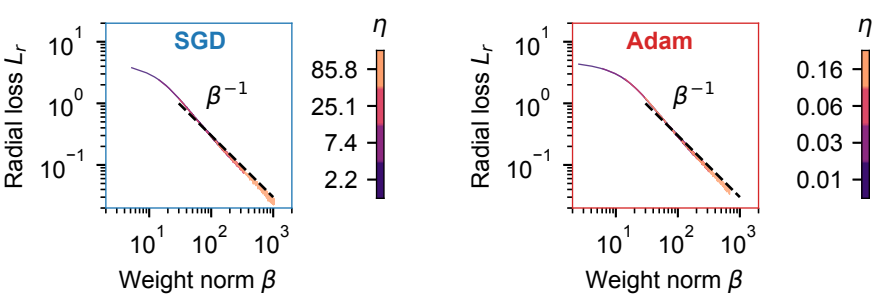

SGD
Radial loss L_r
10^1
10^0
10^{-1}
β^{-1}
10^1
10^2
10^3
Weight norm β
η
85.8
25.1
7.4
2.2
Adam
Radial loss L_r
10^1
10^0
10^{-1}
β^{-1}
10^1
10^2
10^3
Weight norm β
η
0.16
0.06
0.03
0.01