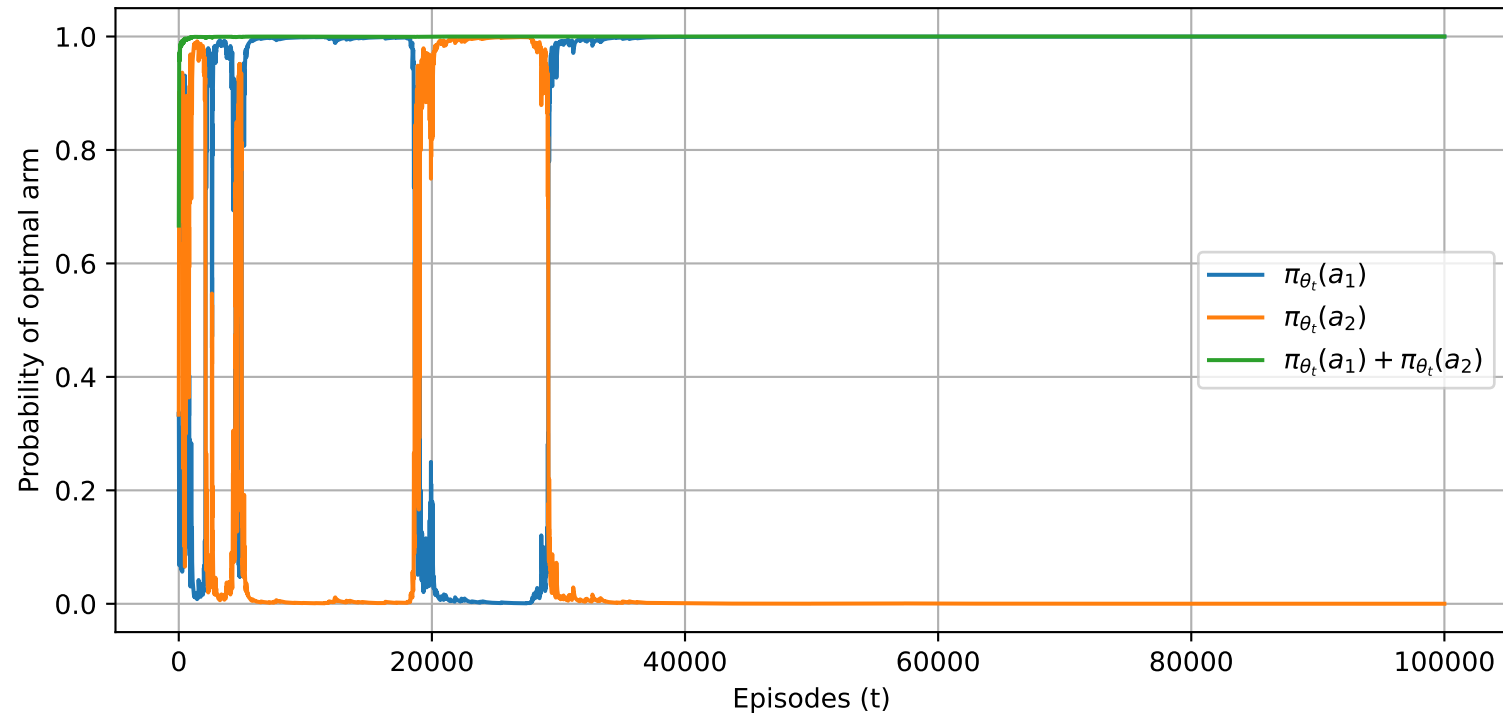


$$\text{Total probability of optimal actions } \sum_{a^*} \pi_{\theta_t}(a^*)$$

 $\eta = 1$  $\eta = 10$ 