



Figure 1: **Regret in the High-Tier/Target Task when Multiple Low-Tier/Source Tasks are Available** We report the regret in learning for  $1e7$  iterations when  $W$  source tasks are available with  $W = 0, 1, 2, 5$ . Here  $W = 0$  means the target task is solved by normal online learning without knowledge transfer from source tasks. The transfer starts from  $0.5e6$ . As we can see, as predicted by theory, our method can reduce regret by knowledge transfer, and when more tasks are available and introduce more transferable states, the benefits of knowledge transfer in target task will be more clear. The shadows indicate 96% confidence interval.