Update on Energy regression

Xin Xia

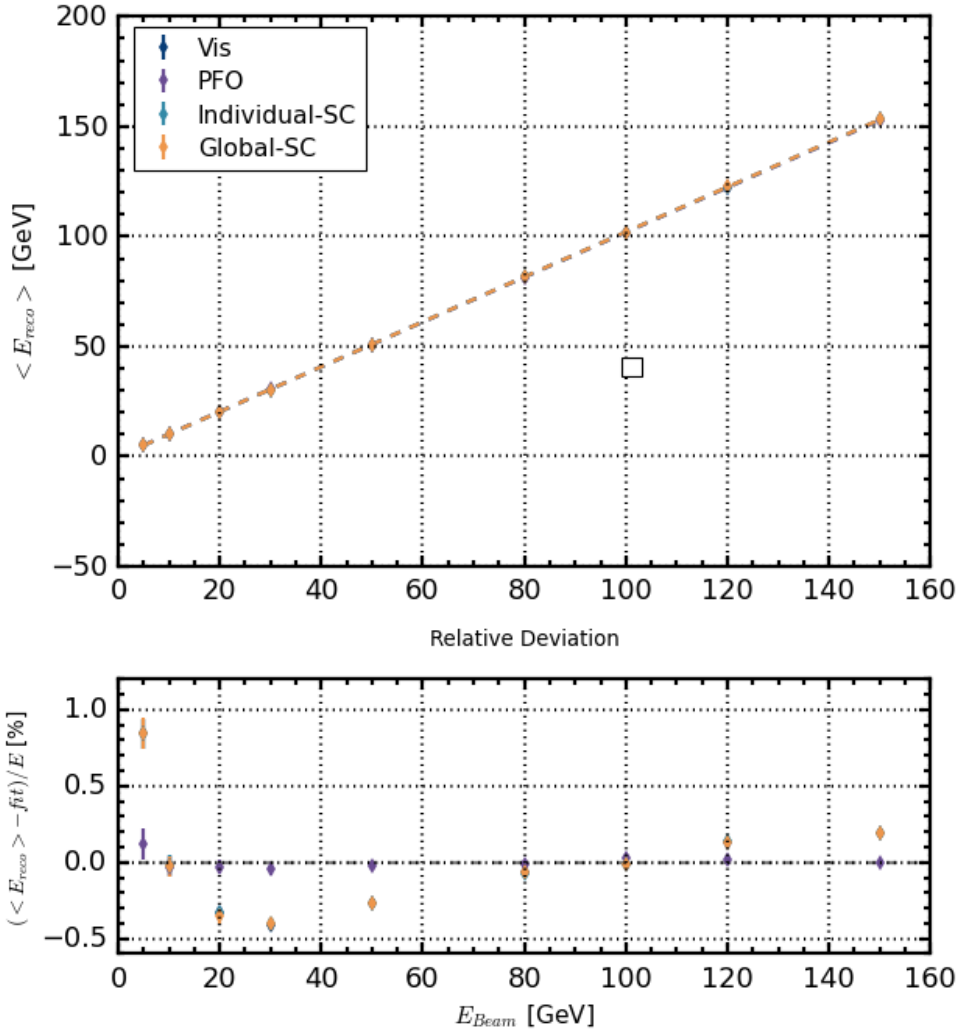
Dataset

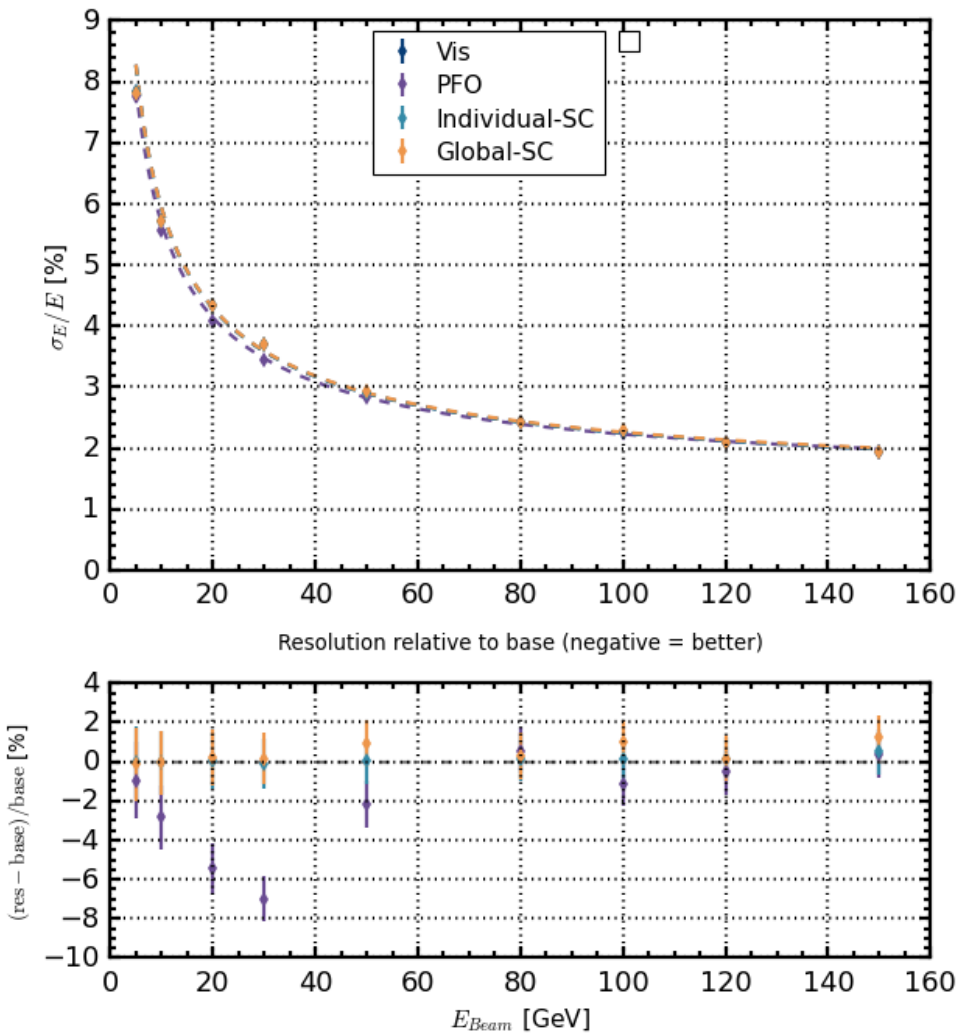
Data set	
Particle	Pi+ / e- / k- in ILD ECAL + AHCAL system ILD/compact/ILD_l5_o1_v02/ILD_l5_o1_v02.xml
Train set	0.2 GeV to 200 GeV (1000 events / 1000 uniform energy points); only pi+
Test set	1,2,3,4,5,6,7,8,9,10,11,12, 13,14,15, 20,30,40,50,80,100,120,150 GeV 10000 events / energy; pi+ / e- / k-

Methods

Methods			
Total energy deposition (Vis)	$E = \sum_{all\_hits} E_{hit}^{ECAL} + \sum_{all\_hits} E_{hit}^{HCAL}$		
Pandora PFA (PFO)	Use standard Pandora PFA reconstruction, choose the PFO with largest energy, ignore other PFOs		
Software Compen (truth fem)	$E_{rec} = \frac{e}{\pi} \cdot E_{dep} = \frac{e}{f_{em} \cdot e + (1 - f_{em}) \cdot h} \cdot E_{dep} = \frac{e/h}{1 + f_{em}(e/h - 1)} \cdot E_{dep}$ (Global-SC)		
	$E_{rec} = \frac{(e/h)_{ECAL}}{1 + f_{em\_ECAL}((e/h)_{ECAL} - 1)} \cdot E_{ECAL} + \frac{(e/h)_{HCAL}}{1 + f_{em\_HCAL}((e/h)_{HCAL} - 1)} \cdot E_{HCAL}$ (Individual-SC)		
CNN(ML w/o time)	(E, x, y, z)		
CNN (ML w time)	(E, T, x, y, z)	Moments	Absolute time
		Bins	sigma = Max(0.04, 0.10 /$\sqrt{E}$) t = t + rng.normal(0, sigma)

Electron Samples - well calibration





CNN Model

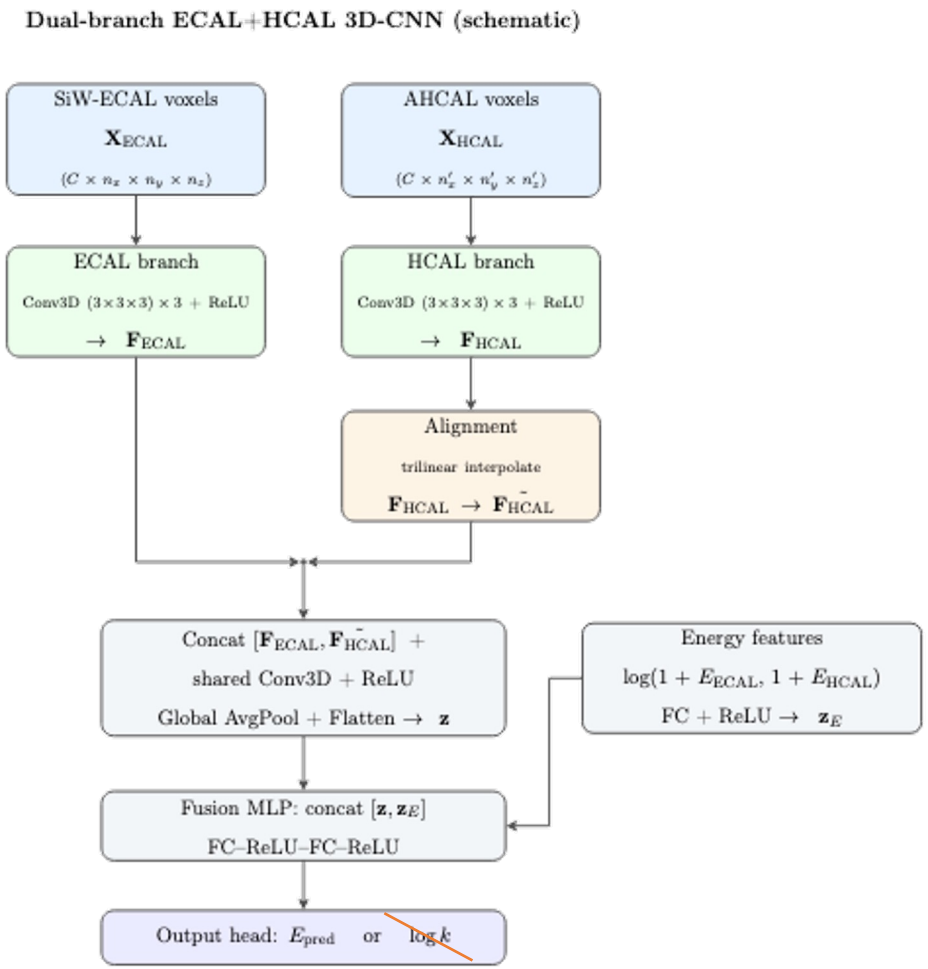


Figure 1: Architecture schematic with parallel feature extraction (ECAL/HCAL) and parallel heads for spatial features $\mathbf{z}$ and global-energy features $\mathbf{z}_E$, fused for regression.

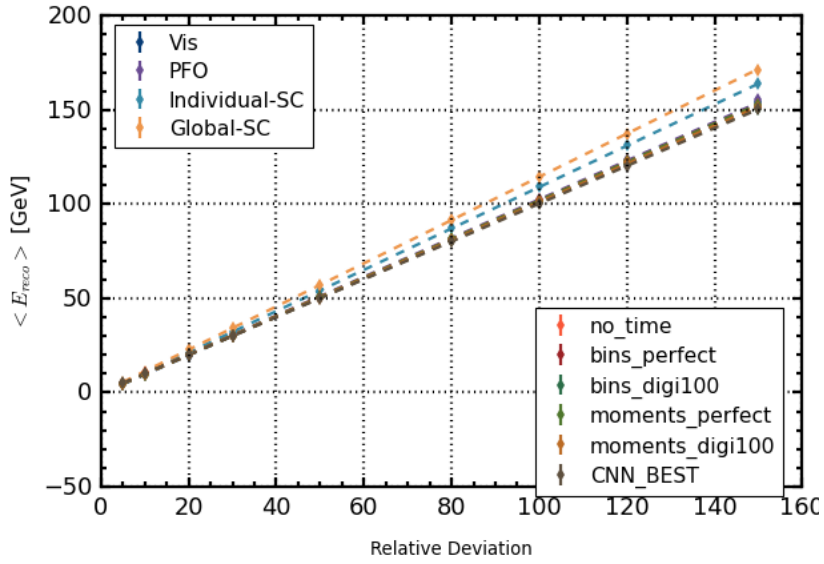
Model	CNN (ML w/o time)	CNN (ML w. time)	
Input information	(E, x, y, z)	(E, T, x, y, z)	
		Moments	Bins
Channels	$[\mathrm{ch}_0, \mathrm{ch}_1]$	$[\mathrm{ch}_0, \mathrm{ch}_1, T_{\mathrm{mean}}, T_{\mathrm{rms}}]$	$[\mathrm{ch}_0, \mathrm{ch}_1, \mathrm{ch2}, \mathrm{ch3}, \mathrm{ch4}]$
	$\mathrm{ch}_0 = \frac{E_{\mathrm{voxel}}}{v_{max}}$, $\mathrm{ch}_1 = \frac{E_{\mathrm{voxel}}}{E_{\mathrm{total}}}$	$T_{\mathrm{mean}} = \frac{\sum_{i\in\mathrm{voxel}} E_i t_i}{\sum_{i\in\mathrm{voxel}} E_i + \epsilon}$, $T_{\mathrm{rms}} = \sqrt{\frac{\sum_{i\in\mathrm{voxel}} E_i t_i^2}{\sum_{i\in\mathrm{voxel}} E_i + \epsilon} - T_{\mathrm{mean}}^2 + \epsilon}$	$\mathrm{ch}_2 = \frac{E_{\mathrm{early}}}{E_{\mathrm{voxel}}}$, $\mathrm{ch}_3 = \frac{E_{\mathrm{mid}}}{E_{\mathrm{voxel}}}$, $\mathrm{ch}_4 = \frac{E_{\mathrm{late}}}{E_{\mathrm{voxel}}}$
			$E_{\mathrm{early}}: t < t_0$, $E_{\mathrm{mid}}: t_0 \le t < t_1$, $E_{\mathrm{late}}: t \ge t_1$
Absolute time, sigma = Max(0.04, 0.10 /$\sqrt{E}$), t = t + rng.normal(0, sigma)			

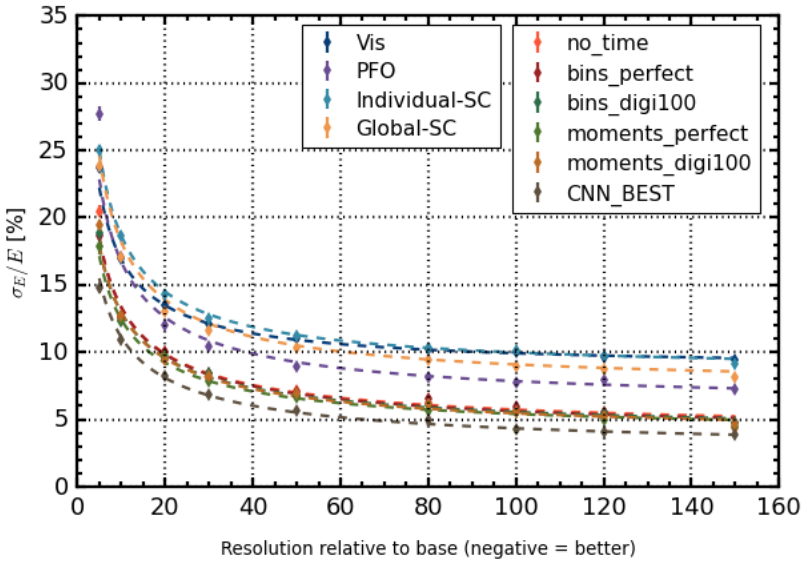
Pion samples

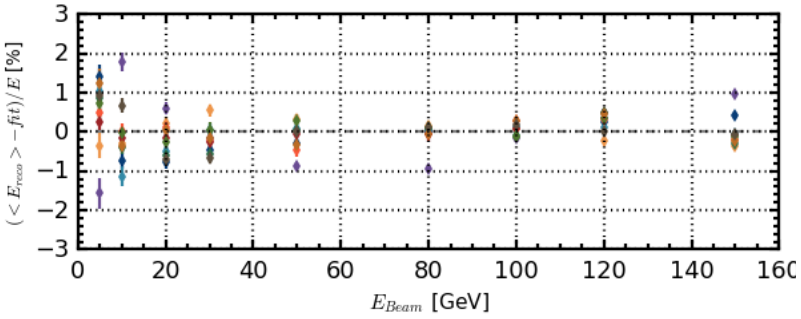
- CNN energy regression improves systematically (linearity: 2%; Resolution:20% - 50%).
- No big difference between w. or w/o time.
- CNN_Best (L1) is the best performance I have achieved.

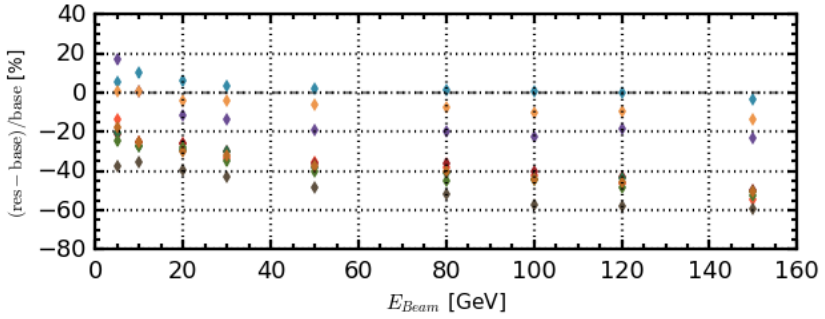
$$L1 = |x|$$

$$SmoothL1(x) = \begin{cases} \frac{x^2}{2\beta}, & |x| < \beta \\ |x| - \frac{\beta}{2}, & |x| \geq \beta \end{cases}$$









Method	a	b	chi2/ndf
Vis	1.0277 ± 0.0006	-0.751 ± 0.013	13.49
PFO	1.0280 ± 0.0005	-0.283 ± 0.015	56.63
HCAL-SC	1.0982 ± 0.0006	-0.737 ± 0.014	5.69
Global-SC	1.1469 ± 0.0006	-0.199 ± 0.014	5.76
no_time	1.0216 ± 0.0005	-0.332 ± 0.013	5.68
bins_perfect	1.0108 ± 0.0005	-0.175 ± 0.010	3.71
bins_digi100	1.0100 ± 0.0005	-0.219 ± 0.012	10.26
moments_perfect	1.0231 ± 0.0006	-0.409 ± 0.010	2.51
moments_digi100	1.0141 ± 0.0005	-0.269 ± 0.012	8.48
CNN_BEST	1.0031 ± 0.0004	-0.066 ± 0.011	18.04

Method	S	C	chi2/ndf
Vis	45.64 ± 0.54%	8.81 ± 0.06%	0.89
PFO	49.22 ± 0.56%	6.10 ± 0.07%	17.09
HCAL-SC	52.57 ± 0.58%	8.48 ± 0.07%	3.15
Global-SC	52.51 ± 0.49%	7.41 ± 0.06%	3.99
no_time	40.22 ± 0.54%	4.07 ± 0.11%	7.99
bins_perfect	40.33 ± 0.47%	3.90 ± 0.10%	5.64
bins_digi100	38.97 ± 0.56%	3.94 ± 0.10%	4.20
moments_perfect	37.32 ± 0.37%	3.91 ± 0.11%	3.54
moments_digi100	39.14 ± 0.44%	3.79 ± 0.10%	3.59
CNN_BEST	34.11 ± 0.40%	2.67 ± 0.09%	2.10

Loss function

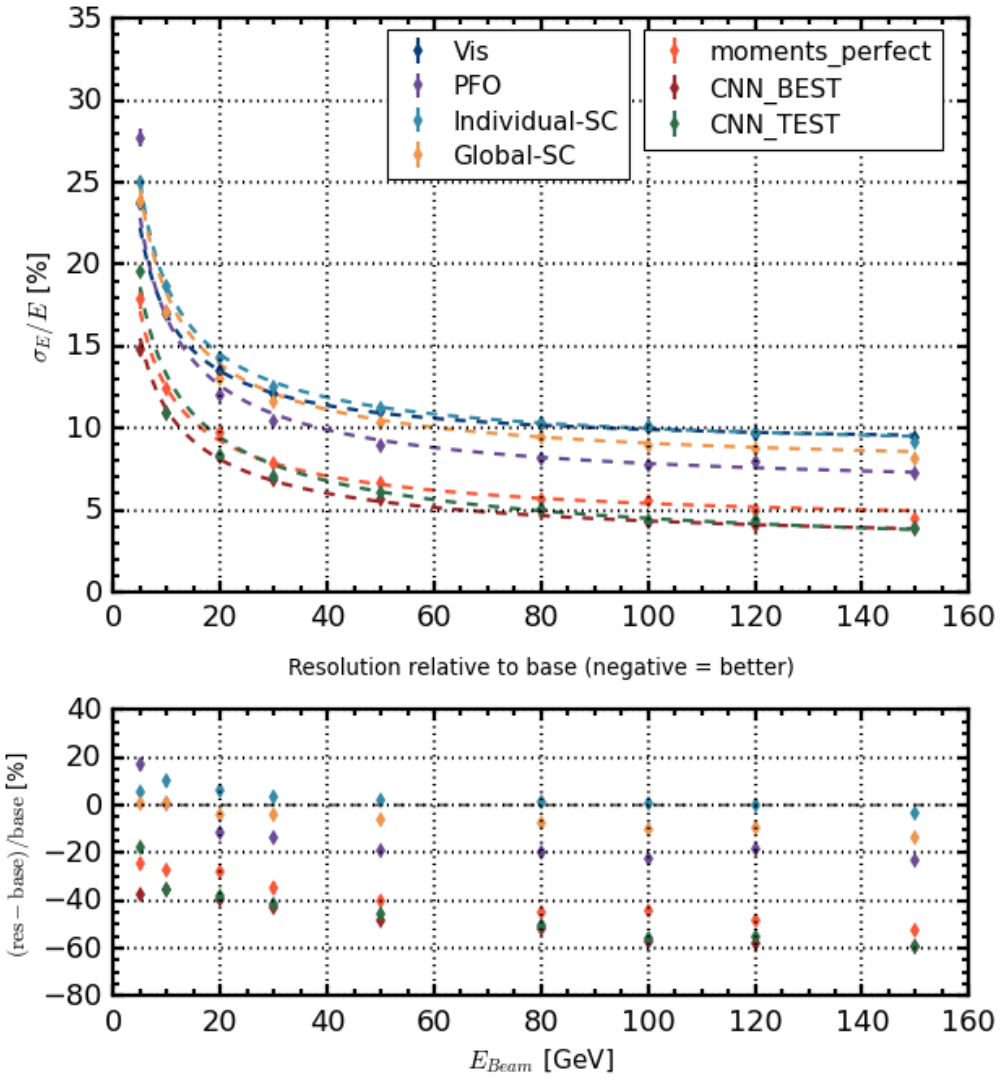
Optimization objective: relative error

$$L_{Huber} = \frac{1}{N}\sum_i \text{Huber}\left(\frac{\hat{E}_i - E_i}{E_i}\right),$$

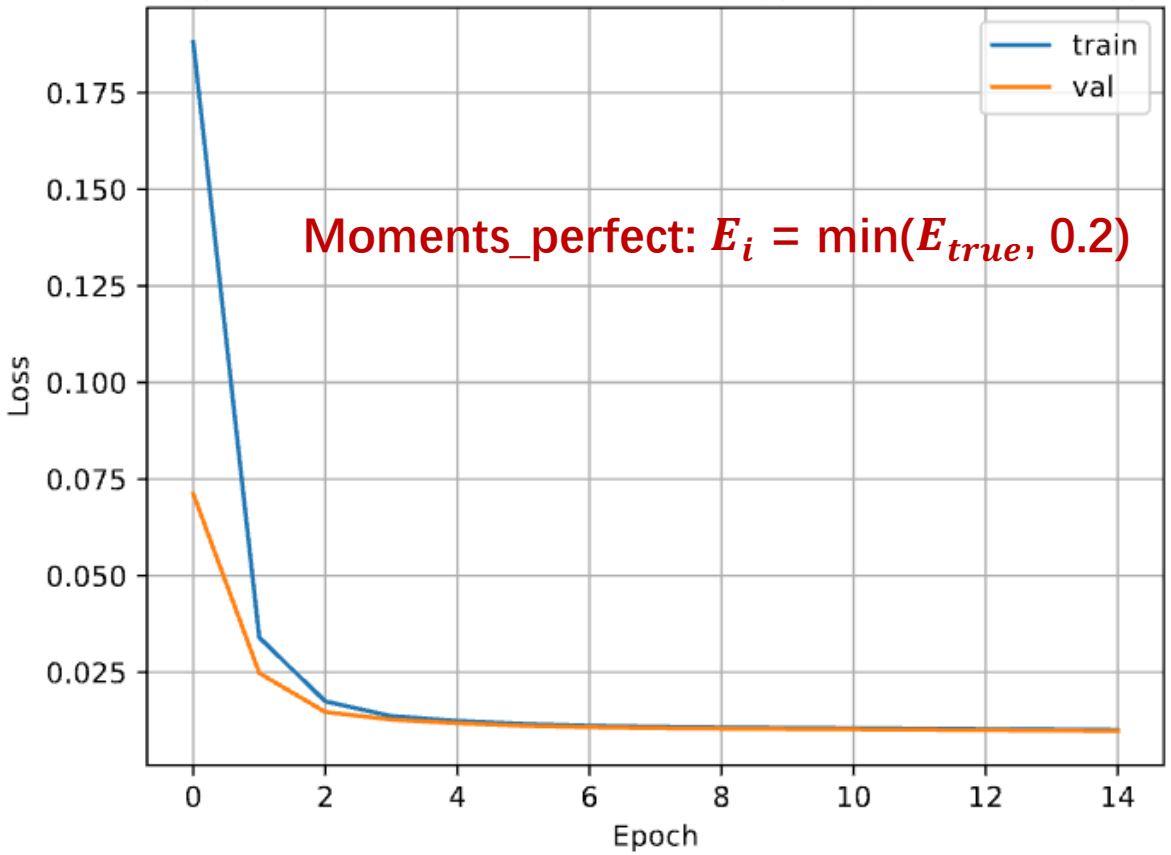
$$\text{where Huber(r) is defined as Huber(r)} = \begin{cases} \frac{1}{2}r^2, |r| < \delta \\ \delta(|r| - \frac{1}{2}\delta), |r| \geq \delta \end{cases}$$

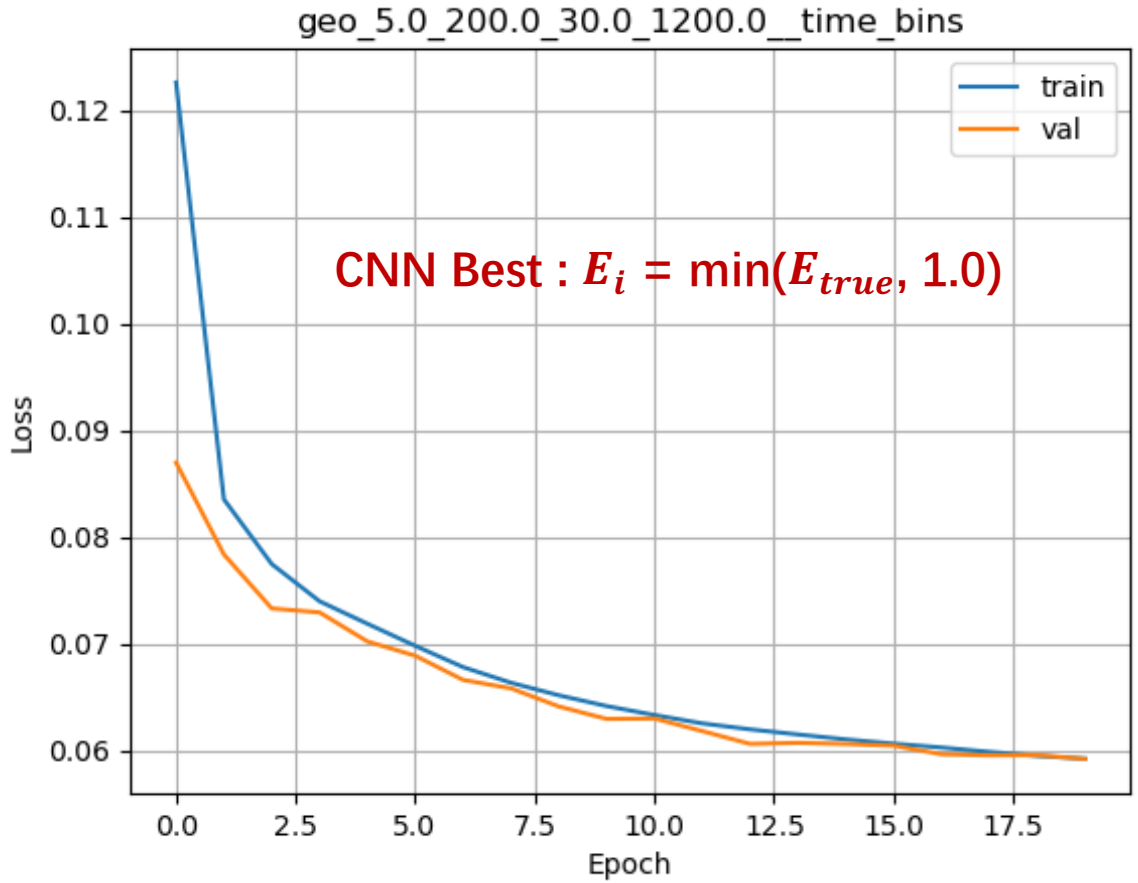
$\hat{E}_i$: predicted energy value of i_th sample
E_i: true energy value of i_th sample
δ: threshold

CNN Best & Test: $E_i = \min(E_{true}, 1.0)$
Moments_perfect: $E_i = \min(E_{true}, 0.2)$



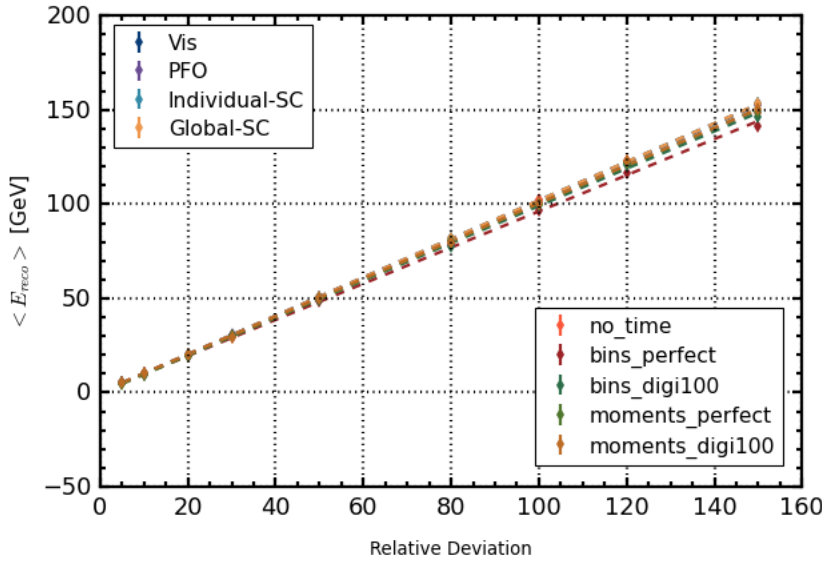
Loss function curve

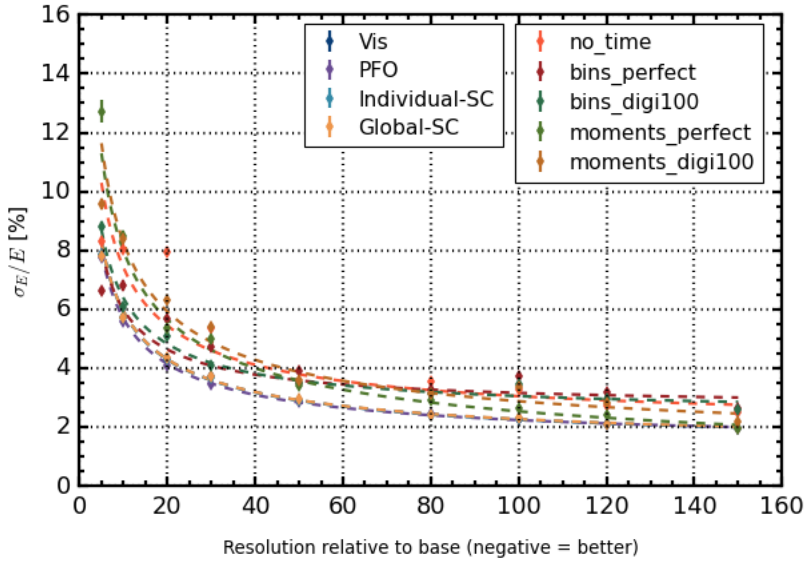


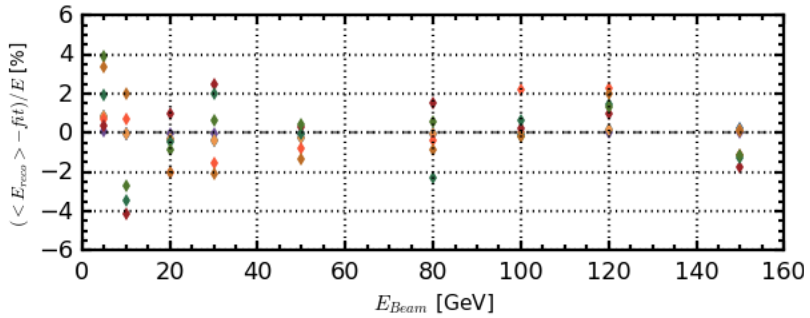


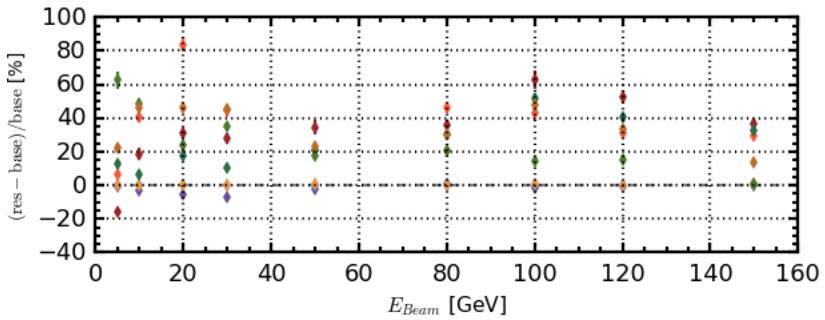
E samples

- CNN model(trained by pion samples) doesn't perform well on electron samples.
- Dedicated compensation for hadronic showers, hadron-specific physics correction
- Over-compensation for electrons







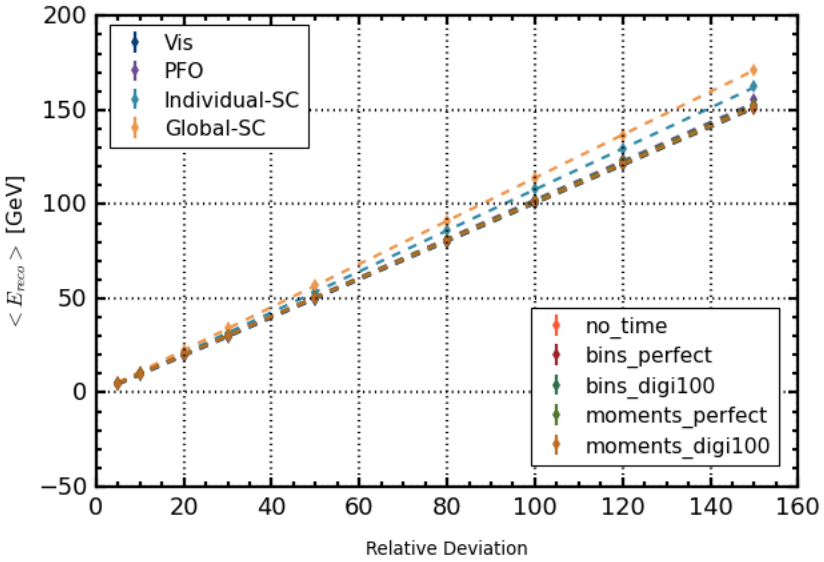


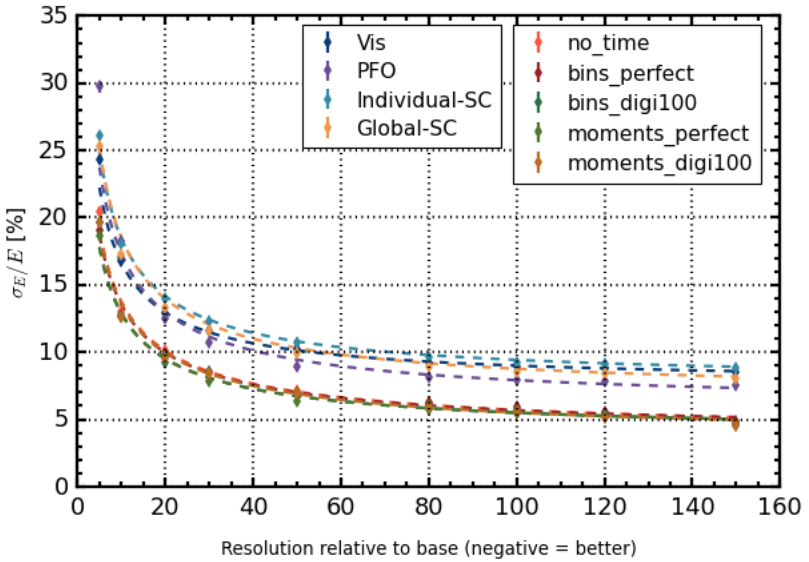
Method	a	b	chi2/ndf
Vis	1.0212 ± 0.0001	-0.121 ± 0.004	59.69
PFO	1.0186 ± 0.0001	-0.027 ± 0.004	0.83
HCAL-SC	1.0212 ± 0.0001	-0.121 ± 0.004	58.99
Global-SC	1.0219 ± 0.0001	-0.122 ± 0.004	58.43
no_time	0.9998 ± 0.0003	0.175 ± 0.006	627.27
bins_perfect	0.9599 ± 0.0003	0.208 ± 0.006	714.82
bins_digi100	0.9902 ± 0.0002	0.070 ± 0.006	950.27
moments_perfect	1.0074 ± 0.0002	-0.488 ± 0.008	415.91
moments_digi100	1.0037 ± 0.0002	0.146 ± 0.007	866.51

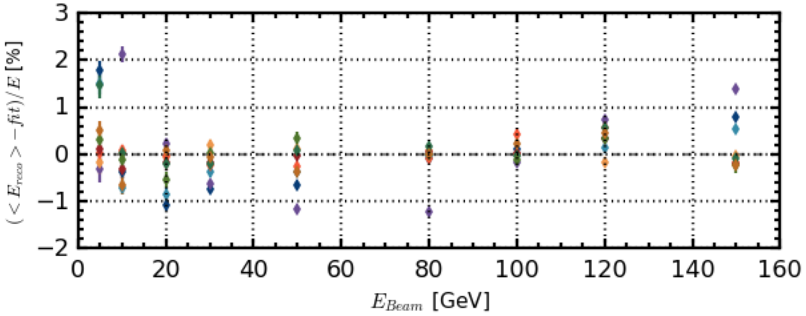
Method	S	C	chi2/ndf
Vis	18.20 ± 0.12%	1.29 ± 0.02%	7.78
PFO	17.29 ± 0.12%	1.39 ± 0.02%	3.20
HCAL-SC	18.17 ± 0.12%	1.30 ± 0.02%	7.38
Global-SC	18.19 ± 0.12%	1.32 ± 0.02%	8.36
no_time	22.18 ± 0.29%	2.08 ± 0.06%	72.94
bins_perfect	16.98 ± 0.30%	2.68 ± 0.05%	41.30
bins_digi100	18.69 ± 0.27%	2.41 ± 0.05%	10.40
moments_perfect	25.36 ± 0.16%	-0.00 ± 63.70%	13.96
moments_digi100	25.55 ± 0.25%	1.30 ± 0.08%	37.22

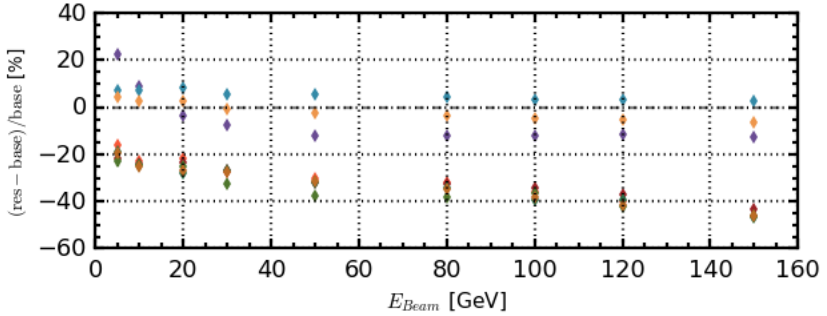
Kaon samples

- CNN model(trained by pion samples) performs well on kaon- samples.
- But a bit worse compared with working on pi+ samples.







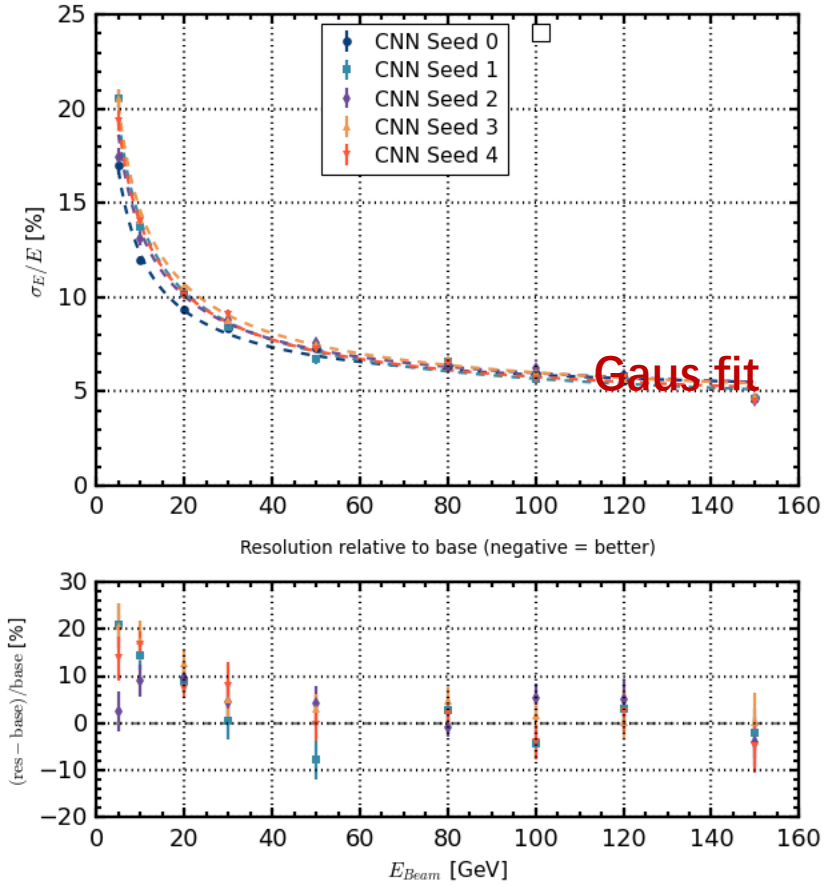


Method	a	b	chi2/ndf
Vis	1.0119 ± 0.0004	-0.811 ± 0.009	62.72
PFO	1.0296 ± 0.0003	-0.625 ± 0.011	193.33
HCAL-SC	1.0862 ± 0.0004	-0.866 ± 0.010	24.51
Global-SC	1.1433 ± 0.0004	-0.345 ± 0.010	1.70
no_time	1.0174 ± 0.0005	-0.259 ± 0.008	5.59
bins_perfect	1.0090 ± 0.0003	-0.104 ± 0.007	6.37
bins_digi100	1.0093 ± 0.0002	-0.241 ± 0.008	24.66
moments_perfect	1.0216 ± 0.0005	-0.325 ± 0.013	5.05
moments_digi100	1.0121 ± 0.0003	-0.193 ± 0.010	16.18

Method	S	C	chi2/ndf
Vis	48.76 ± 0.39%	7.54 ± 0.05%	5.53
PFO	51.96 ± 0.42%	5.92 ± 0.06%	47.19
HCAL-SC	53.42 ± 0.42%	7.61 ± 0.05%	3.72
Global-SC	51.73 ± 0.35%	6.81 ± 0.05%	5.73
no_time	42.18 ± 0.37%	3.76 ± 0.12%	10.47
bins_perfect	41.10 ± 0.34%	3.88 ± 0.07%	7.94
bins_digi100	39.11 ± 0.31%	3.83 ± 0.04%	13.25
moments_perfect	39.13 ± 0.54%	3.82 ± 0.10%	4.06
moments_digi100	41.75 ± 0.41%	3.55 ± 0.08%	8.40

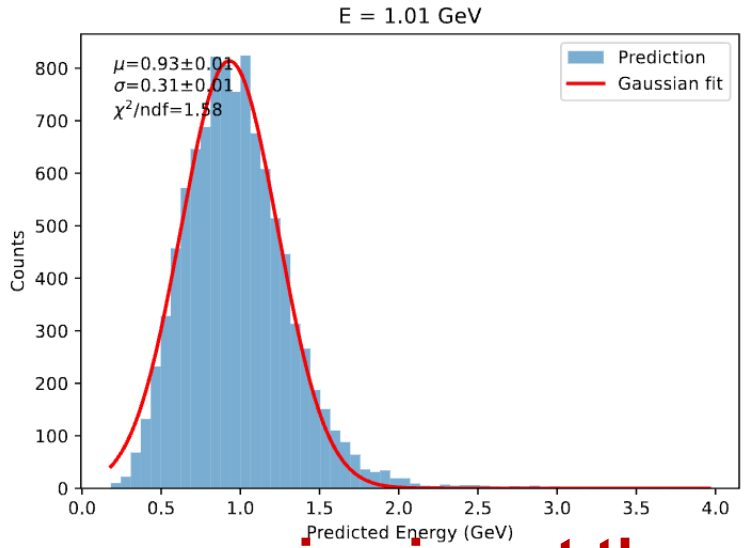
Seeds Problems

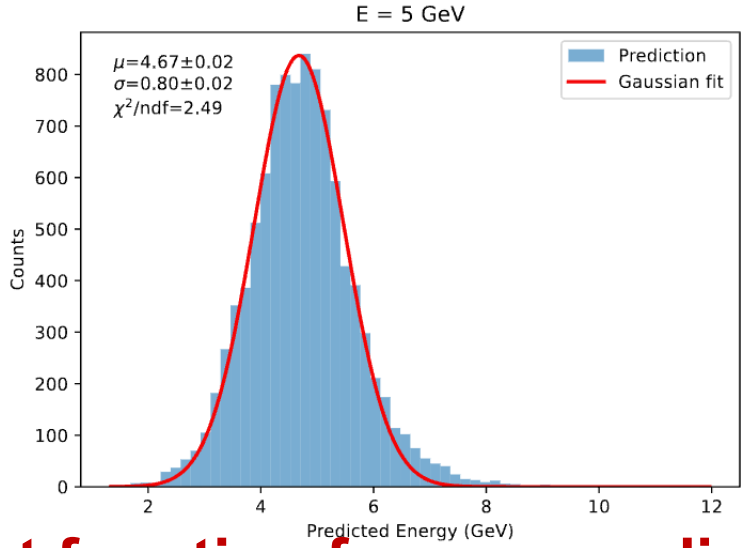
Different seed, different results. Significant in lower energies.



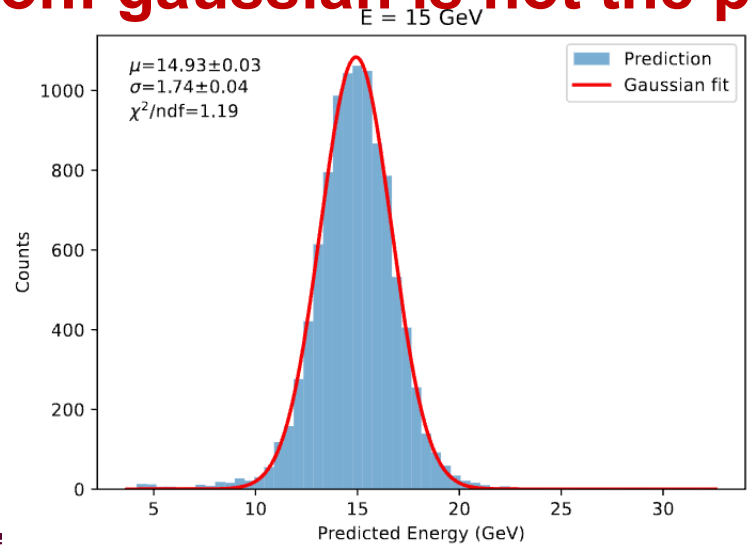
Method	S	C	chi2/ndf
CNN Seed 0	35.99 ± 0.51%	4.65 ± 0.10%	4.02
CNN Seed 1	42.92 ± 0.60%	3.72 ± 0.13%	5.84
CNN Seed 2	40.51 ± 0.69%	4.40 ± 0.11%	9.89
CNN Seed 3	44.32 ± 0.62%	4.05 ± 0.11%	3.30
CNN Seed 4	42.22 ± 0.53%	3.94 ± 0.14%	4.81

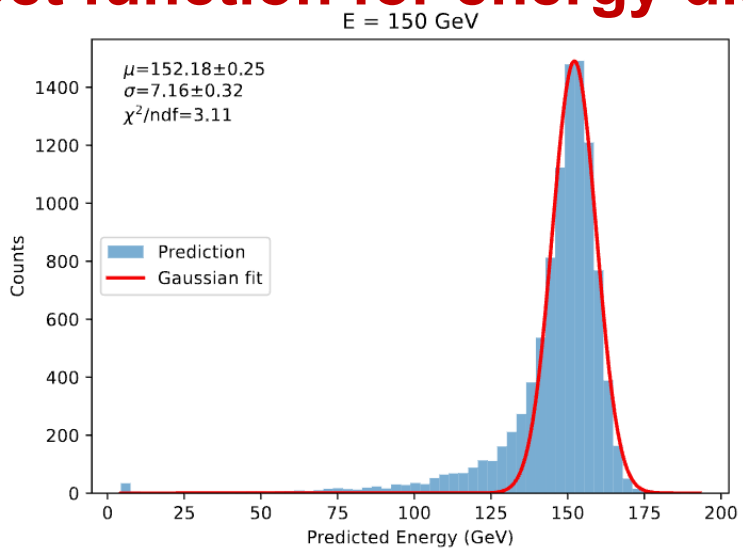
Energy distribution





One reason: gaussian is not the perfect function for energy distribution.

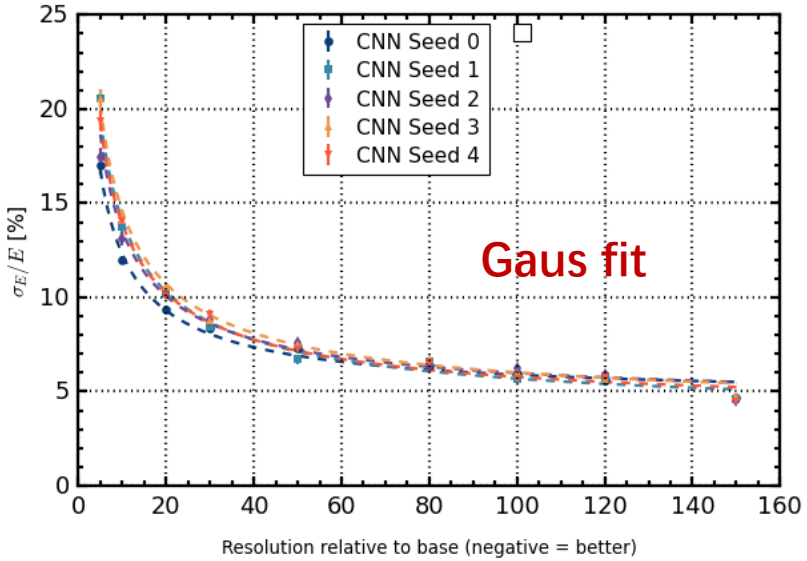




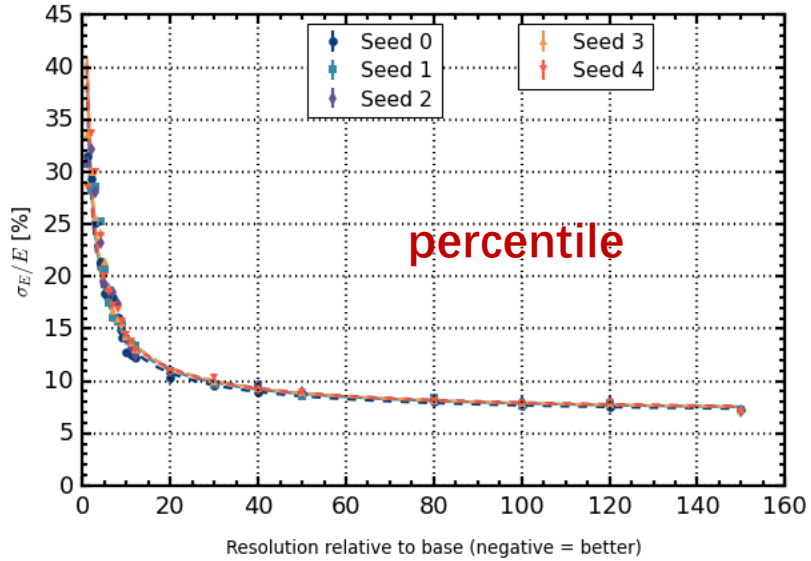
Seeds Problems

One reason: gaussian is not the perfect function for energy distribution.

Method	S	C	chi2/ndf
CNN Seed 0	35.99 ± 0.51%	4.65 ± 0.10%	4.02
CNN Seed 1	42.92 ± 0.60%	3.72 ± 0.13%	5.84
CNN Seed 2	40.51 ± 0.69%	4.40 ± 0.11%	9.89
CNN Seed 3	44.32 ± 0.62%	4.05 ± 0.11%	3.30
CNN Seed 4	42.22 ± 0.53%	3.94 ± 0.14%	4.81



Method	S	C	chi2/ndf
Seed 0	37.79 ± 0.10%	6.68 ± 0.03%	115.50
Seed 1	39.02 ± 0.10%	6.83 ± 0.03%	176.54
Seed 2	39.86 ± 0.11%	6.87 ± 0.03%	188.67
Seed 3	40.43 ± 0.11%	6.77 ± 0.03%	89.66
Seed 4	39.62 ± 0.11%	6.83 ± 0.03%	288.36



Seeds Problems

- The model only predict the energy for each event. And then get the energy distribution for each energy point. Then choose different way to get the sigma/mean value. Then use $\frac{\sigma}{E} = \frac{S}{\sqrt{E}} \oplus C$ to fit the whole energy range. Each step introduces differences, which are reflected in S and C.

Seeds Problems

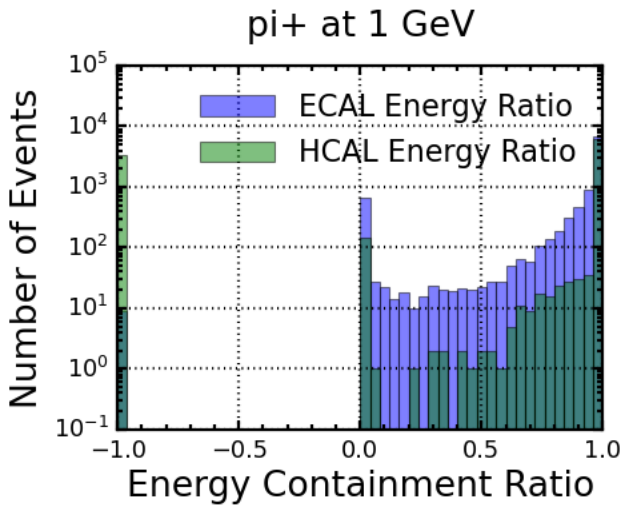
The information of low energy is less than high energy, model can get several solutions from the less information, it's hard to converge to only one solution.

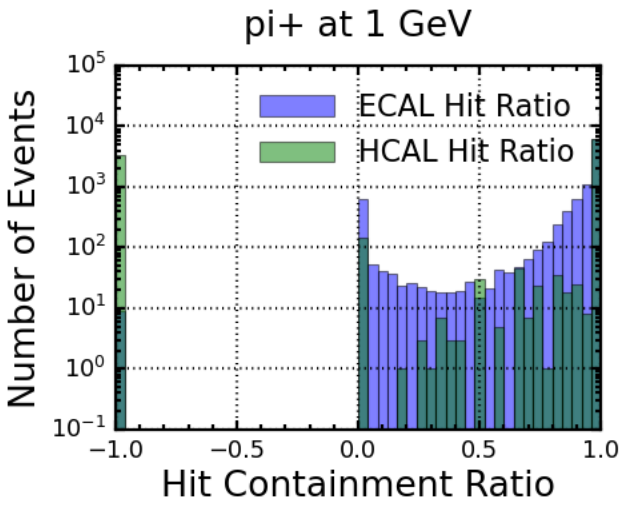
Use different seeds, the first few batches determine which solution the model will choose.

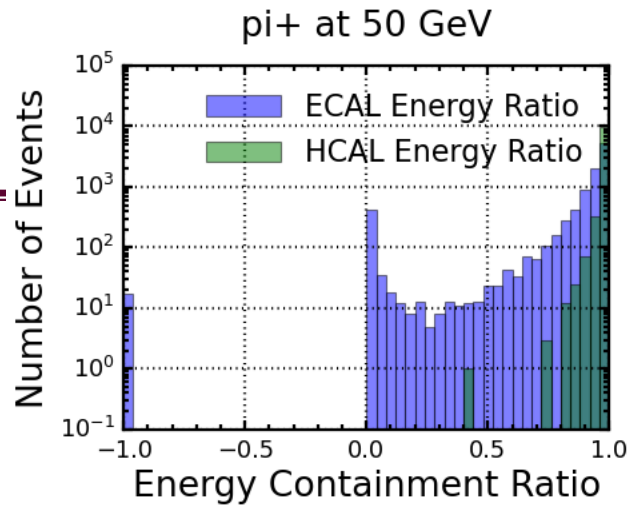
Summary

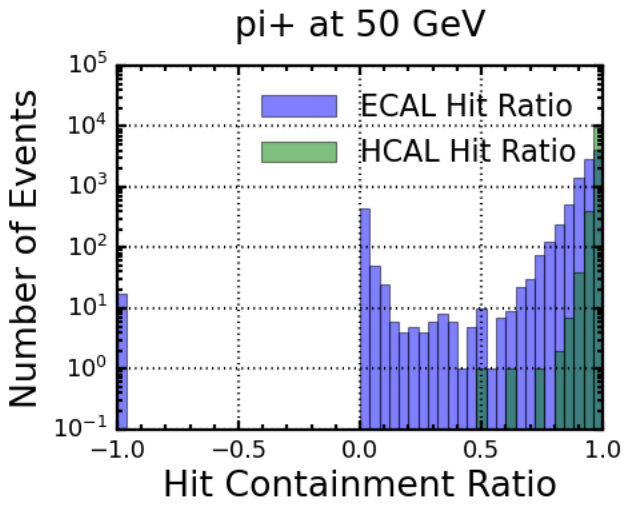
- The CNN significantly reduces the stochastic (S) and constant (C) terms for π, which essentially corresponds to an implicit software compensation.
  - EM sub-shower information generated by π^0
  - Differences in shower topology
  - Differences in the temporal structure of hadronic vs EM
- The performance degradation observed for e^- demonstrates that the model has learned hadronic physics rather than simply performing a trivial fit.
- Timing information looks not very helpful, maybe need further optimization

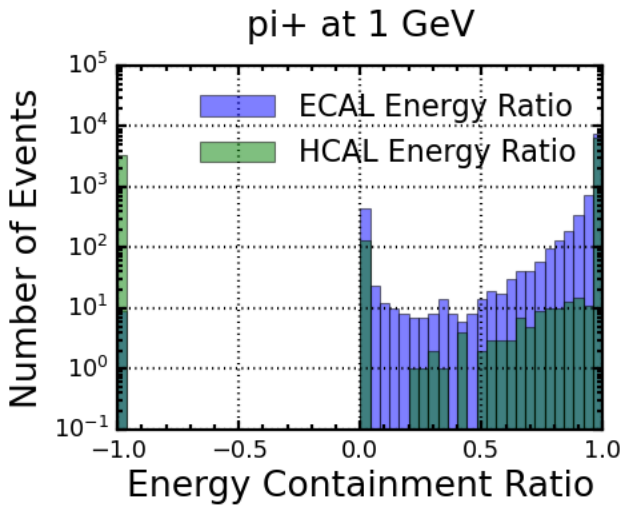
Thanks!

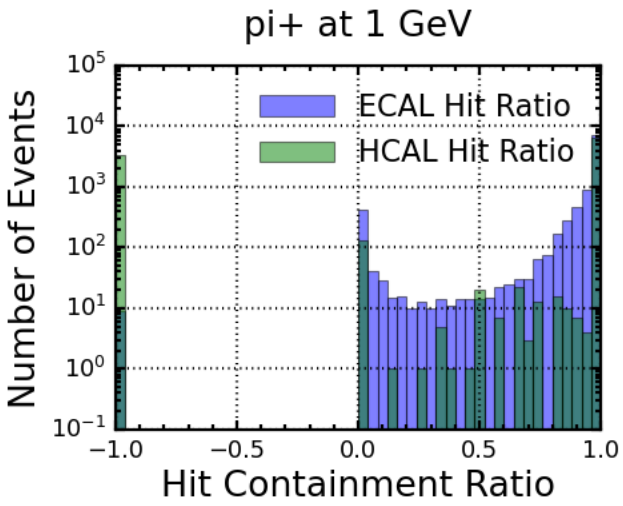

pi+ at 1 GeV
ECAL Energy Ratio
HCAL Energy Ratio
Number of Events
Energy Containment Ratio

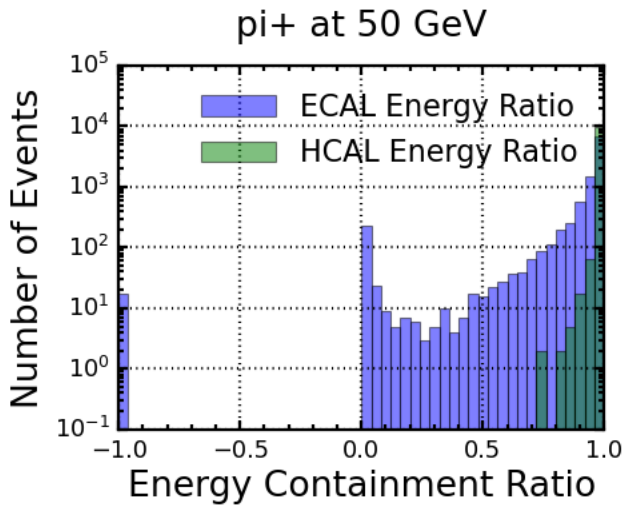

pi+ at 1 GeV
ECAL Hit Ratio
HCAL Hit Ratio
Number of Events
Hit Containment Ratio

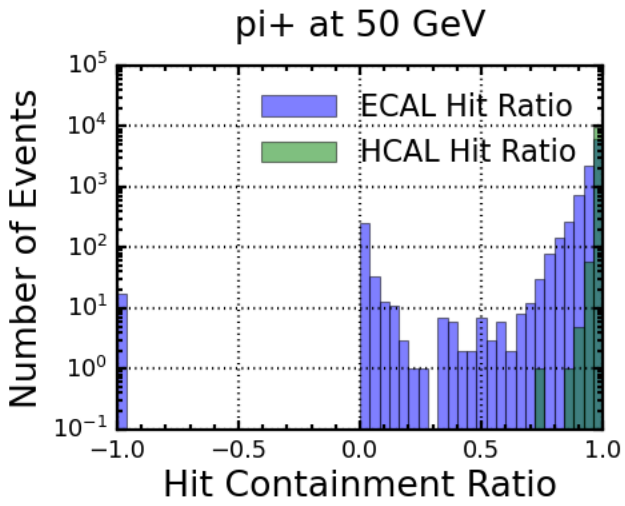

pi+ at 50 GeV
ECAL Energy Ratio
HCAL Energy Ratio
Number of Events
Energy Containment Ratio

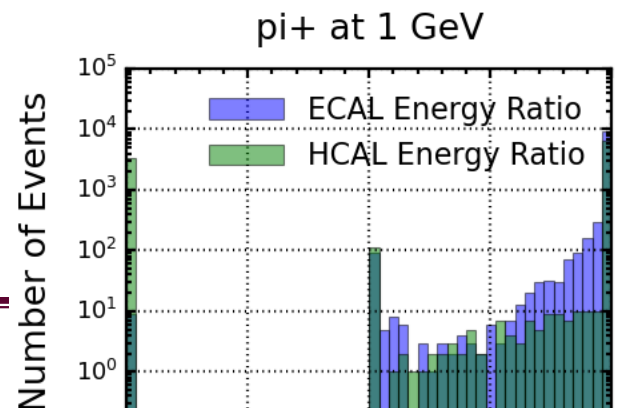

pi+ at 50 GeV
ECAL Hit Ratio
HCAL Hit Ratio
Number of Events
Hit Containment Ratio

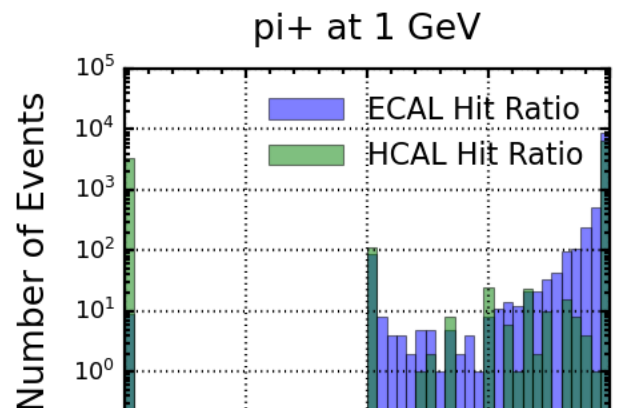

pi+ at 1 GeV
ECAL Energy Ratio
HCAL Energy Ratio
Number of Events
Energy Containment Ratio

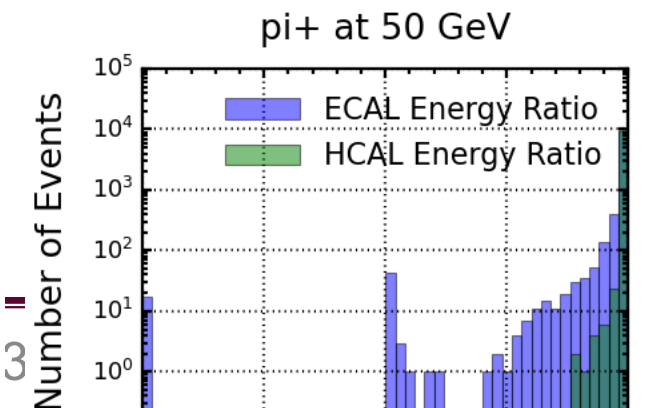

pi+ at 1 GeV
ECAL Hit Ratio
HCAL Hit Ratio
Number of Events
Hit Containment Ratio

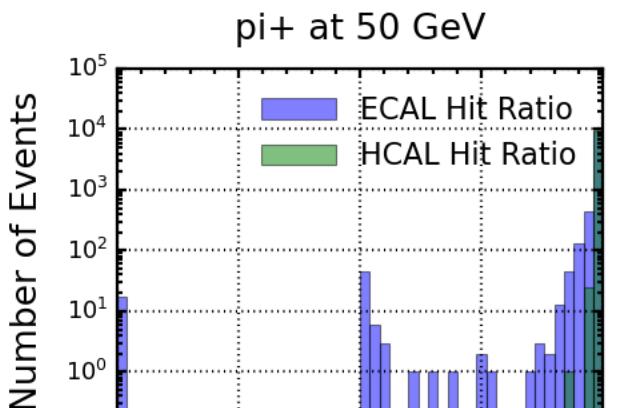

pi+ at 50 GeV
ECAL Energy Ratio
HCAL Energy Ratio
Number of Events
Energy Containment Ratio


pi+ at 50 GeV
ECAL Hit Ratio
HCAL Hit Ratio
Number of Events
Hit Containment Ratio

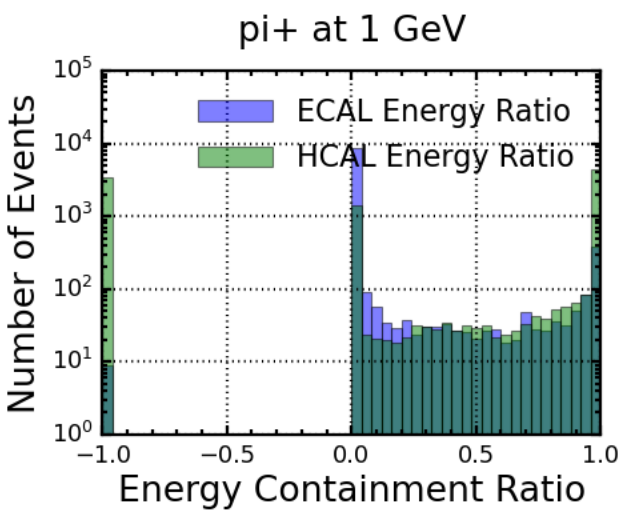

pi+ at 1 GeV
ECAL Energy Ratio
HCAL Energy Ratio
Number of Events

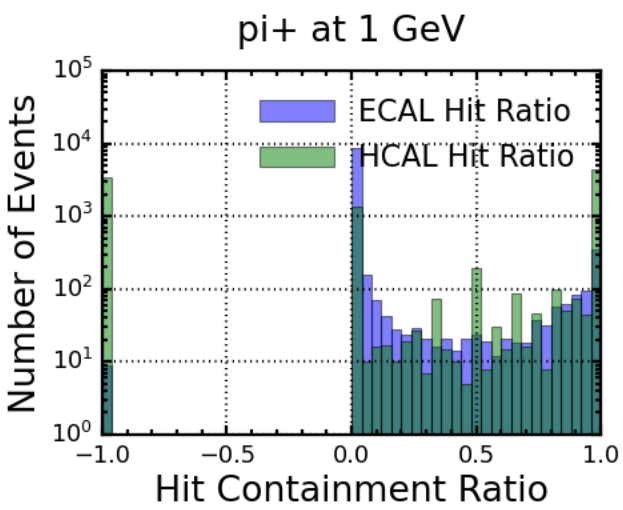

pi+ at 1 GeV
ECAL Hit Ratio
HCAL Hit Ratio
Number of Events

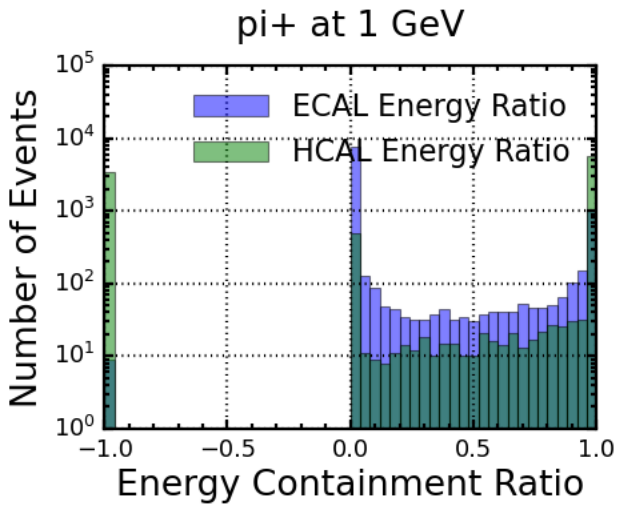

pi+ at 50 GeV
ECAL Energy Ratio
HCAL Energy Ratio
Number of Events

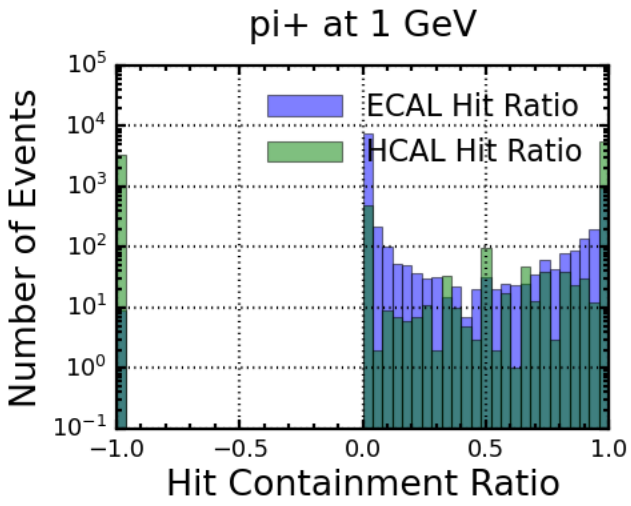

pi+ at 50 GeV
ECAL Hit Ratio
HCAL Hit Ratio
Number of Events

COG

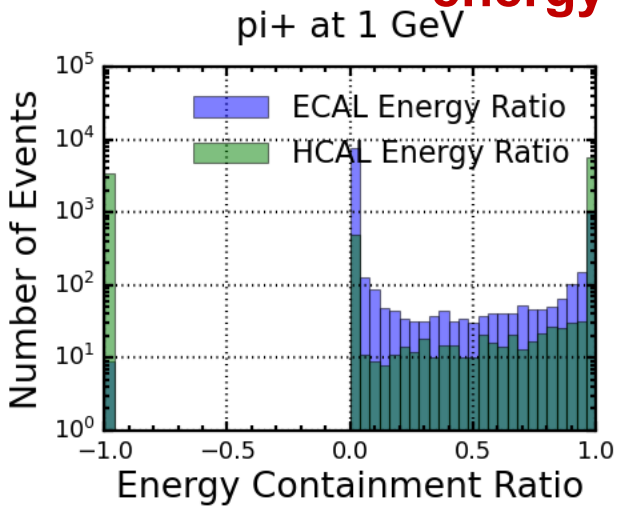


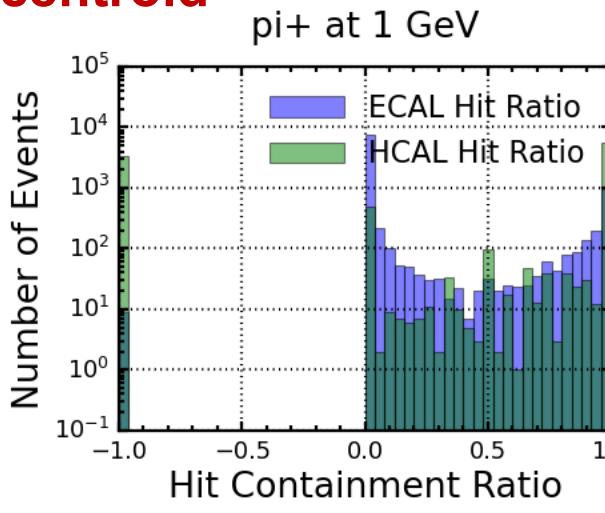


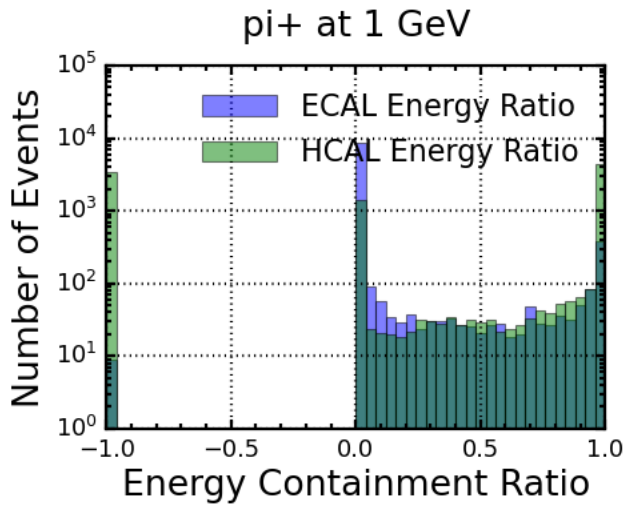


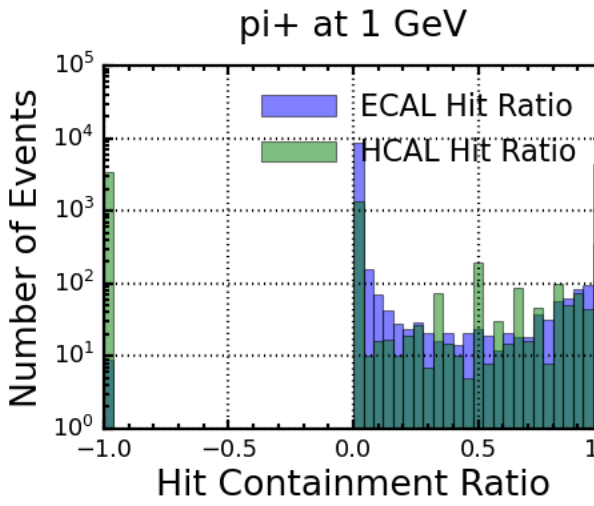


energy centroid

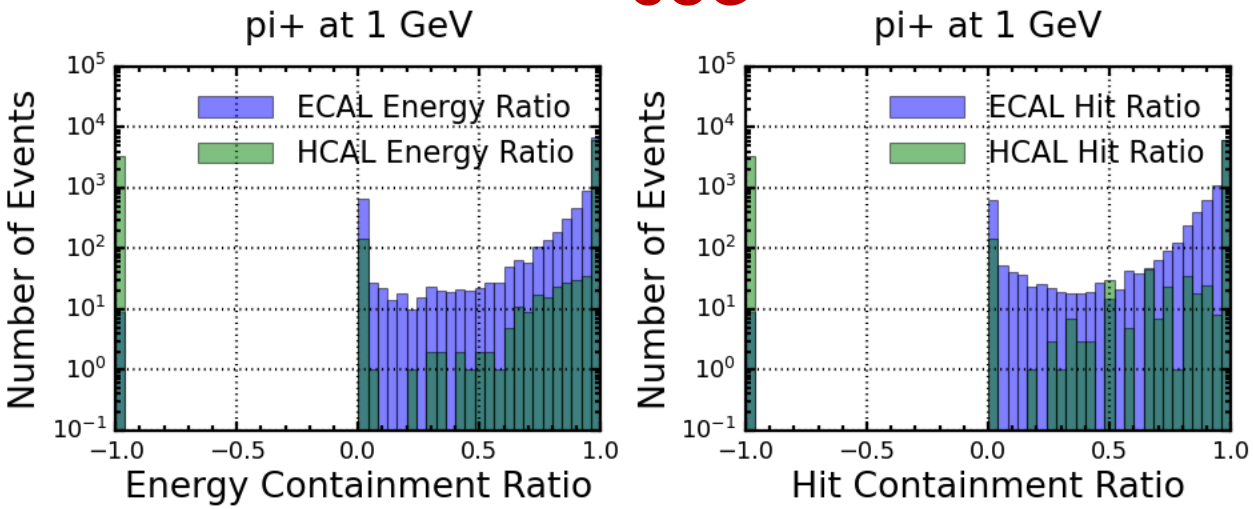


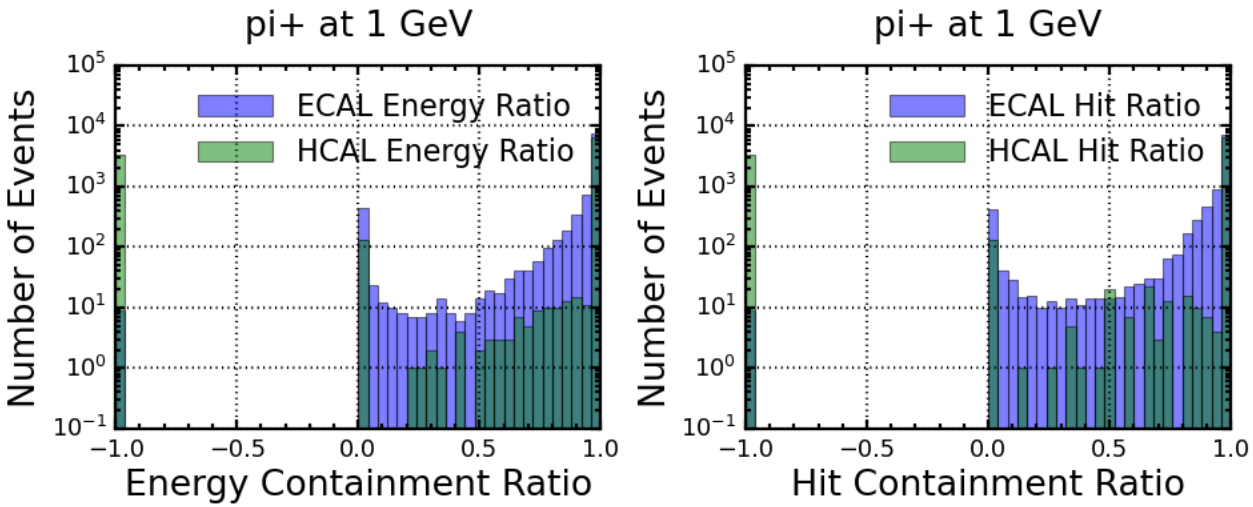




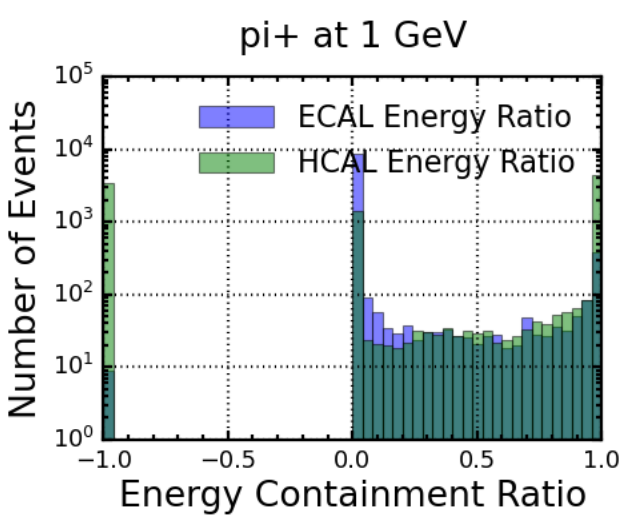


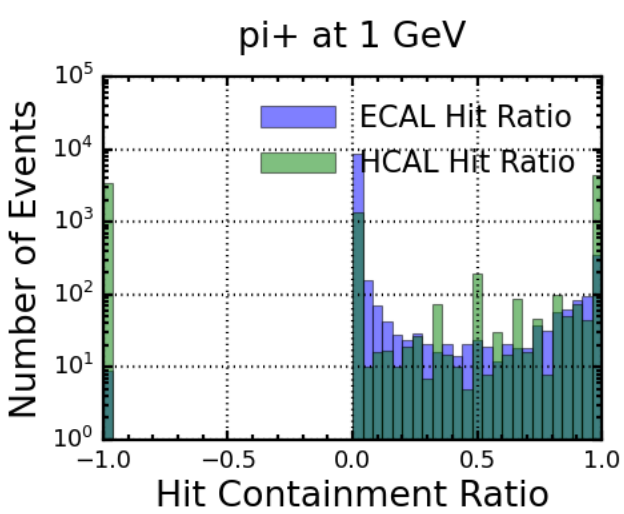
COG

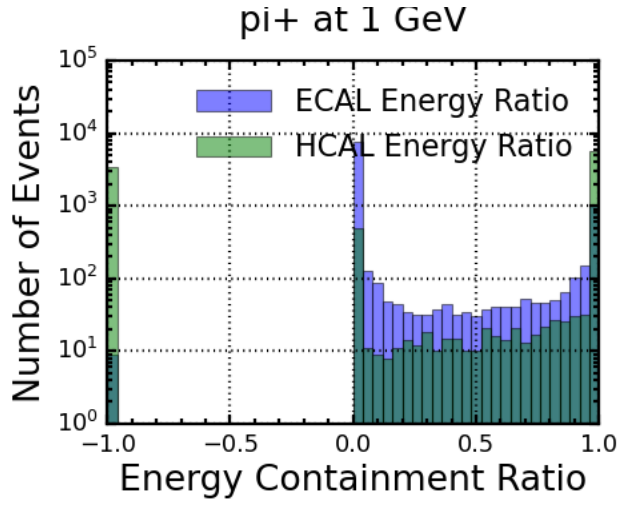


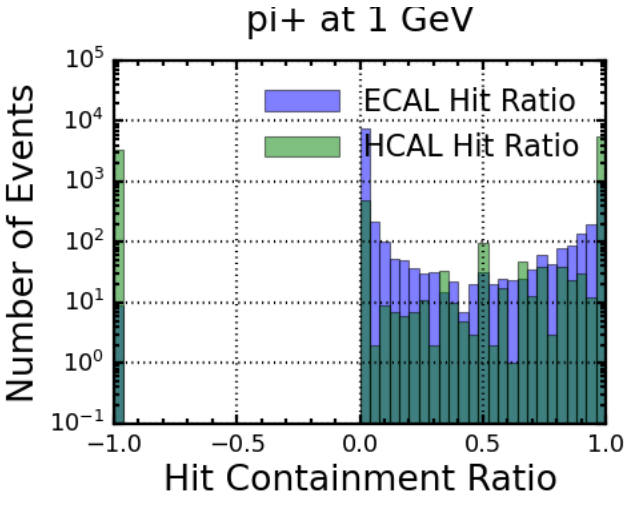


Injection position

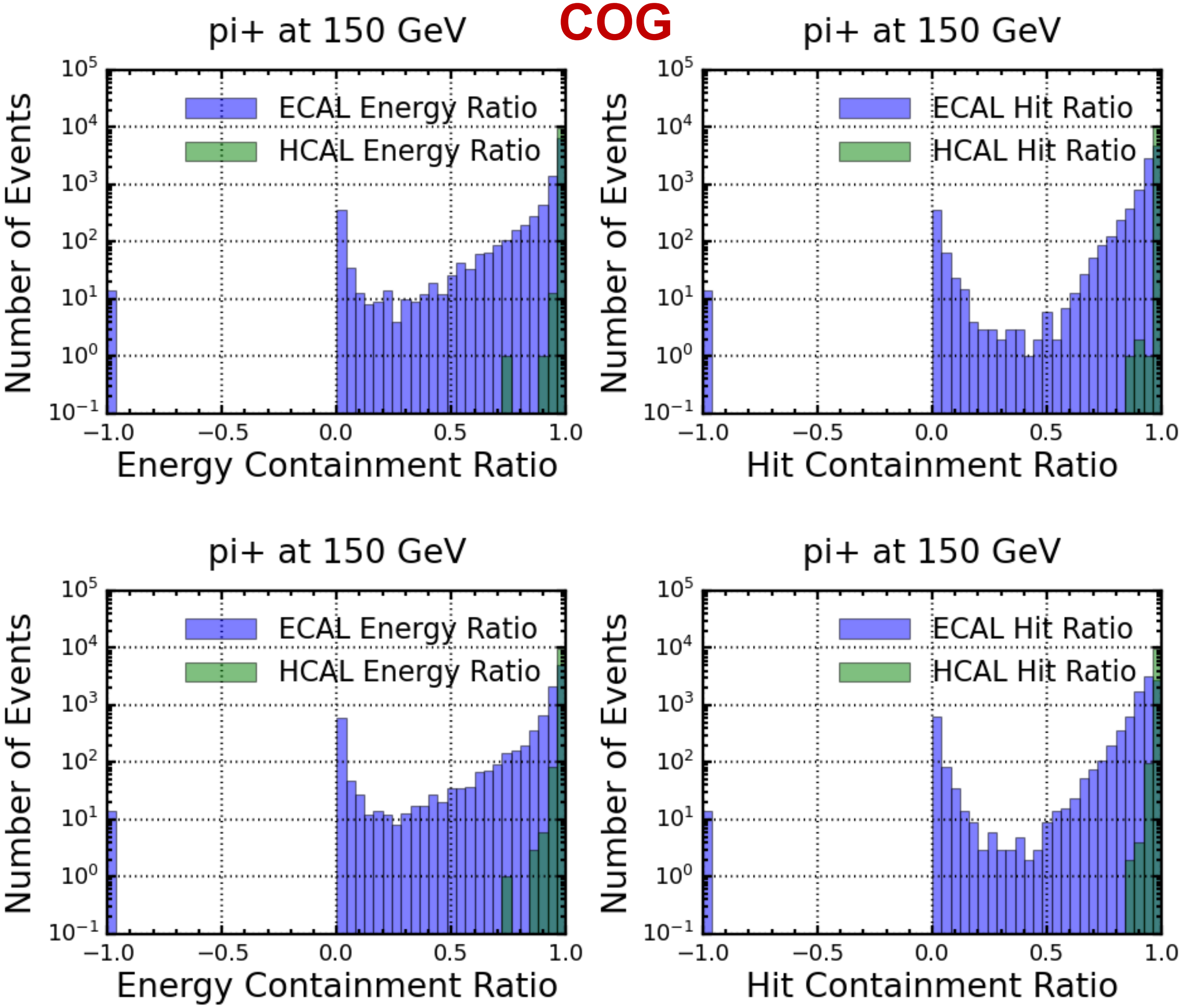




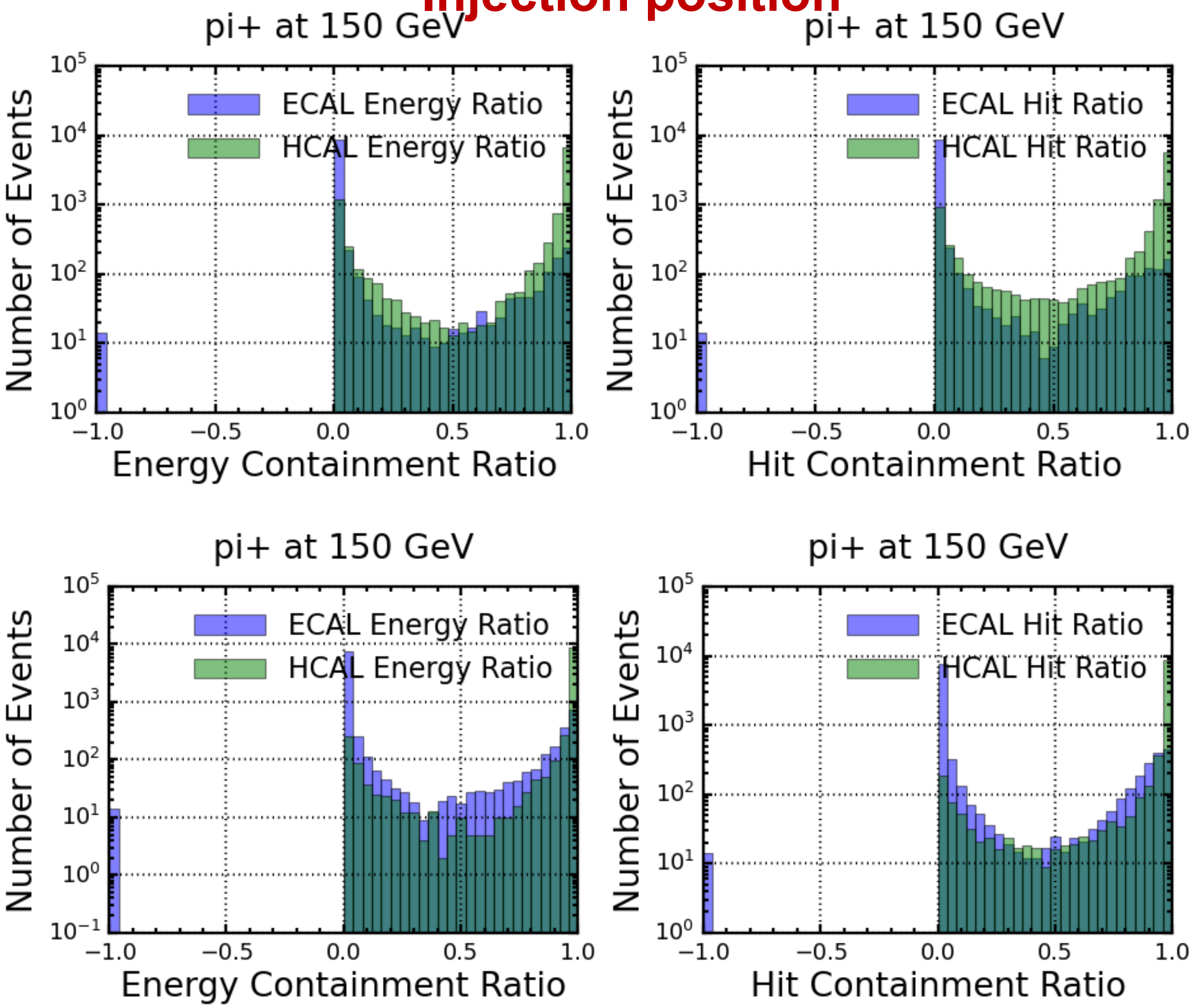


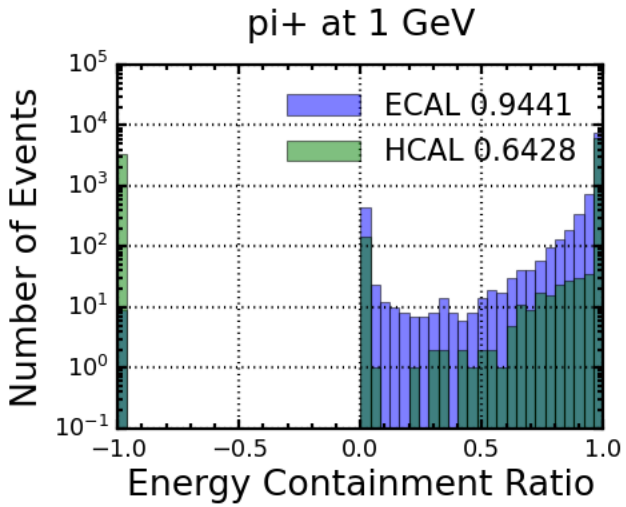


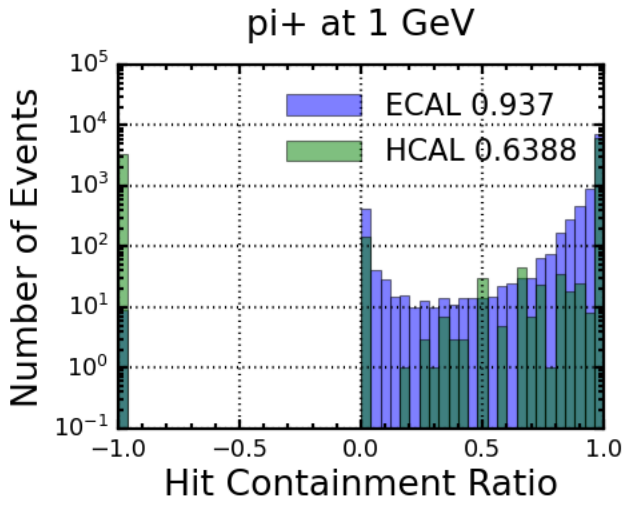
COG

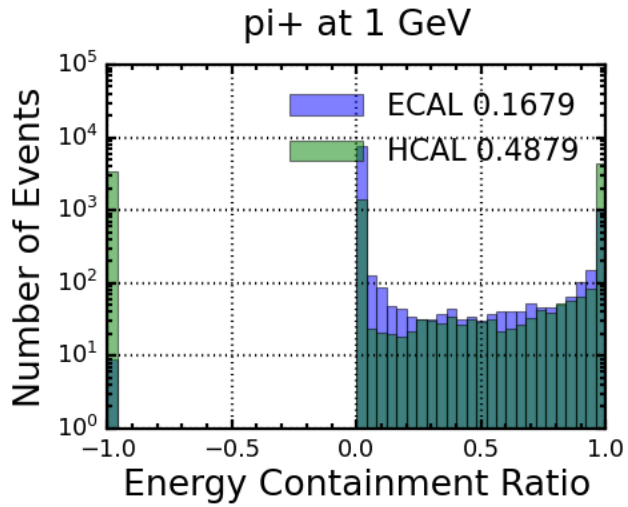


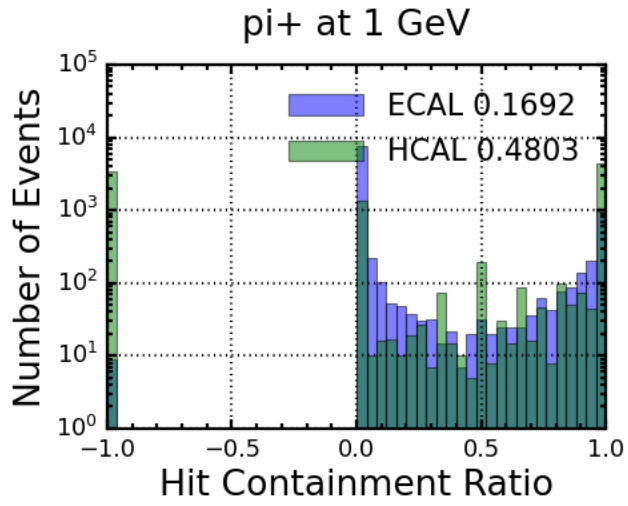
Injection position

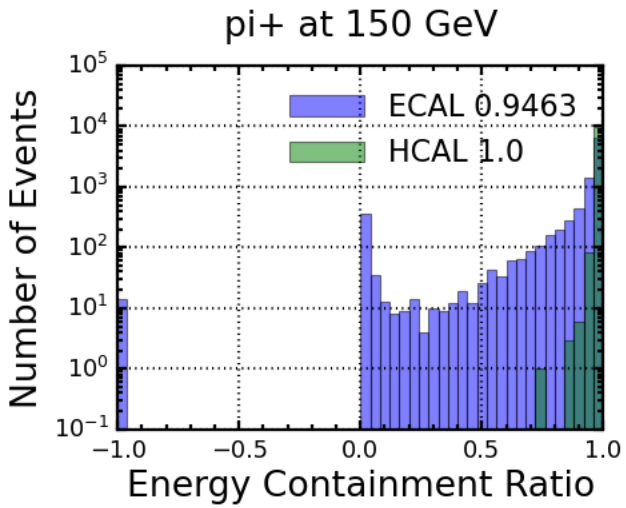


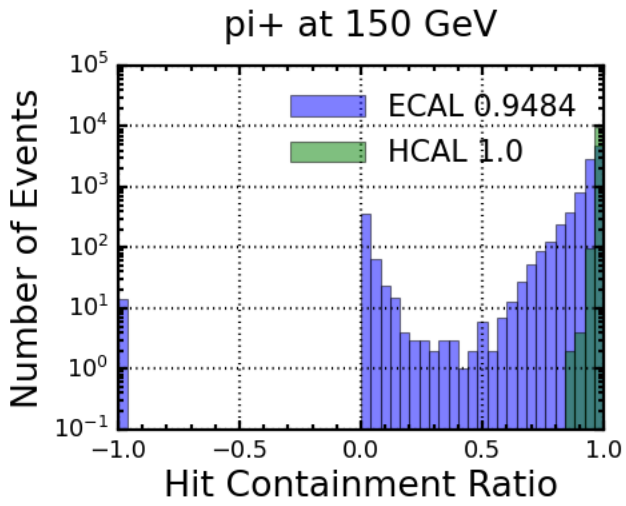

pi+ at 1 GeV
ECAL 0.9441
HCAL 0.6428
Number of Events
Energy Containment Ratio

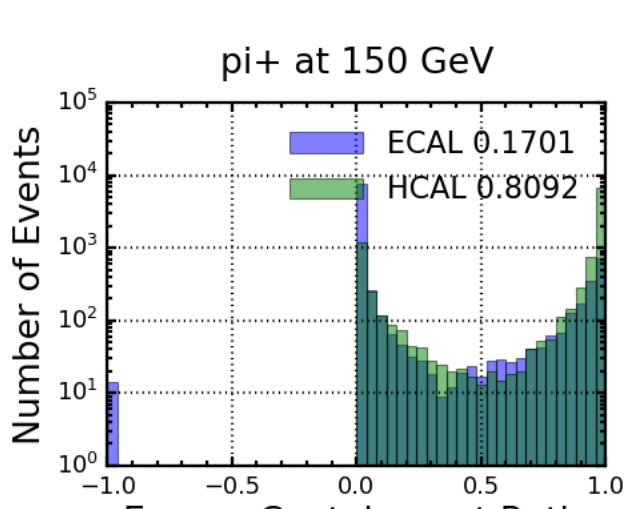

pi+ at 1 GeV
ECAL 0.937
HCAL 0.6388
Number of Events
Hit Containment Ratio

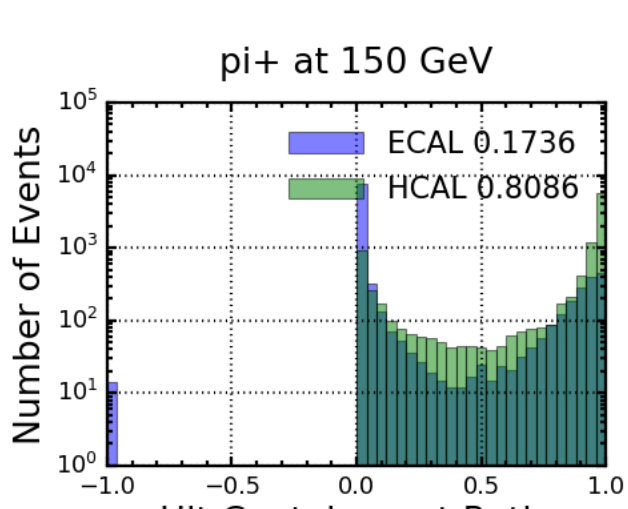

pi+ at 1 GeV
ECAL 0.1679
HCAL 0.4879
Number of Events
Energy Containment Ratio


pi+ at 1 GeV
ECAL 0.1692
HCAL 0.4803
Number of Events
Hit Containment Ratio

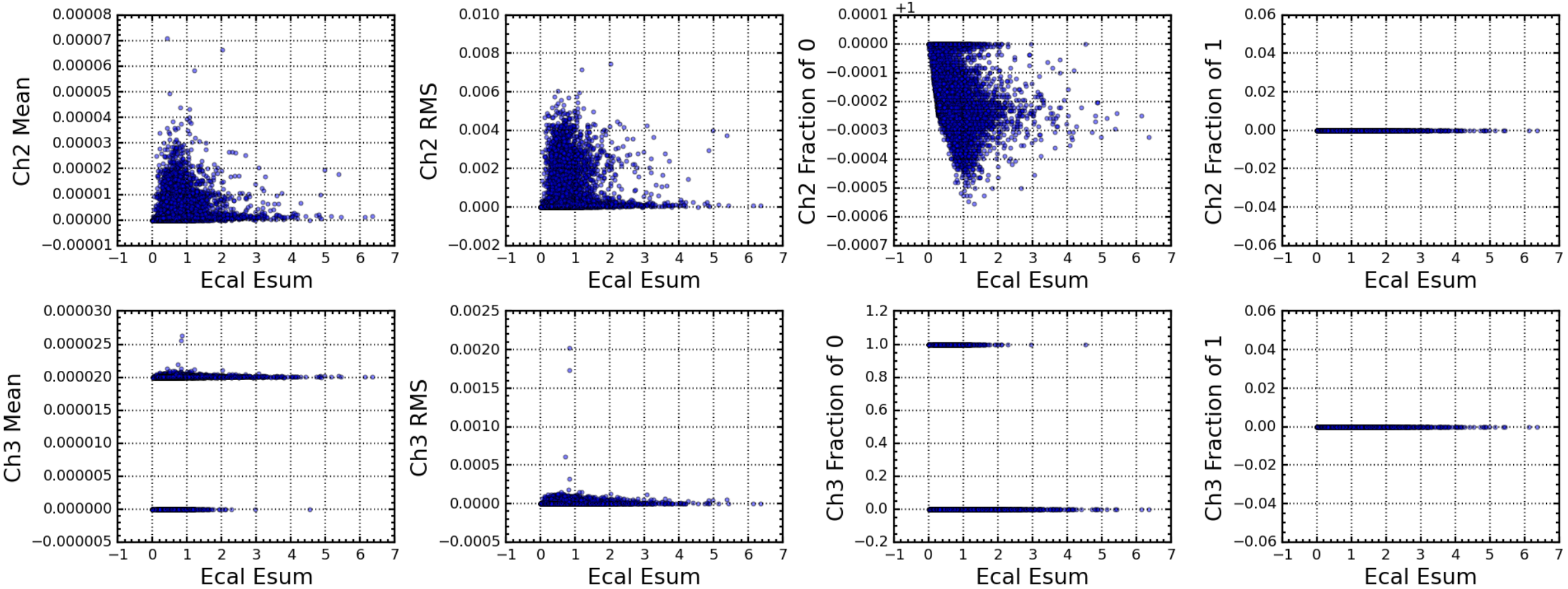

pi+ at 150 GeV
ECAL 0.9463
HCAL 1.0
Number of Events
Energy Containment Ratio


pi+ at 150 GeV
ECAL 0.9484
HCAL 1.0
Number of Events
Hit Containment Ratio

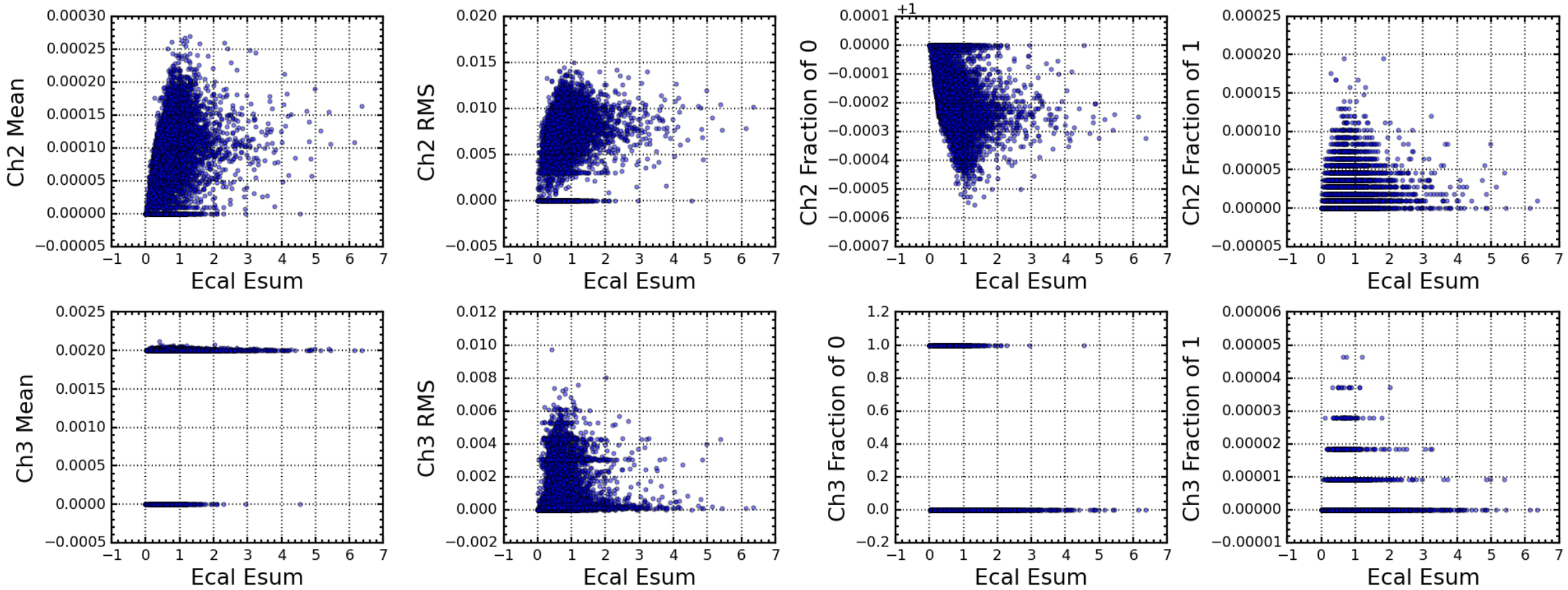

pi+ at 150 GeV
ECAL 0.1701
HCAL 0.8092
Number of Events
Energy Containment Ratio


pi+ at 150 GeV
ECAL 0.1736
HCAL 0.8086
Number of Events
Hit Containment Ratio

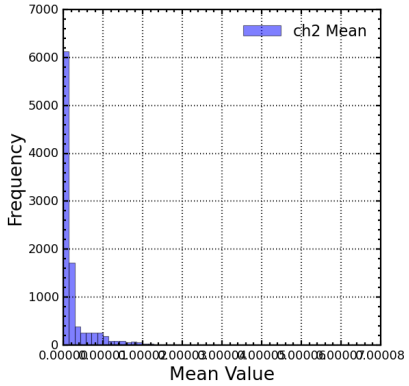
pi+ 1 GeV

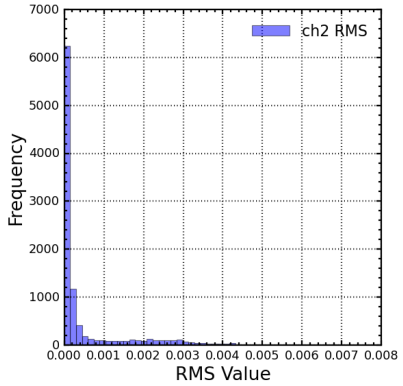

Ch2 Mean
Ch2 RMS
Ch2 Fraction of 0
Ch2 Fraction of 1
Ch3 Mean
Ch3 RMS
Ch3 Fraction of 0
Ch3 Fraction of 1
Ecal Esum
+1

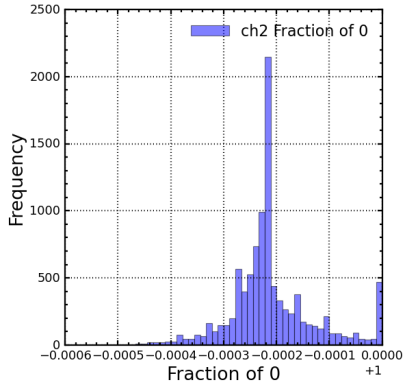
pi+ 1 GeV

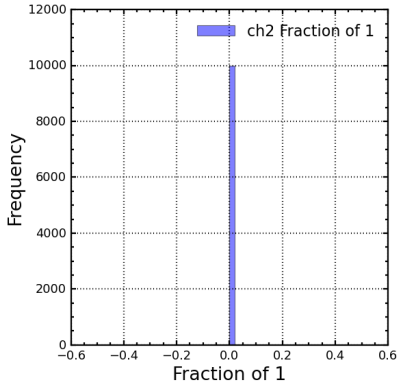

Ch2 Mean
Ch2 RMS
Ch2 Fraction of 0
+1
Ch2 Fraction of 1
Ch3 Mean
Ch3 RMS
Ch3 Fraction of 0
Ch3 Fraction of 1
Ecal Esum

pi+ 1 GeV

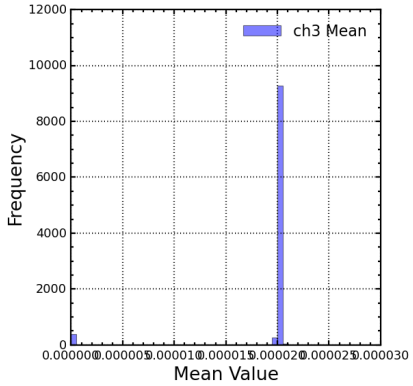


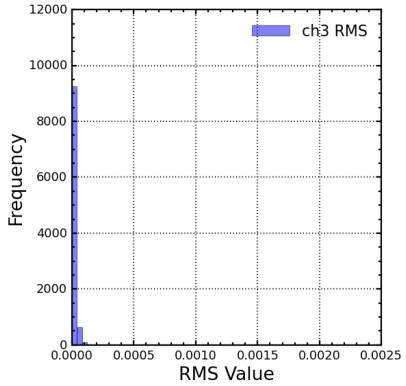


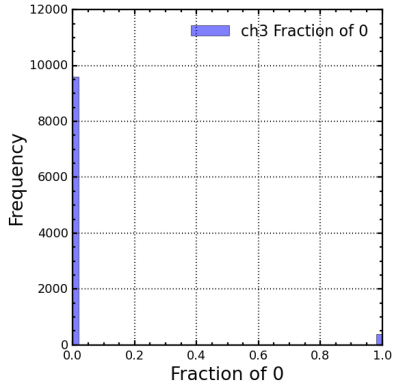


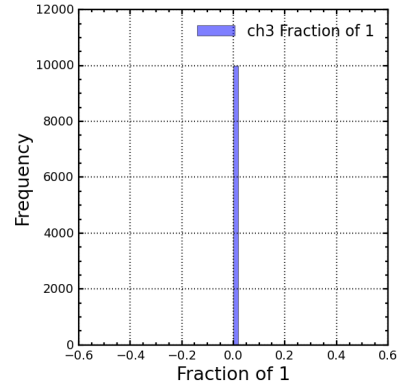


pi+ 1 GeV

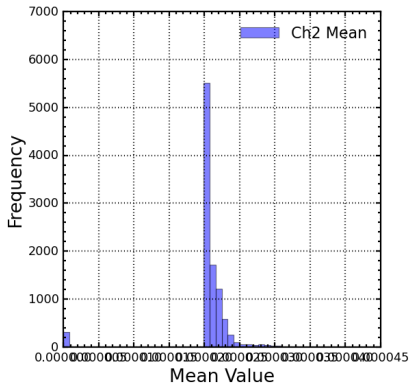


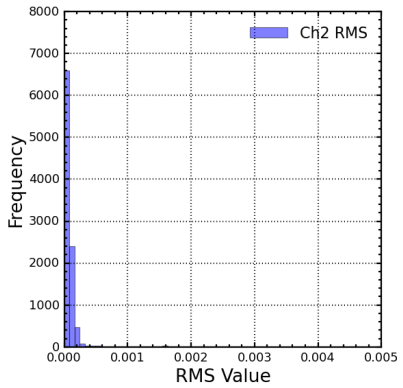


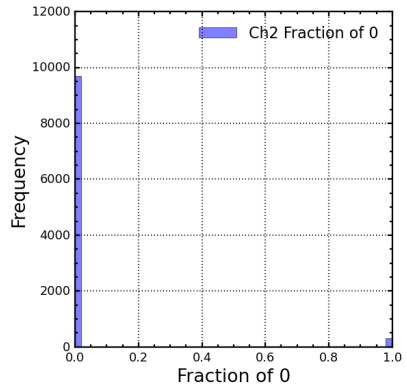


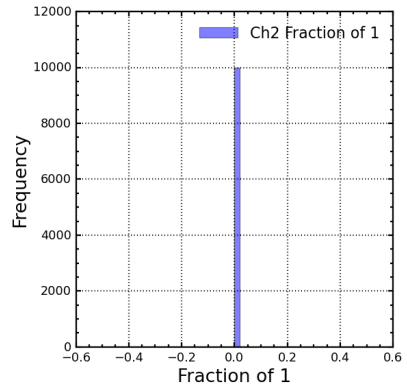


pi+ 150 GeV

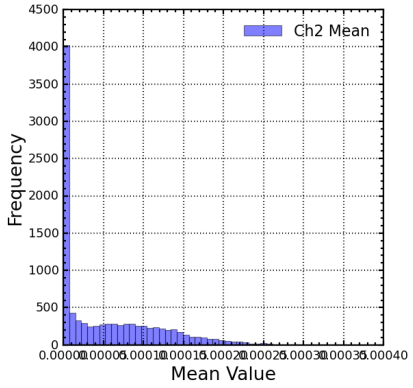


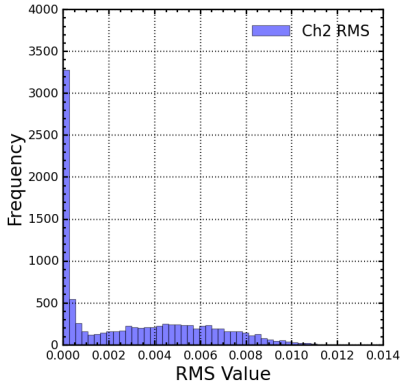


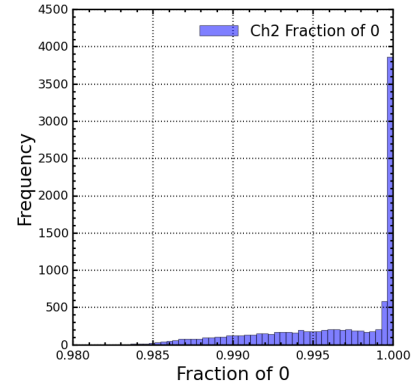


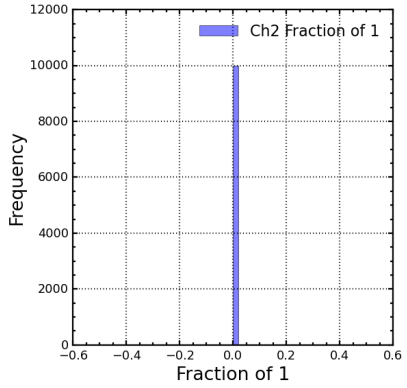


pi+ 150 GeV



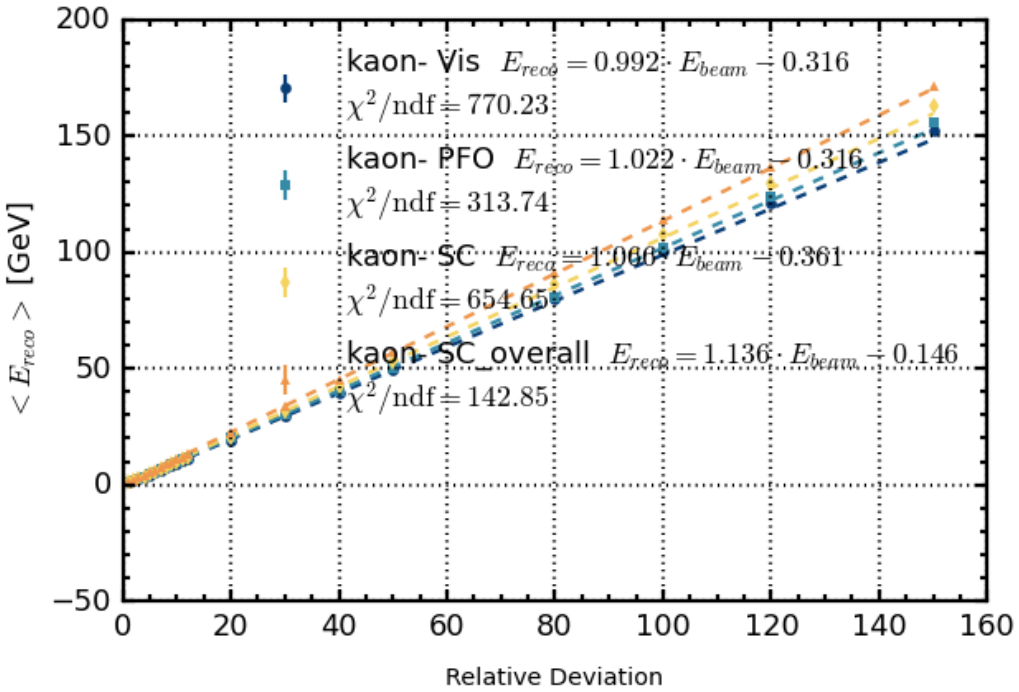


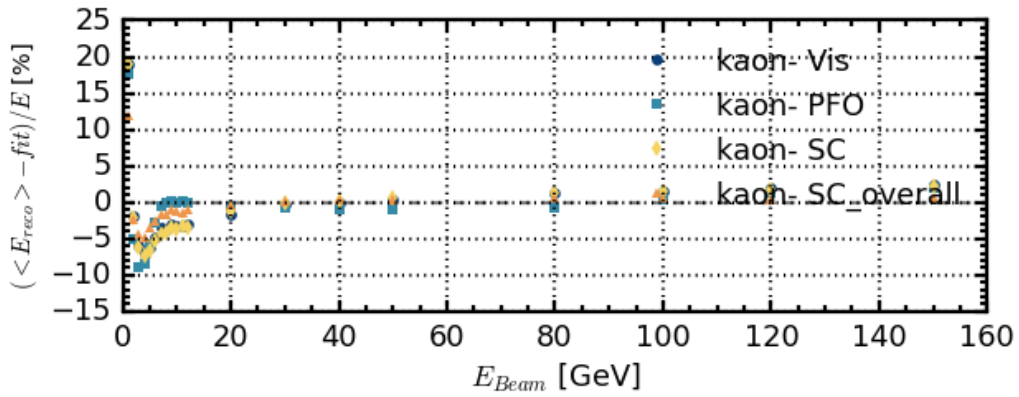


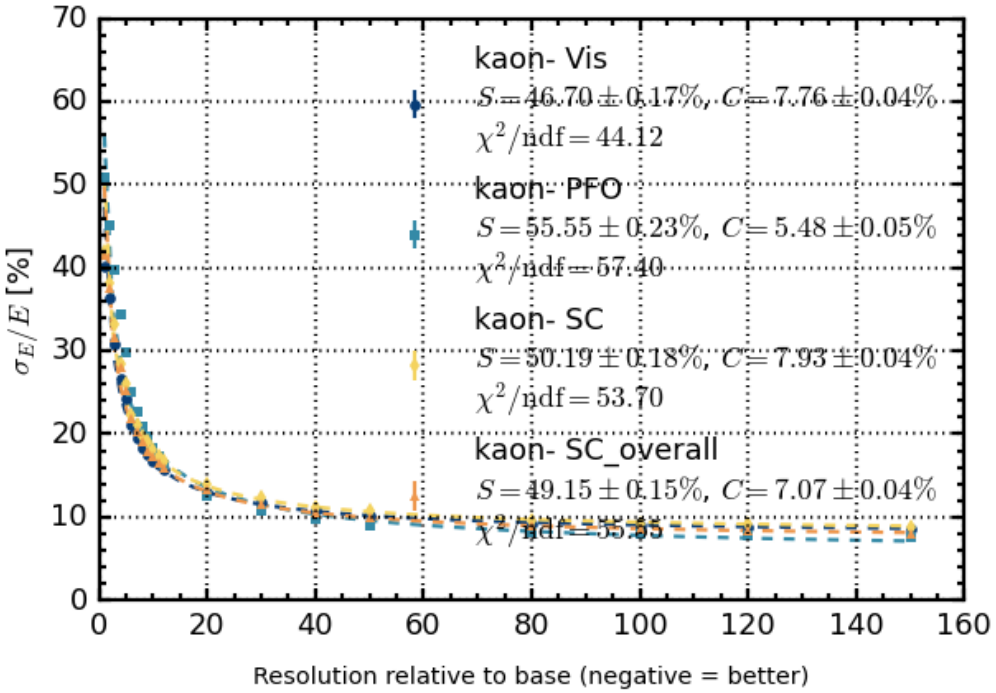


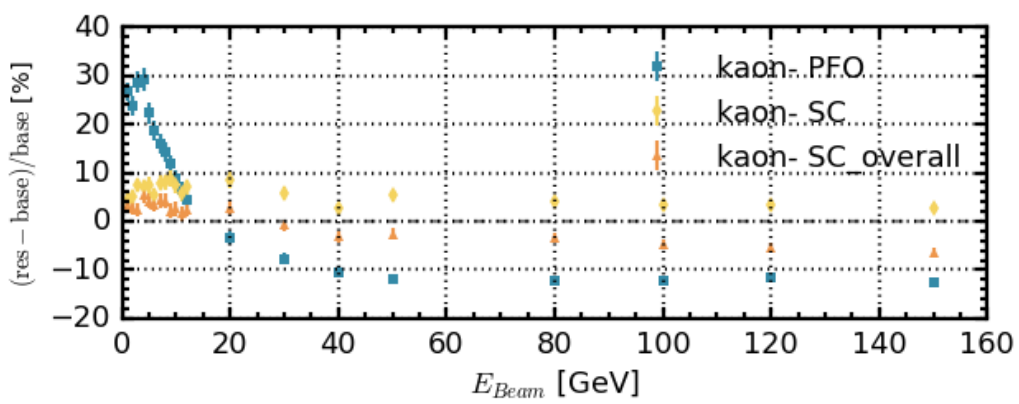
Backup

Kaon- samples

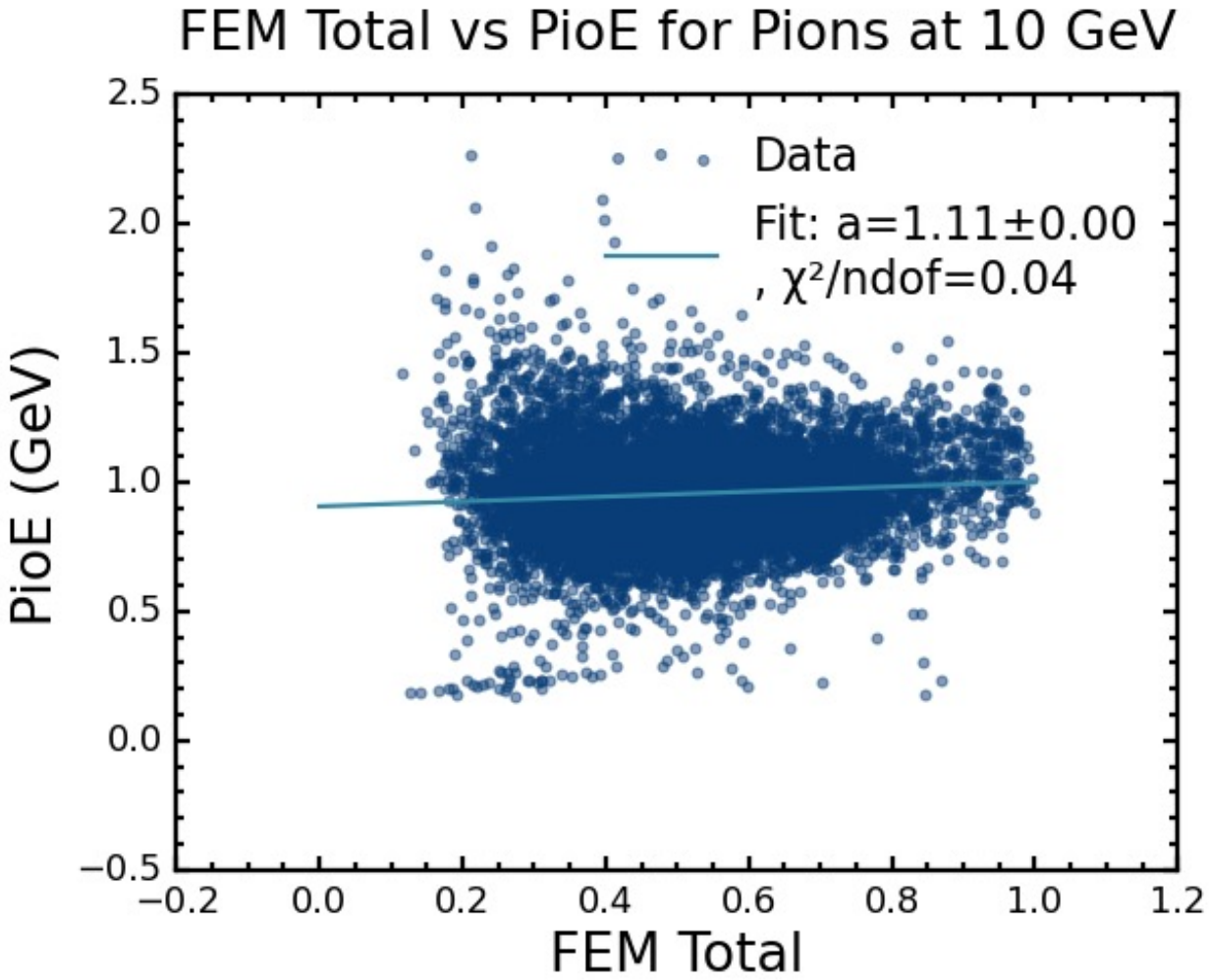








CNN(ML w/o time)	(E, x, y, z)	$\mathrm{ch}_0 = \frac{E_{\mathrm{voxel}}}{v_{max}}$, $\mathrm{ch}_1 = \frac{E_{\mathrm{voxel}}}{E_{\mathrm{total}}}$	
CNN (ML w time)	(E, T, x, y, z)	Moments	$T_{\mathrm{mean}} = \frac{\sum_{i \in \mathrm{voxel}} E_i t_i}{\sum_{i \in \mathrm{voxel}} E_i + \epsilon}$
		Bins	$T_{\mathrm{rms}} = \sqrt{\frac{\sum_{i \in \mathrm{voxel}} E_i t_i^2}{\sum_{i \in \mathrm{voxel}} E_i + \epsilon} - T_{\mathrm{mean}}^2 + \epsilon}$


FEM Total vs PioE for Pions at 10 GeV
Data
Fit: a=1.11±0.00
, χ²/ndof=0.04
PioE (GeV)
FEM Total
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2