ECal reconstruction optimization with neural network(NN)

---Update of EicC ECal Study

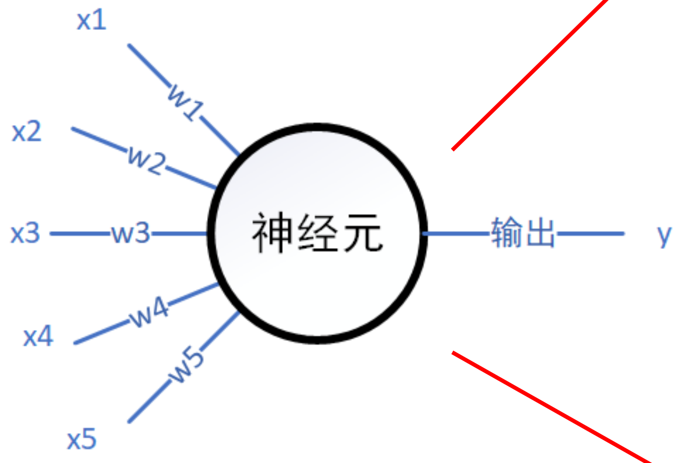
Ye Tian, Dexu Lin, Yutie Liang

Institute of Modern Physics, CAS

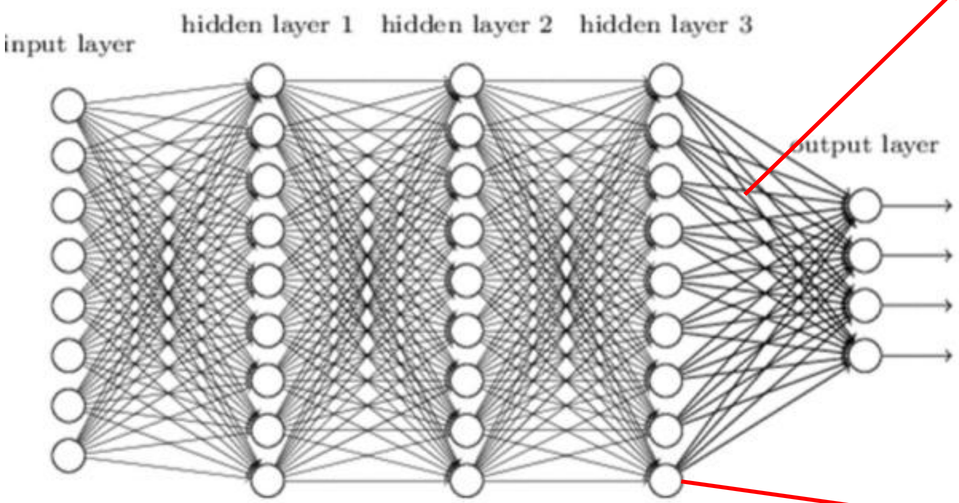
EicC Group bi-week Meeting 2022-08-03

Principle of NN(neural network)

NN abstract: a **block box，**
the parameter is determined by itself with training



Common NN structures:



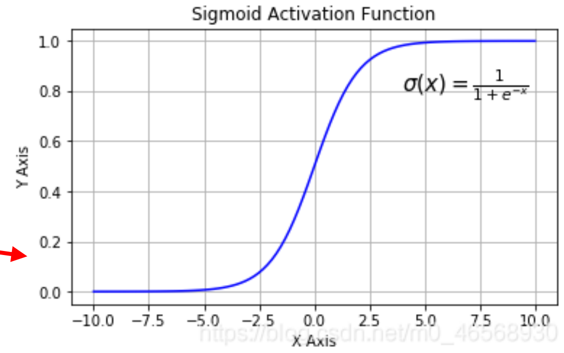
linear

Row

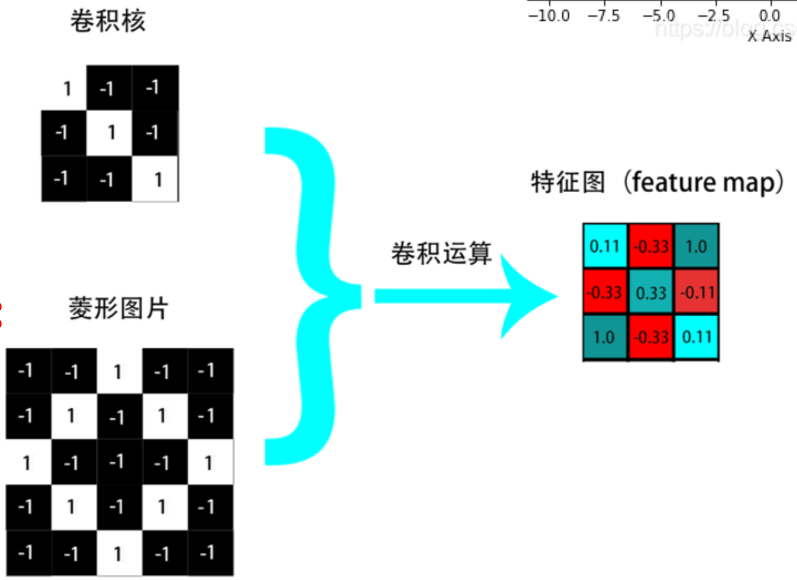
lumn

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2j} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3j} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

Non-linear



CNN (convolutional neural network) :



Why attempt to use NN?

Advantage:

- Future detector reconstruction trend, lots of related articles published in recent few years
- Method is simple, easy for trying
- Could improve the reconstruction performance
- Could deal with complicated logic automatically

Disadvantage:

- The module from training only works for specific data.
- Need amounts of events for training
- The training module influence result, need to tune (work experience)
- CPU/GPU consuming

ATLAS and CMS CNN & GAN （generative adversarial network) to generate ECal shower

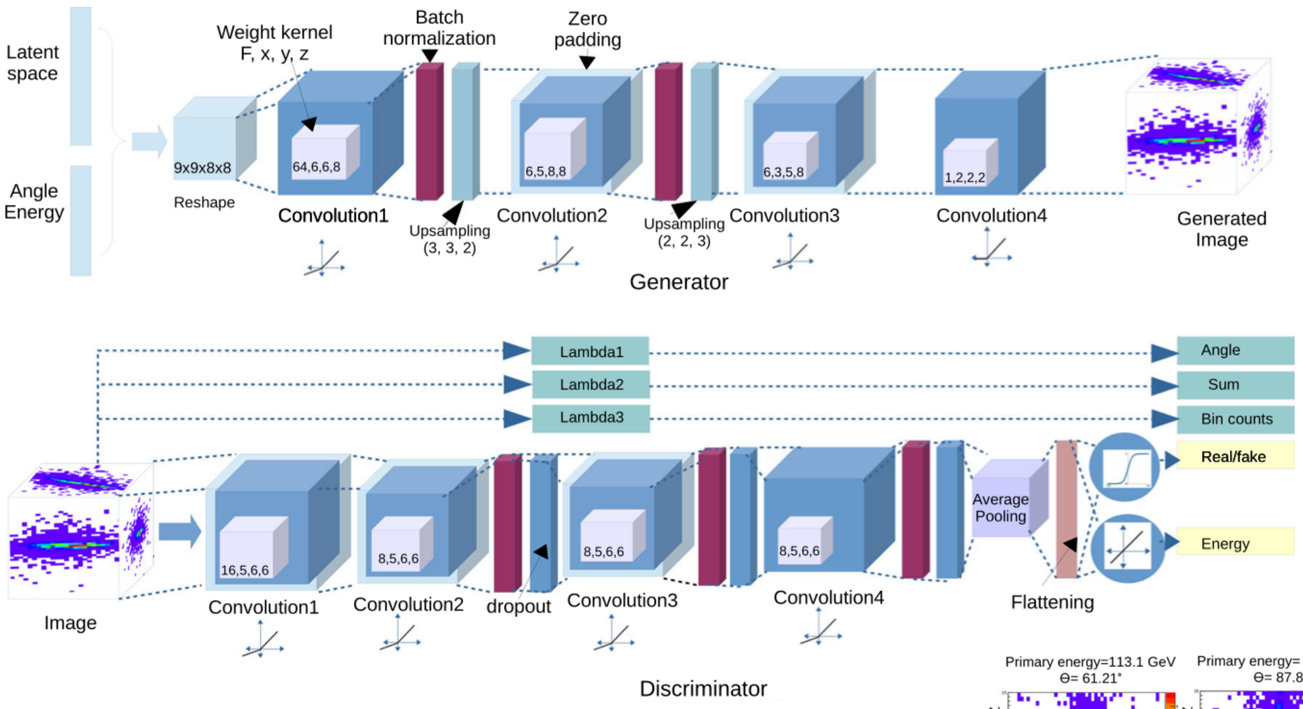
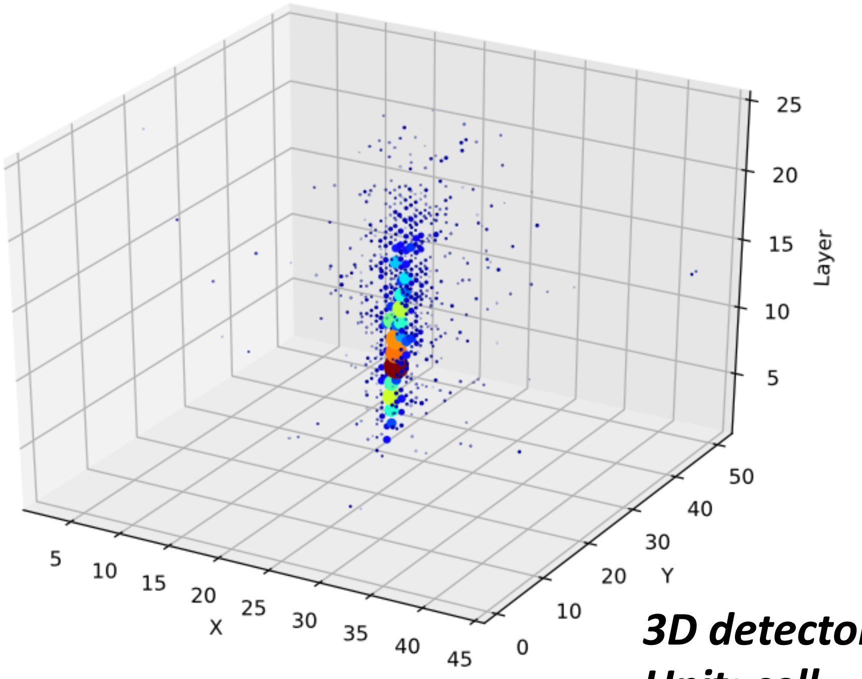


Fig. 6 3DGAN generator and discriminator network architectures



3D detector, Unit: cell

Simulate 100-500GeV e- in ECal

- **Geant4: 17s**
- **GAN: 4ms**(GPU,NVDIA 1080)

For EicC: 0.06s/1 GeV e-, time will be > x100 lower if adding optical transportation…

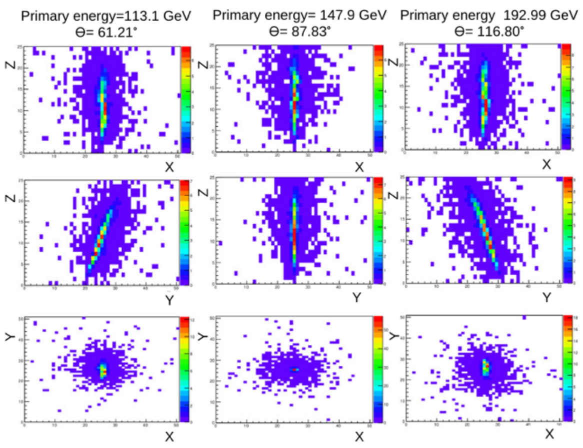


Fig. 7 GEN sample: electrons with different primary particle energies and angles

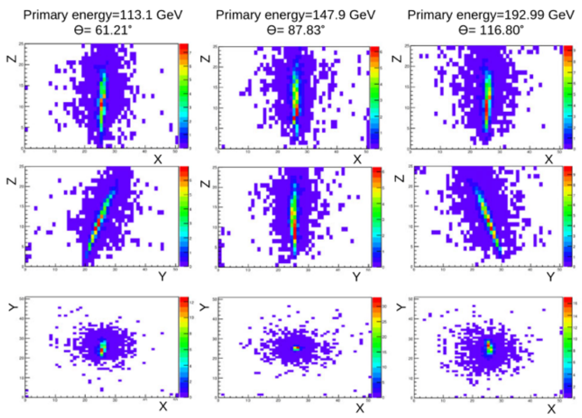
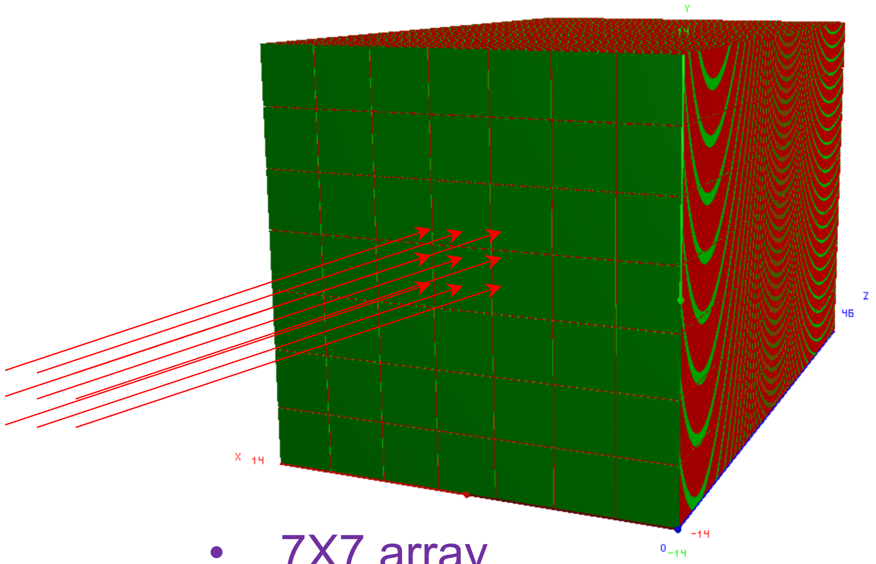


Fig. 8 GAN generated electrons with primary energies and angles corresponding to the electrons showed in Fig. 7

NN method of ECal introduction

- 7*7 Shashlik array simulation, 4*4 cm vertical area source
- Train input: N.P.E. of 49 modules
- Output ： energy, position, incident angle, PID
    - **Energy resolution**
    - **Position resolution**
    - **PID: e/pi separation**



- 7X7 array
- 4X4 cm^2 area source

- Attempt to use neural network for reconstruction
- CNN (convolutional neural network) is NOT applied since no improvement(only tried one time).

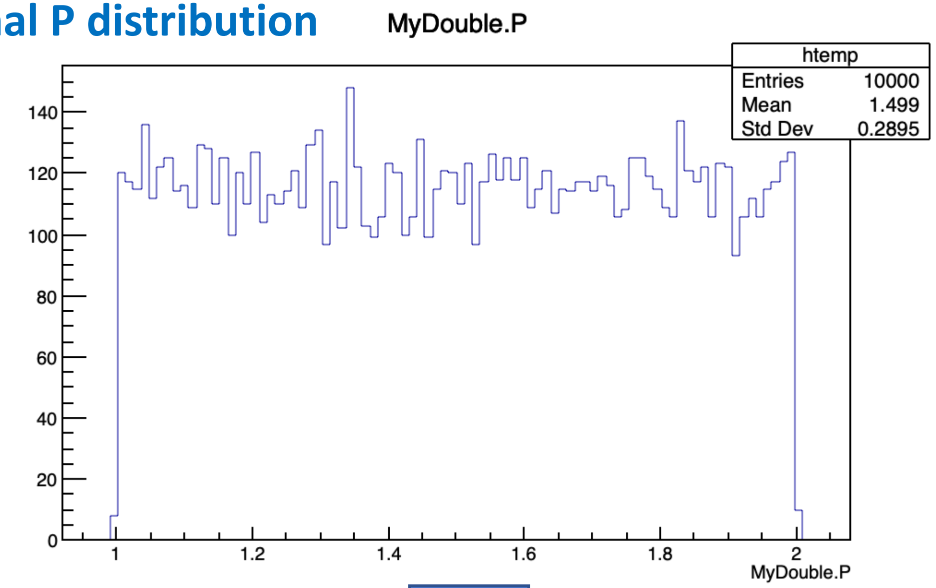
NN structure:

```
self.fc = nn.Sequential(
    nn.Linear(49, 80), #bias is true for default
    nn.ReLU(True),
    nn.Linear(80, 100), #bias is true for default
    nn.ReLU(True),
    nn.Linear(100, 50),
    nn.ReLU(True),
    nn.Linear(50, 15),
    nn.ReLU(True),
    nn.Linear(15, 5),
    nn.ReLU(True),
    nn.Linear(5, n_class)
)
```

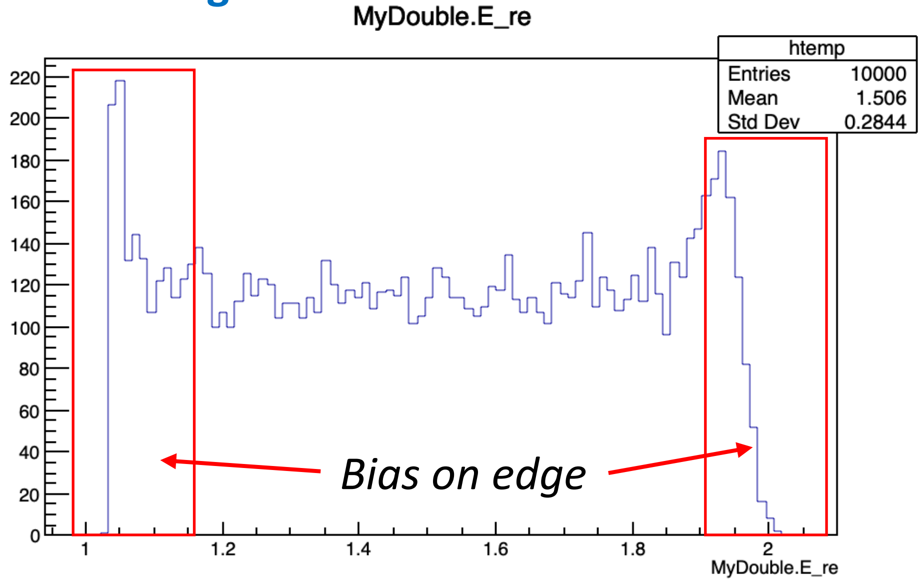
- Train situation: local MacOS, single CPU core, python(PyTorch), training data from simulation

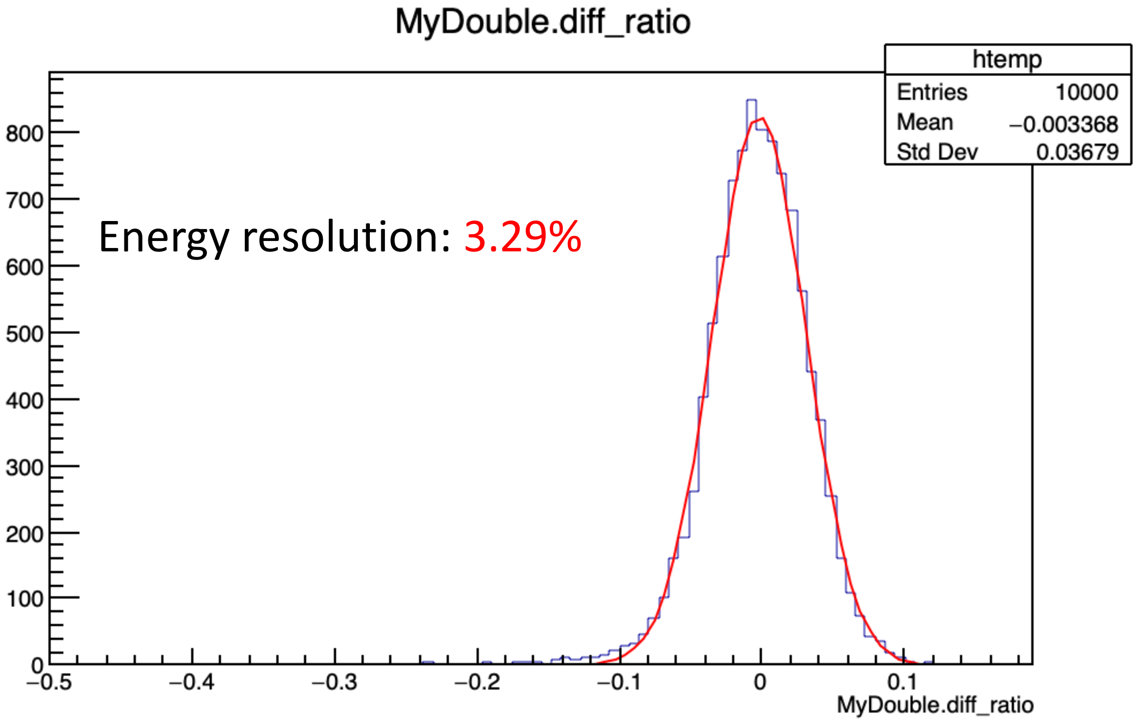
Energy reconstruction (1-2 GeV e-)

Original P distribution

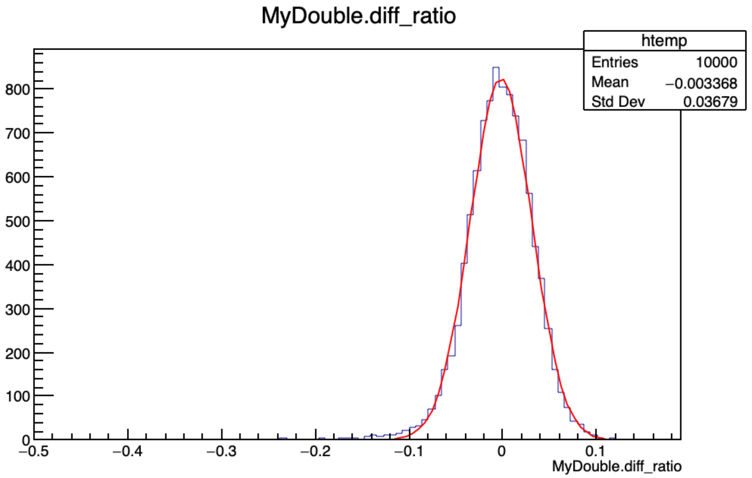


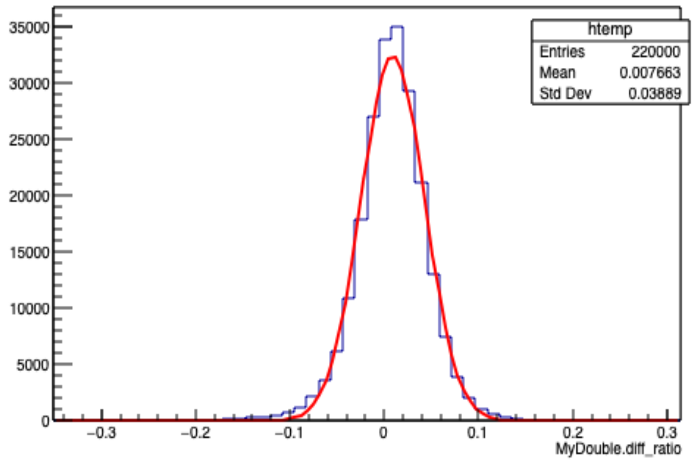
After NN training

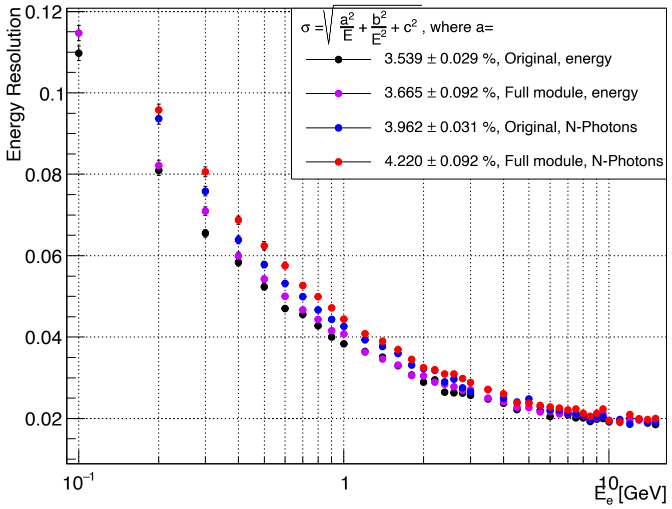




Energy resolution compared with other result







NN training with **[1 GeV, 2 GeV]:**
- 1-2 GeV average: 3.29%

NN training with **[0.5 GeV, 5 GeV]**:
- 0.6 GeV: 5.9%
- 1 GeV: 4.5%
- 2 GeV: 3.5%
- 4 GeV: 2.53%

Add all blocks simply for each energy point: (ordinary method)
- 0.6 GeV: 5.4%
- 1 GeV: 4.25%
- 2 GeV: 3.2%
- 4 GeV: 2.5%

- Larger training energy range lead to worse energy resolution.
- No significant improvement compared to ordinary method.

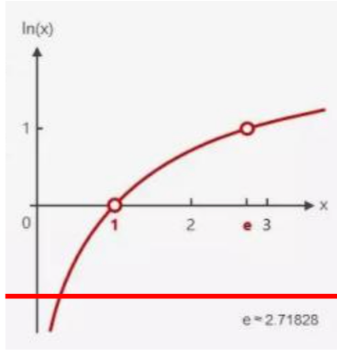
w0 parameter optimization for position reconstruction

The w0 parameter value 4.2 is acquired from article for crystal ECal, and need to be adjusted and optimized for Shashlik.

In earlier studies position corrections were applied to remove this bias (the so-called S-shape correction [6]), but we now use a simpler algorithm which delivers almost as good precision by calculating the weighted mean using the logarithm of the crystal energy:

$$x = \frac{\sum x_i \bullet W_i}{\sum W_i}$$

where x_i the position of crystal i, and W_i is the log weight of the crystal — the log of the fraction of the cluster energy contained in the crystal, calculated with the formula:

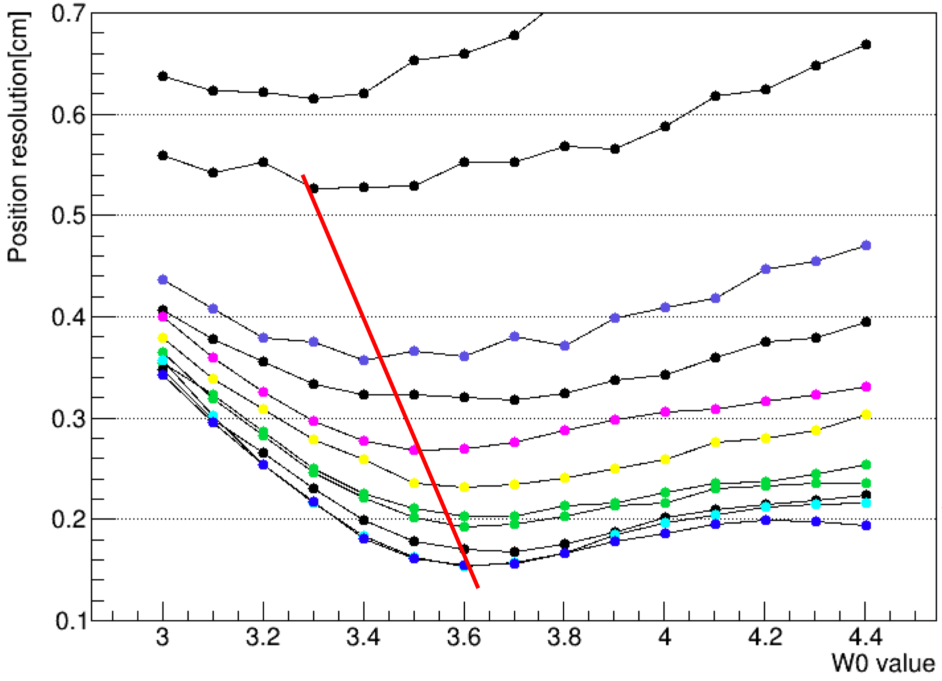


Exp (-W_0)

$$W_i = W_0 + \log\left(\frac{E_i}{\sum E_j}\right)$$

The W_0 value need to be tuned.
For W_0=4.2, equivalent to E_i/E>0.015.

W0=3.5 equivalent to E_i/E>0.03

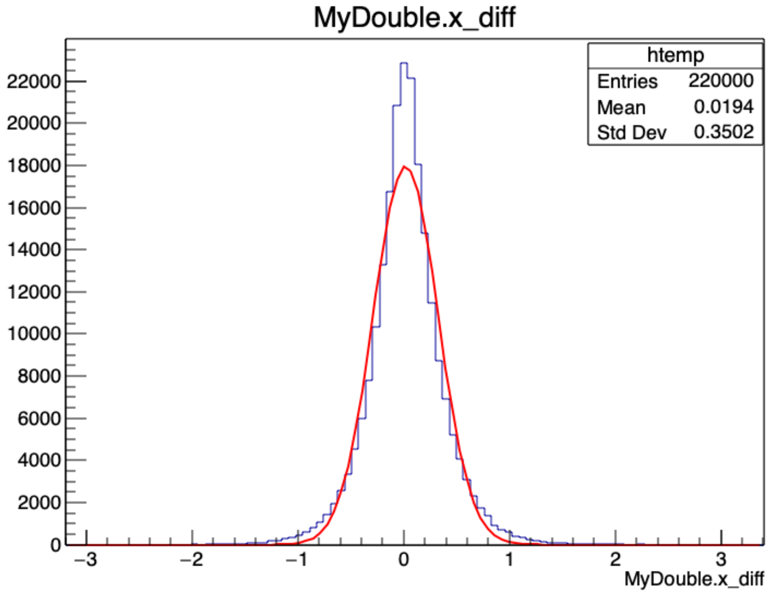


0.6 -15 GeV

w0 optimization value relevant to electron energy.

Position resolution with NN method (0.5-5 GeV)

Both x and y as output simultaneously, and same resolution: 3.07mm



Use NN trained model to test single energy data: [mm]

- 0.6 GeV: 5.77
- 1 GeV: 4.6
- 2 GeV: 3.1
- 4 GeV: 2.52

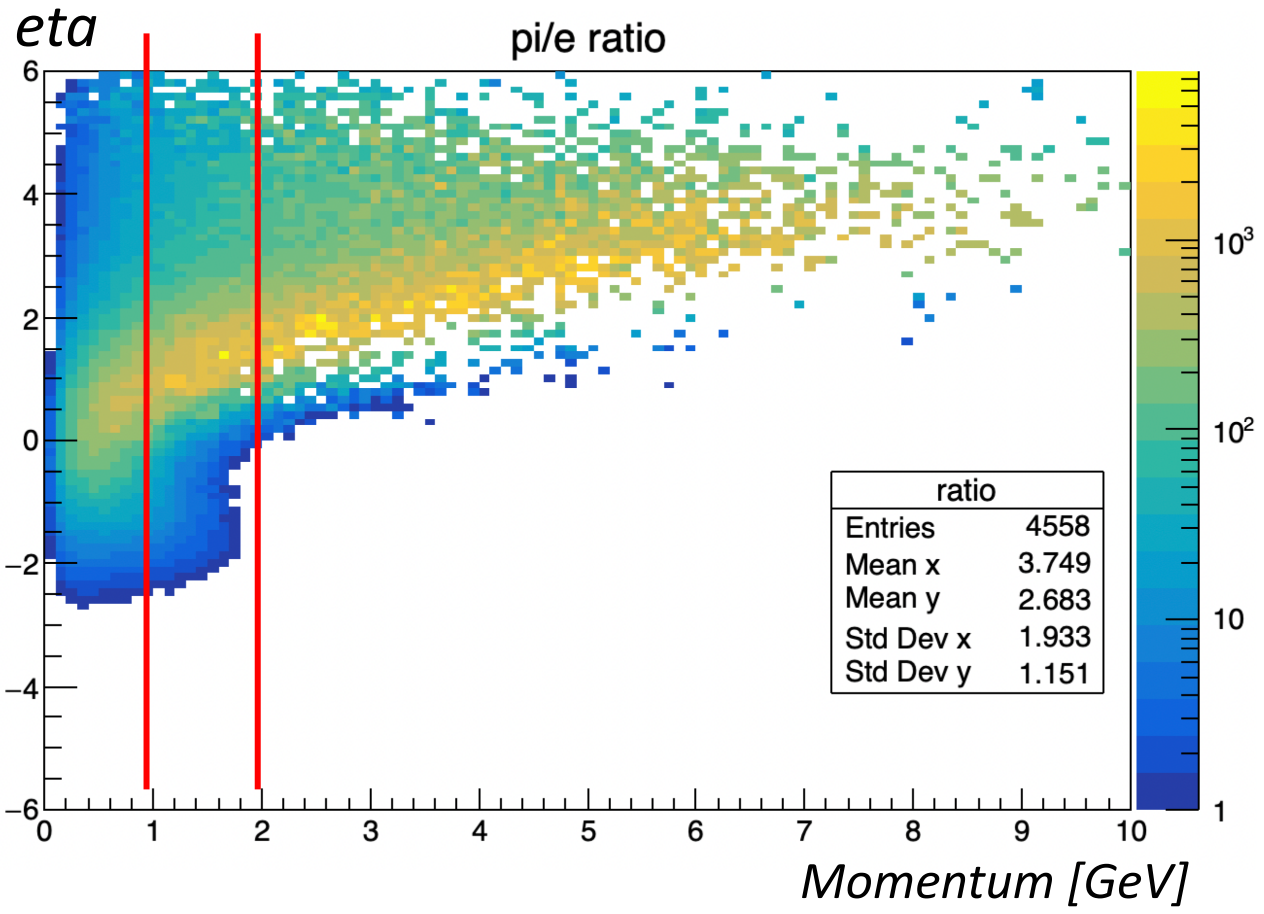
Compared with **the plot shown above**: [mm]

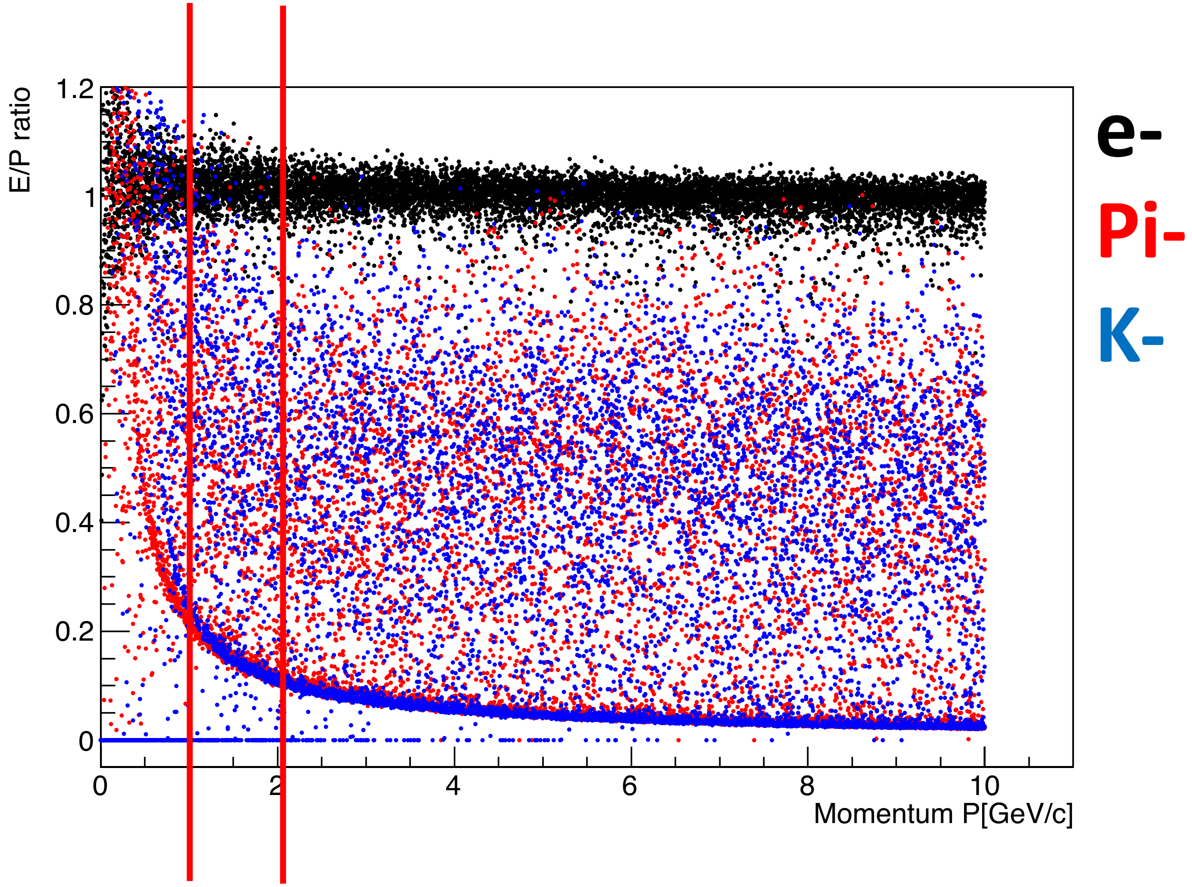
- 0.6 GeV: 6.2
- 1 GeV: 5.3
- 2 GeV: 3.65
- 4 GeV: 2.7

- NN method for position reconstruction is much better than w0 method.
- May exist other better "theoretical calculation" method!

PID: e/π separation of ECal

Ratio from pythia





e-/π- separation with NN method

- Training input: N.P.E. of 49 blocks + REAL momentum (total 50 parameters)
- Training output: the possibility of e and π, the particle with larger possibility is regarded as output.

- Training data: 200k, 1-2 GeV e and π separately
- Result: （10k data for test）

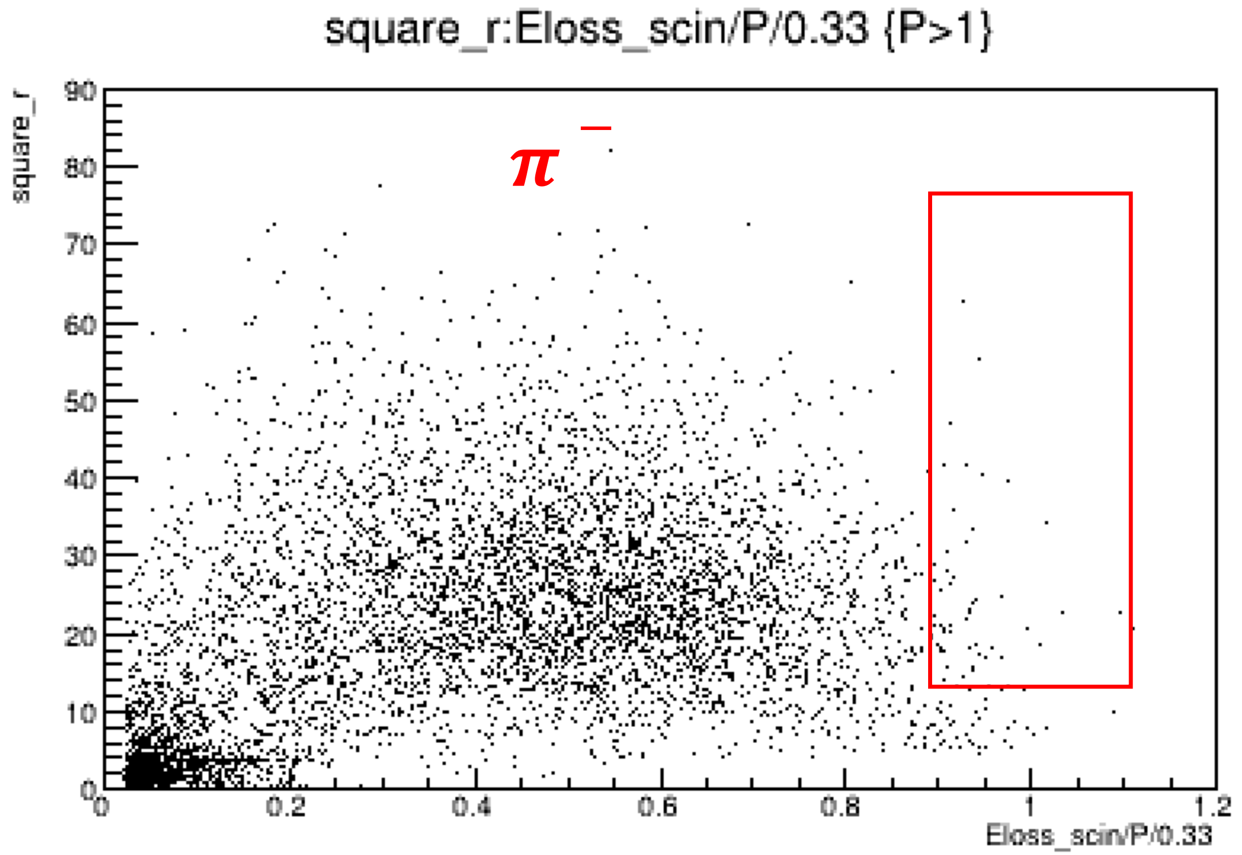
percentage （%）		Real PID	
		e	π
PID result	e	99.2	2.3
	π	0.8	97.7

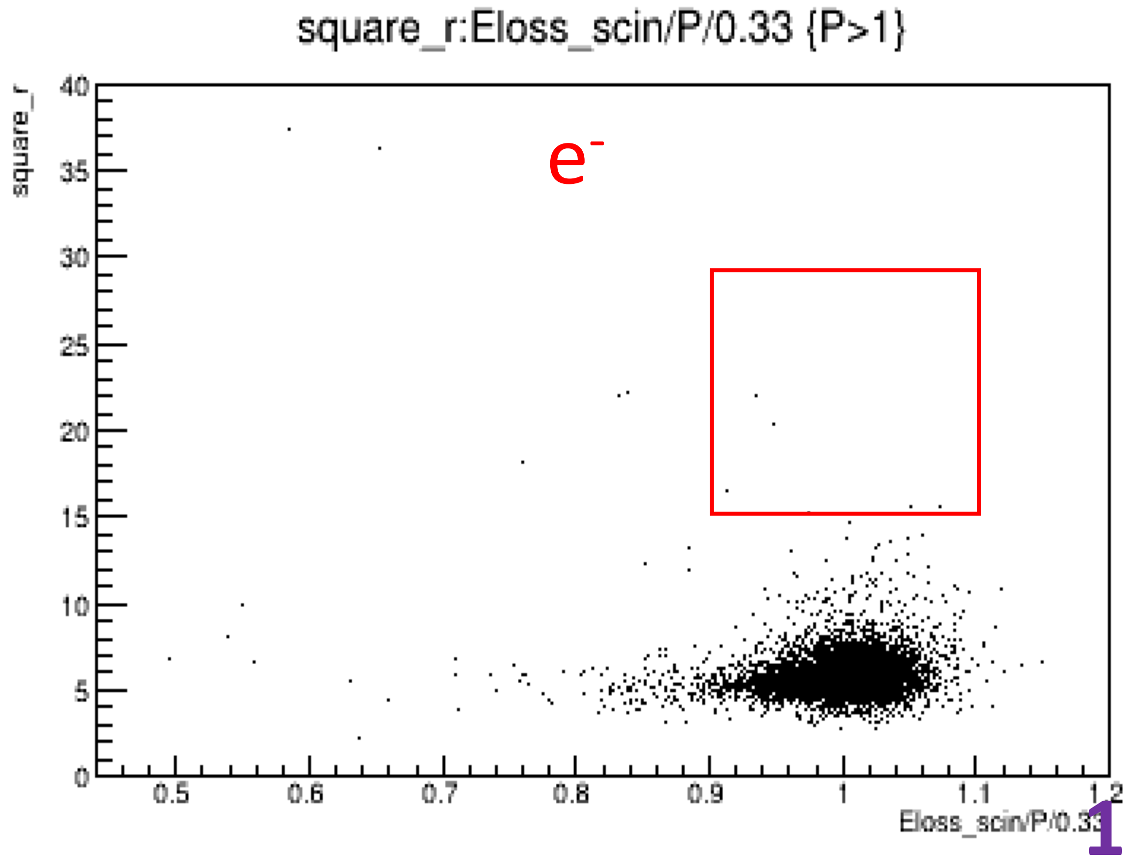
Use shower spread radius to separate e/π (R cut)

Principle: hadron shower has wider shape.

$$V(x) = \sum (x_i - x_{re})^2 \frac{w_i}{w_{all}}$$

$$V(r) = \sqrt{V(x)^2 + V(y)^2}$$

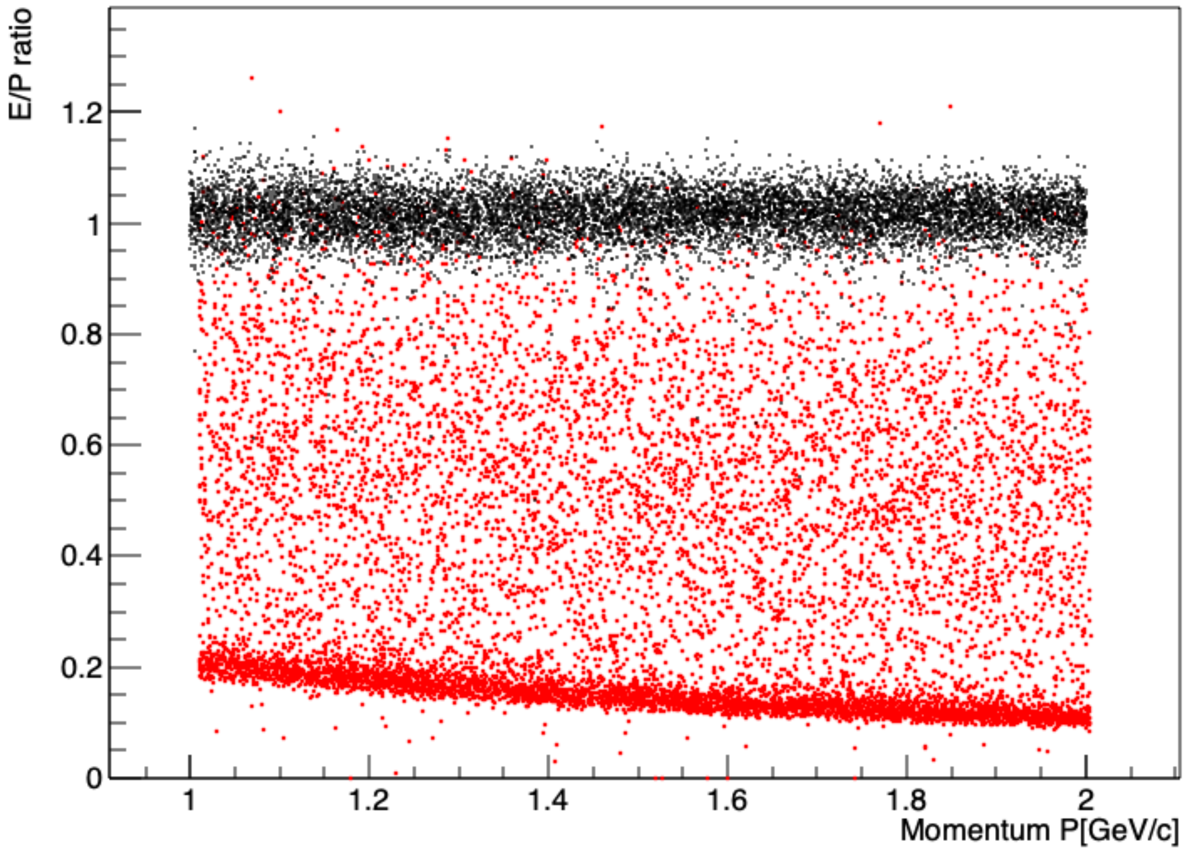


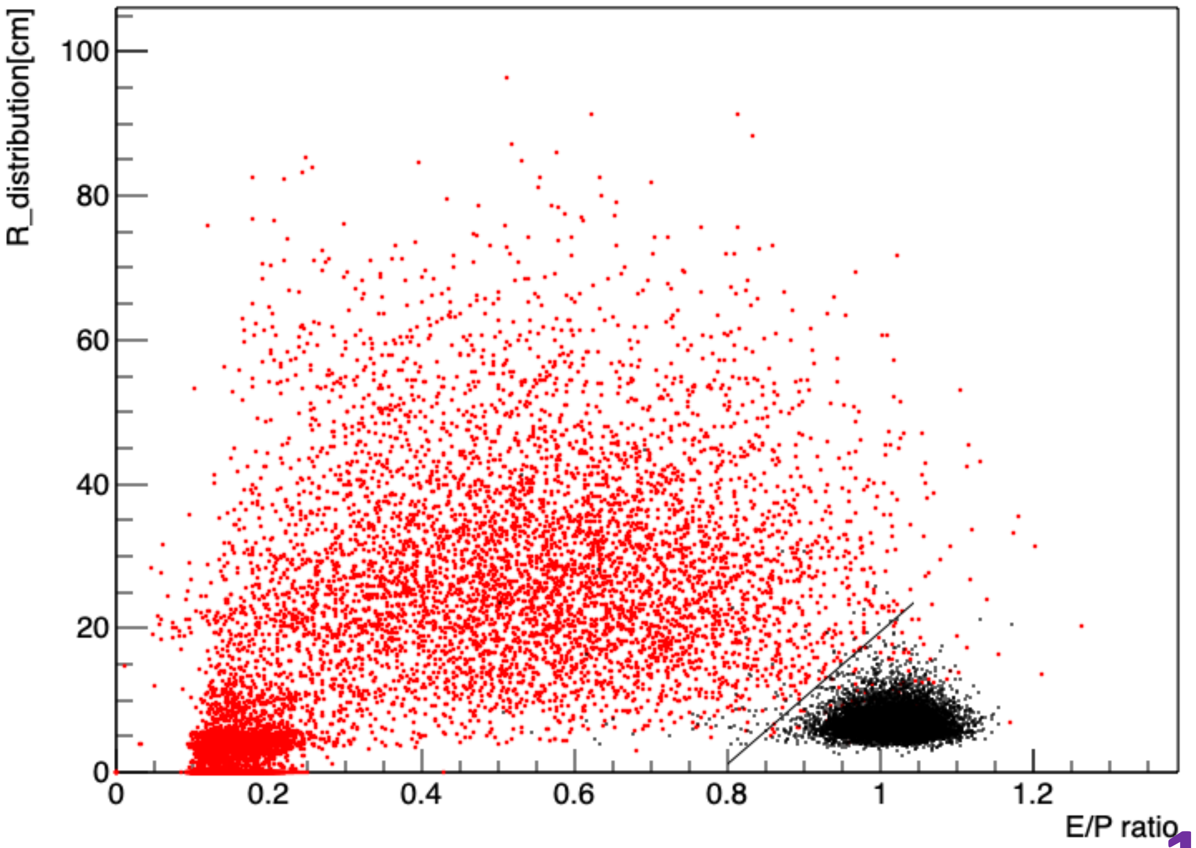


Considering both E/P and R cut for 1-2 GeV/c e/π

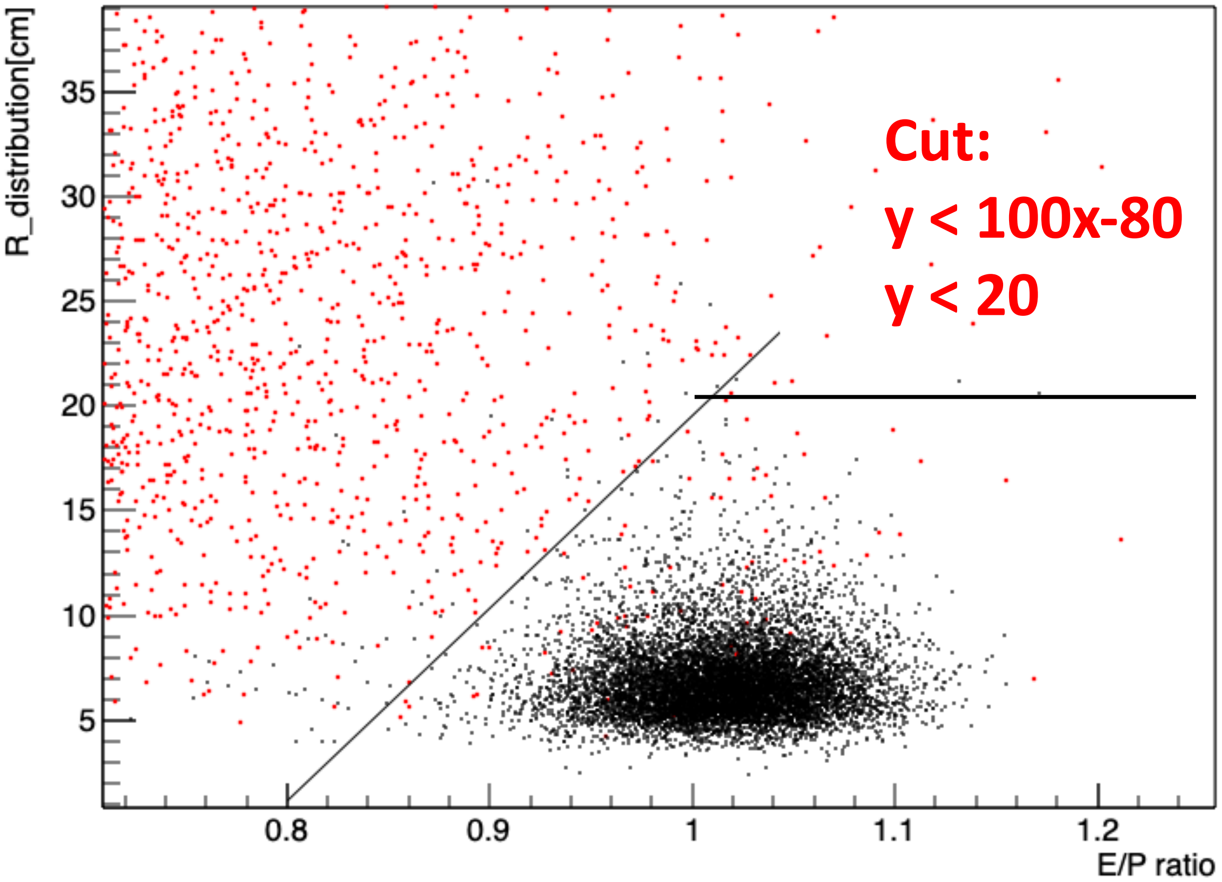
$$V(x) = \sum (x_i - x_{re})^2 \frac{w_i}{w_{all}}$$

$$V(r) = \sqrt{V(x)^2 + V(y)^2}$$





Artificial cut selection and PID result (1-2 GeV/c)



Percentage (%) (10k events)		Real PID	
		e	π
PID result	e	99.2	0.7
	π	0.8	99.3

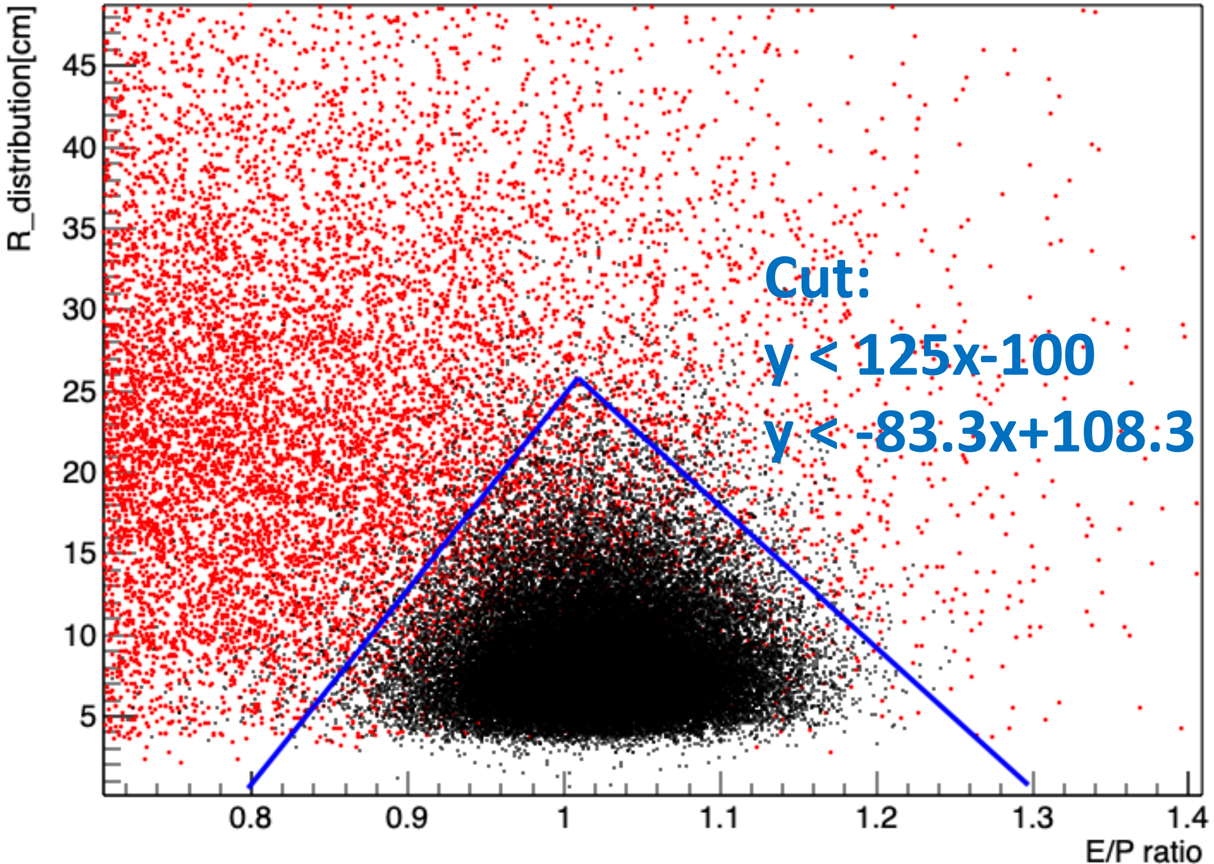
Percentage (%) (200k events)		Real PID	
		e	π
PID result	e	99.25	0.70
	π	0.75	99.30

- Artificial cut is better than NN method for less π
- Good result for PID: e efficiency > 99%, π rejection ~ 100:1
- If considering the real e/π ratio, cut selection and PID result will be much different
- Shower method and MIP/time method result for e/π is independent

Considering both E/P and R cut for 0.5-1 GeV/c e/π

$$P_{\perp min} = r * H * \frac{\frac{e}{c}}{P \to m} = 0.45 * 1.5 * \frac{\frac{1.6}{3}}{1.78} = \mathbf{0.202}\,\frac{\mathbf{GeV}}{\mathbf{c}}$$

If a particle has $P_{\perp}$ less than 0.202 GeV/c, it will never hit barrel, but could hit the endcap.



Percentage (%) (210k events)		Real PID	
		e	π
PID result	e	97.4	3.5
	π	2.6	96.5

Low momentum PID get worse, but not too worse.

NN method summary

- Attempt to apply NN method, which should has better result in principle, and has its advantage
- NN has no artificial bias for reconstruction(any cut and method have bias)

Application in ECal (up to now):

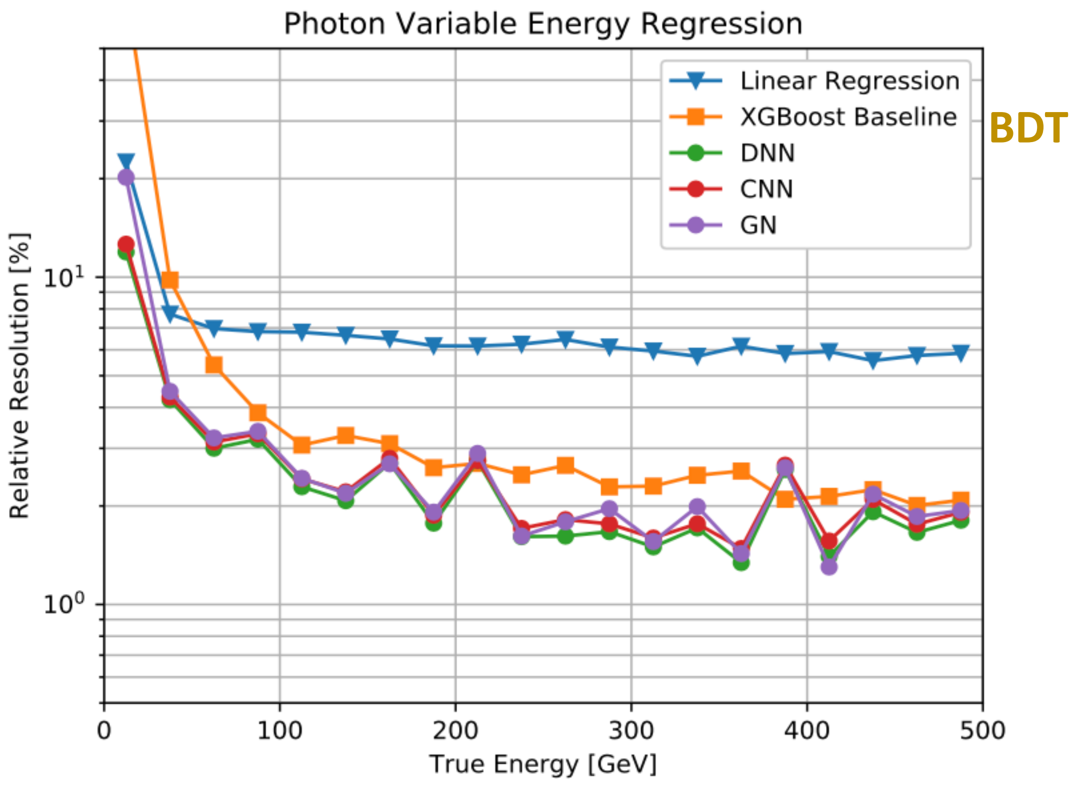
- Significant improvement
  - **Position**(continuous value)
- No improvement and even worse: (result not too bad, could attempt to tune)
  - **Energy** (continuous value)
  - **e/π PID** (separate value)

Other application:

- Crystal ECal
- Incident angle reconstruction

Energy resolution with different method

Simple geometry



ATLAS like geometry