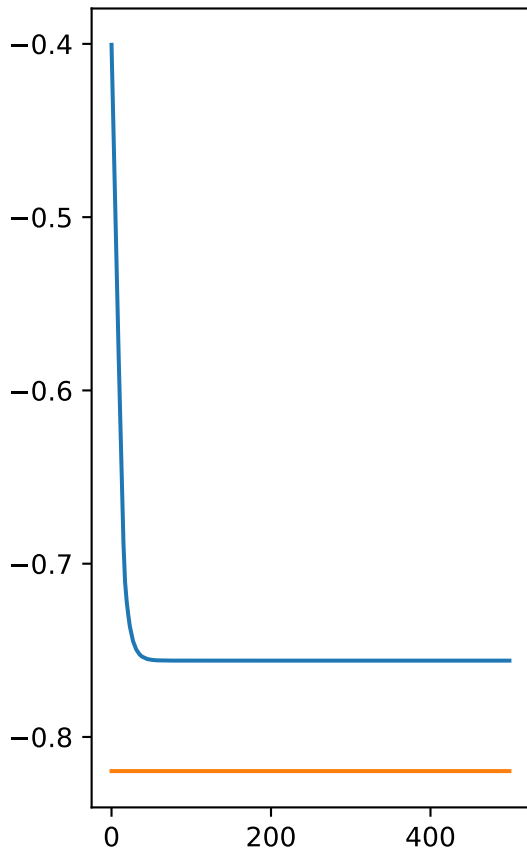
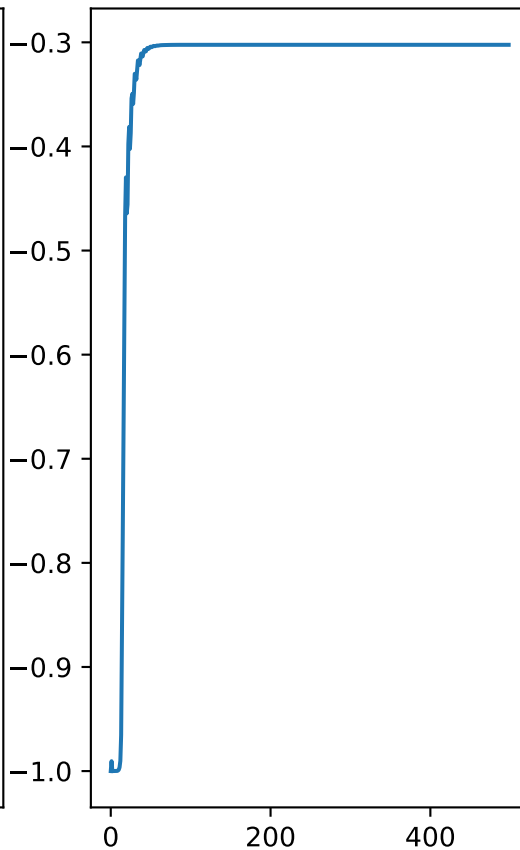


Reward: quadratic

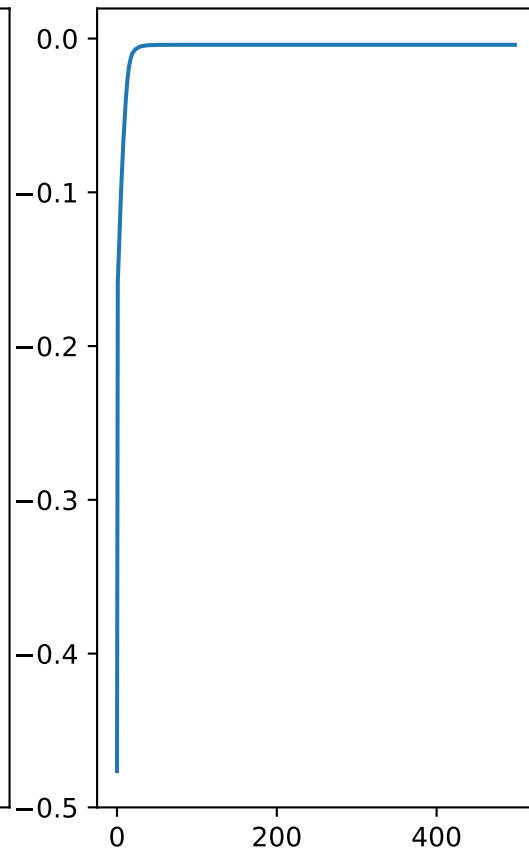
State vs step



Action vs step



Reward vs step



— I — I_{ref}

— V