Goal: $\overset{\overset{\pi^*}{\nearrow}}{\arg\max_{\pi}} \; \overset{\overset{M(\pi^*)}{\nearrow}}{\min_{M \in \mathcal{M}}} \; J_M(\pi)$

Subopt: $J_{M(\pi^*)}(\pi^*) - J_{M(\hat{\pi})}(\hat{\pi})$.

Oracle: given $\pi$: $M(\pi)$.

Alg: NPG on $Q^{\pi}_{M(\pi)}$.

No-regret: $\sum_{i=1}^{T} \left( Q^{\pi_i}_{M(\pi_i)}(s, \pi) - Q^{\pi_i}_{M(\pi_i)}(s, \pi_i) \right) = O(\sqrt{T}$
$\forall \pi$

Output: $\hat{\pi}$ = mixtue of $\pi_{1:T}$.

Analysis: $J_{M(\pi^*)}(\pi^*) - J_{M(\pi_{1:T})}(\pi_{1:T})$

$$= J_{M(\pi^*)}(\pi^*) - \frac{1}{T} \sum_{i=1}^{T} J_{M(\pi_{1:T})}(\pi_i)$$

$$\leq J_{M(\pi_i)}(\pi^*) - \frac{1}{T} \sum_{T=1}^{T} J_{M(\pi_i)}(\pi_i)$$

$$= \frac{1}{T(1-\gamma)} \sum_{i=1}^{T} \mathbb{E}_{d^{\pi^*}_{M(\pi_i)}} \left[ Q^{\pi_i}_{M(\pi_i)}(s, \pi^*) - Q^{\pi_i}_{M(\pi_i)}(s, \pi_i) \right].$$

$$= O\left( \frac{1}{\sqrt{T}} \cdot \frac{1}{1-\gamma} \right).$$