Layer-wise gradients in ResNet50 with dropout epoch:5 epoch:10 10-2 epoch:15 epoch:20 epoch:25 epoch:30 epoch:35 epoch:40 epoch:45 epoch:50 Average gradient 10^{-4} -30 10 40 20 50 0 Convolutional layer