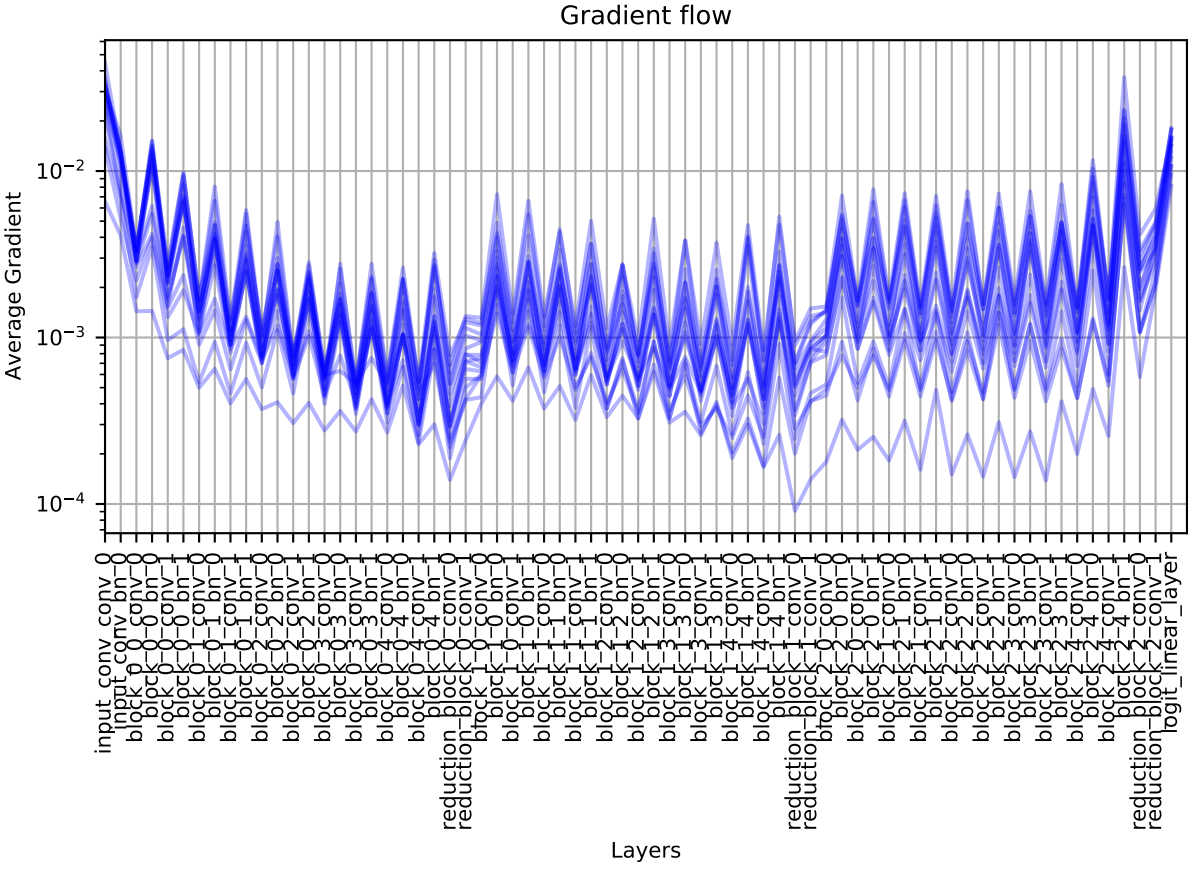
Gradient flow
Average Gradient
$10^{-2}$
$10^{-3}$
$10^{-4}$
Layers
input_conv_0
input_bn_0
block_0_0_conv_0
block_0_0_bn_0
block_0_0_conv_1
block_0_0_bn_1
block_0_1_conv_0
block_0_1_bn_0
block_0_1_conv_1
block_0_1_bn_1
block_0_2_conv_0
block_0_2_bn_0
block_0_2_conv_1
block_0_2_bn_1
block_0_3_conv_0
block_0_3_bn_0
block_0_3_conv_1
block_0_3_bn_1
block_0_4_conv_0
block_0_4_bn_0
block_0_4_conv_1
block_0_4_bn_1
reduction_block_0_conv_0
reduction_block_0_conv_1
block_1_0_conv_0
block_1_0_bn_0
block_1_0_conv_1
block_1_0_bn_1
block_1_1_conv_0
block_1_1_bn_0
block_1_1_conv_1
block_1_1_bn_1
block_1_2_conv_0
block_1_2_bn_0
block_1_2_conv_1
block_1_2_bn_1
block_1_3_conv_0
block_1_3_bn_0
block_1_3_conv_1
block_1_3_bn_1
block_1_4_conv_0
block_1_4_bn_0
block_1_4_conv_1
block_1_4_bn_1
reduction_block_1_conv_0
reduction_block_1_conv_1
block_2_0_conv_0
block_2_0_bn_0
block_2_0_conv_1
block_2_0_bn_1
block_2_1_conv_0
block_2_1_bn_0
block_2_1_conv_1
block_2_1_bn_1
block_2_2_conv_0
block_2_2_bn_0
block_2_2_conv_1
block_2_2_bn_1
block_2_3_conv_0
block_2_3_bn_0
block_2_3_conv_1
block_2_3_bn_1
block_2_4_conv_0
block_2_4_bn_0
block_2_4_conv_1
block_2_4_bn_1
reduction_block_2_conv_0
reduction_block_2_conv_1
logit_linear_layer