Pure Coordination Evaluation Seed Episode Reward Max Config Size 70 Small Medium 60 Large Obstacle 50 Policy Reward Max 20 10 0 2 4 5 1 3 0 1e6 Total steps