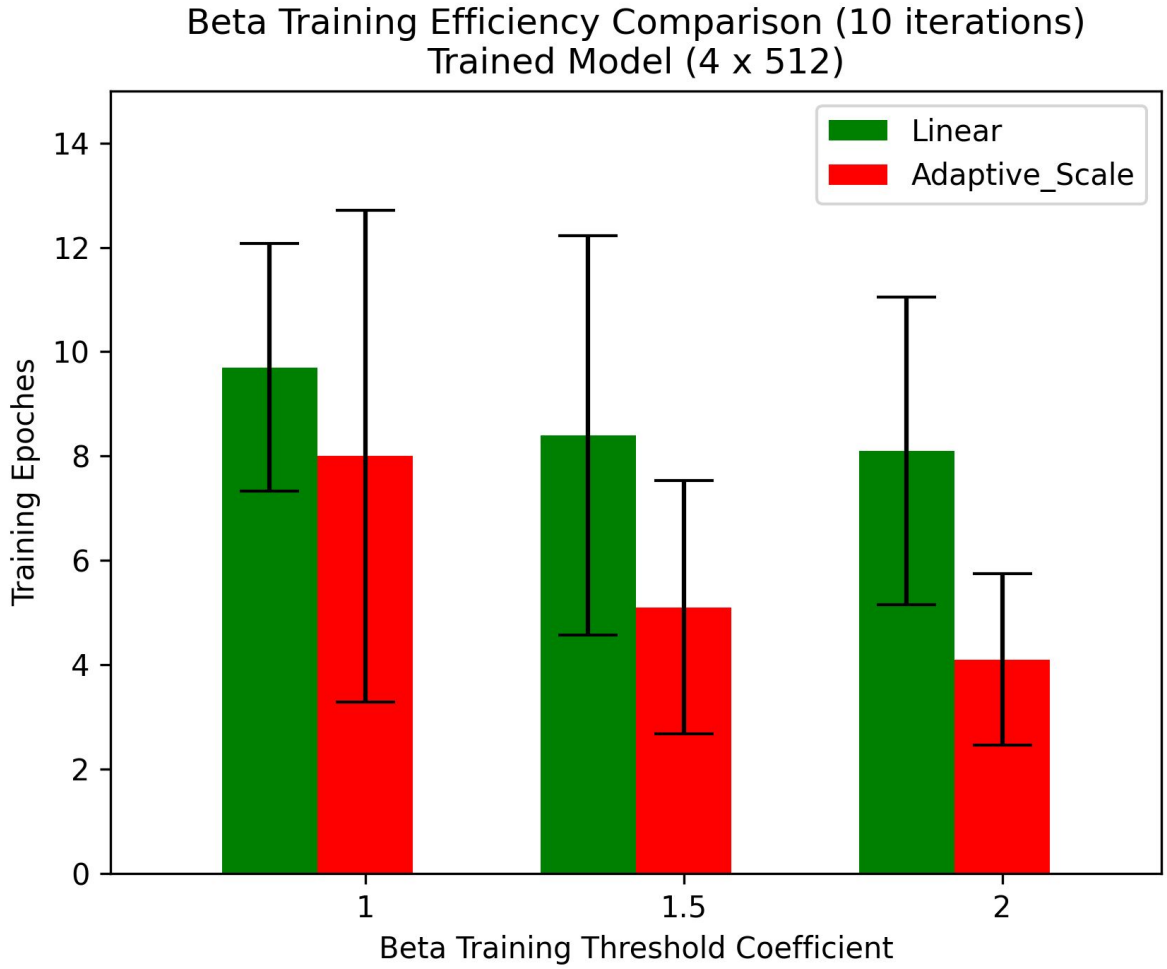
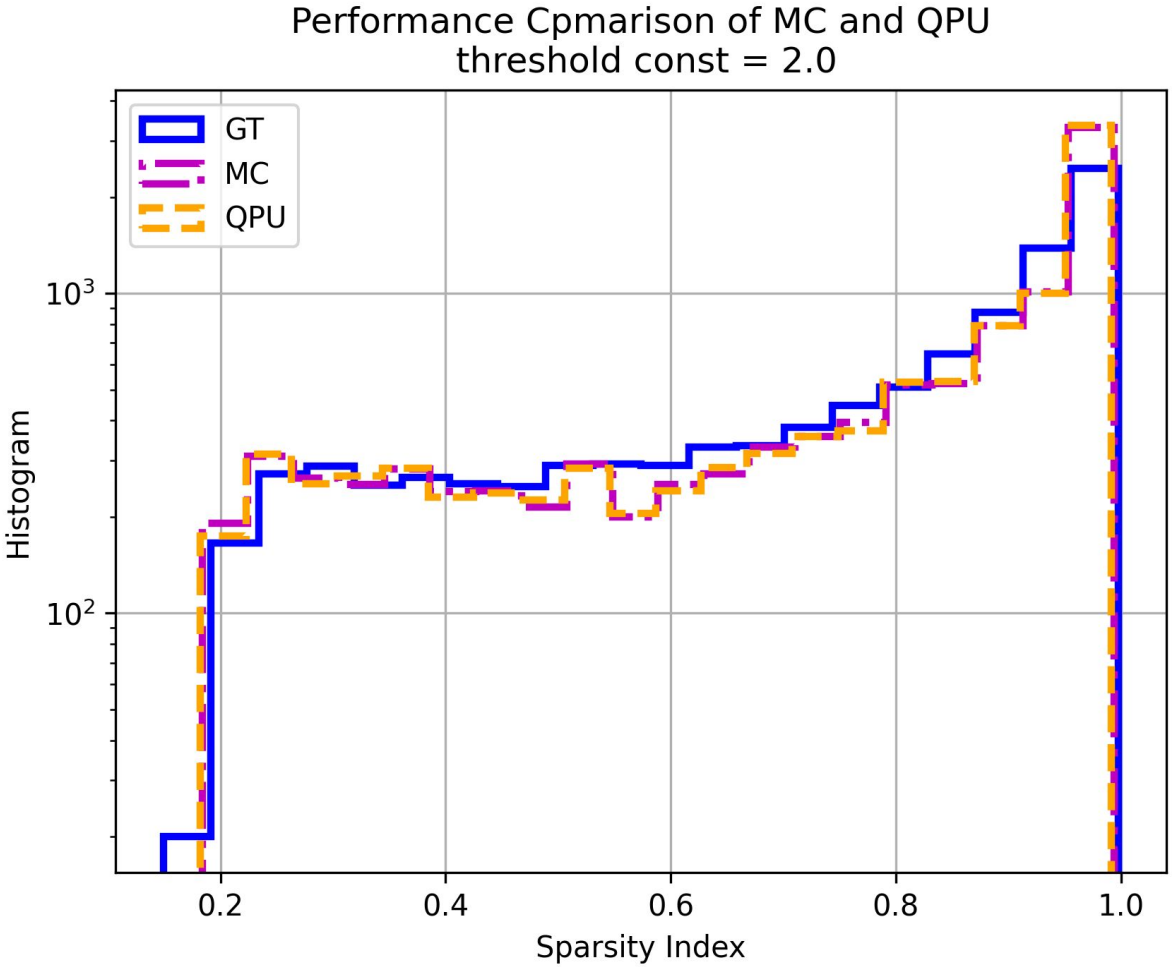
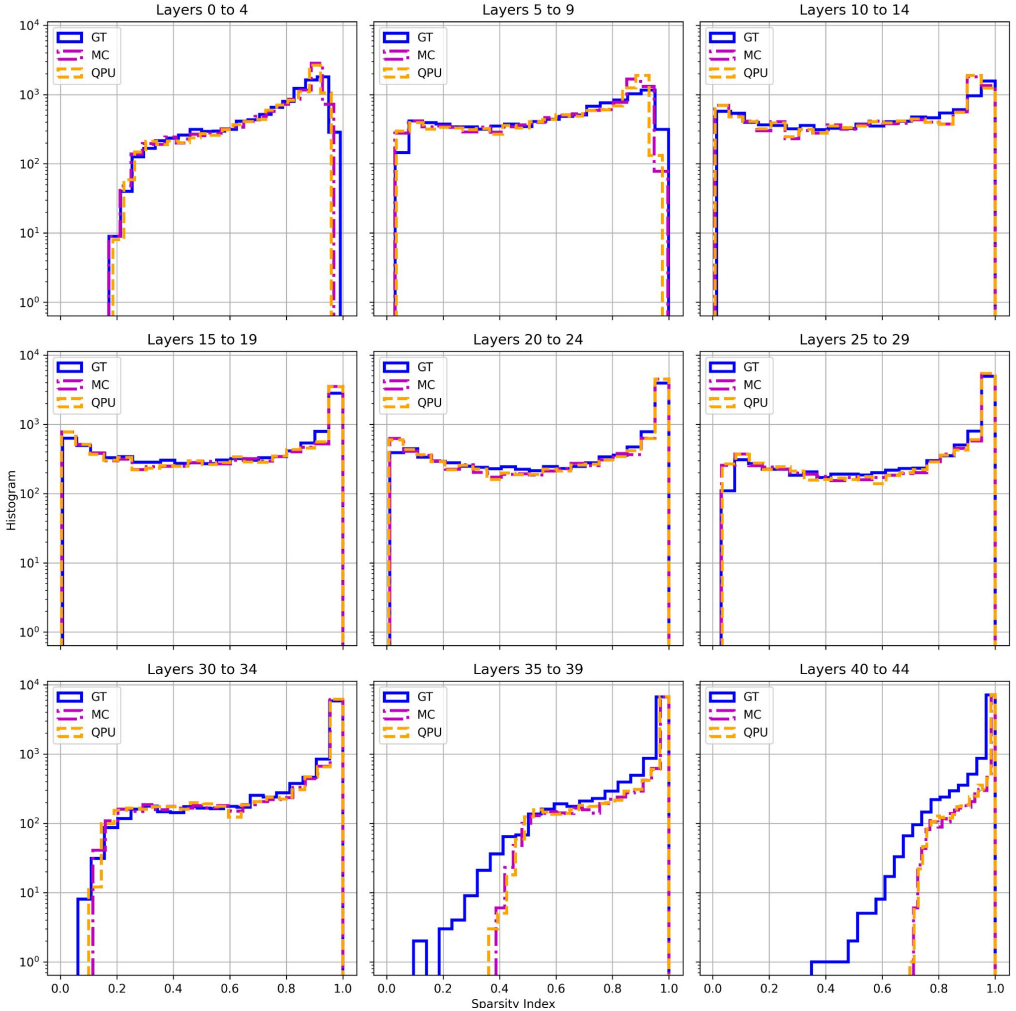
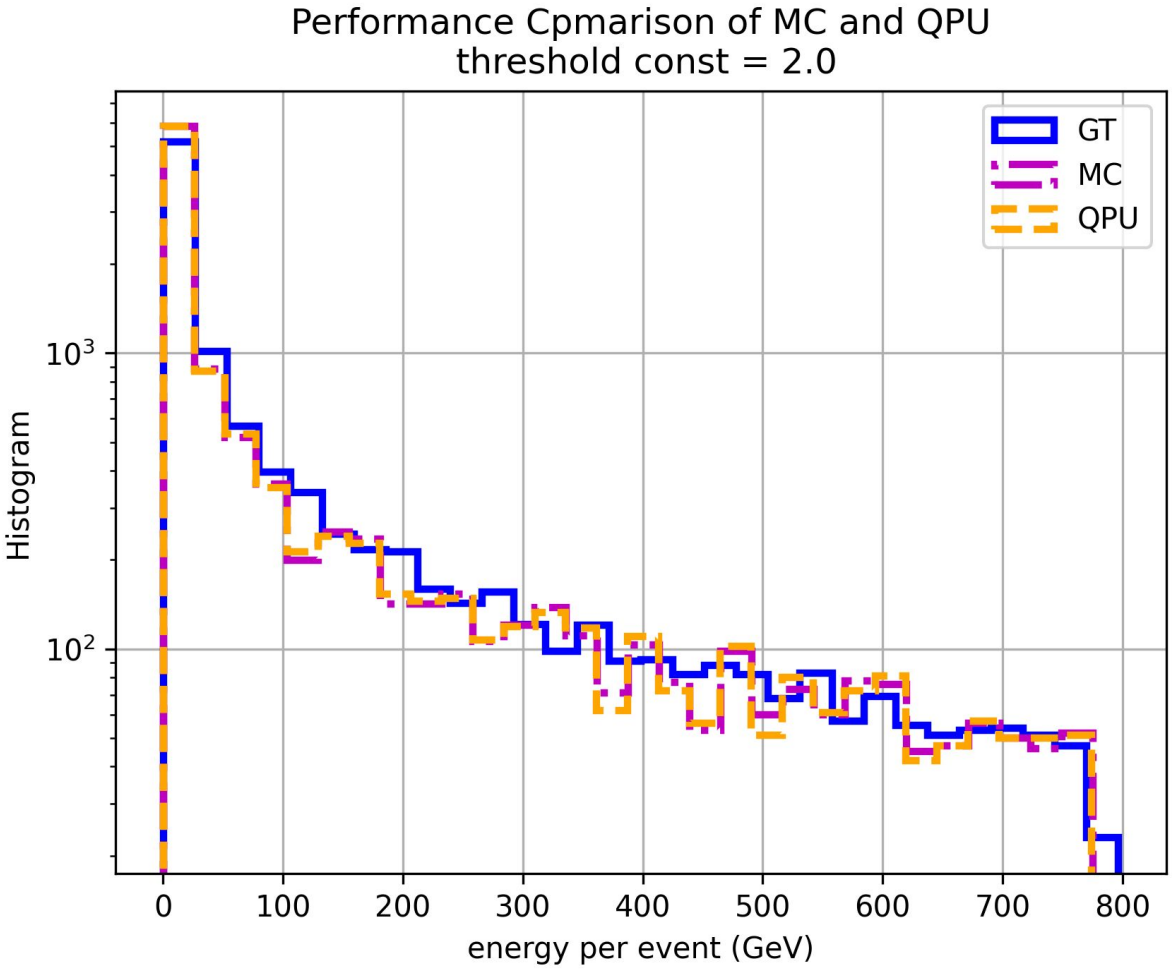
Beta Training Analysis

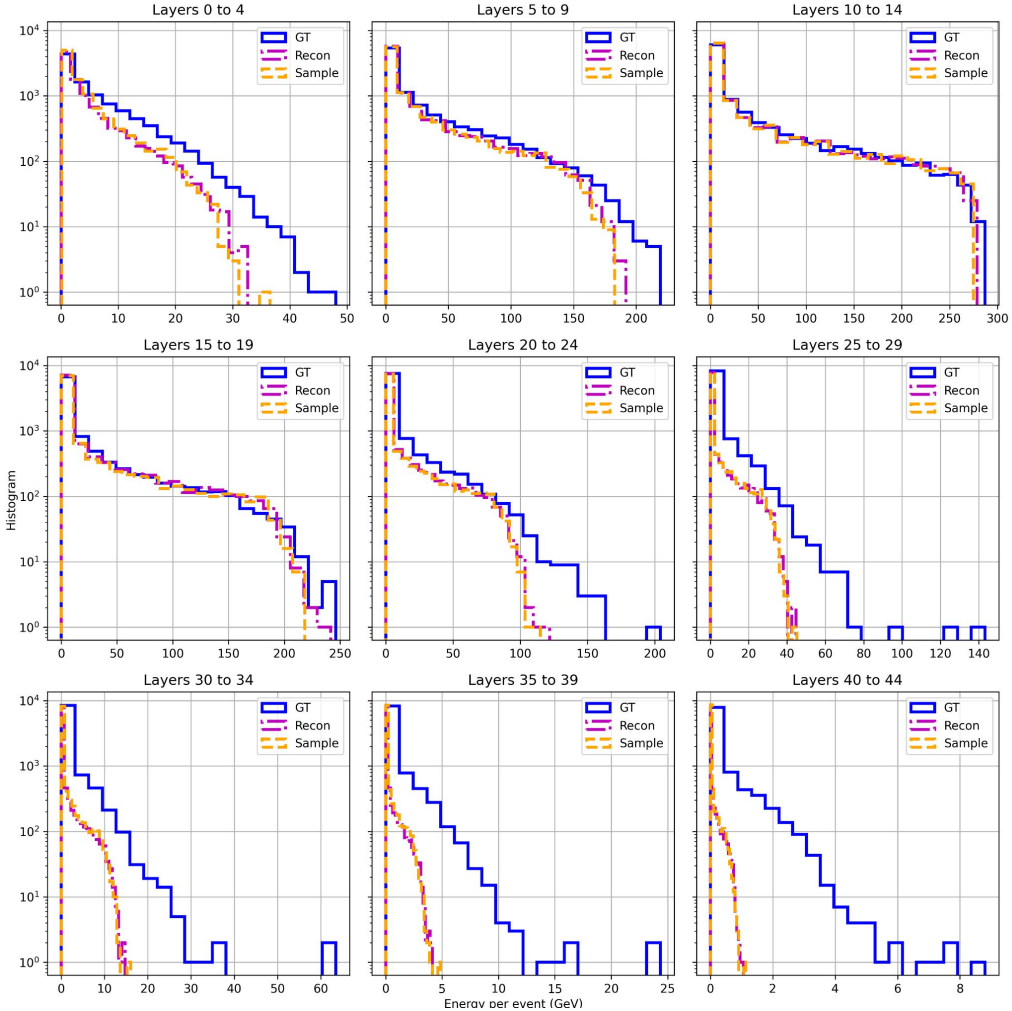
Hao Jia


Beta Training Efficiency Comparison (10 iterations)
Trained Model (4 x 512)
Linear
Adaptive_Scale
Training Epoches
0
2
4
6
8
10
12
14
1
1.5
2
Beta Training Threshold Coefficient


Performance Cpmarison of MC and QPU
threshold const = 2.0
GT
MC
QPU
10^3
10^2
Histogram
0.2
0.4
0.6
0.8
1.0
Sparsity Index


Layers 0 to 4
Layers 5 to 9
Layers 10 to 14
Layers 15 to 19
Layers 20 to 24
Layers 25 to 29
Layers 30 to 34
Layers 35 to 39
Layers 40 to 44
GT
MC
QPU
Histogram
Sparsity Index


Performance Cpmarison of MC and QPU
threshold const = 2.0
GT
MC
QPU
Histogram
10^3
10^2
0
100
200
300
400
500
600
700
800
energy per event (GeV)

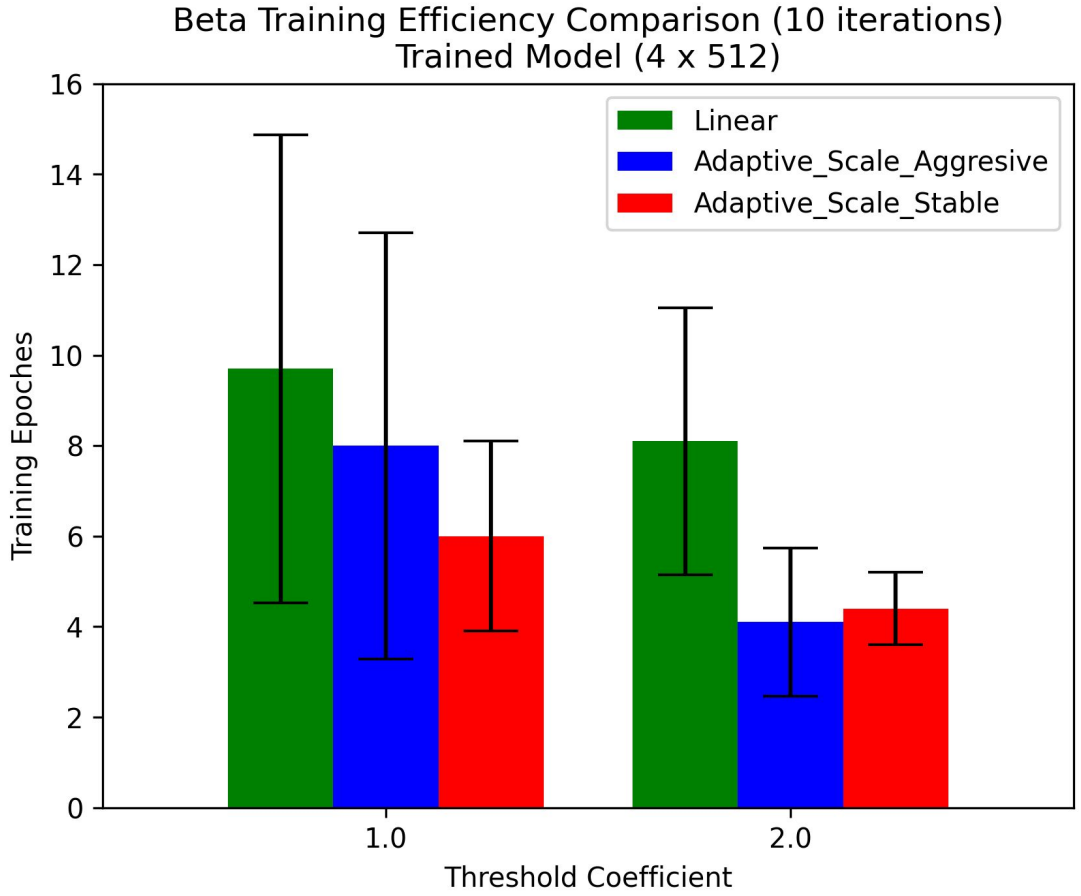

Layers 0 to 4
Layers 5 to 9
Layers 10 to 14
Layers 15 to 19
Layers 20 to 24
Layers 25 to 29
Layers 30 to 34
Layers 35 to 39
Layers 40 to 44
GT
Recon
Sample
Histogram
Energy per event (GeV)

For adaptive scale method:

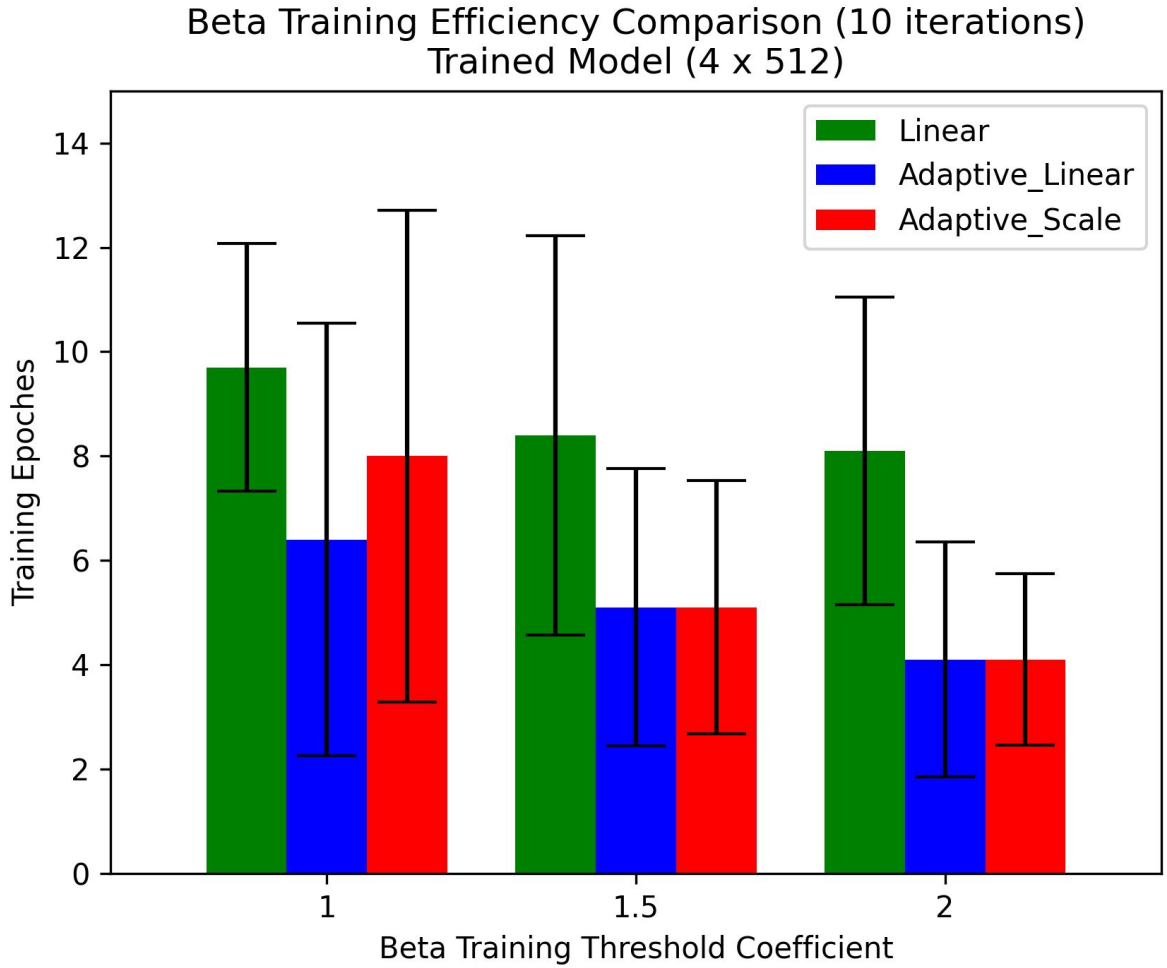
$$|f'_\delta(\beta_{QA})| = \begin{cases} |1 + \frac{\sigma^2_{QA}}{\langle H \rangle_{B(1)}}|, & \delta = 1 \\ |1 + \delta \frac{\sigma^2_{QA}}{\langle H \rangle_{QA}}|, & \delta \neq 1 \end{cases}$$

$$|f'_\delta(\beta_{QA})| < 1$$

When f' = -1: Max training power
When f' = 0 : Most stable



backup


Beta Training Efficiency Comparison (10 iterations)
Trained Model (4 x 512)
Linear
Adaptive_Linear
Adaptive_Scale
Training Epoches
0
2
4
6
8
10
12
14
1
1.5
2
Beta Training Threshold Coefficient