

$(\hat{Q}_h - Q^{\pi_{\text{target}}})^2$  MSE (1024 samples,  $\gamma = 0.99$ )

