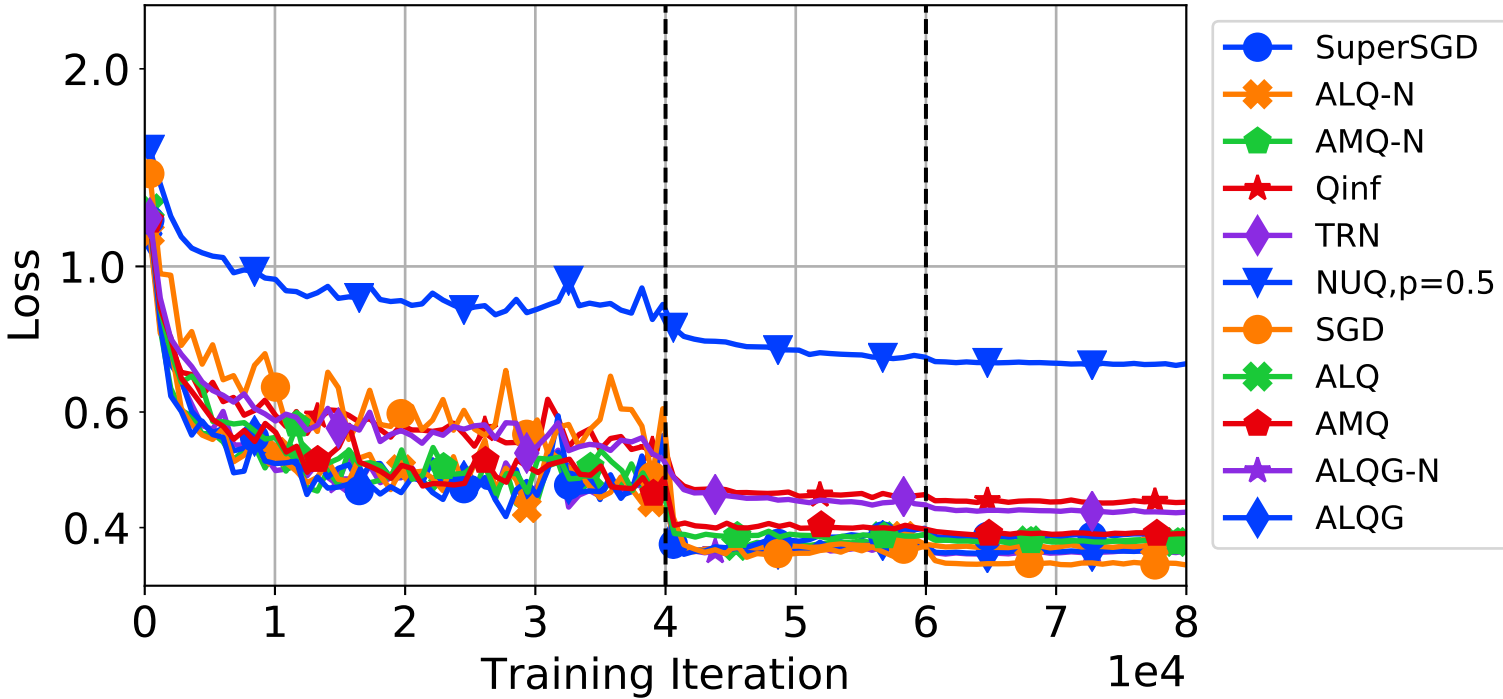
Loss
2.0
1.0
0.6
0.4
0
1
2
3
4
5
6
7
8
Training Iteration
1e4
SuperSGD
ALQ-N
AMQ-N
Qinf
TRN
NUQ,p=0.5
SGD
ALQ
AMQ
ALQG-N
ALQG