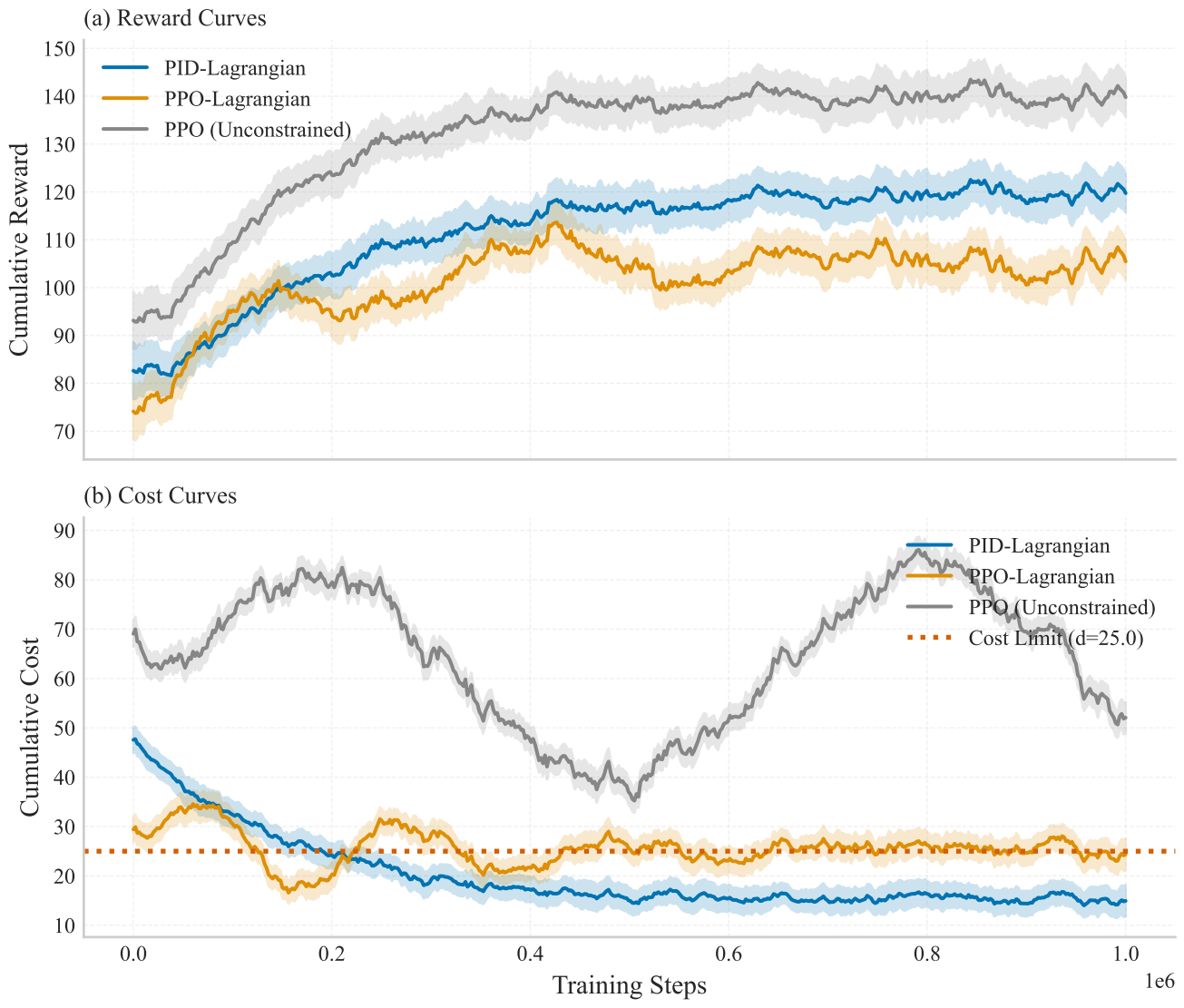
(a) Reward Curves
PID-Lagrangian
PPO-Lagrangian
PPO (Unconstrained)
Cumulative Reward
150
140
130
120
110
100
90
80
70
(b) Cost Curves
PID-Lagrangian
PPO-Lagrangian
PPO (Unconstrained)
Cost Limit (d=25.0)
Cumulative Cost
90
80
70
60
50
40
30
20
10
0.0
0.2
0.4
0.6
0.8
1.0
Training Steps
1e6