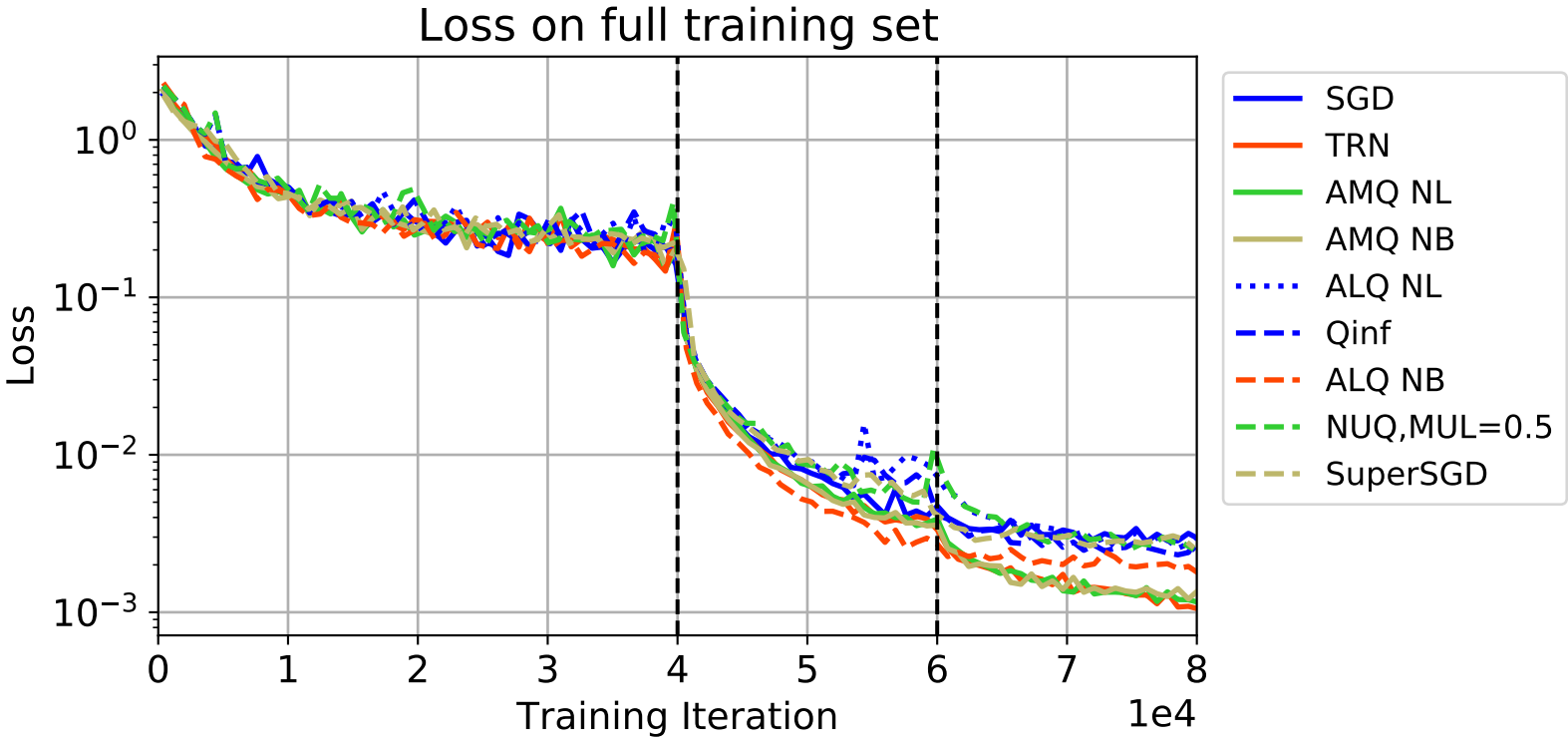
Loss on full training set
Loss
10^0
10^{-1}
10^{-2}
10^{-3}
0
1
2
3
4
5
6
7
8
Training Iteration
1e4
SGD
TRN
AMQ NL
AMQ NB
ALQ NL
Qinf
ALQ NB
NUQ,MUL=0.5
SuperSGD