Multilayer MNIST CNN [Time 765.896s, GPUs 1] (Ir = 0.000100, batch size = 50) Loss Accuracy 1.00 10^{-1} 10^{-2} 0.99 10⁻³ 10^{-4} 0.98 10⁻⁵ 10^{-6} 0.97 - 10^{-7} train train 0.96 -10-8 val val

50

100

200

150

250

300

50

0

100

150

200

250

300