

PPO-Lagrangian Training Dynamics (Safety, Dual Variable, Return, KL)

