

average return of
last 100 episodes

1.2

1.0

0.8

0.6

0.4

0.2

0.0

0K 100K 200K 300K 400K 500K 600K 700K 800K 900K 1000K
time step (in millions)

- $\pi_{\alpha}^{\text{PPO}}$
- π_{β}^{PPO}
- $\pi_{\gamma}^{\text{PPO}}$
- $\pi_{\delta}^{\text{PPO}}$
- $\pi_{\epsilon}^{\text{PPO}}$