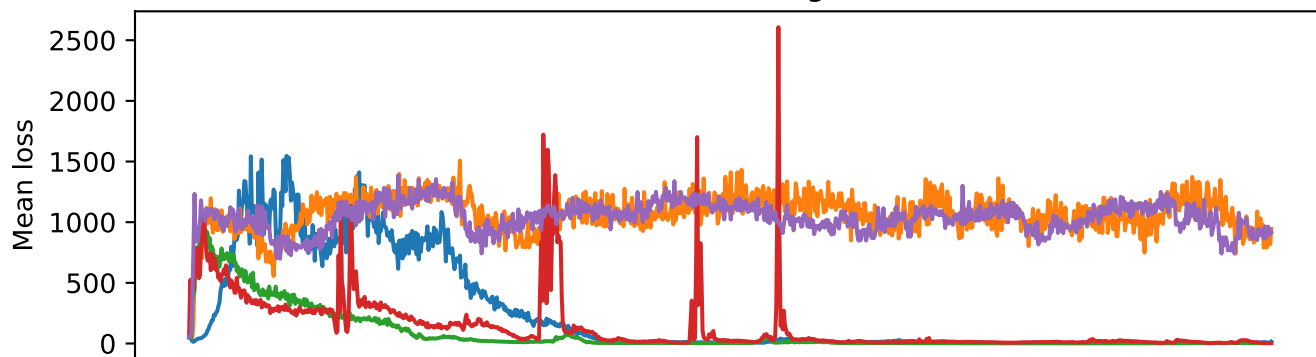
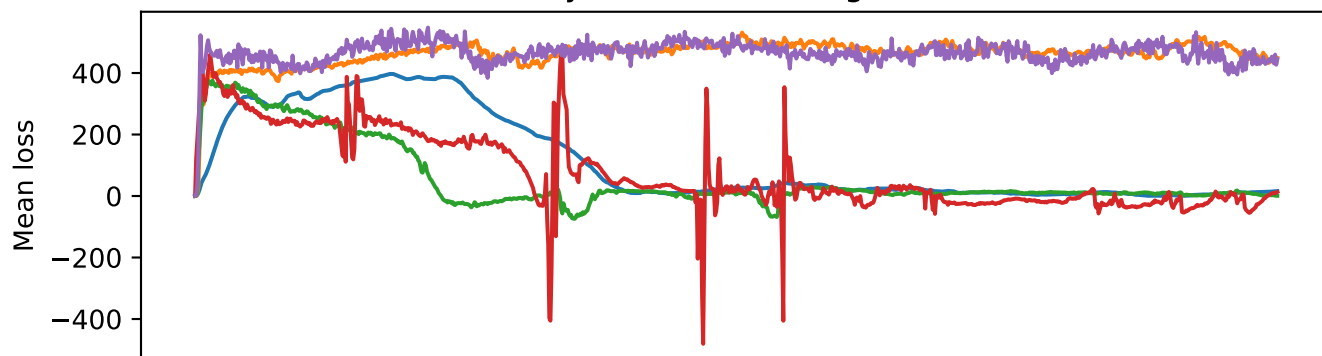


$\tau=0.01$ $\tau=0.2$ $\tau=0.5$ $\tau=0.8$ $\tau=1.0$

Q-network learning curves



Policy network learning curves



Smoothed returns (moving average window = 50)

