Legend: - - - Cumulative reward of π_0 ⋯⋯ Moment of reaching of π_0 performance

PPO, TD3, SAC, DDPG, VPG, REINFORCE, Our agent