

Gradient norm

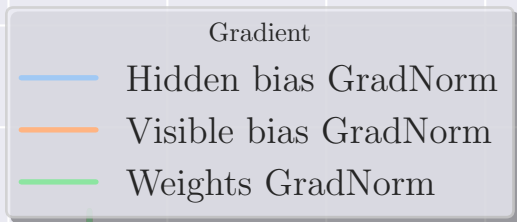
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000

0

100

200

Epoch



Hidden bias GradNorm

Visible bias GradNorm

Weights GradNorm

300

400

500

600