

$$\begin{aligned}\nabla \mathbb{E}_{\tau \sim P_{\pi,T}}[R(\tau)] &= \sum_{\tau} P_{\pi,T}(\tau) R(a) \nabla \log P_{\pi,T}(\tau) \\ &= \mathbb{E}_{\tau \sim P_{\pi,T}} \left[ R(\tau) \nabla \log P_{\pi,T}(\tau) \right]\end{aligned}$$