



Study of Energy Reconstruction with Machine Learning

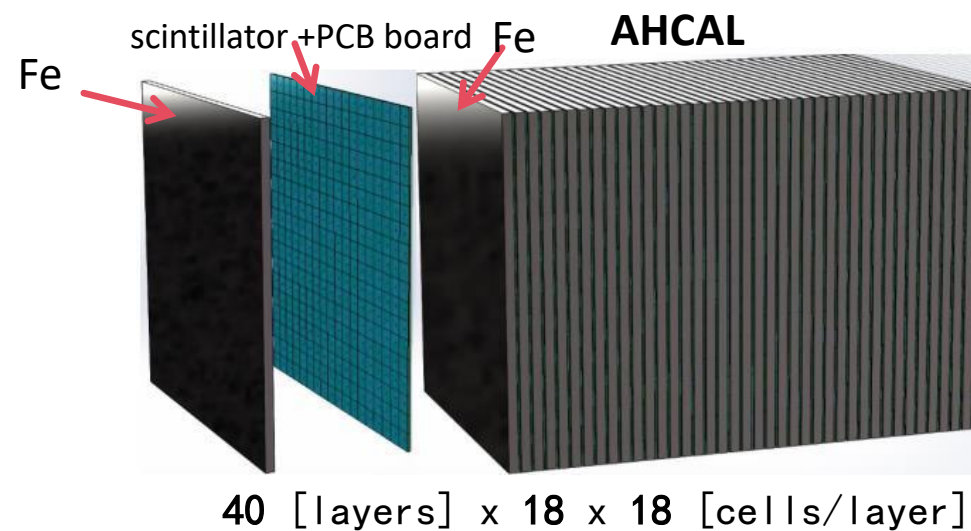
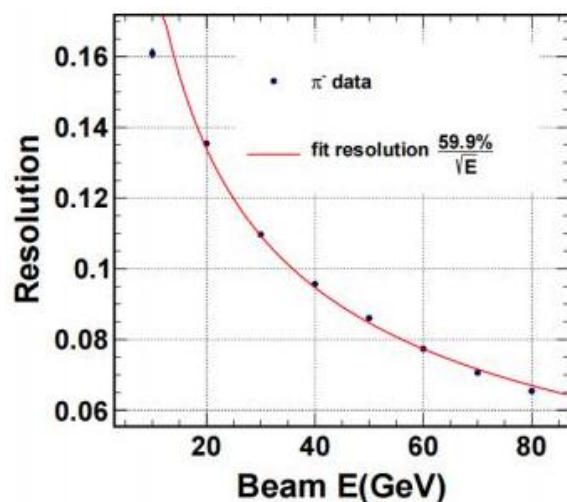
Speaker: Peng Borong

University of Science and Technology of China

2026.10.9

Background

- **CEPC A(nalog) H(adronic) CAL(orimeter)**
- 40 sensitive layers (iron plate + scintillator layer + PCB board + iron plate)
- Energy resolution meets the targets $\left(\frac{60\%}{\sqrt{E} \text{ (GeV)}} \right)$
- Machine Learning——>Improve the Energy Reconstruction Accuracy



plastic scintillator



Beam Tests

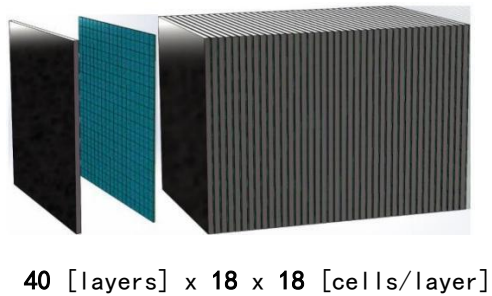
3 beam tests have been done at CERN

- muon position scan(100GeV/c)
- pion (1-120 GeV/c)
- electron (1-120 GeV/c)

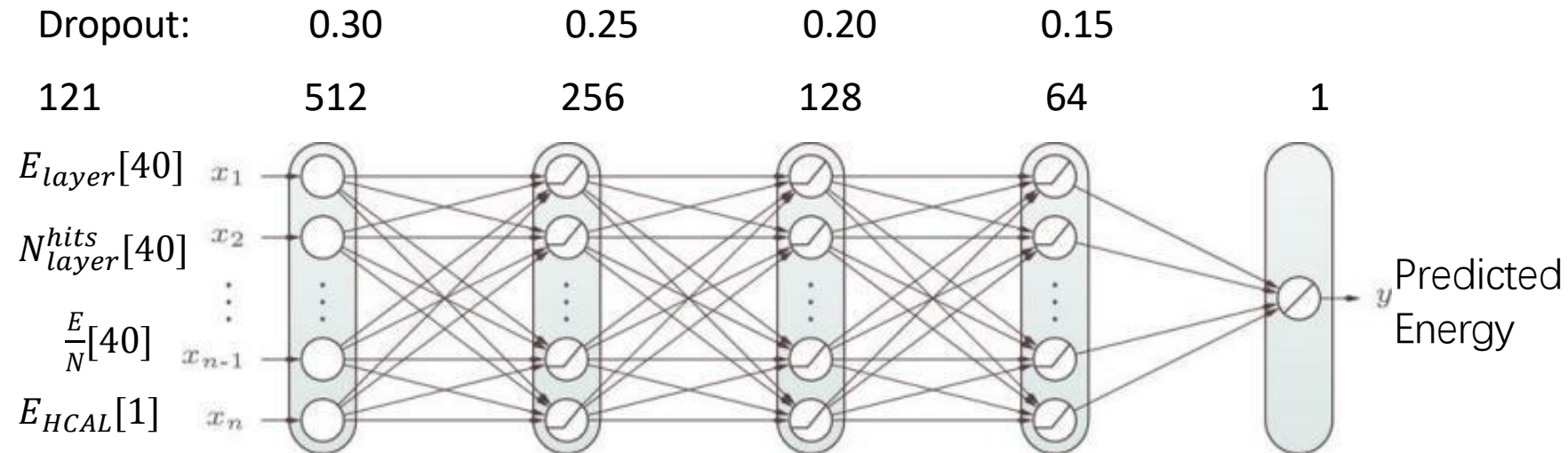


MLP Model

- Input : layer energy(E),layer hits(N),average hit energies(E/N) of 40 layers and total energy in the HCAL(E_{HCAL})
- Output:Predicted energy (1) (40*3+1=121)



Batch_Size	64
Learning Rate	1e-4
Epochs	200



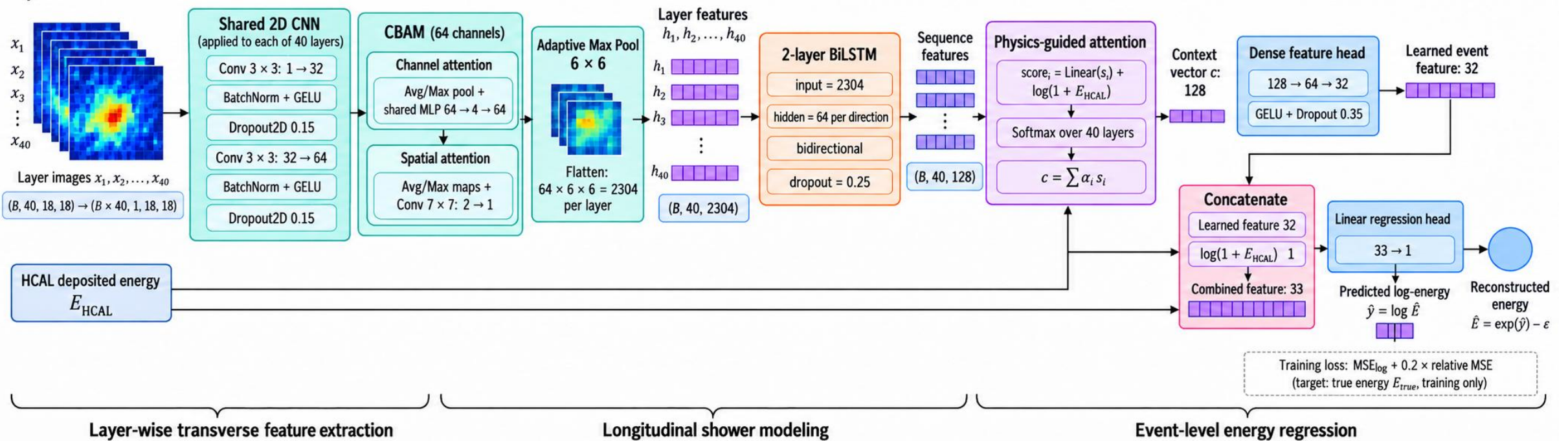
Loss: $L = \frac{1}{N} \sum_n e_n$ $e = \frac{(y_{true} - y_{pred})^2}{y_{true}^2}$ Mean Squared Relative Error (MSRE)

CNN Model

$X \in R^{40 \times 18 \times 18} \rightarrow \text{Reconstructed energy } E$

2D CNN + CBAM + LSTM + Physics-Guided Attention

Input: AHCAL shower $40 \times 18 \times 18$



ParticleNet Model

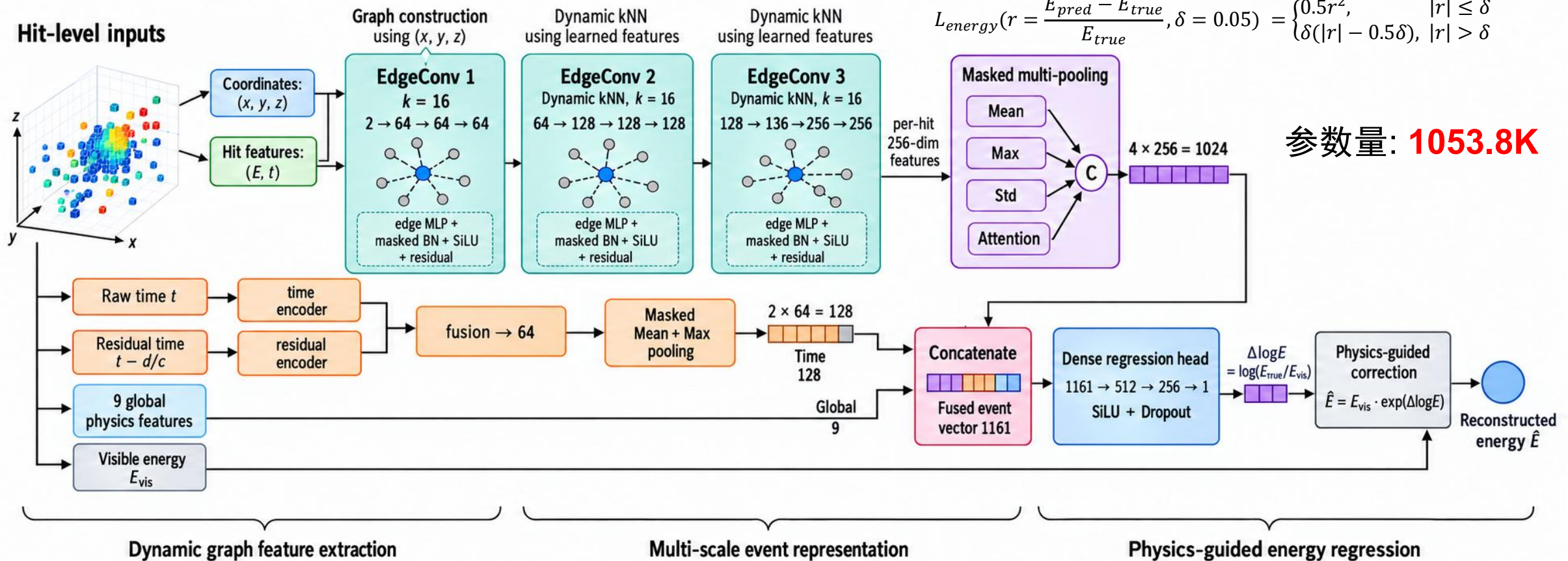
$F_{PN}: \{(x_i, y_i, z_i, E_i, t_i)\}_{i=1}^N \rightarrow \mathbb{R}$ **EdgeConv + T-MLP → Fully Connected → Reconstructed energy E**

particle types : π^+

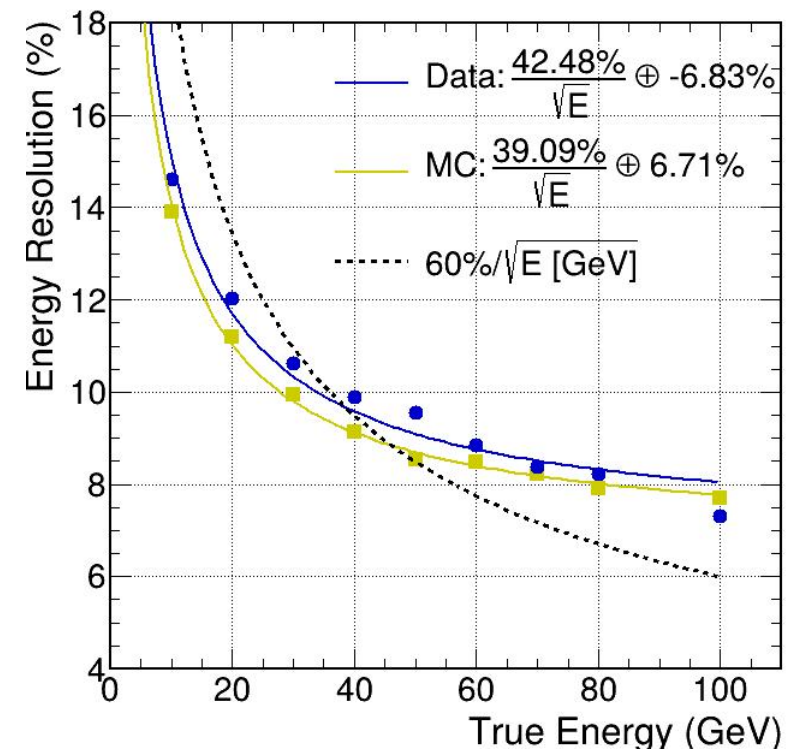
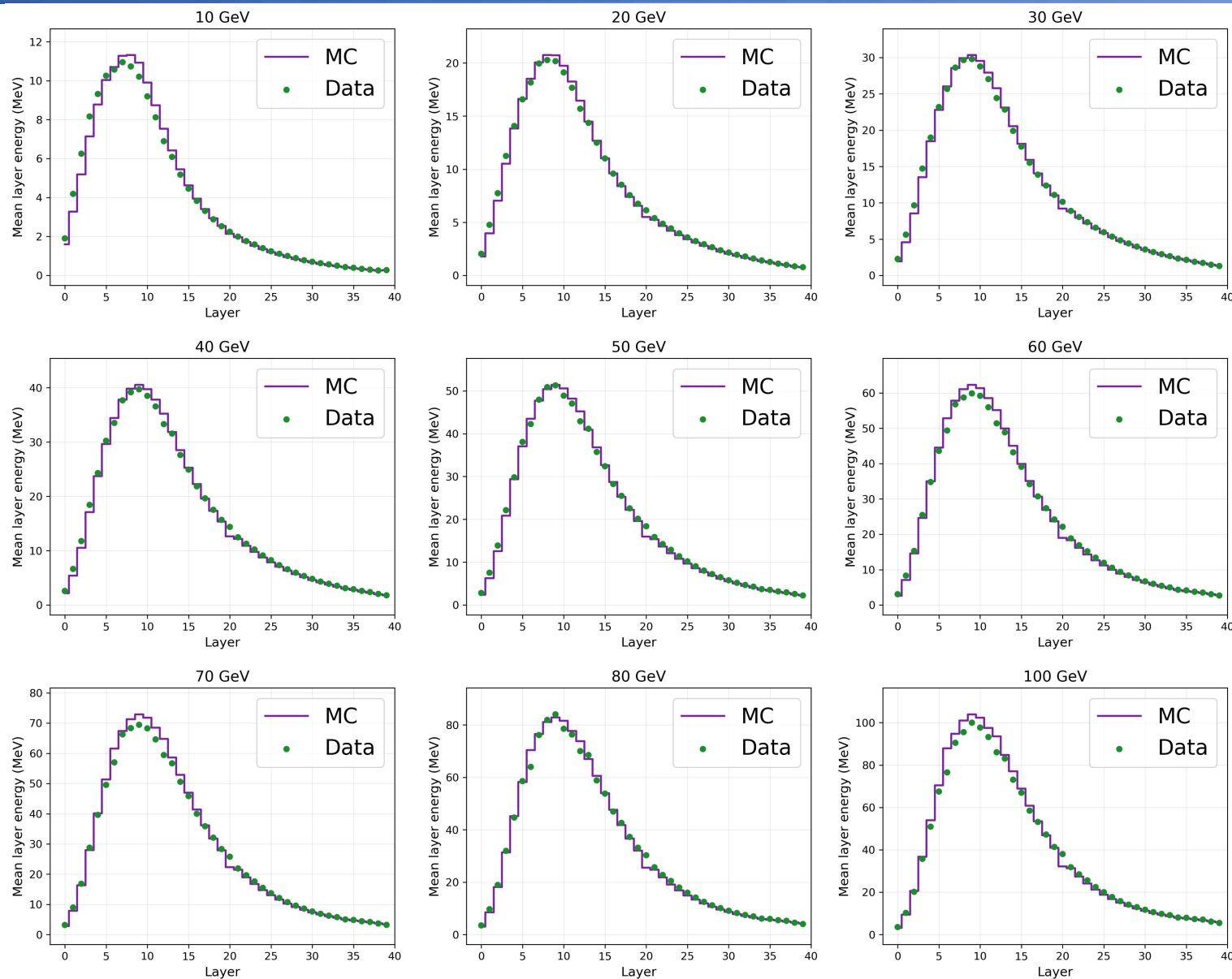
loss: $L_{PN} = 0.7L_{corr} + 0.3L_{energy}$

$$L_{corr}(x = \log \frac{E_{pred}}{E_{true}}, \beta = 0.08) = \begin{cases} 0.5x^2, & |x| \leq \beta \\ \beta(|x| - 0.5\beta), & |x| > \beta \end{cases}$$

$$L_{energy}(r = \frac{E_{pred} - E_{true}}{E_{true}}, \delta = 0.05) = \begin{cases} 0.5r^2, & |r| \leq \delta \\ \delta(|r| - 0.5\delta), & |r| > \delta \end{cases}$$

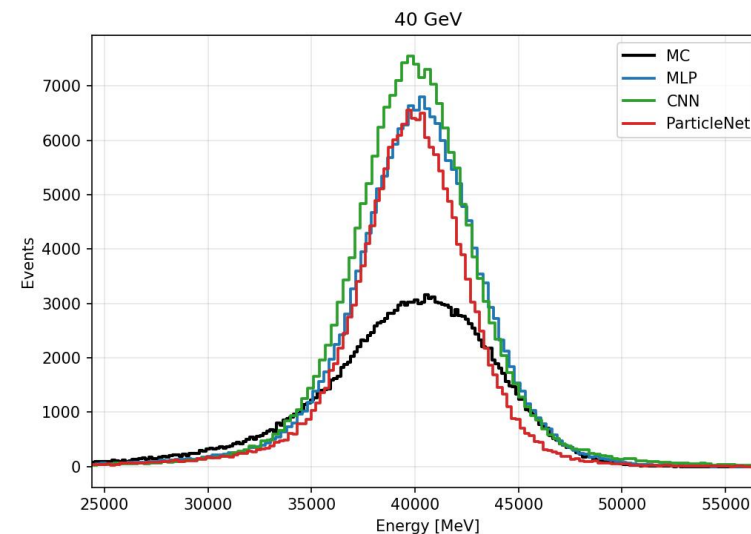
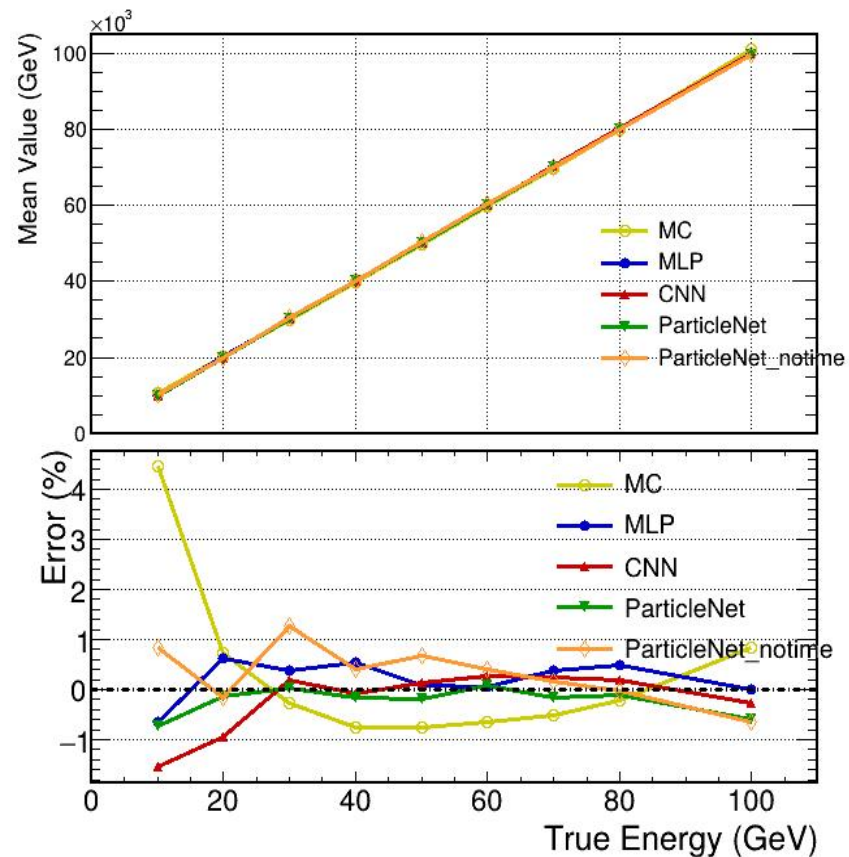
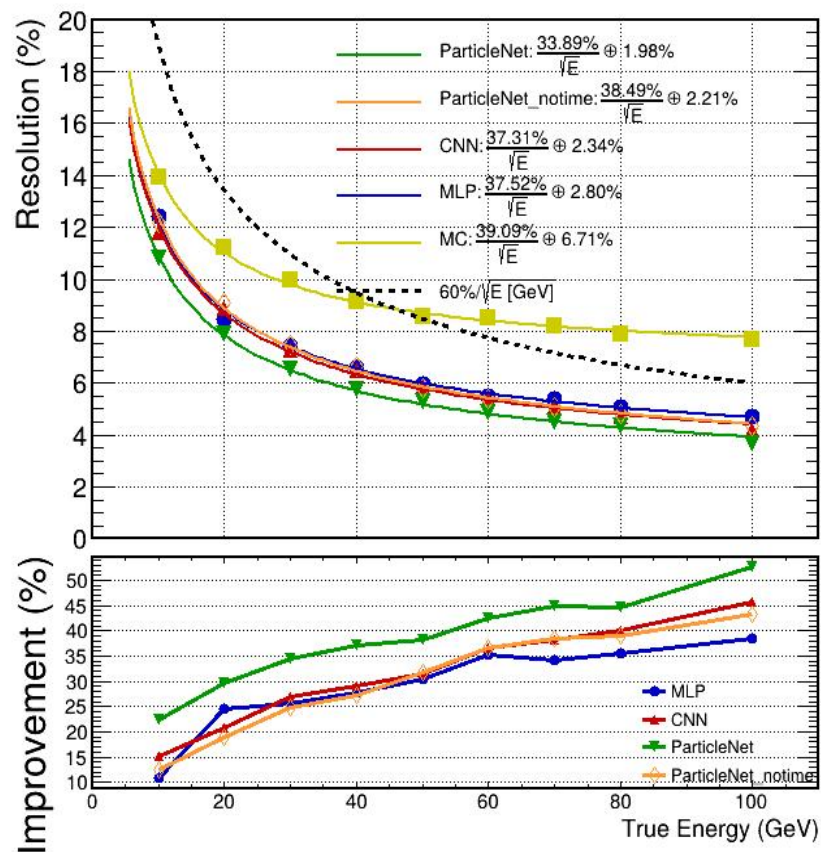


MC and Data



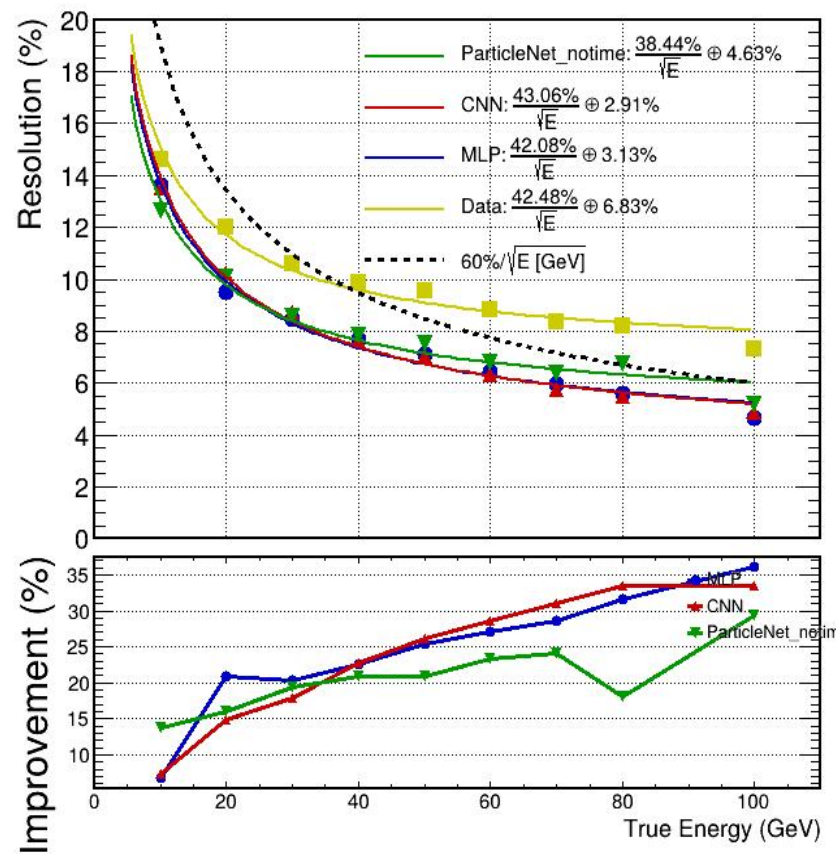
Results on MC

Resolution

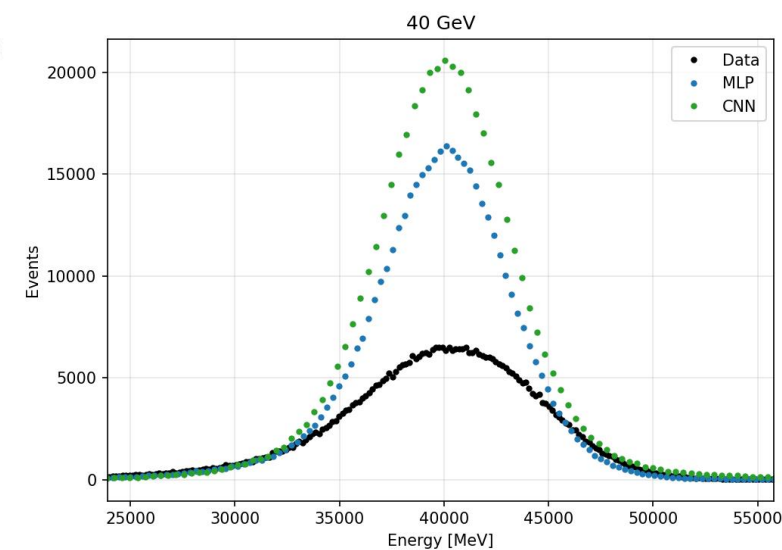
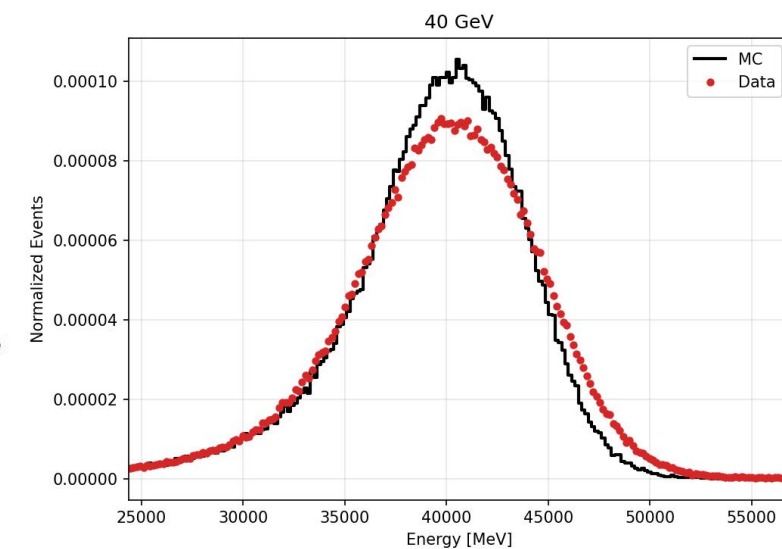
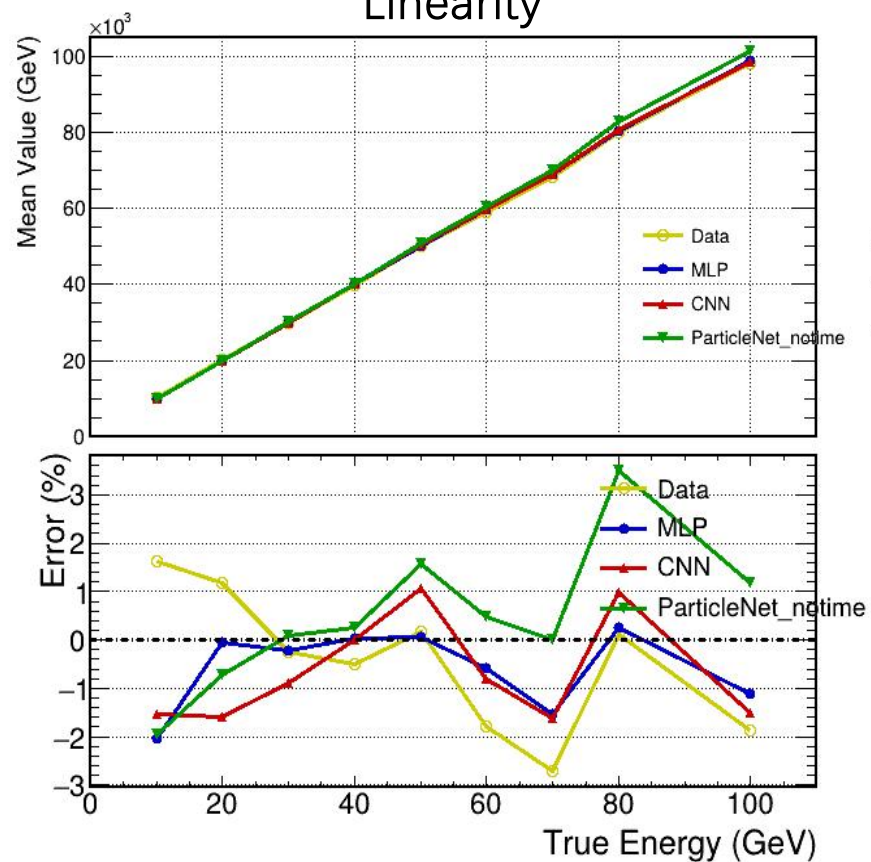


Results on Data

Resolution



Linearity



Summary

- AHCAL : 40 layers $\times$ 18 $\times$ 18 cells/layer

- Model : MLP CNN ParticleNet

- MC results:

Improvement in resolution: ParticleNet(33.4% ,avg) > CNN(26.7% ,avg) $\approx$ MLP(23.8% ,avg)

Error in linearity: ParticleNet(0.29% ,avg) < MLP(0.88% ,avg) < CNN(1.28% ,avg)

- Beam data(pion, 1-120 GeV/c) results:

Improvement in resolution: CNN(24.8% ,avg) $\approx$ MLP(25.2% ,avg)

Error in linearity: MLP(0.64% ,avg) < CNN(1.11% ,avg)

- ParticleNet: Major Advantage on AHCAL !

Thanks for your attention!

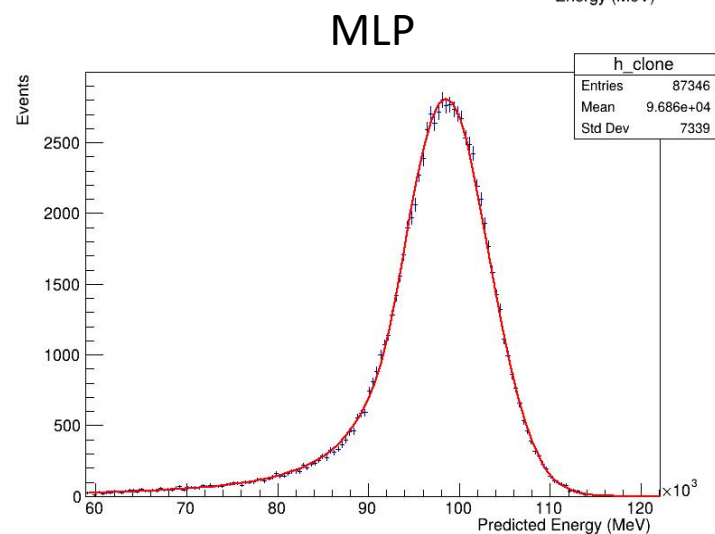
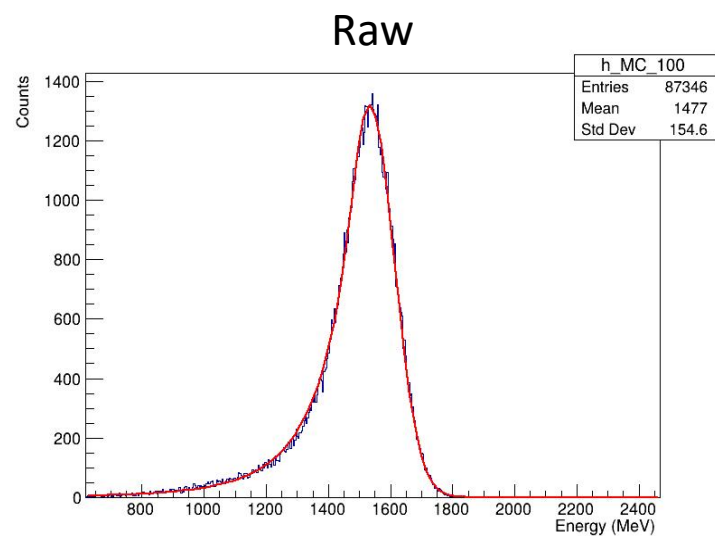
back up

back up

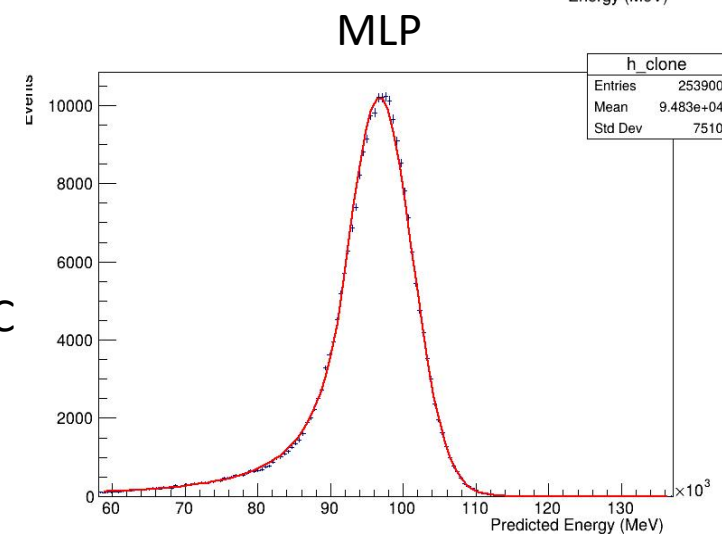
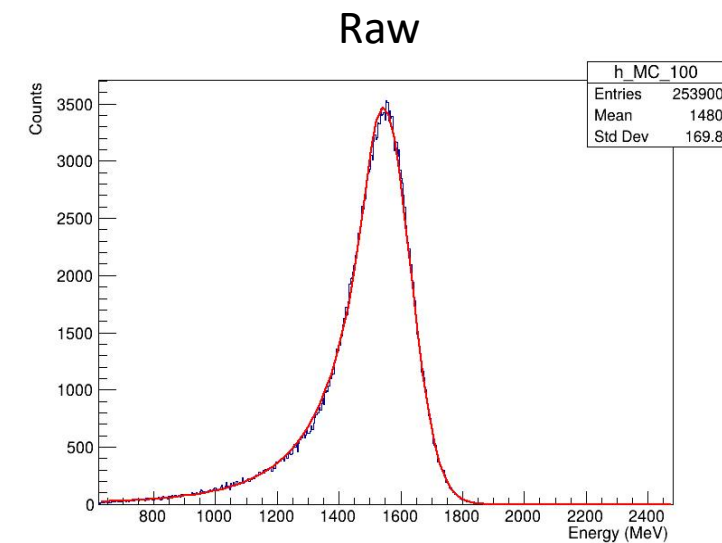
event	energy(G eV)	10	20	30	40	50	60	70	80	100
Data	Pre-cut	1,027,814	749,791	1,081,725	1,038,984	1,258,699	1,034,938	1,058,822	1,073,061	1,286,608
	Post-cut	83,860	302,303	486,653	458,874	498,094	369,281	327,979	283,120	253,900
	Fraction (%)	8.16	40.32	44.99	44.17	39.57	35.68	30.98	26.38	19.73
MC	Pre-cut	400,000	400,000	400,000	400,000	400,000	400,000	400,000	400,000	394,000
	Post-cut	201,676	211,115	199,497	186,119	169,367	152,117	133,628	116,822	87,346
	Fraction (%)	50.42	52.78	49.87	46.53	42.34	38.03	33.41	29.21	22.17
Train : 703,747					Test : 124,035					

Energy Spectrum@100GeV

MC

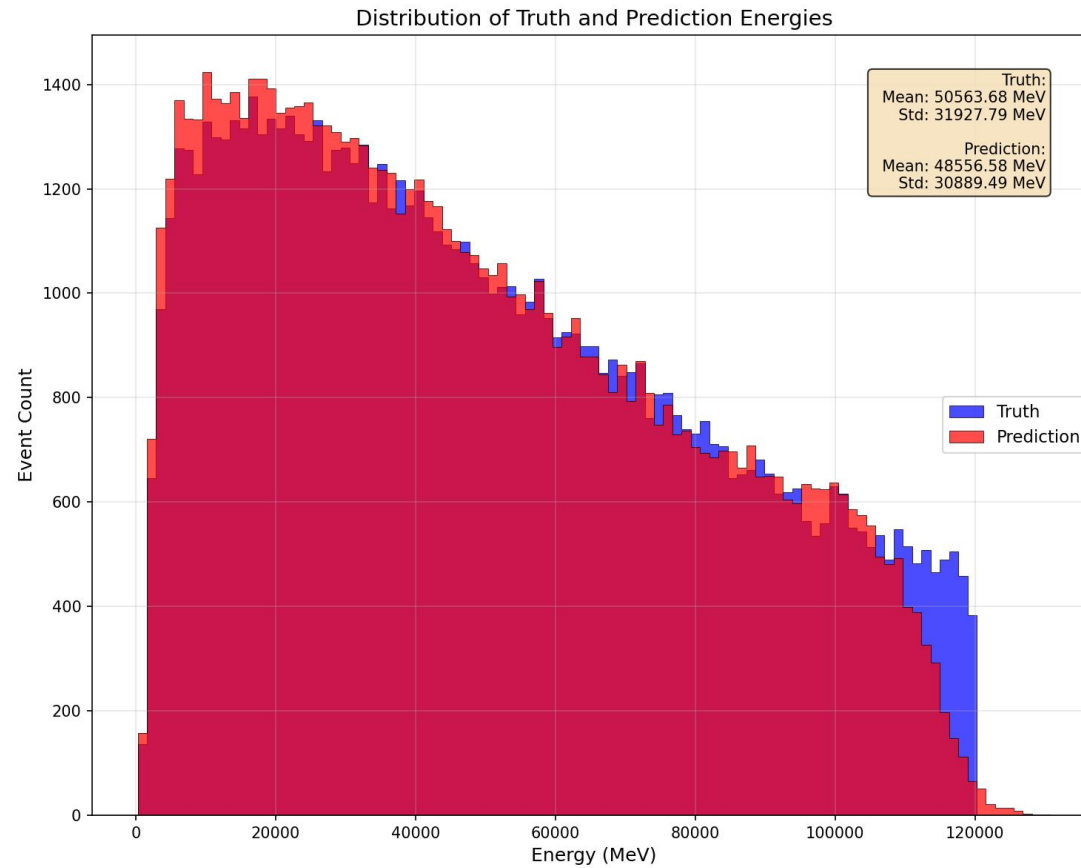


Data



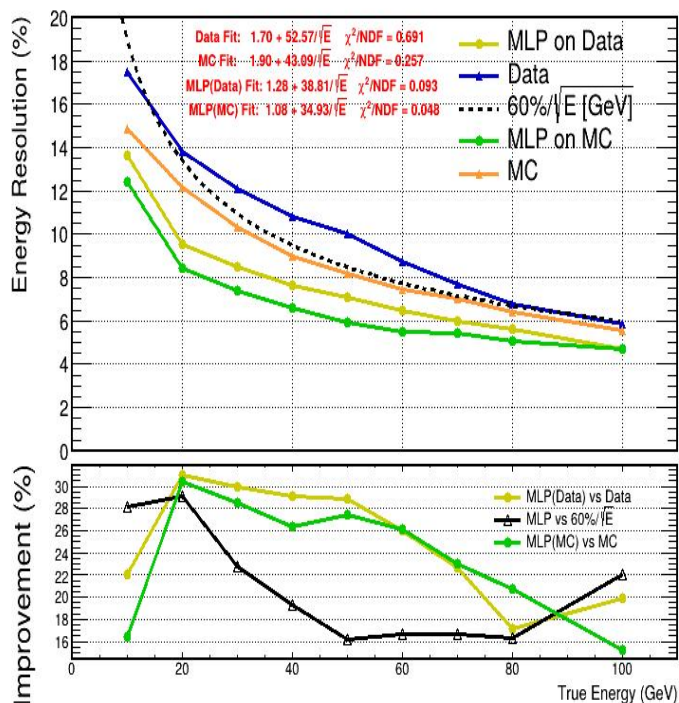
Resolution: Data overall worse than MC
(Larger σ)

back up

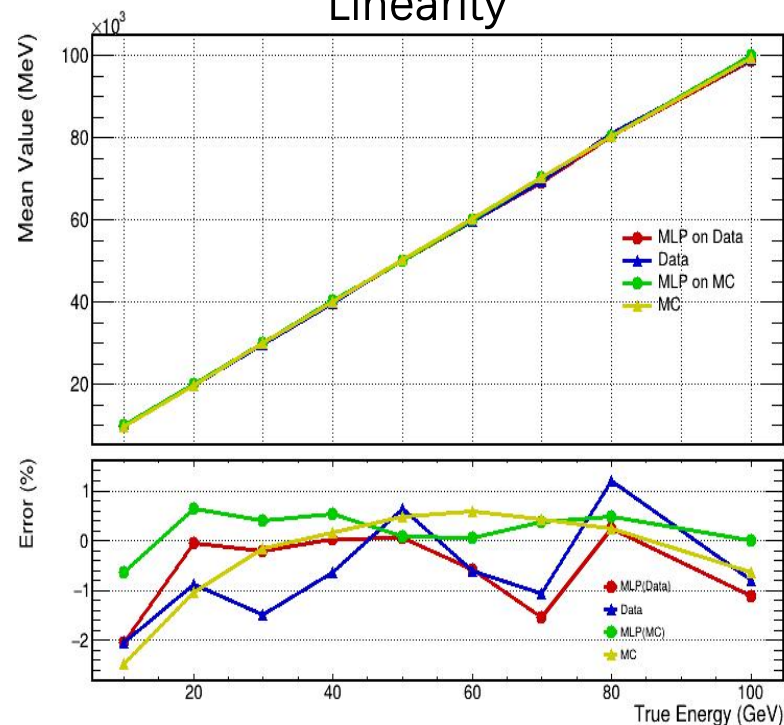


MLP Results

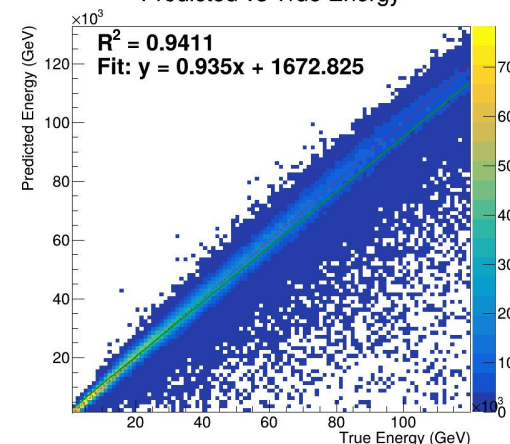
Resolution



Linearity



Predicted vs True Energy



- training sample

- source: pion

range: 1-120GeV

Event: train = 70.3e4

test = 12.4e4

- Data are slightly inferior to MC
(0.83% on average)
- Fewer leakage events
- Resolution: 25% improvement in both MC and Data
- linearity: Data shows notable improvement
(-0.58% on average)

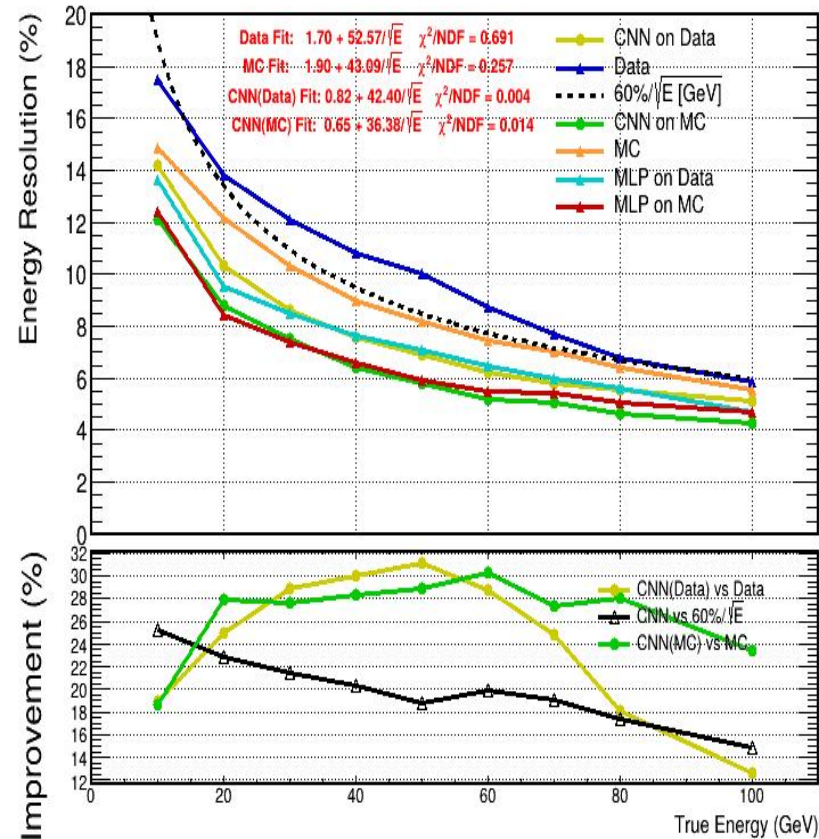
MC remains good (0.21% on average)

Energy (GeV)		10	20	30	40	50	60	70	80	100
Data	Event (left) ($\times 10^4$)	8.3	30.2	48.6	45.8	49.8	36.9	32.7	28.3	25.3
	Percentage (%)	8.1	40.3	44.9	44.1	39.5	35.6	30.9	26.3	19.7
MC	Event (left) ($\times 10^4$)	20.1	21.1	19.9	18.6	16.9	15.2	13.3	11.6	8.7
	Percentage (%)	50.4	52.7	49.8	46.5	42.3	38.0	33.4	29.2	22.1

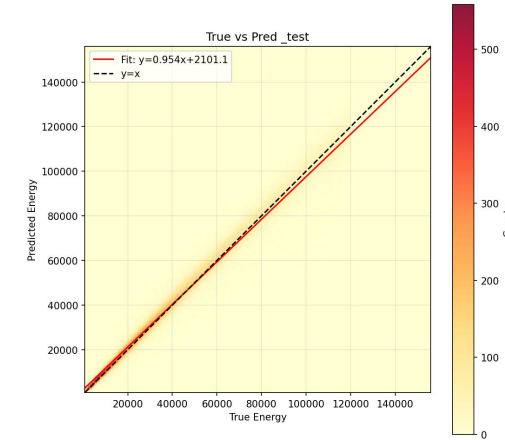
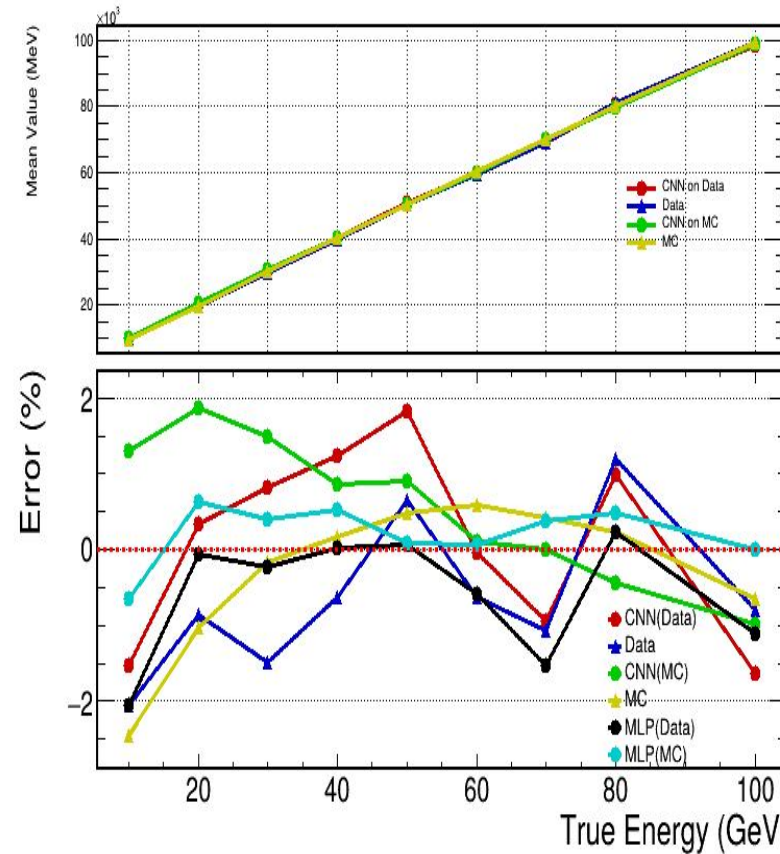
$$Imp = \frac{|R_{MC} - R_{MLP}|}{R_{MC}} \quad Error = \frac{|Mean - True|}{True}$$

CNN Results

Resolution



Linearity



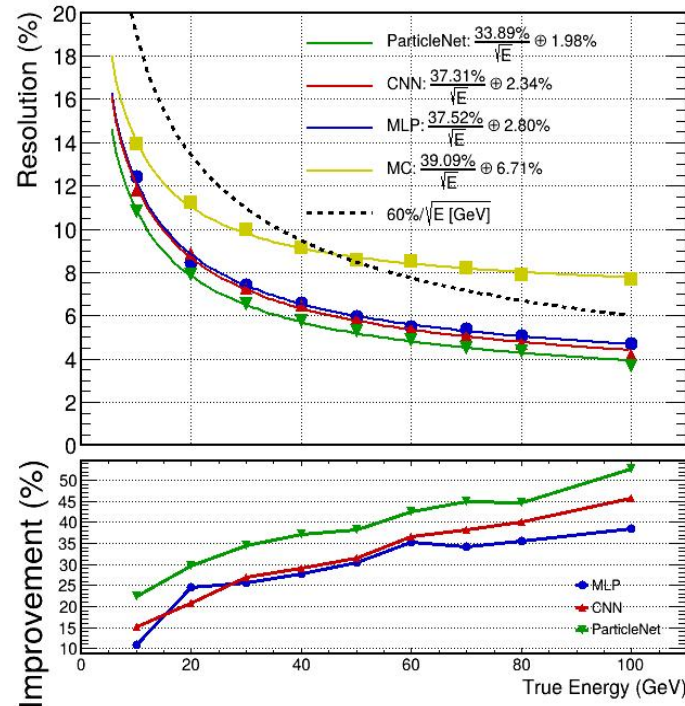
- Resolution
MLP & CNN: ~25% improvement (both)
- Energy Linearity
MLP: notable improvement — 0.51% (avg)
CNN: limited improvement — 0.96% (avg)
- vs. MLP

CNN: no clear advantage on AHCAL yet

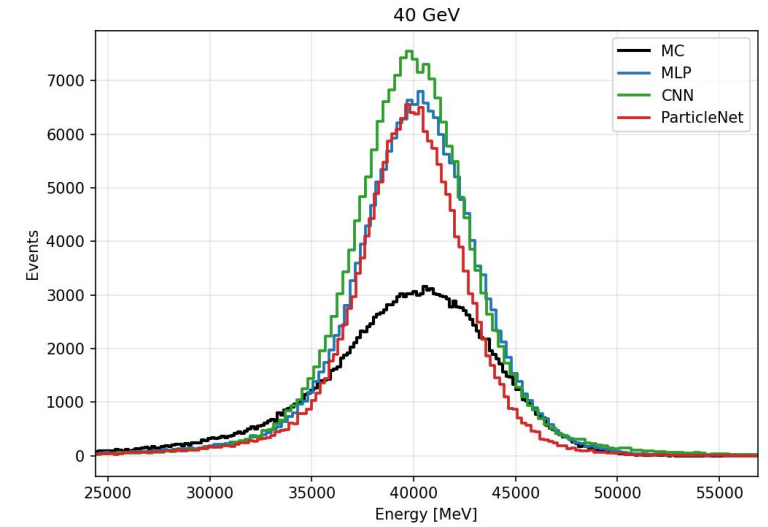
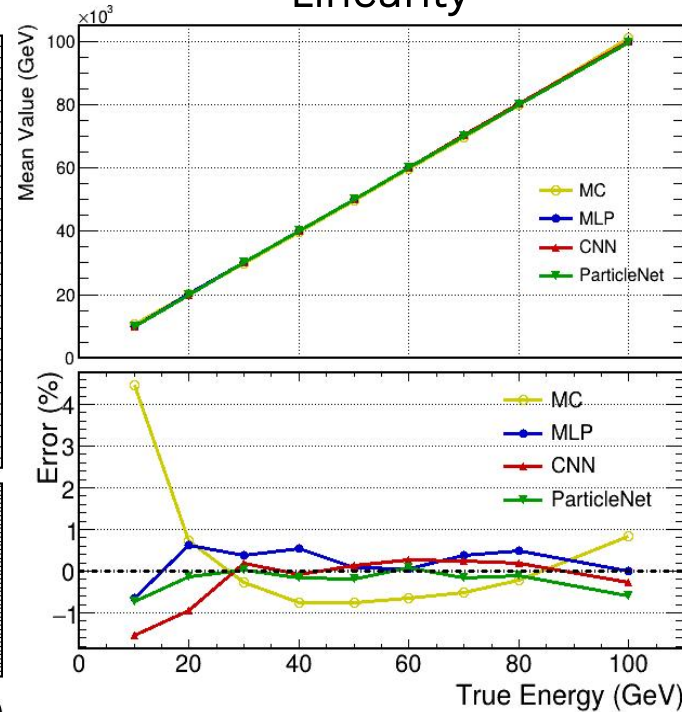
$$Imp = \frac{|R_{MC} - R_{MLP}|}{R_{MC}} \quad Error = \frac{|Mean - True|}{True}$$

Results on MC

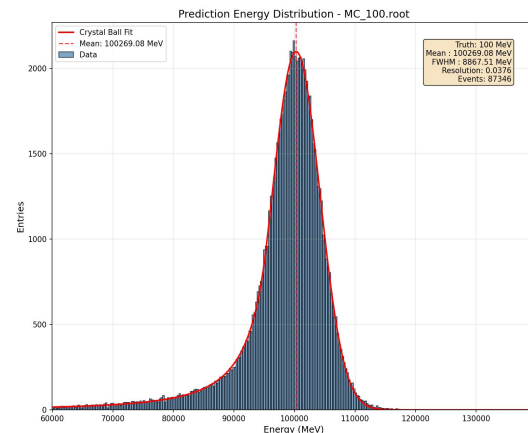
Resolution



Linearity



- training sample
- source: pion
- range: 1-120GeV
- Event: train = 70.3e4
- test = 12.4e4



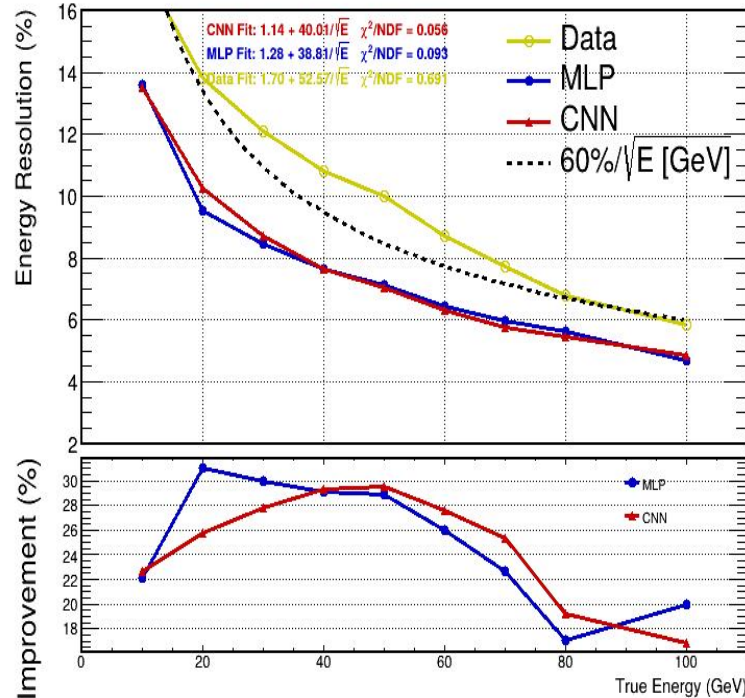
$$Imp = \frac{|R_{MC} - R_{Model}|}{R_{MC}}$$

$$Error = \frac{|Mean - True|}{True}$$

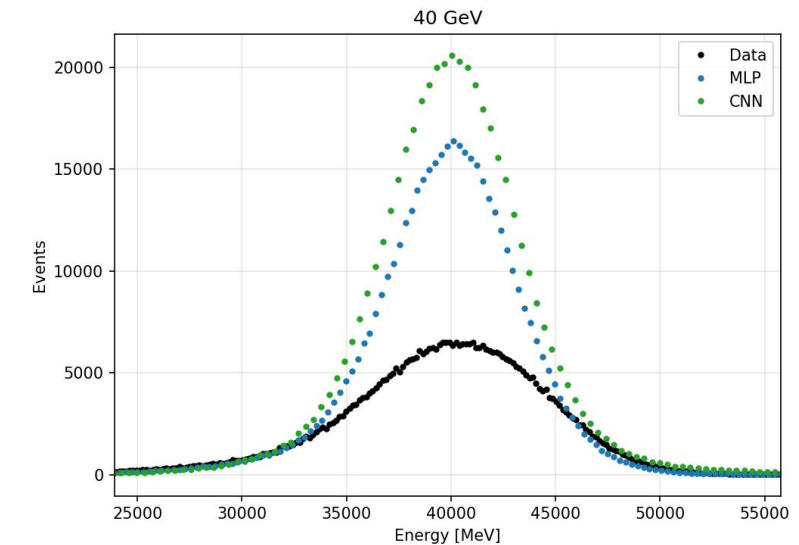
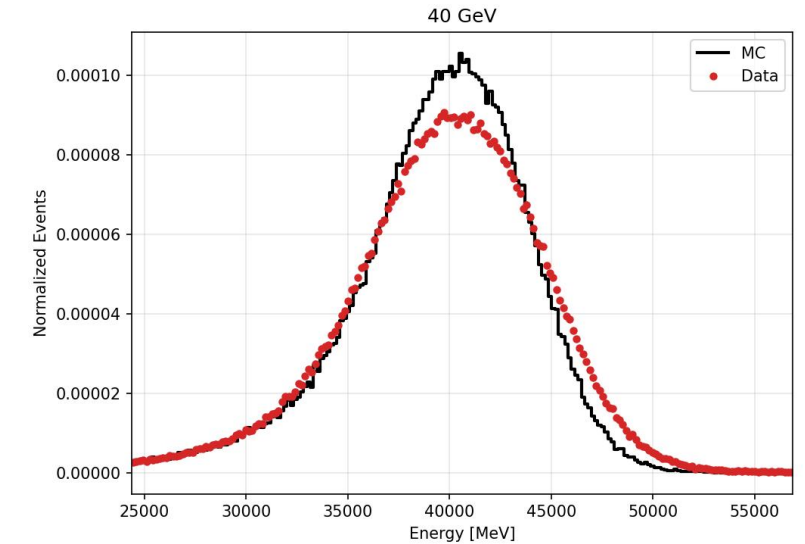
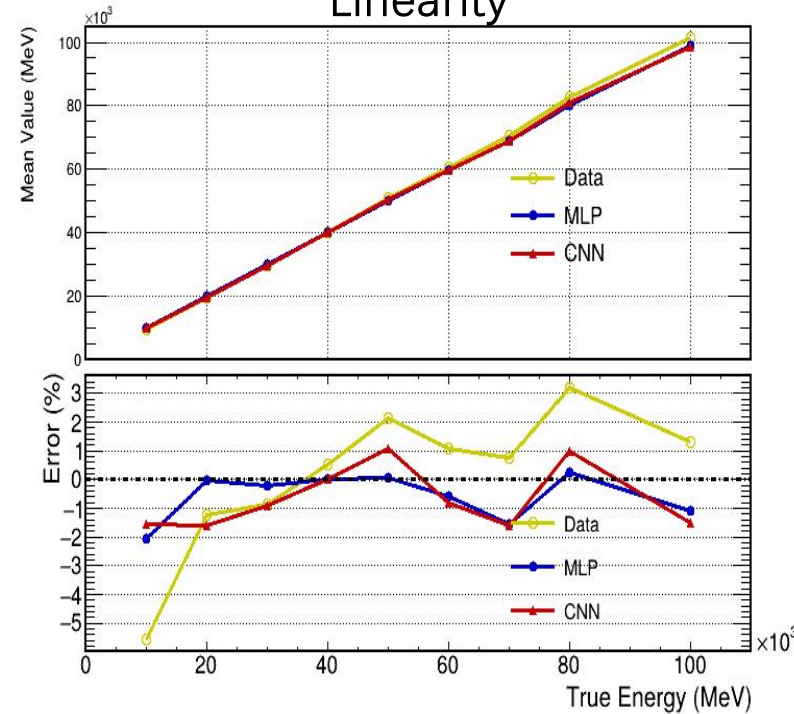
- Resolution
 - MLP & CNN: both ~25% improvement (avg)
 - ParticleNet: 33.4% improvement (avg) — best
- Energy Linearity
 - MLP: 0.88% (avg)
 - CNN: 1.28% (avg) — limited
 - ParticleNet: 0.29% (avg, all within ±1%)
- MLP: solid, stable performance
- CNN: no clear advantage over MLP on MC
- ParticleNet clearly outperforms MLP & CNN

Results on Data

Resolution



Linearity



- Data are slightly inferior to MC (0.83% on average)
- Resolution
MLP & CNN: both ~25% improvement (avg)
- Energy Linearity
MLP: 0.64% (avg)
CNN: 1.11% (avg) — limited