Application of Neural Networks for Energy Reconstruction

J. Damgov and L. Litov
University of Sofia

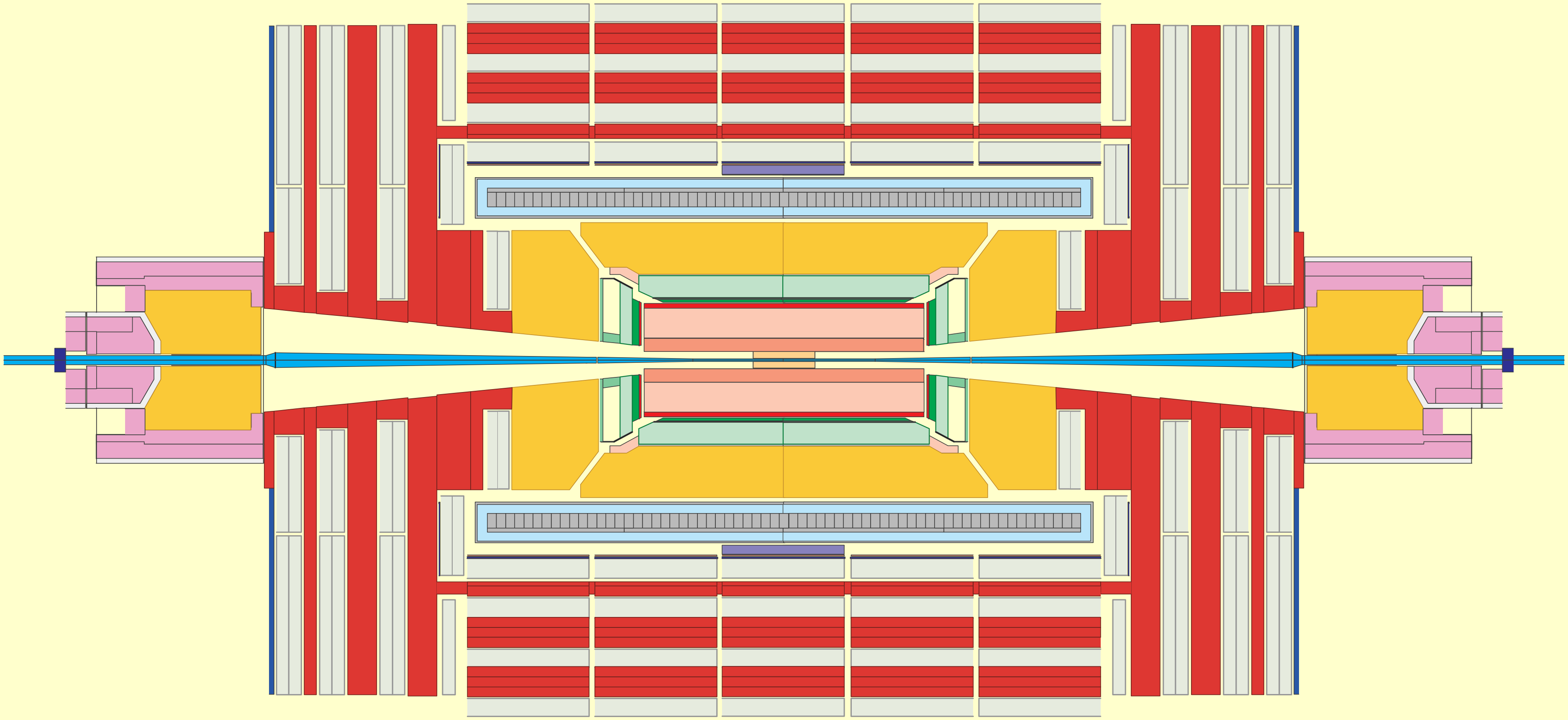
Introduction

- Introduction
- CMS Calorimeter System
- Energy reconstruction
- Energy Reconstruction with Neural Network
- Results
- Conclusions

Introduction

- LHC Physics Program
  - ✓ Search for SM Higgs Boson
  - ✓ $H \rightarrow \gamma\gamma$, $H \rightarrow WW \rightarrow l\nu jj$, $H \rightarrow lljj$
  - ✓ SUSY searches – big E^{t}_{miss}
- Requirement:
  - ✓ Precise measurement of the photon and electron energy – ECAL
  - ✓ Measurement of the jets energy
  - ✓ Good hermetic coverage for measuring E^{t}_{miss}
- LHC experiments
  - Precise Electromagnetic Calorimeters
  - As good as possible Hadron Calorimeters
  - Gaussian response and good linearity

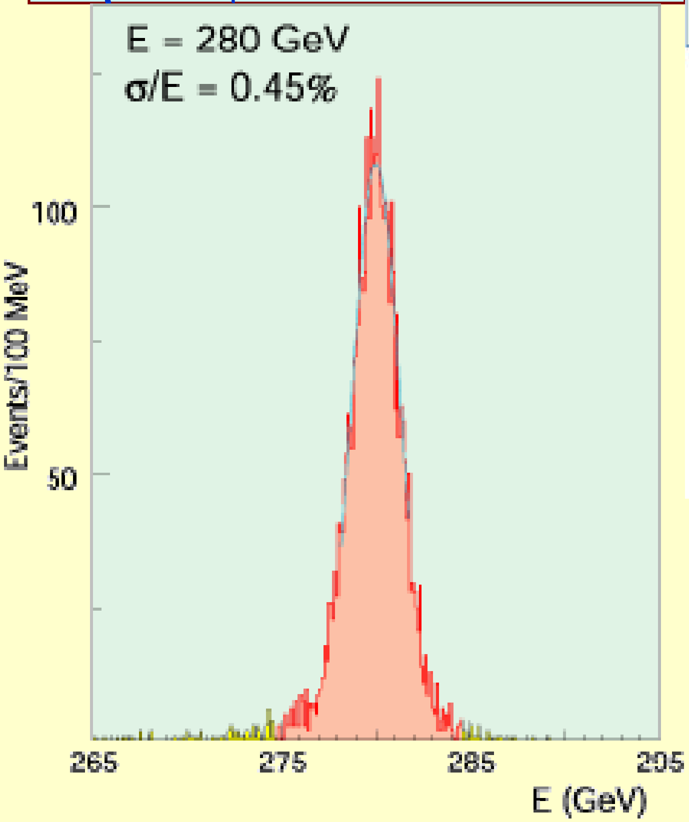
CMS detector

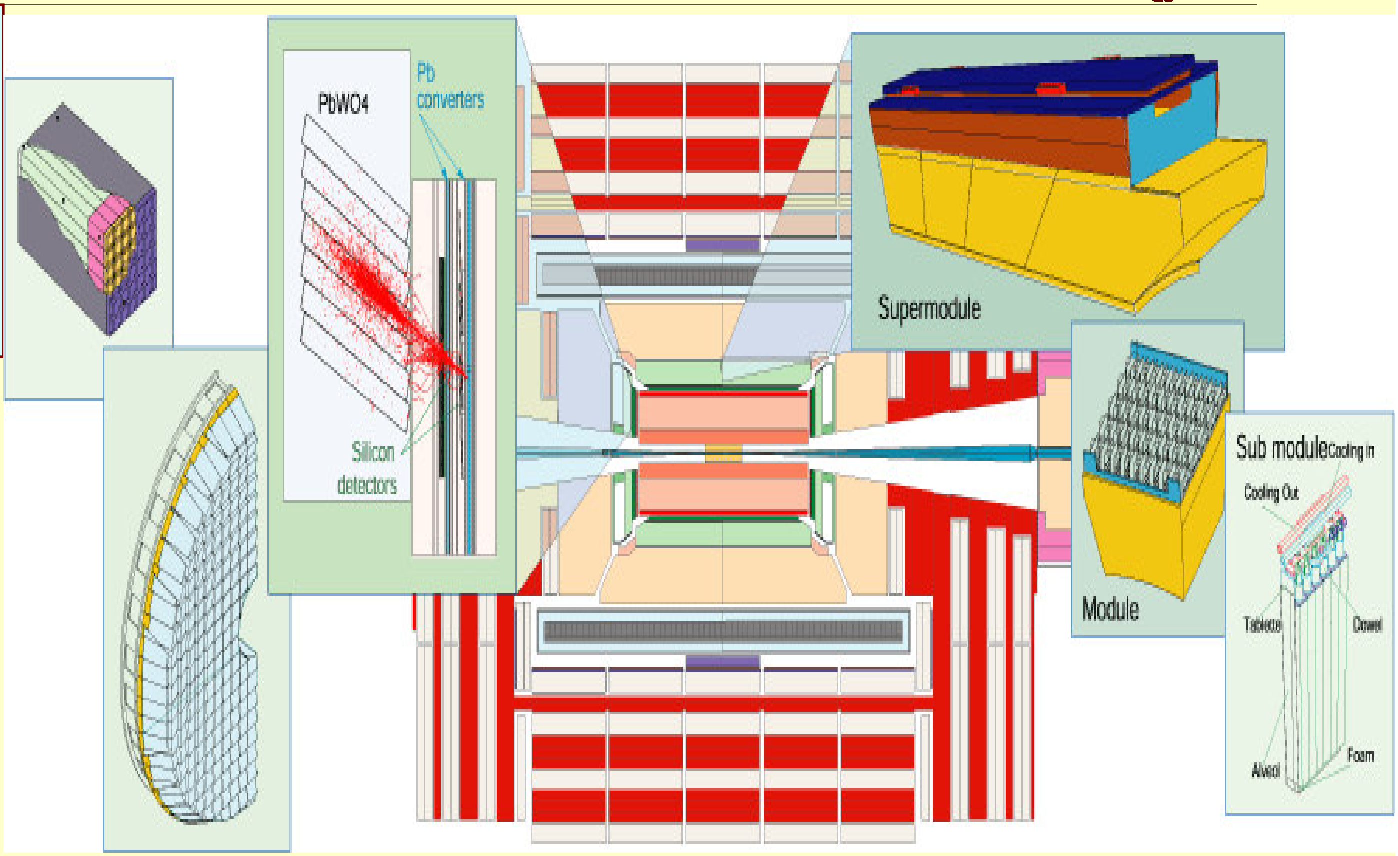


Total weight	: 12500 T	Overall length	: 21.5 m
Overall Diameter	: 15.0 m	Magnetic field	: 4 Tesla

CMS ECAL

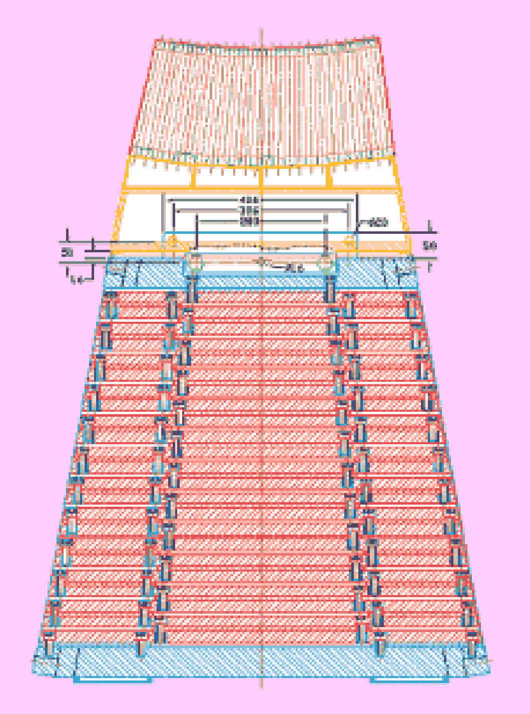
$PbWO_4$ crystals
Barrel: $\eta < 1.479$
23 cm long, 22x22 mm^2
Granularity
$\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$





Endcaps: $1.48 < \eta < 3.0$
Variable granularity
$\Delta\eta \times \Delta\phi = 0.05 \times 0.05$
26 Radiation lengths

CMS HCAL



Endcaps:

Absorber - 8 cm

Lateral segmentation:

$\Delta\eta$ x $\Delta\phi = 0.087x0.087$

Longitudinal:

HE1(1 layer),

HE2(17 layers)

Sampling Calorimeter

Absorber – copper alloy

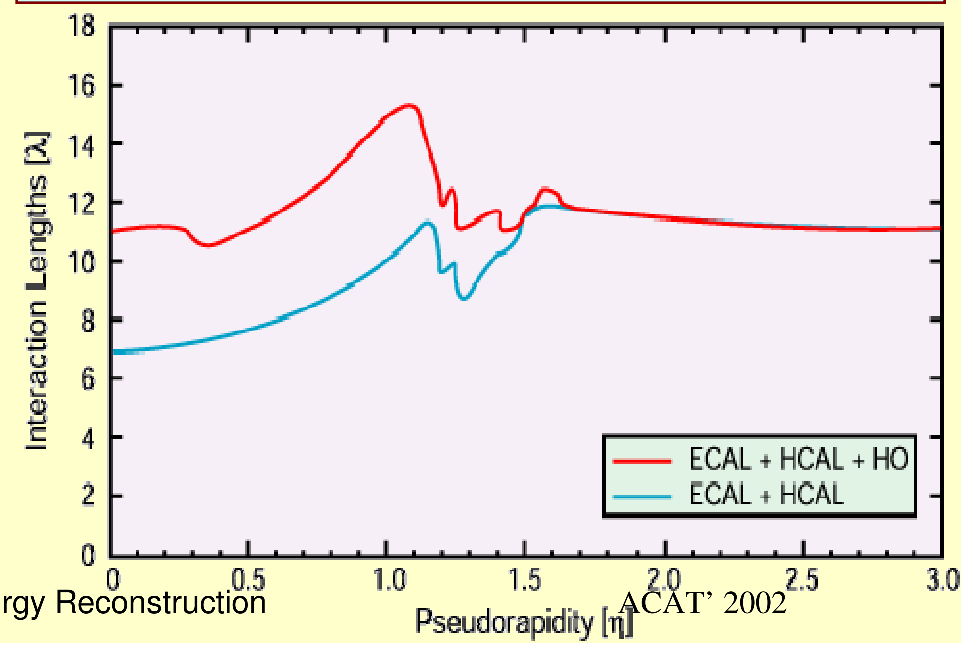
Active elements –

4mm thick scintillator tiles

HB, HE, HO

HO – lateral segmentation as in HB

2 layers;0< η <0.4 – 3 layers



Barrel: (HB)

Absorber plates - 5 cm thick

Lateral segmentation: $\Delta\eta$ x $\Delta\phi = 0.087x0.087$

Longitudinal: HB1(1 layer), HB2(17 layers)

CMS Calorimeter System

Barrel:

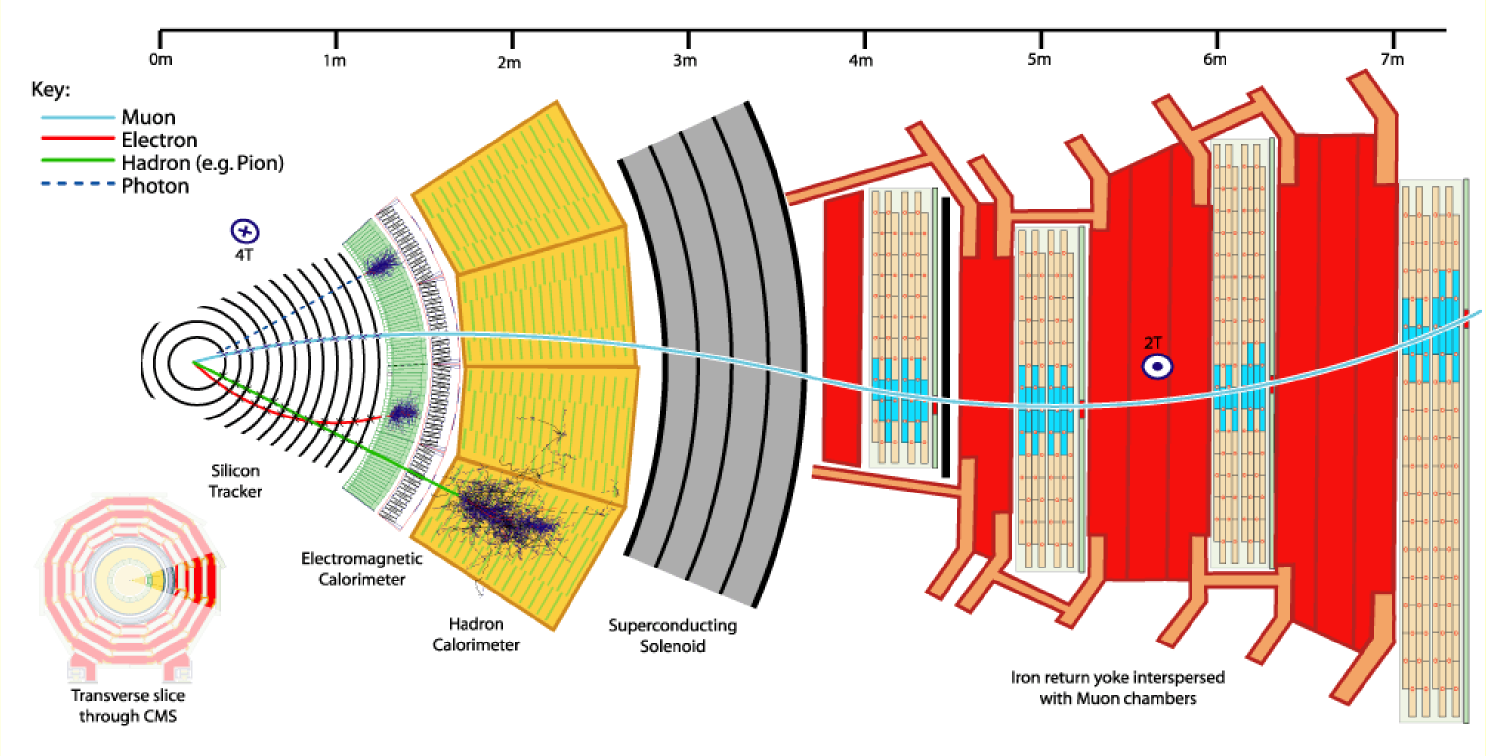
4 longitudinal read-outs

ECAL, HB1,HB2,HO

Endcaps:

3 longitudinal read-outs

ECAL,HE1,HE2



Calibration: ECAL – e-beam scan and in situ calibration – Z → e^+e^-

HCAL calibration – several wedges with hadron and muon beams

Transfer of the calibration to the other wedges with radioactive source.

In situ calibration – obligatory (response depends from magnetic field)

Single track hadrons, photon + jet, dijet resonances W →jj, Z →bb, Z → ττ

Energy reconstruction

- Hadron calorimeters –
  - Intrinsic (stochastic) fluctuations
  - Sampling fluctuations
- EM shower – $E_{vis} \sim E_{inc}$
- Hadron shower:

  $E = E_{EM} + E_h$

  $E_h = E_{ch} + E_n + E_{nuc}$
- Response for e and hadrons is different – $e/\pi > 1$
- Non-compensating Calorimeters
- Response depends on the type of the particle – it is different for e, hadrons and jets

Energy reconstruction

Most common approach (SM):

$$E_{rec} = \sum_j w_j E_j$$

w_j are determined by minimization of the width of the energy distribution with additional constraint

$$\langle E \rangle = E_{inc}$$

Linearity:

$$L = \frac{(E_{rec} - E_{inc})}{E_{inc}}$$

Test – MC events, e and π

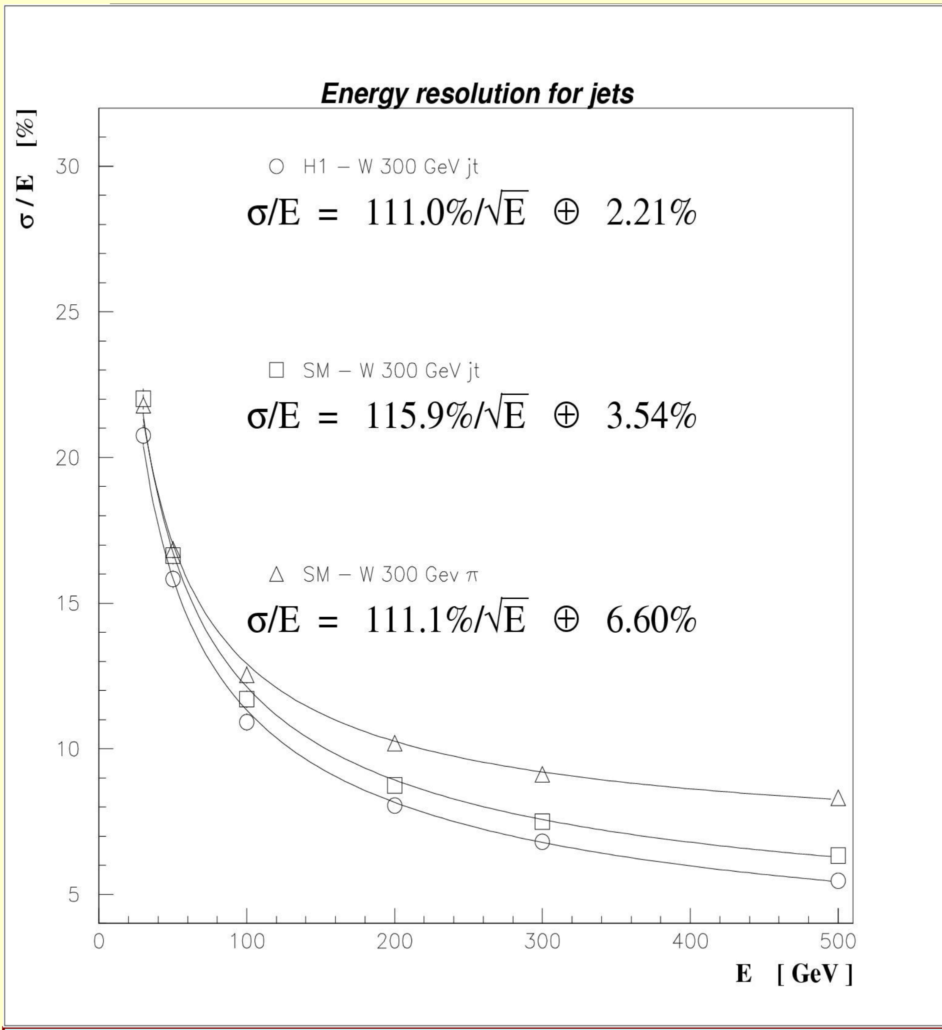
E = 5,10,20,50,100,200,300,500 GeV

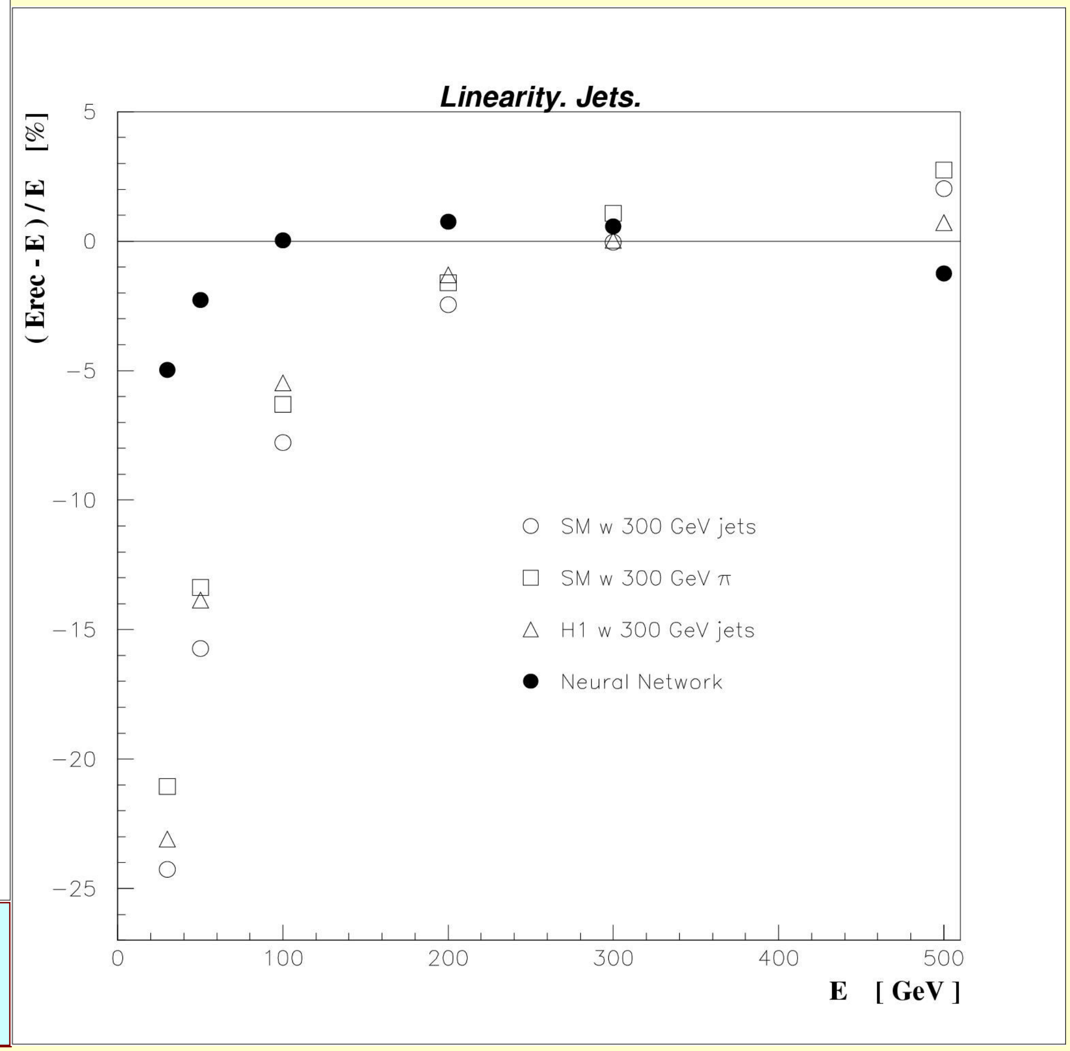
Jets - E = 30,50,100,200,300,500 GeV

w_j are energy dependent

Energy reconstruction

Standard Method

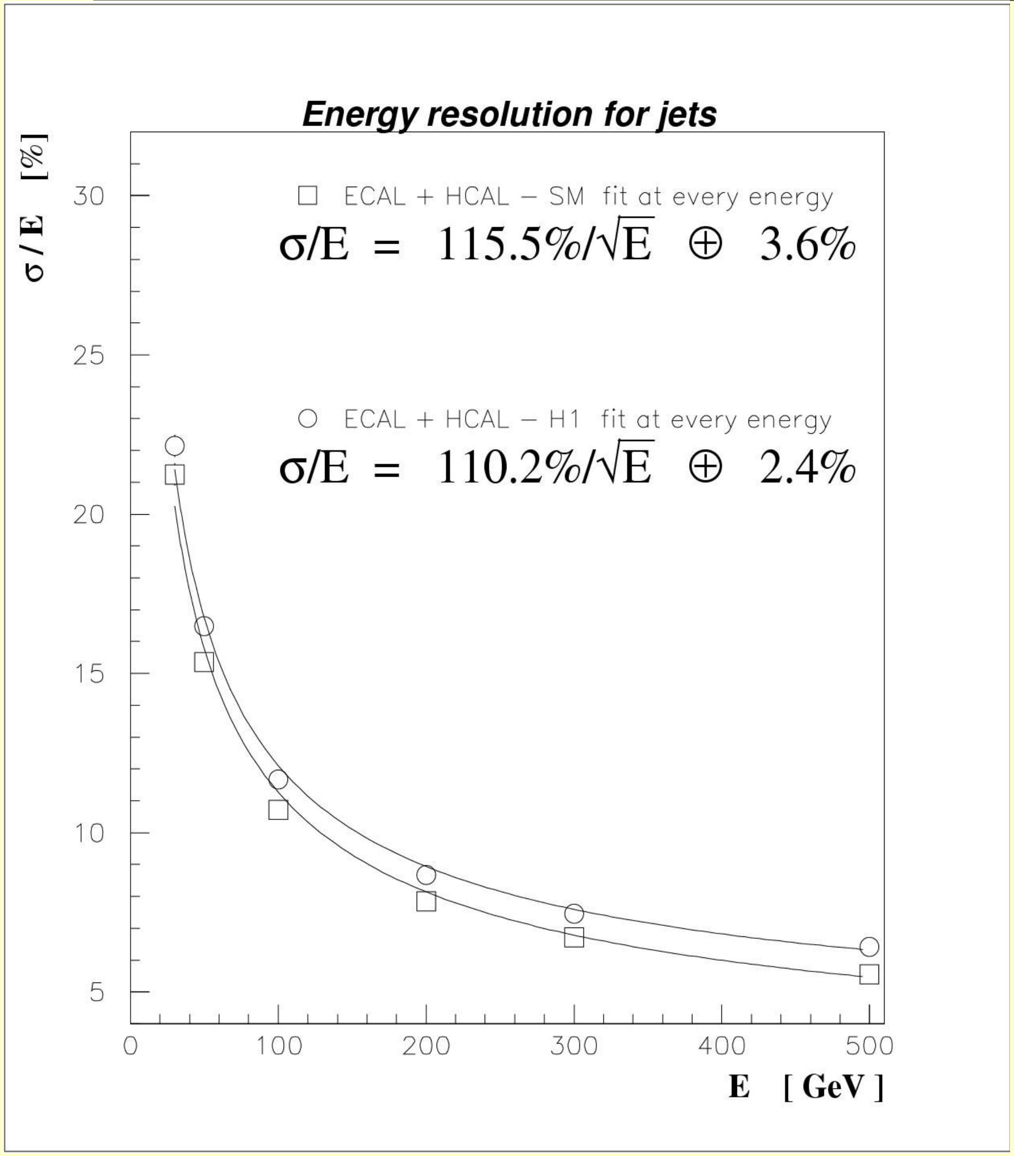




Non-Gaussian tails

Non linear response

Energy reconstruction



Energy dependent weights

- linearity is restored
- no improvement in the energy resolution

- In SM –weights are sensible to the average of fluctuations
- Different correction factor to each event
- Suppression of the EM signal
- Different weighting methods – H1

$$E_{rec} = \sum_i (w_i \sum_j E_{ij} - v_i \frac{\sum_j E_{ij}^2}{\sum_j E_{ij}})$$

Slight improvement – constant term

Energy reconstruction

- To ensure the best possible measurement of the energy
  - To every individual event – different correction factor
  - Using the lateral and longitudinal development - EM part of the hadron shower should be estimated
  - The type of the particle (electron, hadron, jet) should be determined
- We need a method
  - Able to deal with many parameters
  - Sensitive to correlation between them
  - Flexible to react to fluctuations
- Possible solution – Neural Network

Neural Network

Powerful tool for:

- ❖Classification of particles and final states
- ❖Track reconstruction
- ❖Particle identification
- ❖Reconstruction of invariant masses
- ❖Energy reconstruction in calorimeters

Basic computing element - Neuron

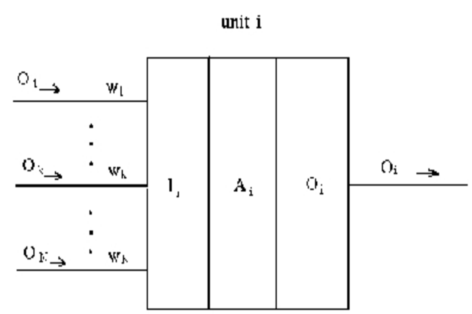


fig 1.NN

neuron performs calculations in three steps

$$I_i = \sum_k w_{ik} O_k, \qquad A_i(I) = \frac{1}{1+e^{-(I_i+b_i)}}, \qquad O_i = \Theta(A_i - A_{0i}), \qquad (1)$$

Neural Network

- Multi-Layer-Feed Forward network consists of:
  - Set of input neurons
  - One or more layers of hidden neurons
  - Set of output neurons
  - The neurons of each layer are connected to the ones to the subsequent layer

- Training
  - Presentation of pattern
  - Comparison of the desired output with the actual NN output
  - Backwards calculation of the error and adjustment of the weights
- Minimization of the error function

$$E = \frac{1}{2}\sum_{j}(t_j - o_j)^2$$

Neural Network

- Backpropagation learning algorithm

$$\Delta w = -\eta \frac{\partial E}{\partial w}$$

- η - learning rate - varies significantly
- Rprop - uses individual learning rate and Manhattan updating rule

$$\Delta w = -\eta sign[\frac{\partial E}{\partial w}]$$

At every step, η is adjusted as:

$$\eta_{w,t+1} = \gamma^{+} \eta_{w,t} \quad \text{if} \quad \partial E_{t+1} . \partial E_t > 0,$$

$$\eta_{w,t+1} = \gamma^{-} \eta_{w,t} \quad \text{if} \quad \partial E_{t+1} . \partial E_t < 0$$

$$0 < \gamma^{-} < 1 < \gamma^{+}$$

Energy reconstruction with NN

- Two possible approaches
- NN directly determined the energy dissipated in the calorimeter
  - GILDA – imaging silicon calorimeter
  - Two steps – first rough classification in of the energy – 6 groups, second step – dedicated net proceeds to discriminate among the different energy values – discrete output – weighted average
  - ATLAS – determine energy correction factors
  - Recurrent neural network with nearest neighbour feedback in the input layer and a single output – works satisfactory
- Second approach
  - Adjustment of the weights w_j on event by event basis

Energy reconstruction with NN

- Data processing in two steps
  - ✓ Identification of the type of the incident particle
  - ✓ mainly EM interacting particles – e, γ
  - ✓ Mainly strong interacting particle – hadrons
  - ✓ Jets
  - ✓ Muons
- Energy reconstruction – with dedicated NN for each class of showers
- Second level NN has four subnets for the for longitudinal read-outs

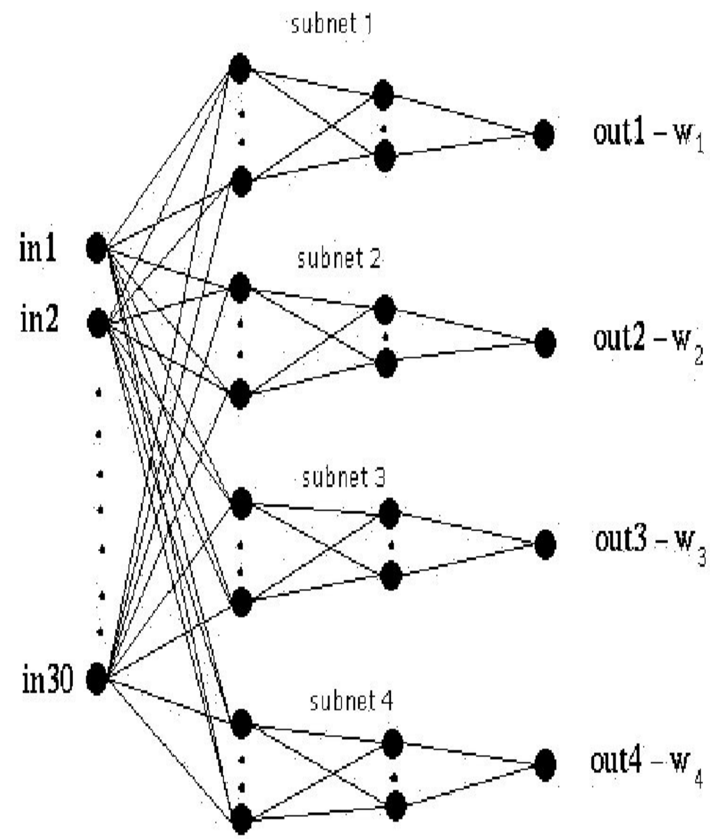


fig 2.NN

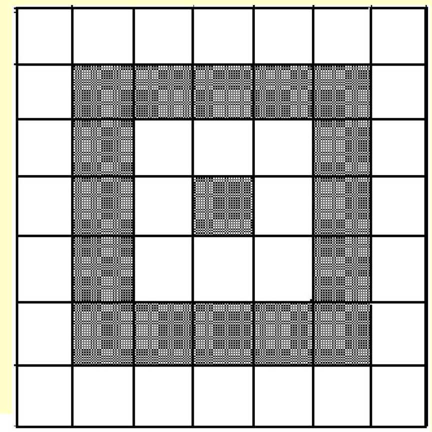
Energy reconstruction with NN

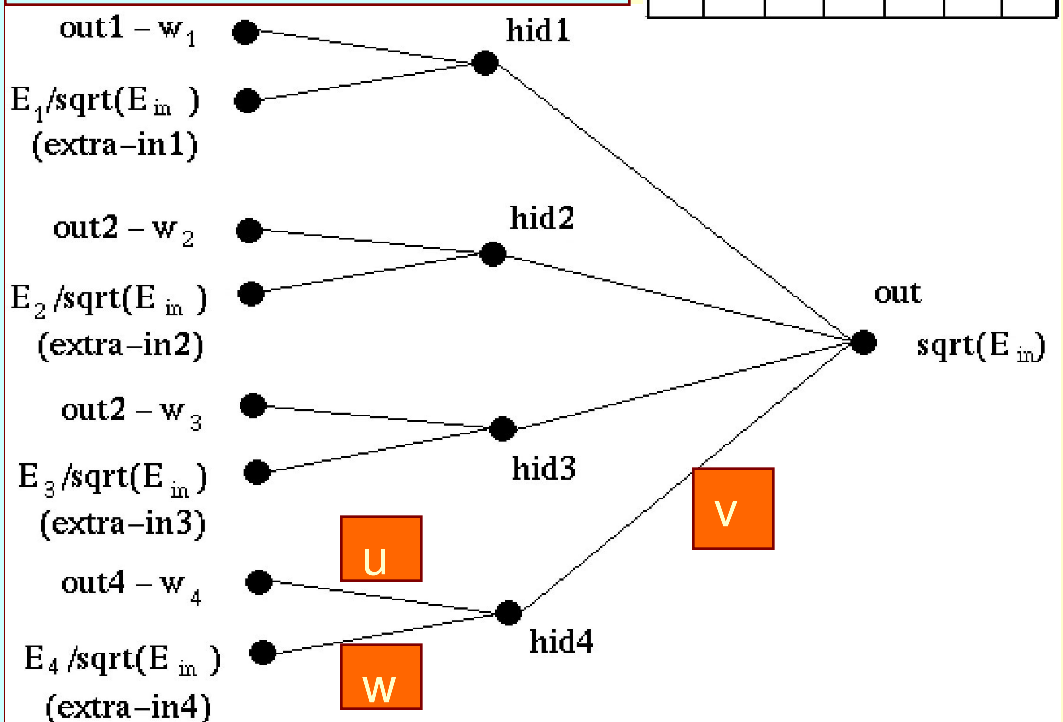
- Inputs – 30
  - E_{rec} – SM with w for 300 GeV
  - $\frac{w_i E_i}{E_{rec}}$, i = 1,2,3,4 ,
  - 13 inputs – ECAL
  - 3x4 inputs HCAL
- Additional neurons – learning
  - hidi – I(O) = O ; A(I) = I
  - Out – sums up signals A(I) = I
  - u,v and w – like all other weights

$$E_{rec} = \sum_i o_i u_i v_i w_i E_i$$

- o_i – takes into account shower fluctuations

(η,φ) – cone ΔR = 0.43
ECAL – 41x41 crystals
HCAL – 7 x 7 towers
Summing energies in concentric squares





Results

- Feed-forward neural network - 30 inputs
- Stuttgart Neural Network Simulator SNNS

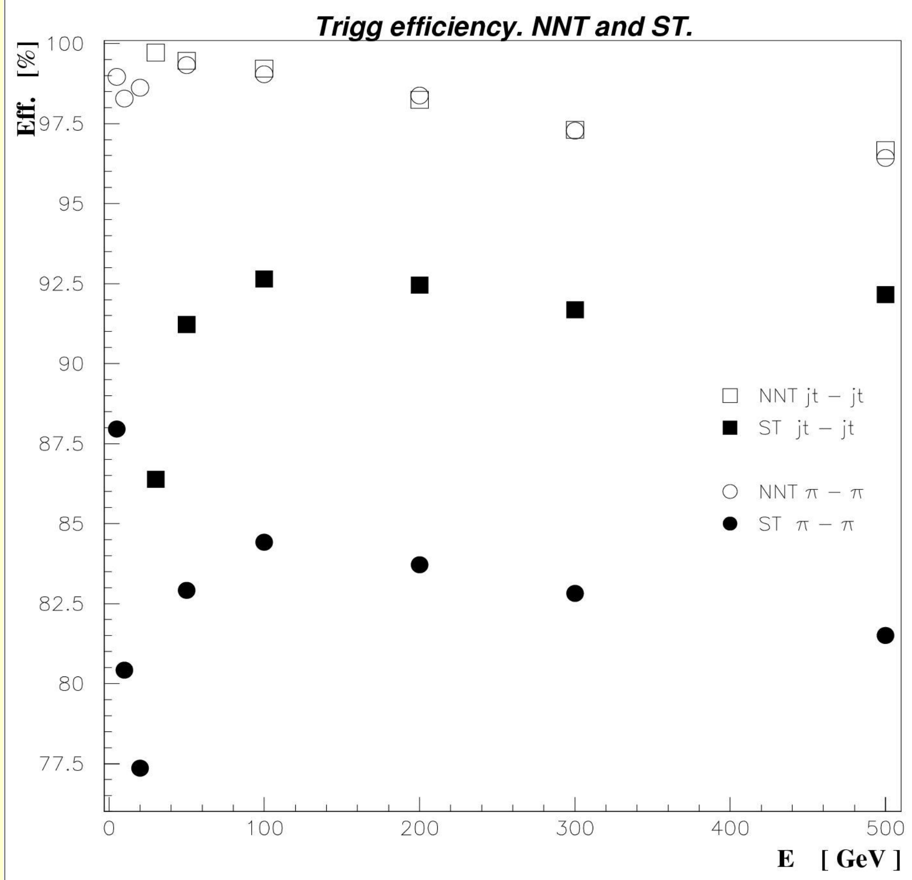
Particle separation with NN

30 – inputs, 4 – outputs for e.h,jet, μ

- Particle identification –two methods
- Using suitably chosen cuts
- Shower pseudo radius – to separate e

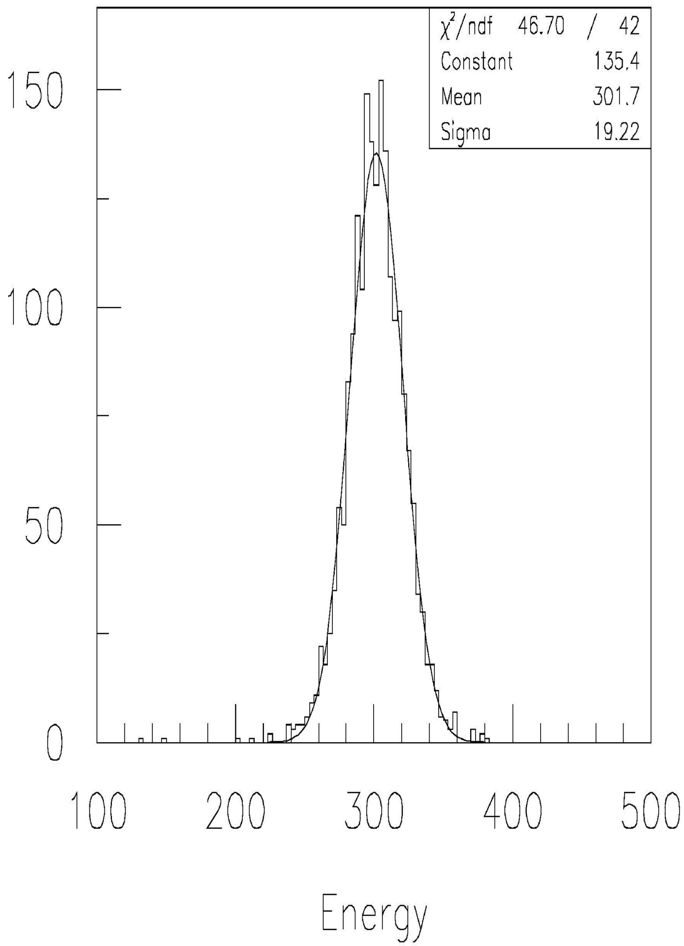
$$R_{sh} = \sqrt{\frac{\sum_{ij} E_{ij}\varphi_{ii}^2 + \sum_{ij} E_{ij}\eta_j^2}{\sum_{ij} E_{ij}} + \frac{(\sum_{ij} E_{ij}\varphi_i)^2 + (\sum_{ij} E_{ij}\eta_j)^2}{(\sum_{ij} E_{ij})^2}}$$

- Single hadron showers from jets
- $R_{sh} < 0.07$
- E_{ECAL} corresponds to MIP
- $mR > 0.332$, $mR = \frac{\sum_{ij} E_{ij}^{1.5}}{\sum_{ij} E_{ij}^{1.63}}$
- $R_2 > 37.5$, $R_2 = E_{HCAL} / E_{ECAL}$



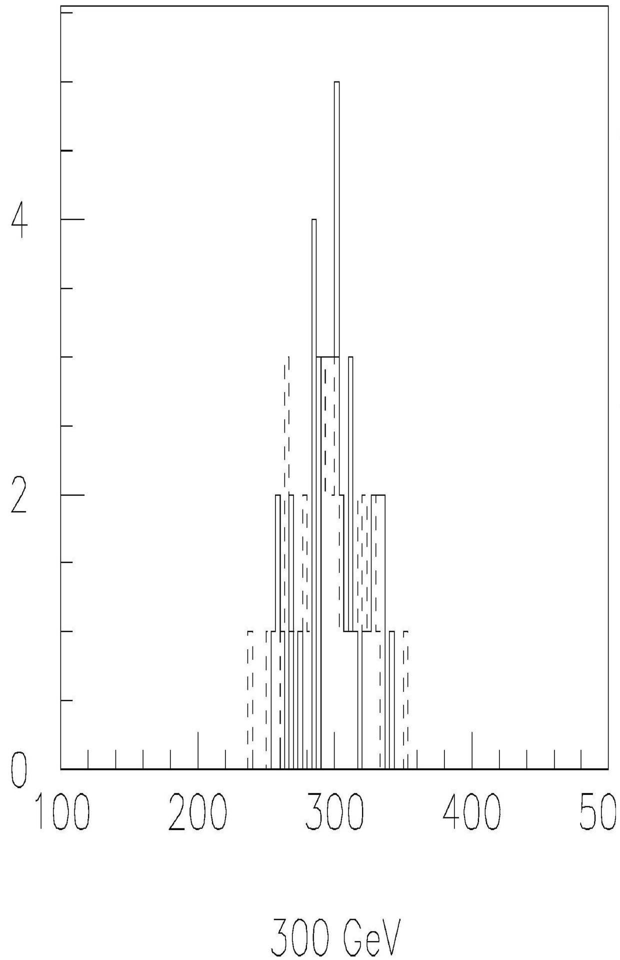
Results

NN performance
Energy distribution - Gaussian

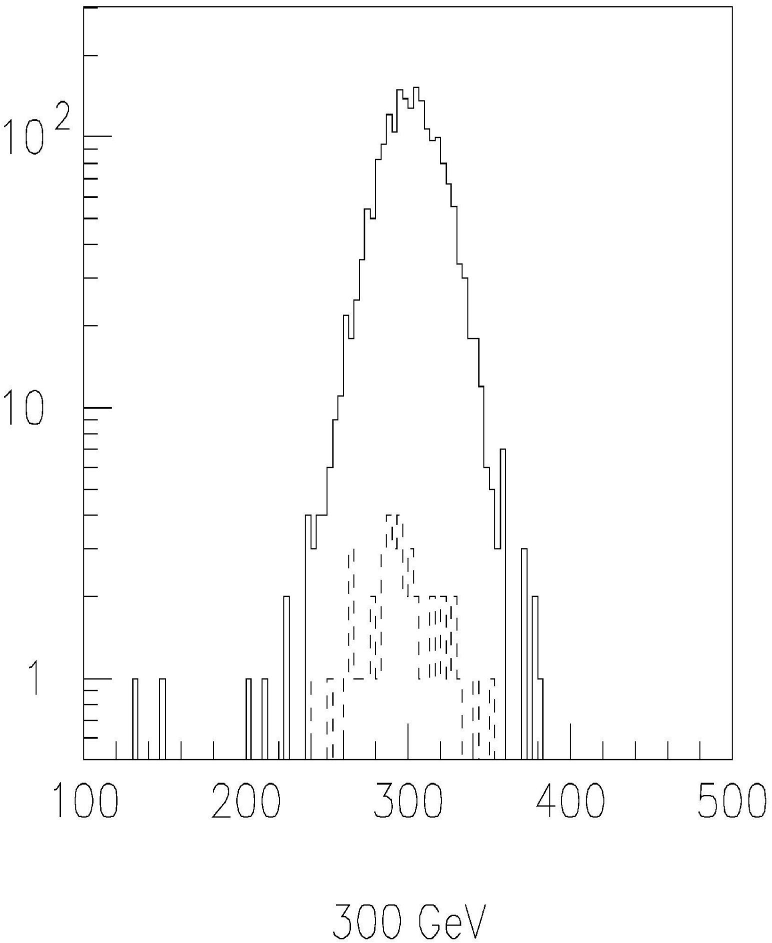


NN performance – energy is well reconstructed

jet → h



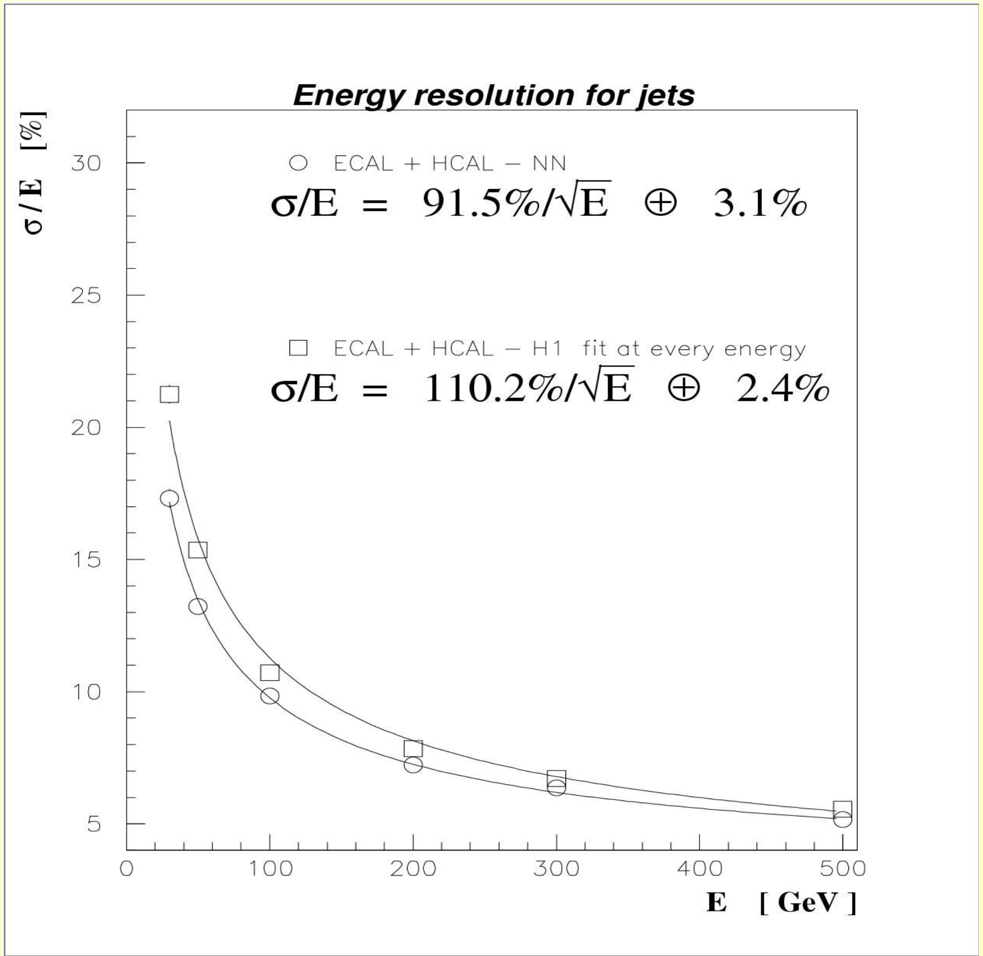
jet → e



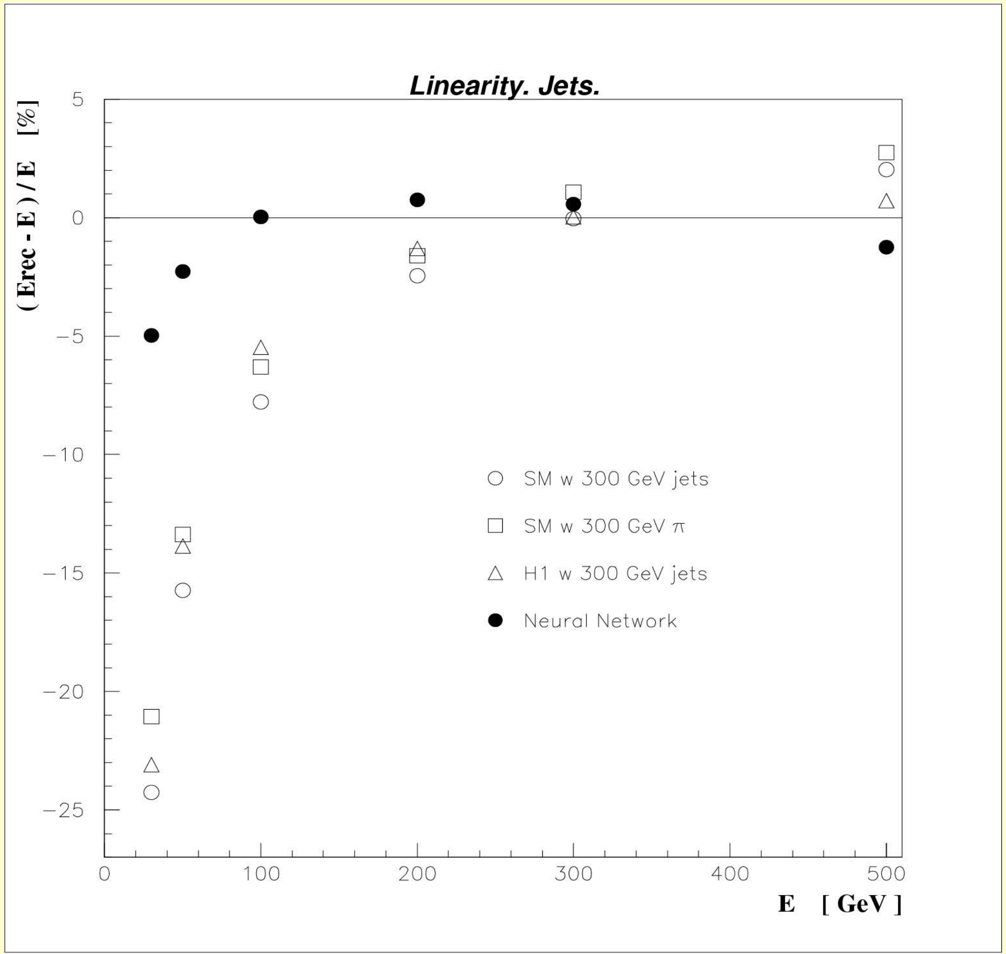
Results

Neural Network performance

Energy resolution for jets



Linearity



Conclusions

- NN has been applied for reconstruction of the energy of single h and jets
- The NN performs reconstruction in two steps
  - Determination of the type of shower initiator – e, hadron, jet
  - If the shower is misidentified, it energy is reconstructed correctly
- NN evaluates the shower energy
  - The energy spectra have Gaussian shape and are free of tails
  - Significant improvement of the energy resolution and linearity

E_{ch}

$$E_{rec} = \sum_i o_i w_i u_i v_i$$

w

$$\frac{w_i E_i}{E_{rec}}$$

$$E_{rec} = \sum_i o_i u_i v_i w_i E_i$$