Link Throughput of Multi-Channel Opportunistic Access with Limited Sensing

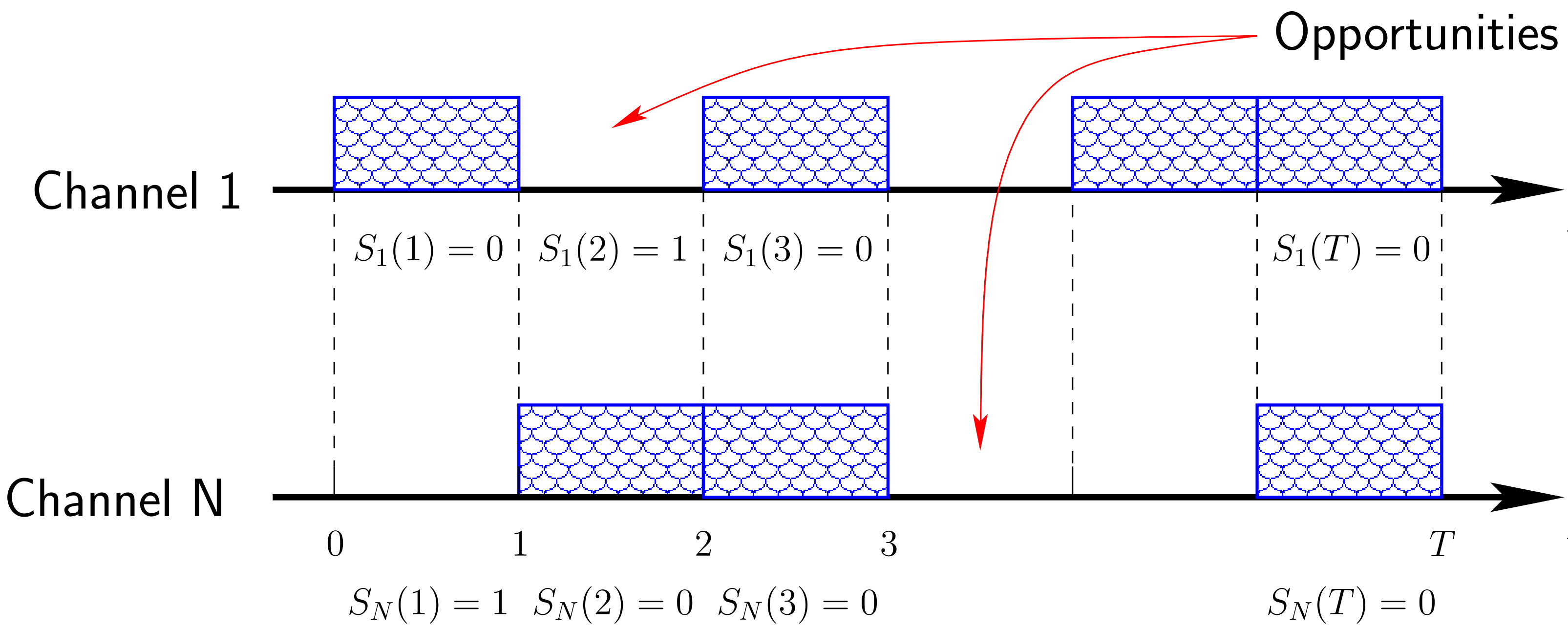
Keqin Liu, Qing Zhao

Department of Electrical and Computer Engineering

University of California, Davis, CA 95616

Supported by NSF and ARL-CTA.

Multi-Channel Opportunistic Access

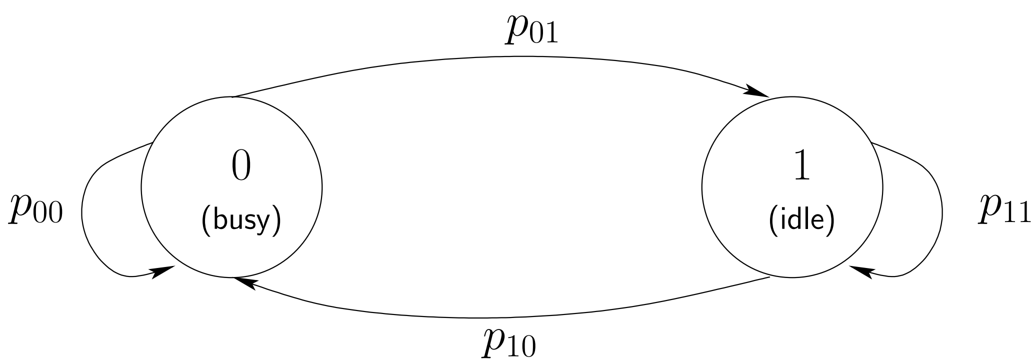


- Opportunistic Access: adapt to time-varying channel state.
- Channel State: fading condition, presence of primary users, presence of jammers (jamming/anti-jamming).
- Limited Sensing: can only sense and access one channel in each slot.

Sensing Strategy: Which channel to sense?

Gilbert-Elliot Channel Model[1]

- N i.i.d. Gilbert-Elliot channels.



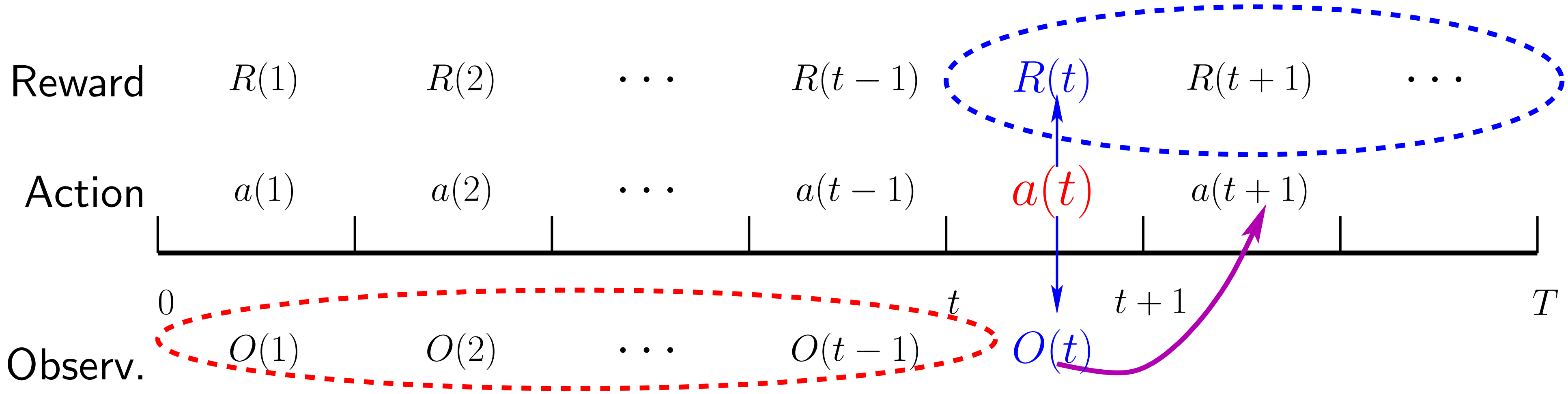
- **Sensing Policy π_s**
  - Choose the sensing action $a(t)$ in each slot t
- **Immediate Reward**
  - If the chosen channel a is idle, a unit reward is accrued.
  - If the chosen channel a is busy, no reward; wait until the next slot.
- **Objective:** choose sensing policy π_s to

$$\max \mathbb{E}[\text{throughput}]$$

[1] E.N. Gilbert, "Capacity of burst-noise channels," Bell Syst. Tech. J., vol. 39, pp. 1253-1265, Sept. 1960.

Sensing Policy: Gaining Access vs. Gaining Information

Optimal Sensing Policy: sequential decision-making



- Each sensing outcome provides information on the state of the system.
- $a(t)$ should be based on the entire observation history.
- $a(t)$ results in an immediate reward $R(t)$ and an observation $O(t)$ that affects future actions and reward.
- Optimal $a(t)$ achieves the best tradeoff between gaining immediate reward and gaining spectrum information.

Restless Multi-armed Bandit

Restless Multi-armed Bandit Problem

- A bandit consists of N independent arms.
- At each slot t, the state of arm i is $Z_i(t)$.
- Given system state $\mathbf{Z}(t) = \{Z_1(t), \cdot\cdot Z_N(t)\}$, activate an arm i and get a reward $R_i(Z_i(t))$ in slot t.
- The state of active arm i transits according to a Markov chain, the state of each passive arm transits according to another Markov chain.

Objective

Decide which arm to activate in each slot to maximize the expected long-term total reward.

Restless Multi-armed Bandit Formulation

- Each channel is considered as an arm.
- If channel i is sensed, then it is "activated".
- The state of each arm should be the observation history of that channel.
- Sufficient statistic: the a posterior distribution (belief vector) $\Omega(t)$ that exploits the entire observation history.

$$\begin{aligned}\Omega(t) &= [\omega_1(t), \cdots, \omega_N(t)] \\ \omega_i(t) &= \Pr[\text{channel } i \text{ is idle in slot } t \mid \underbrace{O(1), \cdots, O(t-1)}_{\text{observations}}]\end{aligned}$$

- The state of arm i in slot t is $\omega_i(t)$.
- The expected immediate reward obtained when activate arm i is $\omega_i(t)$.

Markovian Transition of Belief and Value Function

- The belief vector transits according to Markov processes.
- If channel i is activated in slot t:

$$\omega_i(t+1) = \begin{cases} p_{11}, & \text{if} \quad O(t)=1 \\ p_{01}, & \text{if} \quad O(t)=0 \end{cases}. \tag{1}$$

- If channel i is made passive in slot t:

$$\omega_i(t+1) = \omega_i(t)p_{11} + (1-\omega_i(t))p_{01}. \tag{2}$$

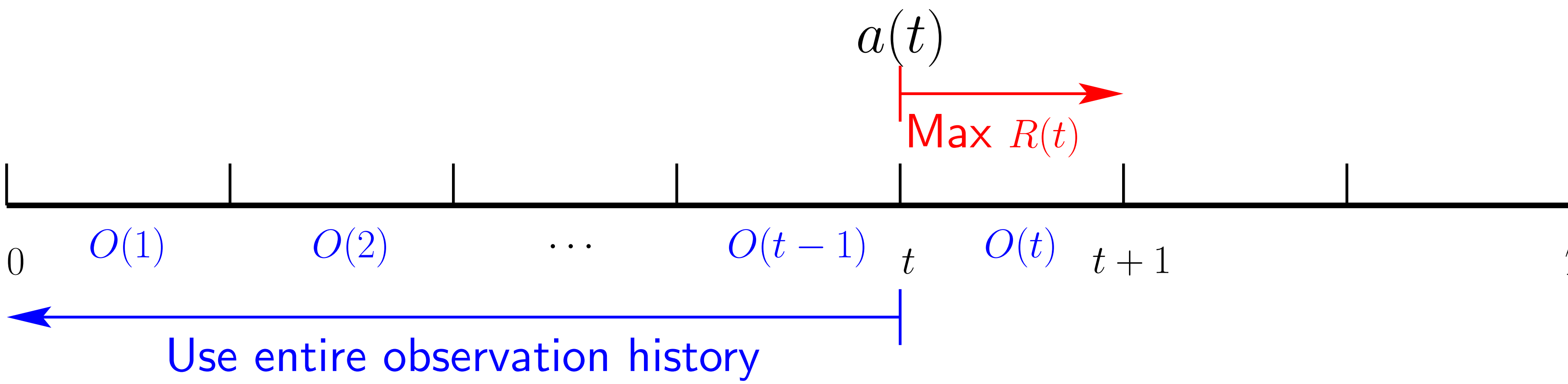
- Value Function: let $V_t(\Omega)$ be the maximum expected total reward accumulated from slot t given the current belief vector Ω.

$$V_t(\Omega) \;=\; \max_a(\omega_a + \mathbb{E}_{S_a} V_{t+1}(\mathcal{T}(\Omega|a,S_a))). \tag{3}$$

- Throughput as Average Reward:

$$U(\Omega) = \lim_{T\to\infty} \frac{1}{T} V_1(\Omega) \tag{4}$$

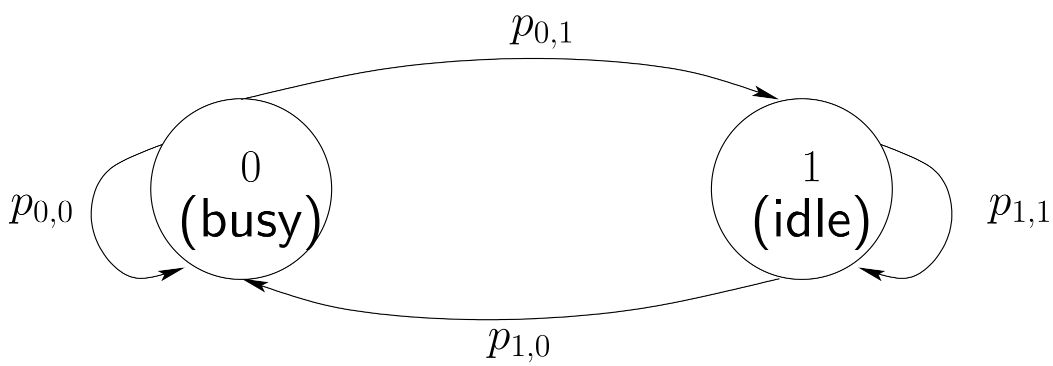
The Myopic Sensing Policy



- Myopic policy: maximize immediate reward $R(t)$

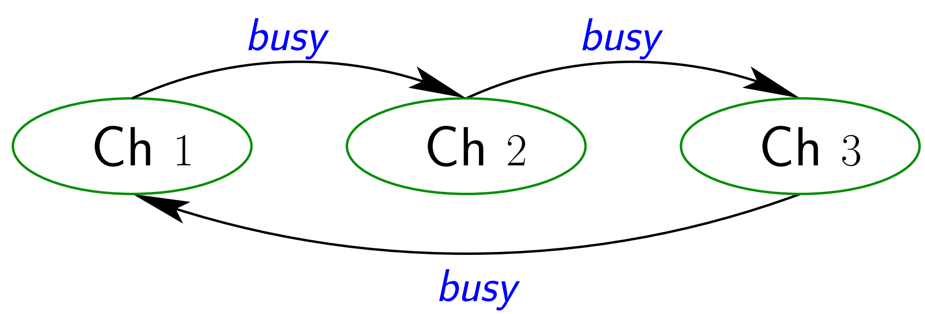
$$
\begin{aligned}
a(t) &= \arg\max_{a=1,\cdots,N} R_a(t) \\
&= \arg\max_{a=1,\cdots,N} \Pr[a \text{ is idle} \mid \underbrace{O(1),\cdots,O(t-1)}_{\text{observations}}] \\
&= \arg\max_{a=1,\cdots,N} \omega_a \qquad (5)
\end{aligned}
$$

Structure of Myopic Sensing for i.i.d. Markov Processes[2]



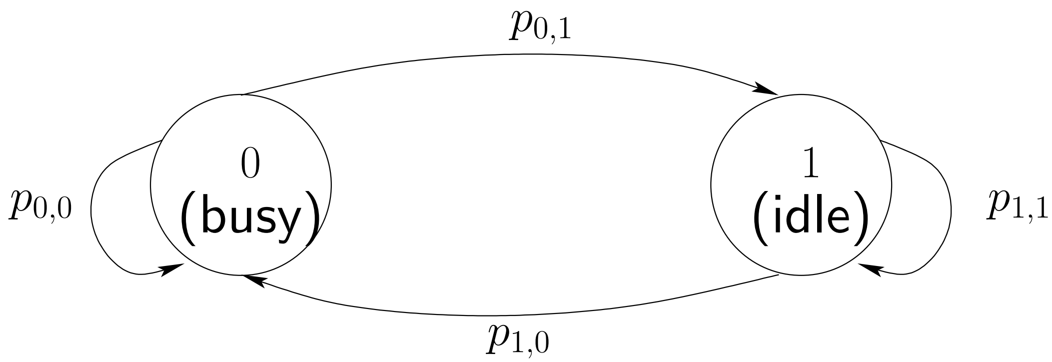
The Structure of Myopic Sensing Policy: $p_{1,1} > p_{0,1}$

- Stay in the same channel if it is idle and switch if it is busy.
- Switch to the channel visited the longest time ago.
- A sufficient statistic: current observation (no belief update).
- No need to know the transition probabilities.
- Automatically tracks model variations.



[2] Q. Zhao and B. Krishnamachari, "Structure and optimality of myopic sensing for opportunistic spectrum access," in *Proc. of IEEE CogNet*, pp. 6476–6481, June, 2007.

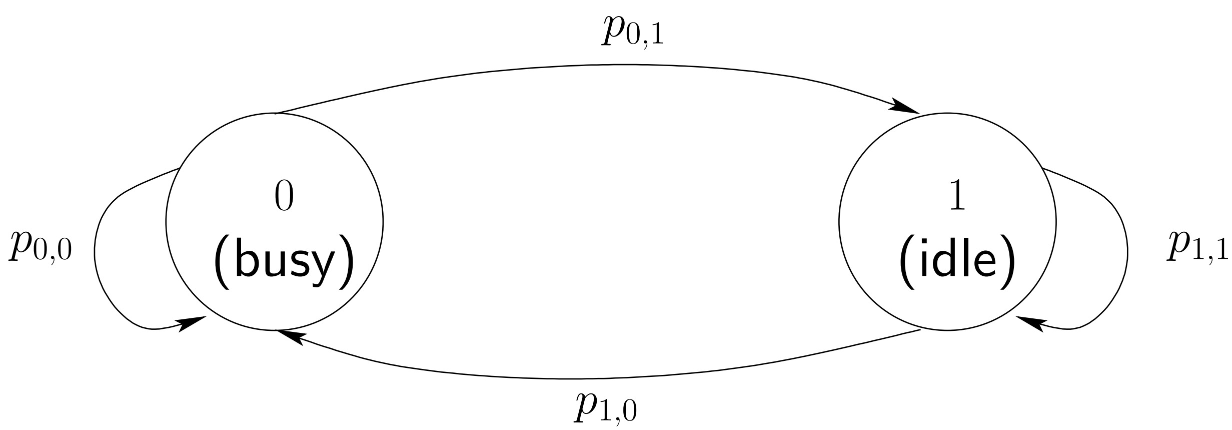
Structure of Myopic Sensing for i.i.d. Markov Processes



The Structure of Myopic Sensing Policy: $p_{0,1} > p_{1,1}$

- Stay in the same channel if it is busy and switch if it is idle.
- Among channels visited an even number of slots ago, choose the most recent.
- If no such channels, choose the one visited the longest time ago.
- A sufficient statistic: current observation and last visit to each channel.
- No need to know the transition probabilities.
- Automatically tracks model variations.

Optimality of Myopic Sensing for i.i.d. Markov Processes



The Optimality of Myopic Sensing Policy[3,4]

- Proven to be optimal for i) $N = 2$ and $N = 3$; ii) $N > 3$ with $p_{11} > p_{01}$.

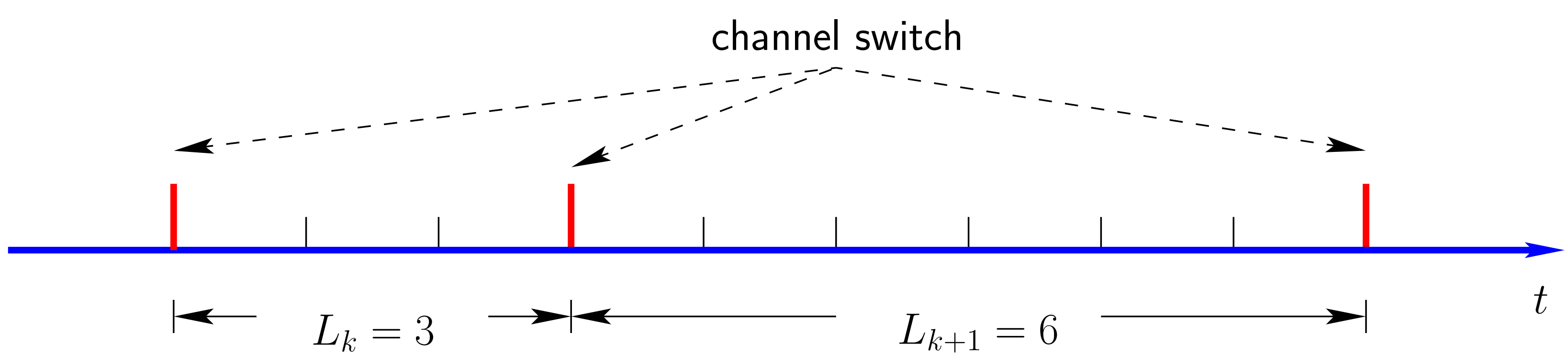
[3] Q. Zhao and B. Krishnamachari, "Structure and Optimality of Myopic Sensing for Opportunistic Spectrum Access," CogNet 2007.
[4] T. Javidi, B. Krishnamachari, Q. Zhao, and M. Liu, "Optimality of Myopic Sensing in Multi-Channel Opportunistic Access," ICC 2008.

Link Throughput Limit

- Throughput limit is the maximum average reward.

$$U(\Omega_o) = \lim_{T\to\infty} \frac{V_1(\Omega_o)}{T}, \quad (6)$$

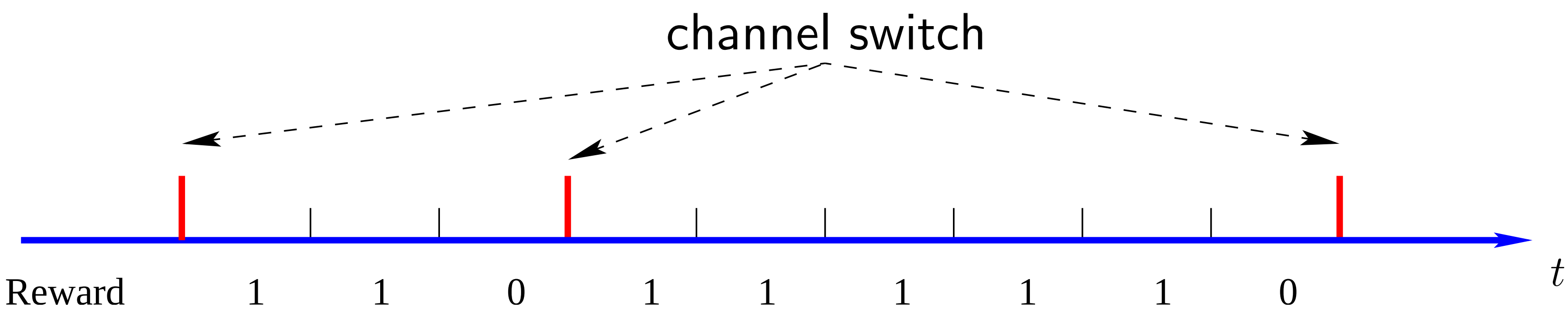
- Transmission Period: The time the user stays in the same channel.



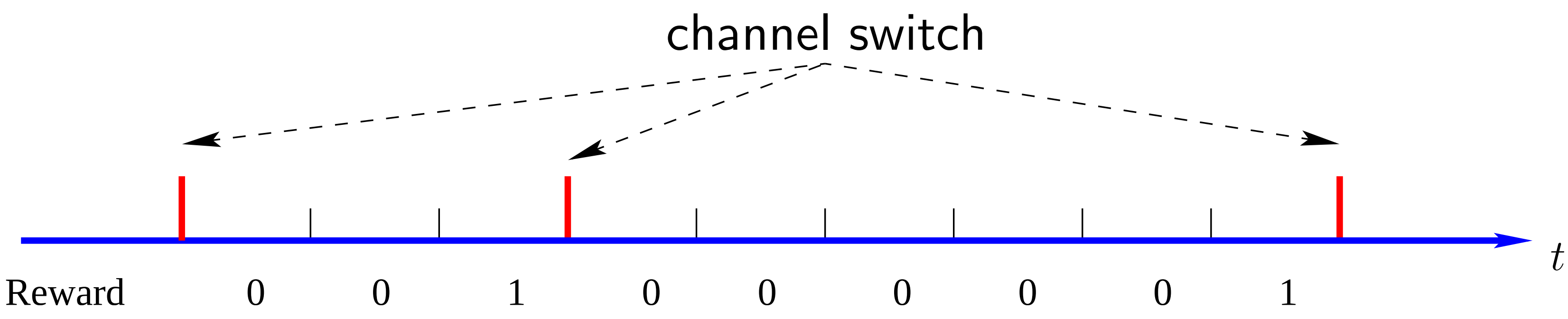
- For $N = 2$, $\{L_k\}_{k=1}^{\infty}$ is a first-order ergodic Markov chain.
- For $N > 2$ and $p_{11} \geq p_{01}$, $\{L_k\}_{k=1}^{\infty}$ is an $(N-1)$th-order ergodic Markov chain.

Reward Obtained in A Transmission Period

The reward obtained in a transmission period: $p_{11} > p_{01}$



The reward obtained in a transmission period: $p_{11} < p_{01}$



- Throughput limit is determined by the average length of a transmission period.

Link Throughput Limit

Theorem 1 *Let* $\bar{L} = \lim_{k\to\infty} \frac{\Sigma_{k=1}^{K} L_k}{K}$ *denote the average length of a transmission period. The throughput limit* $U(\Omega_o)$ *is given by*

$$U(\Omega_o) = \begin{cases} 1 - 1/\bar{L}, & p_{11} \geq p_{01} \\ 1/\bar{L}, & p_{11} < p_{01} \end{cases}. \quad (7)$$

Proof. When $p_{11} \geq p_{01}$, the user collects $(L_k - 1)$ units reward during each transmission period L_k.

$$U(\Omega_o) = \lim_{K\to\infty} \frac{\Sigma_{k=1}^{K}(L_k - 1)}{\Sigma_{k=1}^{K} L_k} = 1 - \frac{1}{\lim_{K\to\infty} \frac{\Sigma_{k=1}^{K} L_k}{K}} = 1 - \frac{1}{\bar{L}}, \quad (8)$$

When $p_{11} < p_{01}$, the user collects 1 unit reward during each transmission period L_k.

$$U(\Omega_o) = \lim_{K\to\infty} \frac{\Sigma_{k=1}^{K} 1}{\Sigma_{k=1}^{K} L_k} = \frac{1}{\lim_{K\to\infty} \frac{\Sigma_{i=1}^{K} L_k}{K}} = \frac{1}{\bar{L}}. \quad (9)$$

Link Throughput Limit for $N = 2$

- If $N = 2$, the average length $\bar{L}$ of a transmission period is obtained by the stationary distribution of $\{L_k\}_{k=1}^{\infty}$.

Theorem 2 *For $N = 2$, the throughput limit U is given by*

$$U = \begin{cases} 1 - \frac{1-p_{11}}{1+\bar{\omega}-p_{11}}, & p_{11} \geq p_{01} \\ \frac{p_{01}}{1-\bar{\omega}'+p_{01}}, & p_{11} < p_{01} \end{cases}, \tag{10}$$

$$\bar{\omega} = \frac{p_{01}^2}{1 + p_{01}^2 - A}, \tag{11}$$

where $A = \frac{p_{01}}{1+p_{01}-p_{11}}(1 - \frac{(p_{11}-p_{01})^3(1-p_{11})}{1-(p_{11})^2+p_{11}p_{01}})$.

$$\bar{\omega}' = \frac{B}{1 - p_{11}^2 + B}, \tag{12}$$

where $B = \frac{p_{01}}{1+p_{01}-p_{11}}(1 + \frac{(p_{11}-p_{01})^3(1-p_{11})}{1-(1-p_{01})(p_{11}-p_{01})})$.

Link Throughput Limit for $N > 2$

- For $N > 2$, $\{L_k\}_{k=1}^{\infty}$ is a higher order Markov chain; the average length $\bar{L}$ of a transmission period is difficult to obtain.

- Construct first-order Markov chain that stochastically dominates or being dominated by $\{L_k\}_{k=1}^{\infty}$

Link Throughput Limit for $N > 2$

Theorem 3 *For $N > 2$, we have the following lower and upper bounds on the throughput limit U.*

- *Case 1:* $p_{11} \geq p_{01}$

$$\frac{C}{C + (1 - D + C)(1 - p_{11})} \leq U \leq \frac{\omega_o}{1 - p_{11} + \omega_o}, \tag{13}$$

where $\omega_o = \frac{p_{01}}{p_{01}+p_{10}}$, $C = \omega_o(1 - (p_{11} - p_{01})^N)$, $D = \omega_o(1 - \frac{(p_{11}-p_{01})^{N+1}(1-p_{11})}{1-(p_{11})^2+p_{11}p_{01}})$.

- *Case 2:* $p_{11} < p_{01}$

$$\frac{p_{10}^2}{p_{01}H - E} + 1 \leq U(\Omega_o) \leq \frac{p_{10}^2}{p_{01}G - E} + 1 \tag{14}$$

where $E = p_{10}^2(1 + p_{01}) + p_{01}(1 - F)$,

$$F = (1 - p_{01})(1 - \omega_o)(\frac{1}{2-p_{01}} - \frac{p_{01}(p_{11}-p_{01})^4}{1-(p_{11}-p_{01})^2(1-p_{01})^2}),$$

$$G = (1 - \omega_o)(\frac{1}{2-p_{01}} - \frac{p_{01}(p_{11}-p_{01})^6}{1-(p_{11}-p_{01})^2(1-p_{01})^2}),$$

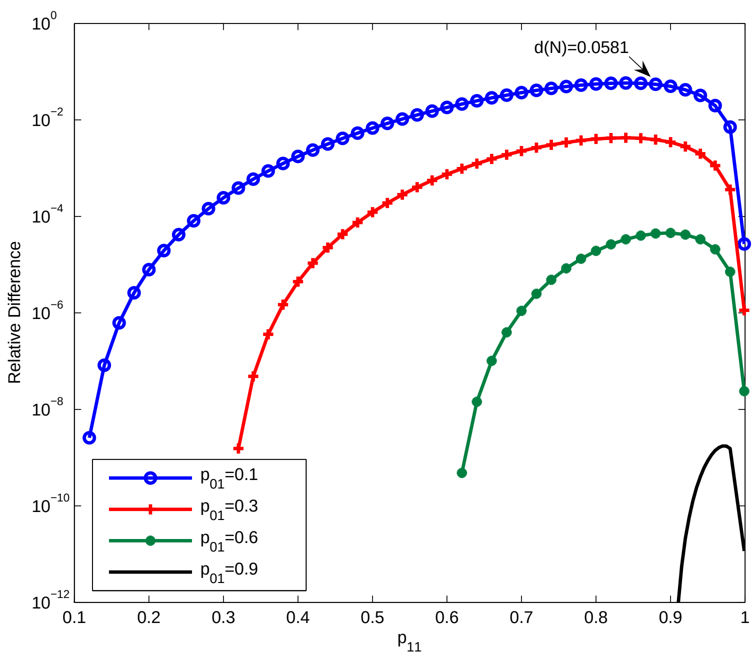
$$H = (1 - \omega_o)(\frac{1}{2-p_{01}} - \frac{p_{01}(p_{11}-p_{01})^{2N-1}}{1-(p_{11}-p_{01})^2(1-p_{01})^2}).$$

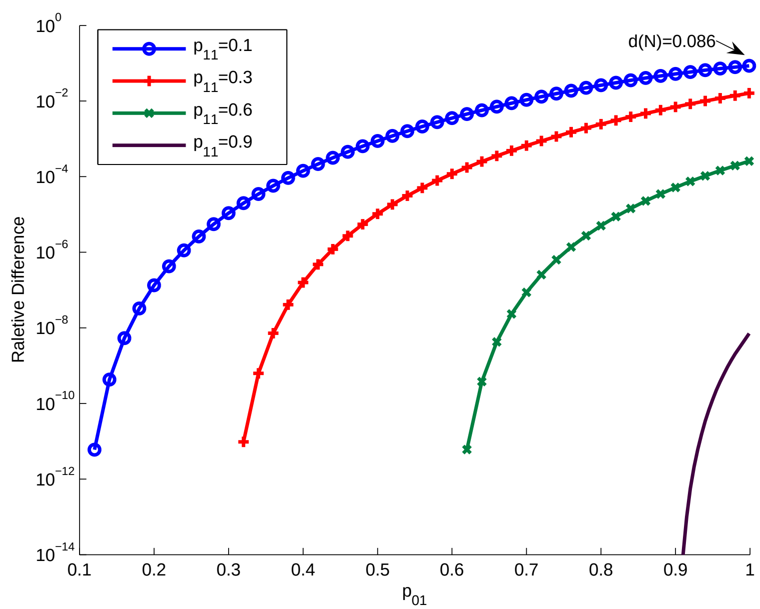
Link Throughput Limit for $N > 2$

Monotonicity

For both cases, the difference between the upper and lower bounds monotonically decreases with N; for $p_{11} \geq p_{01}$, the lower bound converges to the upper bound as $N \to \infty$.

Tightness of bounds ($N = 5$)

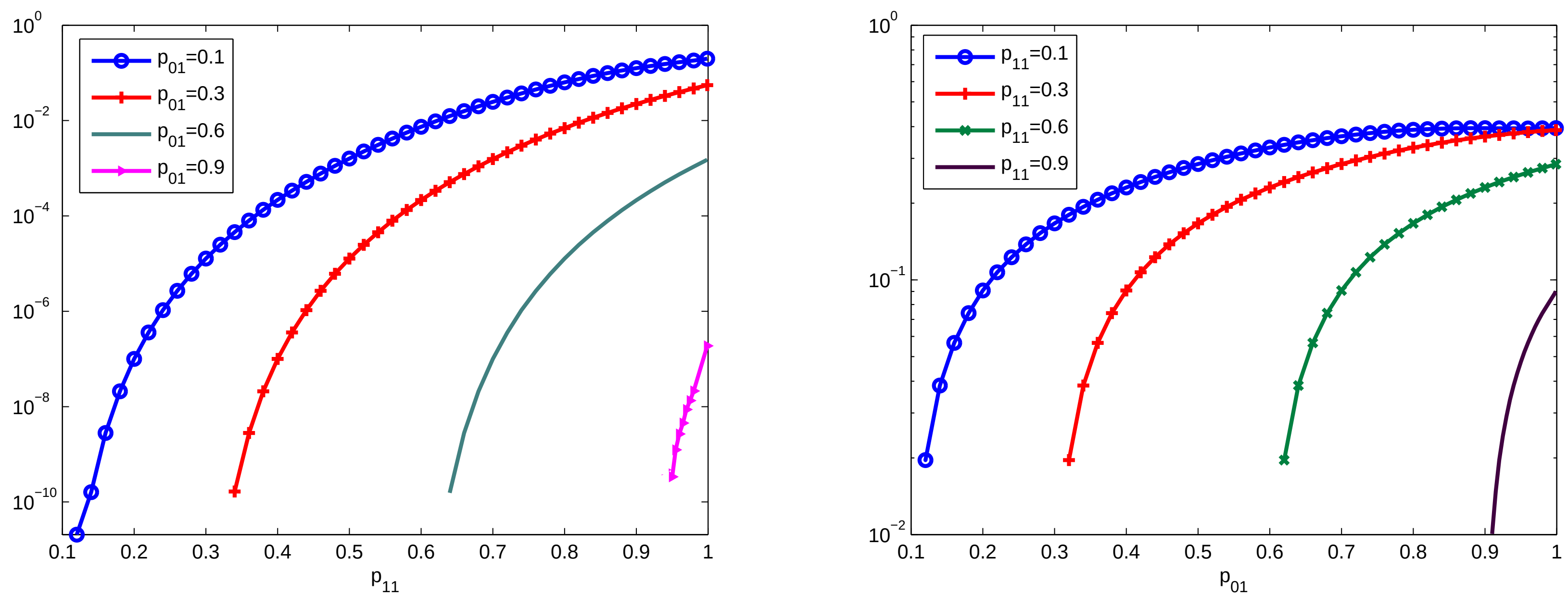




Link Throughput Limit for $N > 2$

The rate that the lower bound approaches to the upper bound

The following graph shows the ratio of the difference between lower and upper bounds when $N = 10$ to that when $N = 3$.



- The throughput of a multi-channel opportunistic system with single-channel sensing quickly saturates as N increases.

Conclusion

Main Results

- A Restless Multi-armed Bandit Formulation.
- Closed-form expressions for link throughput limit when $N = 2$.
- Tight upper and lower bounds of link throughput limit when $N > 2$.
- Link throughput limit quickly saturates as N increases.