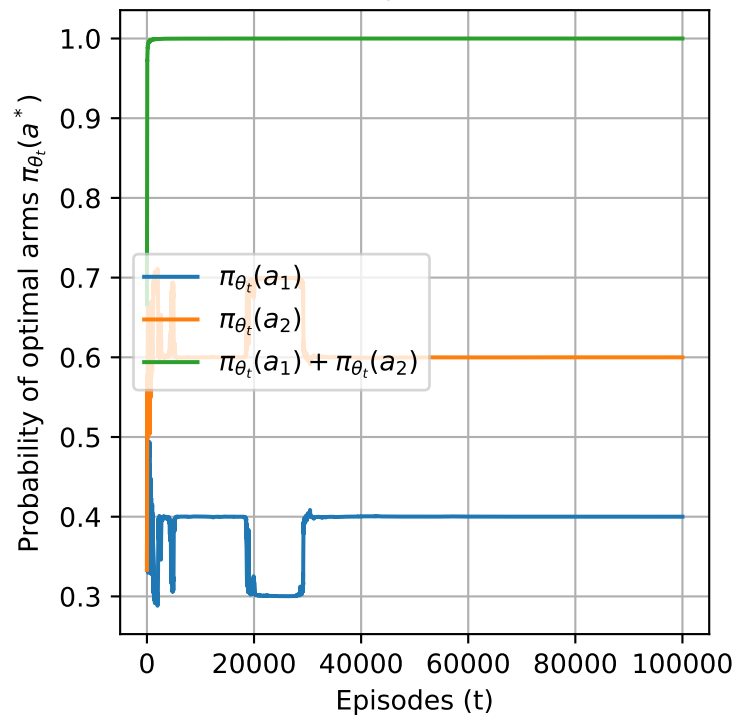
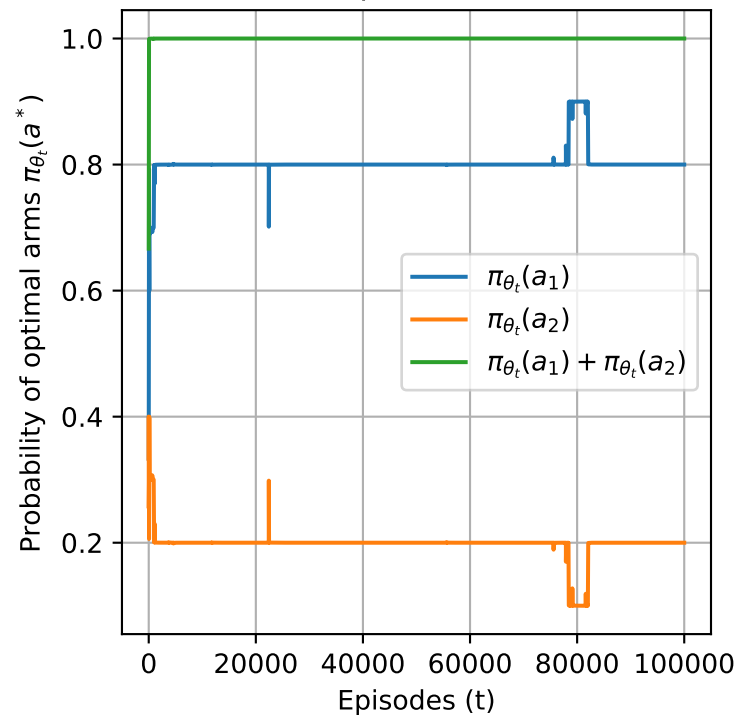


Total probability of optimal actions $\sum_{a^*} \pi_{\theta_t}(a^*)$

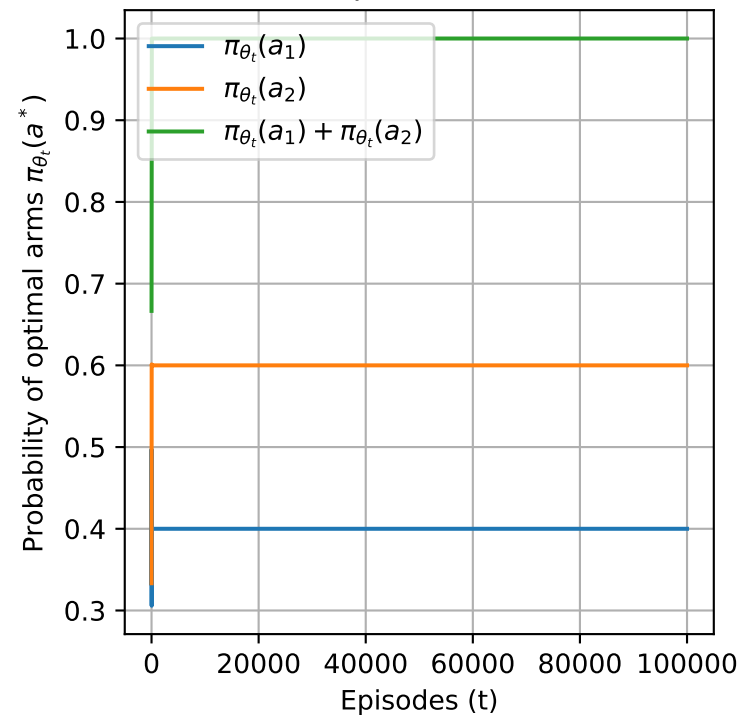
$\eta = 1$



$\eta = 10.0$



$\eta = 100.0$



$\eta = 1000.0$

