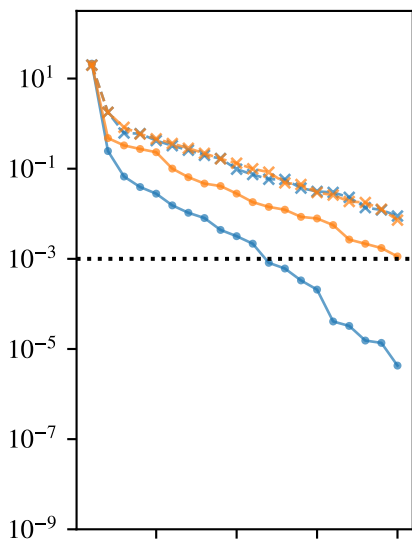
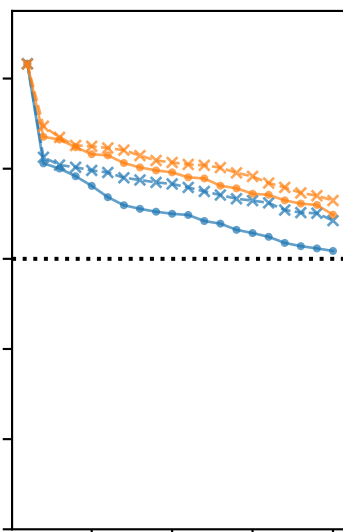
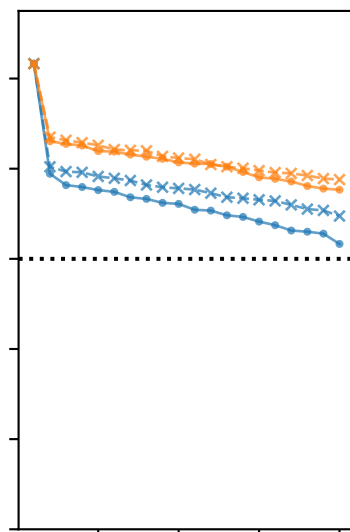
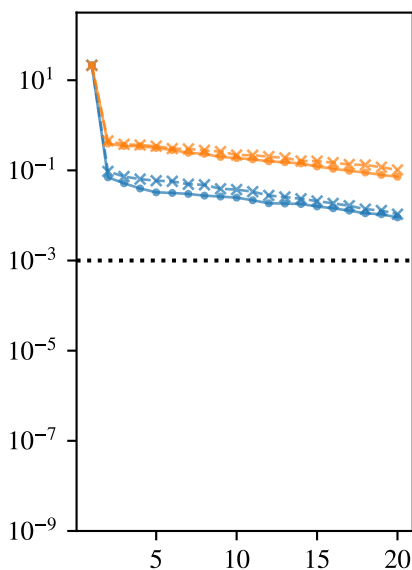
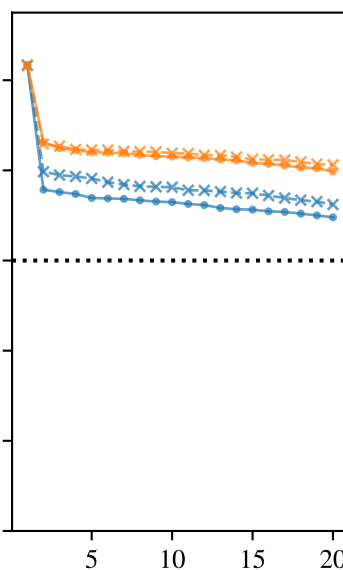
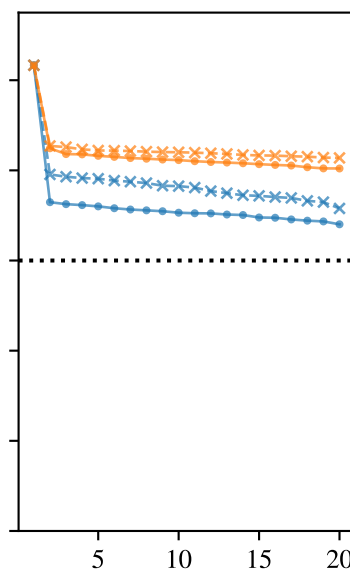


Singular Values of Trained Networks, $\sigma_k(\hat{f})$ $r=1, n=64$  $r=1, n=128$  $r=1, n=256$  $r=1, n=512$  $r=1, n=1024$  $r=1, n=2048$  $\sigma=0$ $\sigma=0.25$

without extra layers

- * - with extra layers

..... effective rank tolerance, $\varepsilon=10^{-3}$