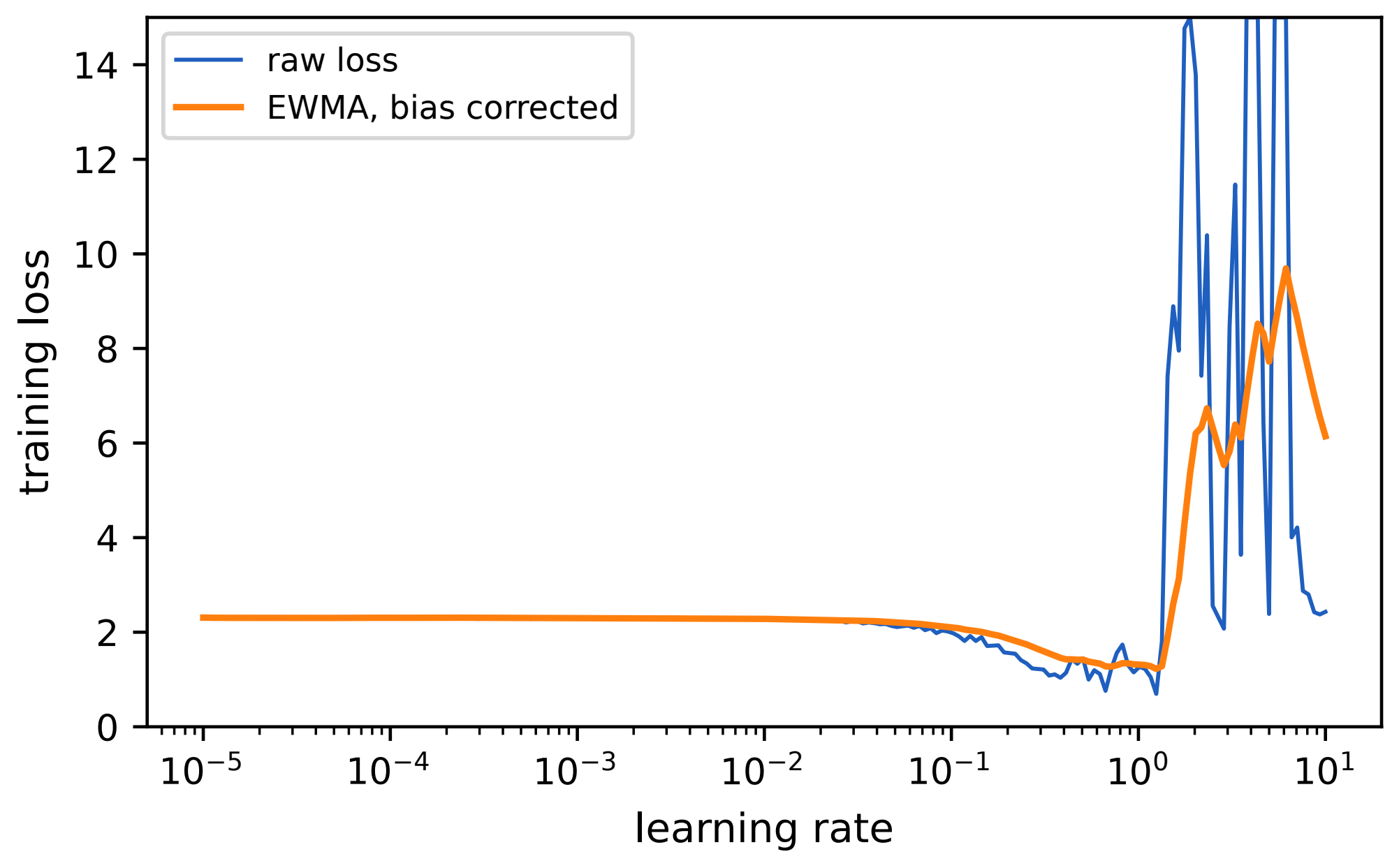

raw loss
EWMA, bias corrected
training loss
learning rate
0
2
4
6
8
10
12
14
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}