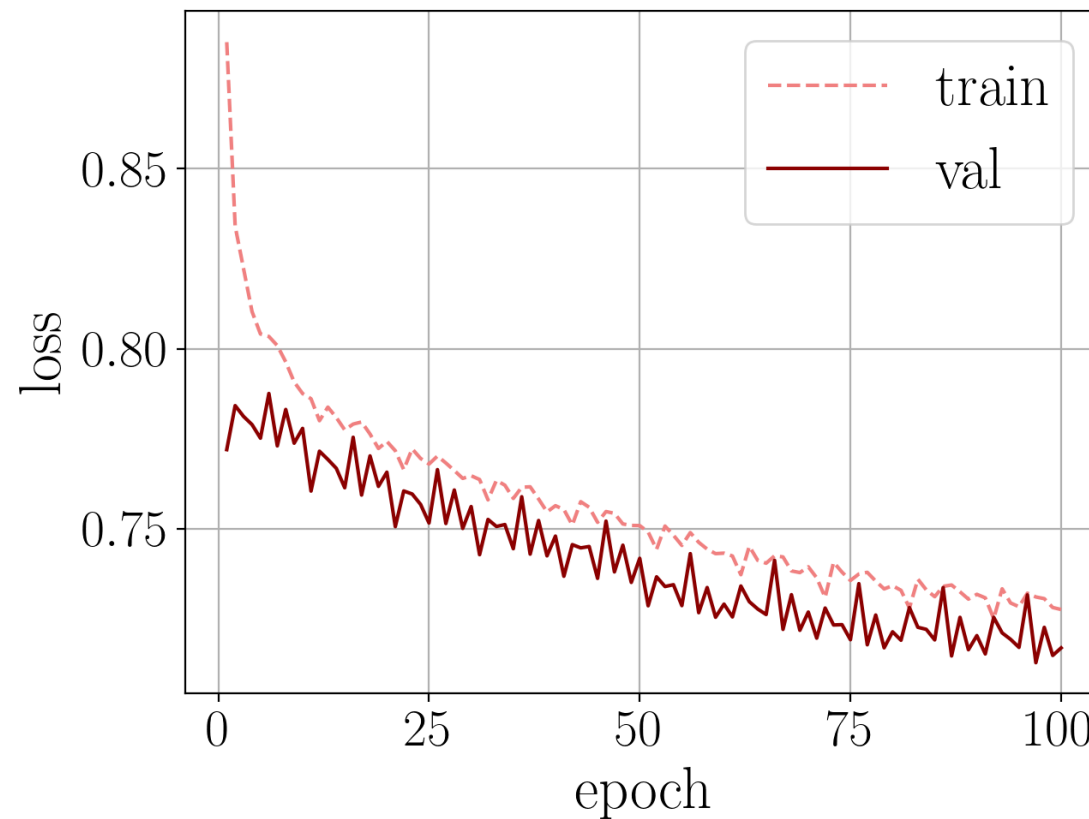
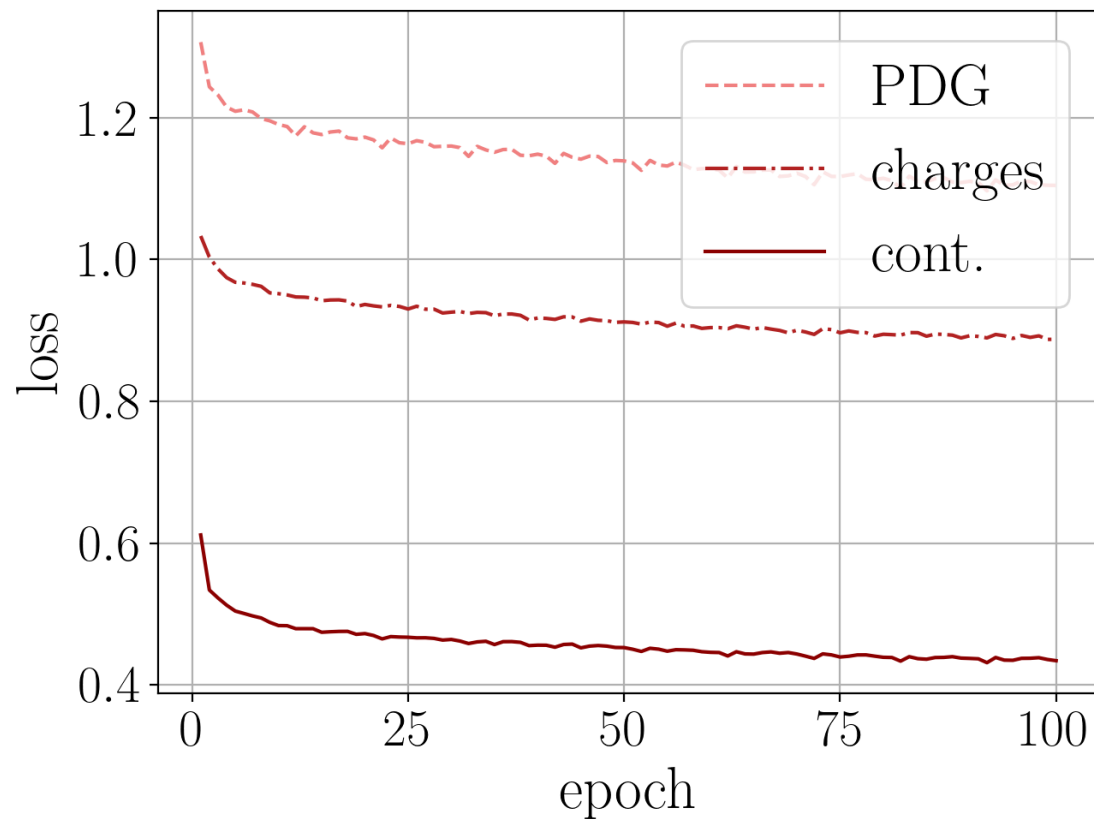


# ILC Weekly Meeting

06.12.2024

# First model

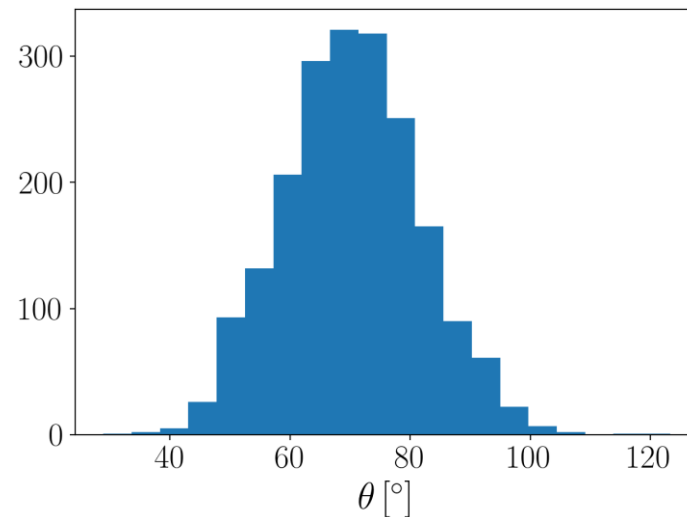
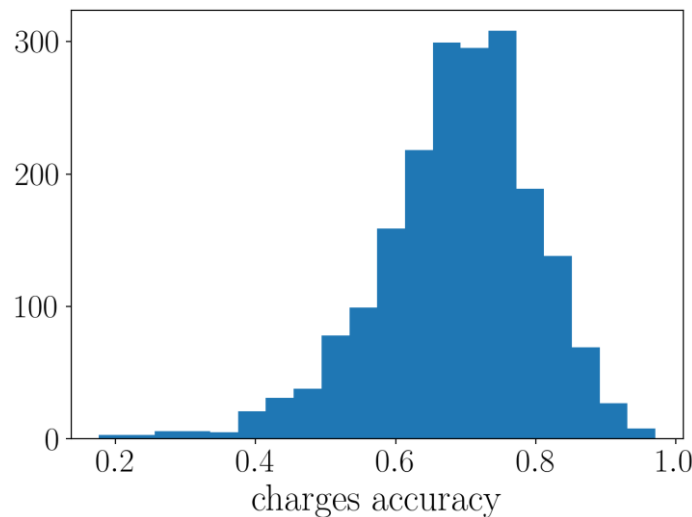
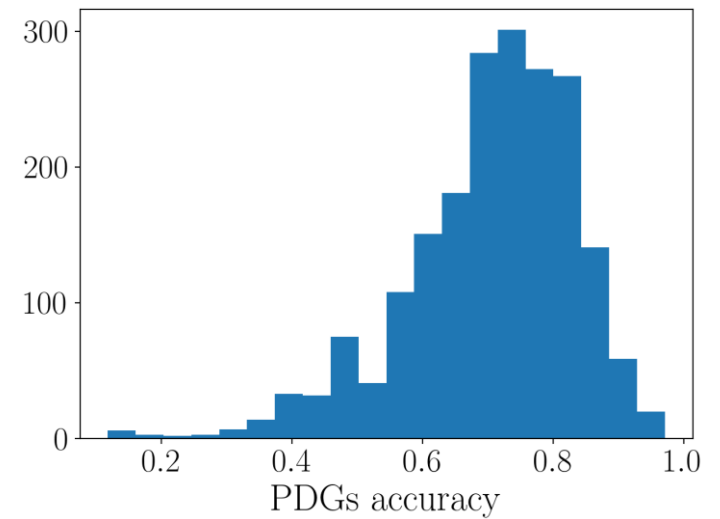
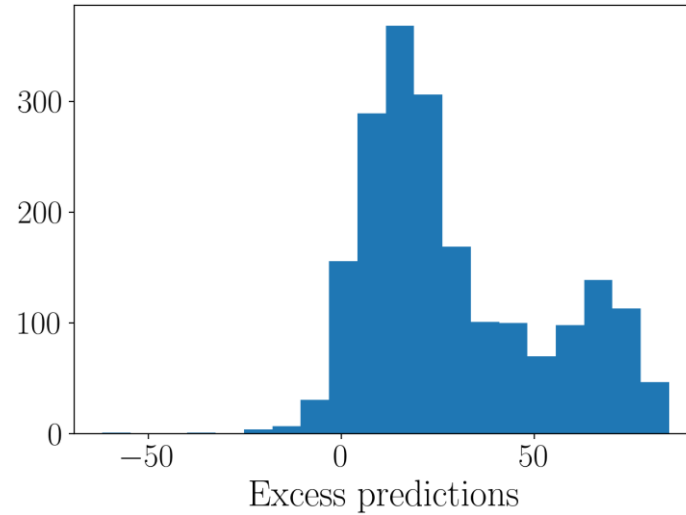
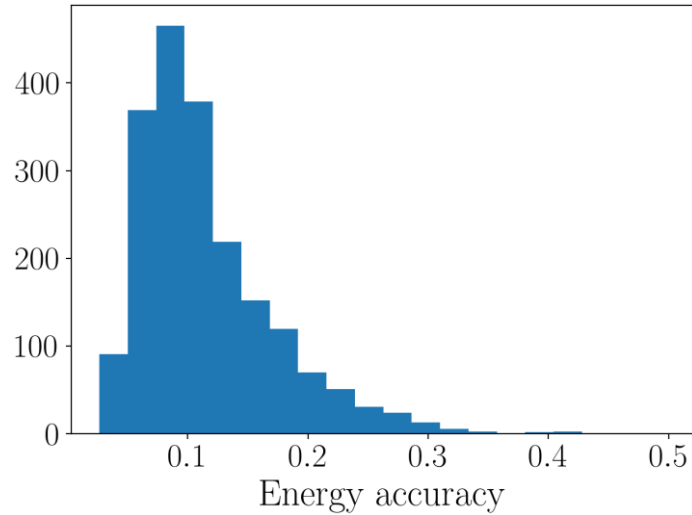
dmodel	nhead	Dim FF	Layers encoder	Layers decoder	Layers embedder	Parameters	Weights loss	Learning rate
256	8	512	1	1	2	1.6 M	[0.3, 0.3, 0.3]	1e-4



Training time per epoch: ~660 s

# First model

dmodel	nhead	Dim FF	Layers encoder	Layers decoder	Layers embedder	Parameters	Weights loss	Learning rate
256	8	512	1	1	2	1.6 M	[0.3, 0.3, 0.3]	1e-4

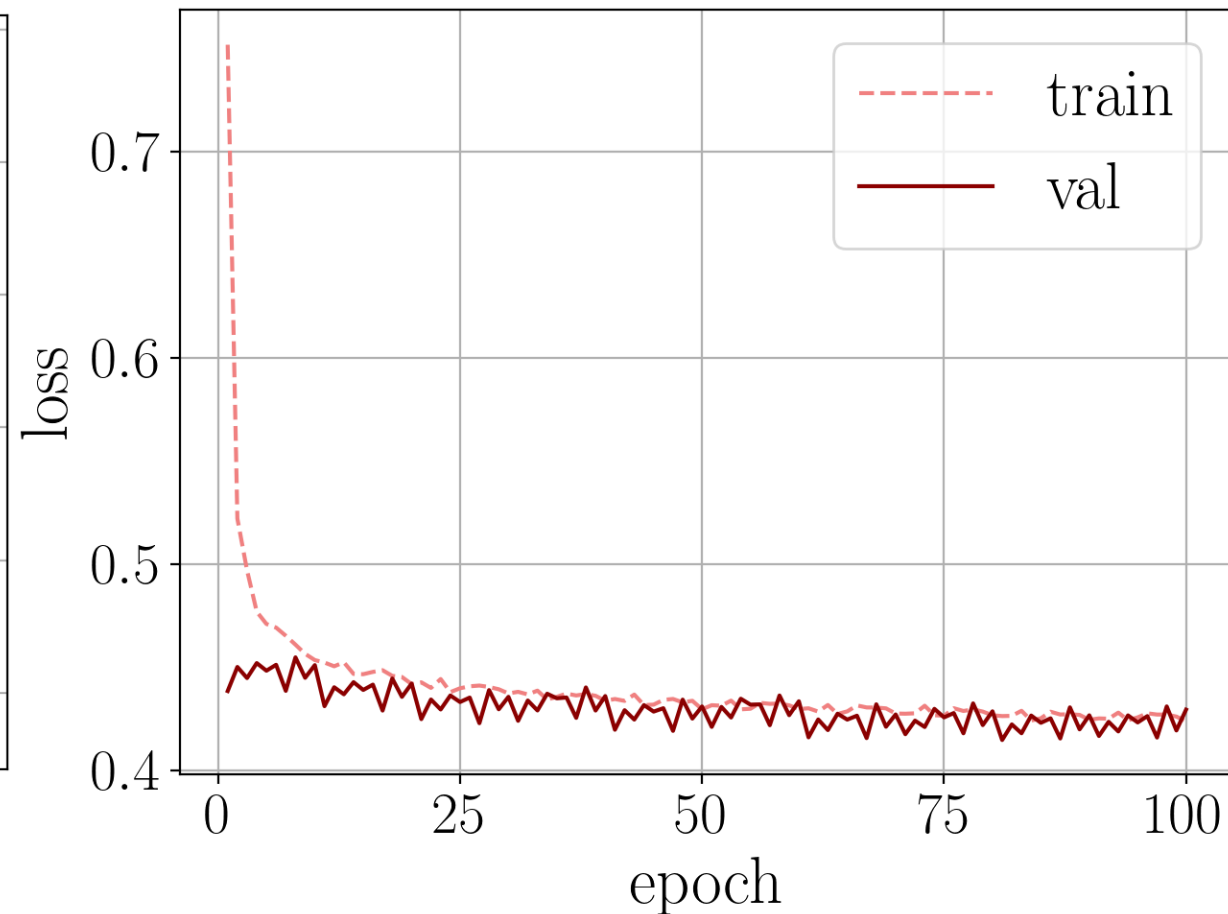
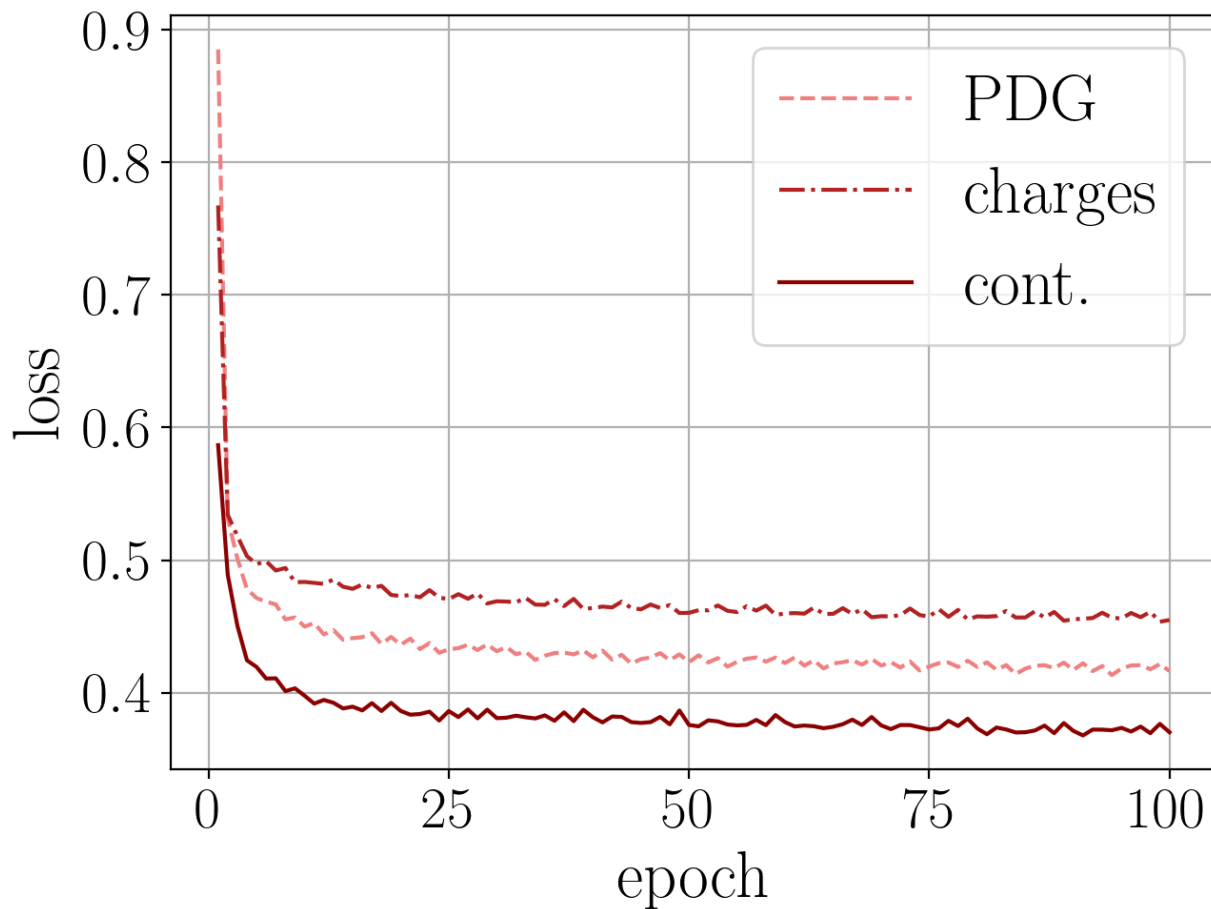


Average over 2000 events:

- Charge: 0.696
- PDGs: 0.718
- Energy: 0.07
- Theta: 70.29
- Excess: 29.25

# Second model

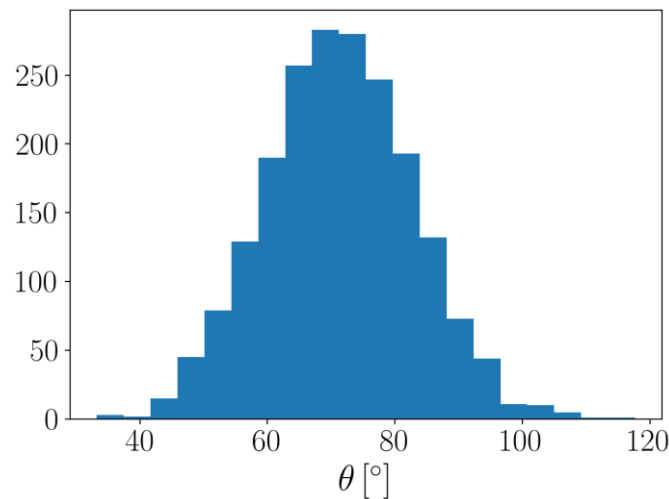
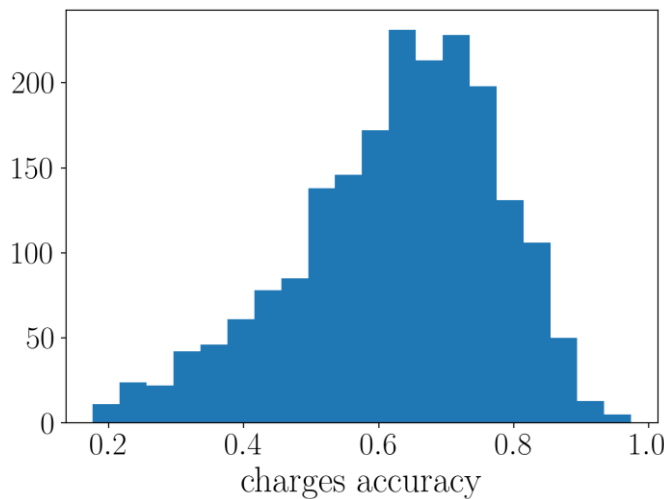
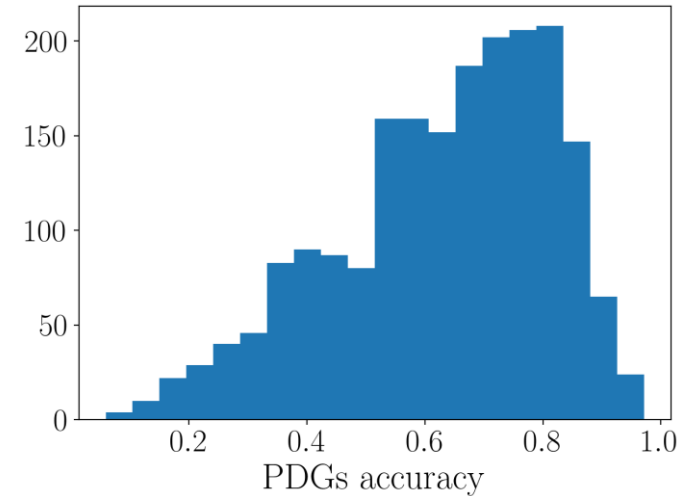
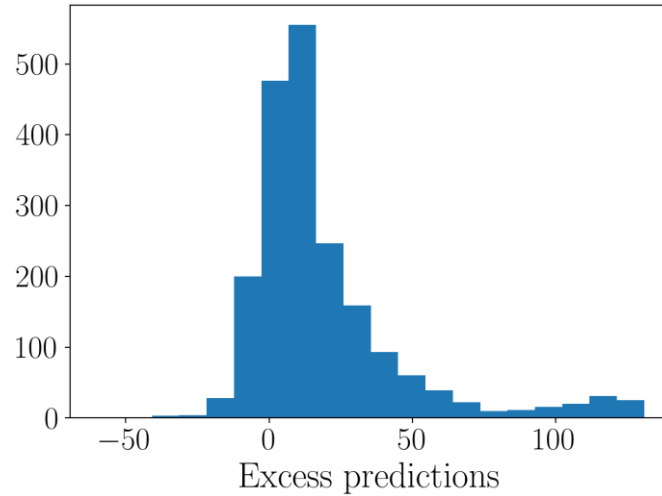
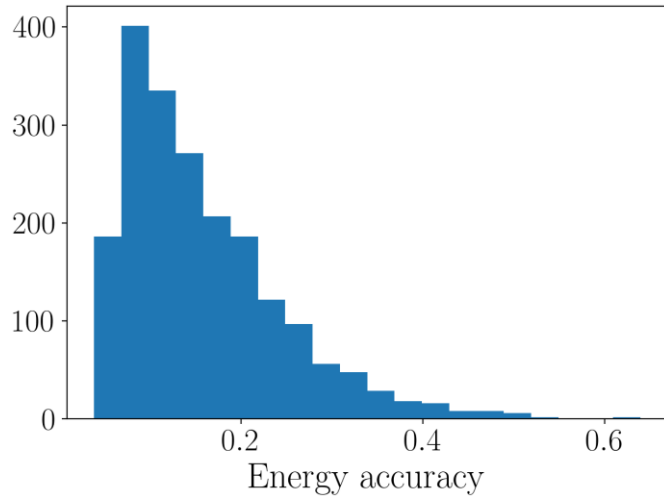
dmodel	nhead	Dim FF	Layers encoder	Layers decoder	Layers embedder	Parameters	Weights loss	Learning rate
128	8	256	3	3	2	1.06 M	[0.5,0.25,0.25]	1e-4



Training time per epoch: ~1850 s

# Second model

dmodel	nhead	Dim FF	Layers encoder	Layers decoder	Layers embedder	Parameters	Weights loss	Learning rate
128	8	256	3	3	2	1.06 M	[0.5,0.25,0.25]	1e-4



Average over 2000 events:

- Charge: 0.637
- PDGs: 0.649
- Energy: 0.102
- Theta: 71.3
- Excess: 18.95

# Comparison

dmodel	nhead	Dim FF	Layers encoder	Layers decoder	Layers embedder	Parameters	Weights loss	Learning rate
256	8	512	1	1	2	1.6 M	[0.3, 0.3, 0.3]	1e-4
128	8	256	3	3	2	1.06 M	[0.5,0.25,0.25]	1e-4

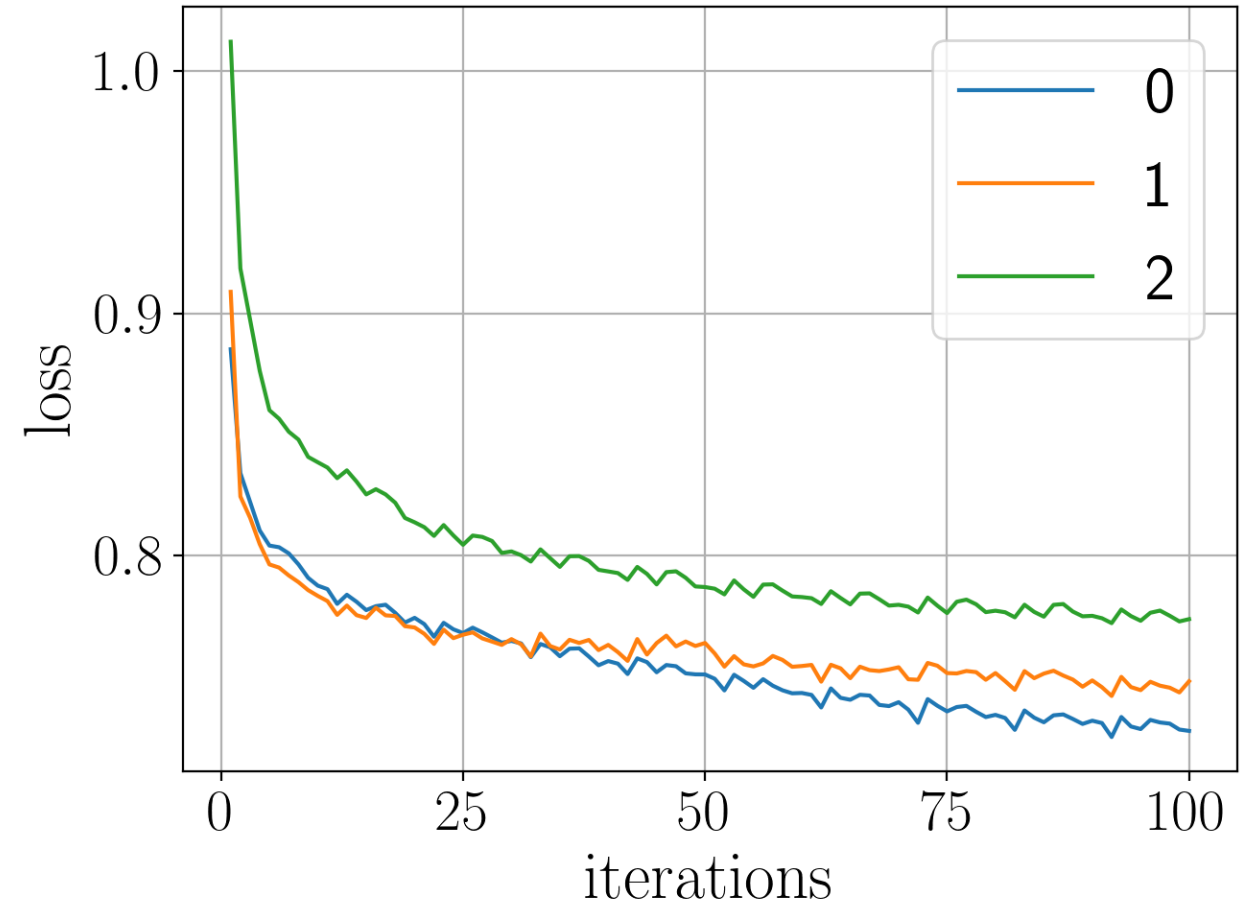
Charges	PDGs	Energy	Direction	Excess
0.696	0.718	0.07	70.29	29.25
0.637	0.649	0.102	71.3	18.95

# Learning rate

dmodel	nhead	Dim FF	Layers encoder	Layers decoder	Layers embedder	Parameters	Weights loss
256	8	512	1	1	2	1.6 M	[0.3, 0.3, 0.3]

0: lr = 1e-4  
1: lr = 1e-3  
2: lr = 1e-5

- Done on 2% of the Dataset (16'000 events)
- 10 epochs



# Other remarks

- Loading the files get really slow for larger dataset due to awkward array concatenation
- Changed loading method -> better performances but still too slow
- Also pre-processing will take a lot of time
- Loading and preprocessing for each batch?
  - Makes it more difficult to inverse normalization on energy