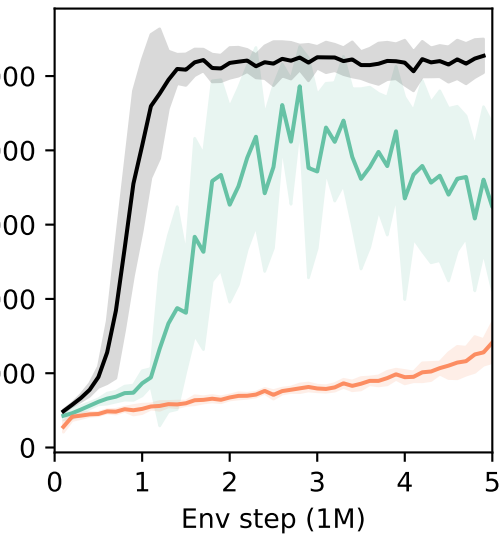
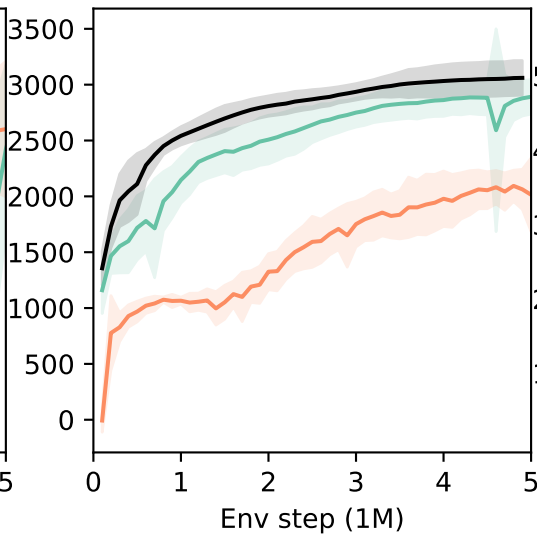
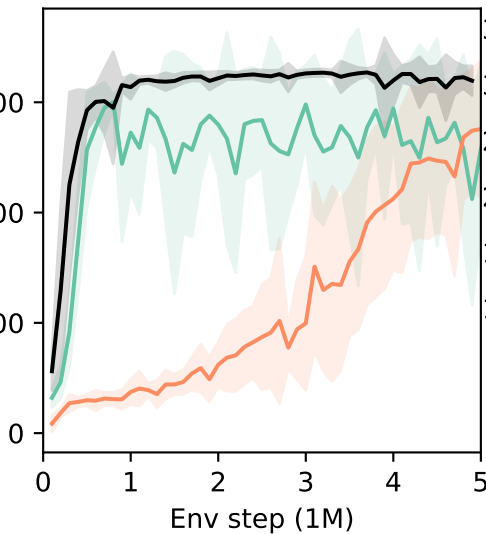


Hopper-v4

HalfCheetah-v4

Humanoid-v4

Undiscounted policy return



- $M=1, lr=10^{-5}$, large NN
- $M=1, lr=10^{-6}$, large NN
- $M=300$, default NN