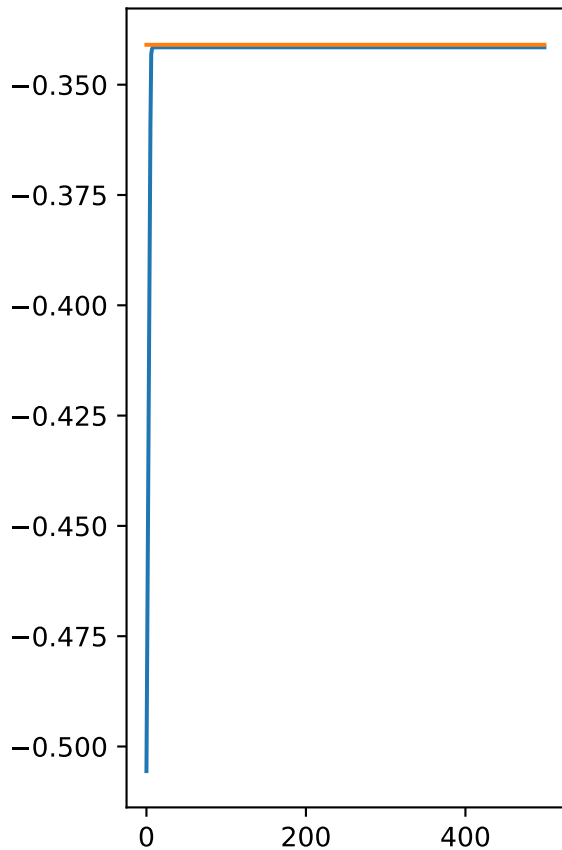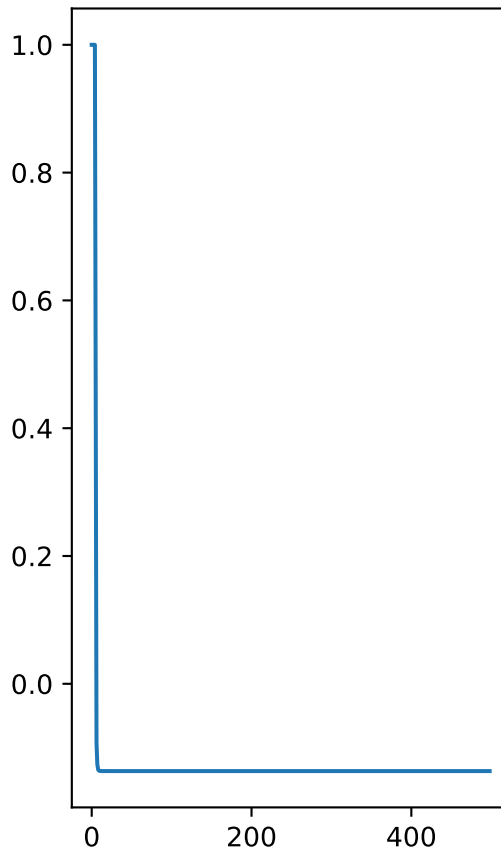Reward: square_root
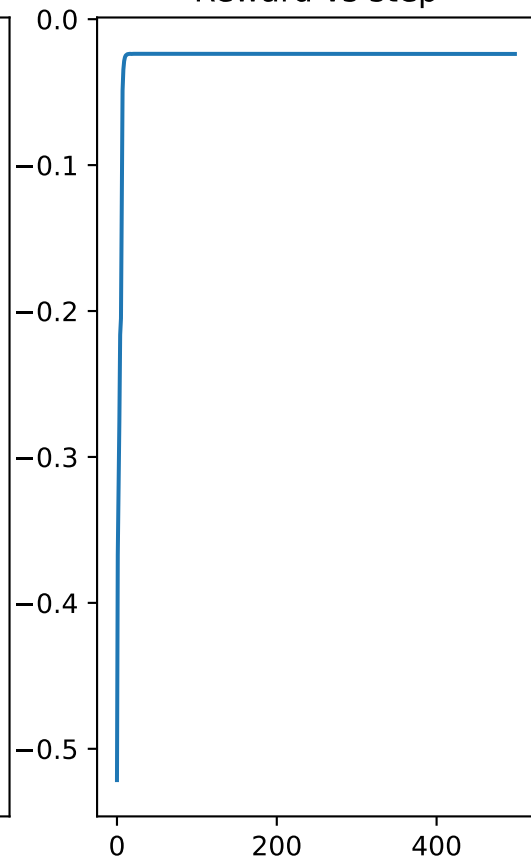
State vs step   Action vs step   Reward vs step

I   Iref   V