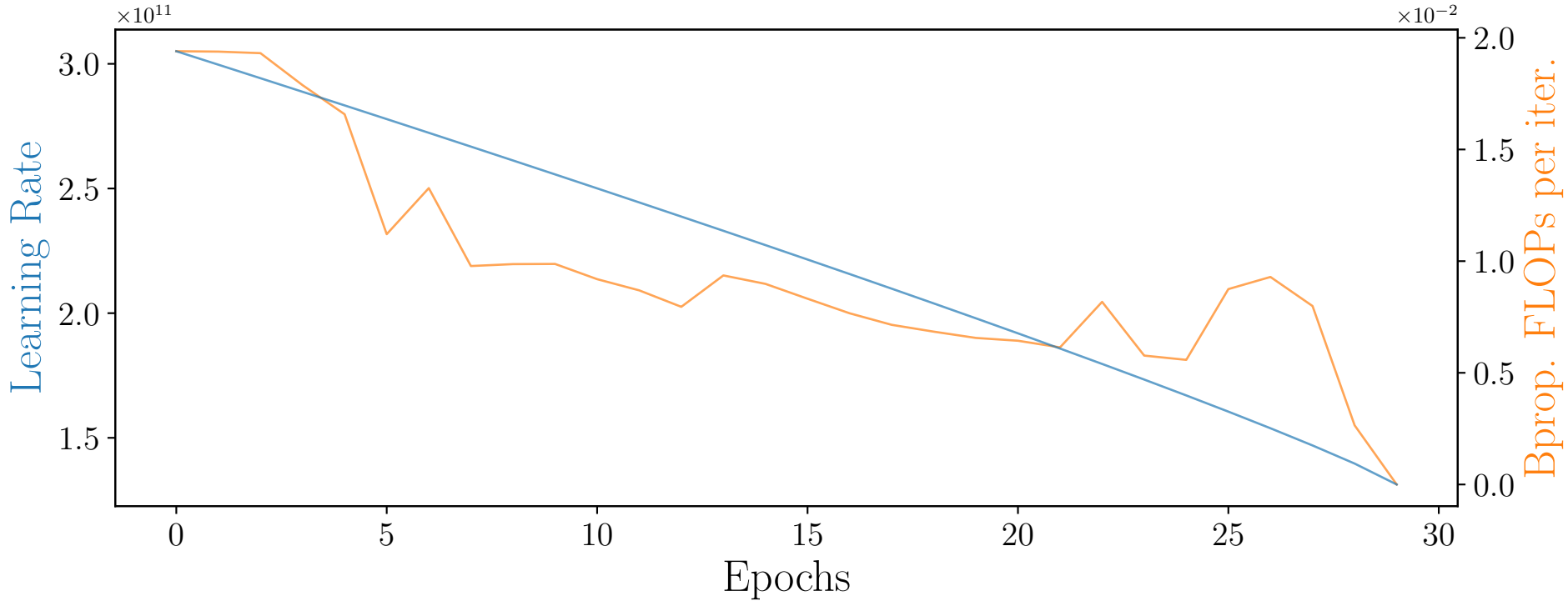
×10^11
×10^−2
3.0
2.5
2.0
1.5
2.0
1.5
1.0
0.5
0.0
0
5
10
15
20
25
30
Epochs
Learning Rate
Bprop. FLOPs per iter.