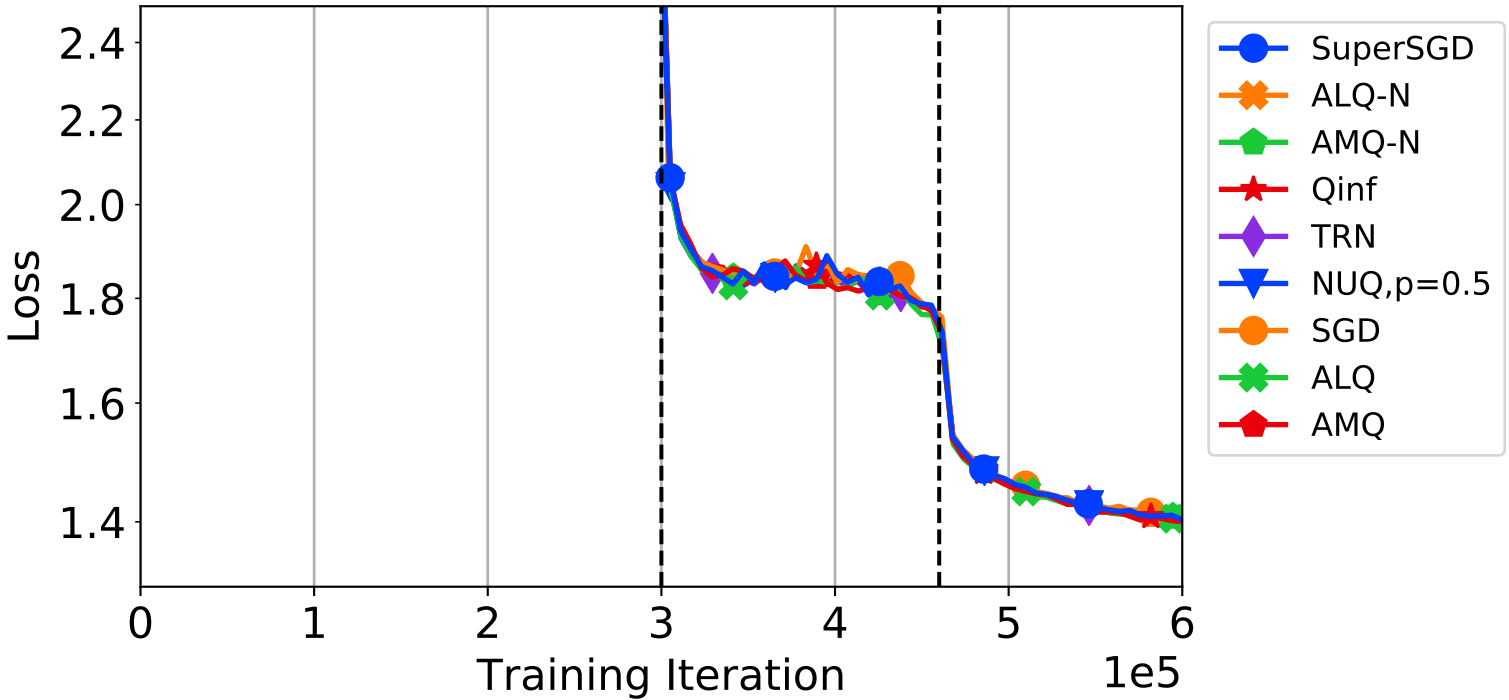
Loss
2.4
2.2
2.0
1.8
1.6
1.4
0
1
2
3
4
5
6
Training Iteration
1e5
SuperSGD
ALQ-N
AMQ-N
Qinf
TRN
NUQ,p=0.5
SGD
ALQ
AMQ