СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

Дубна

E10-95-476

A.Astvatsaturov, J.Budagov, I.Chirikov-Zorin,
D.Pantea[1], A:Paplevka[2], V.Shigaev, S.Sushkov[3],
M.Bosman[4], M.Nessi[5]

IMPROVEMENT IN SEPARATION OF ISOLATED MUONS AND PIONS AT LOW p_T IN ATLAS HADRON CALORIMETER USING ARTIFICIAL NEURAL NETWORKS TECHNIQUE

[1]Institute of Atomic Physics, P.O.Box MG-6, Bucharest, Romania
[2]On leave from BSU, Minsk, Belarus
[3]On leave from GSU, Gomel, Belarus
[4]IFAE, Barselona, Spain
[5]CERN, Geneva, CH-1211, Switzerland

1995

СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

Дубна

E10-95-476

A.Astvatsaturov, J.Budagov, I.Chirikov-Zorin,
D.Pantea[1], A:Paplevka[2], V.Shigaev, S.Sushkov[3],
M.Bosman[4], M.Nessi[5]

IMPROVEMENT IN SEPARATION OF ISOLATED MUONS AND PIONS AT LOW p_T IN ATLAS HADRON CALORIMETER USING ARTIFICIAL NEURAL NETWORKS TECHNIQUE

[1]Institute of Atomic Physics, P.O.Box MG-6, Bucharest, Romania
[2]On leave from BSU, Minsk, Belarus
[3]On leave from GSU, Gomel, Belarus
[4]IFAE, Barselona, Spain
[5]CERN, Geneva, CH-1211, Switzerland

1995

1 Introduction

Artificial Neural Networks (ANN) have already found many applications in High Energy Physics [1]. Due to their inherent parallelism, robustness and good statistical properties the ANN are used both in off-line and on-line analysis.

The main goal of the present article is to show advantages of ANN approach in handling data from highly granulated hadron calorimeter (HC) in comparison with other techniques. To demonstrate the potential of ANN approach the task of isolated low p_T π/μ separation was chosen. In solving the task we can get an insight into what are the most relevant inputs to an effective neural network classifier (discriminator) and what are its performance limits.

Having its own value, the effective solution of isolated low p_T π/μ separation task may be considered as an auxiliary step towards tackling a more difficult problem – tagging b-jets with low p_T muons using HC information. Muons in the range $3 < p_T < 5$ GeV have a significant probability to be absorbed in the calorimeter and therefore they cannot be reliably registered by the muon detector. In [2] it was shown that identification of b-jets with muon's $p_T \geq 3$ GeV might increase the statistics of the observed events by a factor of 2.5 in searching for and measurement of CP violation in $B_d^0 \to J/\psi K_s^0$ channel with $J/\psi \to \mu^+\mu^-$ decay – the problem mentioned in the ATLAS Technical Proposal [3].

In our investigation we restricted ourselves to testing discrimination power of two types of discriminators: linear threshold discriminators (LTD) as used in [2] and neural net discriminators (NND) built using the package JETNET [4]. The present work is based on simulated data; two HC designs were considered – with 4 and 3 longitudinal samples. Distributions of deposited energies in each section of ATLAS calorimeter are shown in Fig.1.

2 Neural networks application scheme and simulation data

In what follows the muon events (calorimeter response to muons) will be refered to as signal events, and the pion events – as background events respectively.

Neural net discriminators being nonlinear nonparametric extensions of conventional classifiers exploit knowledge of joint probability distribution of different features of registered events. Approximation of joint probability distribution is attained through a procedure called neural net training on the basis of a training set of events (simulated or real). Under certain conditions neural net classifiers realize asymptotically optimal, Bayasian decision [5], [6].

To formulate the problem under study closer to identification of low p_T

muons in the b-jet context we do not consider information from electromagnetic calorimeter (EMC) nor from any track detectors, thus entirely relying on hadron calorimeter response data.

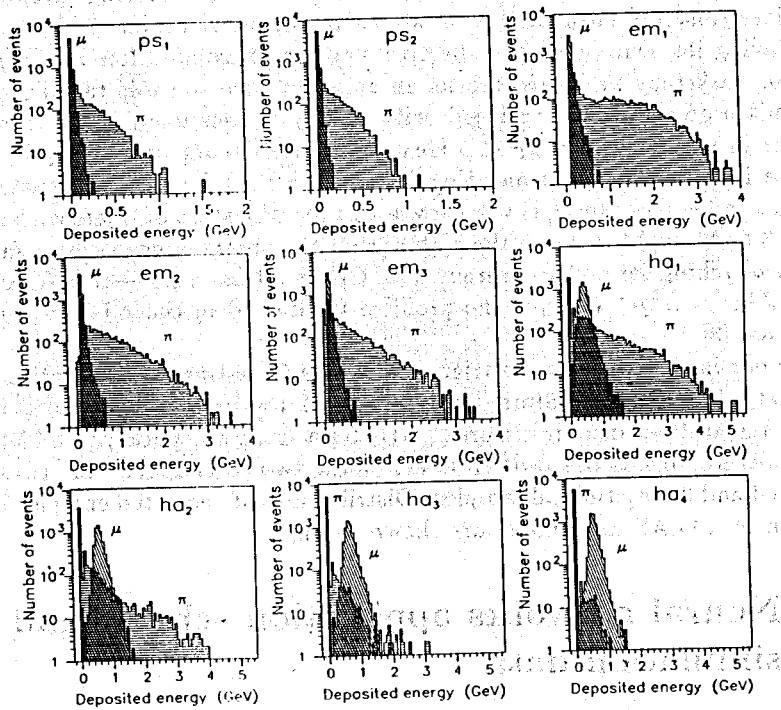

Figure 1: *Distributions of deposited energies in sections of preshower detector (ps_1, ps_2), EM calorimeter ($em_1 - em_3$), Hadron calorimeter ($ha_1 - ha_4$) for muons and pions at $\eta = 0.3$ and p_T values uniformly distributed within (3 - 5) GeV interval*

We adopted the following investigation scheme which consisted of eight distinct steps:

1. Define an interval of interest for p_T value (p_T working interval). In our case it is $3.0 < p_T < 5.0$ GeV.
2. Form training set of events. It comprises both signal and background events generated at p_T values within the p_T working interval (see below the details).
3. Train the neural net discriminator.
4. Test the neural net discriminator. Testing is performed using another set of events (test events) with p_T within the working interval. Quality of the trained discriminator (its characteristics as a classifier) is evaluated as a function of the threshold level applied to neural net output signal.
5. Estimate discriminator quality dependence on p_T (for events both within p_T working interval and outside it).
6. Execute steps 2 - 4 for different levels of photostatistics (in the range 10 - 80 photoelectrons per GeV).
7. Execute step 6 for different values of a cut applied for thresholding energy depositions in HC cells.
8. Execute steps 2 - 4 for two HC designs: a) with four longitudinal samples and b) with three longitudinal samples (samples 2 and 3 grouped togethger).

The standard ATLAS programs (DICE and ATRECON) were used to simulate calorimeter response to isolated μ and π at $\eta = 0.3$ for p_T values uniformly distributed within p_T working interval (3.0, 5.0) GeV. In total 6000 muon events and 6000 pion events were simulated. Actually the p_T working interval was subdivided into four nonoverlapping subintervals of 0.5 GeV width, with 3000 events in each. To evaluate the discriminator quality outside the p_T working interval, we have prepared additional data files for muon and pion events generated at $p_T = 2.0$ and 10.0 GeV (4000 events in total). Noise effects were taken into consideration in a simplified way using a cut of 0.1 GeV for thresholding the simulated energy depositions in HC cells.

The resultant trained neural net discriminator depends on p_T distribution in the training set within both classes of events (signal and background). General case of nonuniform p_T distributions is easily simulated by proper adjustments in a procedure that performs access to event patterns during neural net training phase.

3 Discriminators and their performance in low p_T π/μ separation

In this paper results for four models of π/μ discriminators are presented in the order of their increasing discrimination power.

- LTD – linear threshold discriminator that checks the deposited energy E_4 in the last HC sample against the threshold value.
- NND_{l1} – neural net discriminator operating on the energies E_i, $i = \overline{1,4}$, deposited in four HC sections (i.e. on all longitudinal samples).
- NND_{l2} – neural net discriminator operating on values of event features estimated as functions of arguments E_i.
- NND_{3d} – neural net discriminator operating on 3-dimensional pattern of energy deposition in HC (i.e. on energy deposition in cells).

Three-layered perceptrons with n input neurons (nodes) in the first layer, n_h neurons in a hidden layer and one output neuron in the third layer were selected for constructing neural net discriminators. Adjacent layers of the perceptrons are fully interconnected. A formula (n, n_h, 1) will be used to depict the structure of such perceptrons.

Inputs to the first layer of NND may be thought as components of n-dimensional vector that represents an event in n-dimensional feature space. Dimension n and ordering of input components are fixed for a particular NND. For neurons in the hidden and output layers the nonlinear neuron activation function $g(a) = (1 + \exp(-2a))^{-1}$ was chosen; hence the perceptrons perform nonlinear mappings of n-dimensional space into (0, 1) interval. During training phase the target value of the output neuron was put to 1 for muons and 0 for pions. Training procedure iteratively adjusts weigths of connections between neurons in order to minimize mean fit error MFE, i.e. mean squared deviation of actual net output values $O_{NN}(p)$ from the target values t(p) over the whole training set of events:

$$MFE = \frac{1}{2N_p}\sum_{p=1}^{N_p}(t(p) - O_{NN}(p))^2 \qquad (1)$$

where p denotes events.

Using a trained perceptron one gets one-dimensional distributions of net output values for muons and pions, and the subsequent part of π/μ separation task becomes similar to that of LTD discriminator which deals with one-dimensional distributions of E_4 values ($0 \le E_4 < \infty$).

In Fig.2(a) distributions of E_4 for signal (μ) and background (π) events are presented, and in Fig.2(b) – distributions of neural net output values for the same events (the neural net is that of NND_{l2}).

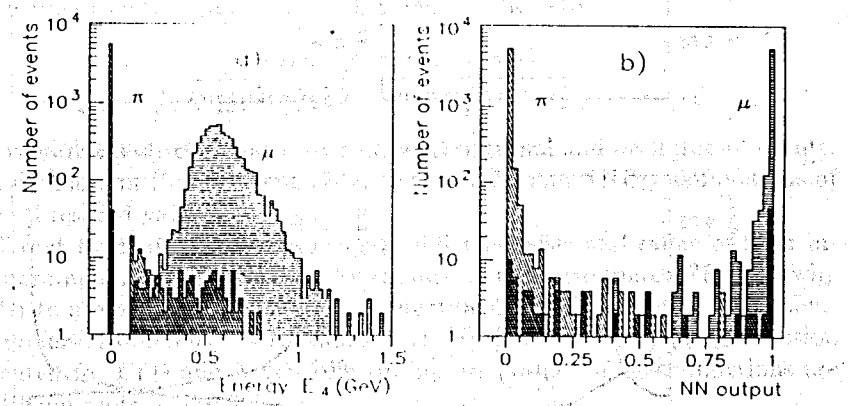

Figure 2: *(a) Distributions of E_4 for μ and π events at $\eta = 0.3$ and p_T values uniformly distributed within (3 - 5) GeV interval, (b) Distributions of NND_{l2} neural net output values for the same events*

A fixed point on x-axis (decision point or threshold) dichotomizes these distributions. Counting events on both sides of the threshold and normalizing the results one gets accumulated probabilities for an event to be correctly classified or misclassified. Applying variable thresholds we get estimates of important characteristics of discriminators:

- ε_μ – efficiency of signal events recognition, i.e. the probability that a muon event be correctly classified,
- α_μ – inefficiency of signal events recognition, i.e. the probability that a muon event be misclassified ($\alpha_\mu = 1 - \varepsilon_\mu$),
- β_π – survival probability for background events, i.e the probability that a pion event be misclassified.

These characteristics for LTD and NND_{l2} discriminators are presented in Fig.3 as functions of discriminator's internal parameter (threshold value for energy E_4 in case of LTD; threshold value for neural net output signal in case of NND).

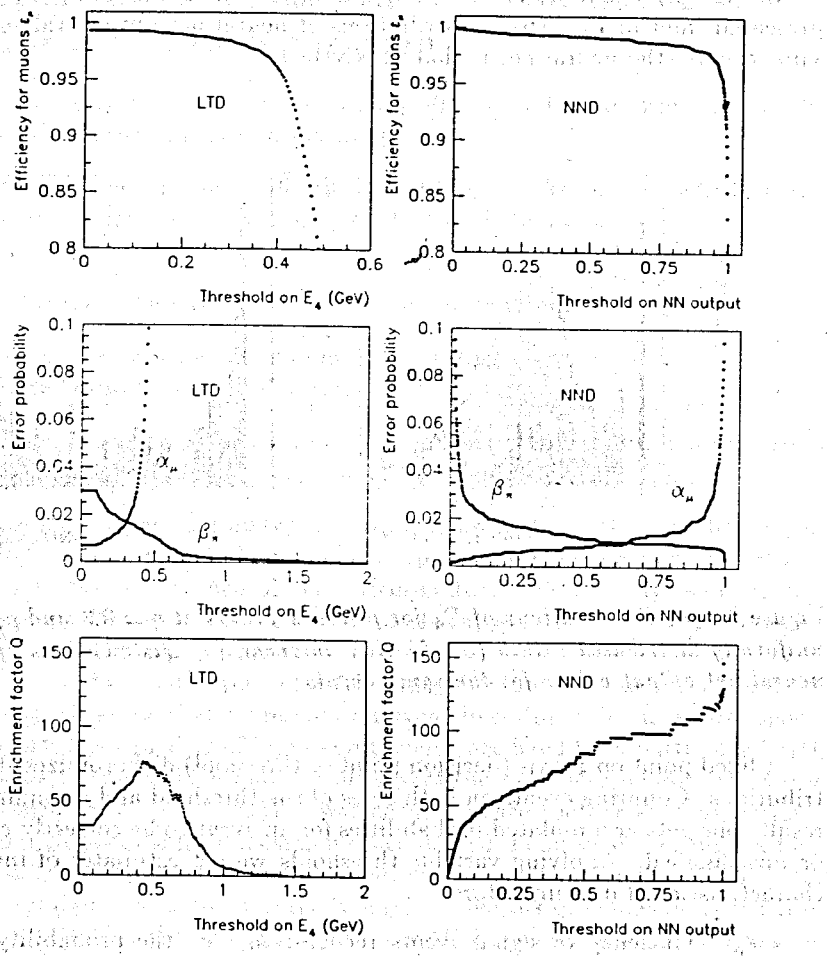

Figure 3: *Characteristics for LTD and NND_{l2} as functions of discriminator's internal parameter (threshold value for energy E_4 in case of LTD; threshold value for neural net output signal in case of NND)*

Two other characteristics are defined as follows:

- $R_\pi = 1/\beta_\pi$ – rejection factor for background events,
- $Q = \varepsilon_\mu \cdot R_\pi$ – enrichment factor.

Enrichment factor Q indicates the change in the ratio

$$\frac{(number\ of\ signal\ events)}{(number\ of\ background\ events)}$$

after applying the discriminator to a mixture of signal and background events. Two lower plots in Fig.3 present Q-factors for LTD and NND_{l2} as functions of variable threshold values.

Different types of discriminators may differ in sense and range of their internal parameters which control performance of a discriminator. That is why we prefer to use parameter independent function $Q(\varepsilon_\mu)$ for comparing functional behaviour of different discriminators [7], [8], [9]. In Fig.4 $Q(\varepsilon_\mu)$ function is presented for LTD and NND_{l2} discriminators (values of these functions are derived from plots in Fig.3).

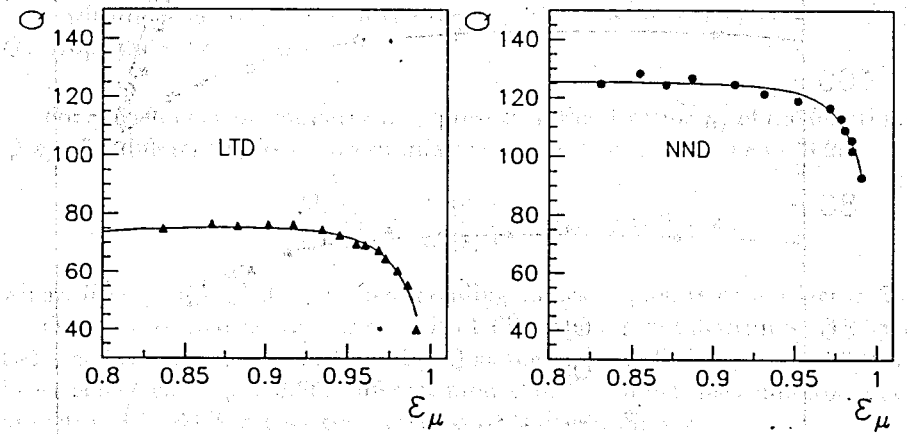

Figure 4: *Q vs ε_μ for LTD and NND_{l2}*

Before presenting and commenting functional behavior of the four discriminators we shall look at what is the difference between neural nets of three NND

discriminators. All three nets being perceptrons of the formula (n, 40, 1) differ in two respects:

1) dimension n of input vectors $\bar{X}_n$,
2) sense of components of $\bar{X}_n$.

Components of $\bar{X}_n$ vector are usually called event features. Features are functions of raw data items (cell energies of HC response in our case). It is worth noting that E_i samples are also features: each E_i is a weighted sum of all energies deposited in separate cells of i-th section of HC, all weights being set to 1). Evaluation of feature values is an operation called preprocessing of measurement data (or source data). Note that operation of reordering features in input vector $\bar{X}_n$ is another example of preprocessing if this operation is event dependent.

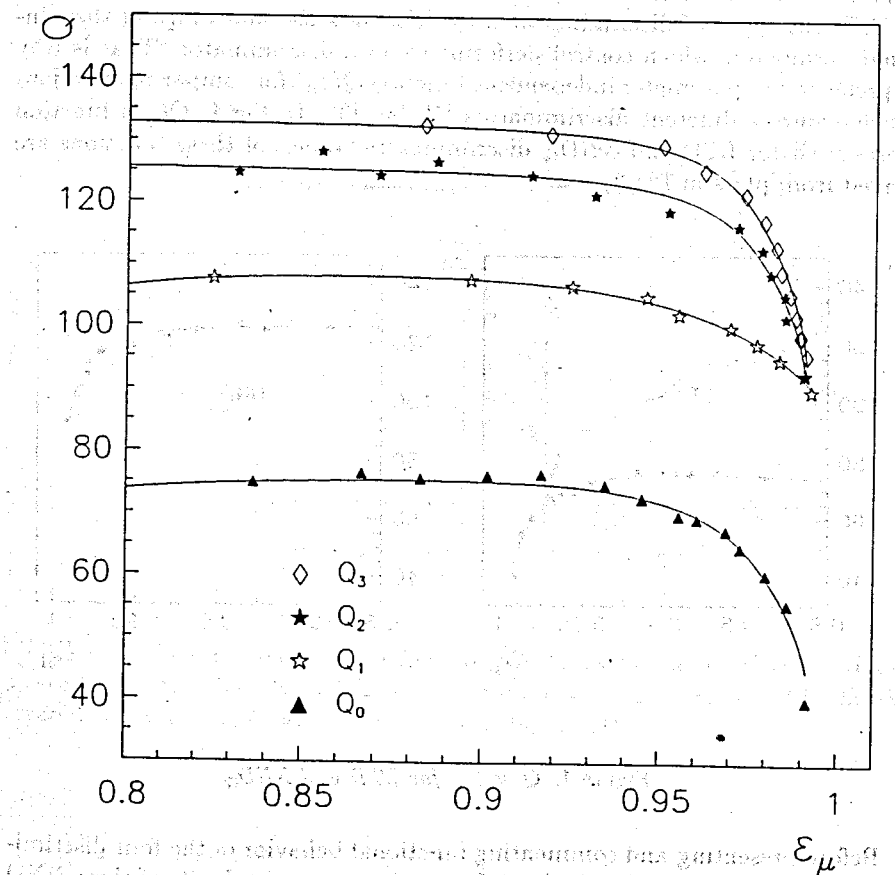

Figure 5: *Q vs ε_μ for discriminators: Q_0 – LTD, Q_1 – NND_{l1}, Q_2 – NND_{l2}, Q_3 – NND_{3d}*

It is well known that definition of a feature space is the most critical stage in pattern classification process. Various features of the events were evaluated and many sessions of neural net training and testing were carried out in search for the most effective subsets of features used as inputs to neural nets.

The neural net of NND_{l1} discriminator is a (4, 40, 1) - perceptron which uses four longitudinal samples E_i, $i = \overline{1,4}$, as components of input vector $\bar{X}_4$. The order of samples E_i in the vector $\bar{X}_4$ is fixed: i-th component of $\bar{X}_n$ is assigned E_i value. In Fig.5 the line labeled by Q_1 presents $Q(\varepsilon_\mu)$ curve for NND_{l1} discriminator. It follows from the figure that for muon registration efficiencies $\varepsilon_\mu = 0.80 - 0.97$ the enrichment factor Q_1 is in the range 100 – 105. $Q_1(\varepsilon_\mu)$ is a decreasing funcnion for larger ε_μ values, and at $\varepsilon_\mu = 0.99$ it drops to ~ 95.

NND_{l2} discriminator is also based on the (4, 40, 1) - perceptron and also uses four longitudinal samples E_i, $i = \overline{1,4}$. In contrast to NND_{l1}, assignment of a particular E_k to a component of $\bar{X}_4$ is dependent on the event itself. Here components of $\bar{X}_4$ are an ordered set of longitudinal samples ordered by their values in descending way. $Q(\varepsilon_\mu)$ curve for NND_{l2} discriminator is presented in Fig.5 by the line labeled Q_2. For muon registration efficiencies $\varepsilon_\mu = 0.80 - 0.97$ the enrichment factor Q_2 is in the range 120 – 125. At efficiency $\varepsilon_\mu = 0.99$ Q_2 drops to ~ 95.

For conveniency of comparison we present in Fig.6 ratios q_{ij} of the functions $Q(\varepsilon_\mu)$ for different pairs of discriminators at efficiencies $\varepsilon_\mu = 0.88 - 0.99$:

$$q_{ij}(\varepsilon_\mu) = \frac{Q_i(\varepsilon_\mu)}{Q_j(\varepsilon_\mu)} = \frac{R_i(\varepsilon_\mu)}{R_j(\varepsilon_\mu)}\ ,\ i > j\ ,\ j = 0, 1, 2\ ,$$

where $R_i(\varepsilon_\mu)$, $R_j(\varepsilon_\mu)$ denote corresponding values of pion rejection factor R_π.

One can see that in comparison with LTD all neural net discriminators have twice as high enrichment factor value Q at the highest efficiency $\varepsilon_\mu = 0.99$. At lower efficiencies ($\varepsilon_\mu < 0.96$) different models of neural net discriminators hold Q factors 40 - 80 % higher compared to LTD discriminator.

In search for effective NND_{3d} discriminator we tried a number a ways to extract additional important features by preprocessing clusters of cells in each HC section. A cluster is defined as the 3 x 3 cells window where the maximum summed energy is deposited. The central cell (η_c, φ_c) of a cluster is that with maximum cell energy.

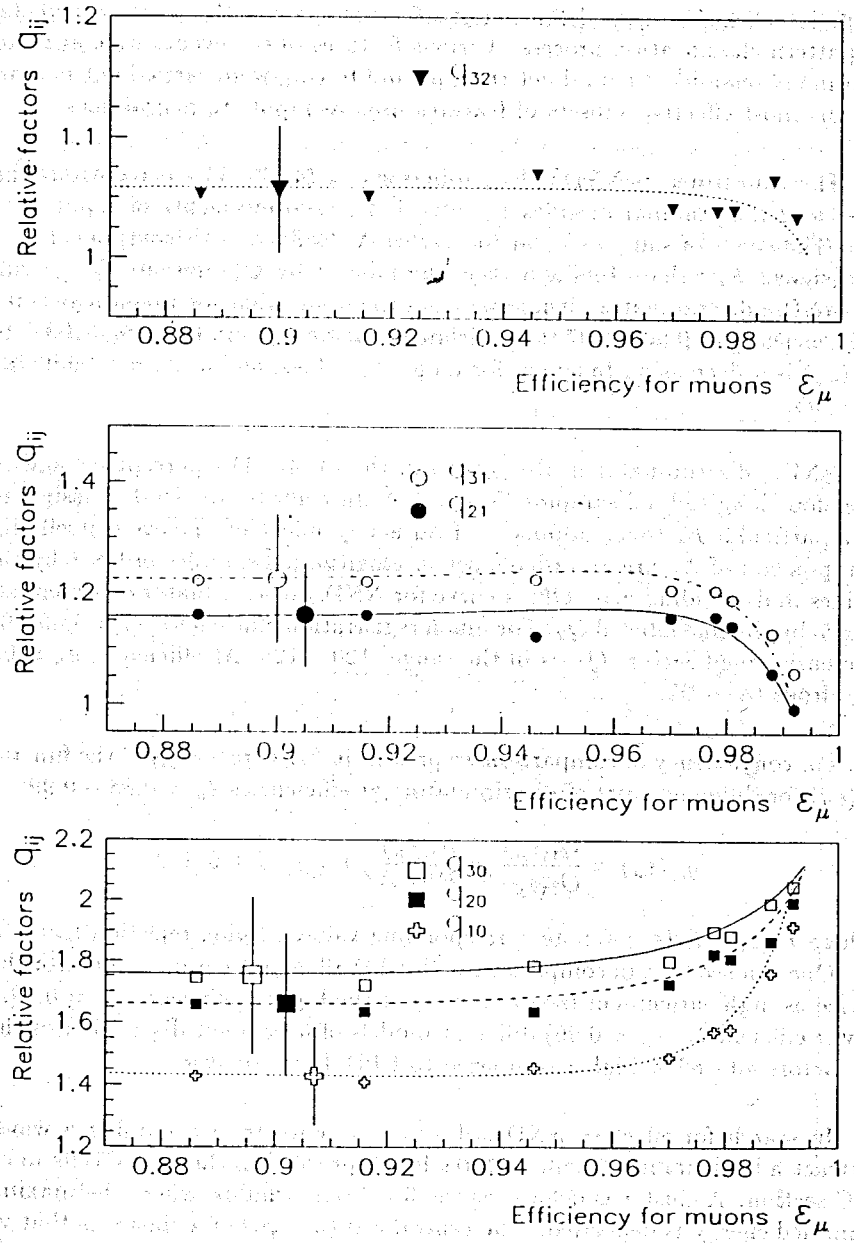

Figure 6: *Ratios $q_{ij}=Q_i/Q_j$ vs ε_μ for different pairs of discriminators D_i, D_j: D_0 - LTD, D_1 - NND_{l1}, D_2 - NND_{l2}, D_3 - NND_{3d}*

Some of the tested features are:

1. energy of the leading cell in a cluster,
2. summed energies in increasing square bands around the centre of a cluster,
3. summed energy of the cells outside a cluster,
4. ordered sample of cell energies in a cluster,
5. ordered sample of cell energies normalized by the total energy in a cluster,
6. m – number of cells with energy deposition above preset thresholds, i.e. multiplicity of active cells in a cluster,
7. m_η – number of active cells with $\eta_{cell} \neq \eta_c$ in a cluster, (η - multiplicity),
8. m_φ – number of active cells with $\varphi_{cell} \neq \varphi_c$, ($\varphi$ - multiplicity),

Additional features of "longitudinal" type were tested in order to take into account nonuniformity V_E of energy depositions in consequitive HC sections. V_E is defined as following:

$$V_E = \sum_{i=1}^{3}(v_i \cdot E_i - v_{i+1}E_{i+1})^2/E_{tot}^2\,,\quad E_{tot} = \sum_{1}^{4} E_i \qquad (2)$$

Three sets of v_i constants were used to prepare three versions of V_E feature:

a) $v_i = 1$,

b) $v_i = 1/d_i$, where d_i – thickness of i-th HC section in nuclear interaction length units,

c) $v_i = 1/M_\mu(E_i)$, where $M_\mu(E_i)$ is the mean value of i-th longitudinal sample in HC for muon events (see Fig.1)

During training and testing sessions we retained only those models of NND_{3d} which had higher characteristics and lower dimension of feature vector $\bar{X}_n$. The final version of NND_{3d} is based on the (8, 40, 1) - perceptron. The input features in $\bar{X}_8$ vector are:

— ordered sample of E_i (four features),

— $m_\eta^{(j_k)}$, $m_\varphi^{(j_k)}$, $k = 1, 2$ (four features),

j_1 , j_2 – are indeces of those two HC longitudinal sections where the greatest summed energies were deposited for an event.

Functional behavior of NND_{3d} is presented by $Q_3(\varepsilon_\mu)$ curve in Fig.5. One can see that an increase in Q - factor value is sensible enough (about 25 units)

compared to NND_{l1}, but is relatively small (less then 10 units) compared to NND_{l2}. In comparison with all other three discriminators the relative increase in Q for NND_{3d} is presented in Fig.6 by $q_{3j}(\varepsilon_\mu)$ curves, j = 0, 1, 2.

In our opinion, the moderate increase in Q for NND_{3d} in comparison with NND_{l2} may be justified as follows.

Muons loose their energy mainly by ionization, and the number of active cells in an HC section does not exceed 2. Distributions of their deposited summed energies in each of four sections are of gaussian type, centered at $M_\mu(E_i)$, $i = \overline{1,4}$ with standard deviations $0.102 < \sigma_i < 0.125$ (ref. Fig.1). Big deviations from mean values $M_\mu(E_i)$ in one or more HC sections are used by well trained NND_{l1}, NND_{l2} discriminators as signatures of a pion. Number of pions misclassified by NND_{l1}, NND_{l2} is not greate and equals to $m_\pi = N \cdot \beta_\pi = N/R_\pi$. To substantially increase classification power the NND_{3d} discriminator should correctly classify a part of m_π pions using information on cell distribution of the deposited energy in HC sections. The rise in multiplicity above 2 active cells is with high probability accompanied by the increase in summed energy deposition by an ammount that is abnormal to a muon event; meanwhile the observed multiplicity of active cells in the subset of m_π misclassified pion events is similar to that in muon events. The little difference in characteristics between NND_{3d} and NND_{l2} shows that using information on active cell multiplicity permits NND_{3d} to lower m_π number only by 5%. This result gives rise to an assumption that some of m_π pions – all exhibiting deep penetration ability with nonzero energy deposition in the last HC section – most likely did not take part in nuclear interactions at all. Obvious contradiction between the actually observed fraction of misclassified pions (~ 0.01) and the fraction of pions (< 0.0001) that could escape nuclear interactions in ATLAS calorimeter at $\eta = 0.3$ leads us to a conclusion that at least a part of the observed m_π cases of HC response is most probably not produced by particles entering HC as pions.

Indeed, pion decay process $\pi^\pm \to \mu^\pm + \bar{\nu}_\mu$ tends to make HC response to a background event (π) look like that to a signal event (μ). The probability of the decay is not negligible in our π/μ separation task: for p_T uniformly distributed in 3 – 5 GeV interval at $\eta = 0.3$ about 0.83% of pions decay prior to the first nuclear interaction and should in average produce muon-like HC responses. At high muon registration efficiencies and 3000 pion events in a test sample we arrive at a limit value $R_\pi = 120^{+30}_{-20}$. It follows from this estimate that longitudinal samples in HC contain enough information for NND_{l1} and NND_{l2} to approach the limit values of R_π and Q. Hence the subsequent improvements in R_π and Q attained by NND_{3d} could not be high.

Simulated events (12000 in total) have been split into two equal parts: one part used for training neural net and another part - for testing its generalization ability. With 3000 + 3000 events in the test sample and high values of background rejection factors R_π attained by discriminators ($R_\pi \sim 100$) the statistical errors in estimation of Q can not be low: $\sigma(Q) \approx 15 - 20$. Nevertheless, the difference in characteristics of any two discriminators may be estimated with higher precision because the common test sample of events is used for evaluating these characteristics which consequently become correlated.

Let D_i, D_j be two discriminators tuned to operate at the given fixed efficiency $\varepsilon_\mu = \varepsilon_0$, and R_i, Q_i, R_j, Q_j – their background rejection factors and enrichment factors at $\varepsilon_\mu = \varepsilon_0$. Assume without loss of generality that $R_i \geq R_j$. It can be shown that maximum likelihood estimation of variance of the ratio $q_{ij} = Q_i/Q_j = R_i/R_j$ may be reduced to the following expression:

$$\hat{var}(q_{ij}) = q_{ij} \cdot \frac{R_i}{N} \cdot \left(q_{ij} + 1 - 2\frac{R_i}{R_{ij}}\right) \tag{3}$$

or

$$\hat{var}(q_{ij}) = q_{ij} \cdot \frac{Q_i}{\varepsilon_0 N} \cdot \left(q_{ij} + 1 - 2\frac{Q_i}{Q_{ij}} \cdot \frac{\varepsilon_{ij}}{\varepsilon_0}\right) \tag{4}$$

where

N – the number of background events (pions) in the test sample,

R_{ij}, Q_{ij}, ε_{ij} – background rejection factor, enrichment factor and muon registration efficiency of the compound discriminator D_{ij} based on D_i, D_j that are operating in parallel (each at $\varepsilon_\mu = \varepsilon_0$) and whose output logical signals 0/1 (classification signals) are processed by "AND" logical function to form output signal of the compound discriminator.

Note that in general case the next inequalities hold:

$$R_{ij} \geq R_i \geq R_j \ , \ \ \varepsilon_{ij} \leq \varepsilon_0$$

Error bars in Fig.6 correspond to estimates according to (3), (4).

Neural nets were thoroughly trained using up to 7 - 10 thousands epochs in a training session. To reach lower event classification error we tested neural net versions with different forms of neuron activation function, used fixed and variable learning rates in a training session, used back-propagation and Rprop training procedures [4] and varied starting values of weigts and thresholds of the neural net when initializing a training sesion.

Output of a short summary after each epoch proved very useful for supervising the process of neural network training. The summary contains: epoch

number, mean fit error MFE (1) in the current epoch, muon and pion recognition efficiency in training and test sets of events, four values of enrichment factor Q at efficiencies $\varepsilon_\mu = 0.99, 0.95, 0.90, 0.85$.

As an example of a training session we present in Fig.7 the dynamics of enrichment factor Q (at $\varepsilon_\mu = 0.99$) and mean fit error MFE as functions of the current epoch number in the session.

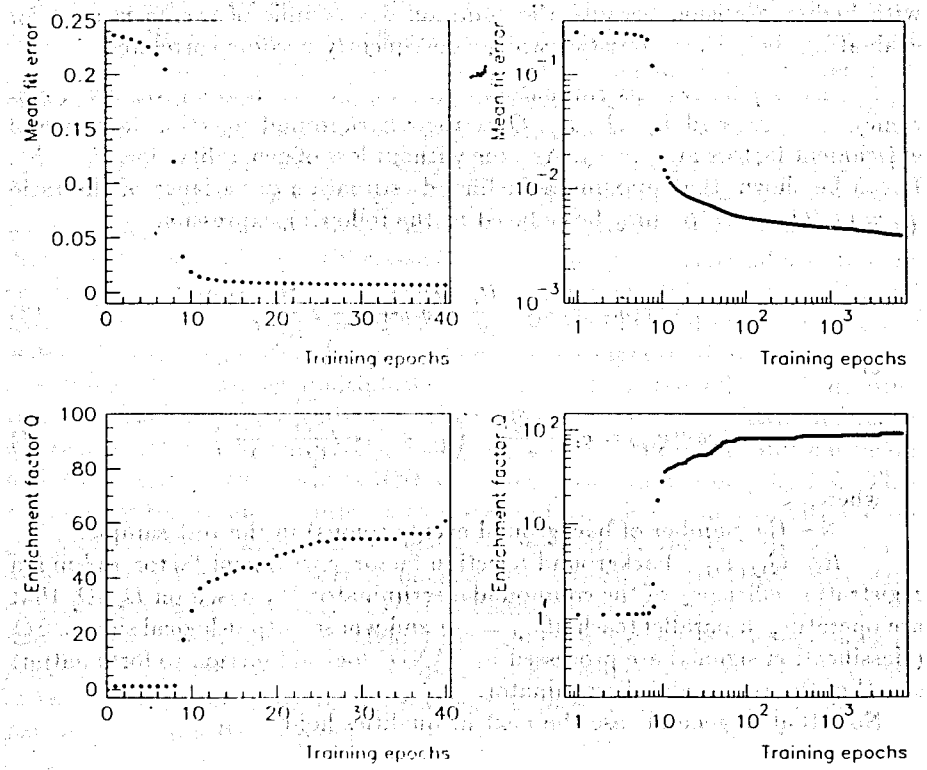

Figure 7: *Dynamics of enrichment factor Q (at $\varepsilon_\mu = 0.99$) and mean fit error MFE in a training session for NND_{l1}*

As the next step in our investigation scheme we estimated the dependence of discriminator characteristics on p_T values of events being tested. We evaluated characteristics of different discriminators at four p_T values inside the working interval (3.0 , 5.0) GeV and at p_T = 2 and 10 GeV outside it. It should be noted that results for p_T outside the working interval are highly sensitive to singularities of NND versions and to the threshold values applied to the neural net output signal. Efficiency ε_μ and pion rejection factor R_π inside the

working interval are constant within statistical errors for all discriminators. At p_T = 10 GeV ε_μ remains as high whereas at p_T = 2 GeV ε_μ drops to the values 0.2 – 0.8 depending on the version of a discriminator.

Due to poor statistics at p_T = 2 and 10 GeV (1000 pion events at each p_T value) only qualitative conclusions can be drawn from the set of estimates of R_π for different discriminators at these two p_T values. The least degradation in R_π and ε_μ outside the working interval is shown by discriminators of NND_{3d} family, the biggest degradation – by LTD. To reach good performance at p_T = 2 GeV one should include events simulated at $2 \leq p_T \leq 3$ GeV into the training set of events.

To investigate the influence of photostatistics on discriminator's performance we repeated steps 2 - 4 of our investigation scheme (see page 3) for seven different photostatistics levels (PSL) in the range 10 - 80 photoelectrons per GeV. For the fixed muon recognition efficiency $\varepsilon_\mu = 0.99$ the dependence of Q - factor upon photostatistics level is presented in Fig.8 for NND_{l2} and LTD discriminators.

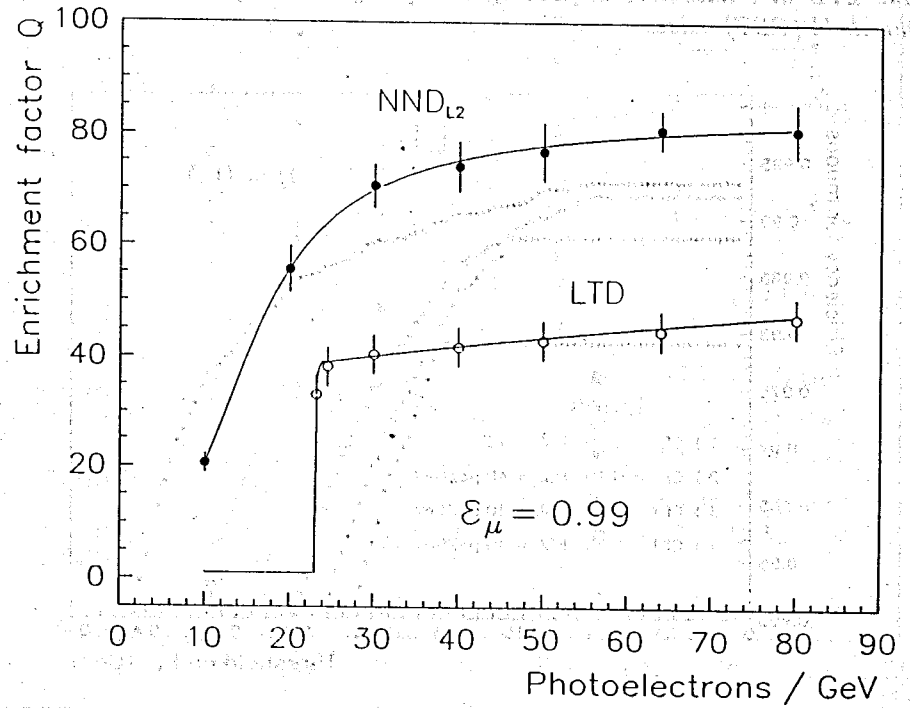

Figure 8: *Enrichment factor Q as function of photostatistics level for NND_{l2} and LTD discriminators at fixed efficiency $\varepsilon_\mu = 0.99$ and cell energy threshold CET =0.1 GeV*

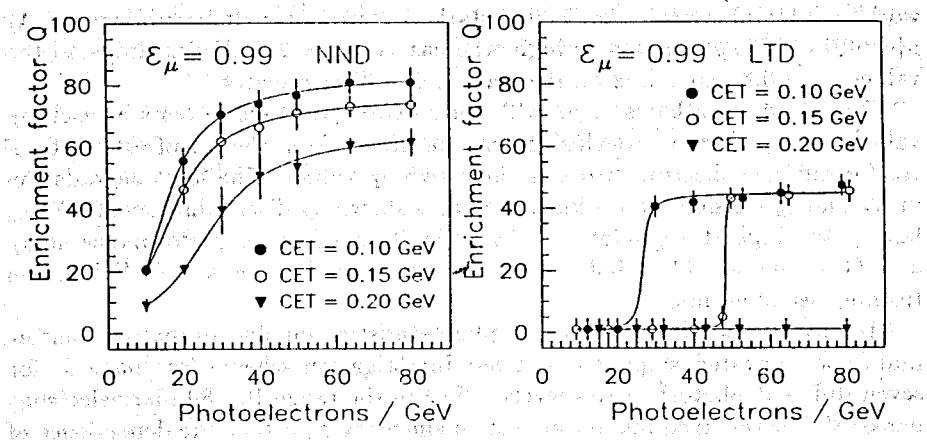

Figure 9: *Enrichment factor Q as function of photostatistics level for NND_{l2} and LTD discriminators at fixed efficiency ε_μ = 0.99 for three cell energy threshold (CET) values*

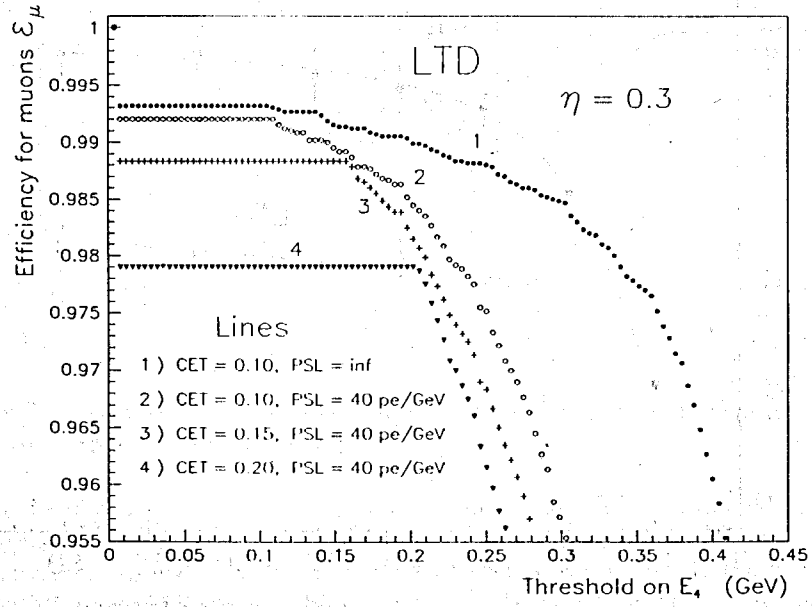

Figure 10: *Efficiency ε_μ versus threshold on E_4 for different values of CET (cell energy threshold) at fixed photostatistics level PSL = 40 photoelectrons per GeV*

It is clearly seen that the neural net discriminator is a more robust classifier which retains its selectivity within the whole range of photostatistics level and whose discrimination power gradually decreases when photostatistics level goes down. In contrast to NND_{l2}, the LTD discriminator cannot retain its selectivity at efficiency ε_μ = 0.99 within the whole range of photostatistics level (25 photoelectrons/GeV is the critical point – ref. Fig.8).

Both characteristics shown in Fig:8 correspond to a cut of 0.10 GeV applied for thresholding energy deposition in a separate HC cell. To examine sensitivity of the characteristics to cell energy threshold values, we have estimated another two pairs of characteristics (for the same ε_μ value) corresponding to cell energy thresholds (CET) of 0.15 and 0.20 GeV. All six curves are presented in Fig.9 separately for NND_{l2} and LTD discriminators. It is seen that for CET = 0.20 GeV the LTD discriminator cannot reach muon registration efficiency ε_μ = 0.99 at any value of photostatistics level without loosing π/μ separation ability.

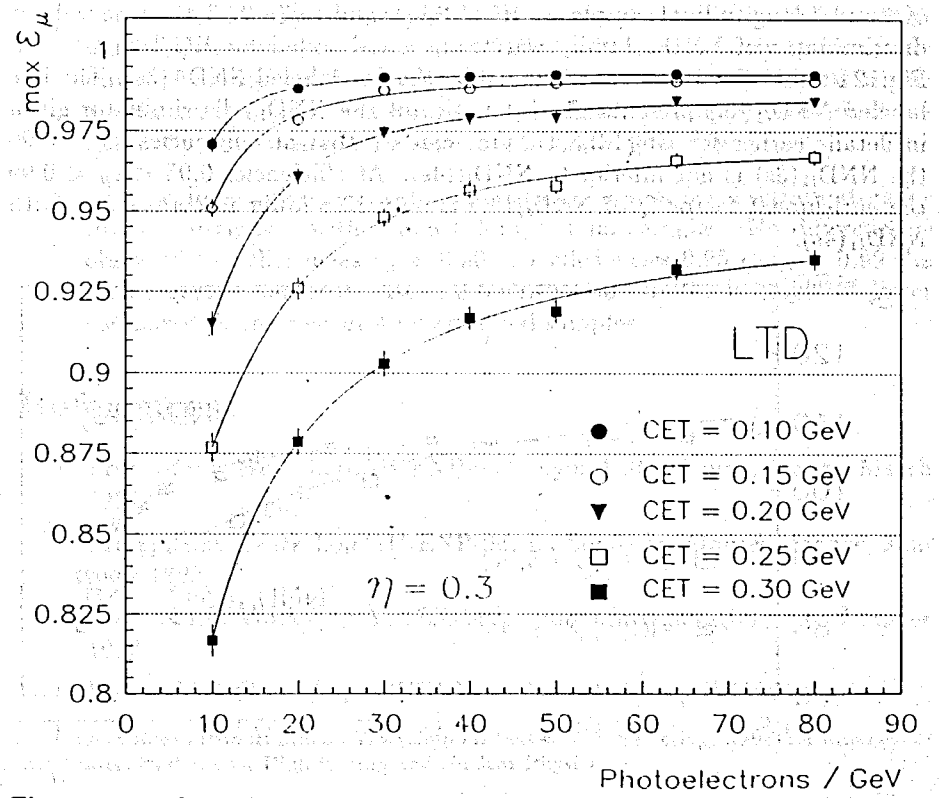

Figure 11: *The maximum efficiency ε_μ attainable without loss of π/μ separation ability as function of photostatistics level at different values of CET (cell energy threshold)*

The dependence of ε_μ on LTD's internal parameter (threshold on E_4 - ref. Fig.3) at the fixed PSL = 40 photoelectrons/GeV is presented in Fig.10 by the lines 2, 3, 4 for different values of CET = 0.10, 0.15, 0.20 GeV. The line 1 corresponds to HC data simulated without taking photostatistics into account (PSL = ∞).

To investigate sensitivity of LTD characteristics to cell energy thresholds, we have additionally evaluated the maximum values of ε_μ attainable by LTD without loosing π/μ separation ability for five CET values in the range 0.10 - 0.30 GeV and seven PSL values in the range 10 - 80 photoelectrons/GeV. The results are presented in Fig.11. They allow one to foresee limitations of LTD expected in solving π/μ separation task in a more realistic environment when higher CET values might be needed to suppress background signals.

According to the ATLAS Technical Proposal [3] the central two sections of hadron calorimeter will be grouped together. We designate the two HC designs of 4 and 3 longitudinal samples as HC(1,2,3,4) and HC(1,2+3,4). A neural net discriminator of NND_{l1} family was trained and tested for HC(1,2+3,4). In Fig.12 its characteristics are presented by the line labeled NND_{l1}(3s). The line labeled NND_{l1}(4s) presents characteristics of the NND_{l1} discriminator given in details earlier for HC(1,2,3,4). One can see that at efficiencies $\varepsilon_\mu < 0.90$ the NND_{l1}(3s) is not inferior to NND_{l1}(4s). At efficiencies $0.95 < \varepsilon_\mu < 0.99$ the enrichment factor Q of NND_{l1}(3s) is only 10% lower in comparison with NND_{l1}(4s).

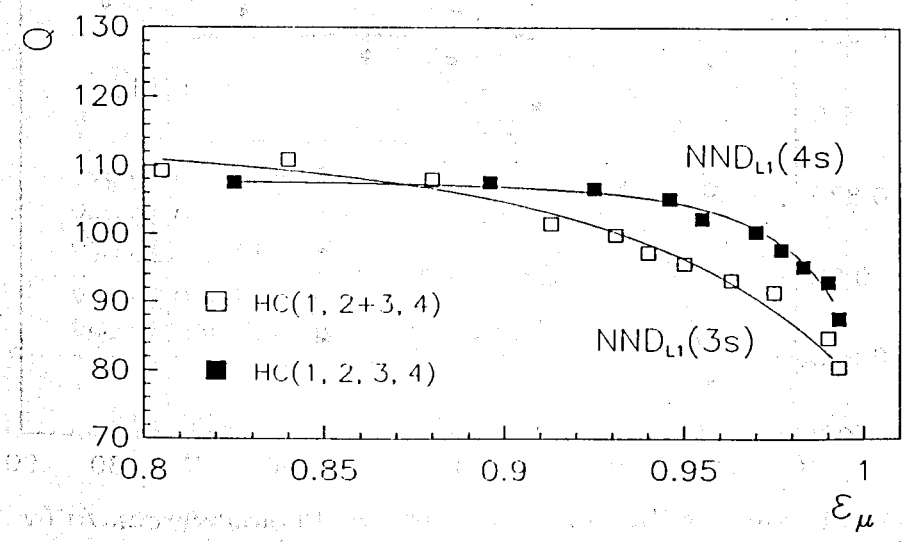

Figure 12: *Characteristics of neural net discriminators for two different HC designs – with 3 and 4 longitudinal samples*

4 Conclusions

1. Neural net discriminators operating on longitudinal or 3-dimensional deposited energy samples and the linear threshold discriminator operating on total deposited energy in the HC last section were applied to low p_T π/μ separation at $\eta = 0.3$ using MC simulated data. Compared to the linear threshold discriminator an increase of 80 – 100% in pion rejection factor at muon recognition efficiency 0.95 – 0.99 was obtained in case of neural network discriminators.

2. Neural net discriminators trained inside the working interval $3 \leq p_T \leq 5$ GeV do not show a sharp deterioration of their performance outside the working interval at $p_t = 10$ GeV. To keep good performance of neural net discriminators at $p_T = 2$ GeV one should include events with $2 \leq p_T \leq 3$ GeV into the training set of events.

3. Neural net discriminators proved to be robust classifiers that at high muon registration efficiency $\varepsilon_\mu = 0.99$ retain their selectivity in a wide range of photostatistics level (10 - 80 photoelectrons/GeV) and whose discrimination power – in contrast to the linear threshold discriminator – gradually decreases when photostatistics level goes down.

4. There is little difference in characteristics of neural net discriminators for two HC designs – with 4 and 3 longitudinal samples. No difference is observed for efficiencies $\varepsilon_\mu < 0.90$. At efficiencies $0.95 < \varepsilon_\mu < 0.99$ the pion regection factor in case of 3 longitudinal samples is only 10% lower compared to the case of 4 longitudinal samples.

References

[1] Proc. of Int. Workshop AIHENP-90 [1], Lyon Villeurbanne, France, March 1990

Proc. of Int. Workshop AIHENP-92, La Londe les Maures, France, January 1992

Proc. of Int. Workshop AIHENP-93, Oberammergau, Germany, October 1993

Proc. of Int. Workshop AIHENP-95, Pisa, Italy, April 1995

[1] Proceedings of International Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics

[2] J.Budagov, I.Chirikov-Zorin, A.Pantea, D.Pantea, O.Pukhov, M.Bosman, M.Nessi, "B-meson tagging improvement using HCAL information", ATLAS Internal Note, TILECAL-NO-023

[3] ATLAS, "Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN", CERN/LHCC/94-43, LHCC/P2 (1994)

[4] C.Peterson, T.Rögnvaldsson and L.Lönnblad, "JETNET 3.0 - A versatile Artificial Neural Network Package", *Comp. Phys. Commun.* **81** (1994) 185

[5] White H., "Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mapping", *Neural Networks,* **3** (1990) 535

[6] Thomas D.S. and Mitiche A. "Asymptotic Optimality of Pattern Recognition by Regression Analysis", *Neural Networks,* **7** (1994) 313

[7] V.V.Palichik, V.M.Severyanov, V.N.Shigaev, Ju.A.Budagov, N.O.Poroshin, "Model Neural Network and Threshold Calorimeter Triggers for b-Quark Events Selection in Fixed Target Experiments", Proc. of Int. Workshop AIHENP-93, Oberammergau, Germany, October 1993 p.347

[8] J.Budagov, V.Palichik, N.Poroshin, V.Severyanov, V.Shigaev "Assesment of Characteristics of Threshold and Neural Network Triggers for b-Events Selection Using Simulation Data for the Forward Calorimeter of "Multiparticle Spectrometer" Installation on UNK 3 TeV Proton Beam", *Commun. of JINR*, P10-93-140, Dubna, 1993 (in Russian)

[9] J.Budagov, I.Chirikov-Zorin, D.Pantea, V.Severyanov, V.Shigaev, S.Sushkov, P.Stavina, M.Bosman, M.Nessi, "Application of Artificial Neural Networks to Low p_T Muon Identification in ATLAS Hadron Calorimeter", to be published in Proc. of Int. Workshop AIHENP-95, Pisa, Italy, April 1995

Received by Publishing Department
on November 23, 1995.

Аствацатуров А.Р. и др.
Применение искусственных нейронных
эффективности разделения изолирован
с малым поперечным импульсом в адро

Показаны преимущества методики
обработки данных с адронного калориме
по разделению изолированных π-мезо
импульсом в интервале $3 < p_T < 5$ ГэВ и
таких p_T имеют заметную вероятность
лориметра, вследствие чего они не м
мюонным детектором. Представлен ср
теристик ряда нейросетевых дискри
дискриминатора, контролирующего эне
ронного калориметра. Анализ основан
при помощи стандартных программ мод
Работа выполнена в Лаборатории яд

Сообщение Объединенного института

Astvatsaturov A.R. et al.
Improvement in Separation of Isolated Mu
at Low p_T in ATLAS Hadron Calorimeter
Artificial Neural Networks Technique

Advantages of artificial neural networ
granulated ATLAS hadron calorimeter (H
π/μ separation task in the range $3 < p_T <$
low p_T muons have a significant probabili
therefore they cannot be reliably registered
analysis of main characteristics is presented
a linear threshold discriminator operating
HC. The analysis is based on MC data obt
The investigation has been performed
JINR.

Communication of the Joint Institute