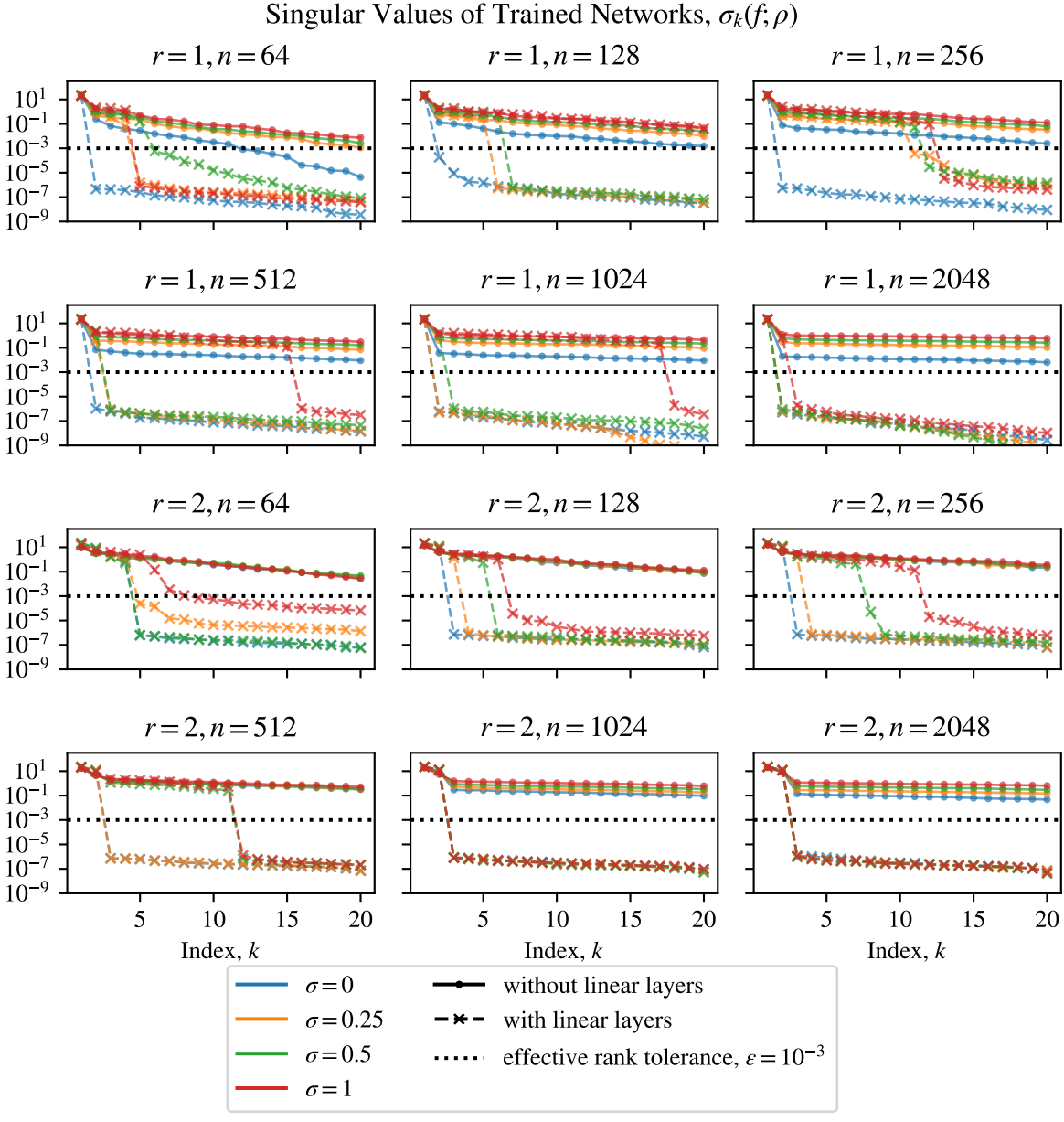
Singular Values of Trained Networks, $\sigma_k(\hat{f};\rho)$
$r=1, n=64$
$r=1, n=128$
$r=1, n=256$
$r=1, n=512$
$r=1, n=1024$
$r=1, n=2048$
$r=2, n=64$
$r=2, n=128$
$r=2, n=256$
$r=2, n=512$
$r=2, n=1024$
$r=2, n=2048$
10^{1}
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
5
10
15
20
Index, k
$\sigma=0$
$\sigma=0.25$
$\sigma=0.5$
$\sigma=1$
without linear layers
with linear layers
effective rank tolerance, $\varepsilon=10^{-3}$