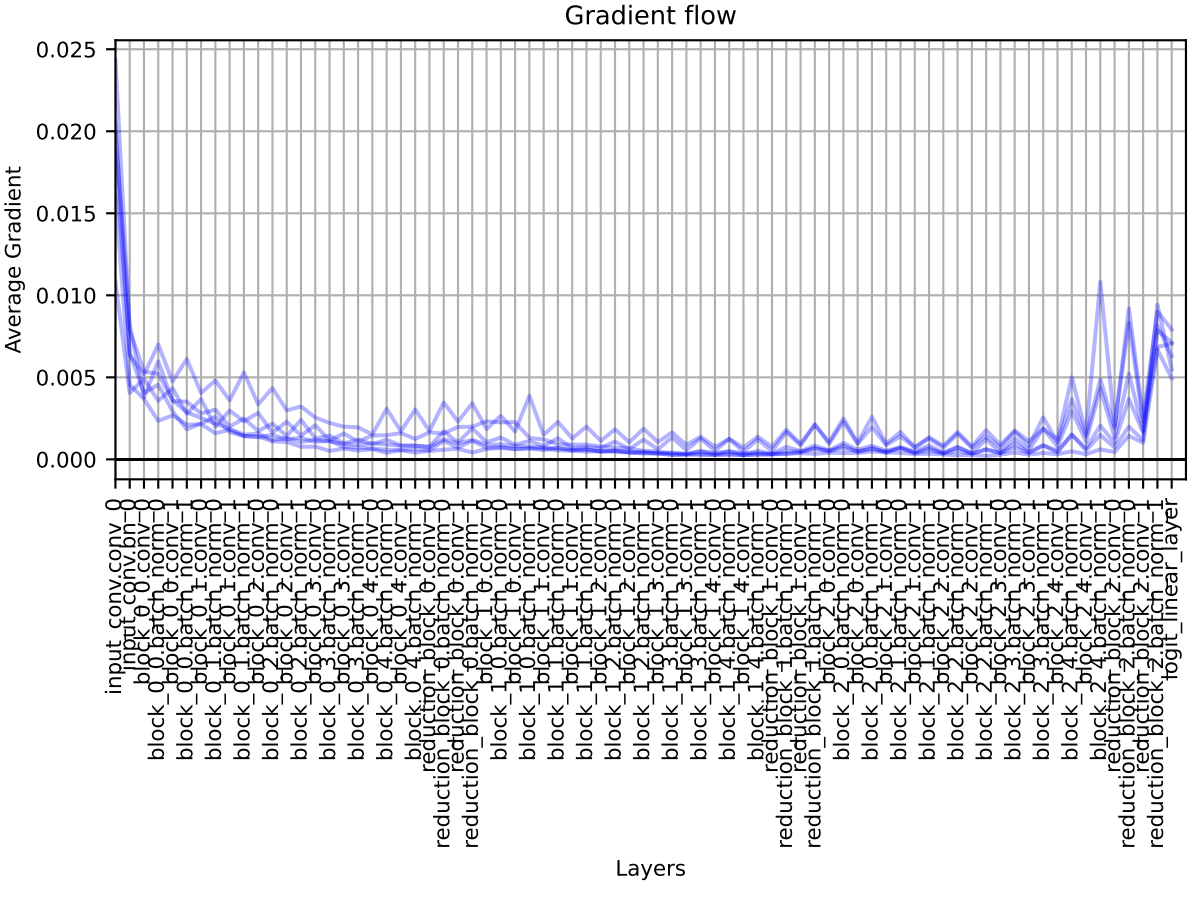
Gradient flow
Average Gradient
0.025
0.020
0.015
0.010
0.005
0.000
Layers
input_conv.conv_0
input_conv.bn_0
block_0_0.conv_0
block_0_0.batch_norm_0
block_0_0.conv_1
block_0_0.batch_norm_1
block_0_1.conv_0
block_0_1.batch_norm_0
block_0_2.conv_0
block_0_2.batch_norm_0
block_0_2.conv_1
block_0_2.batch_norm_1
block_0_3.conv_0
block_0_3.batch_norm_0
block_0_3.conv_1
block_0_3.batch_norm_1
block_0_4.conv_0
block_0_4.batch_norm_0
block_0_4.conv_1
block_0_4.batch_norm_1
reduction_block_0.conv_0
reduction_block_0.batch_norm_0
reduction_block_0.conv_1
reduction_block_0.batch_norm_1
block_1_0.conv_0
block_1_0.batch_norm_0
block_1_0.conv_1
block_1_0.batch_norm_1
block_1_1.conv_0
block_1_1.batch_norm_0
block_1_1.conv_1
block_1_1.batch_norm_1
block_1_2.conv_0
block_1_2.batch_norm_0
block_1_2.conv_1
block_1_2.batch_norm_1
block_1_3.conv_0
block_1_3.batch_norm_0
block_1_3.conv_1
block_1_3.batch_norm_1
block_1_4.conv_0
block_1_4.batch_norm_0
block_1_4.conv_1
block_1_4.batch_norm_1
reduction_block_1.conv_0
reduction_block_1.batch_norm_0
reduction_block_1.conv_1
reduction_block_1.batch_norm_1
block_2_0.conv_0
block_2_0.batch_norm_0
block_2_0.conv_1
block_2_0.batch_norm_1
block_2_1.conv_0
block_2_1.batch_norm_0
block_2_1.conv_1
block_2_1.batch_norm_1
block_2_2.conv_0
block_2_2.batch_norm_0
block_2_2.conv_1
block_2_2.batch_norm_1
block_2_3.conv_0
block_2_3.batch_norm_0
block_2_3.conv_1
block_2_3.batch_norm_1
block_2_4.conv_0
block_2_4.batch_norm_0
block_2_4.conv_1
block_2_4.batch_norm_1
reduction_block_2.conv_0
reduction_block_2.batch_norm_0
reduction_block_2.conv_1
reduction_block_2.batch_norm_1
logit_linear_layer