Lecture 7.3

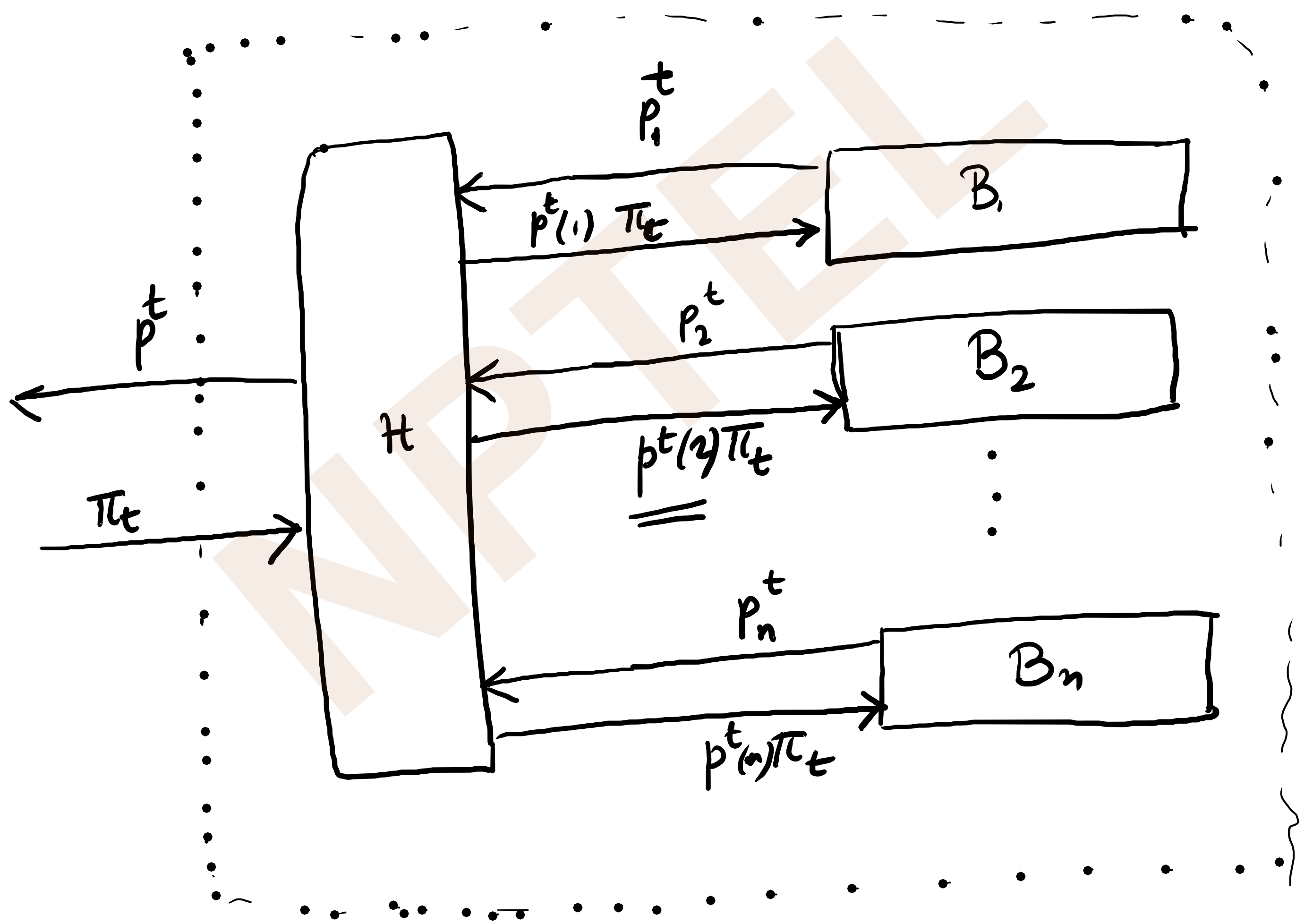
Q. Does there exist a no-swap-regret algorithm?

Black box reduction from no-swap regret to no external regret:

<u>Theorem:</u> Let $|A| = n$. Suppose there exists an algorithm with time averaged external regret $R(T, n)$. Then there also exists an algorithm with time averaged swap regret $n\,R(T, n)$. In particular, if there exists a

no external regret algorithm (i.e. $\lim_{T\to\infty} R(T,n) = 0$), then there also exists a no swap regret algorithm.

<u>Proof</u>: Let the action set of the player (of the no swap regret algorithm) be $A = \{1, \dots, n\}$. Let $B$ be an algorithm with time-averaged external regret $R(T,n)$. Take $n$ copies of $B$; let us call them $B_1, \dots, B_n$. We will build a "master algorithm" $H$ using $B_1, \dots, B_t$



$p^t$
$\pi_t$
$\mathcal{H}$
$p_1^t$
$p^t(1)\,\pi_t$
$B_1$
$p_2^t$
$p^t(2)\pi_t$
$B_2$
$p_n^t$
$p^t(n)\pi_t$
$B_n$

The time-averaged expected pay off of the master algorithm is

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{i \in A} p^t(i) \cdot \pi_t(i)$$

Let $\delta : A \to A$ be any switching function. Then the time-averaged expected pay off of the master algorithm (modified by $\delta$) is

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{i \in A} p^t(i)\, \pi_t(\delta(i))$$

We need to show,

$$\frac{1}{T}\left[\sum_{t=1}^{T}\sum_{i\in A} p^t(i)\, \pi_t(\delta(i)) - \sum_{t=1}^{T}\sum_{i\in A} \boxed{p^t(i)}\, \pi_t(i)\right] \leq n \cdot R(T,n)$$

Since the external regret of each $B_i$ is $R(T,n)$, we have the following

$$\frac{1}{T}\left(\sum_{t=1}^{T} p^t(i)\, \pi_t(\lambda) - \sum_{t=1}^{T}\sum_{j\in A} p_i^t(j)\, p^t(i)\, \pi_t(j)\right) \leq R(T,n)$$

$$\forall\ i \in A,\ \lambda \in A$$

Put $\lambda = \delta(i)$ in the above inequality.

$$\frac{1}{T}\left( \sum_{t=1}^{T} p^t(i)\, \pi_t(\delta(i)) - \sum_{t=1}^{T} \sum_{j \in A} p_i^t(j)\, p^t(i)\, \pi_t(j) \right) \leq R(T,n)$$

$\forall\, i \in A.$

Adding all the above $n$ inequalities,

$$\frac{1}{T}\left( \sum_{t=1}^{T} \sum_{i \in A} p^t(i) . \pi_t(\delta(i)) - \sum_{t=1}^{T} \sum_{i \in A} \sum_{j \in A} p_i^t(j)\, p^t(i)\, \pi_t(j) \right) \leq n R(T,n)$$

$$\Rightarrow \frac{1}{T}\left( \underline{\sum_{t=1}^{T} \sum_{i \in A} p^t(i)\, \pi_t(\delta(i))} - \underline{\sum_{t=1}^{T} \sum_{i \in A} \pi_t(i)} \boxed{\sum_{j \in A} p_j^t(i)\, p^t(j)} \right) \leq n R(T,n)$$

Set $p^t$ as a solution to the following system of linear equation.

$$p^t(i) = \sum_{j \in A} \underline{\underline{p^t_j(i)}} \, p^t(j) \qquad \forall \, i \in A$$

like constants.

$$\sum_{i \in A} p^t(i) = 1$$

<u>Corollary</u>: $|A| = n$. There exists an algorithm with swap regret $O\left(n \cdot \sqrt{\frac{\ln n}{T}}\right)$.