Multilayer MNIST CNN [Time 285.038s, GPUs 1] (Ir = 0.000100, batch size = 5000) Loss Accuracy train 1.0 val 0.9 10^{0} 8.0 0.7 - 10^{-1} 0.6 -0.5 10^{-2} 0.4 train 0.3 val

100

150

200

250

300

50

50

0

100

200

250

300

0

150