

Eigenvalues $\sim \lambda$

$(B, 9)$

MLPB 1

Linear \rightarrow LayerNorm \rightarrow ReLU \rightarrow Dropout
#Params: 660

$(B, 55)$

MLPB 2

Linear \rightarrow LayerNorm \rightarrow ReLU \rightarrow Dropout
#Params: 1.2K

$(B, 20)$

MLPB 3

Linear \rightarrow LayerNorm \rightarrow ReLU \rightarrow Dropout
#Params: 552

$(B, 24)$

Linear

$(B, 24)$



$(B, 24)$

Linear

#Params: 150

$(B, 6)$

Logits