Multilayer MNIST CNN [Time 72.439s, GPUs 2] (Ir = 0.000100, batch size = 50) Loss Accuracy 10^{-1} train 1.000 train val val 0.995 -0.990 -10-2 0.985 -0.980 -0.975 - 10^{-3} 6 8 10 12 14 6 8 10 12 14