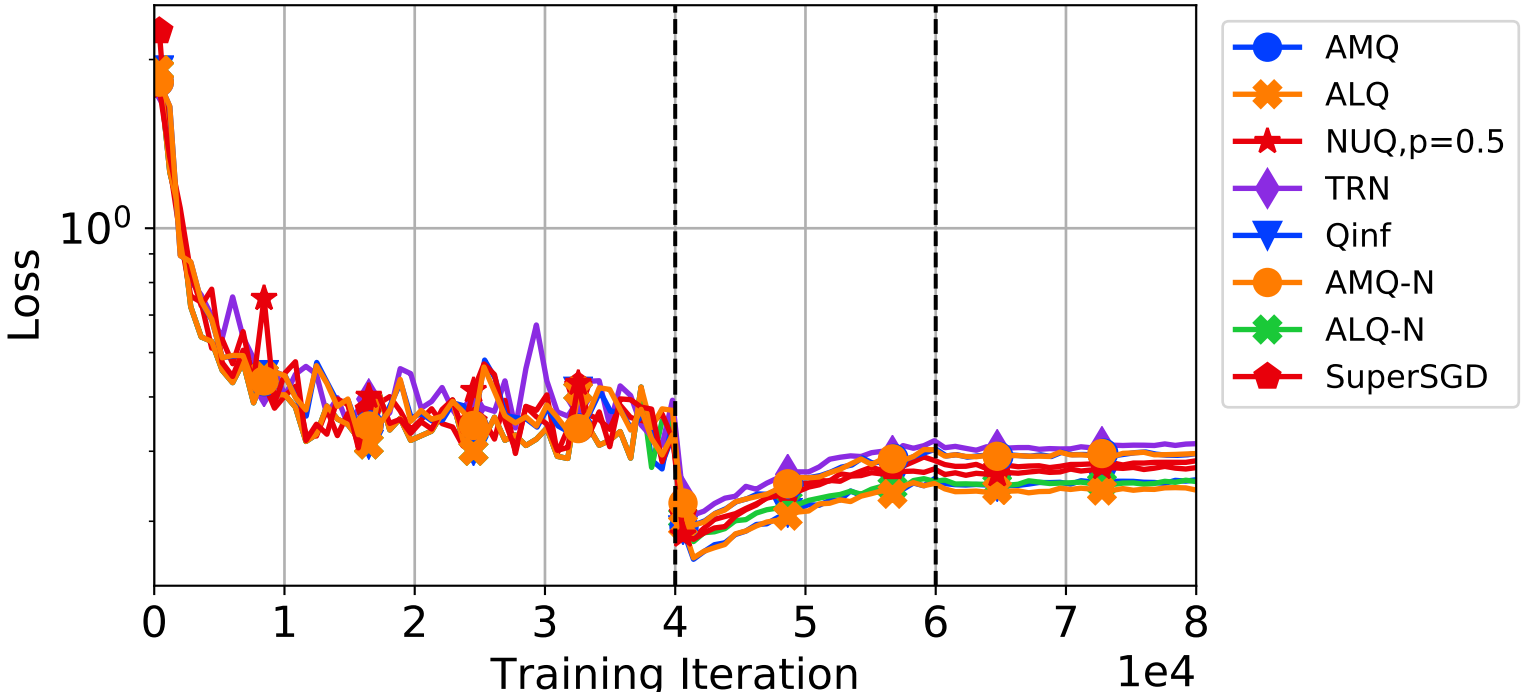
Loss
10^0
0
1
2
3
4
5
6
7
8
Training Iteration
1e4
AMQ
ALQ
NUQ,p=0.5
TRN
Qinf
AMQ-N
ALQ-N
SuperSGD