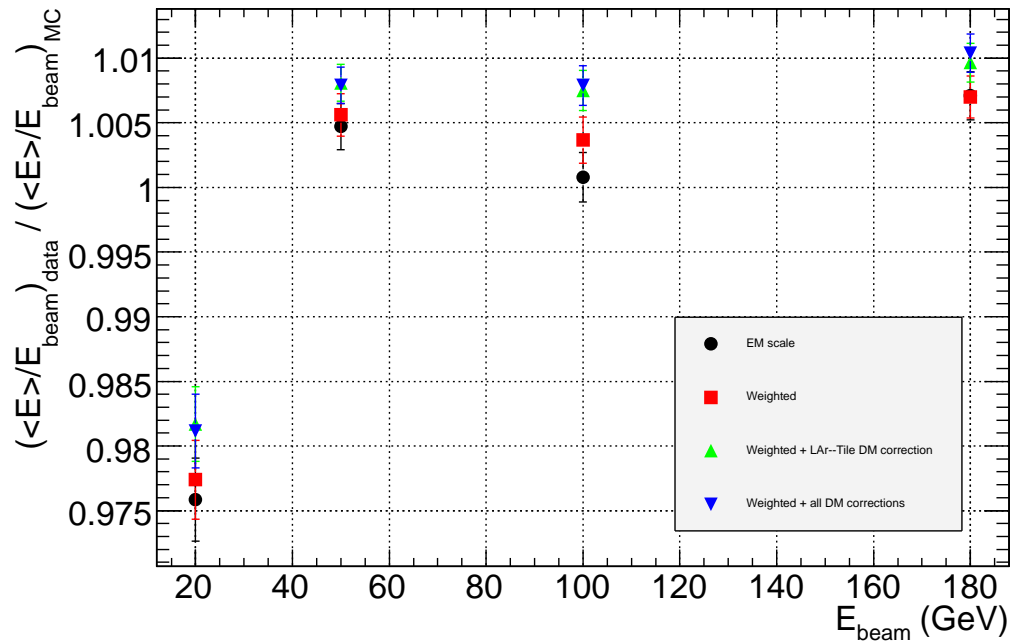
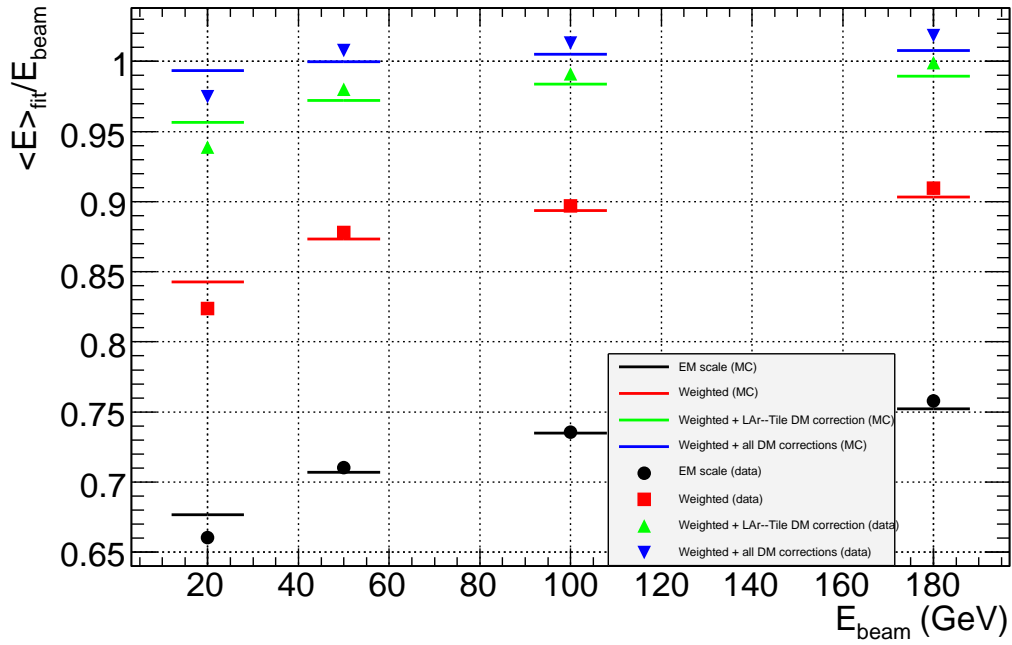


Figure 15: Data and Monte Carlo “mixed” samples are compared for linearity at all stages of the corrections. The upper plot shows the superposed absolute values, while the lower plot shows the evolution of the data-to-Monte Carlo ratio. See text for details.

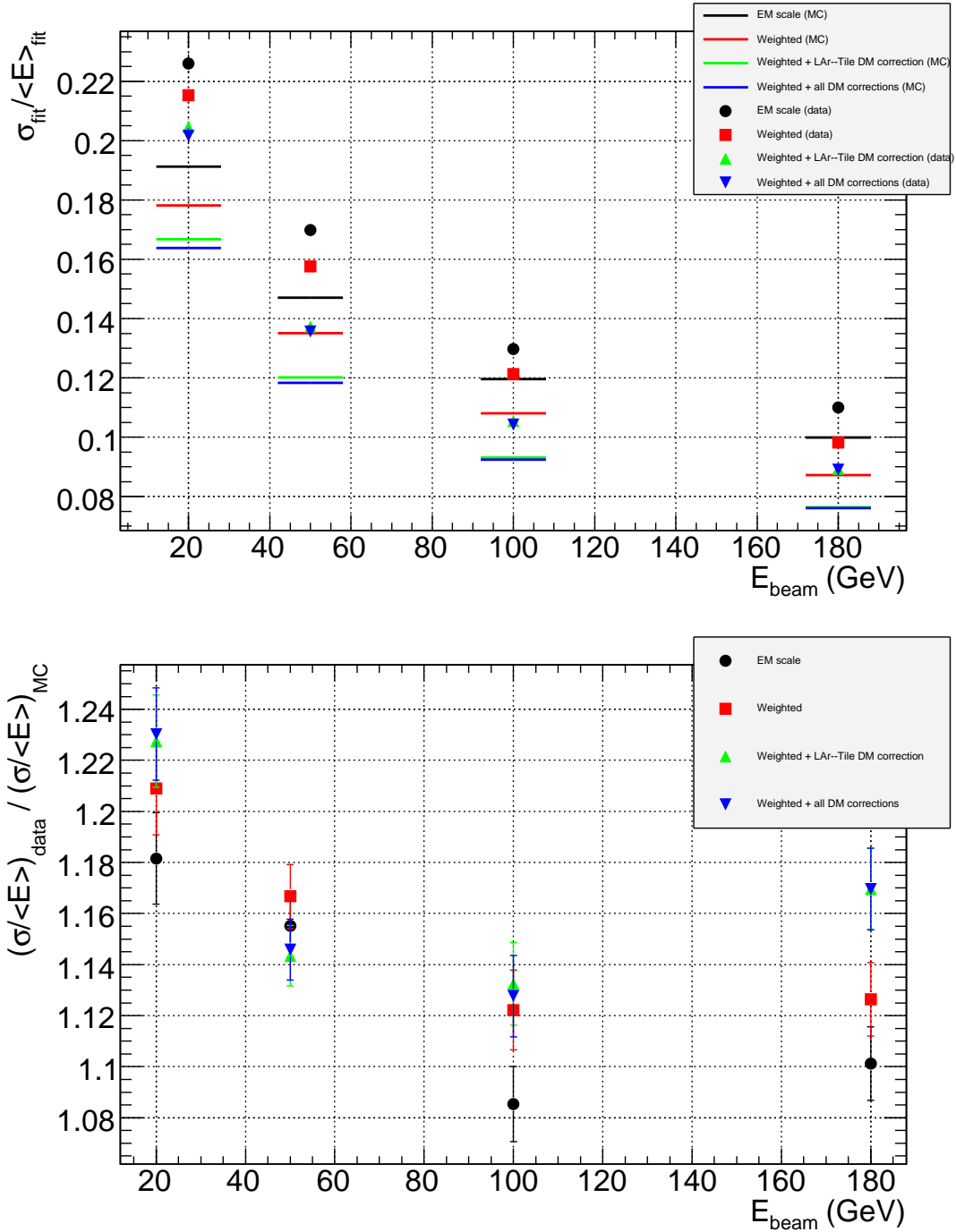


Figure 16: Data and Monte Carlo “mixed” samples are compared for relative resolution at all stages of the corrections. The upper plot shows the superposed absolute values, while the lower plot shows the evolution of the data-to-Monte Carlo ratio. See text for details.

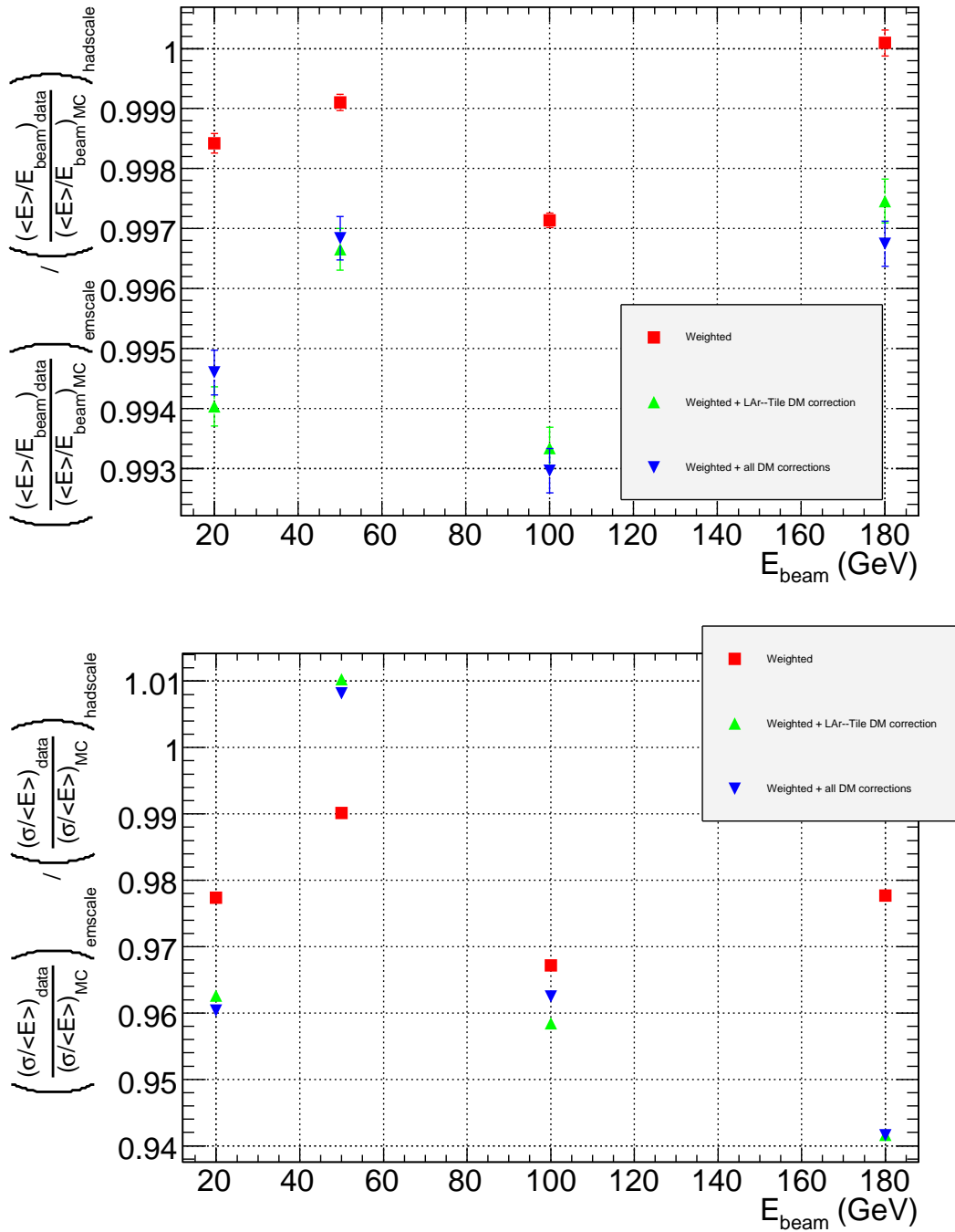


Figure 17: Double ratios of data to Monte Carlo for response and resolution as defined in equations 18 and 19, at three different correction stages, as functions of beam energy. The uncertainties on the electromagnetic and corrected ratios are assumed to be fully correlated.

7 Conclusions

An energy calibration technique was developed to deal in a coherent manner with both compensating the response to hadrons and their most significant dead material losses in a segmented calorimeter. The technique is based on the sensitivity of the correlation between the deposited energies in the different calorimeter layers to hadronic and electromagnetic deposits.

The calibration technique was applied to the energy reconstruction of pions impinging on a subset of the central ATLAS calorimeters during the ATLAS combined beam test in 2004. When taking into account the beam composition, linearity is recovered within 2% for $E_{beam} > 30$ GeV (3% overall) and relative resolution is improved by 17 to 22%. Consistency with the expectation from Monte Carlo studies is good for both the linearity and the percentage improvement in relative resolution. The absolute value of the relative resolution (after all corrections) is larger in data compared to Monte Carlo by 6% to 24%.

The discrepancies between data and Monte Carlo are inherited from the reconstructed energy at the electromagnetic scale and they are not altered significantly when applying the calibration. Additional improvement in the data description by Monte Carlo simulation can help to fulfill the expected absolute value for the relative resolution.

References

- [1] C. Leroy and P. Rancoita, *Physics of cascading shower generation and propagation in matter: Principles of high-energy, ultrahigh-energy and compensating calorimetry*, Rept. Prog. Phys. **63** (2000) 505.
C. W. Fabjan and F. Gianotti, *Calorimetry for particle physics*, Rev. Mod. Phys. **75** (2003) 1243.
- [2] R. Wigmans, *Calorimetry : Energy Measurement in Particle Physics*, (Oxford:Clarendon Press, USA, 2000).
- [3] D. E. Groom, *Energy flow in a hadronic cascade: Application to hadron calorimetry*, Nucl. Instrum. Meth. A **572** (2007) 633 [arXiv:physics/0605164] and references therein.
- [4] N. Nakajima, H. Miyata, A. Sanchez, H. Ono and S. Iba. *Correlation Matrix Method for Pb/Scint Sampling Calorimeter* in Proceedings of the 2005 International Linear Collider Workshop, 18th-22nd March 2005, Stanford, California (USA) and references therein. Available at <http://www.slac.stanford.edu/econf/C050318/papers/1023.PDF>.
- [5] A. Gupta, F.S. Merritt and J. Proudfoot, *Jet Energy Correction Using Longitudinal Weighting*, ATL-COM-PHYS-2006-062.
- [6] B. Di Girolamo et al., *Beamline instrumentation in the 2004 combined ATLAS testbeam*, ATL-TECH-PUB-2005-001, ATL-COM-TECH-2005-001 (CERN,2005).
- [7] C. Amsler et al., *The Review of Particle Physics*, Physics Letters **B667**, 1 (2008).
- [8] T. Petersen *et al.*, *Inner detector electron ID*, presentation at the ATLAS Trigger and Physics week, 7th November 2007 available at <http://indico.cern.ch/materialDisplay.py?subContId=4&contribId=17&sessionId=8&materialId=slides&confId=20504&relax>.
- [9] H. Hakobyan, M. Simonyan, T. Carli and A. M. Henriques-Correia, *Measurement of Pion and Proton Longitudinal Shower Profiles up to 20 Nuclear Interaction Lengths with the ATLAS Tile Calorimeter*, ATL-TILECAL-PUB-2007-008.

- [10] S. Agostinelli et al., *Geant4-A simulation toolkit*, Nuclear Instruments and Methods in Physics Research A 506 (2003) 250-303,
J. Allison et al., *Geant4 developments and applications* IEEE Transactions on Nuclear Science 53 No. 1 (2006) 270-278.
- [11] G. Folger and J. P. Wellisch, *String Parton Models in Geant4*, in Proceedings of *Computing in High Energy and Nuclear Physics*, 24-28 March 2003, La Jolla, California and references therein. Available at <http://www.slac.stanford.edu/econf/C0303241/proc/papers/MOMT007.PDF>.
- [12] M. P Guthrie, R.G. Alsmiller and H.W.Bertini, Nucl. Instr. and Meth. 66 (1968) 29;
H.W. Bertini and P. Guthrie, Nucl. Phys. A 169 (1971) 670;
N. V. Stepanov, Preprint: ITEP-55 (Moscow) 1988.
- [13] See <https://twiki.cern.ch/twiki/bin/view/Atlas/TopologicalClustering>.
- [14] J. E. Jackson, *A User's Guide to Principal Components* (p.505), (Newark, NJ: Wiley, USA, 2005).
- [15] See
<http://www.nevis.columbia.edu/~lelchuk/DeadMaterials/M.LeltchoukJan2004.pdf>
http://www.nevis.columbia.edu/~lelchuk/CalibrationHits/P.Loch_Note.txt
Notes collected at <http://www.nevis.columbia.edu/~lelchuk/CalibrationHits>.
- [16] See A. Maslennikov and G. Pospelov *Dead material Corrections Status*, talk presented at the Calorimeter Calibration workshop, 5th-8th September 2006, Costa Brava available at <http://indico.cern.ch/materialDisplay.py?contribId=49&sessionId=s4&materialId=slides&confId=a062339>.