Gradients in a ResNet18 with training epoch Layer1 Layer3 Layer6 Layer9 Layer13 - Layer16 10<sup>-3</sup> Mean Gradient 10-4 10-5 -25 75 50 125 100 150 175 200 0 Epoch