

## SOURCES OF TRAIN / TEST MISMATCH

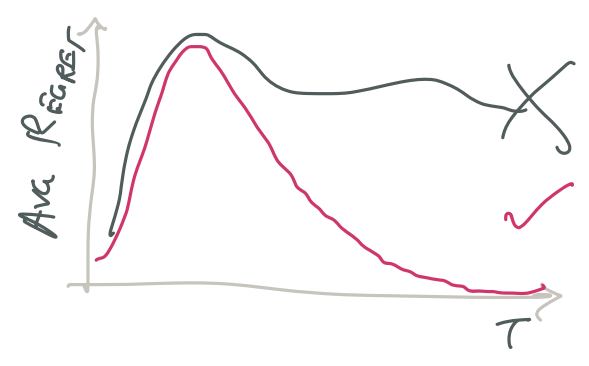
- ① LACK OF COVERAGE IN TRAINING
- ② ADVERSARIAL EXAMPLES (SHADOW)
- ③ CHANGE IN MDP
- ④ FEEDBACK
- ⑤ SIM 2 REAL

	$l_1$	$l_2$	$l_3$	...	$l_T$	$\sum$
$\pi^1$	0.1	$\pi_2 = 1.0$	1.0			11.5
$\pi^2$	$\pi_1 = 1.0$	0.1	0.1			23.0
...						
...			$\pi_3 = 0.5$			9.5
...						
$\pi^N$	1.0	0.5	1.0			100.0

$$\text{Avg. REGRET} = \frac{1}{T} \left[ \sum_{t=1}^T l_t(\pi_t) - \min_{\pi^* \in \Pi} \sum_{t=1}^T l_t(\pi^*) \right]$$

LEARNER BEST POLICY IN HANDSIGHT

No REGRET As  $T \rightarrow \infty$ , Avg. REGRET  $\rightarrow 0$



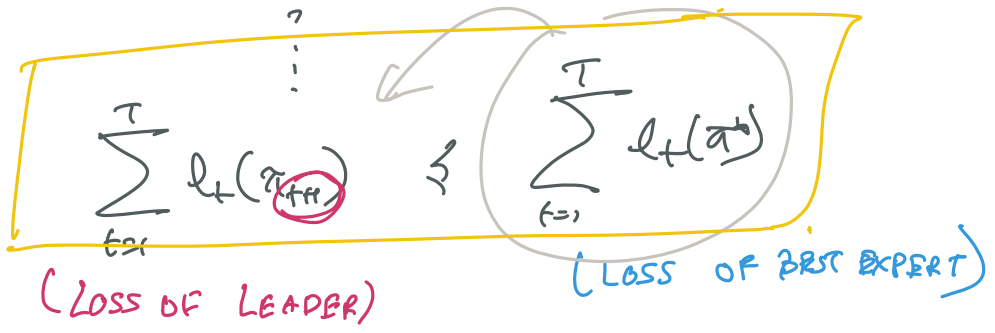
FTL  $\pi_t = \operatorname{argmin}_{\pi \in \Pi} \sum_{i=1}^{t-1} l_t(x)$

REGRET:  $\sum_{t=1}^T l_t(\pi_t) - \min_{\pi^*} \sum_{t=1}^T l_t(\pi^*)$

FUTURE REGRET =  $\sum_{t=1}^T l_t(x_{t+1}) - \min_{\pi^*} \sum_{t=1}^T l_t(\pi^*)$

$l_1(\pi_2) = \min_{\pi} l_1(\pi) \leq l_1(\pi^*)$

$l_1(\pi_2) + l_2(\pi_3) \leq \boxed{l_1(\pi_3) + l_2(\pi_3)} = \min_{\pi} l_1(x) + l_2(\pi) \leq \underline{l_1(\pi^*) + l_2(\pi^*)}$



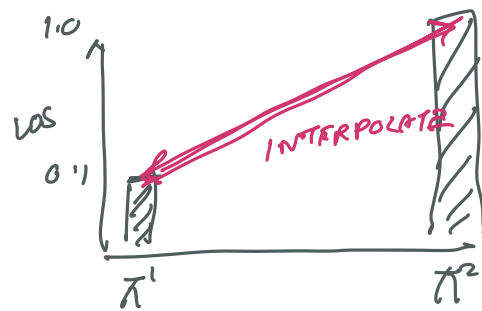
REG =  $\sum_{t=1}^T l_t(\pi_t) - \min_{\pi^*} \sum_{t=1}^T l_t(\pi^*)$

$\leq \sum_{t=1}^T l_t(\pi_t) - \sum_{t=1}^T l_t(\pi_{t+1})$

$\leq \sum_{t=1}^T ( \underset{\text{PRED AT } t}{l_t(\pi_t)} - \underset{\text{PRED AT } t+1}{l_t(\pi_{t+1})} )$

As  $t \rightarrow \infty$ ,  $\pi_t \rightarrow \pi_{t+1} \Rightarrow$  No REGRET!

EXPERTS	$l_1$	$P_1$	$l_2$	$P_2$
$\pi^1$	0.1	0.5	1.0	0.5
$\pi^2$	1.0	0.5	0.1	0.5



EXPECTED REGRET =

$$\sum_{t=1}^T E_{\pi_t, P_t} l_t(\pi_t)$$

⋮

$P_t \cdot l_t$

$$- \min_{\pi^*} \sum_{t=1}^T l_t(\pi^*)$$

"LIFTED TO A

LINEAR PROGRAM"