

$L^2$  error norm with  $\eta = 0.5$  and 3 hidden layers

