Electron/Jet Neural Discrimination based on Nonlinear Independent Components for ATLAS Second-Level Trigger

Eduardo Simas, José Seixas and Luiz Calôba

Signal Processing Laboratory

Federal University of Rio de Janeiro – Brazil

Presenter: Eric Leite

ACAT 2010 - Jaipur, India, 22-27 February, 2010

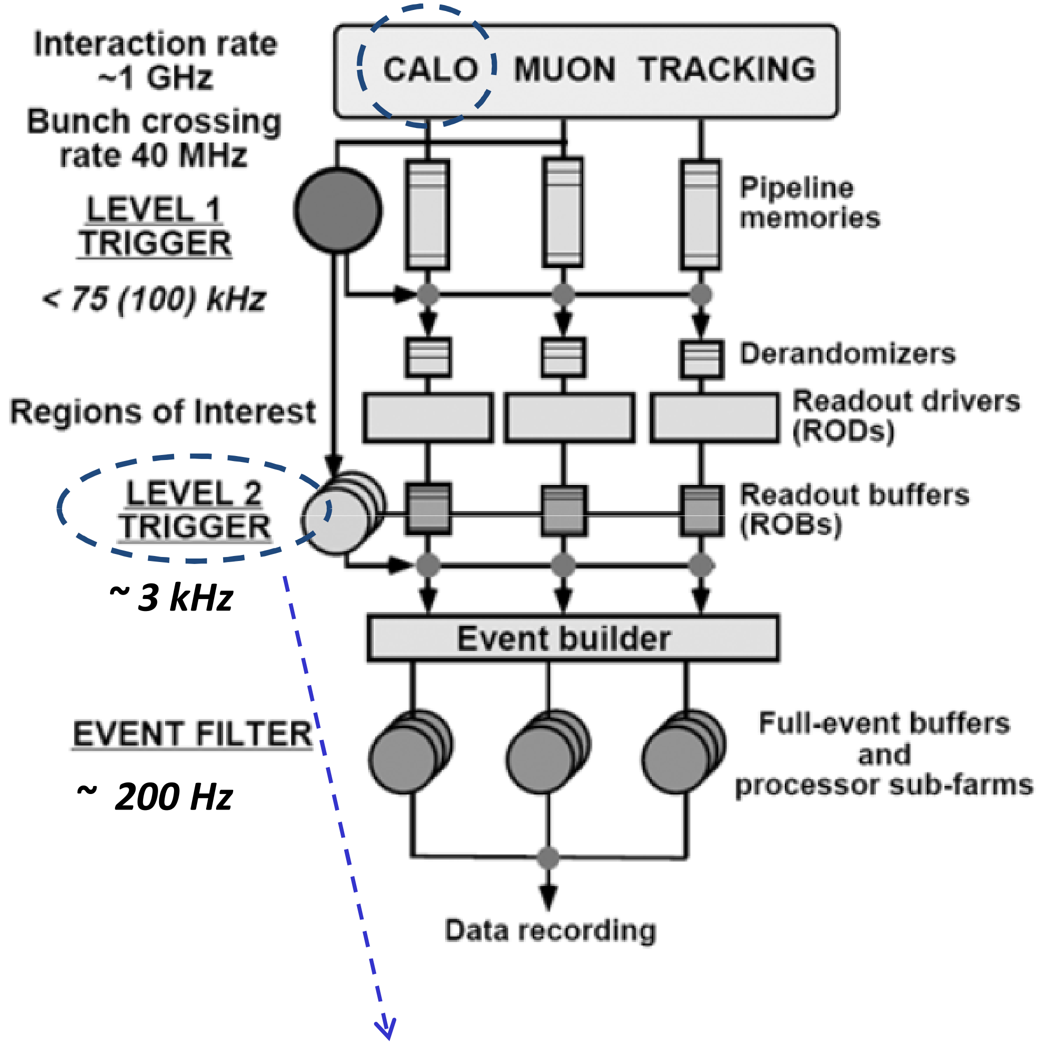
Topics

- Introduction
- Signals pre-processing
- Independent component analysis
  - Post-Nonlinear (PNL) Mixture Model
- Results
- Conclusions

Introduction

ACAT 2010

- ATLAS is a general purpose detector at LHC;
- Considering LHC interaction rates and ATLAS segmentation, an input data rate of ~60 TBytes/s is expected;
- ATLAS online trigger: three levels;
- Uses information from three subdetectors (calorimeters, muon system and tracking);
- L1 is implemented in hardware and selects Regions of Interest - RoI (regions where interesting events may have occurred);
- L2 and EF run software triggers;
- L2 operates over RoI information selected by L1.



This work proposes an alternative algorithm for L2 electron/jet separation based on calorimetry !

Introduction

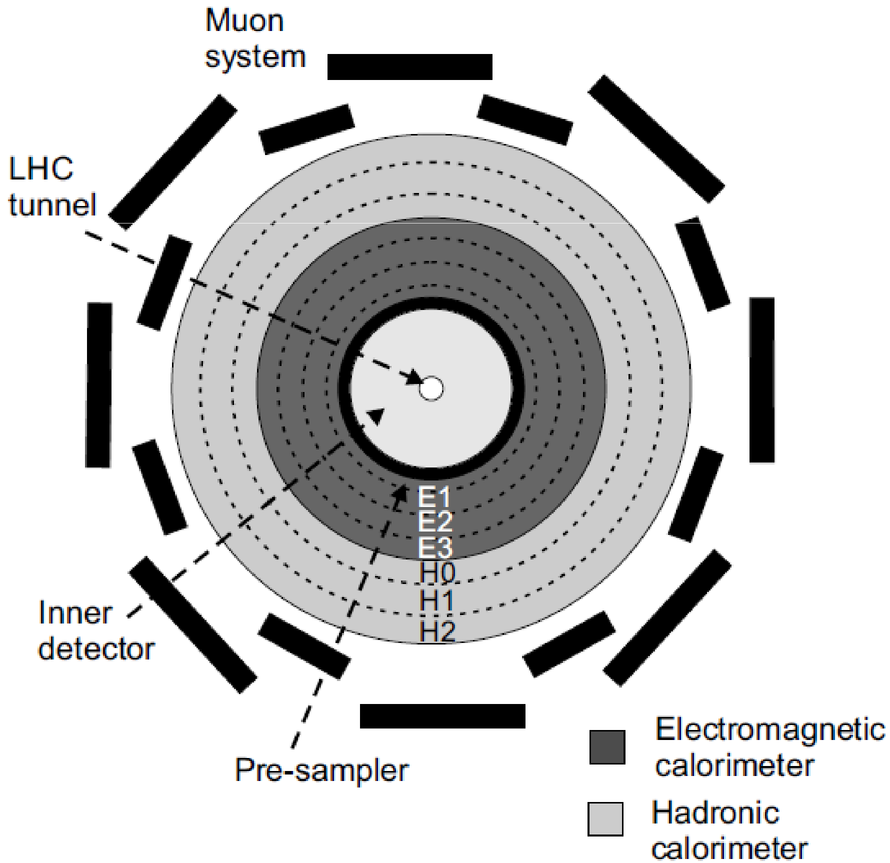
ACAT 2010

Electron / Jet channel at L2:

- Important signatures for the experiment (Higgs, SUSY, etc) may produce electrons in their decays;
- Electron identification uses the energy deposition profile measured by ATLAS calorimeter system;
- Some QCD jets may present energy deposition profiles similar to electrons;
- Electron/jet channel: ~35% of the L2 input rate (27 kHz of 75kHz);
- L2 latency time: 40 ms.

ATLAS calorimeters:

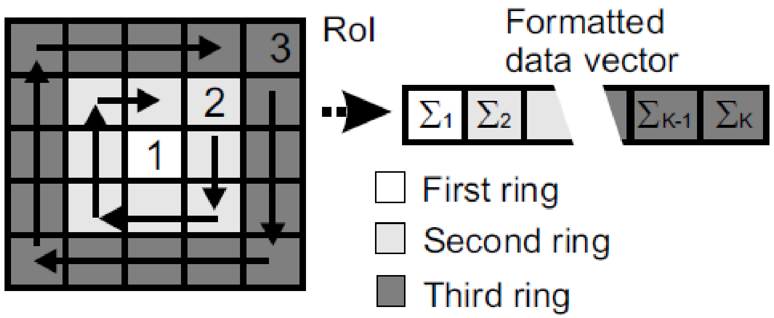
- Seven layers with different characteristics and granularity.



Signal Pre-processing

ACAT 2010

- Here the calorimeter signals are pre-processed using a ring-like structure:



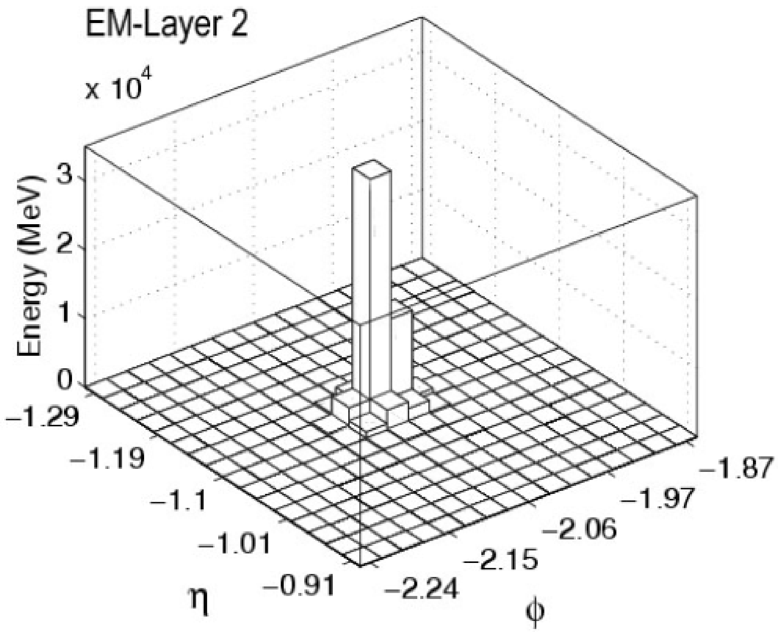
a: select a calorimeter layer;

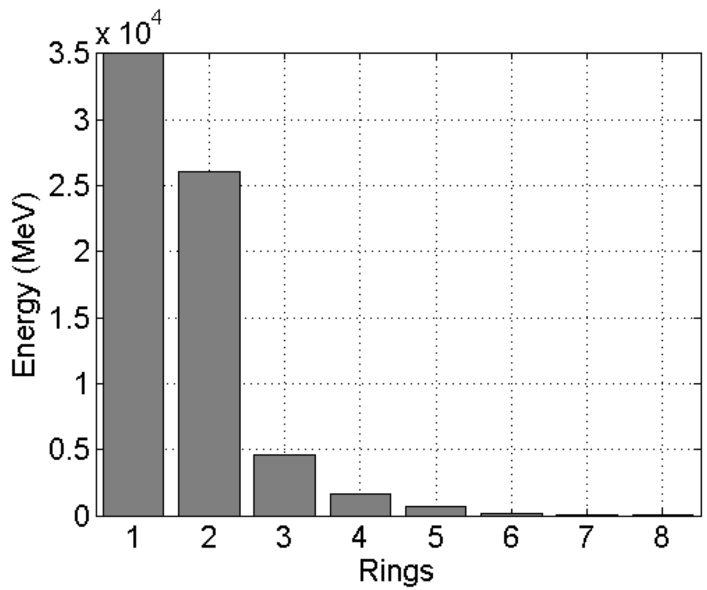
b: find the hottest cell $\rightarrow$ Ring 1;

c: select the cells around the first ring $\rightarrow$ Ring 2;

d: repeat this procedure over the RoI .

Example:





Ring Signals

ACAT 2010

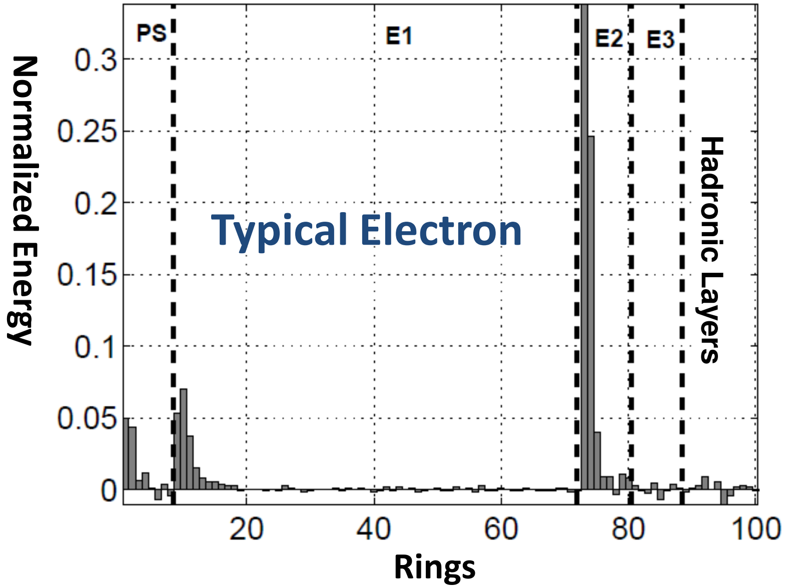
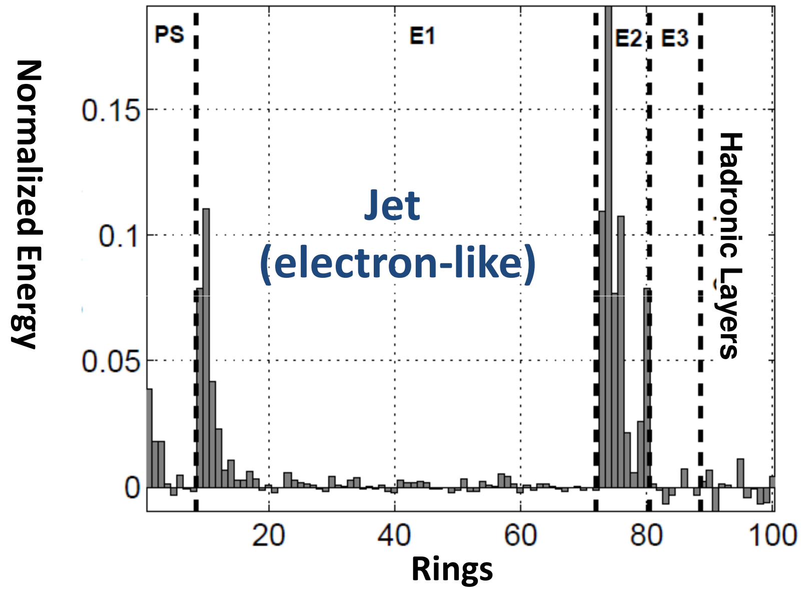
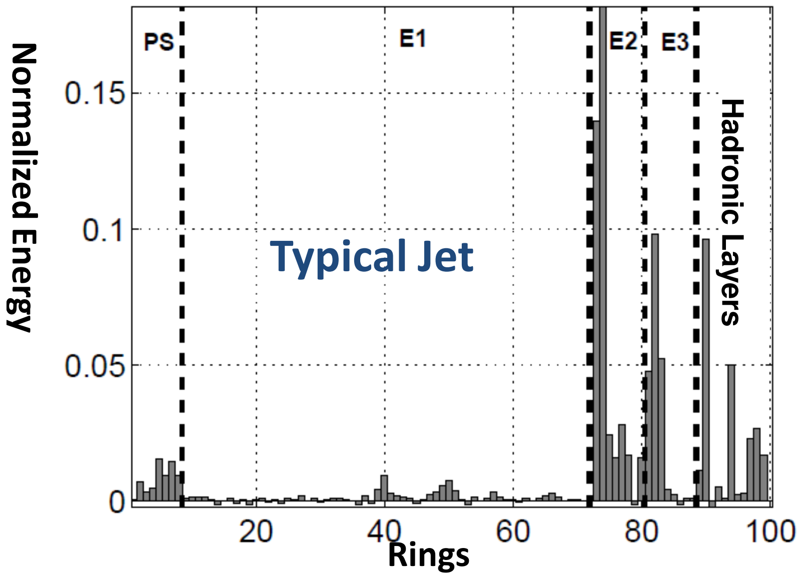


Table 1: Number of rings built up from each calorimeter layer.

Layer	PS	E1	E2	E3	H0	H1	H0	Total
Rings	8	64	8	8	4	4	4	100



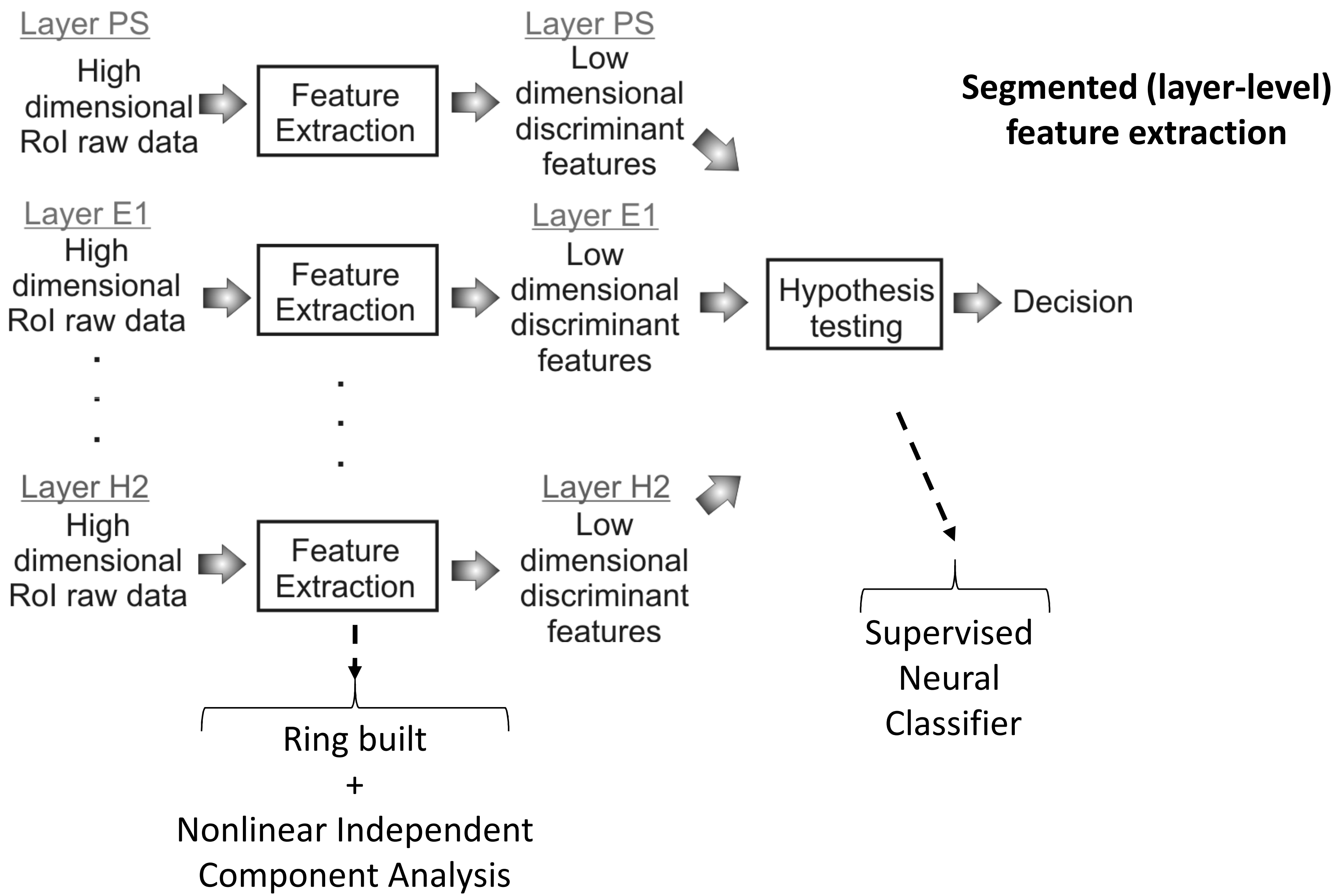


Simulated Data:

- ~160k single electrons between 7 and 80 GeV;
- ~100k QCD dijets events;
- Datasets initially pre-filtered by L1, considering energy, EM and HAD isolation.;
- ~140k electron and ~13k fake electron (jets) RoI reached L2.

Signal Processing Chain proposed for ATLAS L2 Trigger

ACAT 2010



Independent Component Analysis

- Independent component analysis (ICA) is a statistical signal processing technique which aims at finding underlying independent features in a dataset.

- The ICA model considers that a N-dimensional observed vector **x**=$[x_1, x_2, \ldots, x_N]^T$ was generated as independent source signals **s**=$[s_1, s_2, \ldots, s_N]^T$ propagated through a "mixing system":

$$\mathbf{x} = \mathbf{A}\mathbf{s}$$

- Nonlinear ICA model:

$$\mathbf{x} = \mathrm{F}(\mathbf{s})$$

F(.) is a $R^N \rightarrow R^N$ nonlinear mapping

- Statistical independence measures such as negentropy, mutual information or the Kullback-Leibler divergence may be applied to search for an inverse transformation:

$$\hat{\mathbf{s}} = \mathrm{G}(\mathbf{x})$$

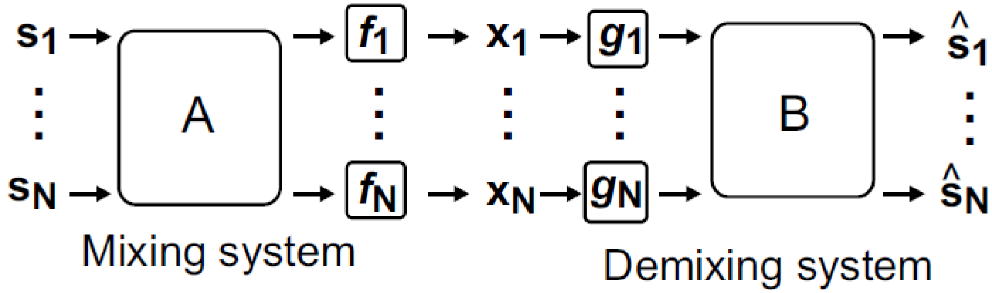
G(.) is a $R^N \rightarrow R^N$ (inverse) nonlinear mapping

- In which the components of $\hat{\mathbf{s}}$ are as independent as possible.

Post-Nonlinear ICA

ACAT 2010

- The post-nonlinear (PNL) ICA model is described through:



- For *i=1, ... , N*, the measured signals are defined as:

$$x_i = f_i\left(\sum_{j=1}^{N} a_{ij} s_j\right)$$

- An estimation of the independent signals:

$$\hat{s}_i = \sum_{j=1}^{N} b_{ij} g_j(x_j)$$

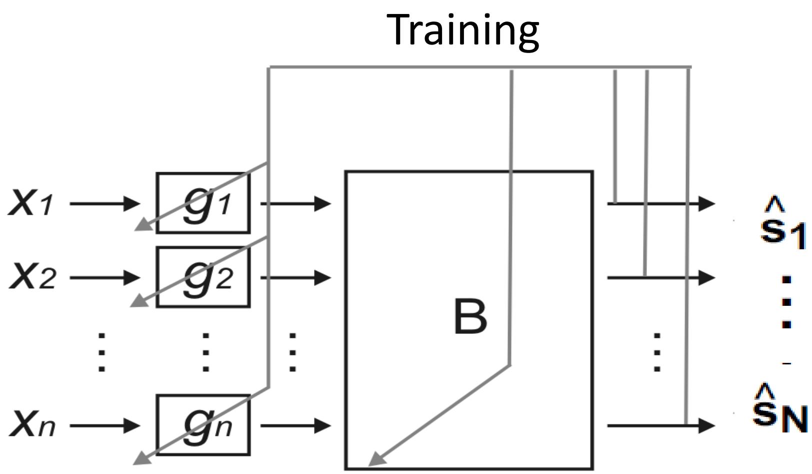
- Each nonlinear function g_k is modeled by a two-layer MLP neural network:

$$g_k(x_k) = \sum_{h=1}^{N_H} \sigma_h \tanh(\omega_h x_k + \eta_h)$$

Where N_H is the number of hidden neurons.

Post-Nonlinear ICA

ACAT 2010



- **PNL – ICA algorithm short description:**
  - Estimate the mutual information I(**s**) between the outputs $\hat{\mathbf{s}}$;

$$\mathrm{I}(\hat{\mathbf{s}}) = \mathrm{I}(\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_n) = \sum_{i=1}^{n} \mathrm{H}(\hat{s}_i) - \mathrm{H}(\hat{\mathbf{s}})$$

Entropy:

$$\mathrm{H}(s) = \sum_i P(s = a_i)\log\big(P(s = a_i)\big)$$

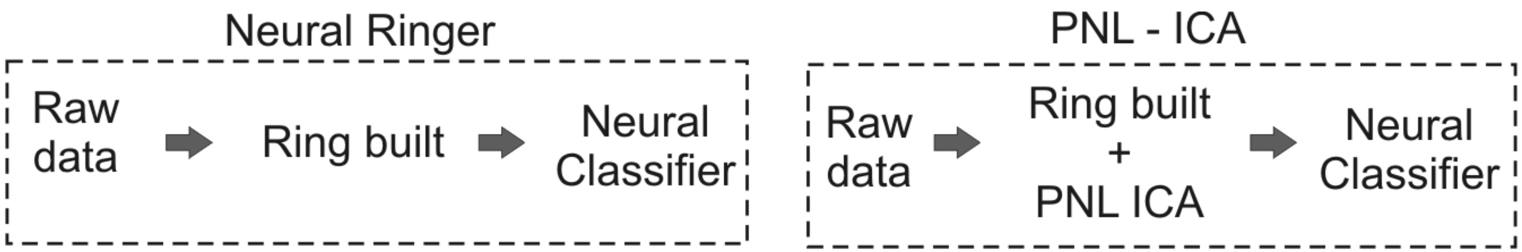
  - The parameters b_i (matrix **B**) and σ_h, ω_h, η_h (neural net estimation of g_i) are adjusted to minimize the mutual information I(**s**);
  - By minimizing I(**s**), the model outputs $\hat{\mathbf{s}}$ converge to the independent components.

Results

ACAT 2010

- Performance comparison:
  - SP index: $$\mathbf{SP} = \sqrt{\frac{Ef_e + Ef_j}{\mathbf{2}} \times \sqrt{Ef_e \times Ef_j}}$$ $Ef_e = P_D$, $Ef_j = 1 - P_F$; P_D = Electron Efficiency, P_F = Jet Acceptance
  - Receiver Operating Characteristics (ROC) curve: illustrates how both detection and false alarm probabilities (respectively P_D and P_F) vary as the decision threshold changes.

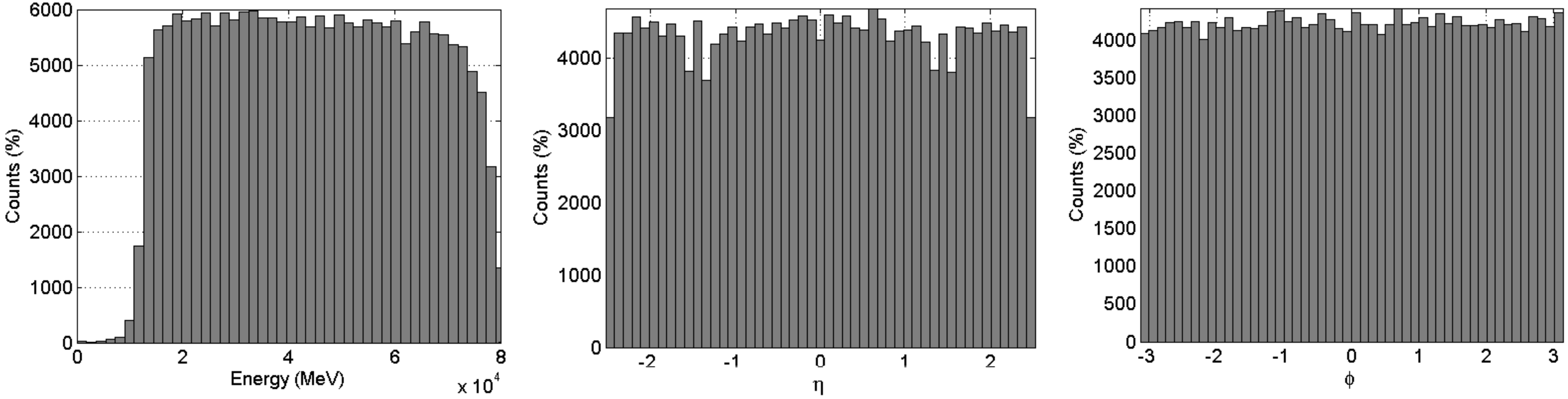
- The proposed method is compared to the Neural Ringer discriminator, which consists on a supervised neural discriminator operating directly over the ring signals



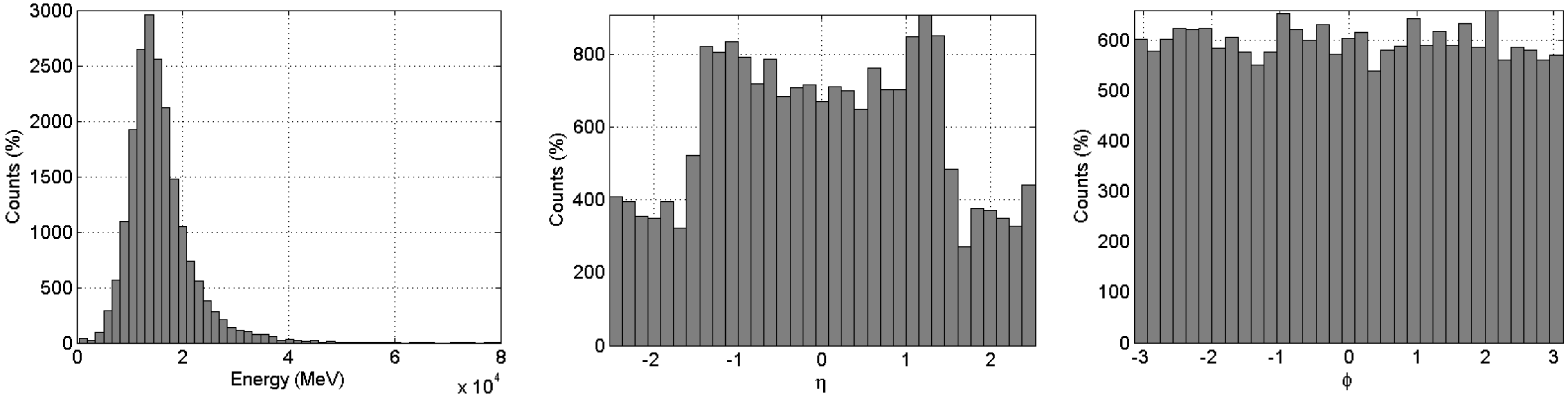
Results – Simulated L2 Input Data

ACAT 2010

Electrons



Jets

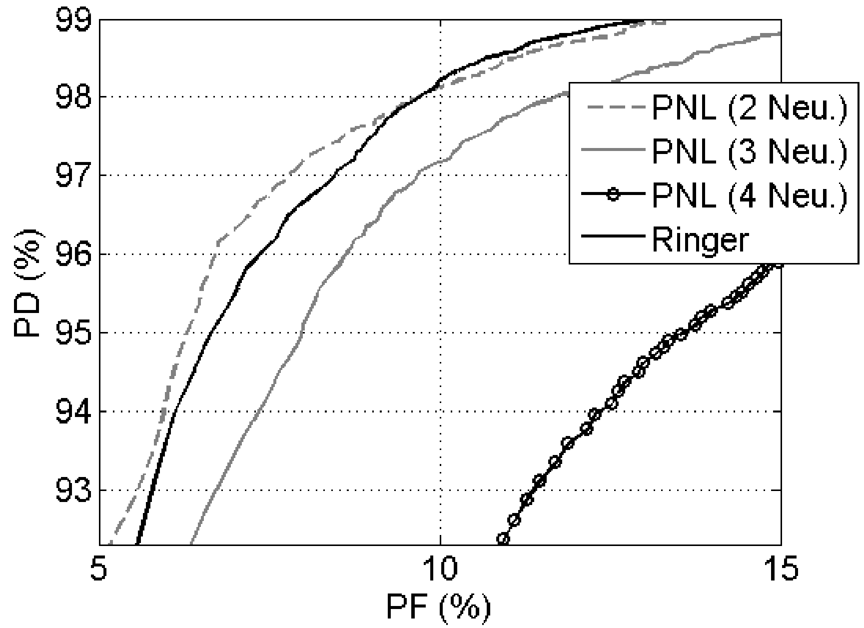


Results

ACAT 2010

- Number of neurons used to estimated each nonlinear function in the PNL model:
  - The nonlinearities are expected to be smooth (the calorimeter is approximately linear) ;
  - The same number of hidden neurons was used to estimate each nonlinearity;
  - By increasing the number of neurons the discrimination efficiency decreases.

Discriminator	Best SP x 100	PF for PD=97%
Ringer	94.35	(8.67 ± 0.20) %
PNL (2 Neur.)	94.70	(7.69 ± 0.35)%
PNL (3 Neur.)	93.70	(9.67 ± 0.38)%
PNL (4 Neur.)	90.83	(17.39 ± 0.40)%



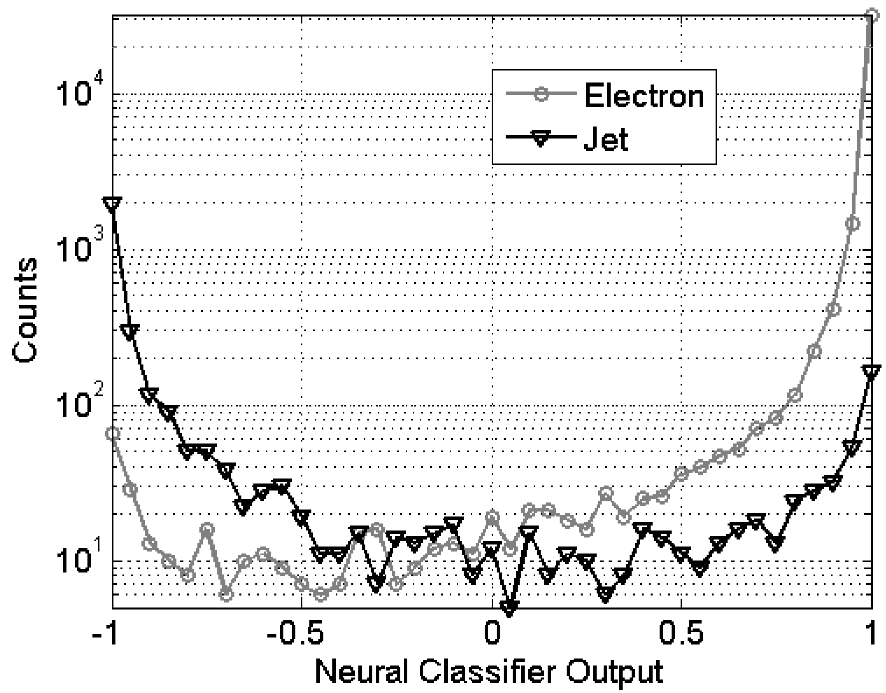
P_D – Probability of Detection (Electron Efficiency)
P_F – Prob. of False Alarm (Jet Acceptance)

Results - (PNL model / 2 neurons)

- The neural classifier (Multi-layer perceptron architecture) :
  - one hidden layer ;
  - one output neuron ;
  - tanh activation function;
  - training: error back-propagation.

- Topology: 31 x 7 x 1

- Neural classifier output:

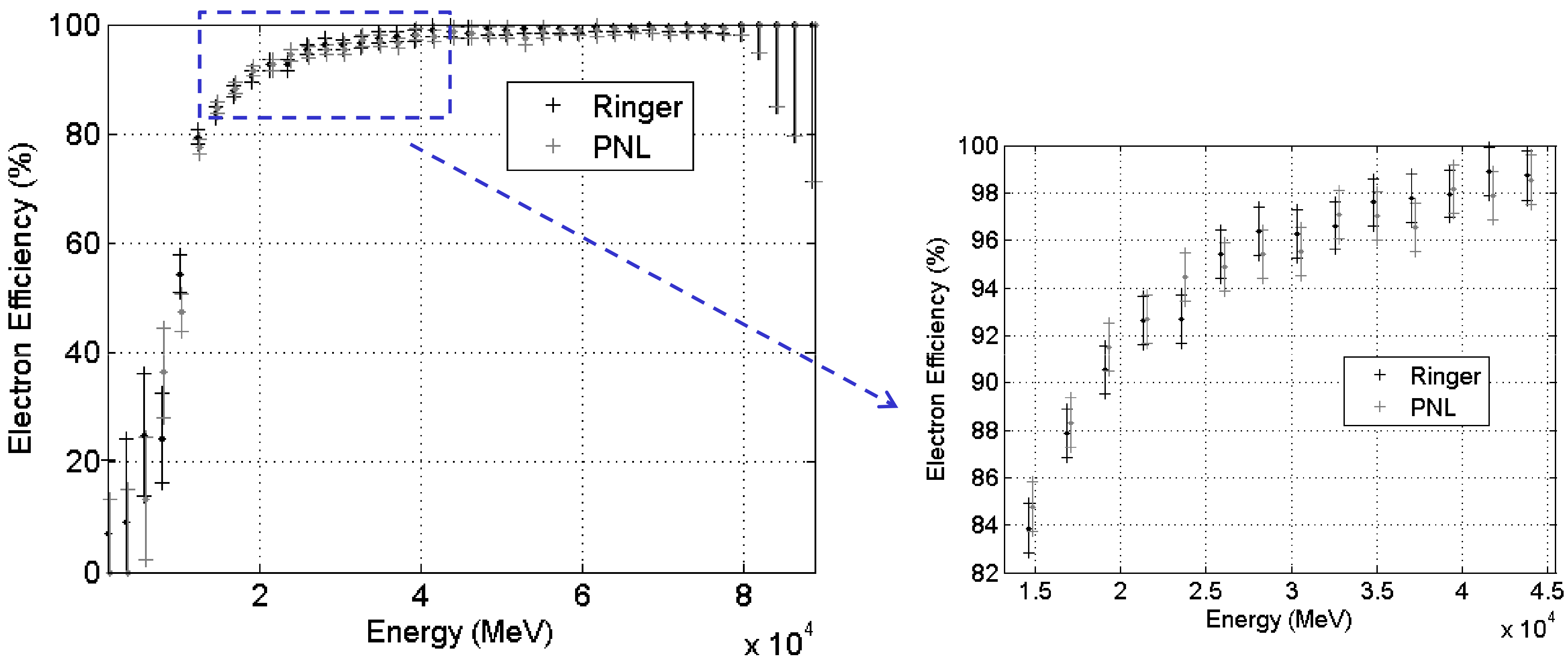


Applied cut:

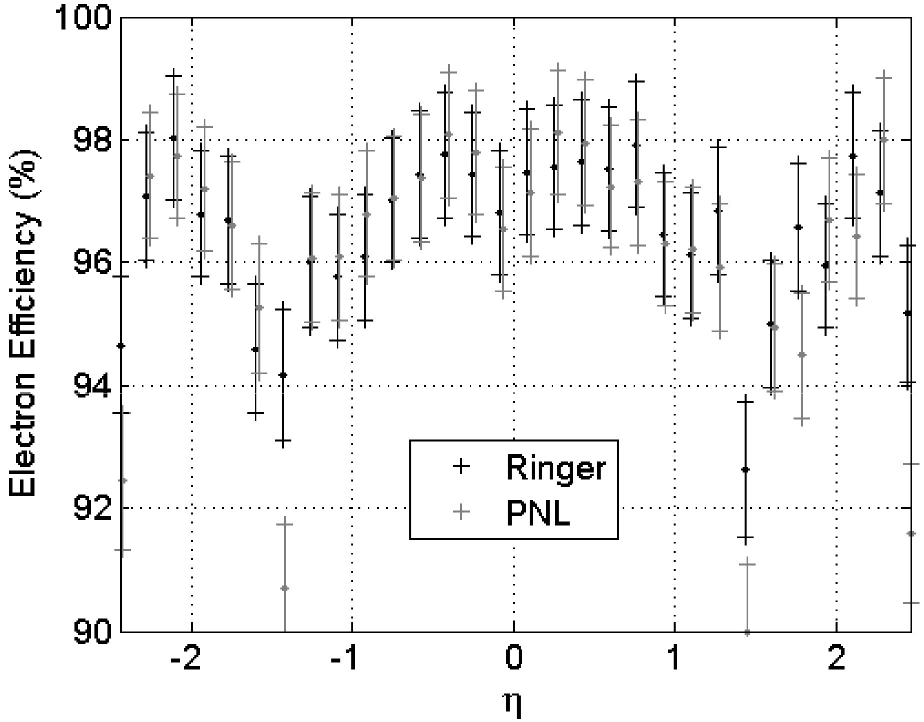
The threshold which produced maximum SP was used to operate the discriminator.

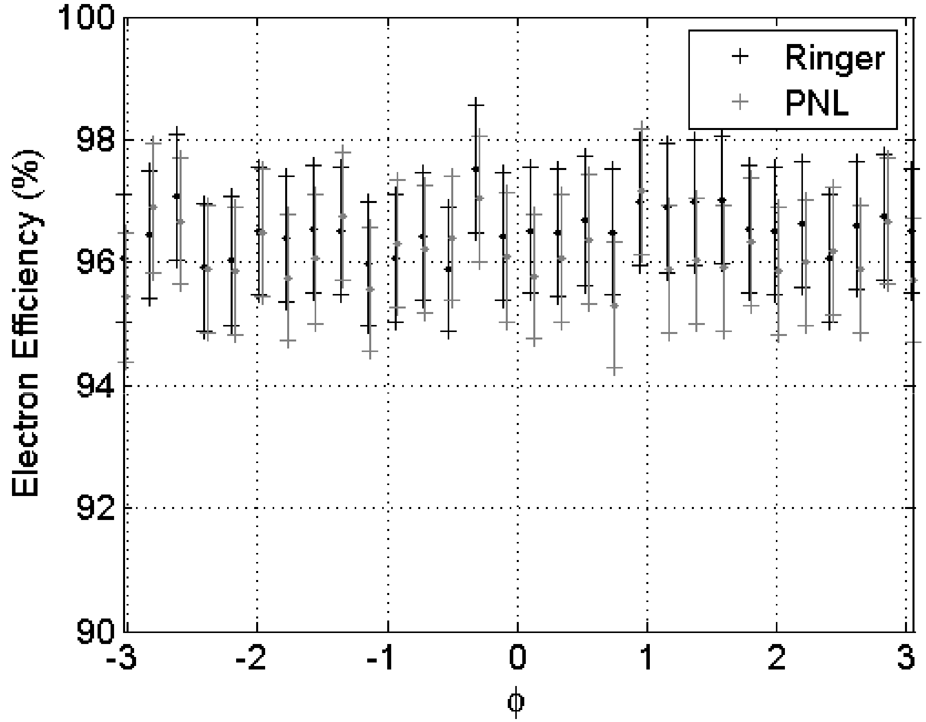
Results – Electron Efficiency (PNL model / 2 neurons)

ACAT 2010



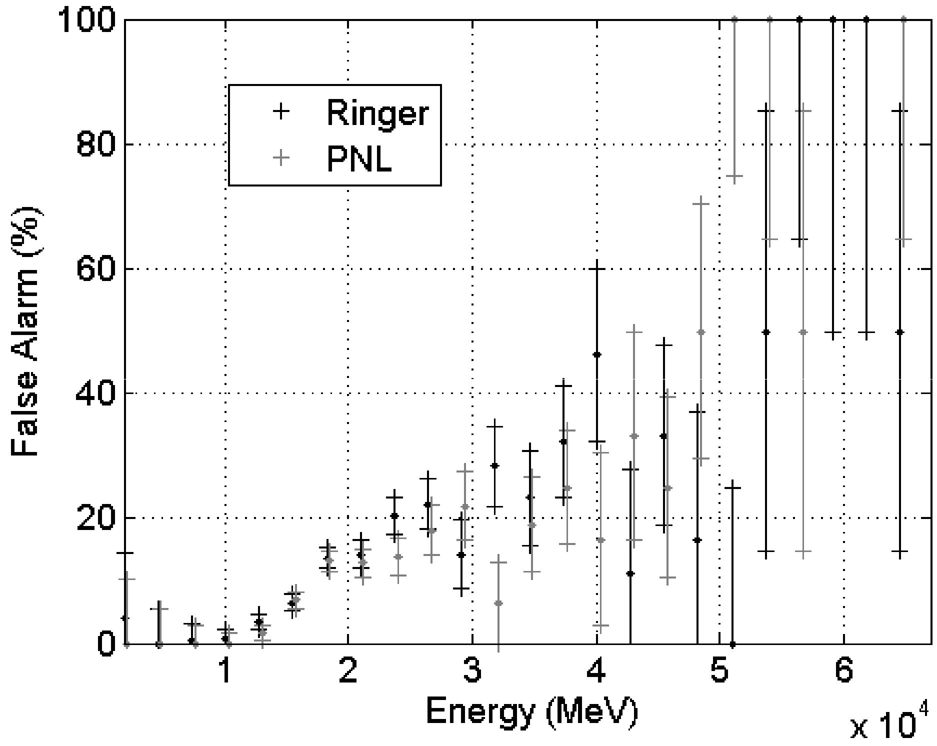
Results – Electron Efficiency (PNL model / 2 neurons)





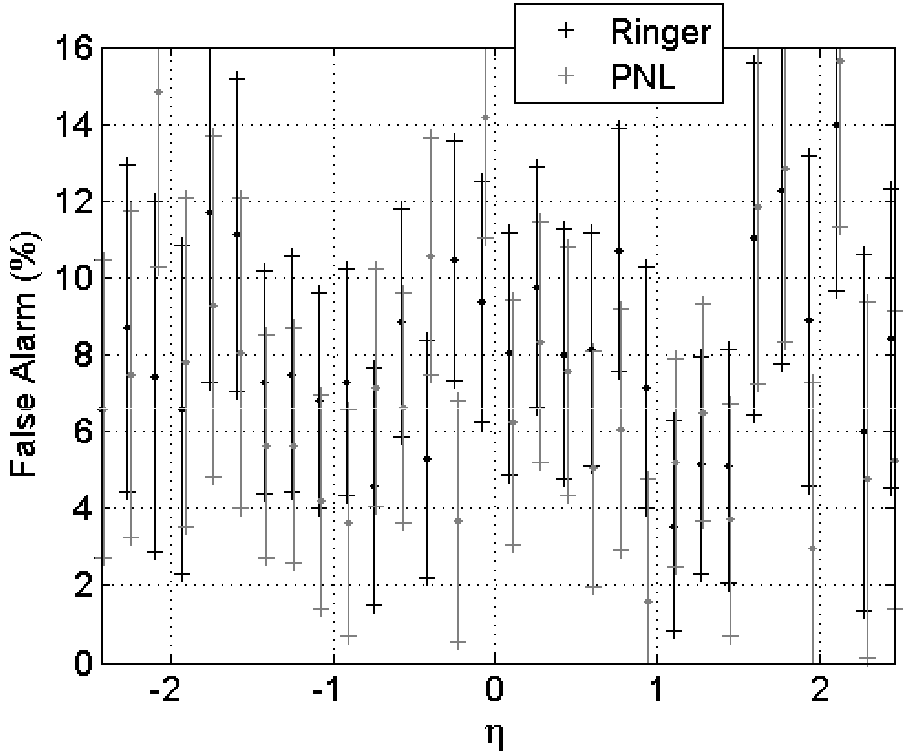
Results – Jet False Alarm (PNL model / 2 neurons)

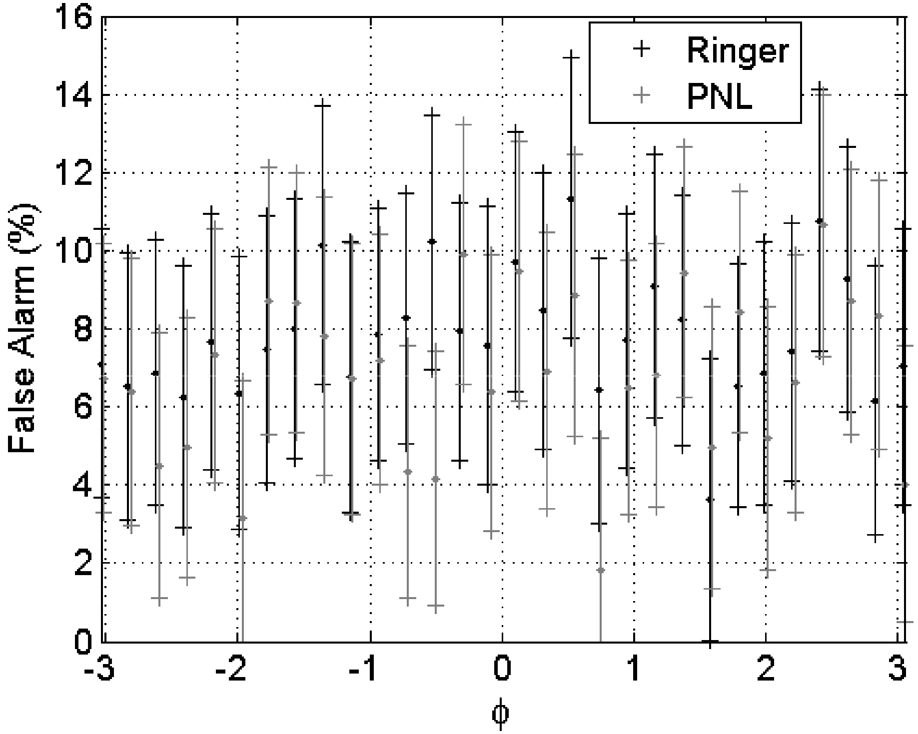
ACAT 2010



Results – Jet False Alarm (PNL model / 2 neurons)

ACAT 2010





Conclusions

- A segmented feature extraction procedure based on NLICA was proposed for the ATLAS L2 trigger (e^-/j channel);

- The nonlinearities estimated by the model were limited to a slight deviation from full linearity, as the calorimeter response is almost linear. Results confirmed such assumption.

- Compared to Neural Ringer discriminator, the false alarm was reduced from (8.67 ± 0.20) % to (7.69 ± 0.35) %, at a detection probability of 97%;

- A proper evaluation of the processing time is underway, but a preliminary comparative analysis indicates that the L2 requirements are satisfied.

Acknowledgments : ATLAS Trigger/DAQ Collaboration for fruitful discussions concerning this work.