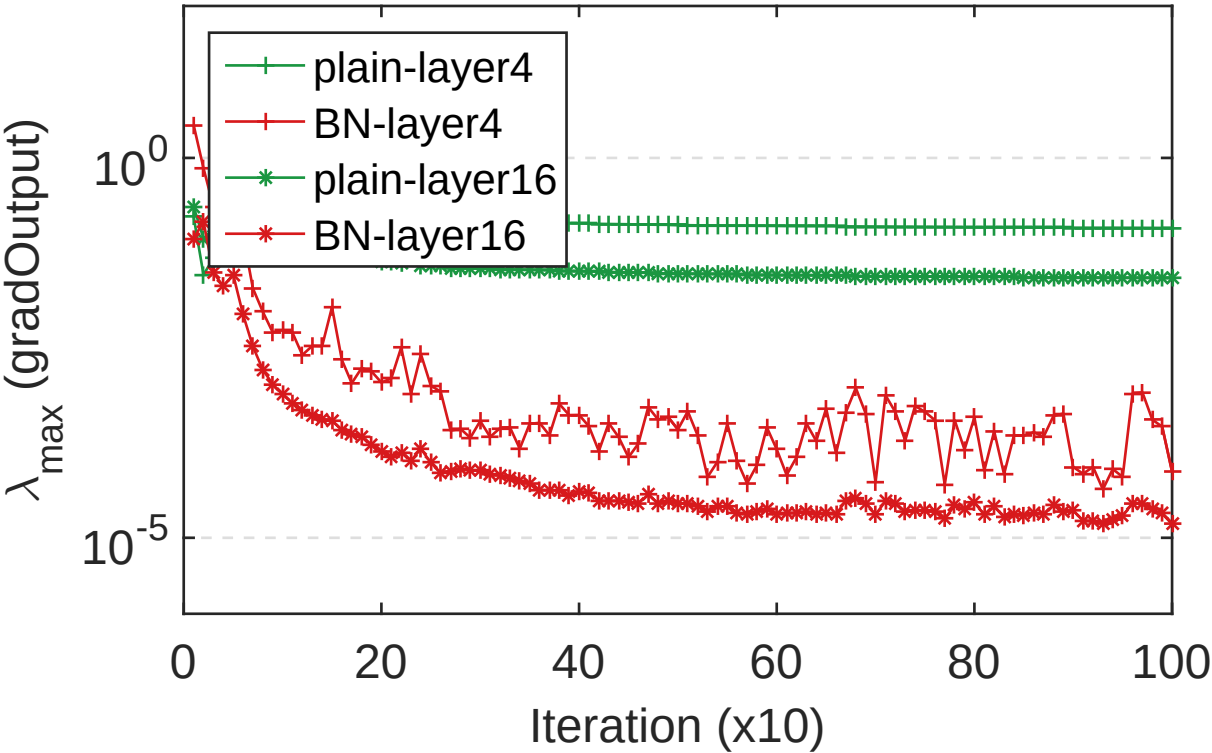
plain-layer4
BN-layer4
plain-layer16
BN-layer16
10^0
10^{-5}
λ_{max} (gradOutput)
0
20
40
60
80
100
Iteration (x10)