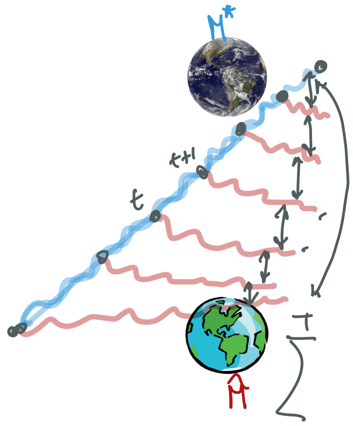
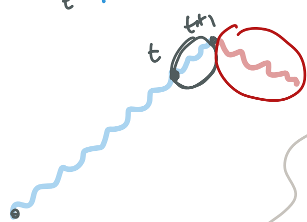
$$J_{M^*}(\pi^*) - J_{M^*}(\bar{\pi})$$

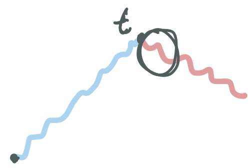
Ba



$$= \frac{\left| J_{M^*}(\pi) - J_{\hat{M}}(\pi) \right|}{}$$

$$\sum_{t=1}^{T} \mathbb{E}_{s_t \sim M^*} \left[\underline{\underline{Q^{M^*+\hat{M}}(s_t, \pi(s_t))}} - \underline{Q^{\hat{M}}(s_t, \pi(s_t))} \right]$$





$$\cancel{c(s_t, \pi(s_t))} + \mathbb{E}_{s_{t+1} \sim M^*} V^{\hat{M}}(s_{t+1}) - \left(\cancel{c(s_t, \pi(s_t))} + \mathbb{E}_{s_{t+1} \sim \hat{M}} V^{\hat{M}}(s_{t+1}) \right)$$

$$\sum_{t=1}^{T} \mathbb{E}_{s_{t+1} \sim M^*} V^{\hat{M}}(s_{t+1}) - \mathbb{E}_{s_{t+1} \sim \hat{M}} V^{\hat{M}}(s_{t+1})$$

$$\leq \sum_{t=1}^{T} \mathbb{E}_{s_t \sim M^*} \left| \sum_{s_{t+1}} M^*(s_{t+1} \mid s_t, a_t) V^{\hat{M}}(s_{t+1}) - \sum_{s_{t+1}} \hat{M}(s_{t+1} \mid s_t, a_t) V^{\hat{M}}(s_{t+1}) \right|_1$$

HOLDER'S INEQUALITY.

$$\leq \sum^{T} \mathbb{E} \left\| V^{\hat{M}}(s_{t+1}) \right\| \left| \sum M^*(s_{t+1} \mid s_t, a_t) - \right.$$

$$\sum_{t=1} \; \underset{s_t \sim M^*}{E} \; \underbrace{\| \ldots \|_\infty}_{V_{MAX}} \left| \sum \ldots \right.$$

$$\left. \sum_{s_{t+1}} \hat{M}(s_{t+1} \mid s_t, a_t) \right|,$$

$$\leq V_{MAX} \sum_{t=1}^{T} \underset{s_t \sim M^*}{E} \left| \sum M^*(s_{t+1} \mid s_t, a_t) - \sum_{s_{t+1}} \hat{M}(s_{t+1} \mid s_t, a_t) \right|,$$

PSINKER'S INEQ

$$\leq V_{MAX} \sum_{t=1}^{T} \underset{s_t \sim M^*}{\overset{\pi}{E}} \sqrt{KL(M^*, \hat{M})} \; .$$

STATES VISITED BY LEARNER $\hat{\pi}$

IN THE REAL WORLD

$$\underset{s_t \sim d_{M^*}^{\pi^*}}{E} \ldots KL(M^*, \hat{M})$$