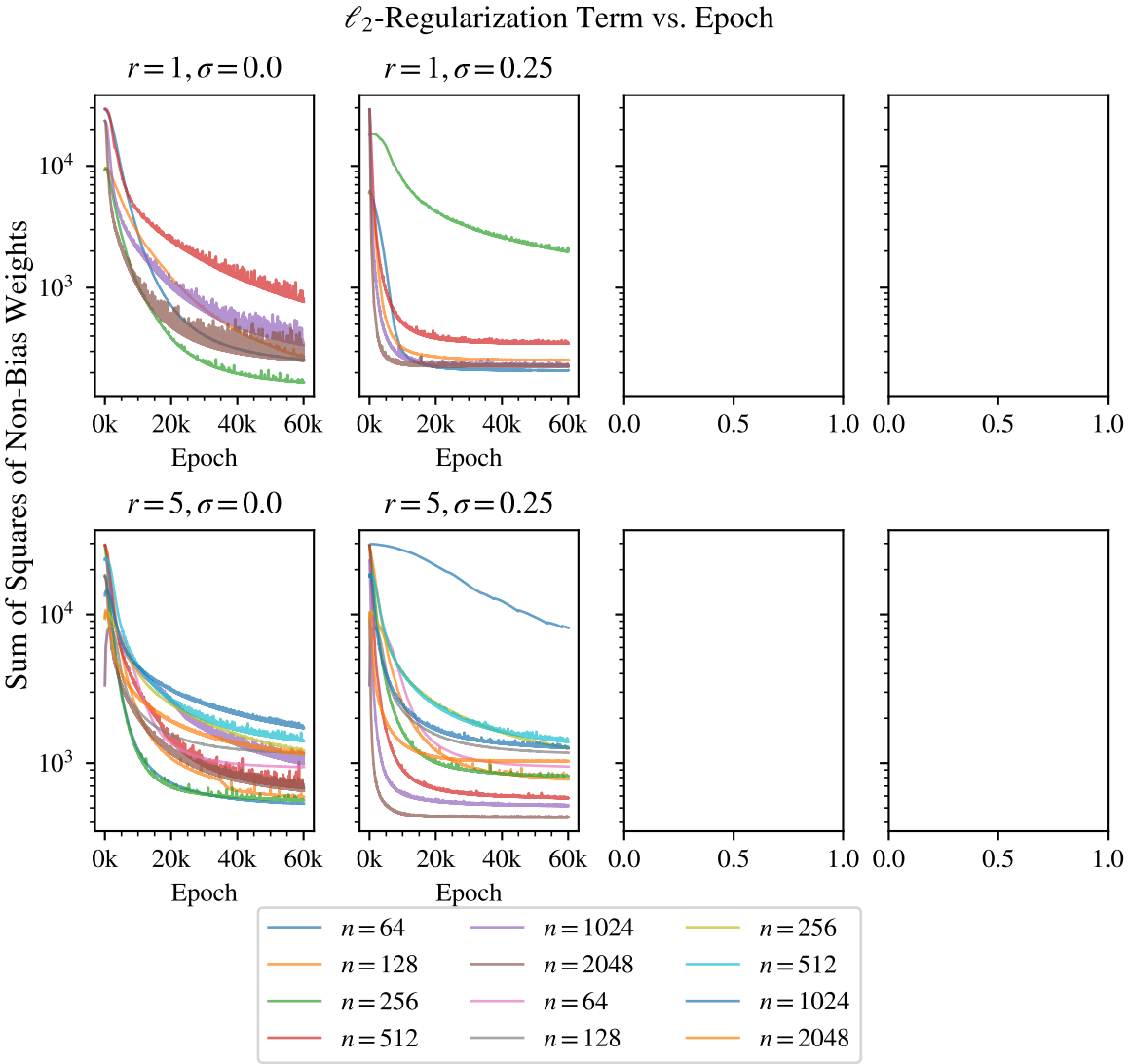
ℓ_2-Regularization Term vs. Epoch
$r=1, \sigma=0.0$
$r=1, \sigma=0.25$
$r=5, \sigma=0.0$
$r=5, \sigma=0.25$
Sum of Squares of Non-Bias Weights
10^4
10^3
0k
20k
40k
60k
Epoch
0.0
0.5
1.0
$n=64$
$n=128$
$n=256$
$n=512$
$n=1024$
$n=2048$
$n=64$
$n=128$
$n=256$
$n=512$
$n=1024$
$n=2048$