

Artificial Neural Networks for reconstruction of energy losses in dead materials between barrel LAr and Tile calorimeters: exploration and results

J.A.Budagov¹⁾, J.I.Khubua^{1), 2)},
Y.A.Kulchitsky^{1), 3)}, N.A.Rusakovich¹⁾,
V.N.Shigaev¹⁾, P.V.Tsiareshka^{1), 3)}

¹⁾ Joint Institute for Nuclear Research, Dubna, Russia

²⁾ HEPI, Tbilisi State University, Tbilisi, Georgia

³⁾ Institute of Physics, National Academy of Sciences, Minsk, Belarus

Abstract

In the course of computational experiments with Monte-Carlo events for ATLAS Combined Test Beam 2004 setup Artificial Neural Networks (ANN) technique was applied for reconstruction of energy losses in dead materials between barrel LAr and Tile calorimeters (Edm). The constructed ANN procedures exploit as their input vectors the information content of different sets of variables (parameters) which describe particular features of the hadronic shower of an event in ATLAS calorimeters. It was shown that application of ANN procedures allows one to reach 40% reduction of the Edm reconstruction error compared to the conventional procedure used in ATLAS collaboration. Impact of various features of a shower on the precision of Edm reconstruction is presented in detail. It was found that longitudinal shower profile information brings greater improvement in Edm reconstruction accuracy than cell energies information in LAr_3 and $Tile_1$ samplings.

1 Objectives of the present investigation

Procedures for energy reconstruction of a hadronic shower usually include an additive term which gives an estimate of energy losses in dead materials between barrel LAr and Tile calorimeters (Edm) [1, 2, 3]. The error of energy loss reconstruction affects directly the overall precision of hadronic energy reconstruction procedures. In the present investigation we intend to get answers to the following questions:

1. How good is the conventional procedure for estimating the energy lost in dead materials between barrel LAr and Tile calorimeters?
2. Is it possible to construct a procedure of appreciably higher precision?
3. How to realize such a procedure?
4. What gain in precision of such a procedure may be reached by exploiting information content of different sets of additional variables (parameters) which describe particular features of a calorimeter response?

2 Artificial Neural Networks

Artificial Neural Networks (ANN) in the form of Multi-Layer-Perceptrons (MLP) with back-propagation updating were applied for function mapping purpose. Generally the mapping is nonlinear. For example, in the case of 4-layered MLP the mapping may be written in the form:

$$F(X) = g \left(\sum_i w_{0i} g \left(\sum_j w_{ij} g \left(\sum_k w_{jk} x_k + \theta_j \right) + \theta_i \right) + \theta_0 \right), \quad (1)$$

where

- g — neuron activation function (linear or sigmoidal),
- x_k — is the k -th element of the input vector X (activation value of k -th neuron in the input layer of MLP),
- w_{0i} — weight (synaptic weight or strength) of the connection between the single output neuron of MLP and the i -th neuron in the second hidden layer,
- w_{ij} — connection weight between i -th neuron in the second hidden layer and j -th neuron in the first hidden layer,
- w_{jk} — connection weight between j -th neuron in the first hidden layer and k -th neuron in the input layer,
- $\{\theta\}$ — thresholds.

3 Some properties of MLPs and technological aspects of MLPs application in the present investigation

The main properties of Multi-Layer-Perceptrons are:

- MLPs are universal function approximators [4, 5].
- MLP is fully defined by its structure and the whole set W of its synaptic weights.
- There exists well elaborated technique for stochastic approximation of unknown function with MLPs through so called "learning" procedure performed on a set of known examples (train set) to define the optimal W set that minimizes the global approximation error for the train set.
- The trained MLP performs non-linear mapping from multidimensional input space of the function arguments onto one-dimensional space of the function value.
- Correctly trained MLPs demonstrate good generalization performance, i.e. ability to model correct mapping of data it has "never seen" before.
- MLPs are computationally efficient, i.e. their computational complexity (measured in number of operations) linearly depends on the number of synaptic weights W in the net.

In the context of neural networks (NN) applications the complexity of the dead matter (DM) energy reconstruction problem was practically tested in the course of computational experiments with the use of a series of neural networks of various topologies and number of nodes (neurons). The number of hidden layers in NNs ranged from 1 to 4, and the number of nodes in a layer – from 2 to 40. The total number of tested networks is around 100. We found that NNs with 2 hidden layers worked satisfactorily, and the use of greater number of hidden layers did not bring better performance. On the other hand, single hidden layer networks with increased number of nodes in a layer showed that they were more likely to stuck at higher minima of NN cost function than networks with 2 or more hidden layers.

To achieve good generalization quality it is important that the number of events for training NN well exceeds the number of weights in NN. During the present investigation we had enough Monte Carlo events to reach the ratio around 20 – 60 events per one connection weight of the NN. Files containing from 20000 to 25000 MC-events were generated for each beam energy of

CTB04. In accordance with the standard procedure each file of MC-events was split onto 2 sets: 1) the test set containing 6000 events and 2) the train set containing the remaining events in the file. No preselection procedure was applied to the events. Further on, we split the test set onto 3 subsets, each subset of test events contains 2000 events.

In order to hold the problem of NN overtraining under control we do not use the so called early stopping method of training. Instead, we preset the length of train/test sessions to be big enough (10000 epochs) to assess asymptotical behavior of 5 train/test curves (train curve for the train set, 3 curves for 3 subsets of test events and one curve for the full set of test events). To admit NN as normally trained network we require that after train curve reached the plateau the following condition must be fulfilled: within the span of subsequent several thousand epochs all 5 curves lie within certain limits $C \pm \delta$. The used values of δ correspond to $\sim 1.5\%$ of the RMS value of DM energy reconstruction by the ATLAS conventional method (see subsequent sections of the present paper). The admitted parameters of the NN are those saved at the beginning of the plateau.

The subdivision of test events onto 3 subsets was introduced in order to facilitate location of rare MC-events (1 – 2 per thousand) which manifest very big energy deposition in the dead matter (about 50% of the incident pion energy). We did not eliminate such events from files but used them to explain slightly different behavior of corresponding test curves for some subsets.

Investigation of the problem of DM energy reconstruction with NN approach was performed for each CTB04 energy independently. This was done to avoid the influence of the properties of a more general NN on solving the main task of the present work: to establish for each energy the limits of accuracy attainable in DM energy reconstruction with the use of certain features of a hadron shower as procedure arguments.

In Section 4 the items of the pool of variables are expressed in natural units. They need preprocessing before using them as components of NN input vectors. Preprocessing of all items which are expressed in energy units (GeV) consists in normalization by a constant equal to the nominal value of the incident pion energy. Other items are normalized by constants which are defined after analysis of their distributions in the files.

In the course of the present investigation we used the artificial neural network package JETNET 3.0 [6] which consists of a set of subroutines that realize a variety of minimization options.

4 Pool of Variables for ANN construction

Here we present the pool of variables which was used for ANN construction:

Sampling energies info:

ELAr1, ELAr2, ELAr3	—	Energies in LAr samplings
ETile1, ETile2, ETile3	—	Energies in Tile samplings

LAr₃ cell energies info:

E1MX_LAr3		
E2MX_LAr3	—	The largest 4 cell energies in <i>LAr₃</i>
E3MX_LAr3		sampling for the current event
E4MX_LAr3		

Tile₁ cell energies info:

E1MX_Tile1		
E2MX_Tile1	—	The largest 4 cell energies in <i>Tile₁</i>
E3MX_Tile1		sampling for the current event
E4MX_Tile1		

Active cells info:

FN2C_LAr3	—	The number of cells in <i>LAr₃</i> with $E_{cell} > 3\sigma_{noise,LAr3}$
FN2C_Tile1	—	The number of cells in <i>Tile₁</i> with $E_{cell} > 3\sigma_{noise,Tile1}$

Energy spread factors in cell energies distribution:

F2LAr3	—	Energy spread factor in <i>LAr₃</i> cell energies distribution
F2Tile1	—	Energy spread factor in <i>Tile₁</i> cell energies distribution

Energy spread factor is the function of cell energies:

$$F2 = \left(\sum_j e_j^2 \right) \cdot \left(\sum_j e_j \right)^{-2}, \quad (2)$$

where e_j is the energy of the j -th cell in a sampling.

5 Data of the computational experiment

The investigation was performed on the basis of MC events generated for 9 energies of the incident pions: 10, 20, 50, 100, 150, 180, 250, 320 and 350 GeV.

For events generation ATHENA, release 12.0.6 was used with QGSP_GN physics list. Events were generated for ATLAS CTB04 setup with beam direction: $\eta = 0.25$, $\phi = 0.0$. From 11000 to 25000 events were generated at each beam energy.

As the first step we performed thorough investigations for 2 energies: 250 and 10 GeV. The obtained results for these 2 energies clarify well the main aspects of the problem to be solved. In what follows we present distributions of differences between true values of Edm (supplied by MC generator) and Edm values reconstructed by procedures under consideration. The form of these distributions is not Gaussian, and as a measure of widths of these distributions we used RMS values.

6 Precision of the conventional procedure

How precise is the conventional procedure for estimating the energy Edm lost in dead materials between barrel LAr and Tile calorimeters? In what follows the variables and parameters of the conventional procedure we mark by "calc" label.

According to the conventional procedure [1, 2, 3] the Edm is estimated as the mean geometric value of two sampling energies:

$$Edm_{calc} = C \cdot \sqrt{E_{LAr3} \cdot E_{Tile1}} \quad (3)$$

where E_{LAr3} and E_{Tile1} are energy depositions in the calorimeter samples which are the nearest to the dead material.

Before estimating RMS of the distribution of differences ($Edm_{true} - Edm_{calc}$) at 250 GeV the calibration constant C in the above expression was adjusted to fulfill the equality:

$$\langle Edm_{true} \rangle = \langle Edm_{calc} \rangle = C \langle \sqrt{E_{LAr3} \cdot E_{Tile1}} \rangle, \quad (4)$$

The value of $\langle Edm_{true} \rangle$ and RMS of ($Edm_{true} - Edm_{calc}$) distribution were evaluated over 25000 events. The obtained results are: $\langle Edm_{true} \rangle = 19.53$ GeV or 7.8% of the nominal beam energy, $RMS(Edm_{true} - Edm_{calc}) = 7.55$ GeV or 3% of the beam energy.

A question arises: can other functions of the 2 arguments ELAr3 and ETile1 yield higher precision in reconstruction of Edm_{true} ?

The answer to this question is given by the results of the following 2 computational experiments.

6.1 Experiment 1.

Approximation of Edm_{true} is done by the linear function

$$Edm_{linear} = A0 + A1 \cdot ELAr3 + A2 \cdot ETile1. \quad (5)$$

This function was realized with the use of the linear perceptron NN 28A whose layers structure is "2 - 1":

- 2 neurons (nodes) in the input layer (for ELAr3 and ETile1 values),
- No hidden layers,
- 1 output neuron (for Edm_{linear} value),
- Linear activation function in each neuron of NN 28A.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length - 500 epochs.

After performing NN 28A train and test sessions we get from $(Edm_{true} - Edm_{linear})$ distribution the following NN 28A result (see Fig. 1):

$$\text{RMS}(Edm_{true} - Edm_{linear}) = 8.76 \text{ GeV}.$$

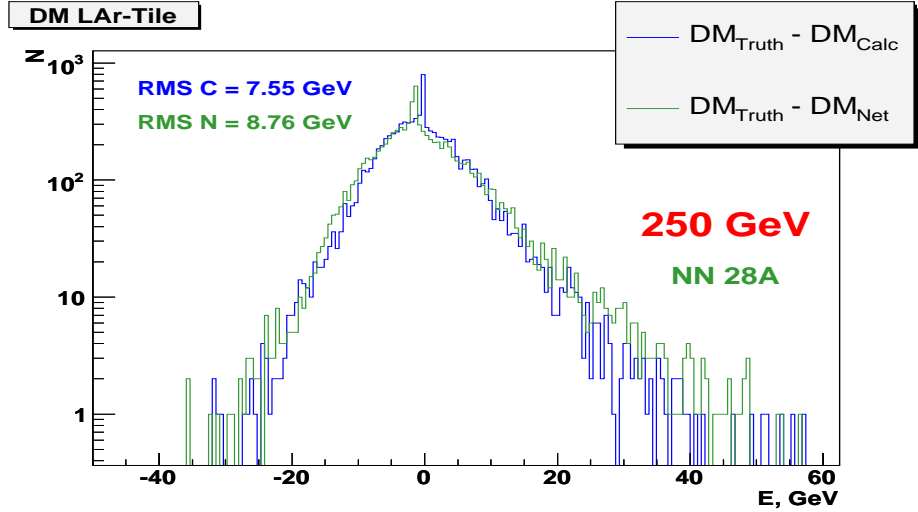


Figure 1: Distribution of Edm reconstruction error for linear approximator NN 28A at 250 GeV.

Compared to the conventional method of Edm_{true} reconstruction NN 28A yields the following **loss** in precision:

Absolute loss in precision = 1.23 GeV,

Relative loss in precision = 16.3%,

were "Relative loss in precision" is $100\% \cdot (Abs.loss) / RMS(Edm_{true} - Edm_{calc})$.

It is seen that the **linear** function of ELAr3 and ETile1 arguments leads to **greater error** in reconstruction of Edm_{true} than the conventional procedure.

6.2 Experiment 2.

The general case of Edm_{true} approximation was realized with the use of NN 28B whose layers structure is "2 - 12 - 1":

- two neurons (nodes) in the input layer (for ELAr3 and ETile1 values),
- one hidden layer with 12 neurons,
- one output neuron (for Edm_{net} value),
- sigmoidal activation function $g(x) = (1 + e^{-2x})^{-1}$ for hidden neurons and linear activation function $g(x) = x$ for the output neuron of NN 28B.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length - 1500 epochs.

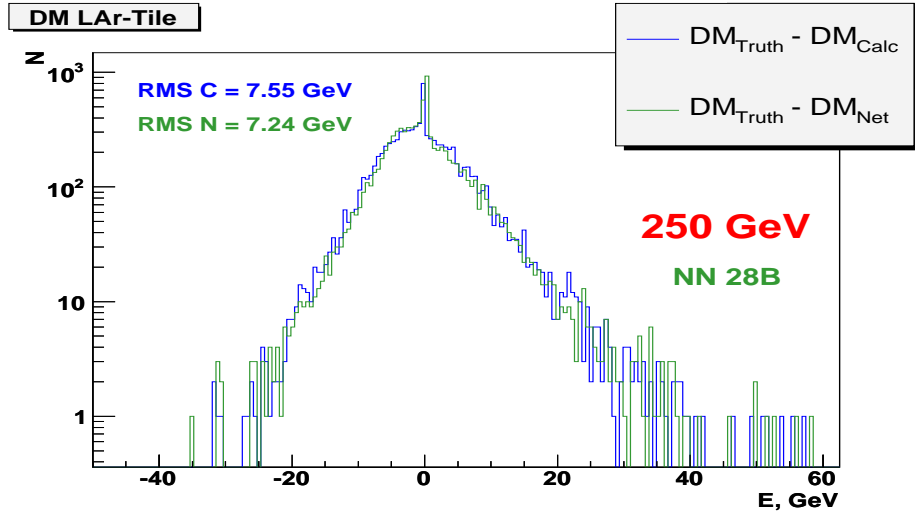


Figure 2: Distribution of Edm reconstruction error for general approximator NN 28B of 2 variables ELAr3 and ETile1 at 250 GeV.

Experiment 2 gave the following result (see Fig. 2):

$$\mathbf{RMS(Edm_{true} - Edm_{net}) = 7.24 GeV}$$

Compared to the conventional method the use of NN 28B gives rather little gain in precision:

$$\mathbf{Absolute\ gain = 0.31 GeV,}$$

$$\mathbf{Relative\ gain = 4.1 \%,}$$

were "Absolute gain" is $RMS(Edm_{true} - Edm_{calc}) - RMS(Edm_{true} - Edm_{net})$, and "Relative gain" is $100\% \cdot (Abs.gain) / RMS(Edm_{true} - Edm_{calc})$.

Results of these 2 experiments lead to the conclusion that among methods using 2 variables (ELAr3, ETile1) as their arguments the conventional method of Edm_{true} reconstruction is a good one.

7 Exploration Strategy

To reach higher precision of Edm_{true} reconstruction in comparison with the conventional method, it is necessary to utilize additional information on hadron showers of events. In the computational experiments described below we applied neural nets of various structures with ever growing dimension of NN input vector. Input vectors were constructed as subsets of items from the "Pool of variables" described above. Our strategy was like this:

1. First we perform detailed exploration of the problem at one fixed value of beam energy (250 GeV).
2. Performing the first group of experiments we restricted ourselves to using only data on 2 samplings — LAr_3 and $Tile_1$. From one experiment to another we step by step added info on energy distribution among cells of these 2 samplings.
3. Performing the second group of subsequent experiments we step by step increased dimension of the NN input vector by adding info on energies of other samplings.
4. In the third group of experiments we explored several NNs which use only sampling energies which represent the longitudinal profile of a shower.
5. Finally the best NN versions were applied to events of some other beam energies (10, 50 and 350 GeV).

In total about 100 versions of NN procedures were tested in the present exploration. A group of most informative NN versions is considered in the current paper.

8 First group of experiments

Performing experiments of this group we restricted ourselves to using only data on 2 samplings – LAr_3 and $Tile_1$. Apart from sampling energies E_{LAr_3} and E_{Tile_1} we have at our disposal (ref. "Pool of Variables for ANN construction") additional data which relates to cell energies distributions in LAr_3 and $Tile_1$.

For convenience' sake the cell energies data from the "Pool" was grouped into 3 sets (Levels).

Level 1. In total **8 items** for NN input vector. Included are:

LAr_3 cell energies info

- The largest 4 cell energies in LAr_3 sampling for the current event.

$Tile_1$ cell energies info

- The largest 4 cell energies in $Tile_1$ sampling for the current event.

Level 2. In total **10 items** for NN input vector Included are:

LAr_3 cell energies info

- The largest 4 cell energies in LAr_3 sampling for the current event.
- The number of cells in LAr_3 with $E_{cell} > 3\sigma_{noise, LAr_3}$ for the current event.

$Tile_1$ cell energies info

- The largest 4 cell energies in $Tile_1$ sampling for the current event.
- The number of cells in $Tile_1$ with $E_{cell} > 3\sigma_{noise, Tile_1}$ for the current event.

Level 3. In total **12 items** for NN input vector Included are:

LAr_3 cell energies info

- The largest 4 cell energies in LAr_3 sampling for the current event.
- The number of cells in LAr_3 with $E_{cell} > 3\sigma_{noise, LAr_3}$ for the current event.
- Energy spread factor in LAr_3 cell energies distribution for the current event.

$Tile_1$ cell energies info

- The largest 4 cell energies in $Tile_1$ sampling for the current event.
- The number of cells in $Tile_1$ with $E_{cell} > 3\sigma_{noise, Tile_1}$ for the current event.
- Energy spread factor in $Tile_1$ cell energies distribution for the current event.

Among results obtained in the 1-st group of experiments we present here those obtained with NN 18, NN 19A and NN 29 networks.

8.1 NN 18 at 250 GeV

In the input vector of NN 18 the cell information is represented by the largest 4 cell energies in each sampling of the central pair (LAr_3 and $Tile_1$).

NN 18 structure: 10 – 20 – 20 – 1.

NN 18 Inputs: 1. ELAr3,
2. ETile1,
3 – 10. Level 1 info on cell energies distribution.

Activation function: $g(x) = \tanh(x)$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 2000 epochs.

NN 18 results (see Fig. 3):

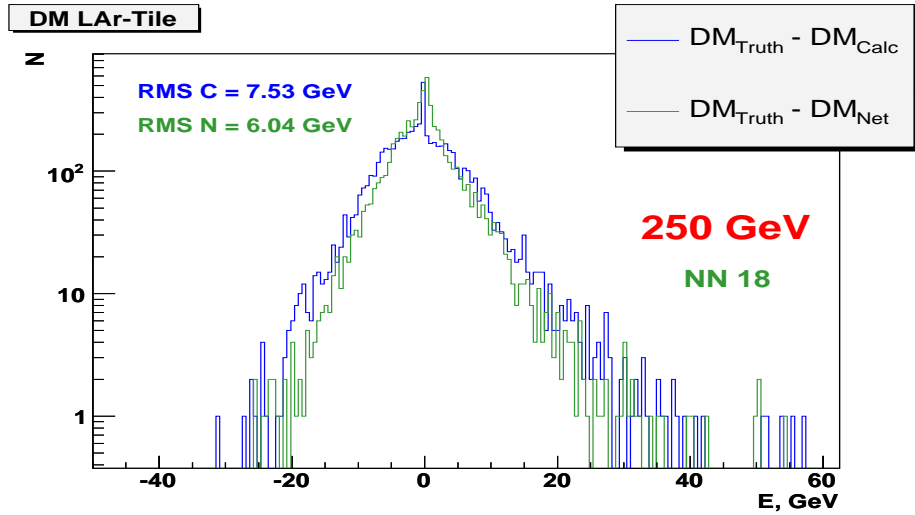


Figure 3: Distribution of Edm reconstruction error for NN 18 approximator at 250 GeV.

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{net}}) = 6.04 \text{ GeV}$$

Compared to the conventional method of Edm_{true} reconstruction NN 18 yields the following gain in precision:

$$\text{Absolute gain} = 1.49 \text{ GeV},$$

$$\text{Relative gain} = 19.7\%.$$

8.2 NN 19A at 250 GeV

For NN 19A the input vector is further augmented by adding info on the number of active cells in LAr_3 and $Tile_1$ for the current event.

NN 19A structure: 12 – 24 – 20 – 1.

NN 19A Inputs: 1. ELAr3,
2. ETile1,
3 – 12. Level 2 info on cell energies distribution.

Activation function: $g(x) = \tanh(x)$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 2000 epochs.

NN 19A results (see Fig. 4):

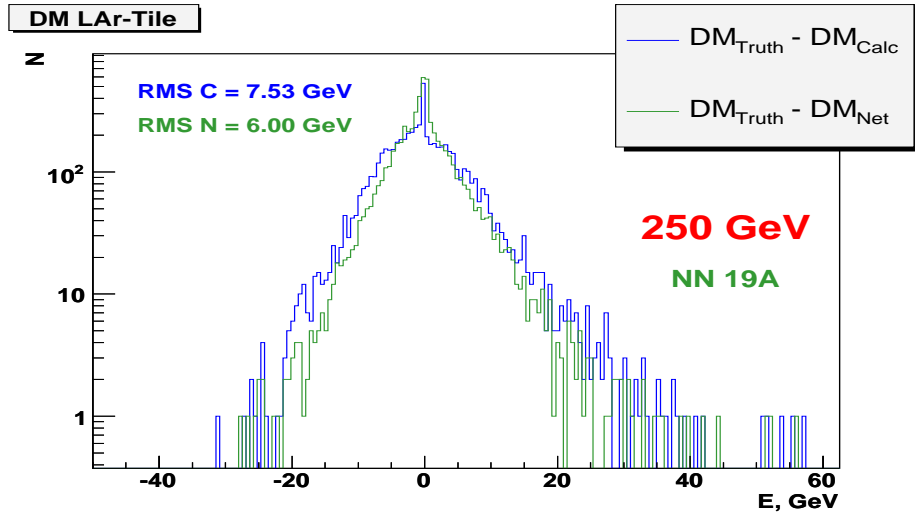


Figure 4: Distribution of Edm reconstruction error for NN 19A approximator at 250 GeV.

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{net}}) = 6.00 \text{ GeV}.$$

Compared to the conventional method of Edm_{true} reconstruction NN 19A yields the following gain in precision:

$$\text{Absolute gain} = 1.53 \text{ GeV},$$

$$\text{Relative gain} = 20.3\%.$$

8.3 NN 29 at 250 GeV

For NN 29 the input vector is further augmented by adding values of cell energies spread factors (in LAr_3 and $Tile_1$) for the current event.

NN 29 structure: 14 – 28 – 20 – 1.

NN 29 Inputs: 1. ELAr3,
2. ETile1,
3 – 14. Level 3 info on cell energies distribution.

Activation function: $g(x) = \tanh(x)$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 2500 epochs.

NN 29 results (see Fig. 5):

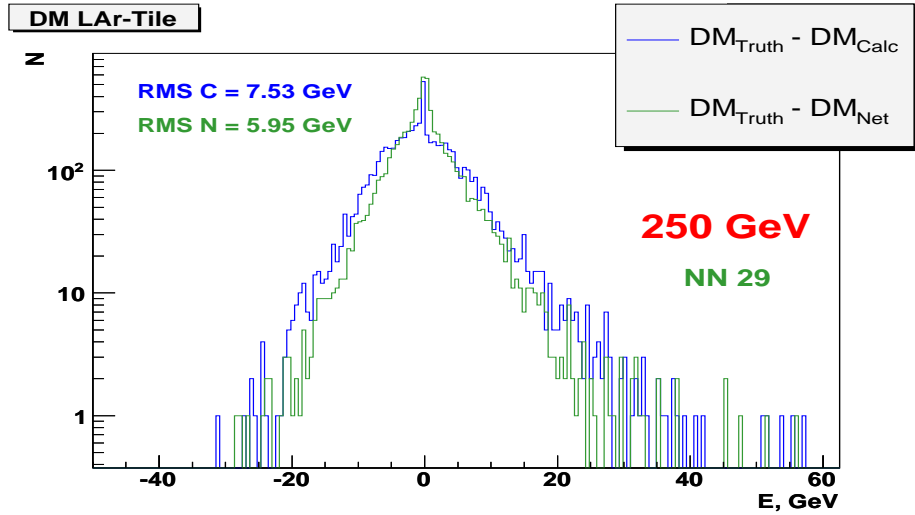


Figure 5: Distribution of Edm reconstruction error for NN 29 approximator at 250 GeV.

$$\mathbf{RMS(Edm_{\text{true}} - Edm_{\text{net}}) = 5.95 \text{ GeV.}}$$

Compared to the conventional method of Edm_{true} reconstruction NN 29 yields the following gain in precision:

$$\mathbf{Absolute \ gain = 1.58 \ GeV,}$$

$$\mathbf{Relative \ gain = 21.0\%.$$

We shall see later that this 21.0% gain in precision is half the maximum gain attained with the use of artificial neural networks in the course of the present investigation.

9 Second group of experiments

In the just considered first group of computational experiments the input vectors were constructed of data retrieved from calorimeter response in 2 central samplings (LAr_3 and $Tile_1$). Now in the following series of 4 experiments the input vector will be gradually augmented by adding energies of other samplings of the ATLAS calorimeter. These 4 experiments were performed with the use of NN 21, NN 20, NN 22 and NN 24 networks.

9.1 NN 21 at 250 GeV

The input vector of this network is augmented by ETile2 sampling energy.

NN 21 structure: 13 – 26 – 20 – 1.

- NN 21 Inputs:
1. ELAr3,
 2. ETile1,
 3. Level 2 info on cell energies distributions in LAr_3 and $Tile_1$,
 4. ETile2.

Activation function: $g(x) = \tanh(x)$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 2500 epochs.

NN 21 results (see Fig. 6):

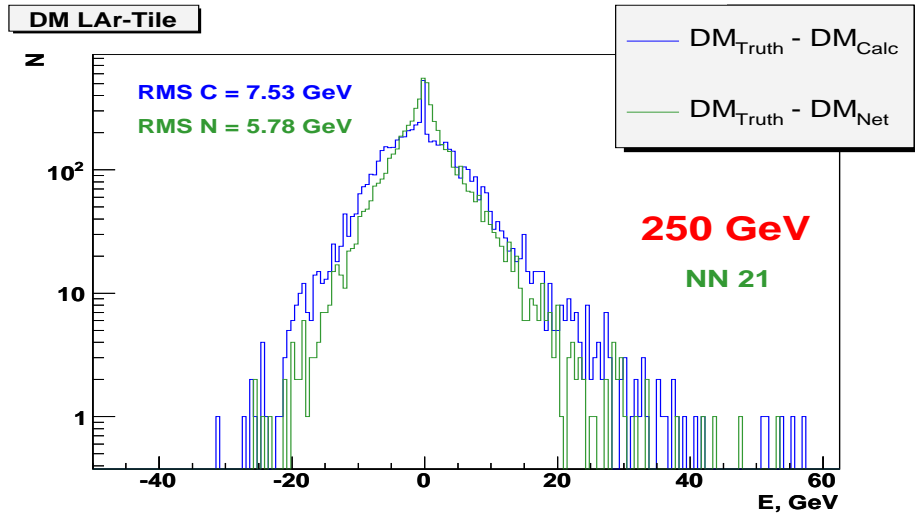


Figure 6: Distribution of Edm reconstruction error for NN 21 approximator at 250 GeV.

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{net}}) = 5.78 \text{ GeV}.$$

Compared to the conventional method of Edm_{true} reconstruction NN 21 yields the following gain in precision:

Absolute gain = 1.75 GeV,

Relative gain = 23.2%.

9.2 NN 20 at 250 GeV

The input vector of this network is augmented by ELAr2 sampling energy.

NN 20 structure: 13 – 26 – 20 – 1.

- NN 20 Input:
1. ELAr3,
 2. ETile1,
 3. Level 2 info on cell energies distributions in LAr_3 and $Tile_1$,
 4. ELAr2.

Activation function: $g(x) = \tanh(x)$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 2500 epochs.

NN 20 results (see Fig. 7):

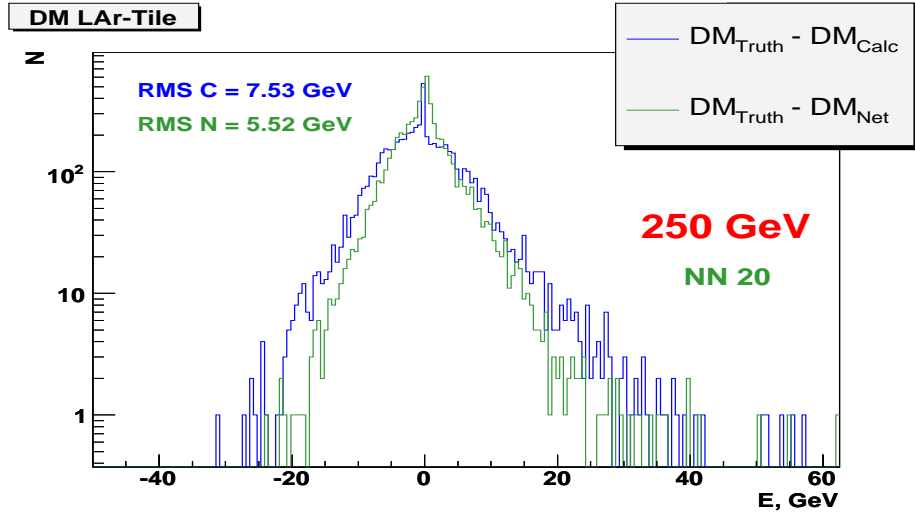


Figure 7: Distribution of Edm reconstruction error for NN 20 approximator at 250 GeV.

$$\mathbf{RMS(Edm_{true} - Edm_{net}) = 5.52 GeV.}$$

Compared to the conventional method of Edm_{true} reconstruction NN 20 yields the following gain in precision:

Absolute gain = 2.01 GeV,

Relative gain = 27.0%.

9.3 NN 22 at 250 GeV

The input vector of this network is augmented by ELAr2 and ETile2 sampling energies. The input vector includes:

1. items on cell energies in LAr_3 , $Tile_1$ samplings and
2. info on the central fragment of longitudinal profile in terms of 4 sampling energies.

NN 22 structure: 14 – 26 – 20 – 1.

1. ELAr3,
2. ETile1,

- NN 22 Inputs:
3. Level 2 info on cell energies distributions in LAr_3 and $Tile_1$,
 4. ELAr2.
 5. ETile2.

Activation function: $g(x) = \tanh(x)$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 2500 epochs.

NN 22 results (see Fig. 8):

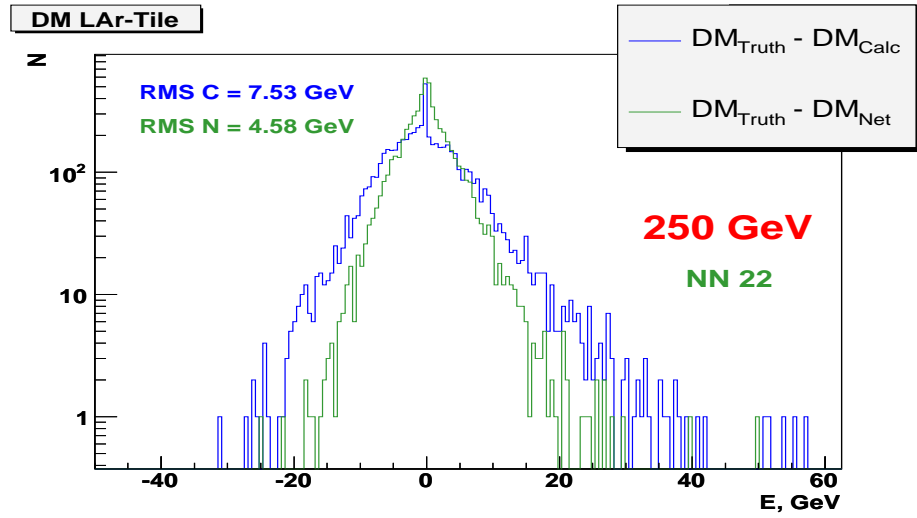


Figure 8: Distribution of Edm reconstruction error for NN 22 approximator at 250 GeV.

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{net}}) = 4.58 \text{ GeV}.$$

Compared to the conventional method of Edm_{true} reconstruction NN 22 yields the following gain in precision:

$$\text{Absolute gain} = 2.95 \text{ GeV},$$

$$\text{Relative gain} = 39.2\%.$$

9.4 NN 24 at 250 GeV

Information on full longitudinal profile of a shower (in terms of 6 sampling energies) together with cell information from LAr_3 , $Tile_1$ is used as input for this neural network procedure.

NN 24 structure: 16 – 28 – 20 – 1.

NN 24 Inputs:

- 1 – 3. ELAr1, ELAr2, ELAr3,
- 4 – 6. ETile1, ETile2, ETile3,
- 7 – 16. Level 2 info on the cell energies distributions in LAr_3 and $Tile_1$.

Activation function: $g(x) = \tanh(x)$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 2500 epochs.

NN 24 results (see Fig. 9):

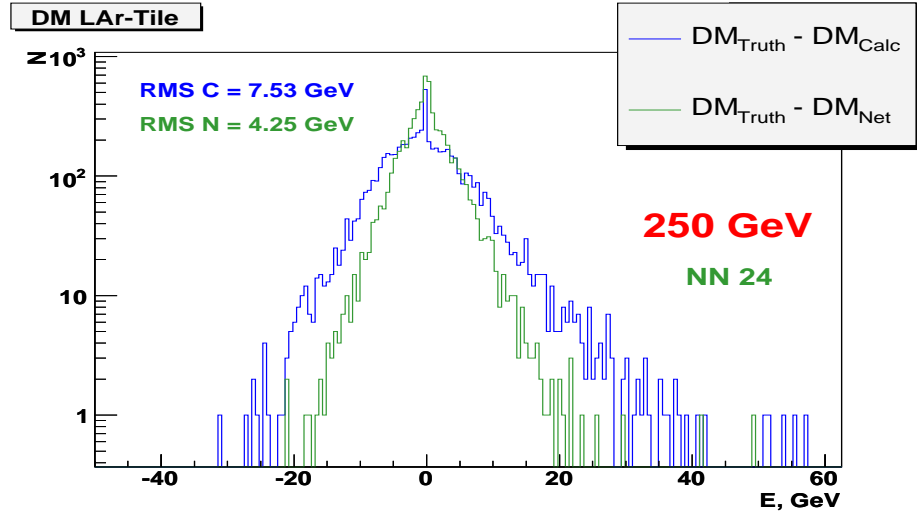


Figure 9: Distribution of Edm reconstruction error for NN 24 approximator at 250 GeV.

$$\mathbf{RMS}(\mathbf{Edm}_{\text{true}} - \mathbf{Edm}_{\text{net}}) = \mathbf{4.25 GeV}.$$

Compared to the conventional method of Edm_{true} reconstruction NN 24 yields the following gain in precision:

$$\mathbf{Absolute\ gain = 3.28 GeV},$$

$$\mathbf{Relative\ gain = 43.6\%}.$$

10 Third group of experiments

In this group of experiments **only sampling energies** of LAr and Tile calorimeters were used as components of NN input vector. We present here results of 3 experiments realized with the use of NN 28B, NN 10 and NN 12 networks.

- **NN 28B** uses energies of 2 central samplings (ELAr3, ETile1) as its input vector.
- **NN 10** uses energies of 4 central samplings (ELAr2, ELAr3, ETile1, ETile2).
- **NN 12** uses full longitudinal profile of a hadronic shower in LAr and Tile calorimeters as its input vector (6 sampling energies: ELAr1, ELAr2, ELAr3, ETile1, ETile2, ETile3).

10.1 NN 28B at 250 GeV

NN 28B uses energies of 2 central samplings (ELAr3, ETile1). This network results were presented earlier in the section 6. We remind NN 28B results at 250 GeV:

$$\mathbf{Absolute\ gain = 0.31 GeV},$$

$$\mathbf{Relative\ gain = 4.1\%}.$$

10.2 NN 10 at 250 GeV

NN 10 structure: 4 – 10 – 10 – 1.

NN 10 Inputs: 1 – 2. ELAr2, ELAr3,
3 – 4. ETile1, ETile2.

Activation function: $g(x) = (1 + e^{-2x})^{-1}$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 5000 epochs.

NN 10 results (see Fig. 10):

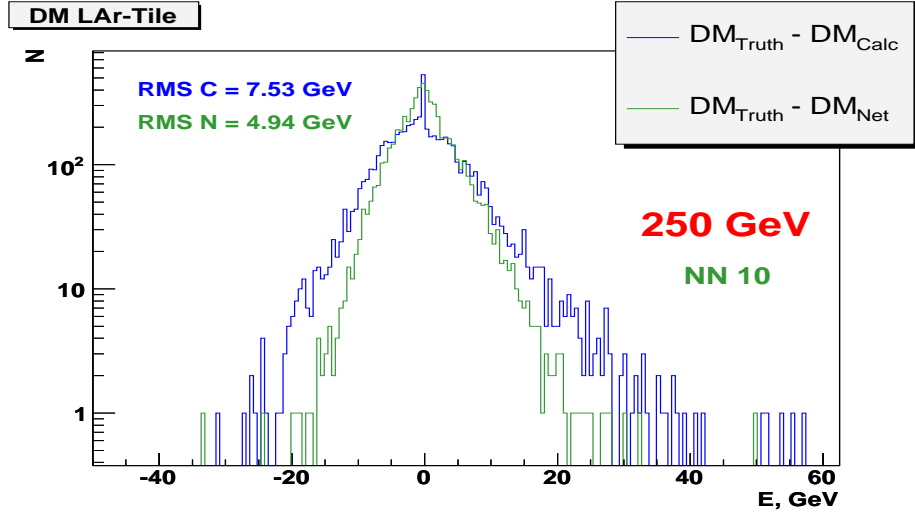


Figure 10: Distribution of Edm reconstruction error for NN 10 approximator at 250 GeV.

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{net}}) = 4.94 \text{ GeV}.$$

Compared to the conventional method of Edm_{true} reconstruction NN 10 yields the following gain in precision:

$$\text{Absolute gain} = 2.59 \text{ GeV},$$

$$\text{Relative gain} = 34.3\%.$$

10.3 NN 12 at 250 GeV

NN 12 structure: 6 – 15 – 15 – 1.

NN 12 Inputs: 1 – 3. ELAr1, ELAr2, ELAr3,
4 – 6. ETile1, ETile2, ETile3.

Activation function: $g(x) = (1 + e^{-2x})^{-1}$.

Files: $NTRAIN = 19159$ events, $NTEST = 6000$ events.

Training length – 3000 epochs.

NN 12 results (see Fig. 11):

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{net}}) = 4.65 \text{ GeV}.$$

Compared to the conventional method of Edm_{true} reconstruction NN 10 yields the following gain in precision:

$$\text{Absolute gain} = 2.85 \text{ GeV},$$

$$\text{Relative gain} = 37.8\%.$$

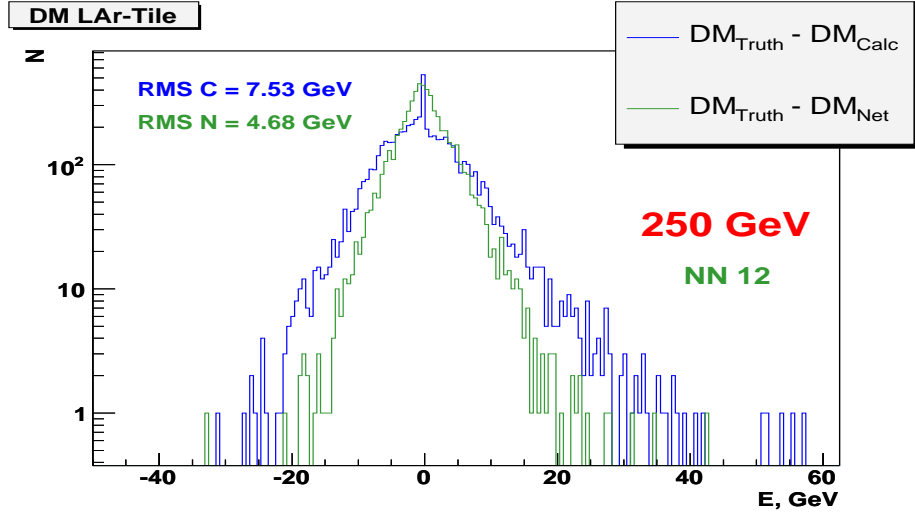


Figure 11: Distribution of Edm reconstruction error for NN 12 approximator at 250 GeV.

11 Edm reconstruction for 10 GeV pions

The 20000 MC events were generated for 10 GeV pions. The conventional method of Edm_{true} reconstruction resulted into:

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{calc}}) = 1.12 \text{ GeV}.$$

The precision attainable at 10 GeV with the use of neural networks is manifested by NN 8 network.

11.1 NN 8 at 10 GeV

NN 8 structure: 11 – 11 – 11 – 1.

- 1 – 3. ELAr1, ELAr2, ELAr3
- 4 – 6. ETile1, ETile2, ETile3
- 7 – 11. Partial info on the cell energies spectrum in LAr_3 and $Tile_1$:

NN 8 Inputs:

1. Energies of 2 leading cells in LAr_3 ,
2. Number of active cells in LAr_3 ,
3. Number of active cells in $Tile_1$,
4. energy spread factor in $Tile_1$.

Activation function: $g(x) = (1 + e^{-2x})^{-1}$.

Files: $NTRAIN = 14000$ events, $NTEST = 6000$ events.

Training length – 2000 epochs.
 NN 8 results (see Fig. 12):

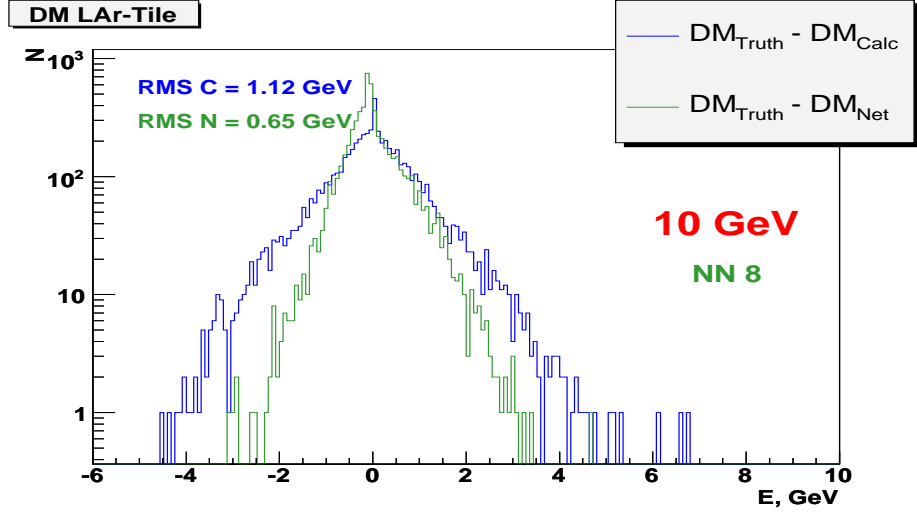


Figure 12: Distribution of Edm reconstruction error for NN 8 approximator at 10 GeV.

$$\mathbf{RMS(Edm_{true} - Edm_{net}) = 0.65 \text{ GeV.}}$$

Compared to the conventional method of Edm_{true} reconstruction NN 8 yields the following gain in precision:

$$\mathbf{Absolute \ gain = 0.47 \ GeV,}$$

$$\mathbf{Relative \ gain = 42.0\%}$$

11.2 NN 12 at 10 GeV

NN 12 structure: 6 – 11 – 11 – 1.

NN 12 Inputs: 1 – 3. ELAr1, ELAr2, ELAr3
 4 – 6. ETile1, ETile2, ETile3

Activation function: $g(x) = (1 + e^{-2x})^{-1}$.

Files: $NTRAIN = 14000$ events, $NTEST = 6000$ events.

Training length – 2000 epochs.

NN 12 results at 10 GeV (see Fig. 13):

$$\mathbf{RMS(Edm_{true} - Edm_{net}) = 0.70 \ GeV.}$$

Absolute gain = 0.42 GeV,
 Relative gain = 37.5%.

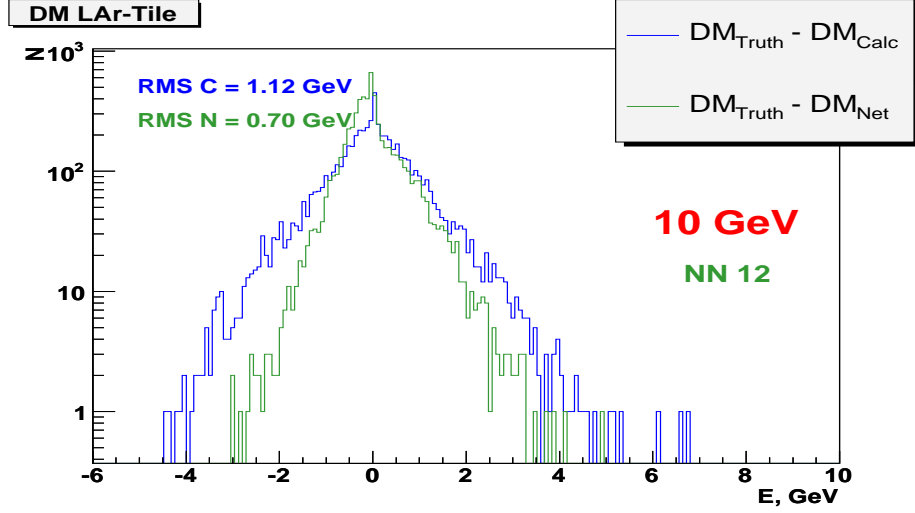


Figure 13: Distribution of Edm reconstruction error for NN 12 approximator at 10 GeV.

12 Edm reconstruction for 50 GeV pions

NN 12 at 50 GeV

The conventional procedure at 50 GeV has

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{calc}}) = 2.62 \text{ GeV}.$$

NN 12 structure: 6 – 15 – 15 – 1.

NN 12 Inputs: 1 – 3. ELAr1, ELAr2, ELAr3
 4 – 6. ETile1, ETile2, ETile3

Activation function: $g(x) = (1 + e^{-2x})^{-1}$.

Files: $NTRAIN = 14000$ events, $NTEST = 6000$ events.

Training length – 3000 epochs.

NN 12 results at 50 GeV:

$$\text{RMS}(\text{Edm}_{\text{true}} - \text{Edm}_{\text{net}}) = 1.69 \text{ GeV}.$$

Absolute gain = 0.93 GeV,

Relative gain = 35.5%.

13 Edm reconstruction for 350 GeV pions

NN 12 at 350 GeV

The conventional procedure at 350 GeV has

$$\mathbf{RMS(Edm_{true} - Edm_{calc}) = 9.50 GeV.}$$

NN 12 structure: 6 – 15 – 15 – 1.

NN 12 Inputs: 1 – 3. ELAr1, ELAr2, ELAr3
4 – 6. ETile1, ETile2, ETile3

Activation function: $g(x) = (1 + e^{-2x})^{-1}$.

Files: *NTRAIN* = 16154 events, *NTEST* = 6000 events.

Training length – 6000 epochs.

NN 12 results at 350 GeV:

$$\mathbf{RMS(Edm_{true} - Edm_{net}) = 5.59 GeV.}$$

$$\mathbf{Absolute\ gain = 3.91 GeV,}$$

$$\mathbf{Relative\ gain = 41.2\%}.$$

14 Summary

The computational experiments performed with MC events (pions at 10, 50, 250 and 350 GeV, $\eta = 0.25$, $\varphi = 0$.) allowed to estimate precision of different procedures for reconstruction of energy losses in the dead material between barrel LAr and Tile calorimeters.

For 250 GeV the results may be summarized as follows:

1. In a general class of procedures that use LAr_3 and $Tile_1$ sampling energies as their arguments the conventional method of Edm reconstruction proved good precision. Its RMS in $(Edm_{true} - Edm_{conv})$ distribution is only slightly ($\sim 4\%$) greater than corresponding RMS of the neural network procedure with the same 2 inputs.
2. At 250 GeV RMS in $(Edm_{true} - Edm_{conv})$ distribution amounts to 7.53 GeV (3% of the beam energy).
3. It was shown that application of neural network procedures for Edm reconstruction may substantially reduce RMS of $(Edm_{true} - Edm_{rec})$ distribution.

4. Among *Edm* reconstruction procedures which use as their arguments 2 central sampling energies (ELAr3, ETile1) together with data on cell energies distributions in *LAr*₃ and *Tile*₁ the lowest RMS attained is 5.95 GeV. Compared to the conventional procedure the attained reduction of RMS amounts to 21%.
5. A level of 4.25 GeV in RMS may be reached by the neural network procedure which uses full longitudinal profile of a hadronic shower in LAr and Tile calorimeters together with data on cell energies distributions in *LAr*₃ and *Tile*₁. Compared to the conventional procedure the RMS of the neural network procedure is 43% less.
6. The use of only a central fragment (4 samplings) of longitudinal profile data of a hadronic shower in LAr and Tile calorimeters (sampling energies in *LAr*₂, *LAr*₃, *Tile*₁, *Tile*₂) allows to reach the RMS at a level of 4.95 GeV. Compared to the conventional procedure the attained reduction of RMS amounts to 34%.
7. Procedures that use only full longitudinal profile of a hadronic shower in LAr and Tile calorimeters (6 sampling energies without cells info) may reach the RMS at a level of 4.65 GeV. Compared to the conventional procedure the attained reduction of RMS amounts to 38%.

Results for other beam energies may be summarized as follows:

- 8 For the conventional procedure of *Edm* reconstruction at 10 GeV the RMS value of ($Edm_{true} - Edm_{conv}$) distribution amounts to 1.12 GeV (11.2% of the beam energy).
- 9 At 10 GeV a level of 0.65 GeV for RMS in ($Edm_{true} - Edm_{net}$) distribution may be reached by a neural network procedure which uses full longitudinal profile of a hadronic shower in LAr and Tile calorimeters together with data on cell energies distributions in *LAr*₃ and *Tile*₁. Compared to the conventional procedure the RMS of the neural network procedure is 42% less.
- 10 In a wide range of beam energies (10, 50, 250 and 350 GeV) neural network procedures which use only full longitudinal profile of a hadronic shower as their inputs reveal 35 – 41% reduction in *Edm* reconstruction error in comparison with the conventional procedure.

15 Conclusion

The computational experiments performed with MC events for CTB04 setup showed that application of ANN technique for reconstruction of energy losses in dead materials between barrel LAr and Tile calorimeters allows one to reach 40% reduction of the energy reconstruction error compared to the conventional procedure used in the ATLAS collaboration.

Contrary to initial expectations it was found that the use of information on longitudinal profile of a hadronic shower brings greater improvement in DM energy reconstruction accuracy than the use of cell energies information in LAr_3 and $Tile_1$ samplings.

Application of ANN technique for evaluation of energy losses in dead materials between barrel LAr and Tile calorimeters should increase the accuracy of pion energy reconstruction in the ATLAS calorimeter.

16 Acknowledgements

The authors are greatly indebted to ATLAS Collaboration for ATLAS software. The authors are thankful to Vladimir Vinogradov for very interesting and valuable discussions and important comments. We appreciate to Jose Manoel de Seixas for careful reading on this paper and useful questions and comments. The presented work was partly supported by ISTC Project No. G-1458.

References

- [1] M.Cobal et al., "Analysis results of the April 1996 combined test of the LArgon and TILECAL barrel calorimeter prototypes" CERN-ATL-TILECAL-98-168, 1998, CERN, Geneva, Switzerland.
- [2] Y.Kulchitsky, M.Kuzmin, V.Vinogradov, "Non-compensation of an Electromagnetic Compartment of a Combined Calorimeter", CERN-ATL-TILECAL-99-021, 1999, CERN, Geneva, Switzerland; JINR-E1-99-303,1999, JINR, Dubna, Russia.
- [3] S.Akhmadaliev et al., "Results from a new combined test of an electromagnetic liquid argon calorimeter with a hadronic scintillating-tile calorimeter", Nucl. Instr. and Meth. A 449 (2000) 461.
- [4] K.Hornik, M.Stinchcombe and H.White. "Multilayer feedforward networks are universal approximators", Neural Networks, 1989, vol. 2, p. 359-366.

- [5] K.Hornik, M.Stinchcombe and H.White. "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks", *Neural Networks*, 1990, vol. 3, p. 551-560.
- [6] Carsten Peterson, Thorsteinn Rögnvaldsson, Leif Lönnblad. "JETNET 3.0 — A versatile artificial neural network package", *Comput. Phys. Commun.* 81 (1994) 185-220.