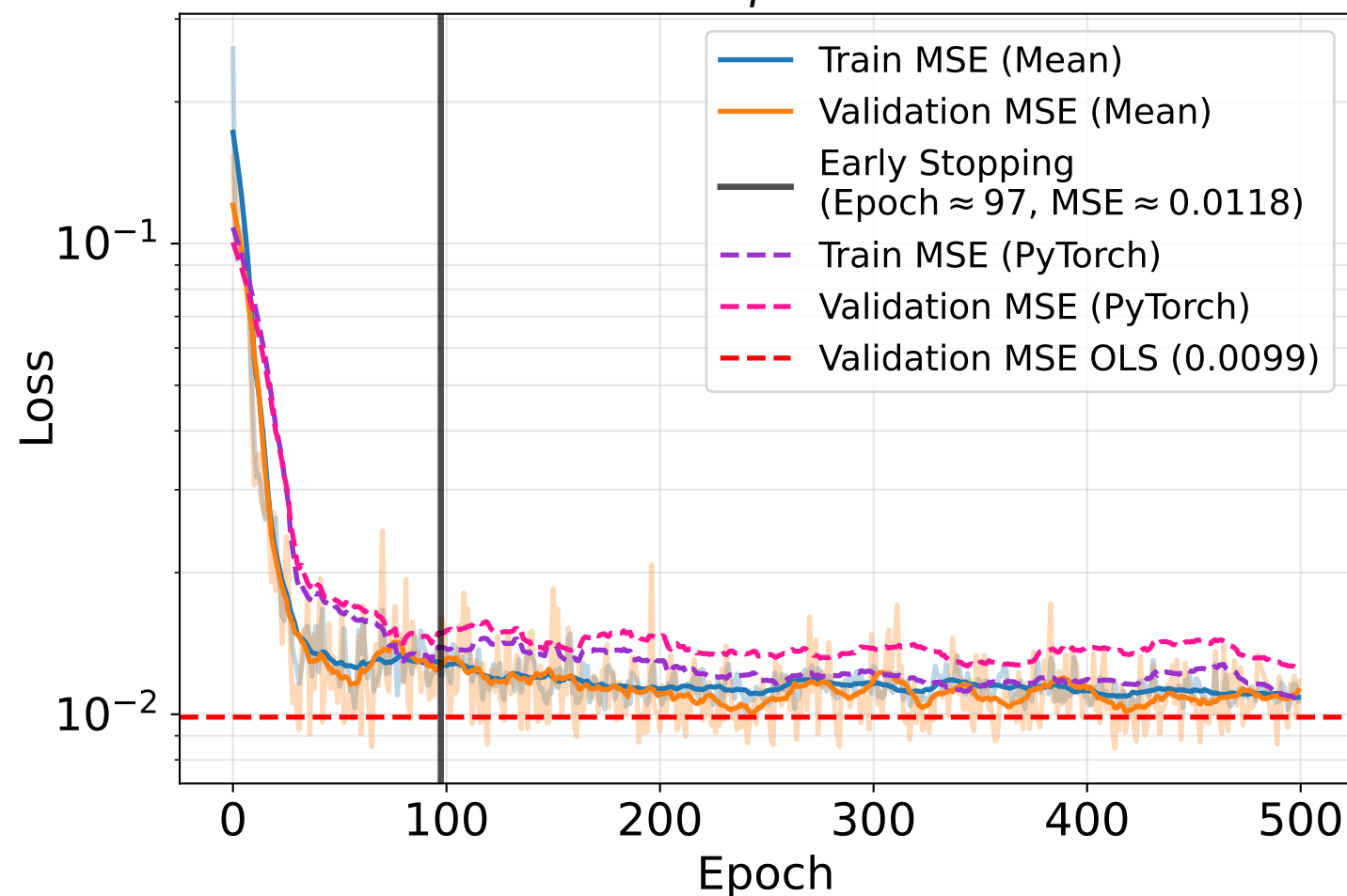


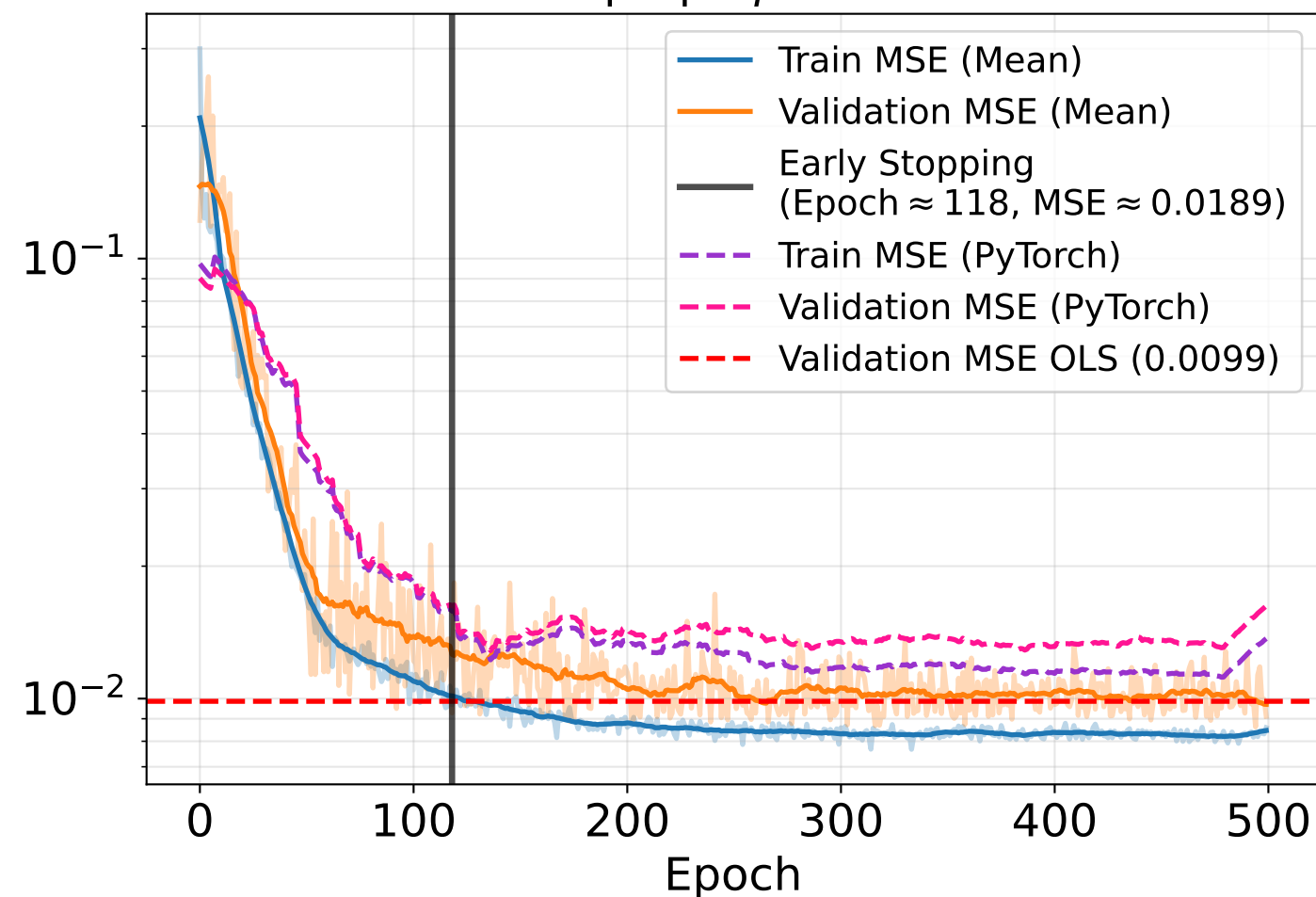
Learning Curves for N=300, $\sigma=0.1$ (Averaged Over 3 Runs)

1 Hidden Layer & 50 Nodes

Adam: $\eta=0.0464$

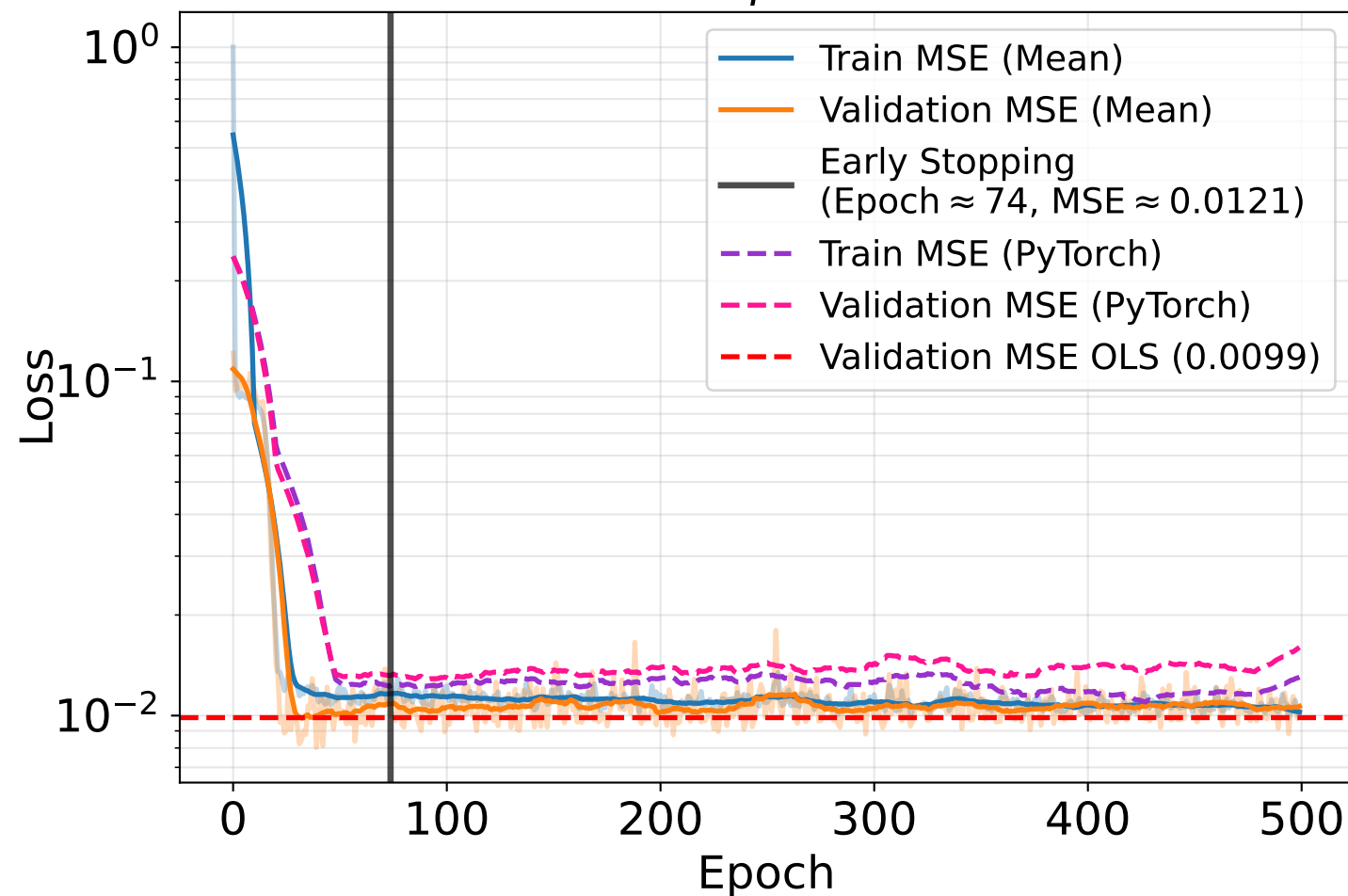


RMSprop: $\eta=0.0215$



2 Hidden Layers & 100 Nodes Each

Adam: $\eta=0.0215$



RMSprop: $\eta=0.0046$

