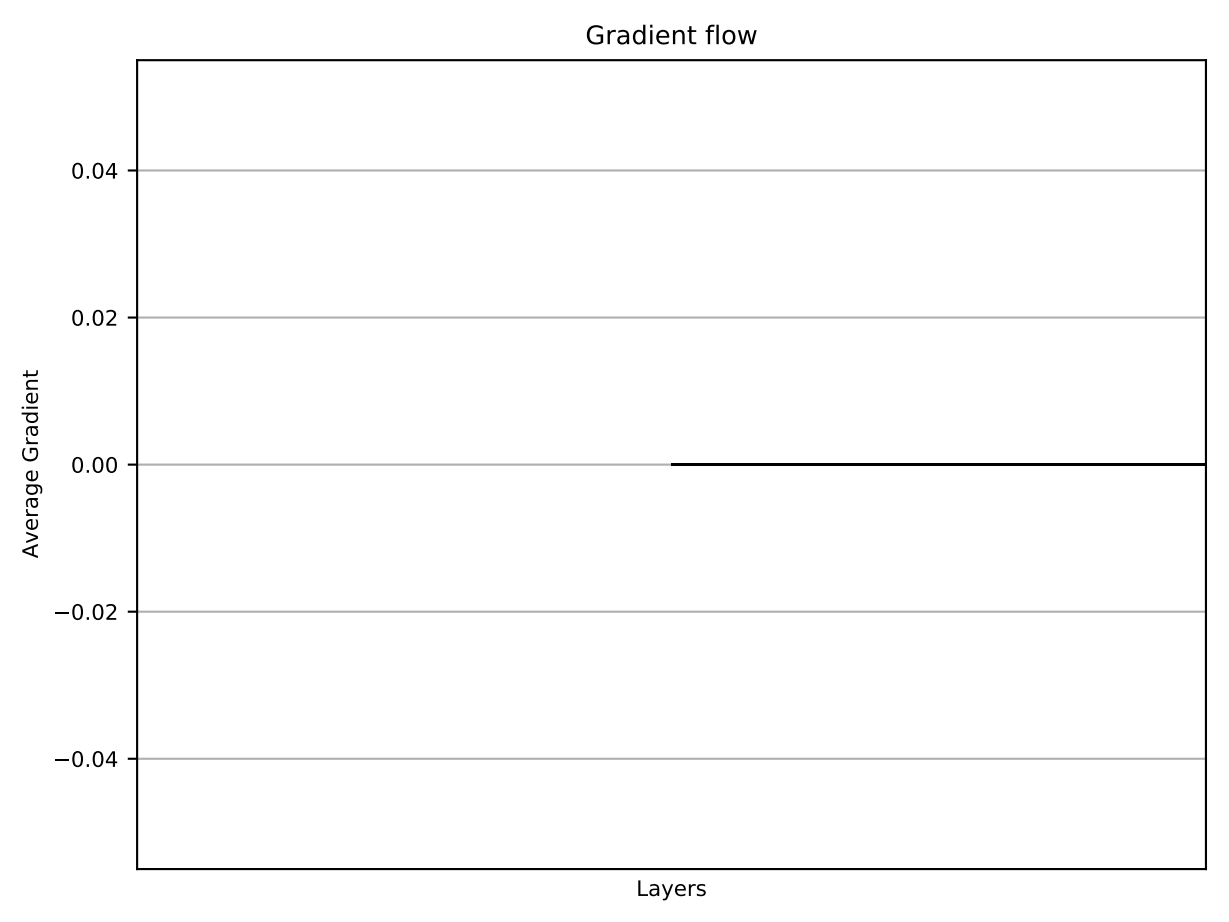
Gradient flow
0.04
0.02
0.00
−0.02
−0.04
Average Gradient
Layers