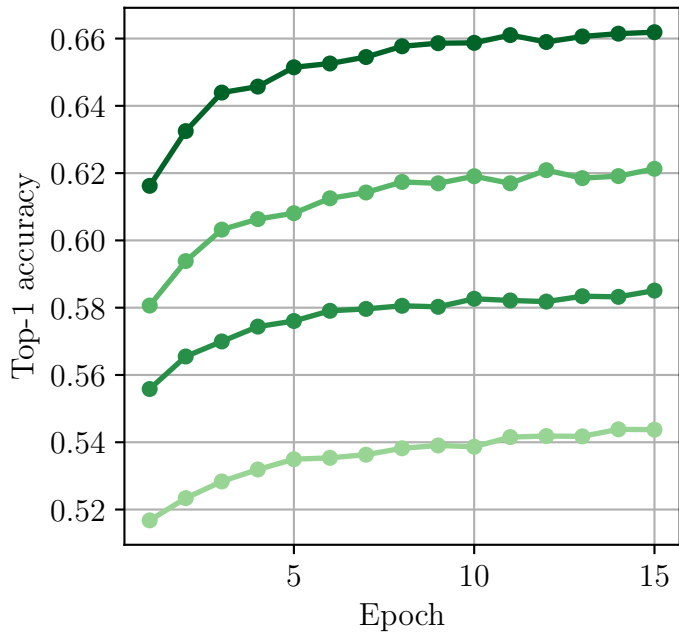
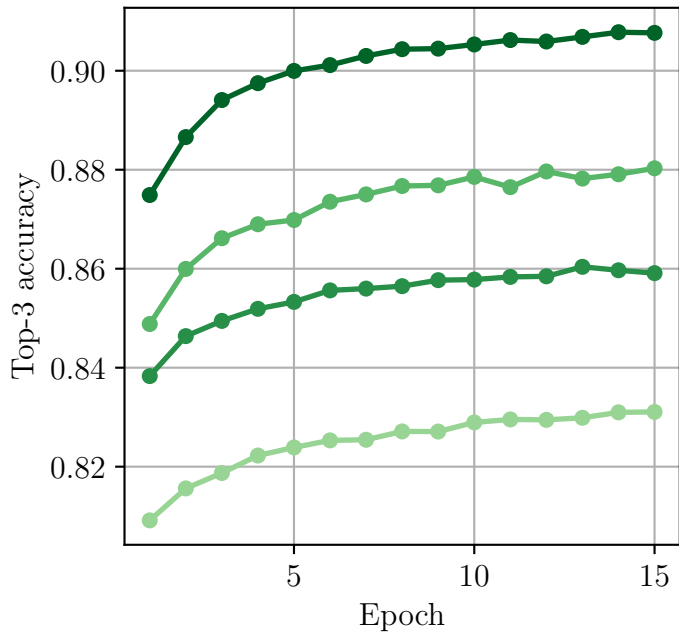


Top-1 validation accuracy



Top-3 validation accuracy



Policy network

 $f_{\pi,(32,5)}$  $f_{\pi,(32,10)}$  $f_{\pi,(64,5)}$  $f_{\pi,(64,10)}$