

Two extra layers, 40000 training samples,  $L_1 = 0.0001$

