

Selected state-action trajectory for REINFORCE on two-tank system

