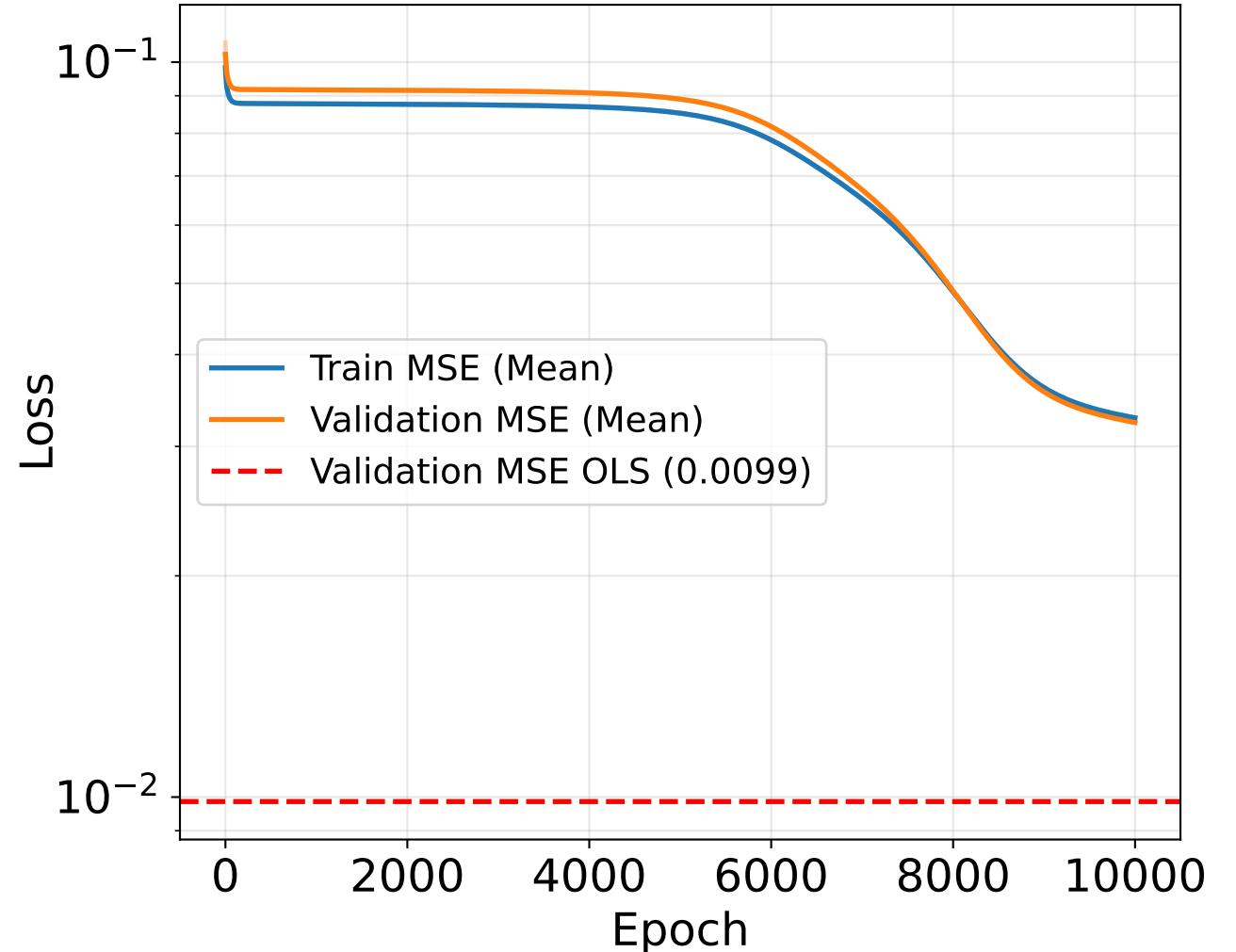


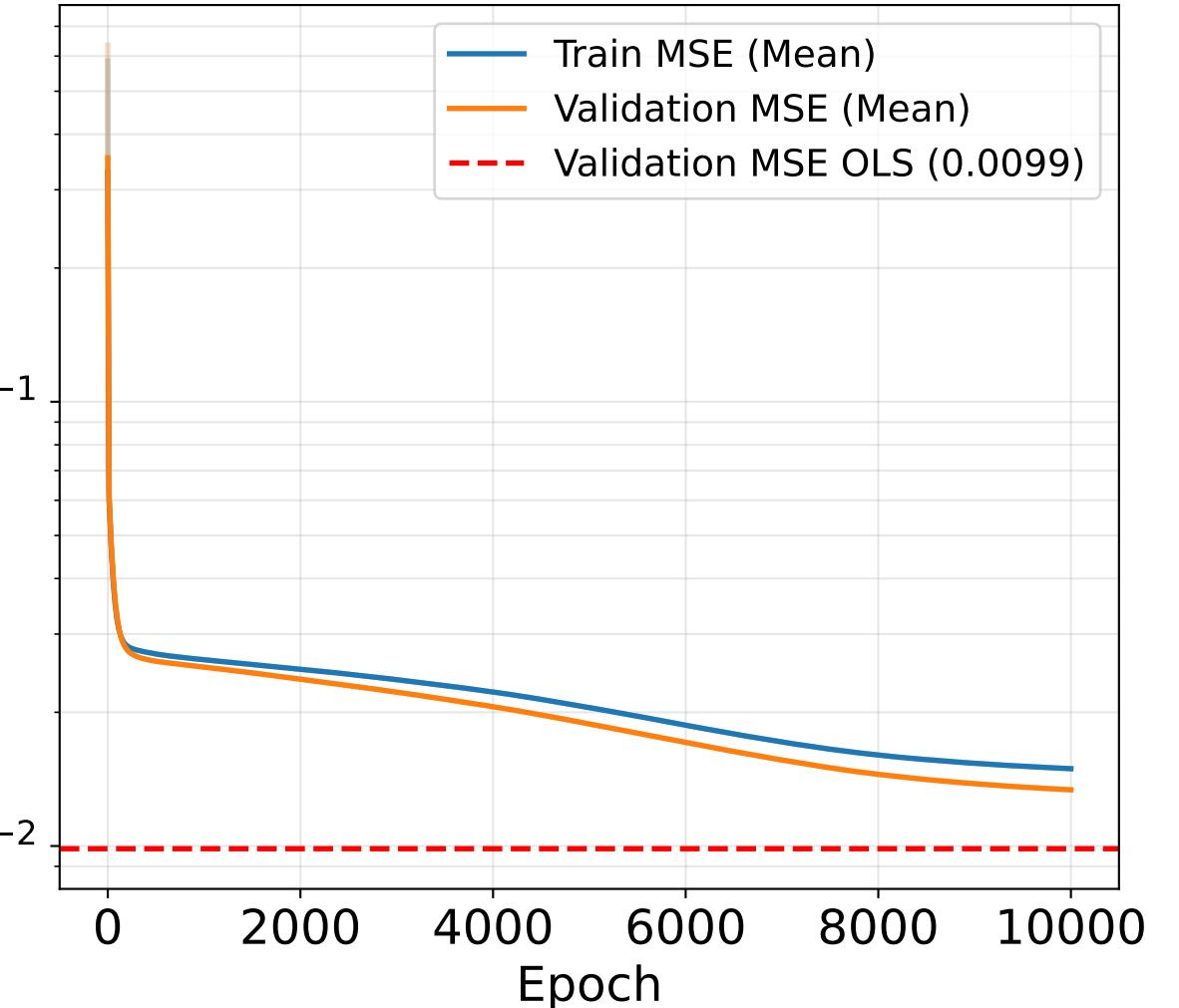
# Learning Curves with GD ( $N=300$ , $\sigma=0.1$ ) - Averaged Over 3 Runs

## 1 Hidden Layer & 50 Nodes

Sigmoid:  $\eta=0.0464$



ReLU:  $\eta=0.0215$



LeakyReLU:  $\eta=0.0215$

