

Training loss for $\eta = 0.25$

