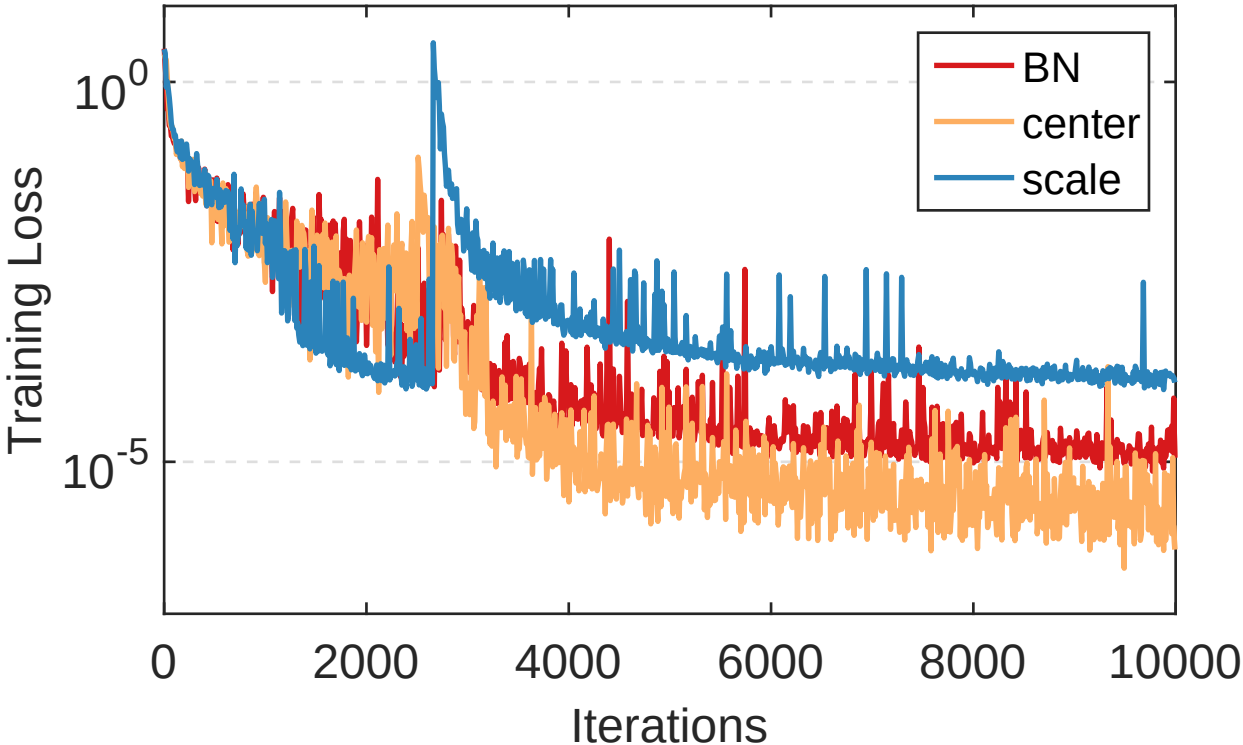
BN
center
scale
Training Loss
10^{0}
10^{-5}
0
2000
4000
6000
8000
10000
Iterations