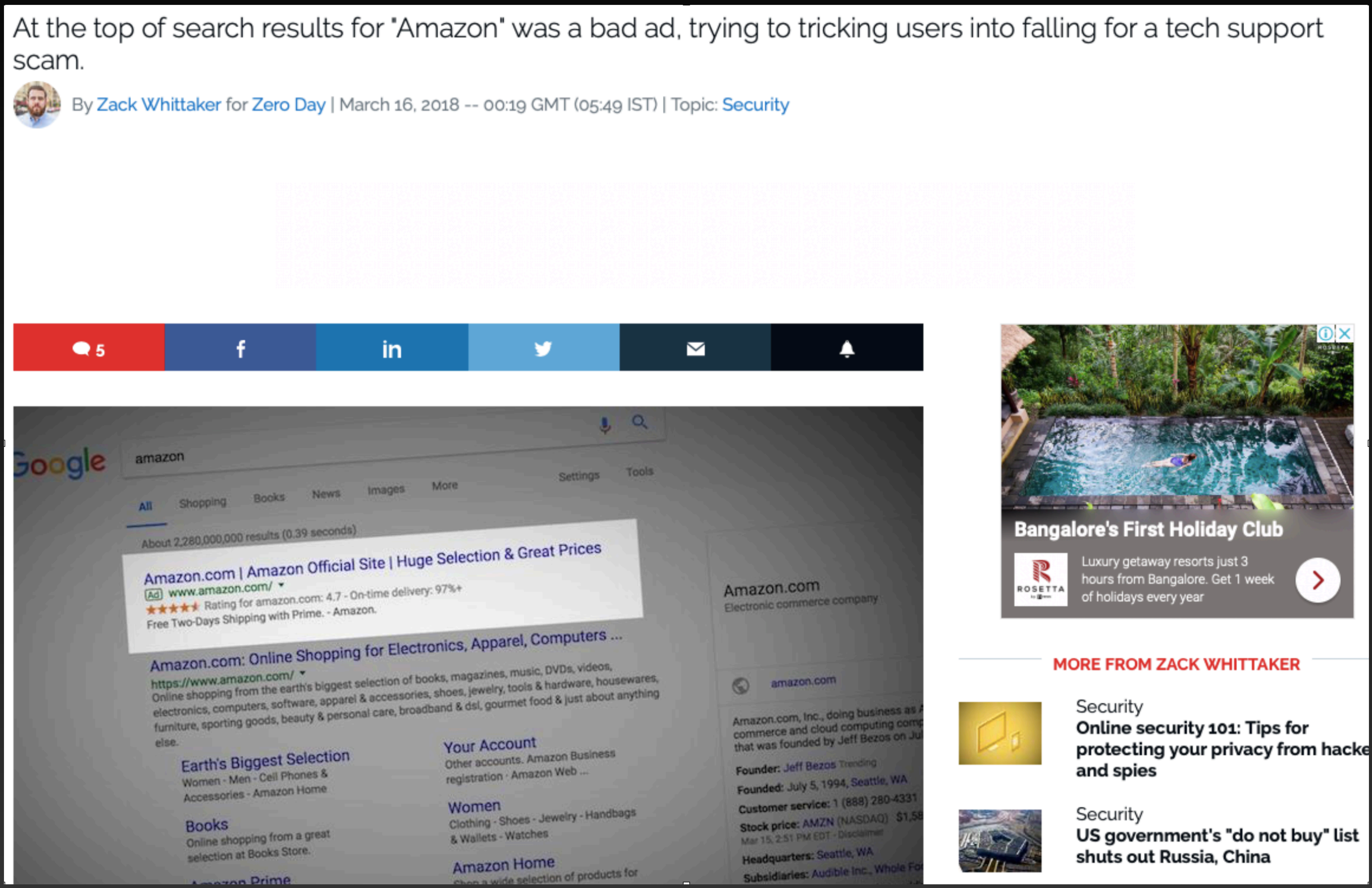
Odd Arm Identification in Multi-armed Bandits with Markov Observations

EECS Research Students Symposium - 2020

Karthik P. N.
Department of Electrical Communication Engineering

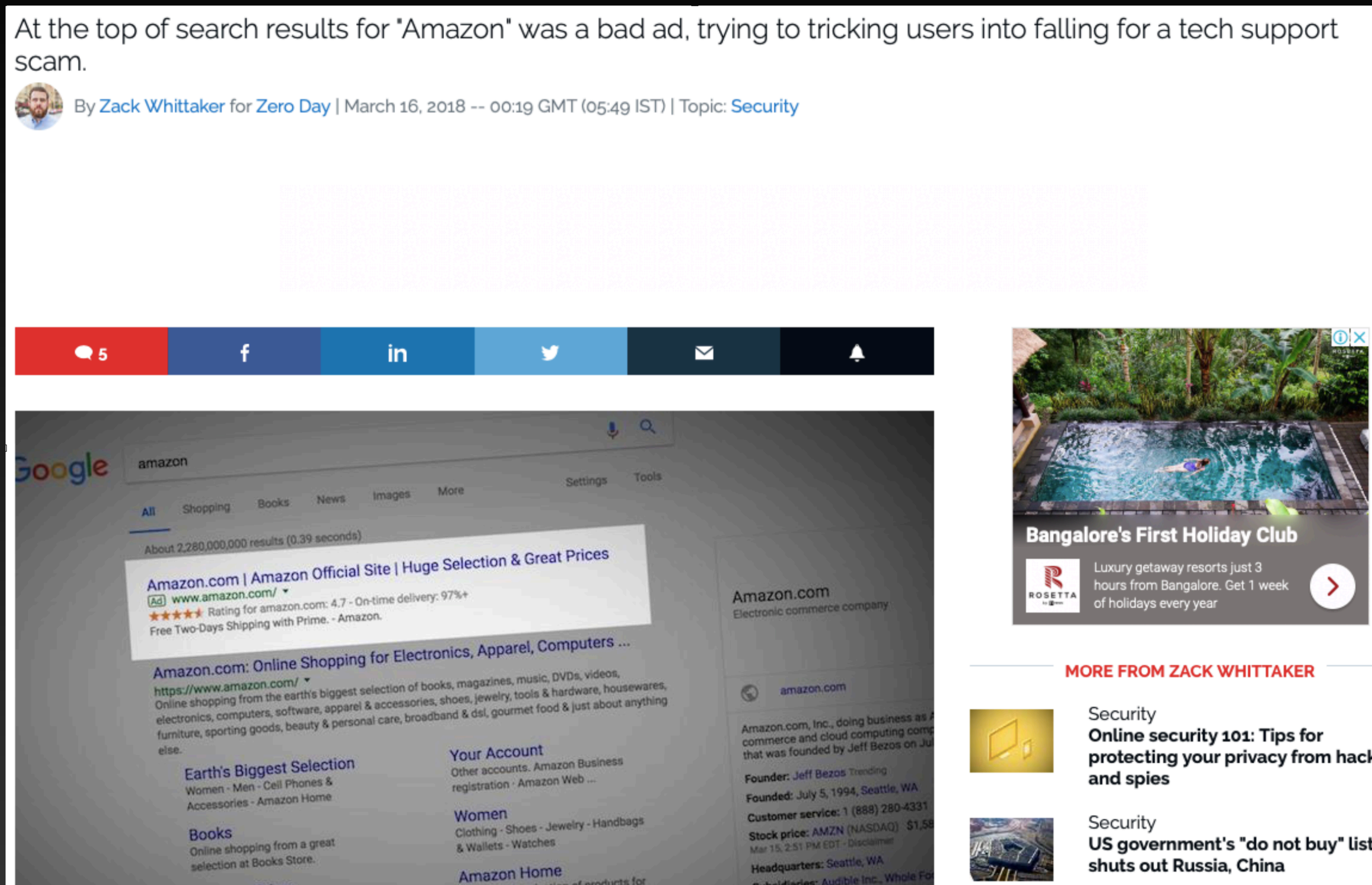
This work was supported by

Identification of a Malicious Ad



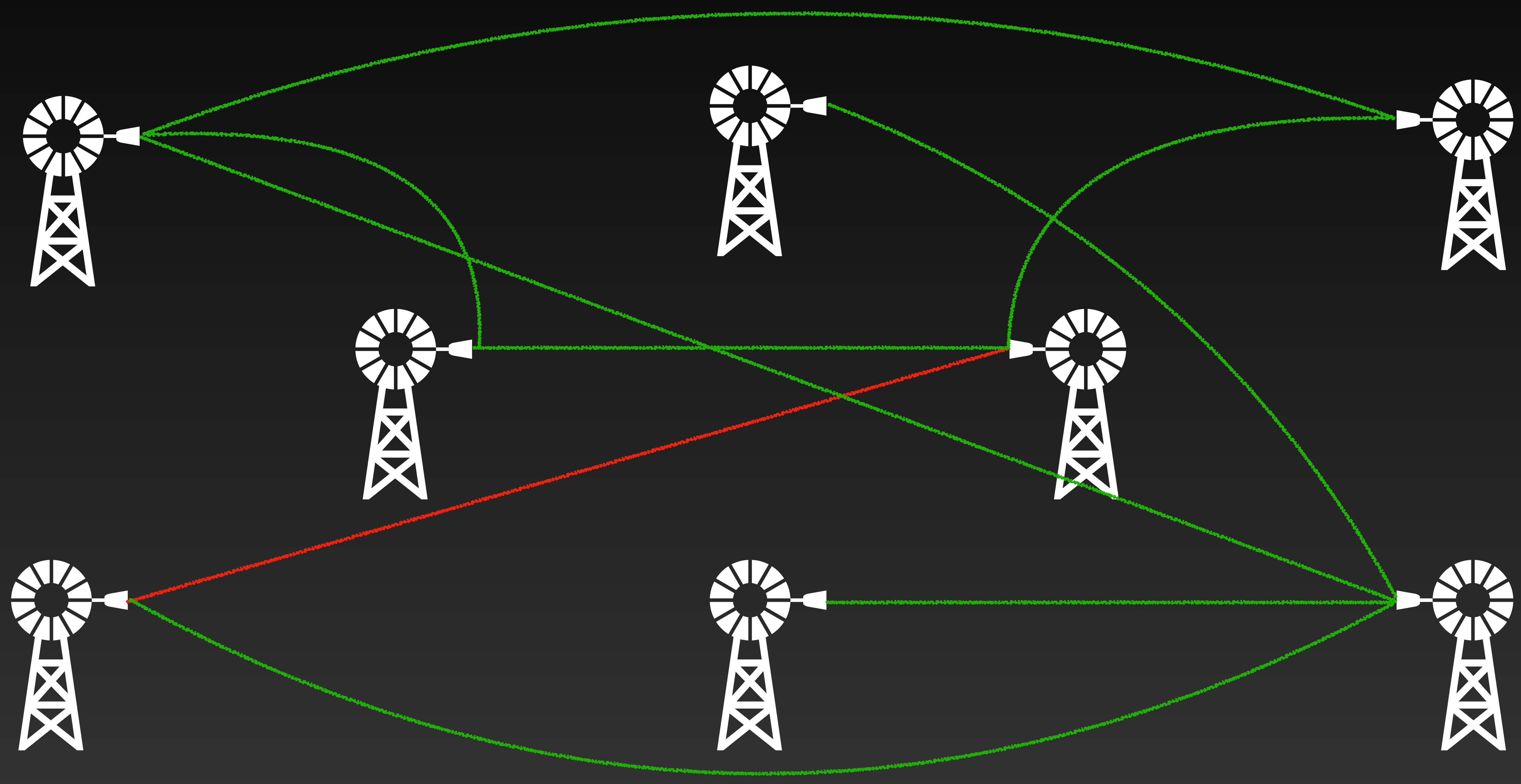
An online post bringing to light a malicious ad on Google.

Identification of a Malicious Ad



An online post bringing to light a malicious ad on Google.

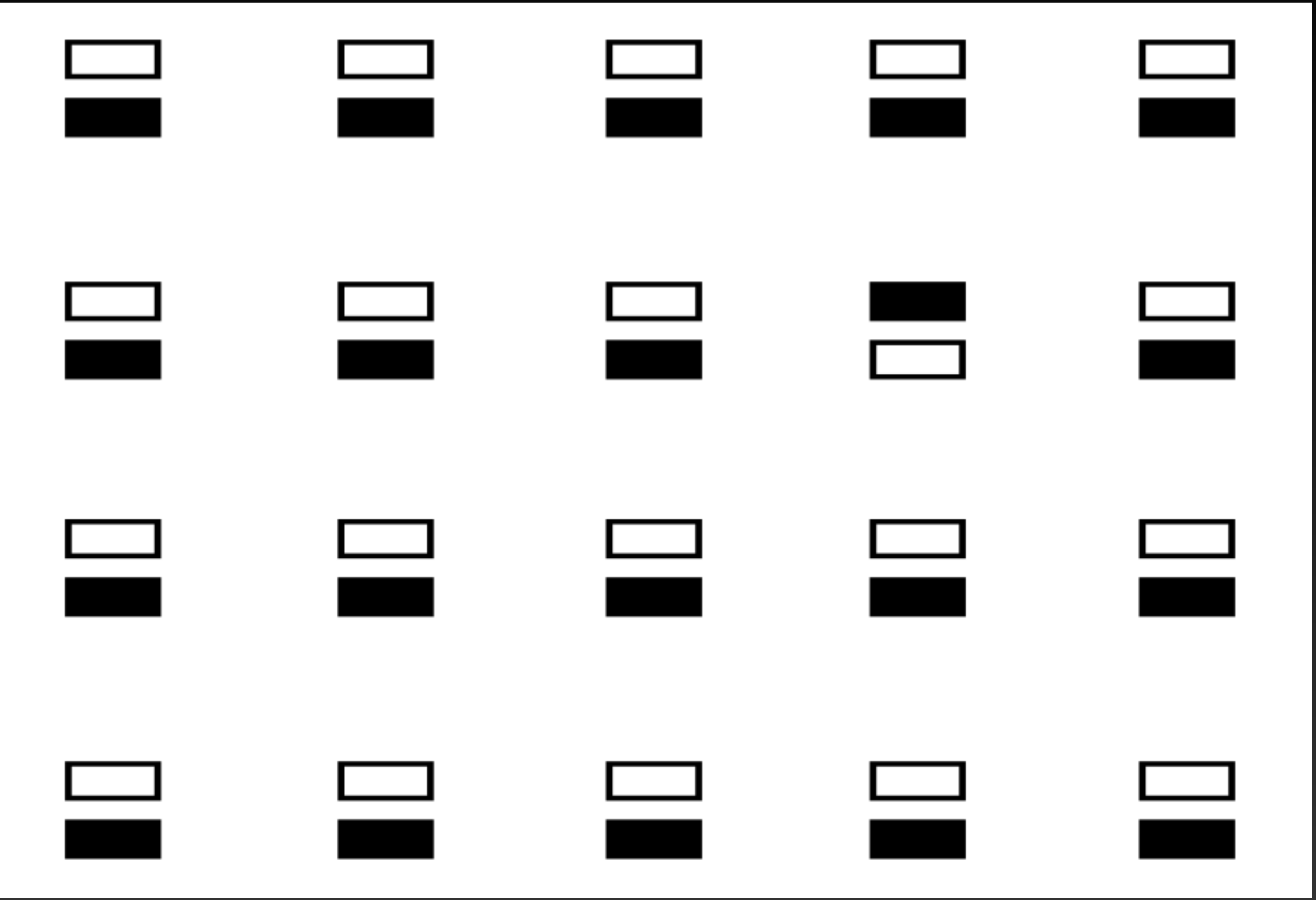
Identification of a Line in Outage in a Power Grid



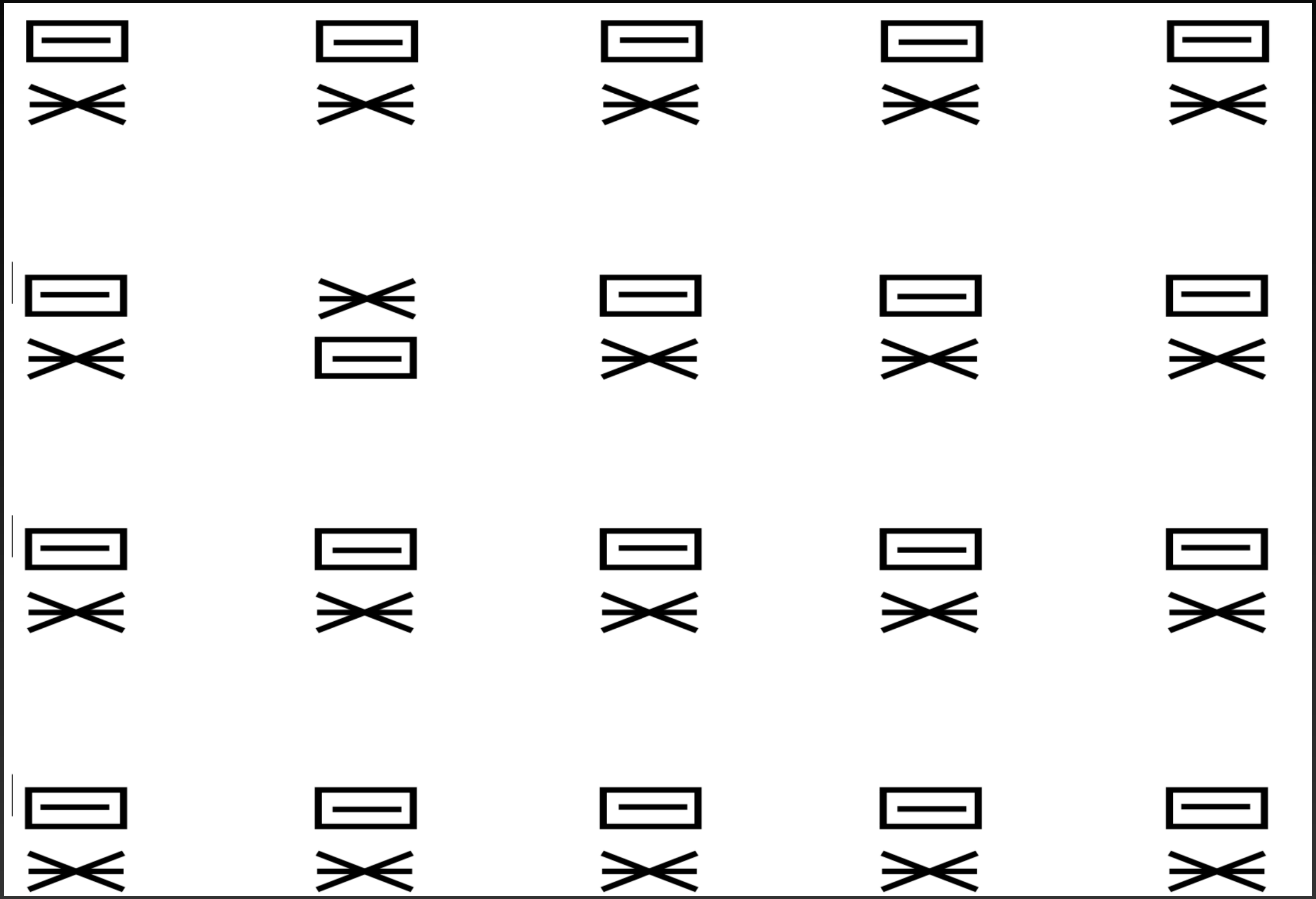
Line working fine

Line in outage

Odd Image Identification

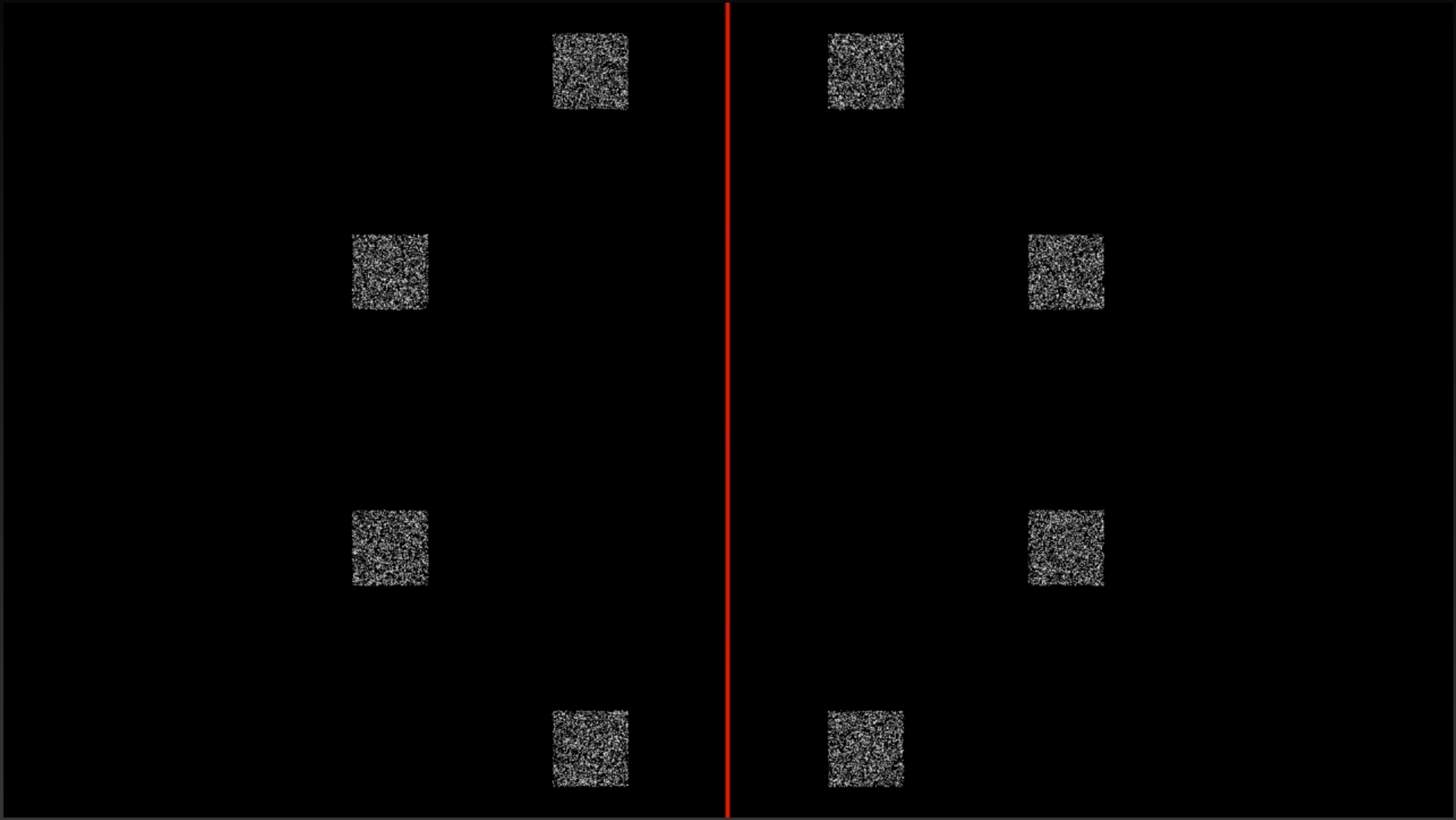


Odd Image Identification

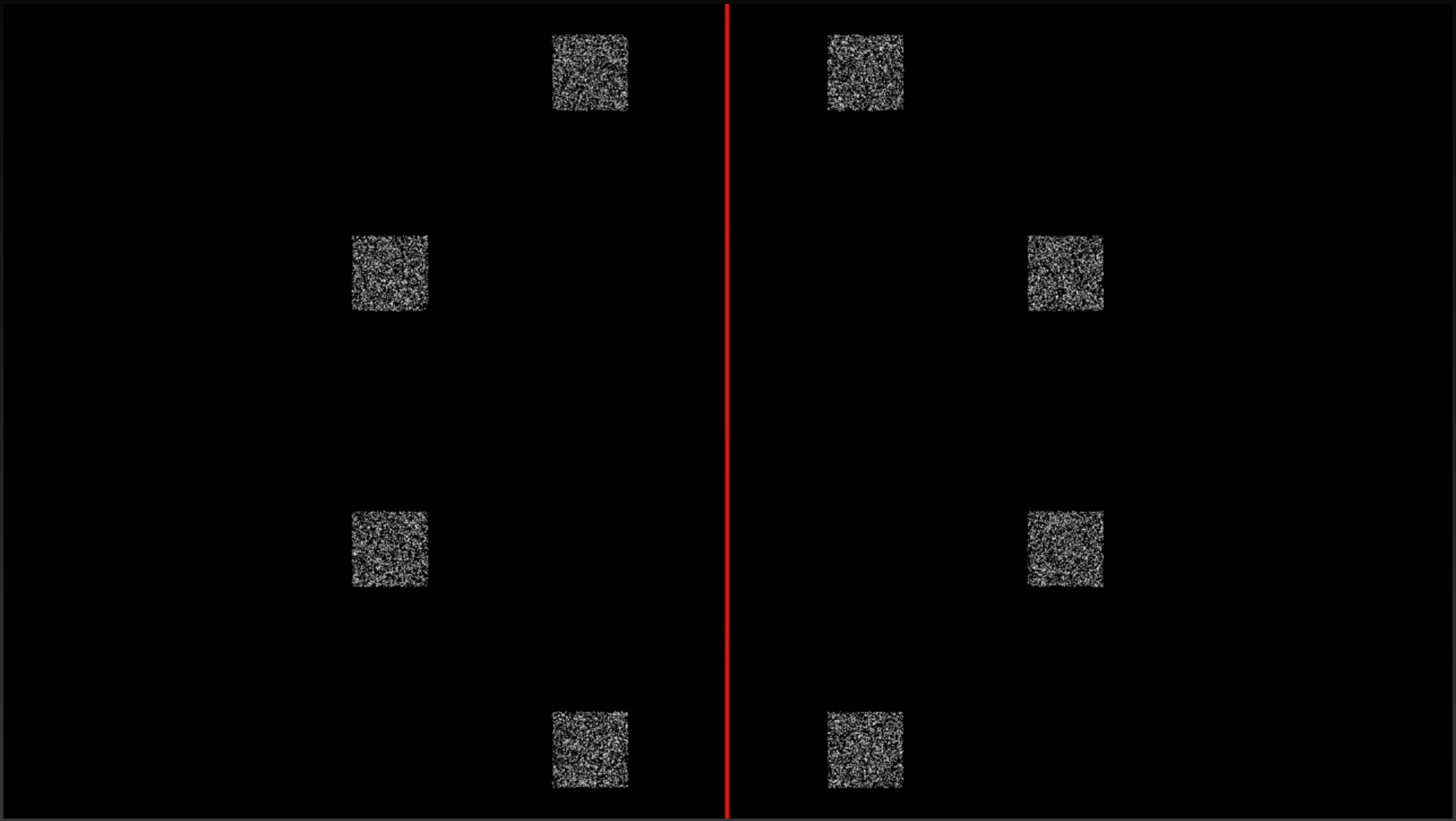


Identify the ‘odd’ image.

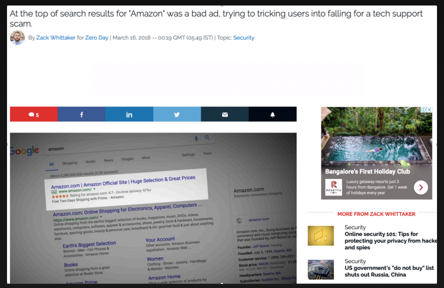
Odd Movie Identification

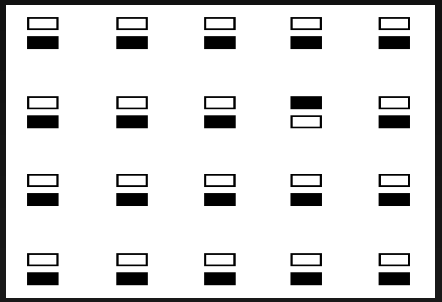


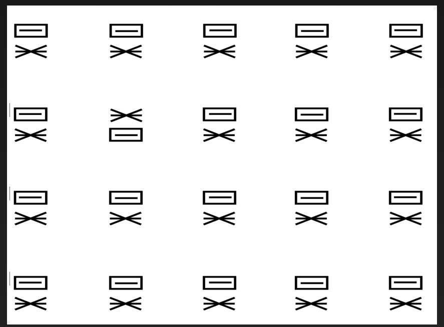
Odd Movie Identification

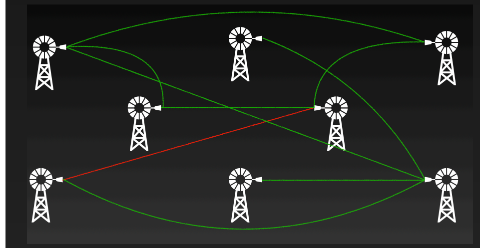


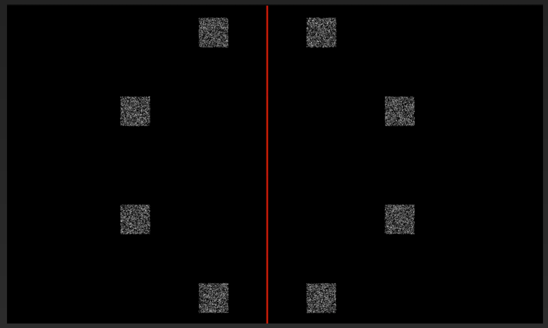
A Closer Look at the Examples





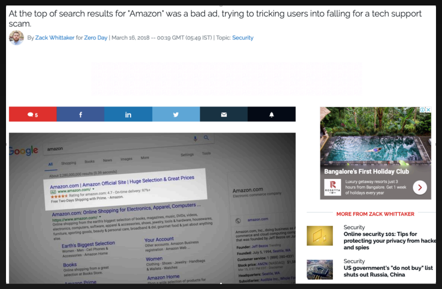


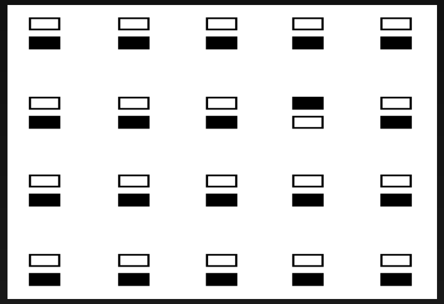


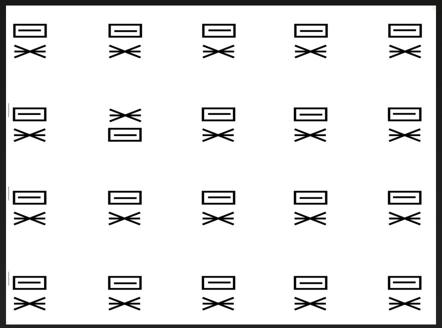


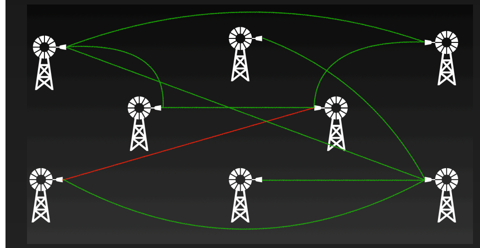
A Closer Look at the Examples

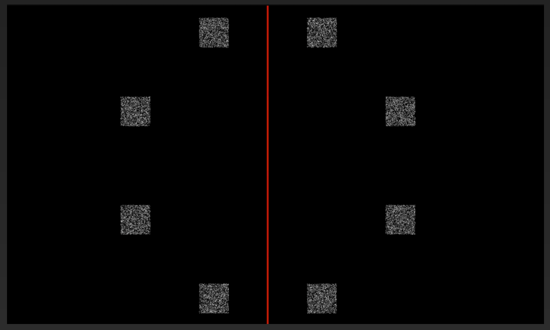
Anomalous or “odd” behaviour











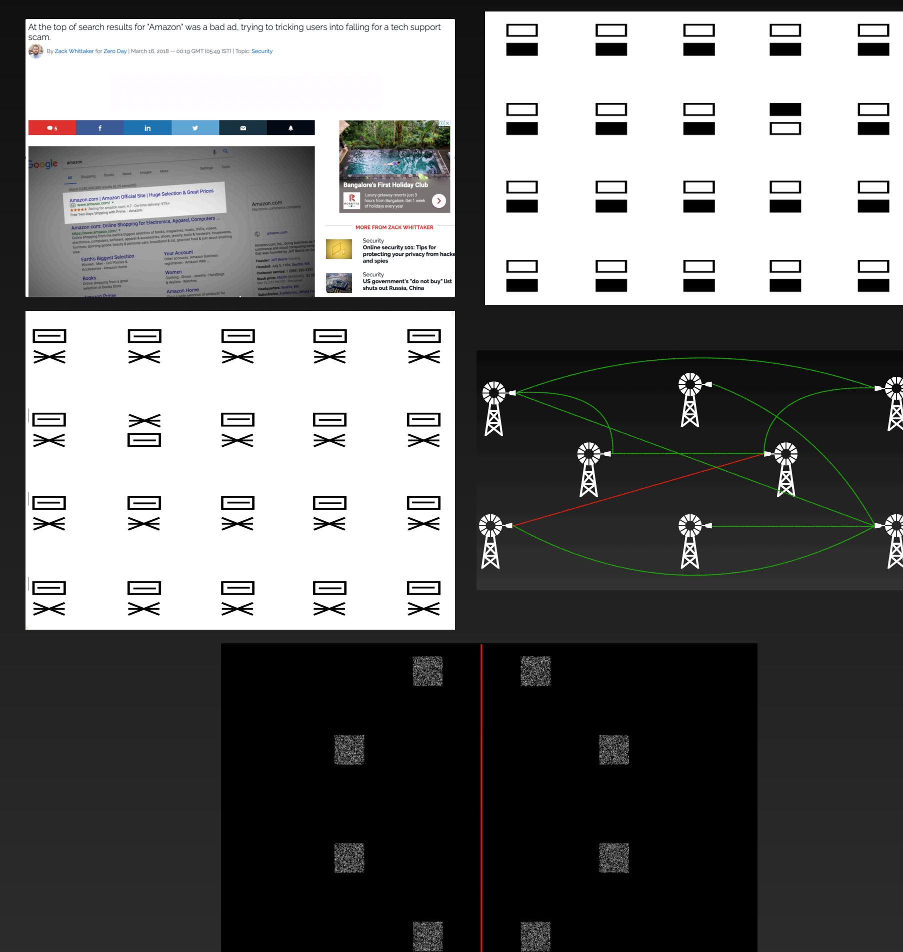
A Closer Look at the Examples

Anomalous or “odd” behaviour

Two quantities of interest

Time to identify the anomaly

Error in reporting the anomaly correctly



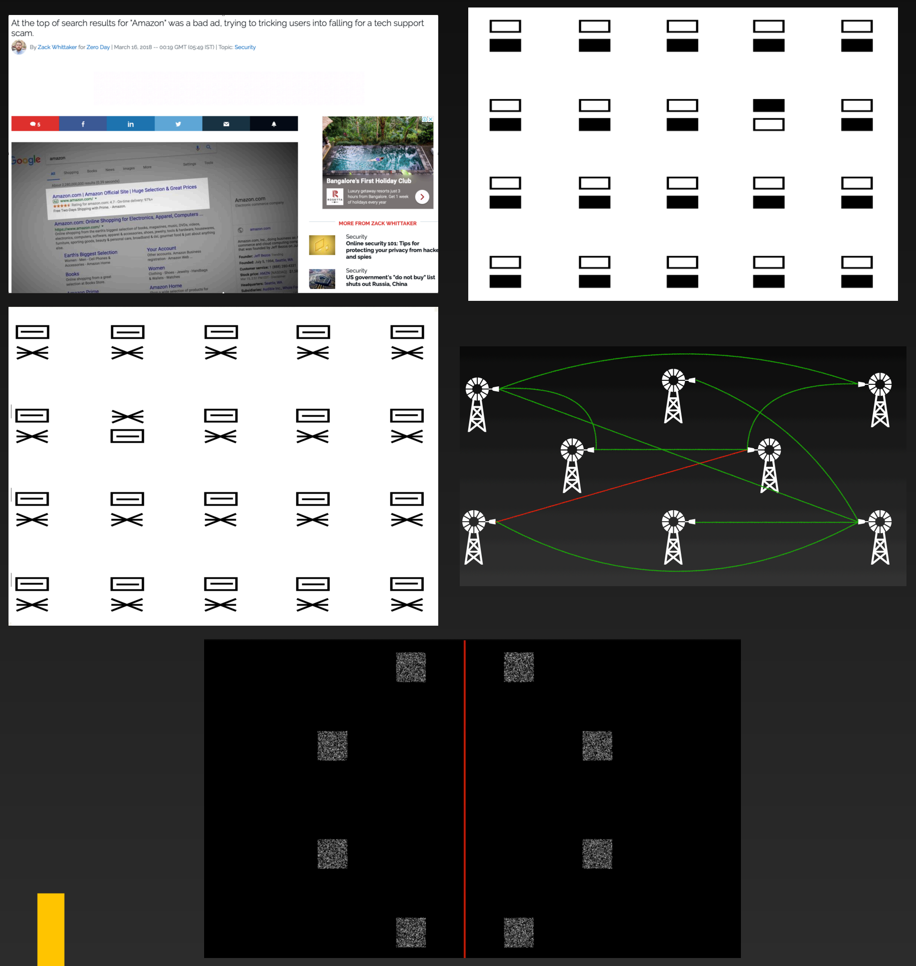
A Closer Look at the Examples

Anomalous or “odd” behaviour

Two quantities of interest

Time to identify the anomaly

Error in reporting the anomaly correctly



A study of the prior works reveals that

$$\text{Time to identify the anomaly} \approx \left(\log \frac{1}{\text{Error in reporting the anomaly correctly}} \right) \frac{1}{D^*}$$

D^* captures the hardness of the problem

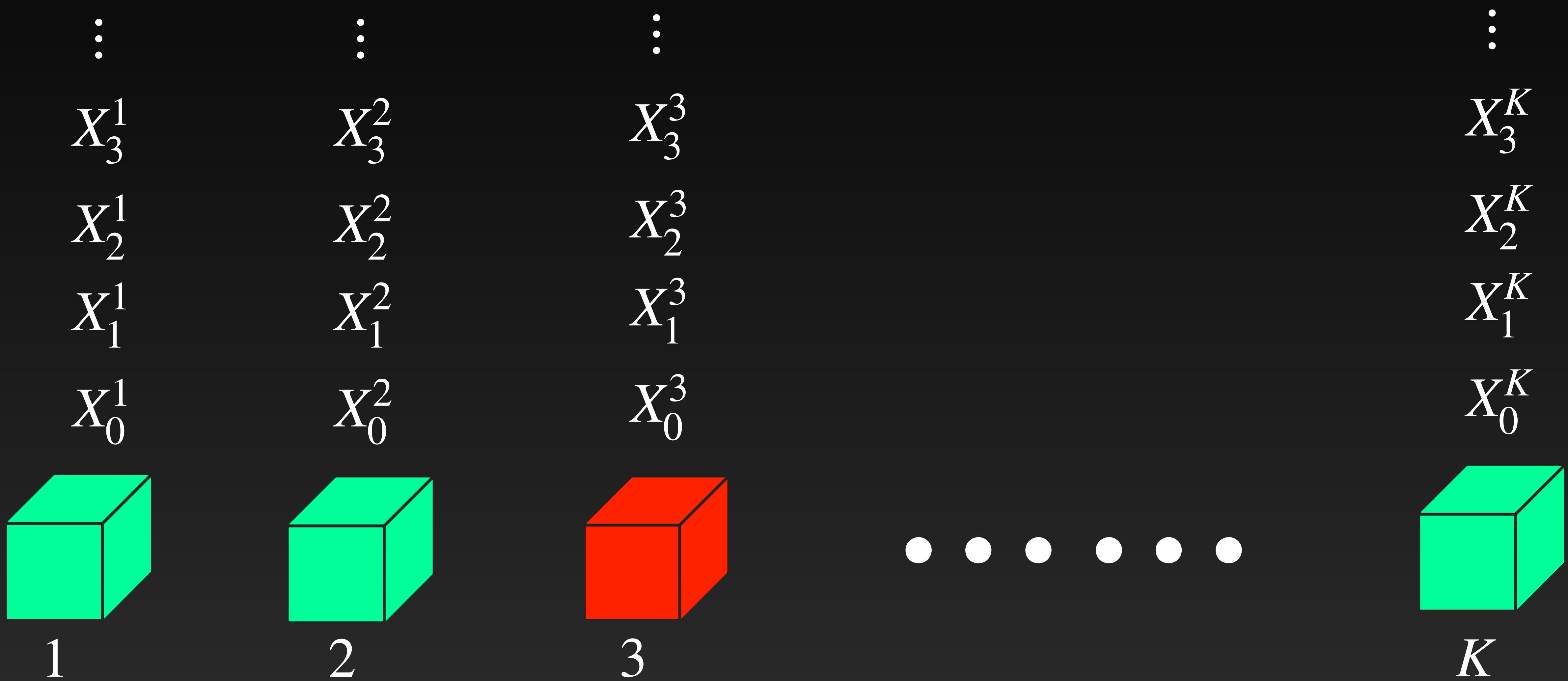
A study of the prior works reveals that

Exact in the asymptotic limit as error vanishes

$$\text{Time to identify the anomaly} \;\boxed{\approx}\; \left(\log \frac{1}{\text{Error in reporting the anomaly correctly}} \right) \frac{1}{D^*}$$

D^* captures the hardness of the problem

Multi-armed Bandits

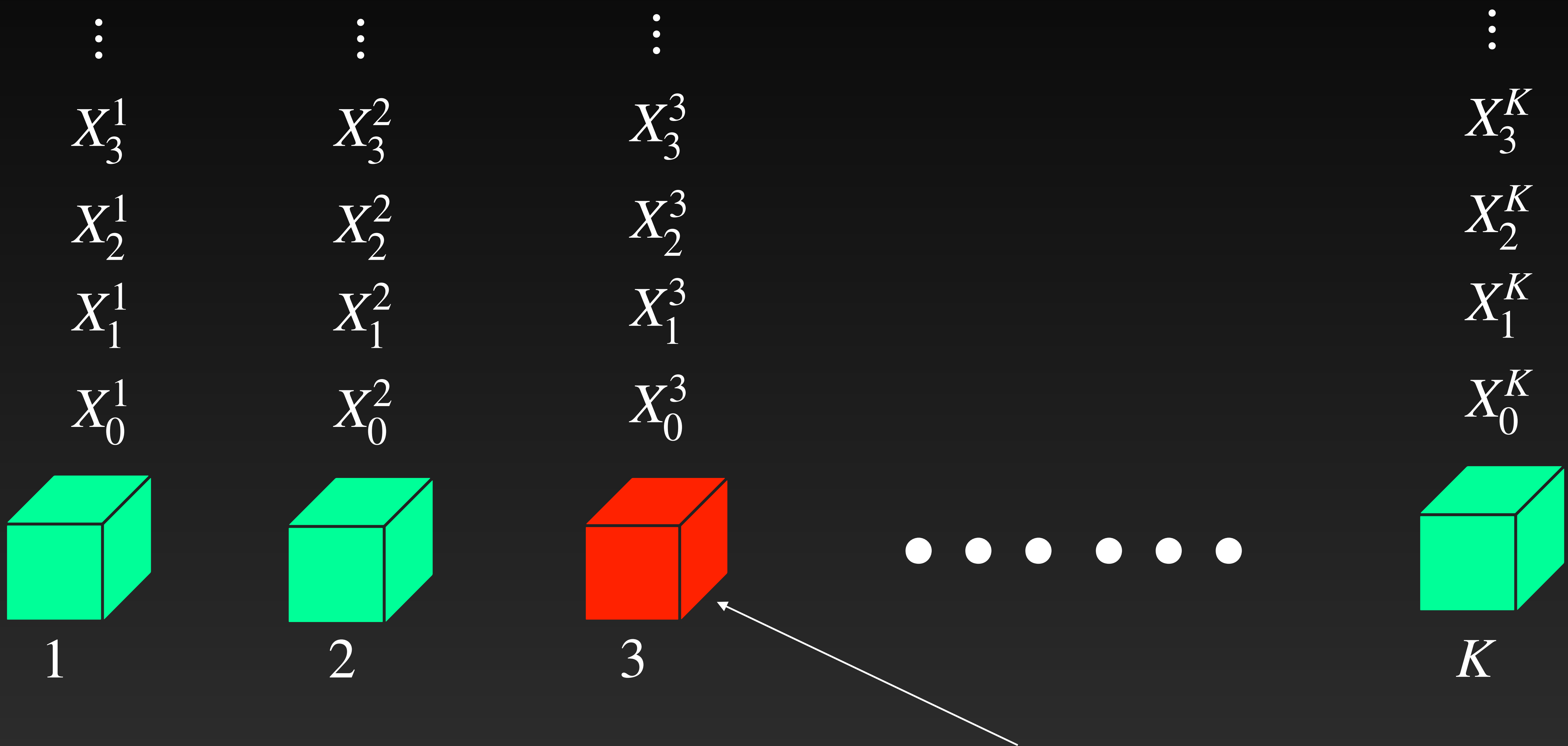


X_t^a

a ← Arm index

t ← Time (starts from 0)

Multi-armed Bandits

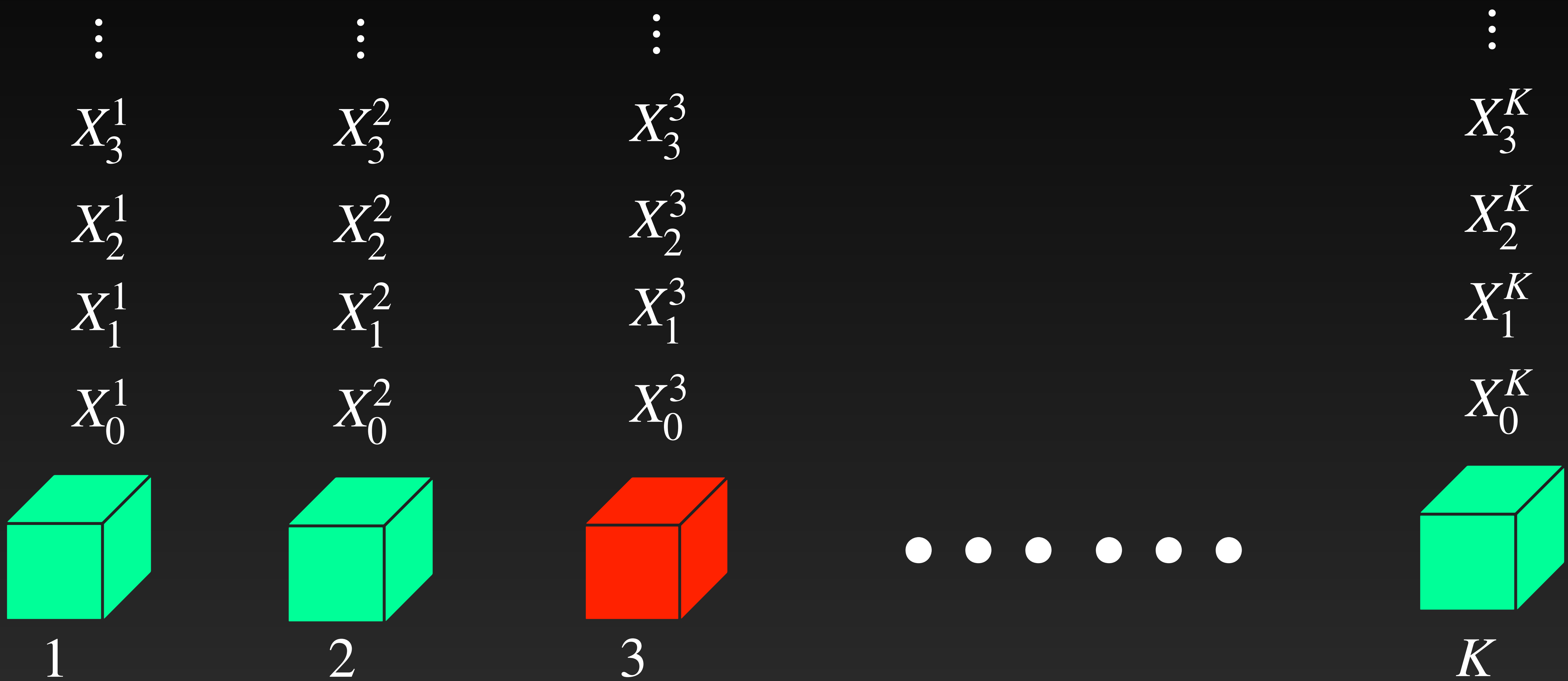


X_t^a

Arm index

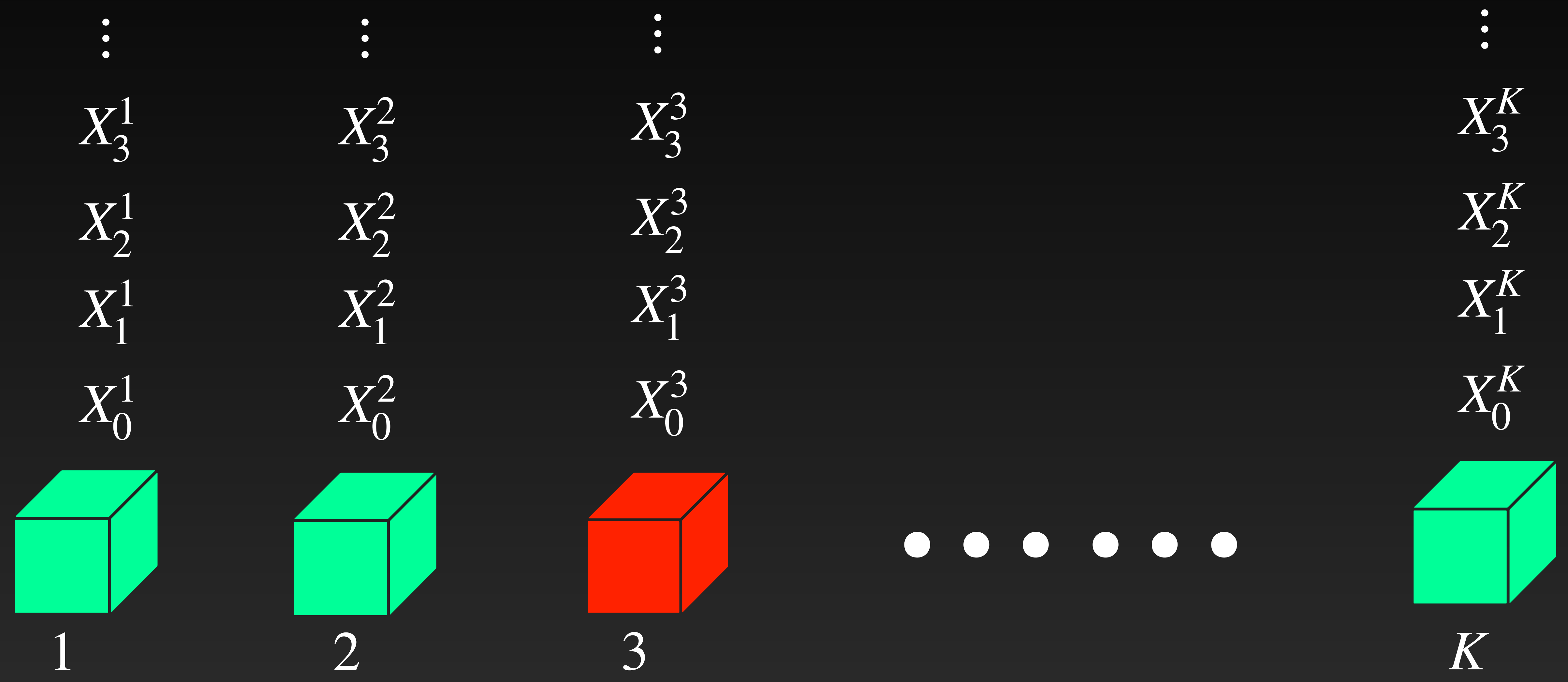
Time (starts from 0)

Multi-armed Bandits



Prior works: $\{X_t^a : t \geq 0\}$ is an iid process

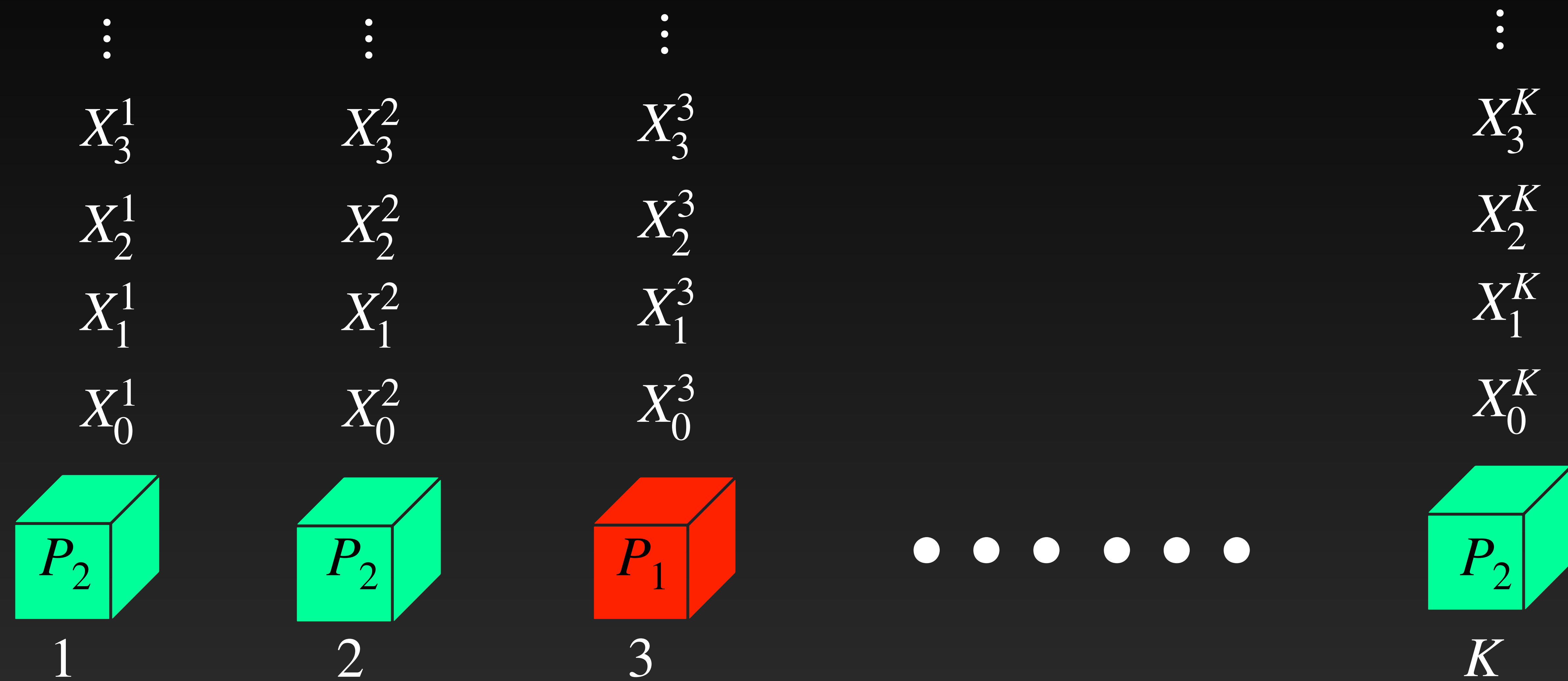
Multi-armed Bandits



Prior works: $\{X_t^a : t \geq 0\}$ is an iid process

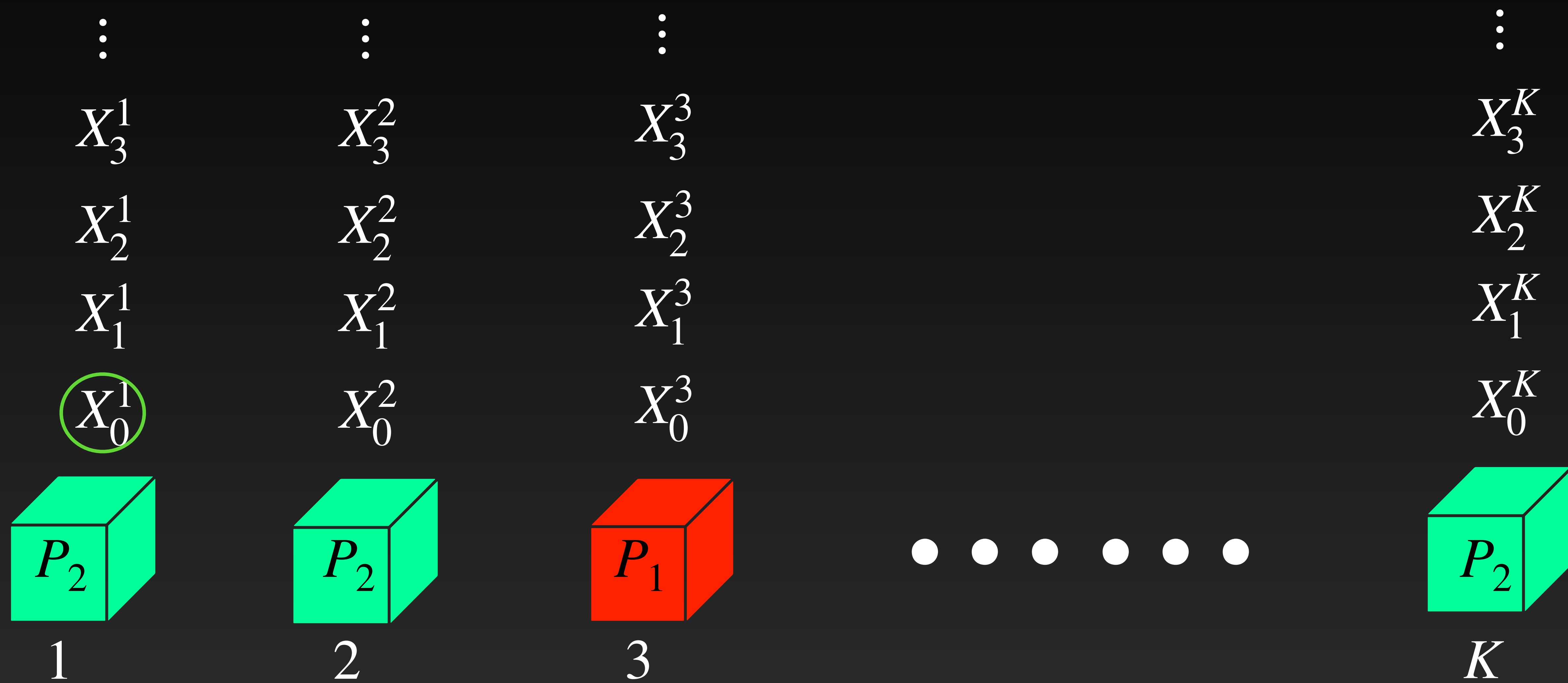
Our work: $\{X_t^a : t \geq 0\}$ is a Markov process

Markov Observations: Rested Case



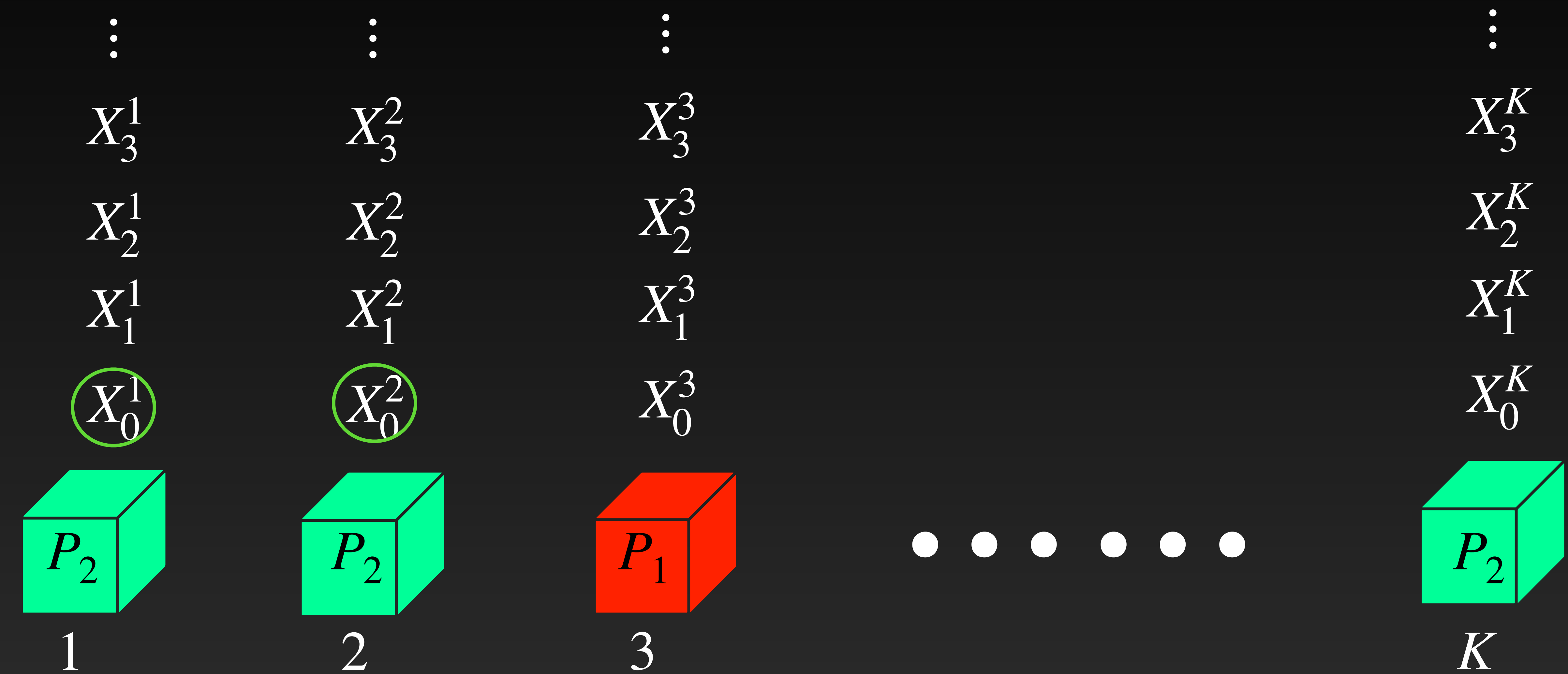
t	**0**	**1**	**2**	**3**	**4**	**5**
Arm	1	2	3	K	3	K
Obs						

Markov Observations: Rested Case



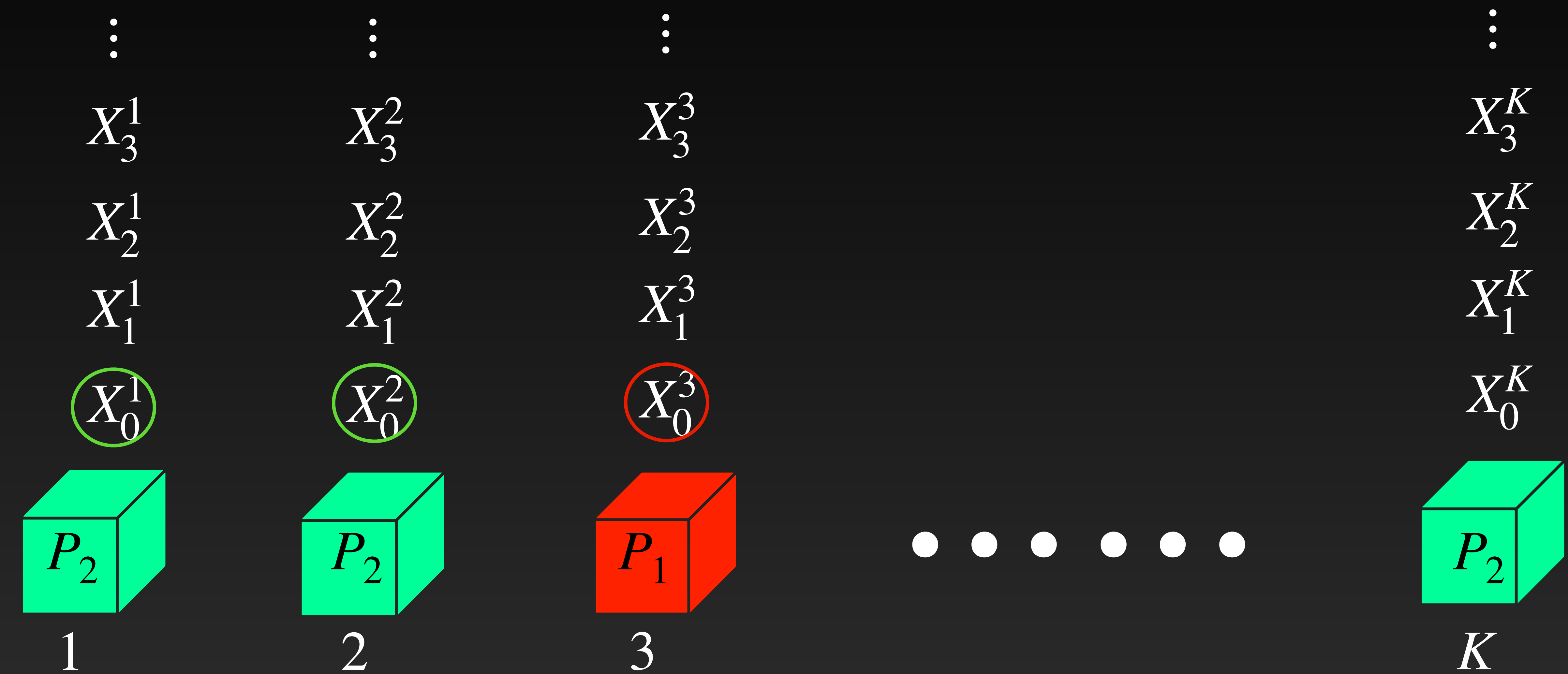
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1					

Markov Observations: Rested Case



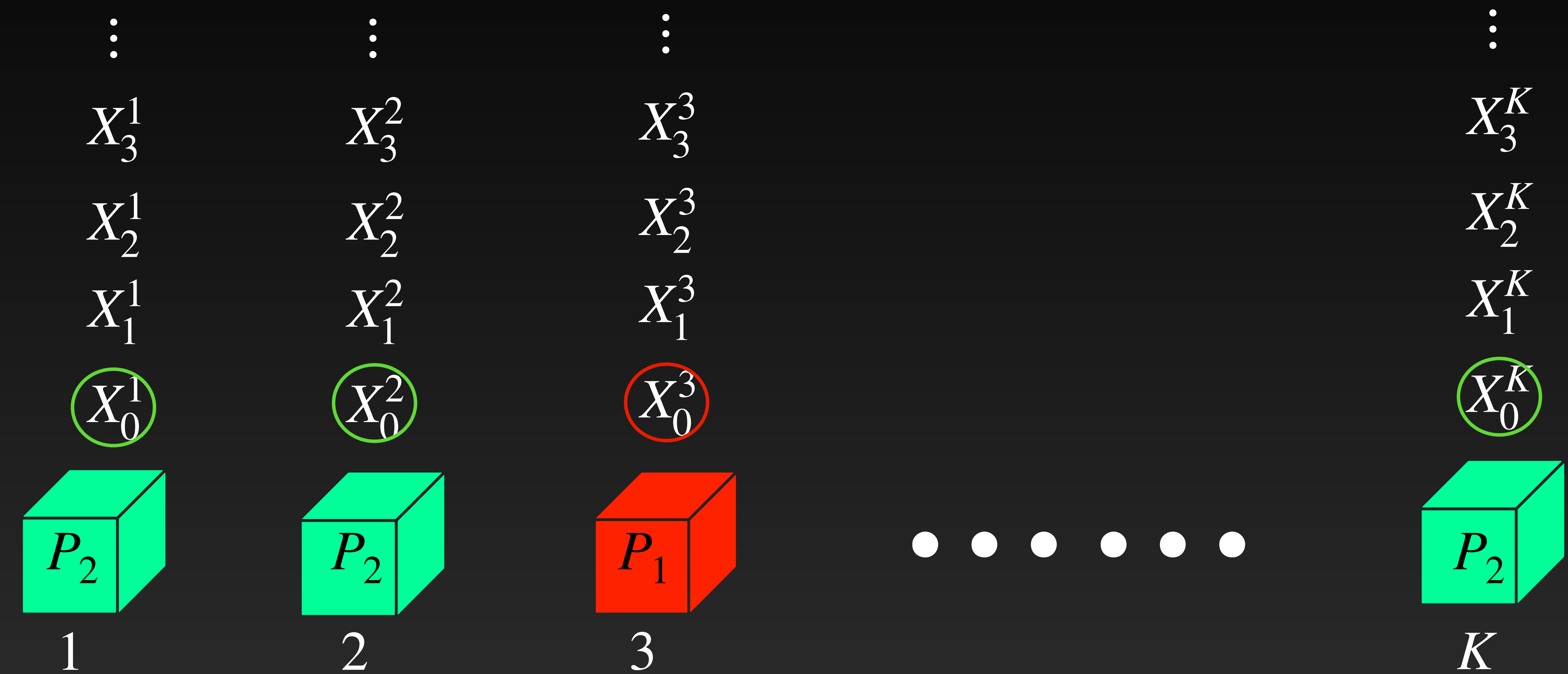
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2				

Markov Observations: Rested Case



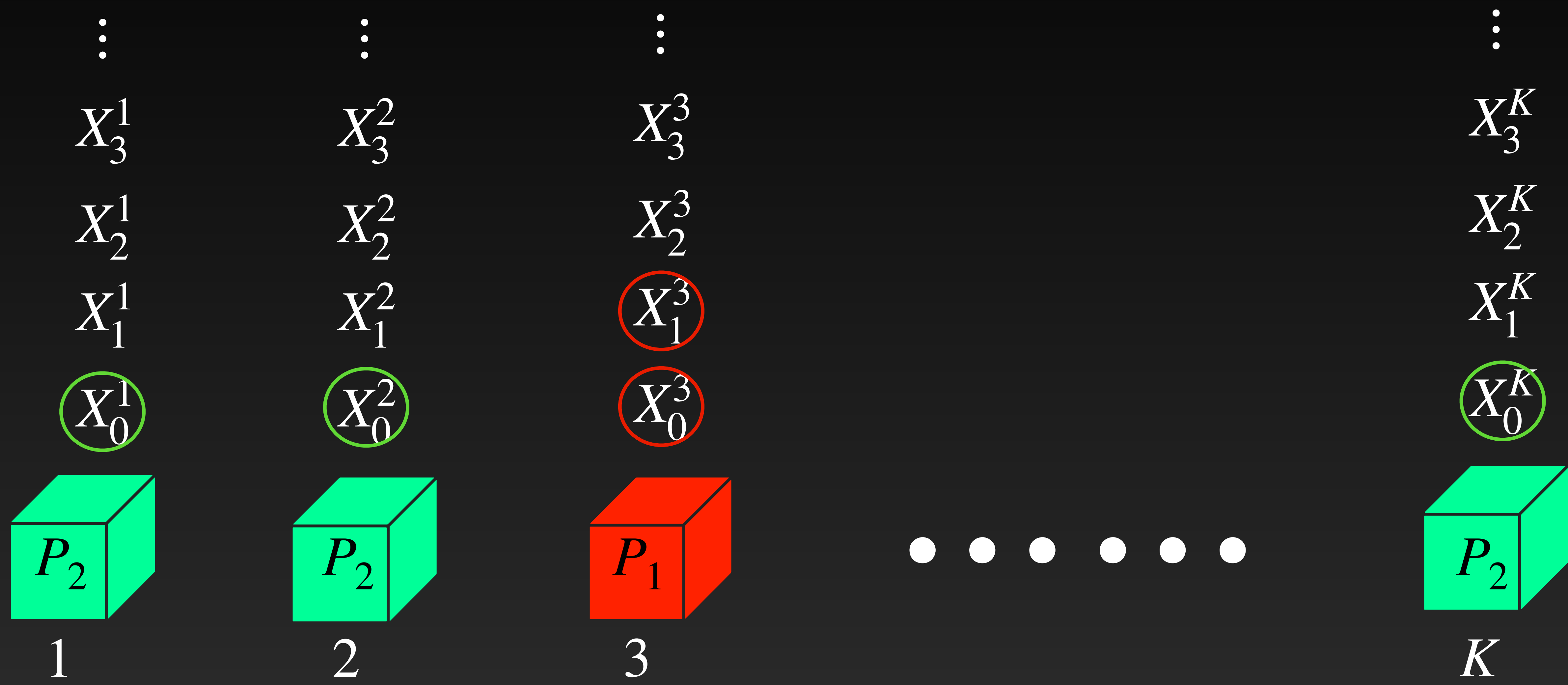
t	**0**	**1**	**2**	**3**	**4**	**5**
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3			

Markov Observations: Rested Case



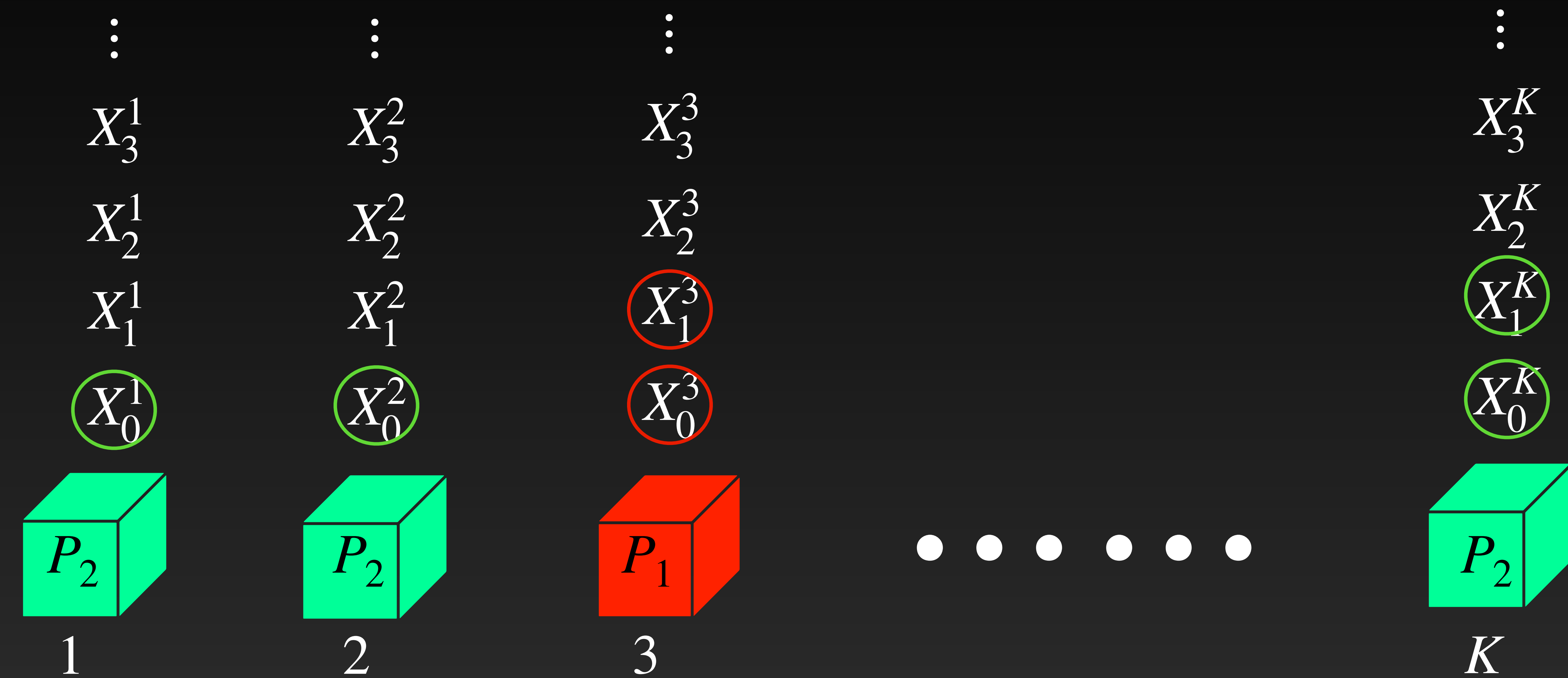
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3	X_0^K		

Markov Observations: Rested Case



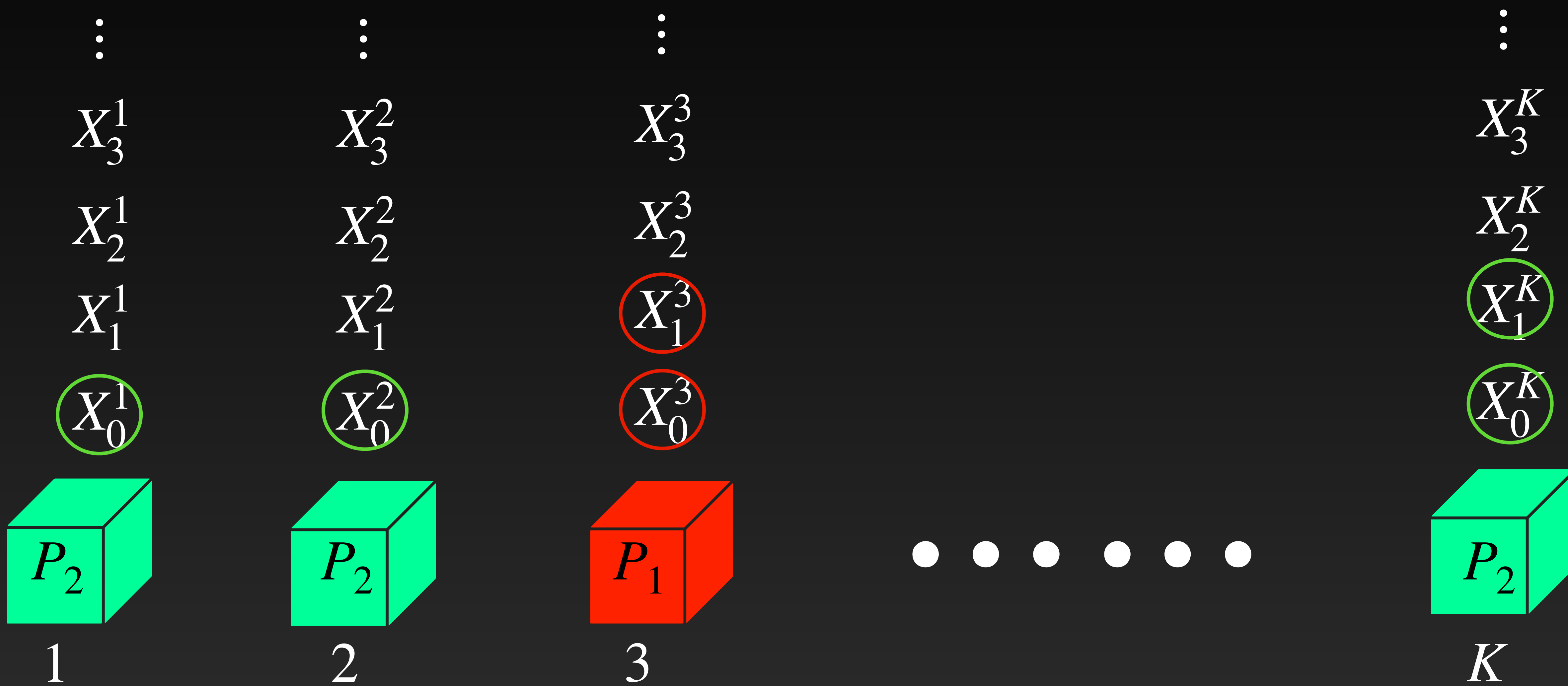
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3	X_0^K	X_1^3	

Markov Observations: Rested Case



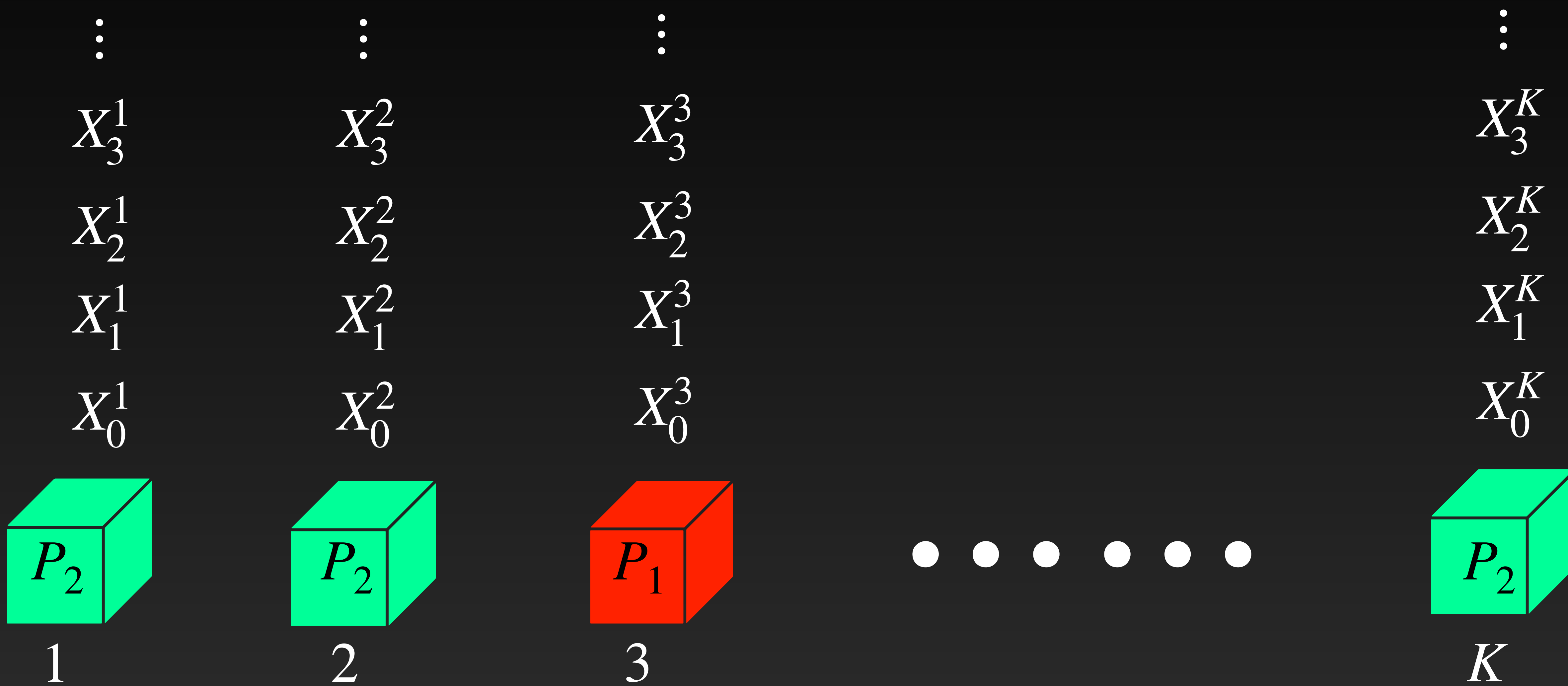
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3	X_0^K	X_1^3	X_1^K

Markov Observations: Rested Case



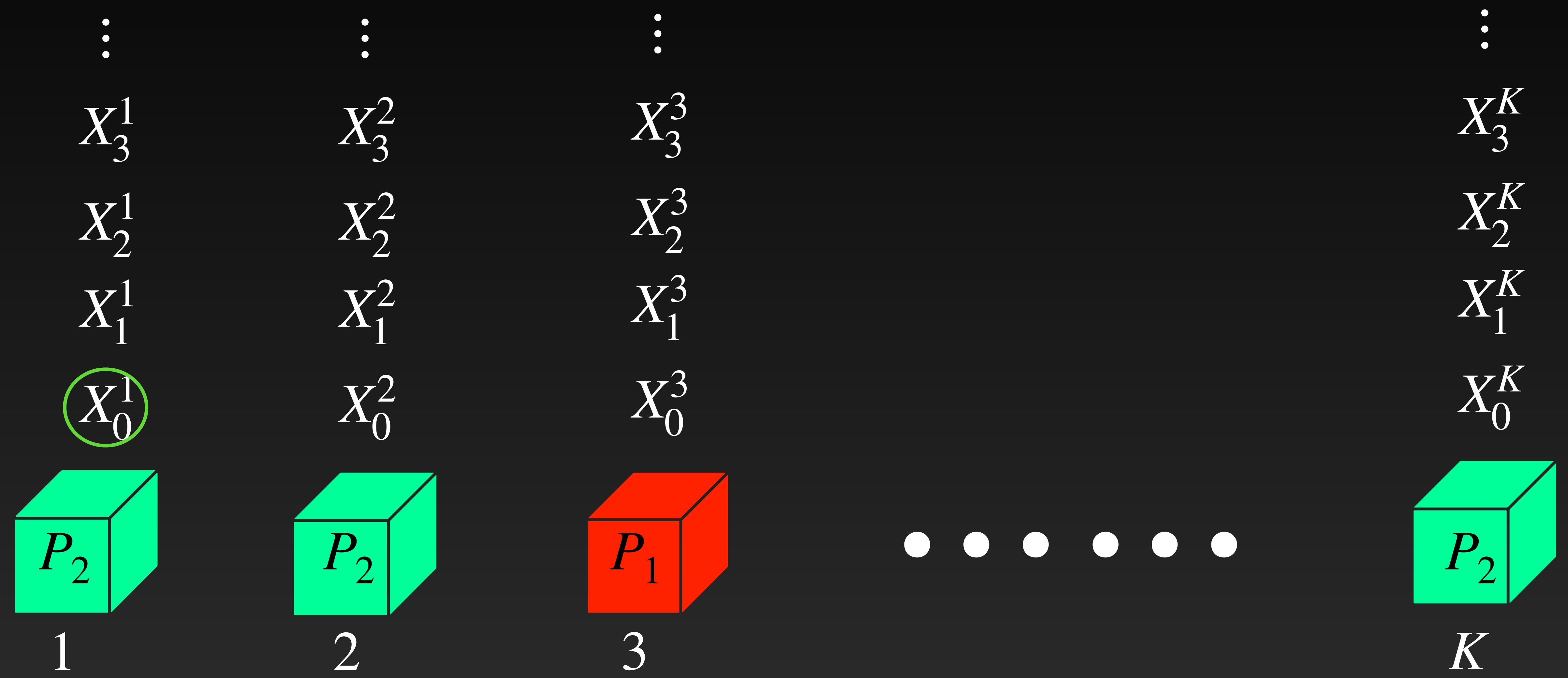
t	**0**	**1**	**2**	**3**	**4**	**5**
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3	X_0^K	X_1^3	X_1^K

Markov Observations: Restless Case



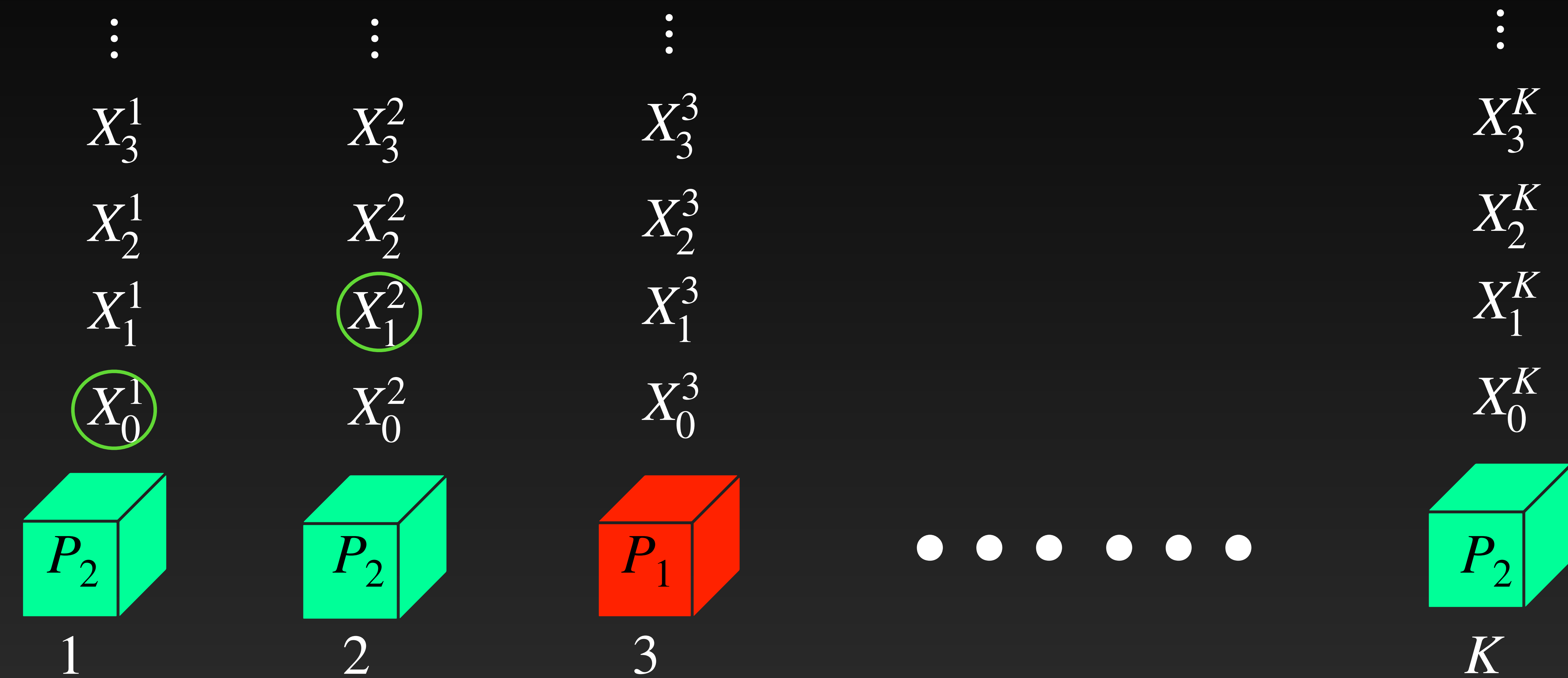
t	**0**	**1**	**2**	**3**	**4**	**5**
Arm	1	2	3	K	3	K
Obs						

Markov Observations: Restless Case



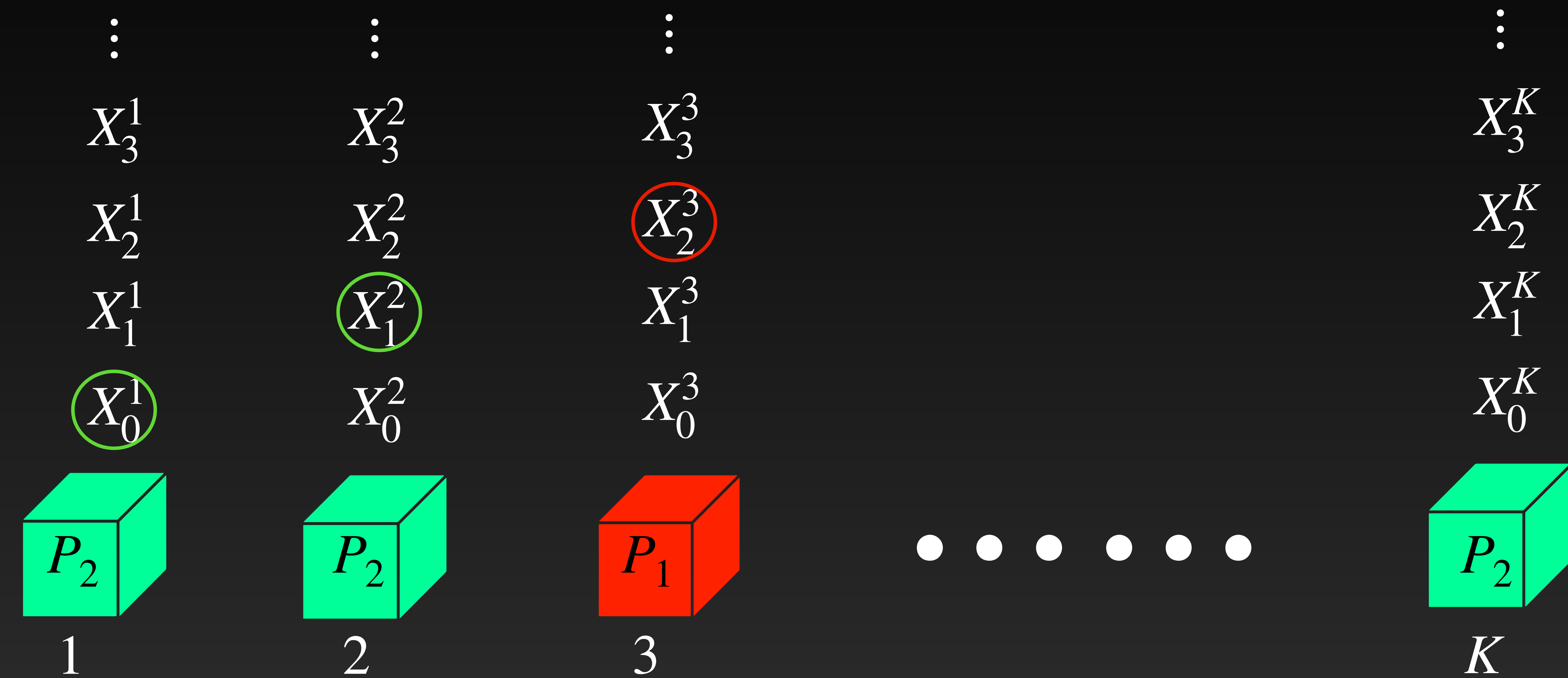
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1					

Markov Observations: Restless Case



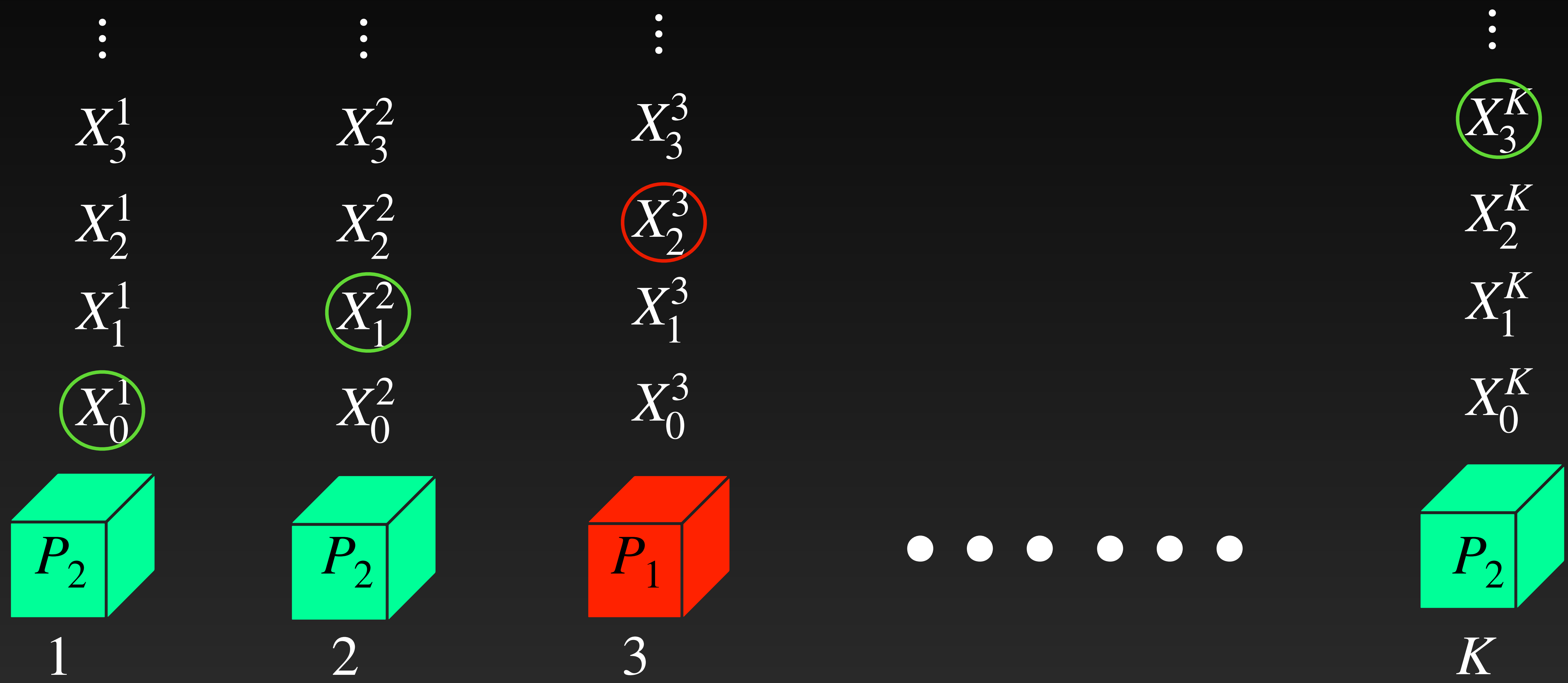
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_1^2				

Markov Observations: Restless Case



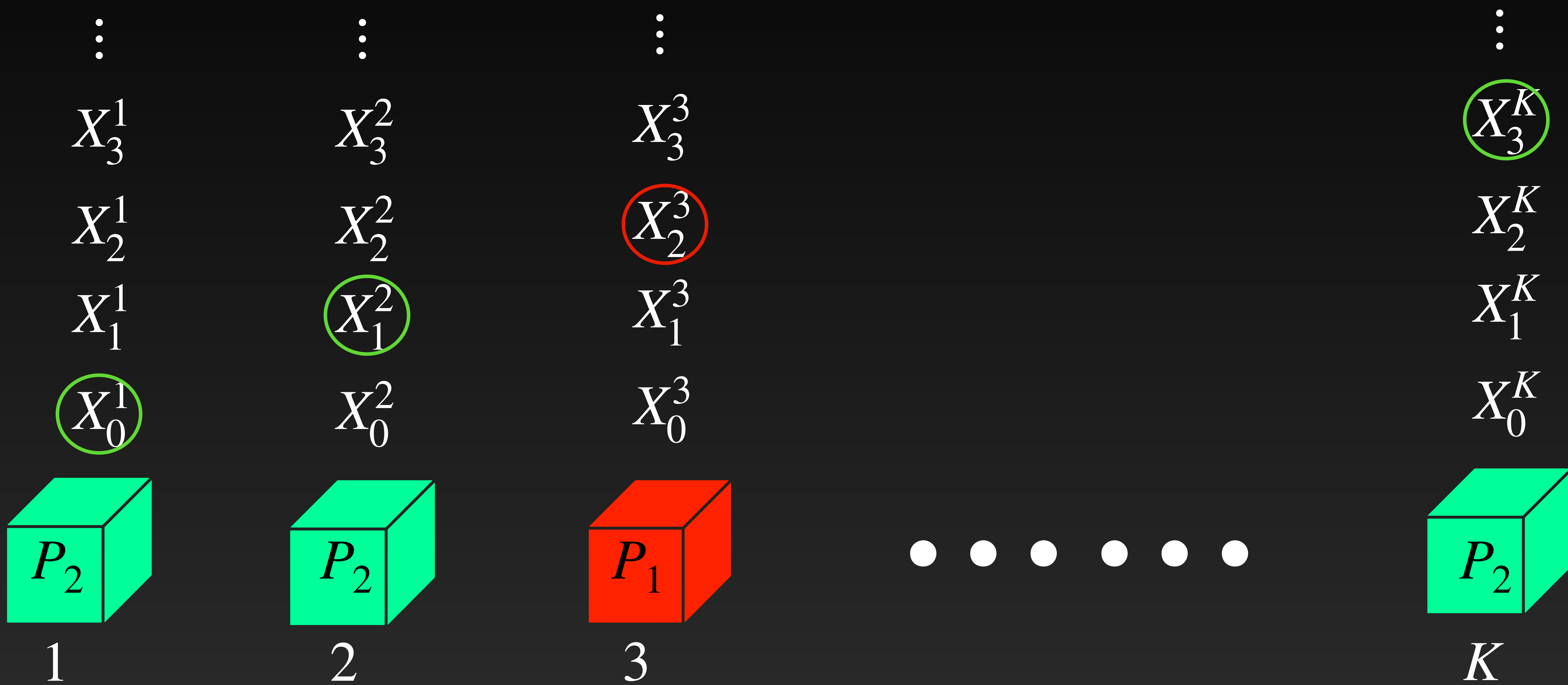
t	**0**	**1**	**2**	**3**	**4**	**5**
Arm	1	2	3	K	3	K
Obs	X_0^1	X_1^2	X_2^3			

Markov Observations: Restless Case



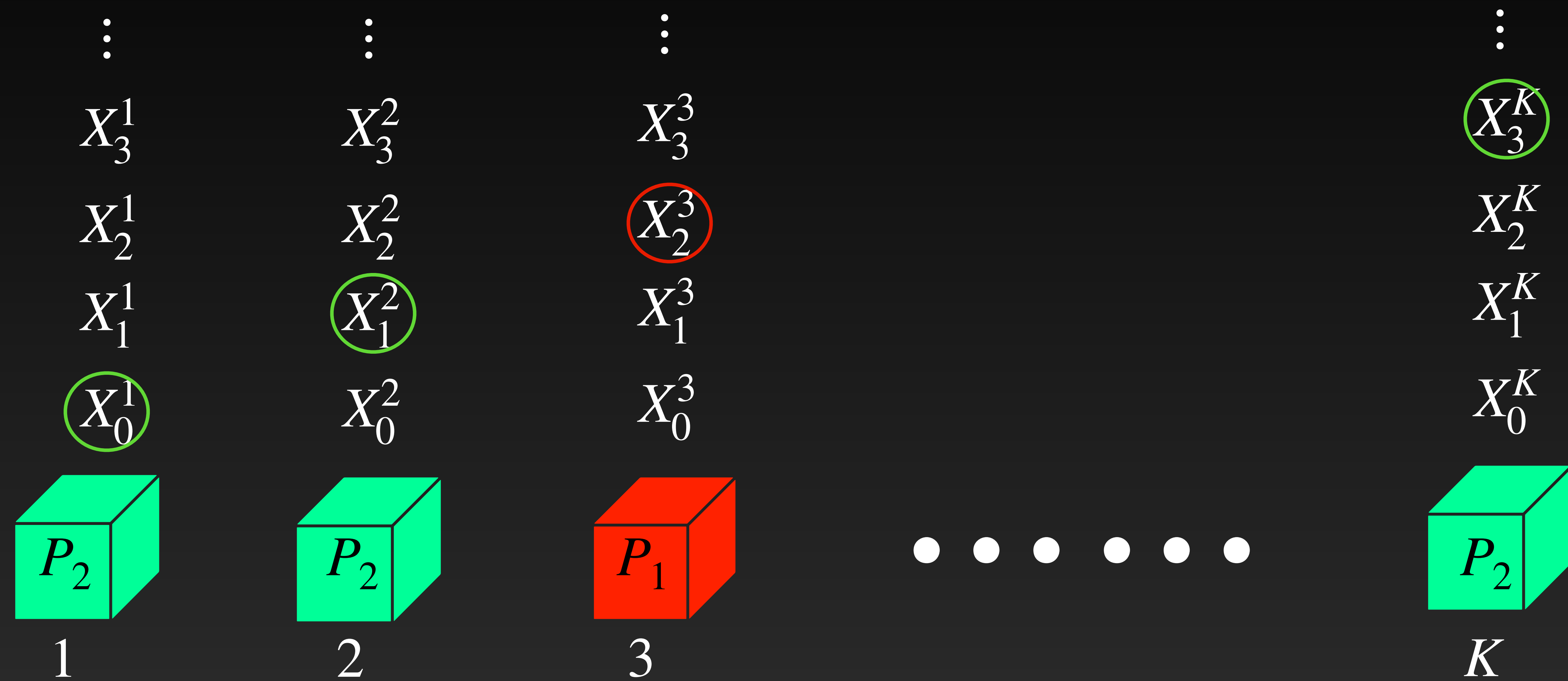
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_1^2	X_2^3	X_3^K		

Markov Observations: Restless Case



t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_1^2	X_2^3	X_3^K	X_4^3	X_5^K

Markov Observations: Restless Case



t	**0**	**1**	**2**	**3**	**4**	**5**
Arm	1	2	3	K	3	K
Obs	X_0^1	X_1^2	X_2^3	X_3^K	X_4^3	X_5^K

When the observations from each arm form a Markov process, does the following still hold?

Exact in the asymptotic limit as error vanishes

$$\text{Time to identify the anomaly} \;\boxed{\approx}\; \left(\log \frac{1}{\text{Error in reporting the anomaly correctly}} \right) \frac{1}{D^*}$$

When the observations from each arm form a Markov process, does the following still hold?

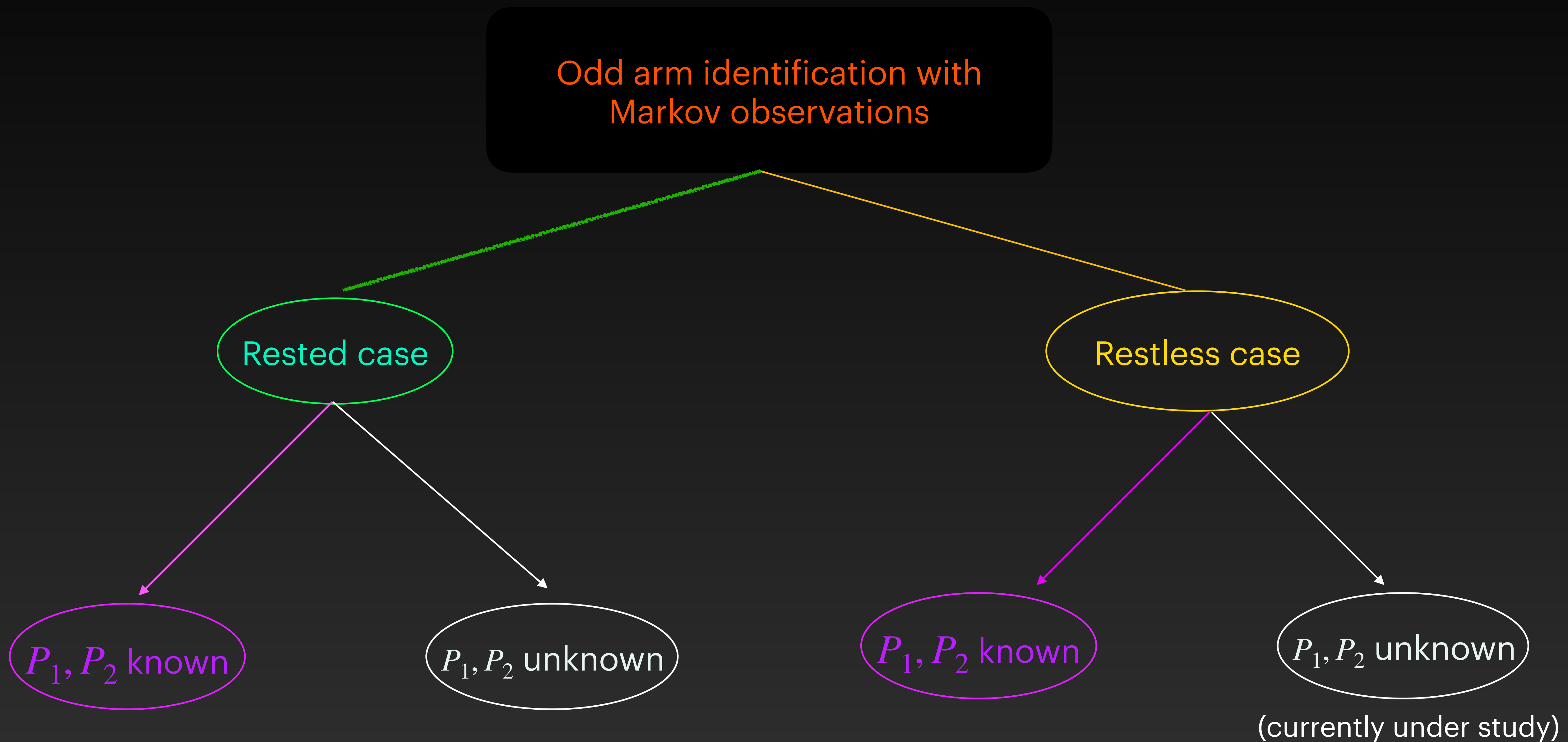
Exact in the asymptotic limit as error vanishes

Time to identify the anomaly $\boxed{\approx}$ $\left(\log \dfrac{1}{\text{Error in reporting the anomaly correctly}} \right) \dfrac{1}{D^*}$

YES!

We characterise D^* explicitly

Organisation



Organisation

Odd arm identification with Markov observations

- Rested case
  - P_1, P_2 known
  - P_1, P_2 unknown
- Restless case
  - P_1, P_2 known
  - P_1, P_2 unknown (currently under study)

$$\log \frac{1}{\text{Error in reporting the anomaly correctly}} \cdot \frac{1}{D^*} \underset{\text{Asymptotic lower bound}}{\overset{\text{Data processing inequality}}{\lesssim}} \text{Time to identify the anomaly}$$

Organisation

Odd arm identification with Markov observations

Rested case

P_1, P_2 known

P_1, P_2 unknown

Restless case

P_1, P_2 known

P_1, P_2 unknown

(currently under study)

$$\log \frac{1}{\text{Error in reporting the anomaly correctly}} \cdot \frac{1}{D^*} \underset{\text{Asymptotic lower bound}}{\overset{\text{Data processing inequality}}{\lesssim}} \text{Time to identify the anomaly} \underset{\text{Asymptotic upper bound}}{\overset{\text{Policy design}}{\lesssim}} \log \frac{1}{\text{Error in reporting the anomaly correctly}} \cdot \frac{1}{D^*}$$

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Supremum is attained for some λ_h^*

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Supremum is attained for some λ_h^*

Keep an estimate $\hat{h}(t)$ of h at time t, and sample according to $\frac{\eta}{K} + (1-\eta)\ \lambda^*_{\hat{h}(t)}(\,\cdot\,)$ for some $\eta > 0$

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Supremum is attained for some λ_h^*

Keep an estimate $\hat{h}(t)$ of h at time t, and sample according to $\boxed{\frac{\eta}{K} + (1-\eta)\ \lambda^*_{\hat{h}(t)}(\,\cdot\,)}$ for some $\eta > 0$

Trembling hand model; ensures sufficient exploration

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Supremum is attained for some λ_h^*

Keep an estimate $\hat{h}(t)$ of h at time t, and sample according to $\boxed{\frac{\eta}{K} + (1-\eta)\, \lambda_{\hat{h}(t)}^*(\,\cdot\,)}$ for some $\eta > 0$

$\hat{h}(t) \to h$ as $t \to \infty$, and so $\lambda_{\hat{h}(t)}^* \to \lambda_h^*$

Trembling hand model; ensures sufficient exploration

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are unknown:

$$D^*_{h,P_1,P_2} = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are unknown:

$$D^*_{h,P_1,P_2} = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Supremum is attained for some λ^*_{h,P_1,P_2}

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are unknown:

$$D^*_{h,P_1,P_2} = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Supremum is attained for some λ^*_{h,P_1,P_2}

Keep estimates $(\hat{h}(t), \hat{P}^t_1, \hat{P}^t_2)$ of (h, P_1, P_2) at time t, and sample according to $\lambda^*_{\hat{h}(t),\hat{P}^t_1,\hat{P}^t_2}$ with η-trembling

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are unknown:

$$D^*_{h,P_1,P_2} = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Supremum is attained for some λ^*_{h,P_1,P_2}

Keep estimates $(\hat{h}(t), \hat{P}^t_1, \hat{P}^t_2)$ of (h, P_1, P_2) at time t, and sample according to $\lambda^*_{\hat{h}(t),\hat{P}^t_1,\hat{P}^t_2}$ with η-trembling

Principle of certainty equivalence

Key Ideas and Challenges

When arm h is the odd arm and the arms are rested:

When P_1 and P_2 are unknown:

$$D^*_{h,P_1,P_2} = \sup_{\lambda(\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Supremum is attained for some λ^*_{h,P_1,P_2}

Keep estimates $(\hat{h}(t), \hat{P}^t_1, \hat{P}^t_2)$ of (h, P_1, P_2) at time t, and sample according to $\lambda^*_{\hat{h}(t),\hat{P}^t_1,\hat{P}^t_2}$ with η-trembling

Principle of certainty equivalence

$(\hat{h}(t), \hat{P}^t_1, \hat{P}^t_2) \to (h, P_1, P_2)$ as $t \to \infty$

Continuity of λ^* (e.g., Berge's maximum theorem) will give the result

Key Ideas and Challenges

When arm h is the odd arm and the arms are restless:

A key MDP formulation makes the problem amenable to analysis. The MDP viewpoint is new and does not appear in the prior works. Trembling hand ensures ergodicity.

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot|\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Key Ideas and Challenges

When arm h is the odd arm and the arms are restless:

A key MDP formulation makes the problem amenable to analysis. The MDP viewpoint is new and does not appear in the prior works. Trembling hand ensures ergodicity.

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot|\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Not clear if the supremum is attained

Key Ideas and Challenges

When arm h is the odd arm and the arms are restless:

A key MDP formulation makes the problem amenable to analysis. The MDP viewpoint is new and does not appear in the prior works. Trembling hand ensures ergodicity.

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot|\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Not clear if the supremum is attained

Supremum can be approached $\delta -$ closely for any $\delta > 0$

Key Ideas and Challenges

When arm h is the odd arm and the arms are restless:

A key MDP formulation makes the problem amenable to analysis. The MDP viewpoint is new and does not appear in the prior works. Trembling hand ensures ergodicity.

When P_1 and P_2 are known:

$$D_h^* = \sup_{\lambda(\cdot|\cdot)} \min_{h' \neq h} \cdots\cdots\cdots$$

Not clear if the supremum is attained

Supremum can be approached $\delta -$ closely for any $\delta > 0$

Stitching the solutions for various δ does the job

Some Future Directions

- General state space for each arm (our work deals with finite state space)
- More than one odd arm
- Number of arms $\rightarrow \infty$
- Time is continuous rather than discrete

Publications

- P. N. Karthik, R. Sundaresan, “Detecting an Odd Markov Arm with a Trembling Hand”, proceedings of the 2020 IEEE International Symposium on Information Theory (virtual).

- P. N. Karthik, R. Sundaresan, “Learning to Detect an Odd Markov Arm”, proceedings of the 2019 IEEE International Symposium on Information Theory, Paris, France.

- P. N. Karthik, R. Sundaresan, “Learning to Detect an Odd Markov Arm”, IEEE Transactions on Information Theory, vol. 66, no. 7, pp. 4324-4348.

Community Members Working Actively on Related Topics

Rajesh Sundaresan
IISc

Srikrishna Bhashyam
IIT Madras

Aditya Gopalan
IISc

Sandeep Juneja
TIFR Mumbai

Jonathan Scarlett
NUS

Aurélien Garivier
ENS

Emilie Kaufmann
INRIA

Venugopal Veeravalli
UIUC

Acknowledgements

- The Science and Engineering Research Board, Department of Science and Technology (grant no. EMR/2016/002503)
- The Robert Bosch Centre for Cyber-Physical Systems, Indian Institute of Science
- The Centre for Networked Intelligence, Indian Institute of Science

Thank You!