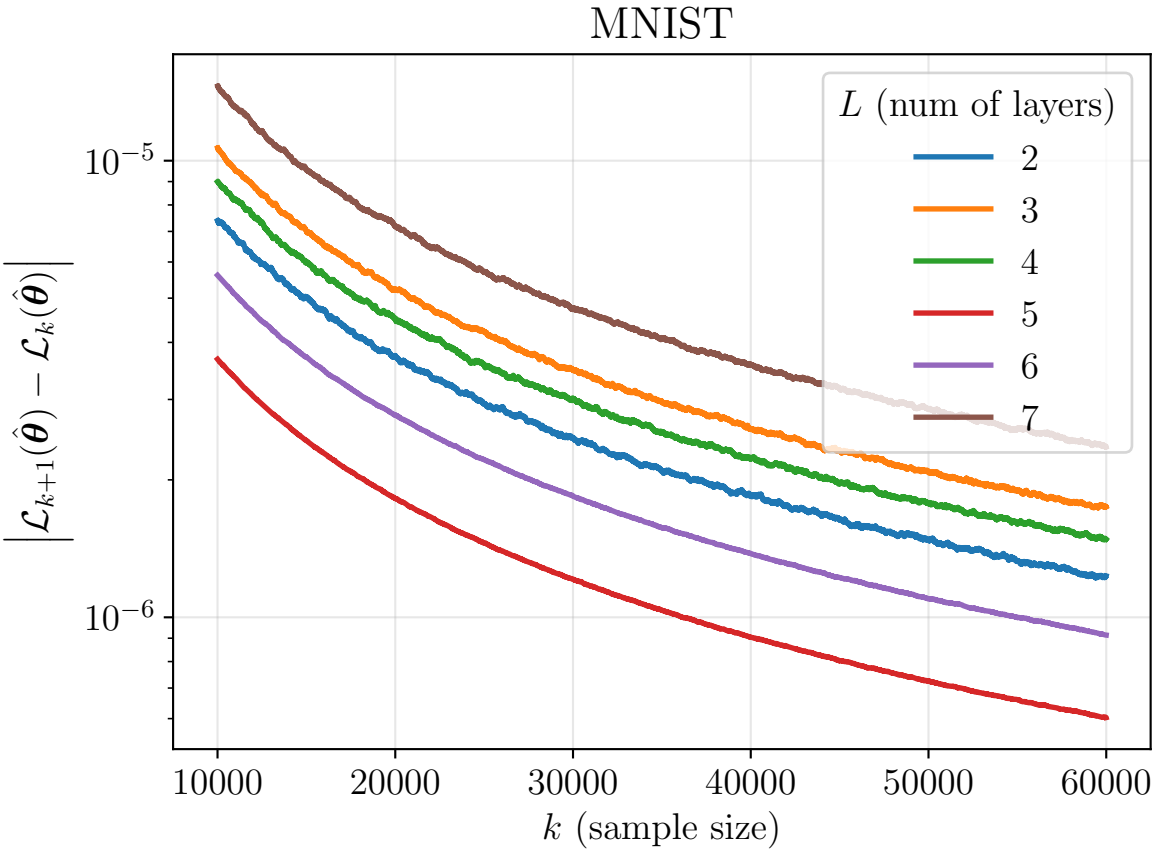
MNIST
L (num of layers)
2
3
4
5
6
7
10^{-5}
10^{-6}
$\left|\mathcal{L}_{k+1}(\hat{\boldsymbol{\theta}}) - \mathcal{L}_k(\hat{\boldsymbol{\theta}})\right|$
10000
20000
30000
40000
50000
60000
k (sample size)