Multilayer MNIST CNN [Time 232.930s, GPUs 2] (Ir = 0.000100, batch size = 5000) Loss Accuracy train 1.0  $10^{0}$ val 0.9 - $10^{-1}$ 0.8 0.7 -10-2 train 0.6 val 25 50 100 125 25 50 0 75 150 0 75 100 125 150