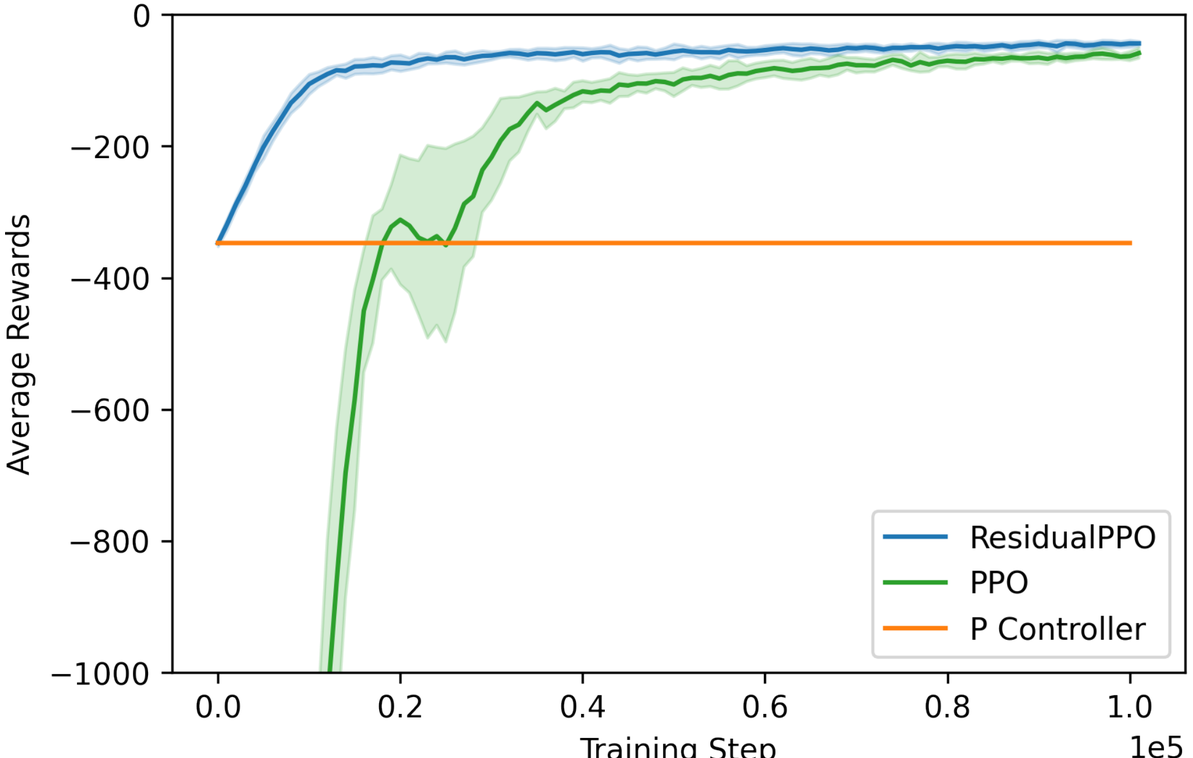
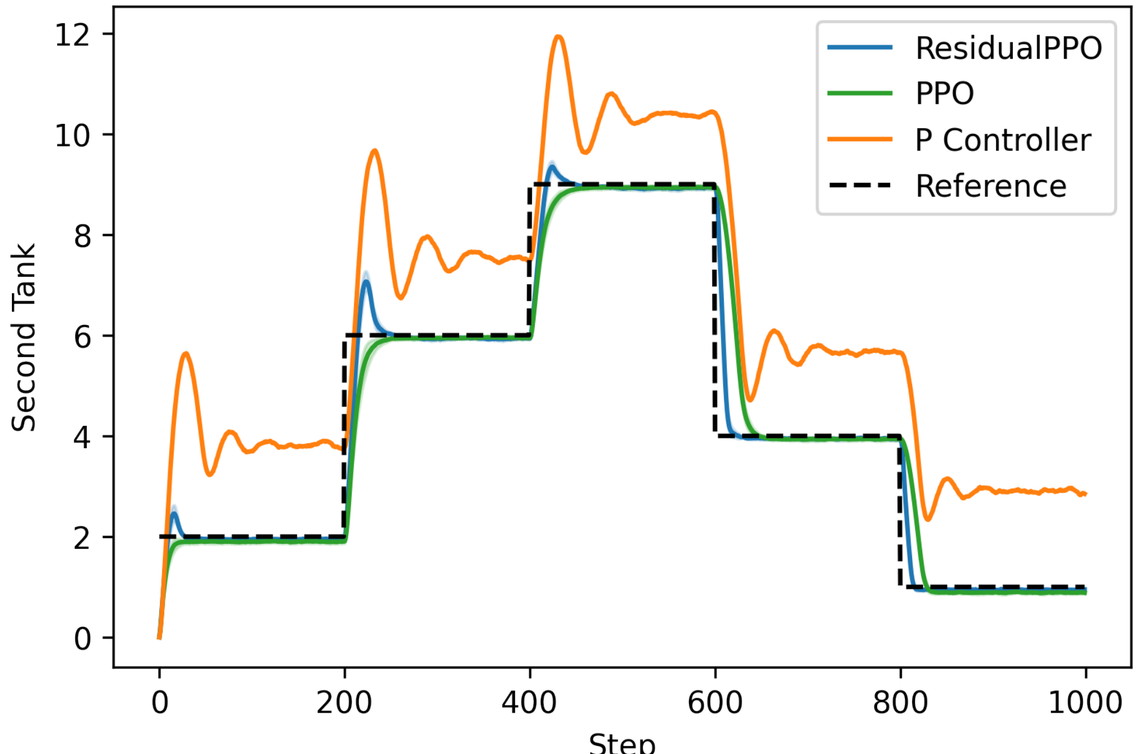

0
−200
−400
−600
−800
−1000
Average Rewards
0.0
0.2
0.4
0.6
0.8
1.0
1e5
Training Step
ResidualPPO
PPO
P Controller


12
10
8
6
4
2
0
Second Tank
0
200
400
600
800
1000
Step
ResidualPPO
PPO
P Controller
Reference