

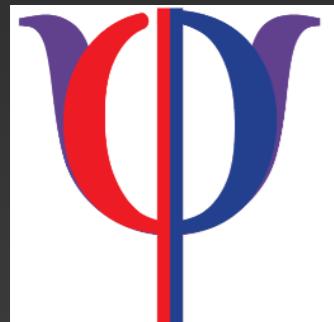
Odd Arm Identification in Multi-armed Bandits with Markov Observations

EECS Research Students Symposium - 2020

Karthik P. N.

Department of Electrical Communication Engineering

This work was supported by



CENTRE FOR NETWORKED
INTELLIGENCE

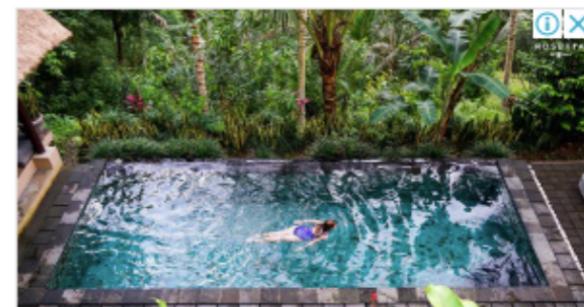
Identification of a Malicious Ad

At the top of search results for "Amazon" was a bad ad, trying to tricking users into falling for a tech support scam.



By Zack Whittaker for Zero Day | March 16, 2018 -- 00:19 GMT (05:49 IST) | Topic: Security

The screenshot shows a Google search results page for the query "amazon". At the top, there are social sharing icons for LinkedIn, Facebook, and Twitter. Below the search bar, the results are displayed. The first result is a sponsored link from ROSETTA, which appears to be a scam. The link is titled "Amazon.com | Amazon Official Site | Huge Selection & Great Prices" and includes a star rating of 4.7 and a delivery guarantee. The main search results for Amazon.com follow, including links for "Earth's Biggest Selection", "Books", and "Amazon Home".



Bangalore's First Holiday Club



Luxury getaway resorts just 3 hours from Bangalore. Get 1 week of holidays every year



MORE FROM ZACK WHITTAKER



Security
Online security 101: Tips for protecting your privacy from hackers and spies



Security
US government's "do not buy" list shuts out Russia, China

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

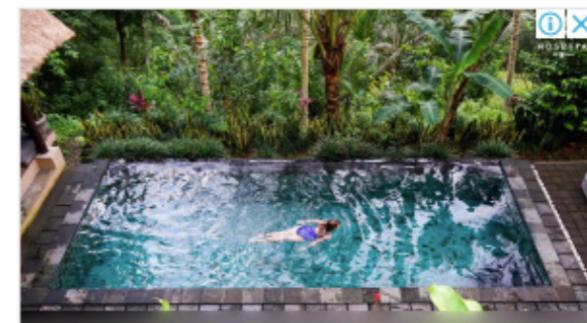
Identification of a Malicious Ad

At the top of search results for "Amazon" was a bad ad, trying to tricking users into falling for a tech support scam.



By Zack Whittaker for Zero Day | March 16, 2018 -- 00:19 GMT (05:49 IST) | Topic: Security

The screenshot shows a Google search results page for the query "amazon". At the top, there are social sharing icons for LinkedIn, Facebook, and Twitter. Below the search bar, the results are displayed. The first result is a sponsored link from ROSETTA, which is highlighted with a red box. The link points to a page titled "Bangalore's First Holiday Club" featuring a woman swimming in a pool. The main search results below are from the official Amazon.com website, including links for "Earth's Biggest Selection", "Books", and "Amazon Home".



Bangalore's First Holiday Club



Luxury getaway resorts just 3 hours from Bangalore. Get 1 week of holidays every year



MORE FROM ZACK WHITTAKER

Security

Online security 101: Tips for protecting your privacy from hackers and spies



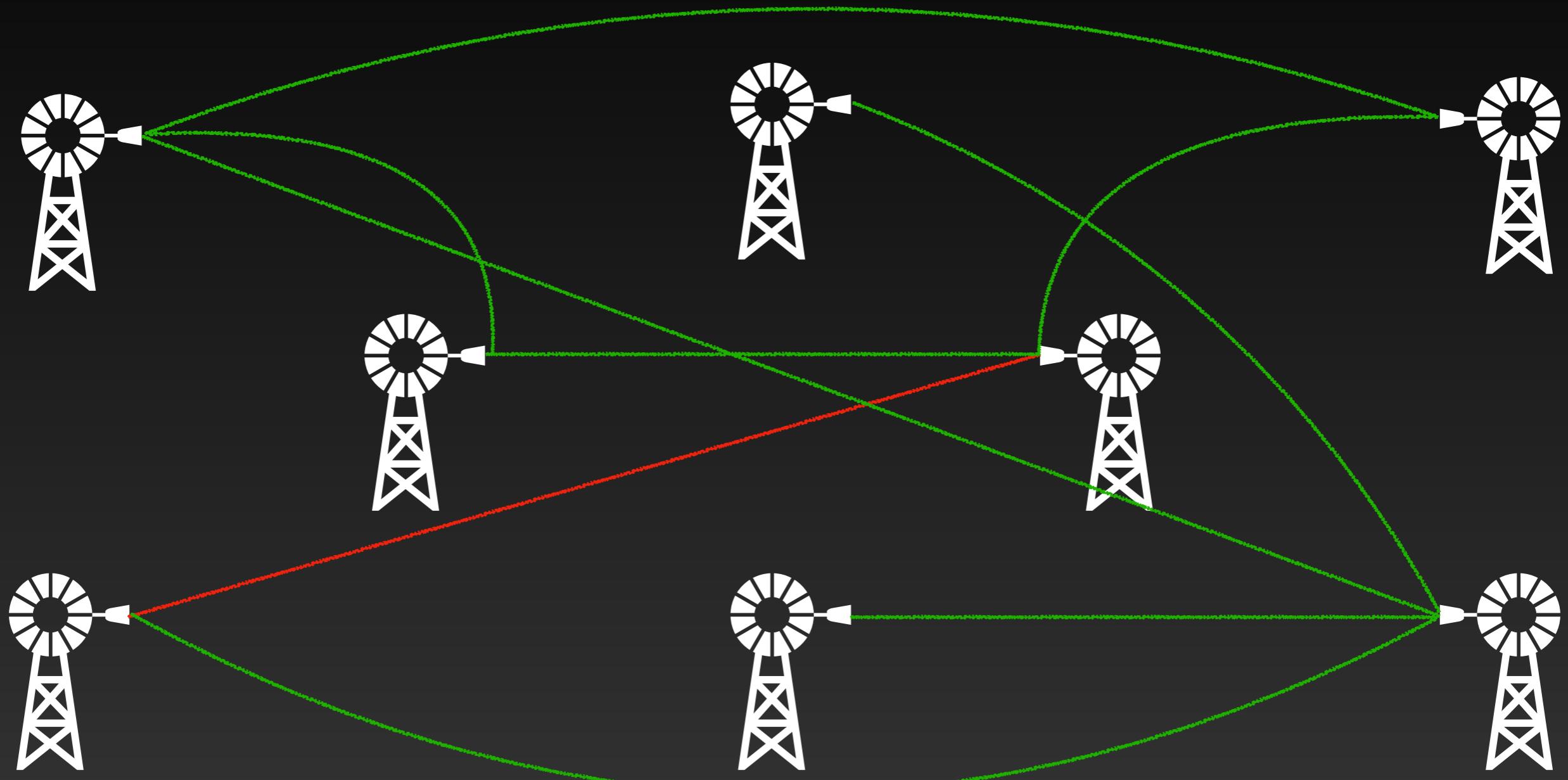
Security

US government's "do not buy" list shuts out Russia, China



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

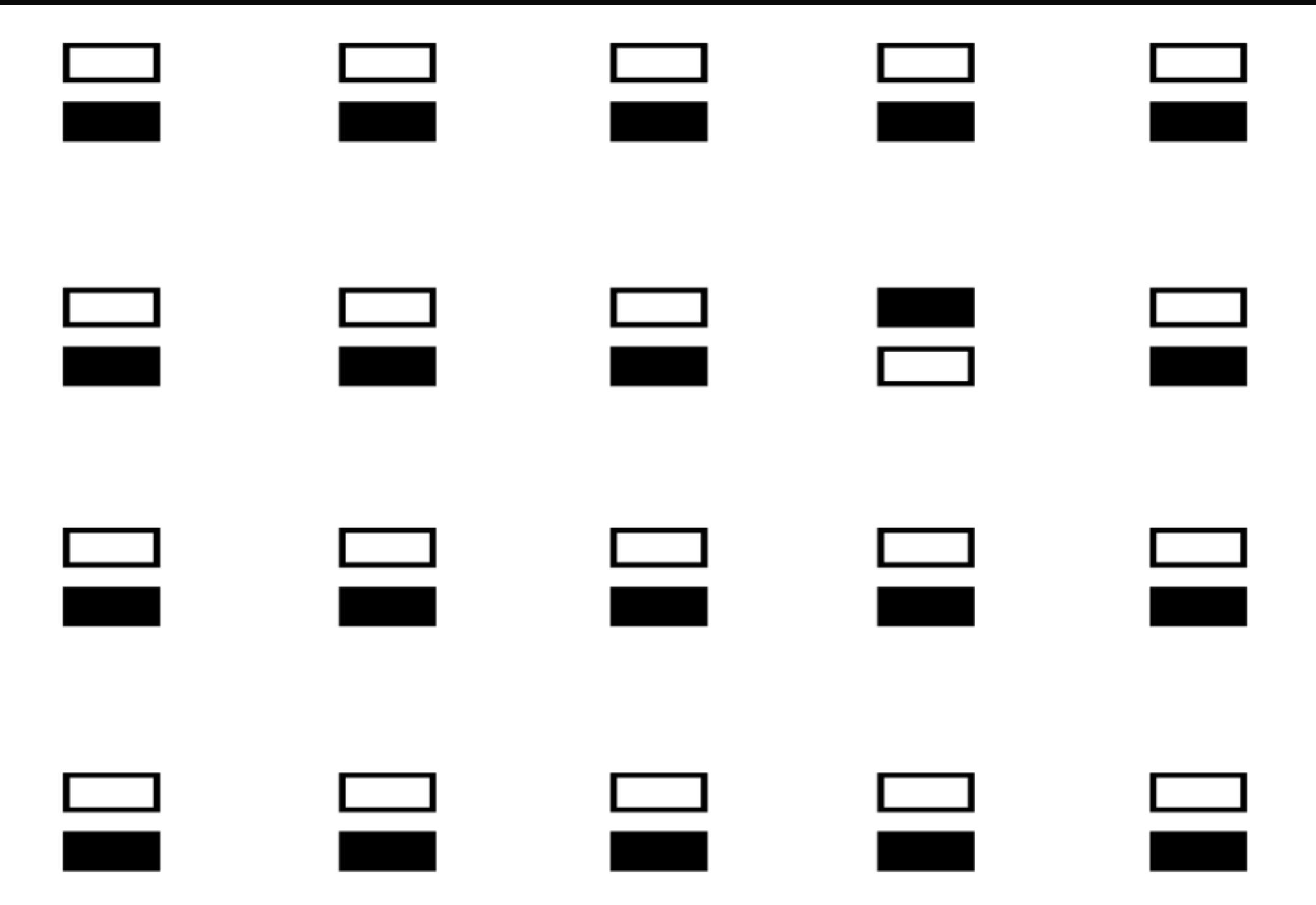
Identification of a Line in Outage in a Power Grid



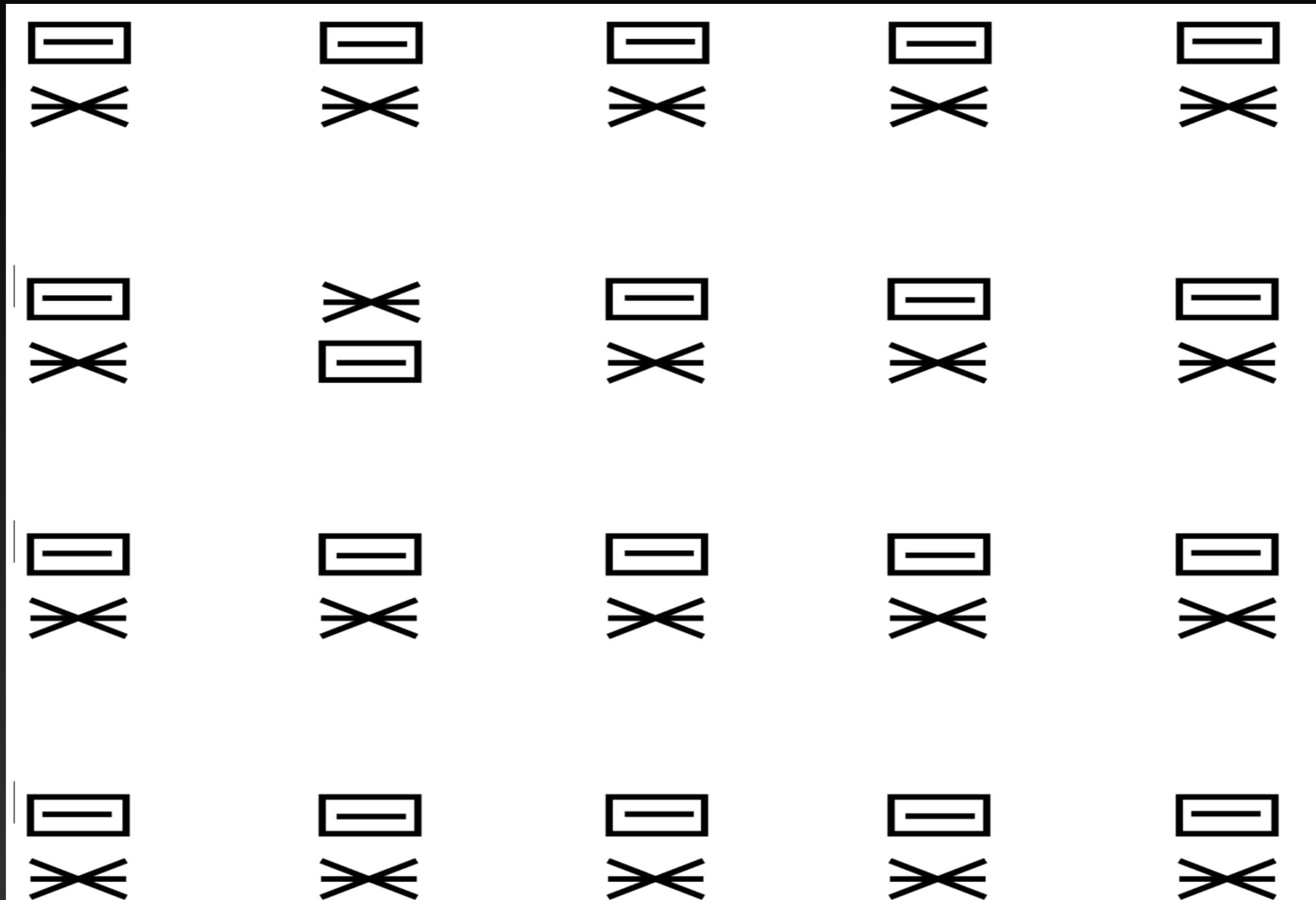
Line working fine

Line in outage

Odd Image Identification

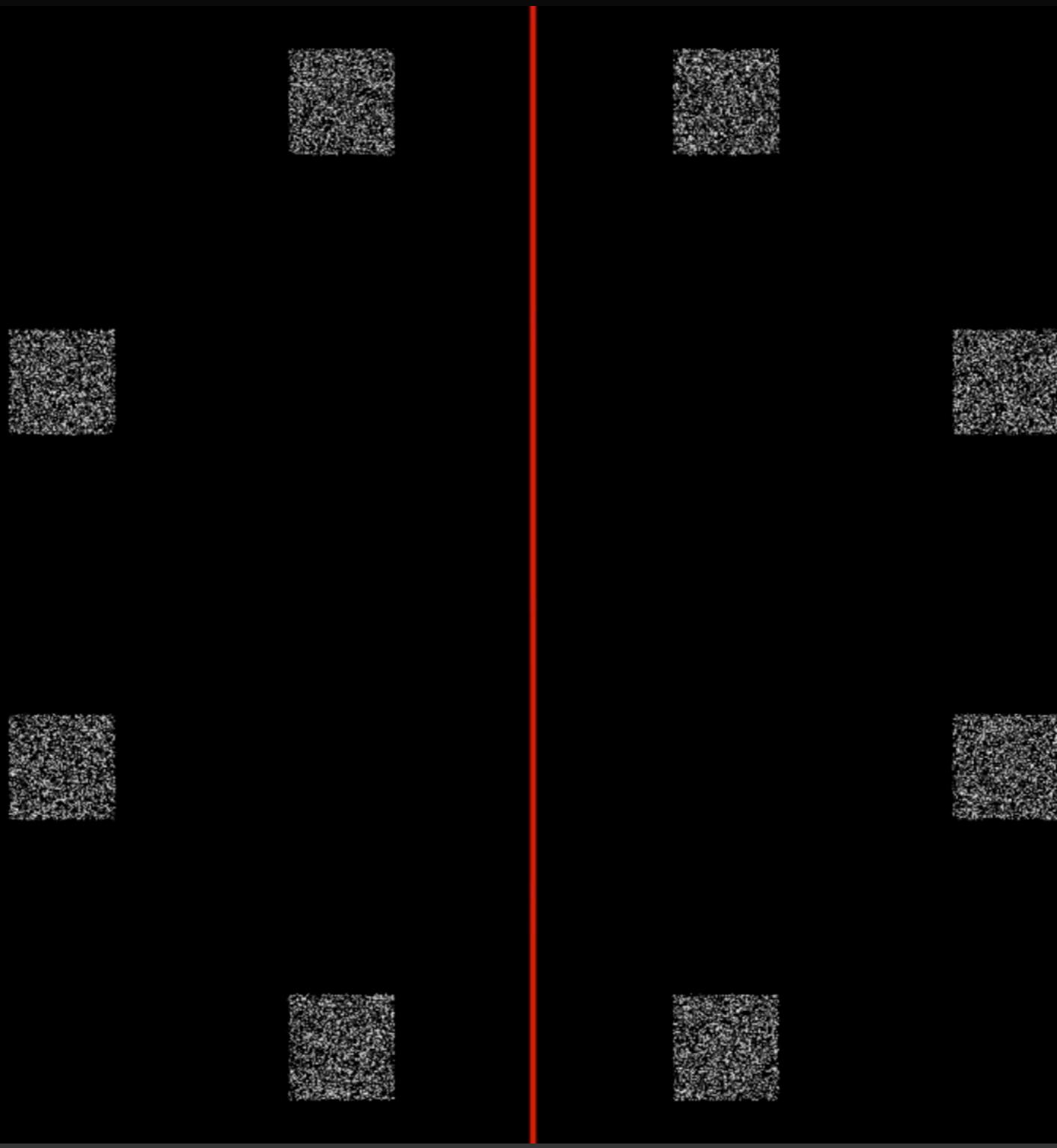


Odd Image Identification

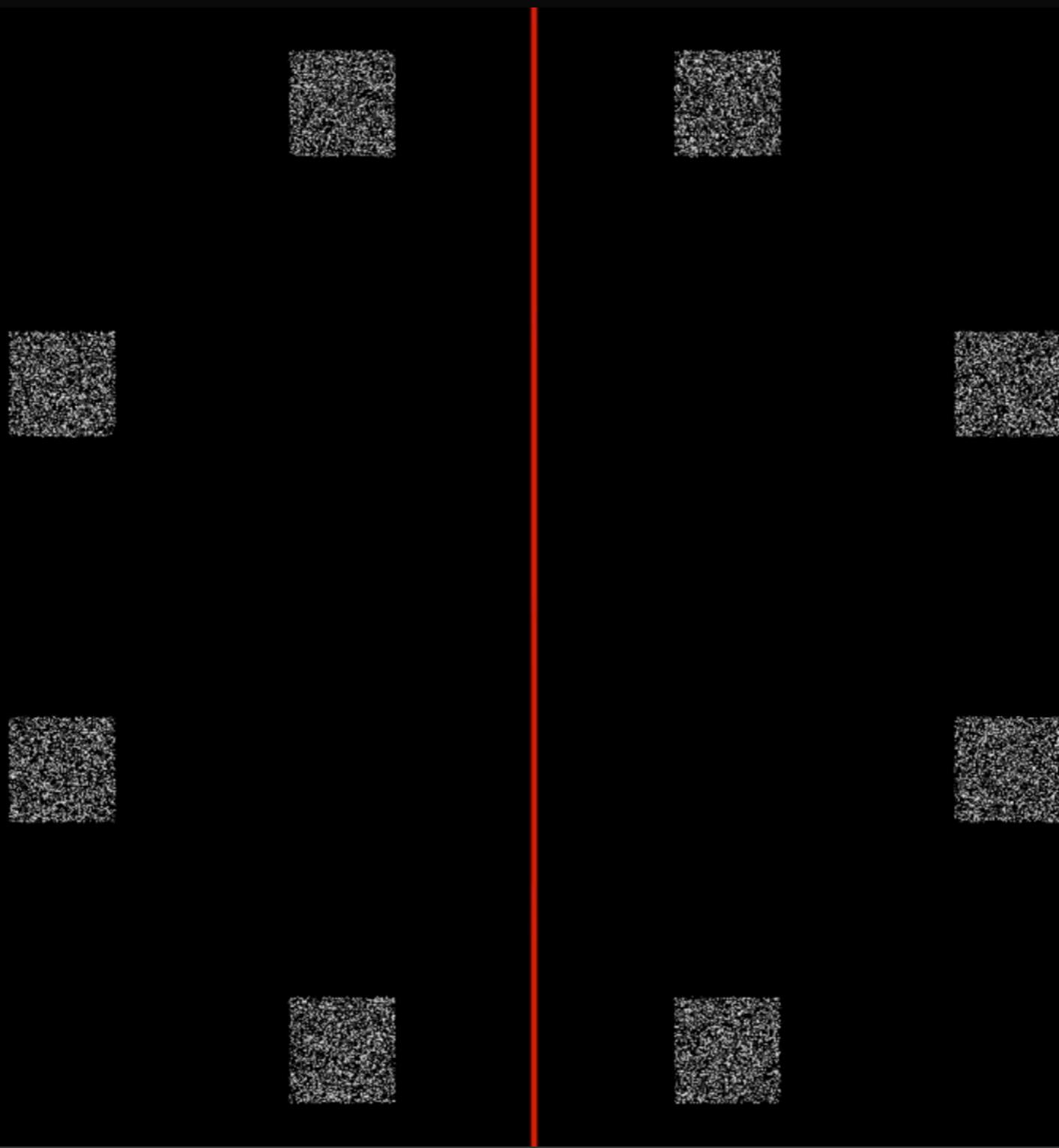


Identify the 'odd' image.

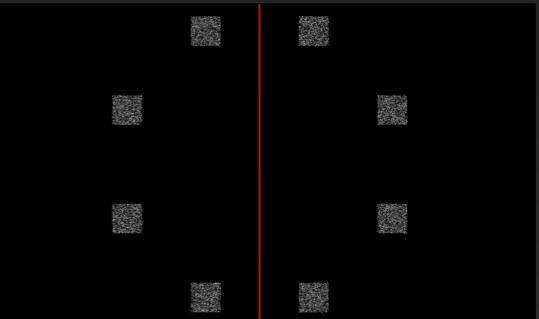
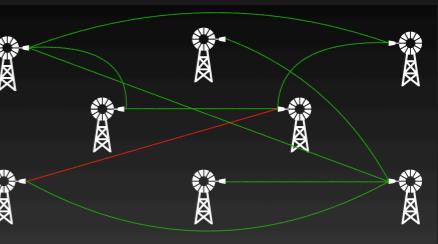
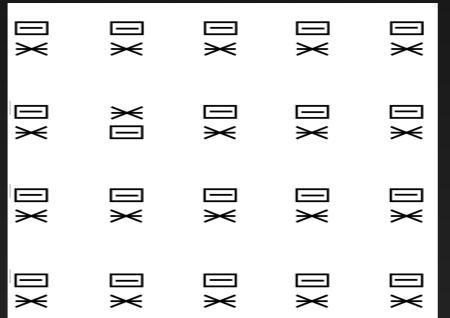
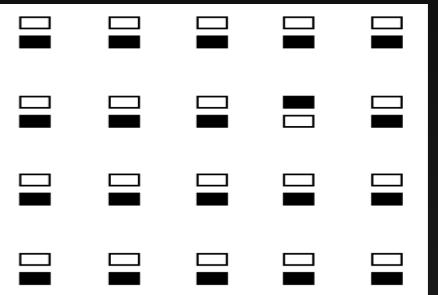
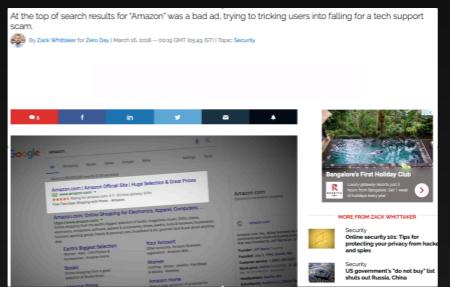
Odd Movie Identification



Odd Movie Identification

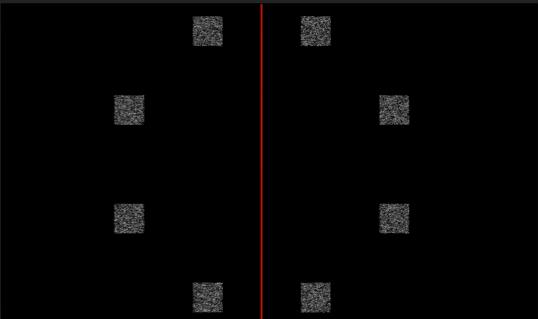
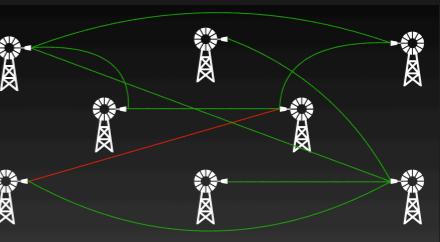
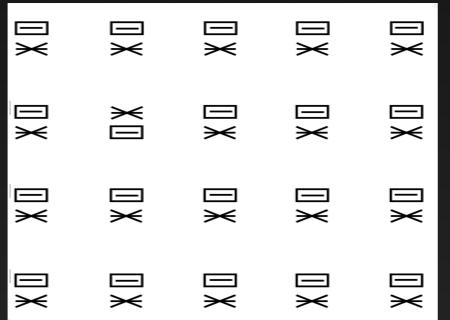
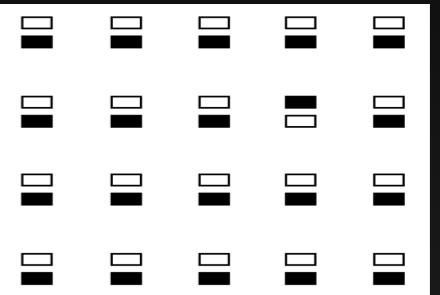
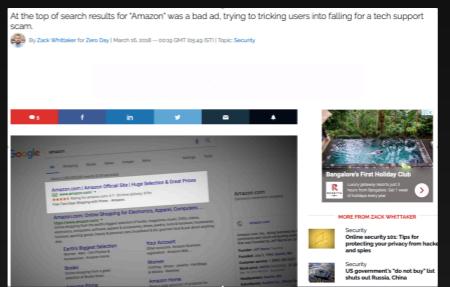


A Closer Look at the Examples

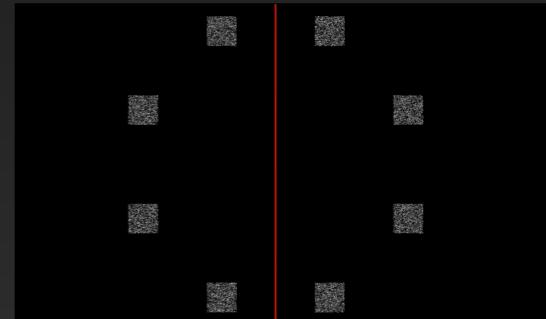
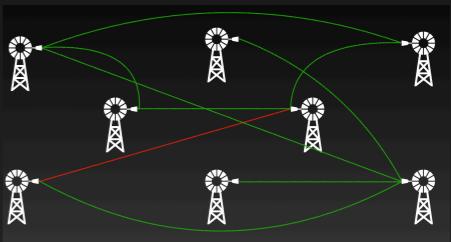
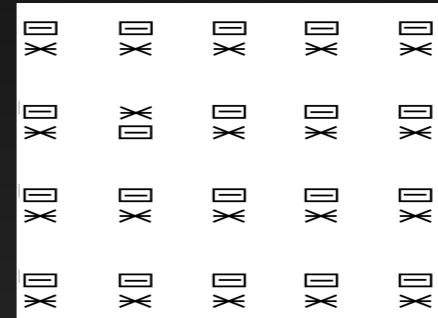
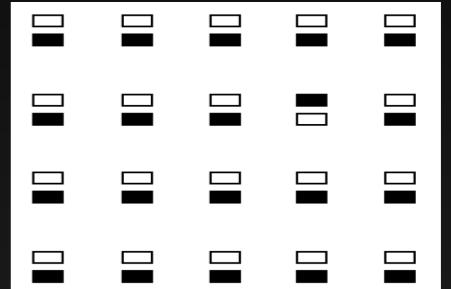
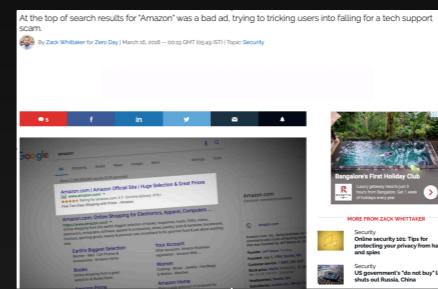
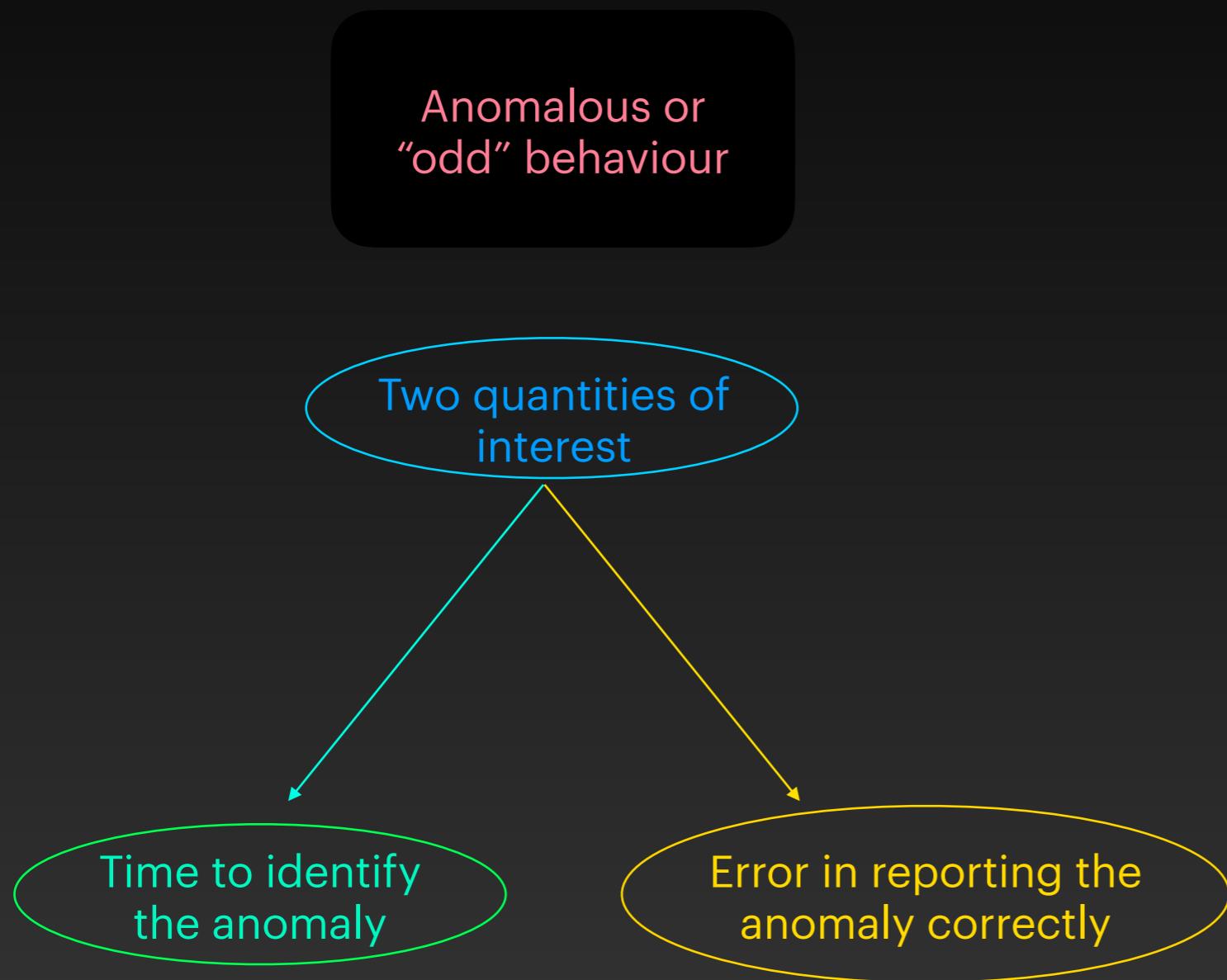


A Closer Look at the Examples

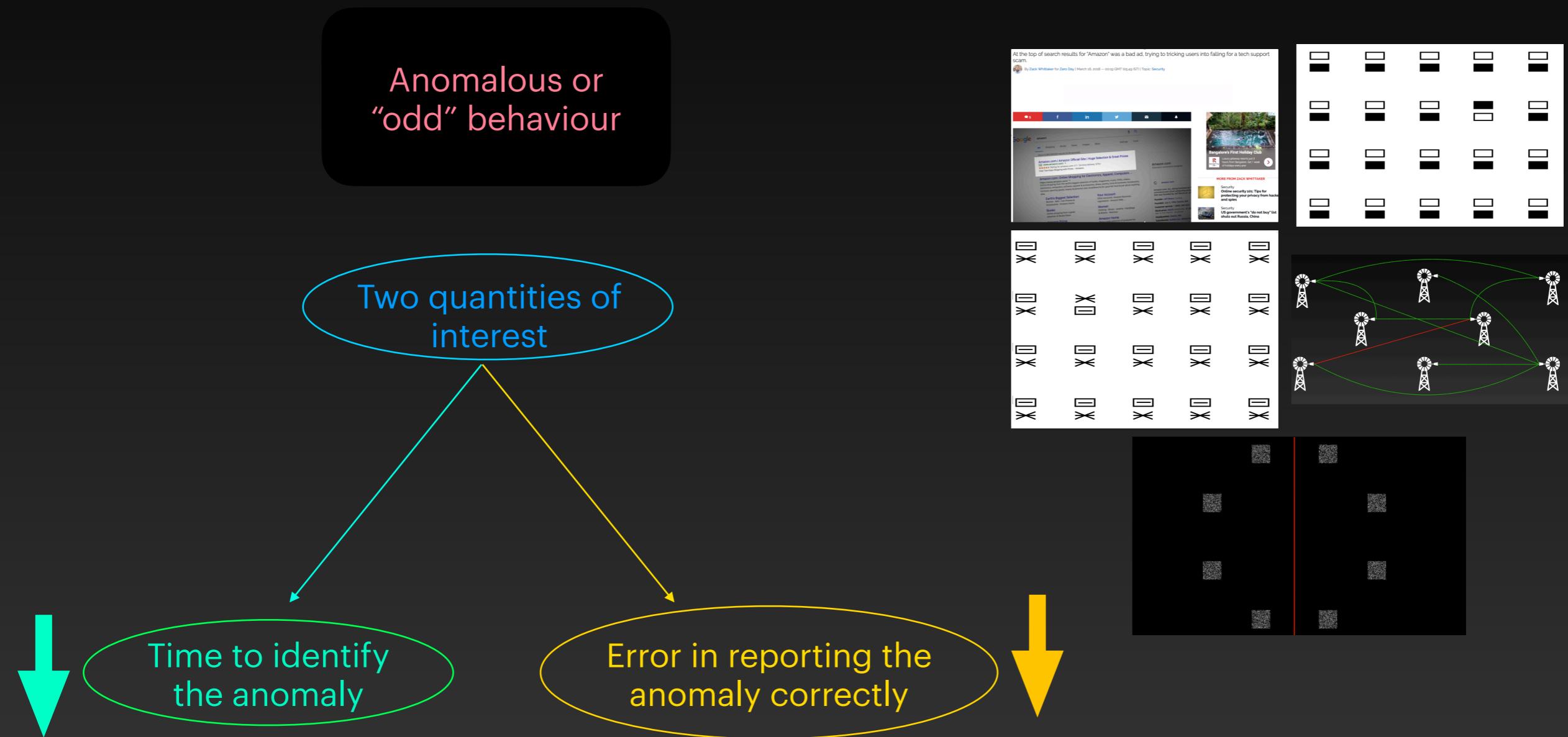
Anomalous or
“odd” behaviour



A Closer Look at the Examples



A Closer Look at the Examples



A study of the prior works reveals that

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

D^* captures the hardness of the problem

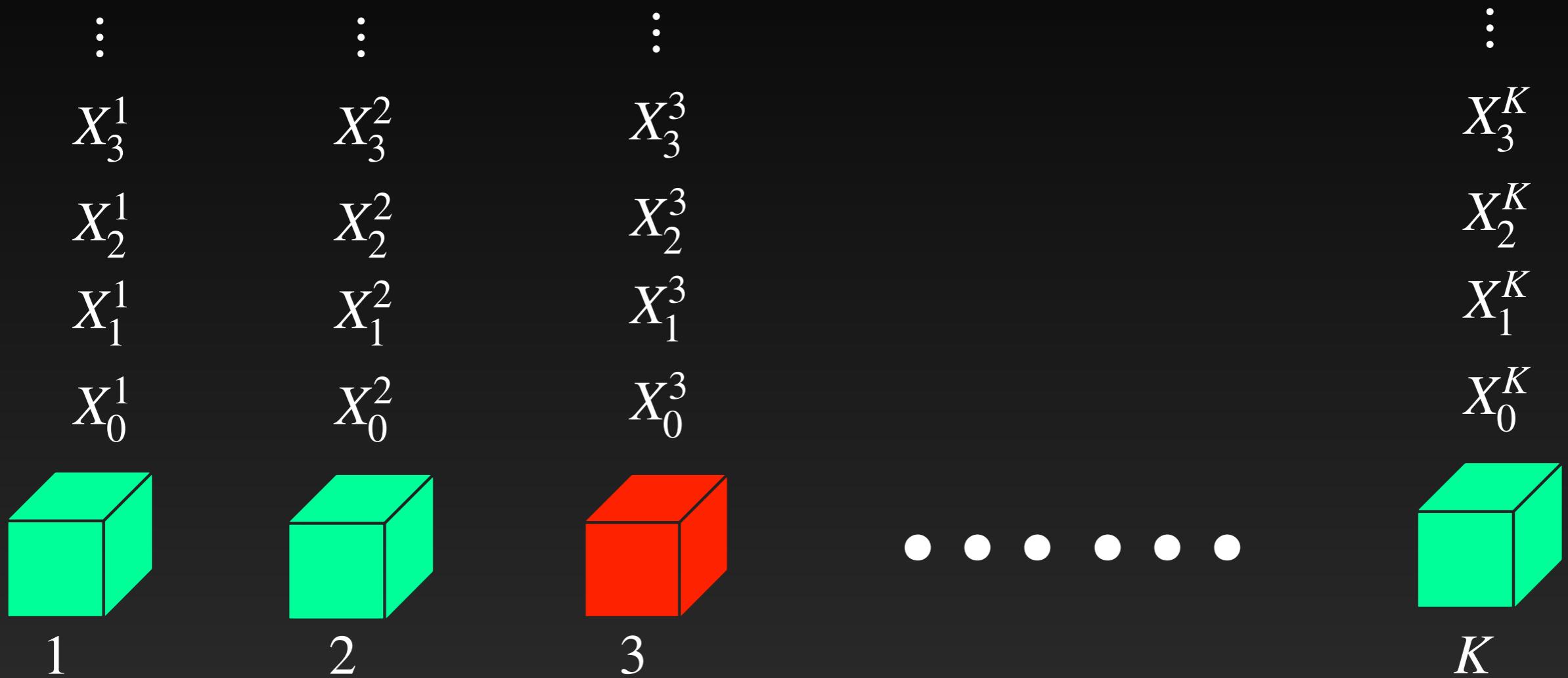
A study of the prior works reveals that

Exact in the asymptotic limit
as **error** vanishes

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

D^* captures the hardness of the problem

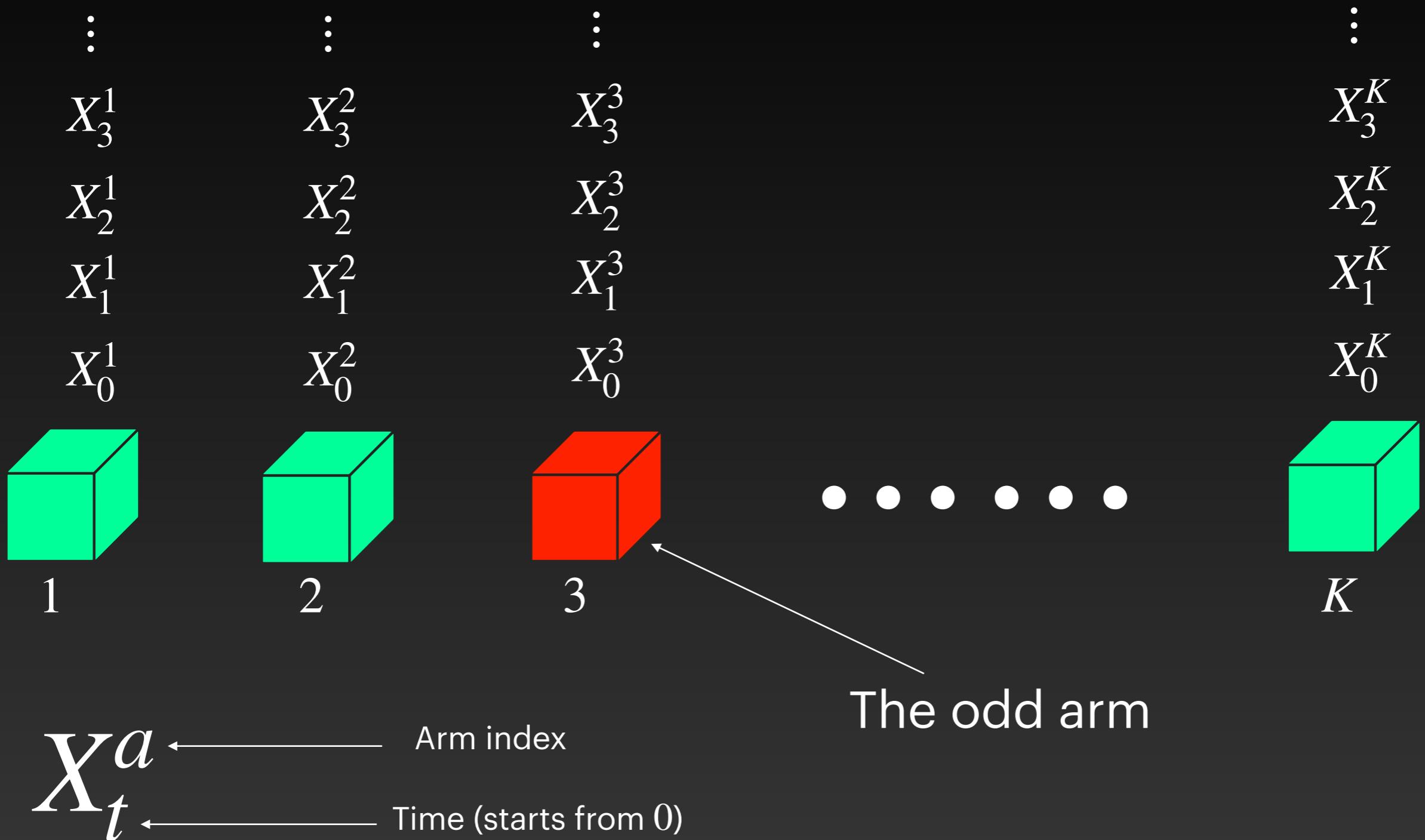
Multi-armed Bandits



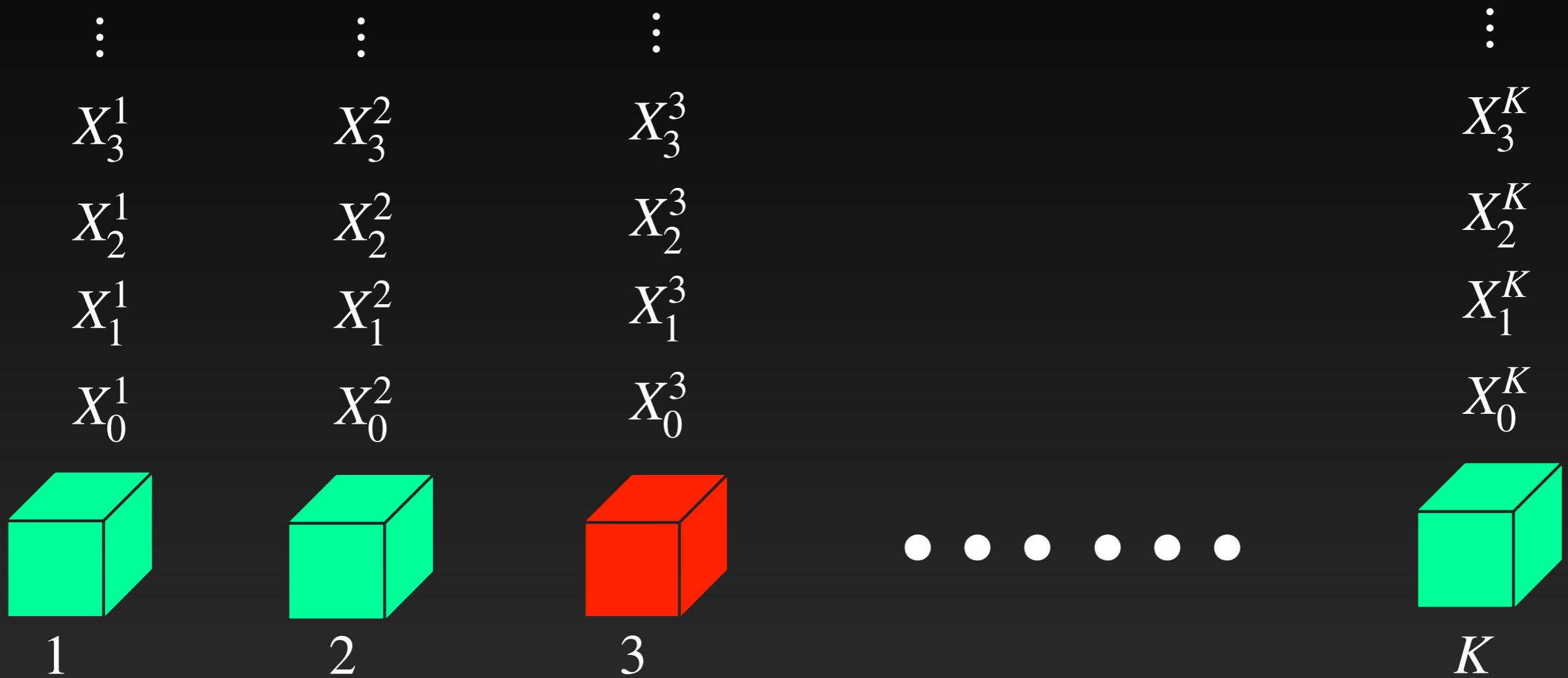
X_t^a ←———— Arm index

X_t ←———— Time (starts from 0)

Multi-armed Bandits

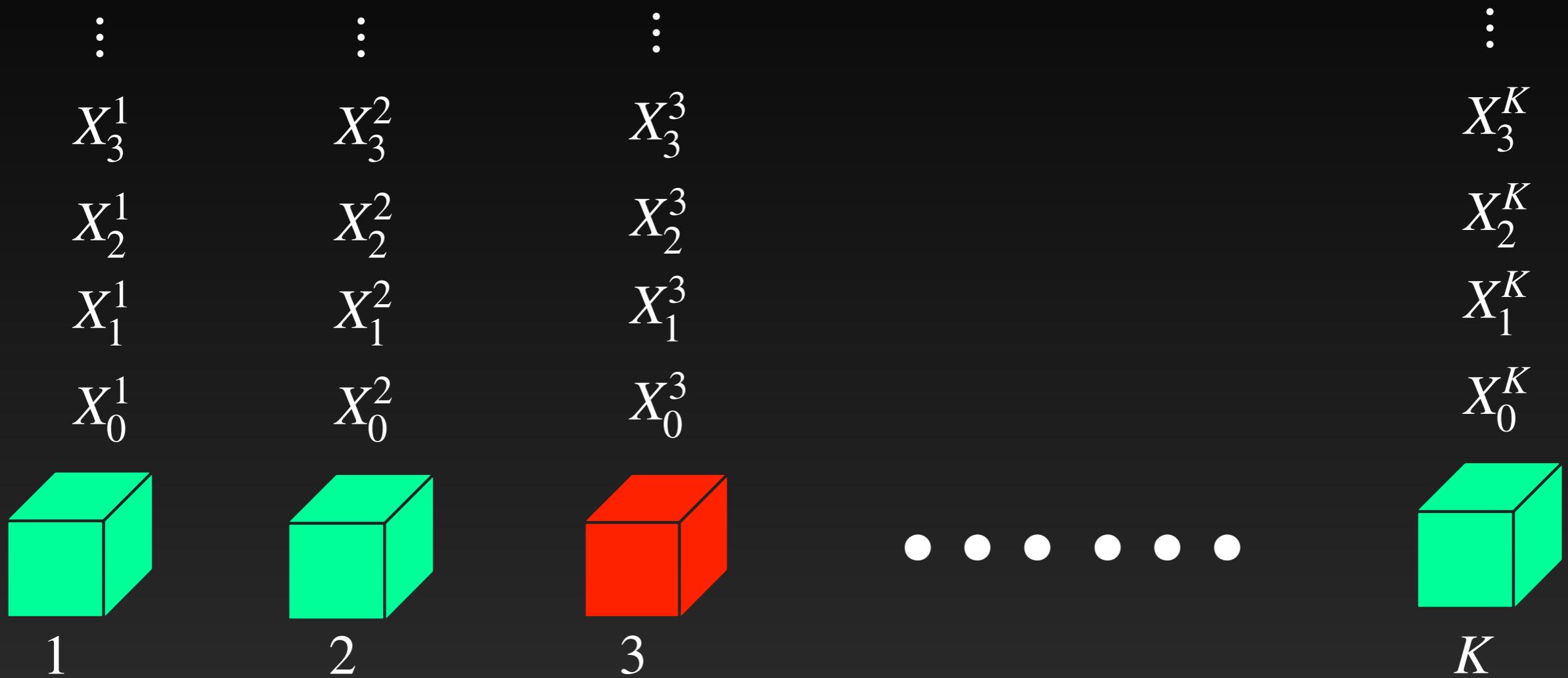


Multi-armed Bandits



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

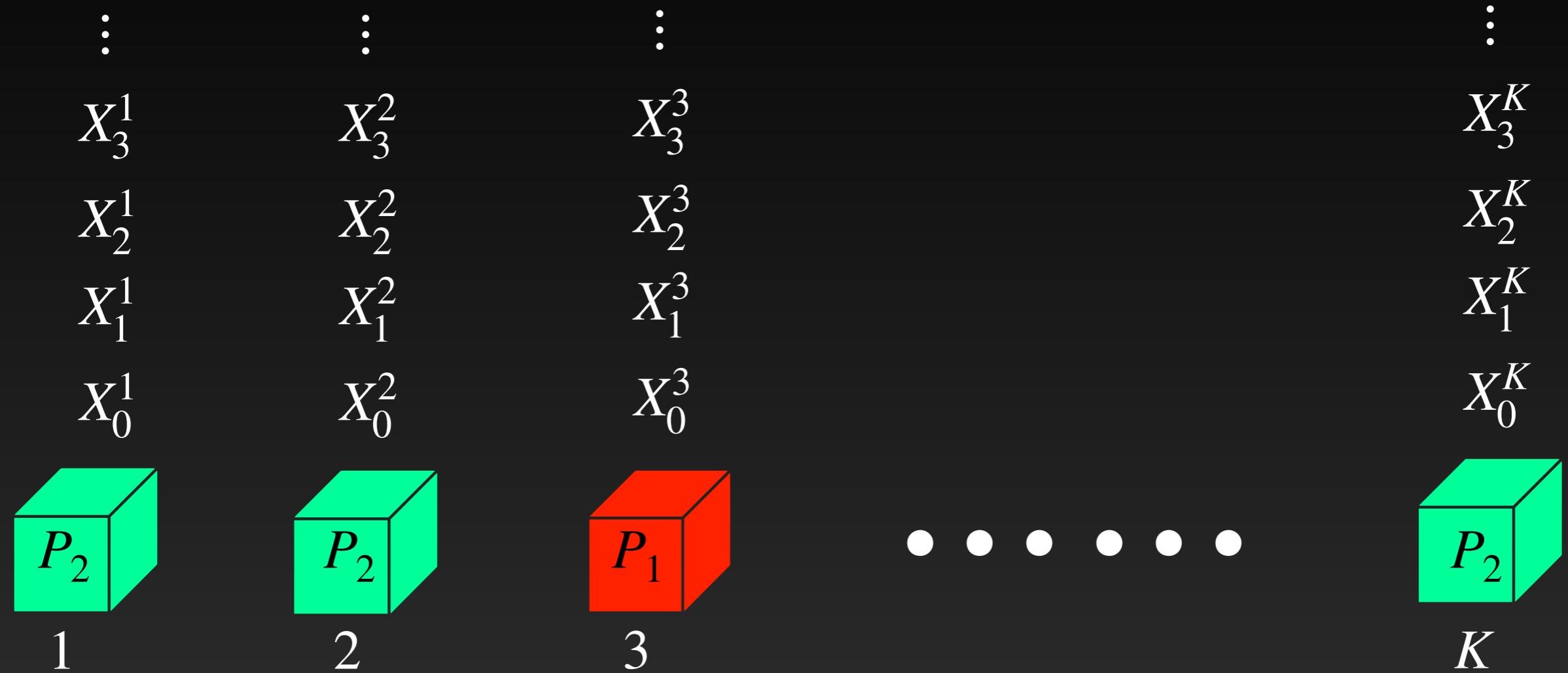
Multi-armed Bandits



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

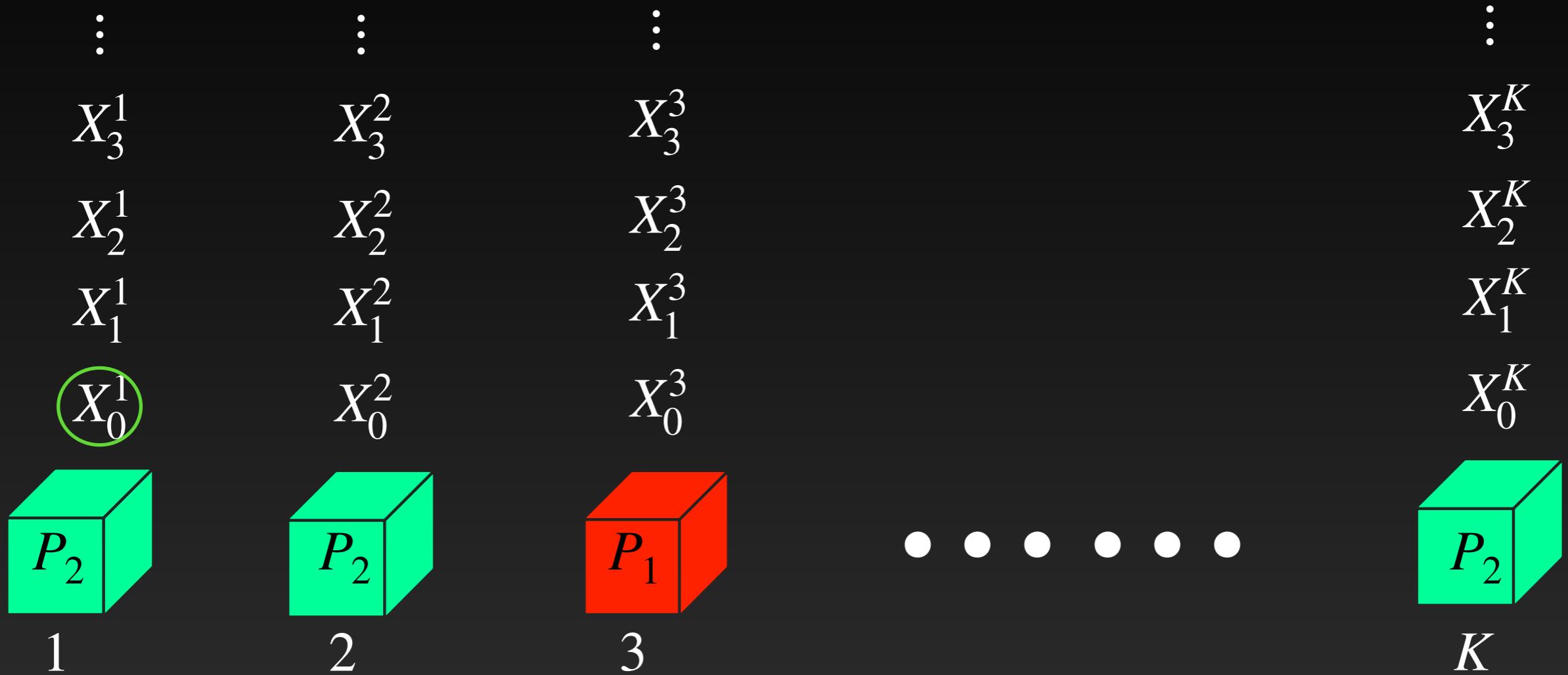
Our work: $\{X_t^a : t \geq 0\}$ is a Markov process

Markov Observations: Rested Case



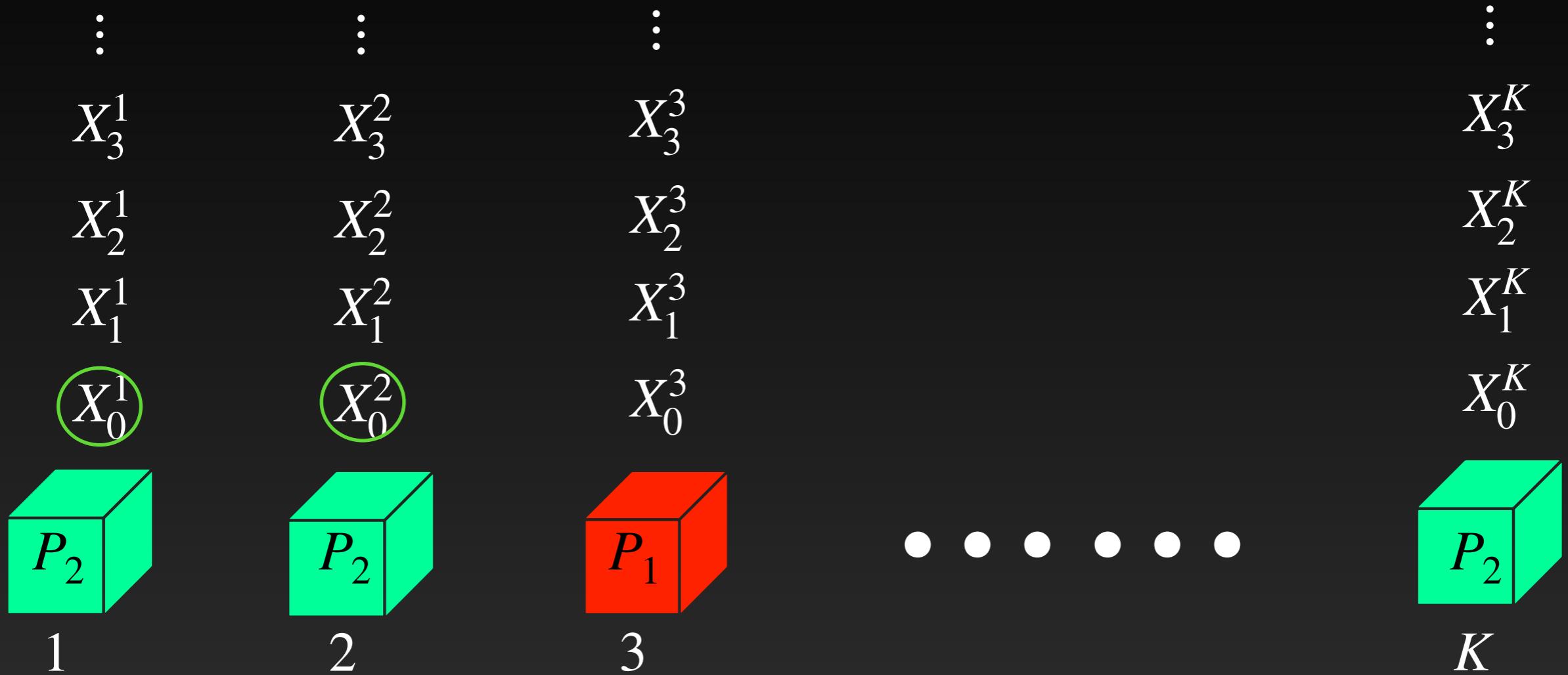
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs						

Markov Observations: Rested Case



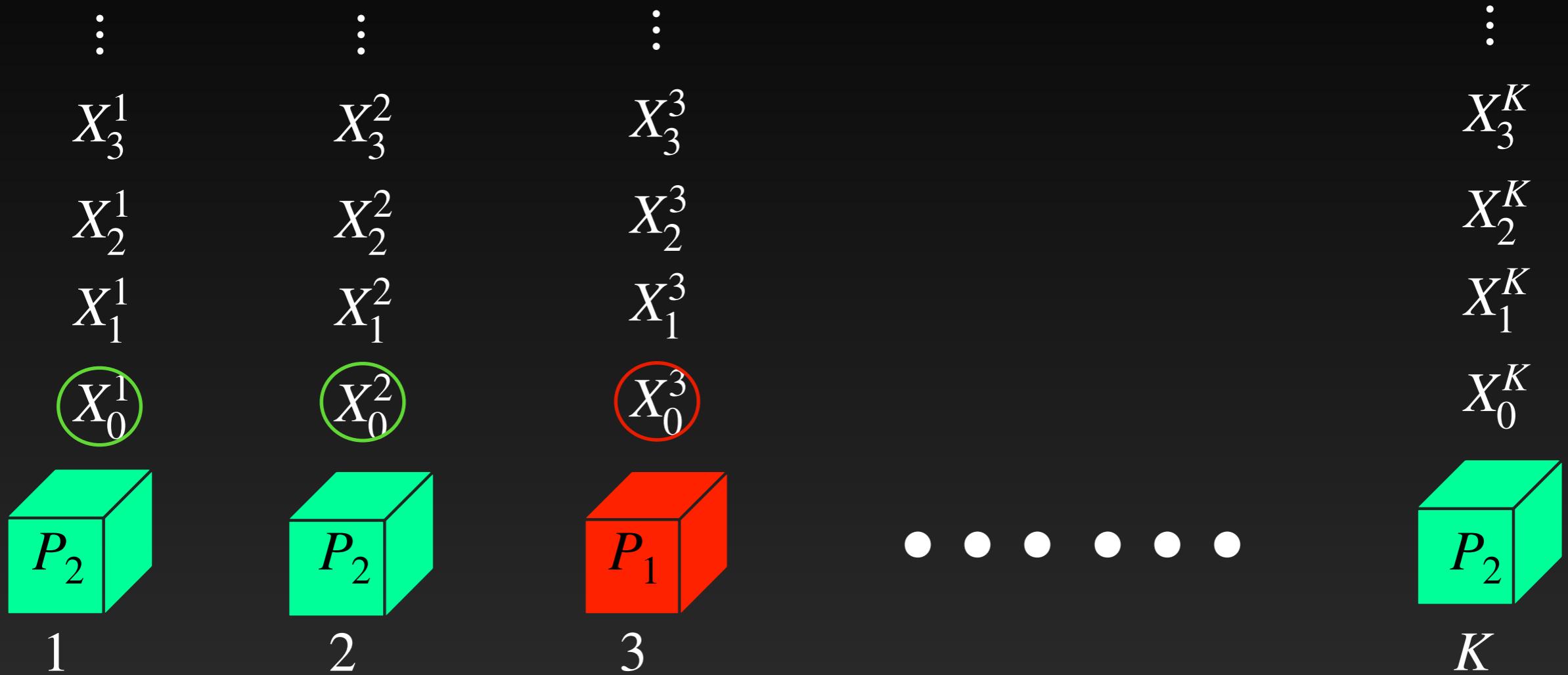
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1					

Markov Observations: Rested Case



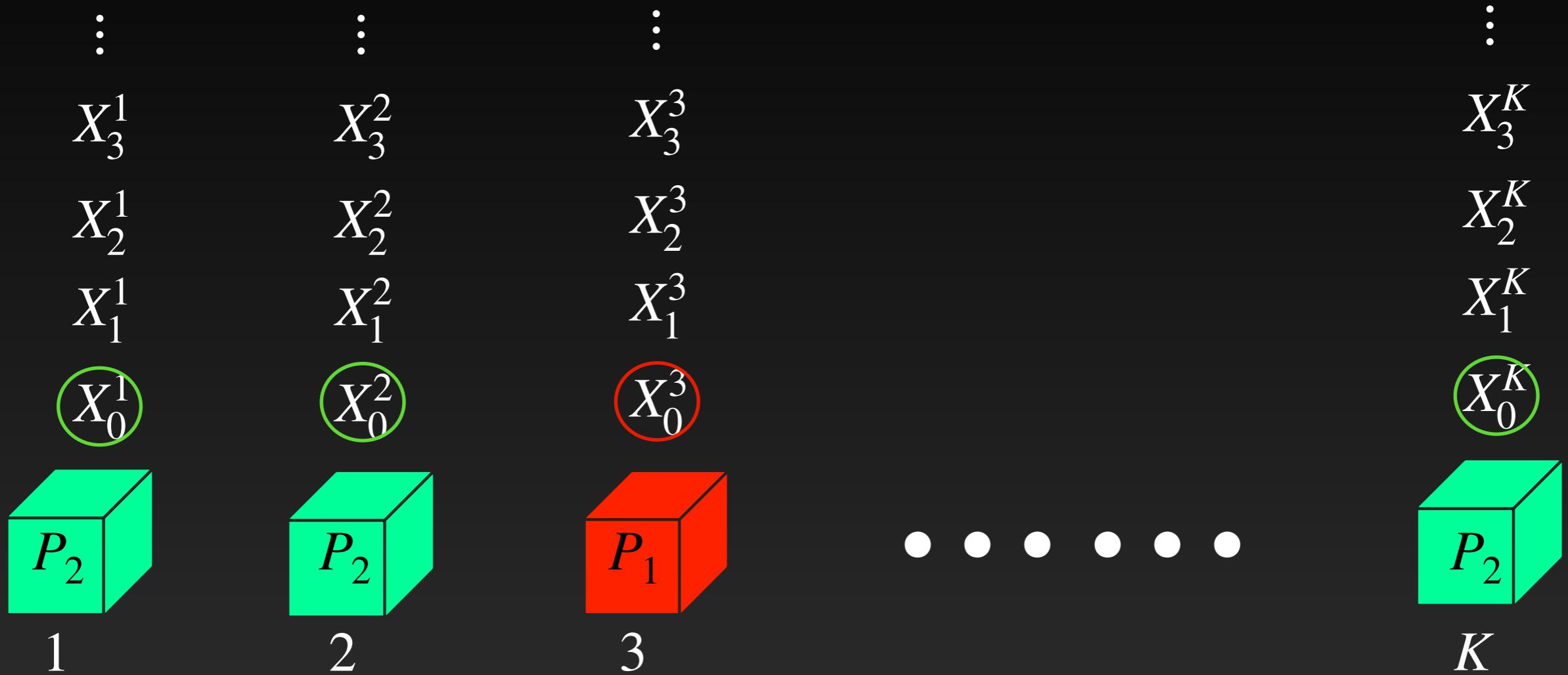
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2				

Markov Observations: Rested Case



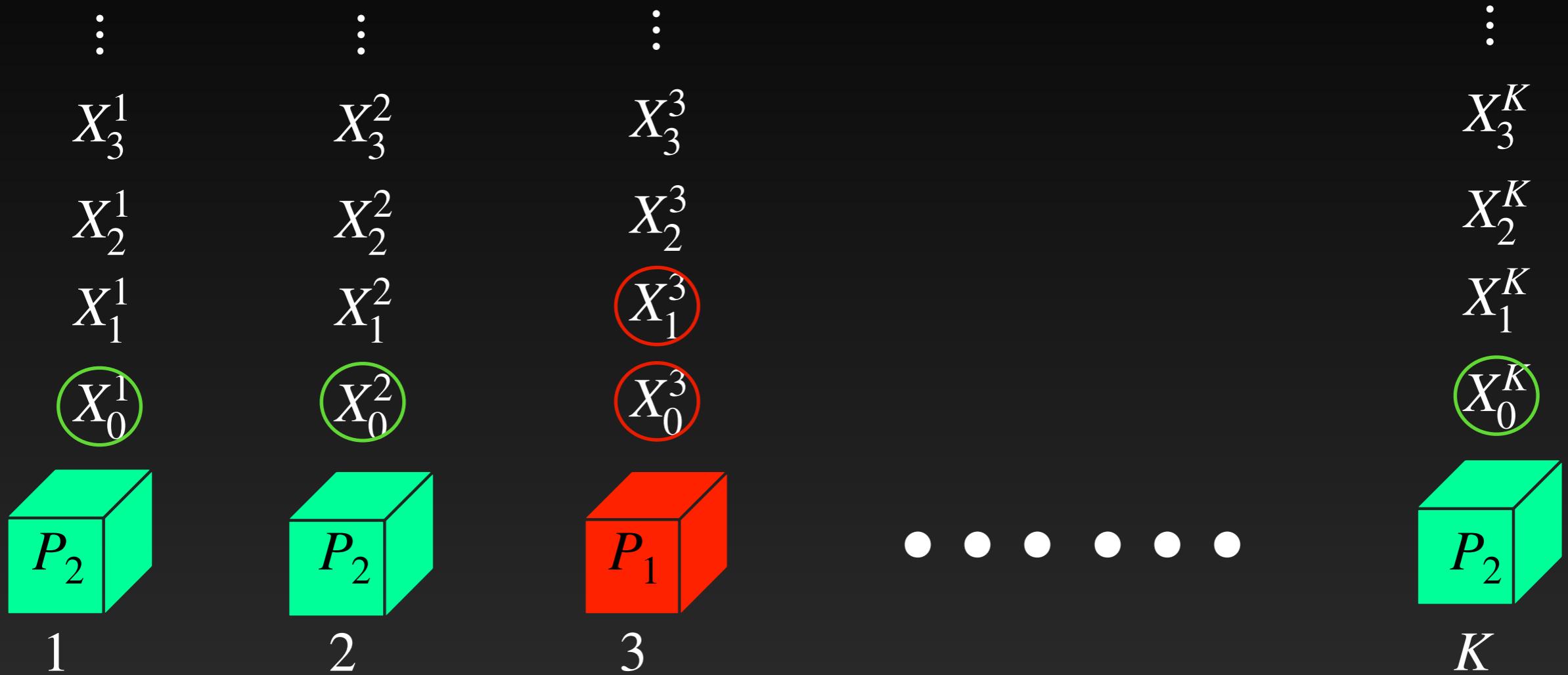
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3			

Markov Observations: Rested Case



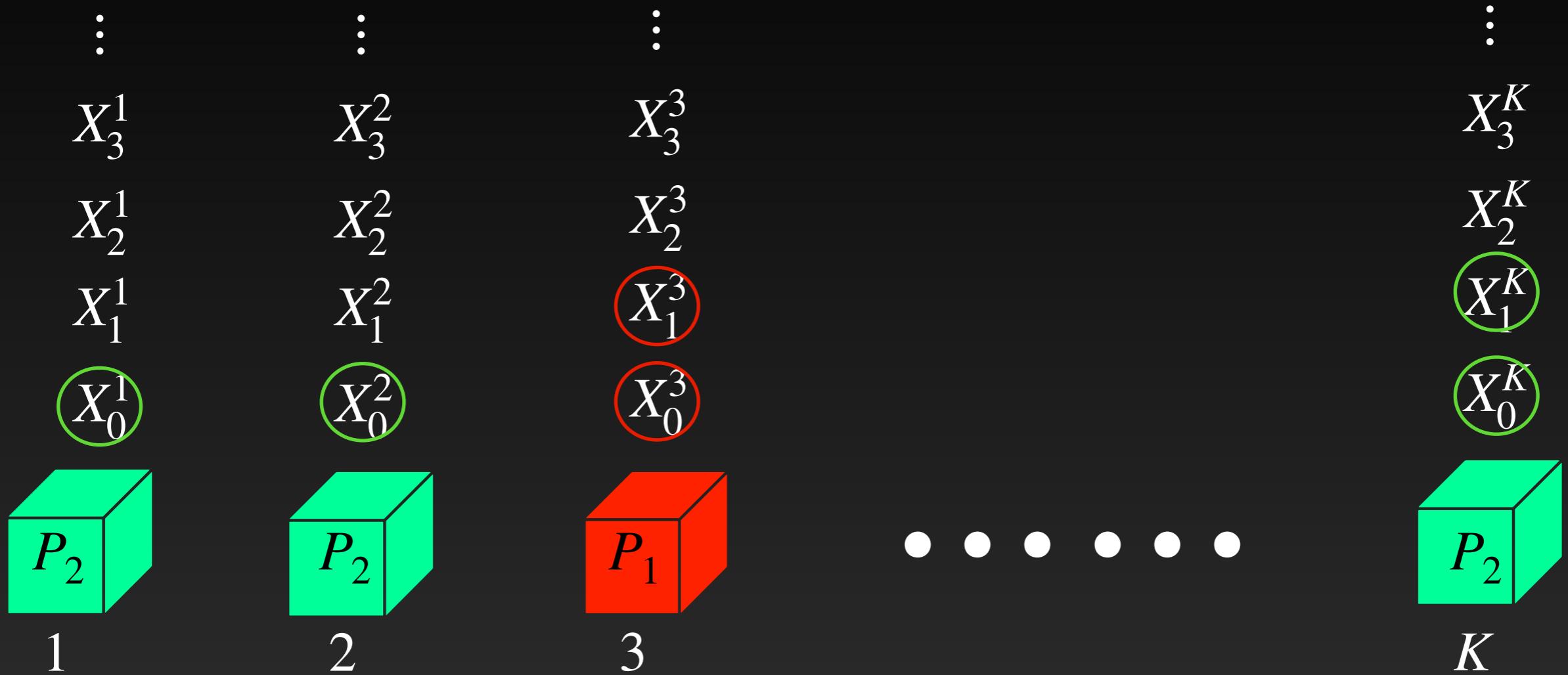
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3	X_0^K		

Markov Observations: Rested Case



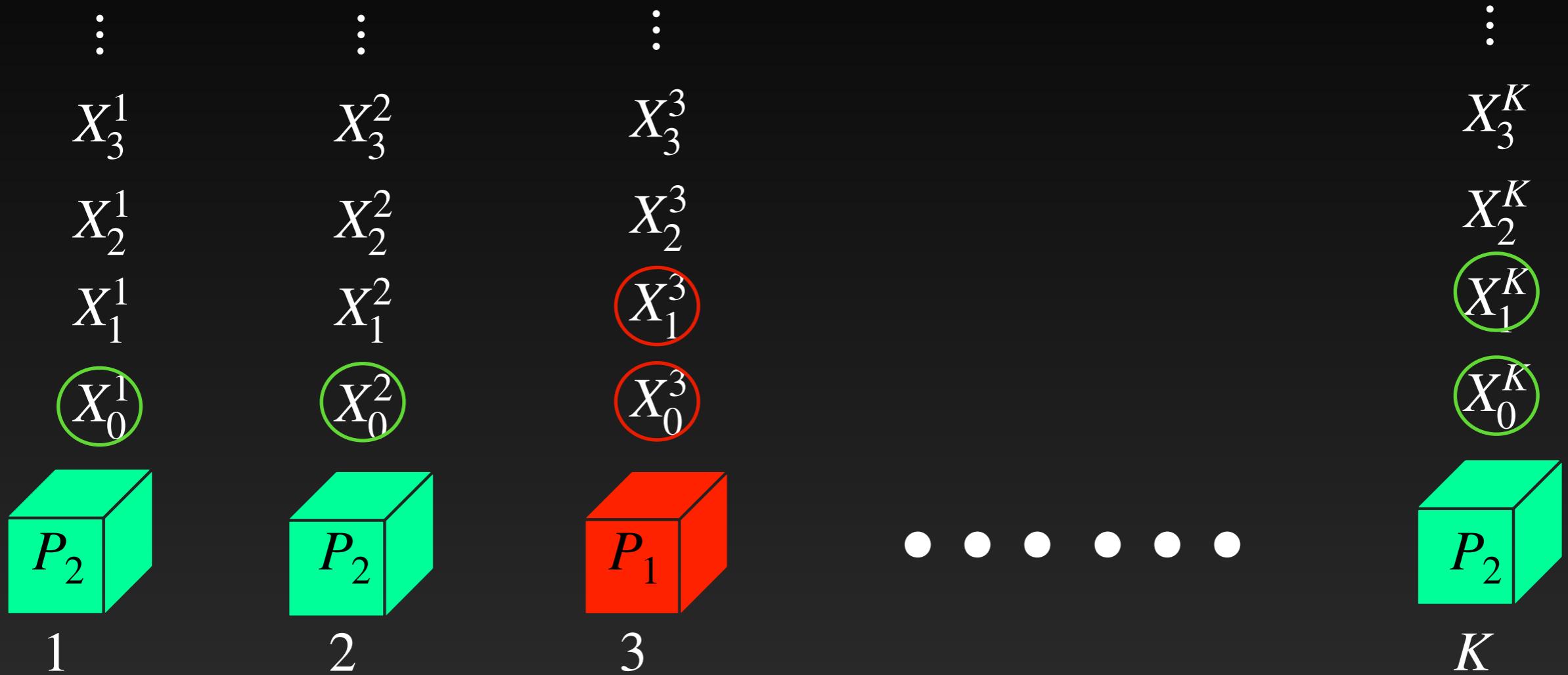
	t	0	1	2	3	4	5
Arm	1	2	3	K	3	K	
Obs	X_0^1	X_0^2	X_0^3	X_0^K	X_1^3		

Markov Observations: Rested Case



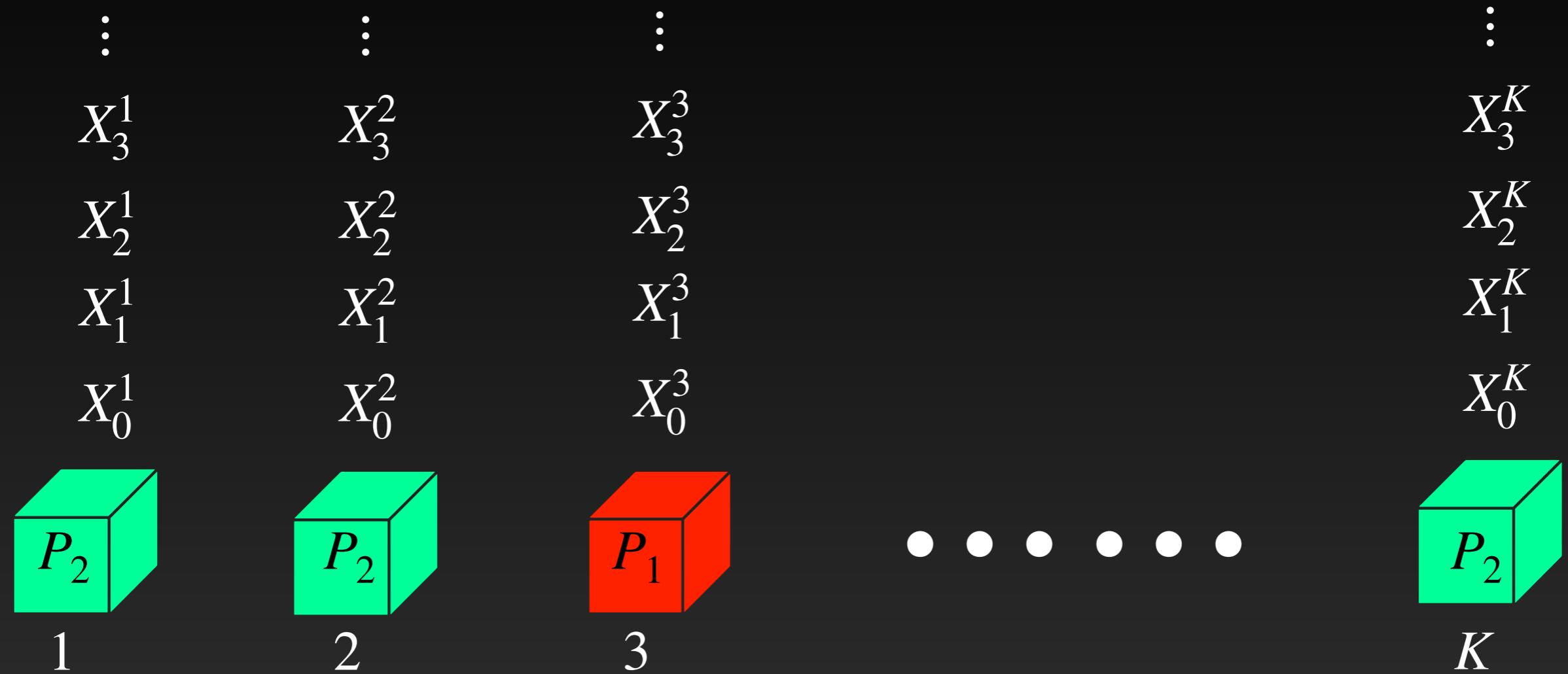
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3	X_0^K	X_1^3	X_1^K

Markov Observations: Rested Case



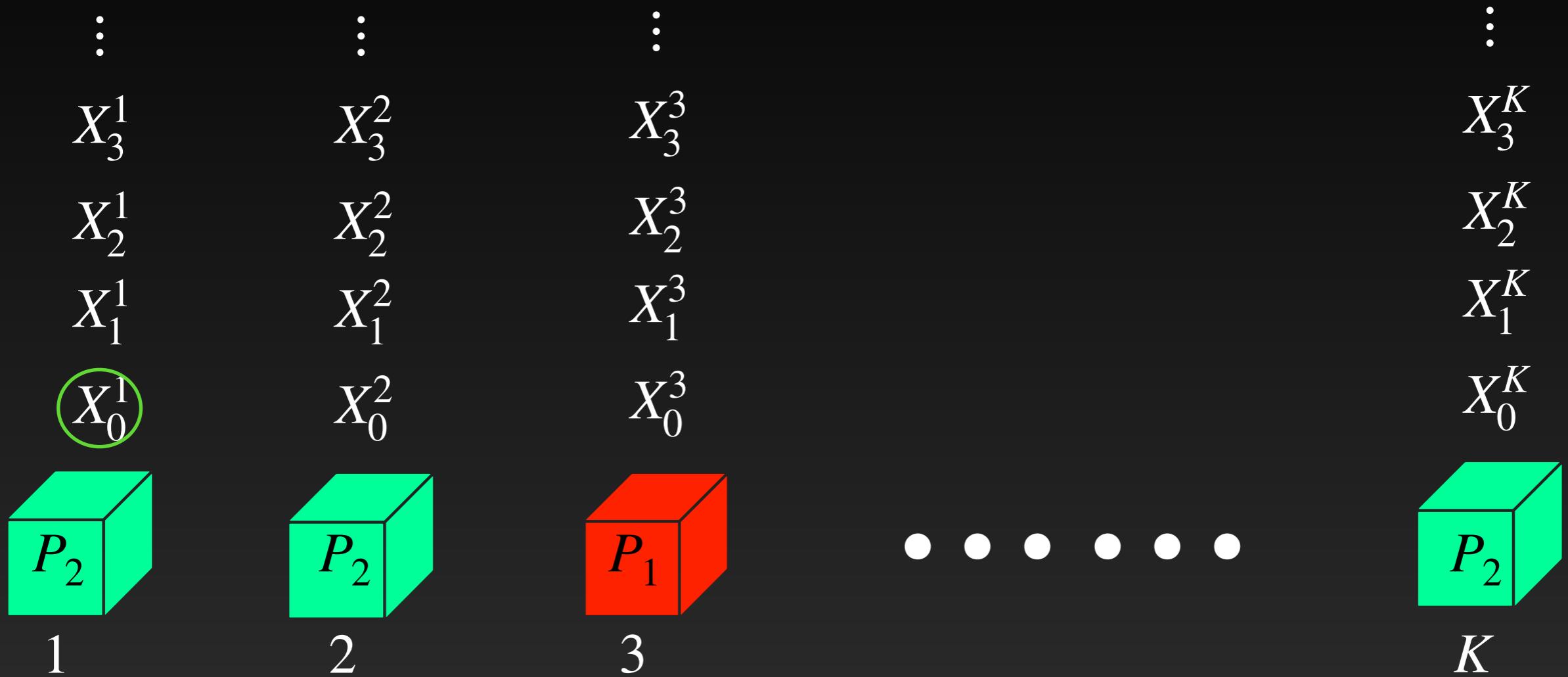
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_0^2	X_0^3	X_0^K	X_1^K	X_1^K

Markov Observations: Restless Case



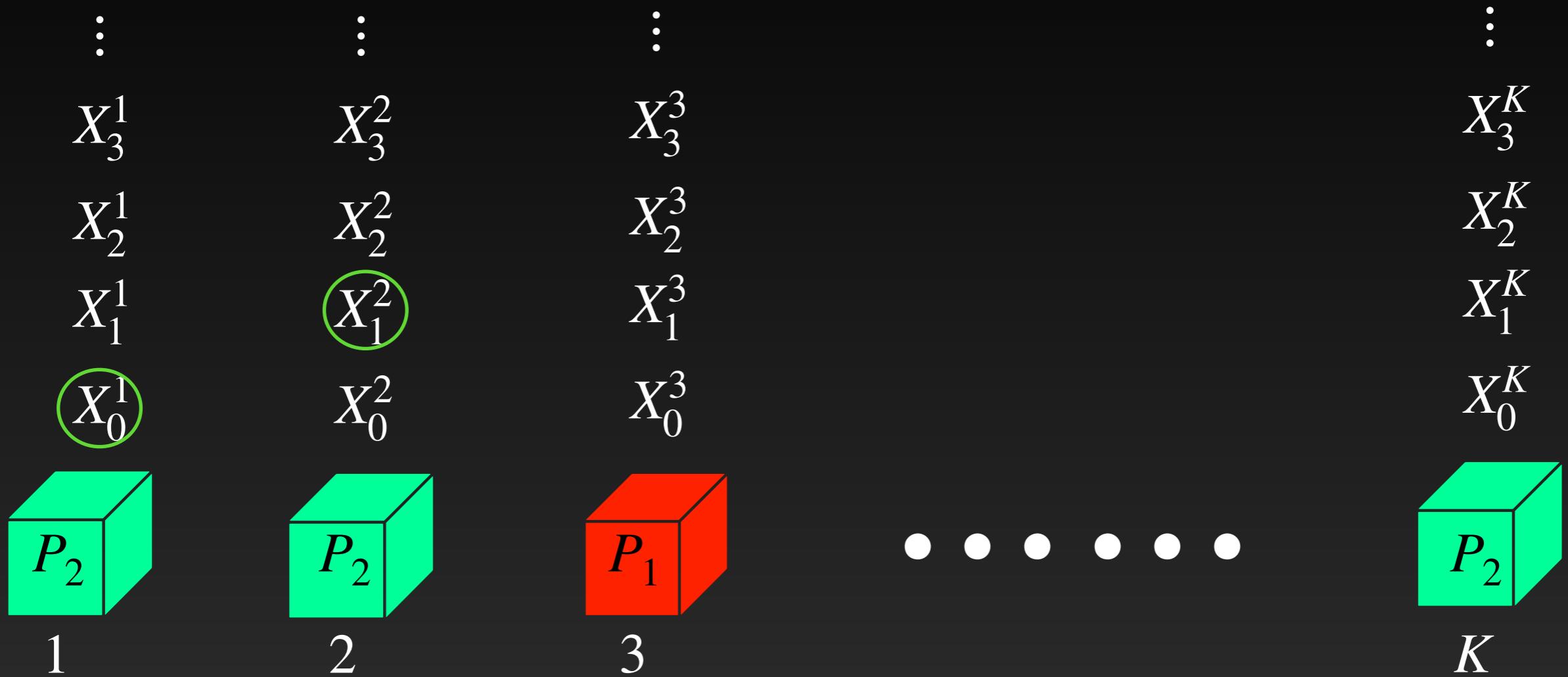
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs						

Markov Observations: Restless Case



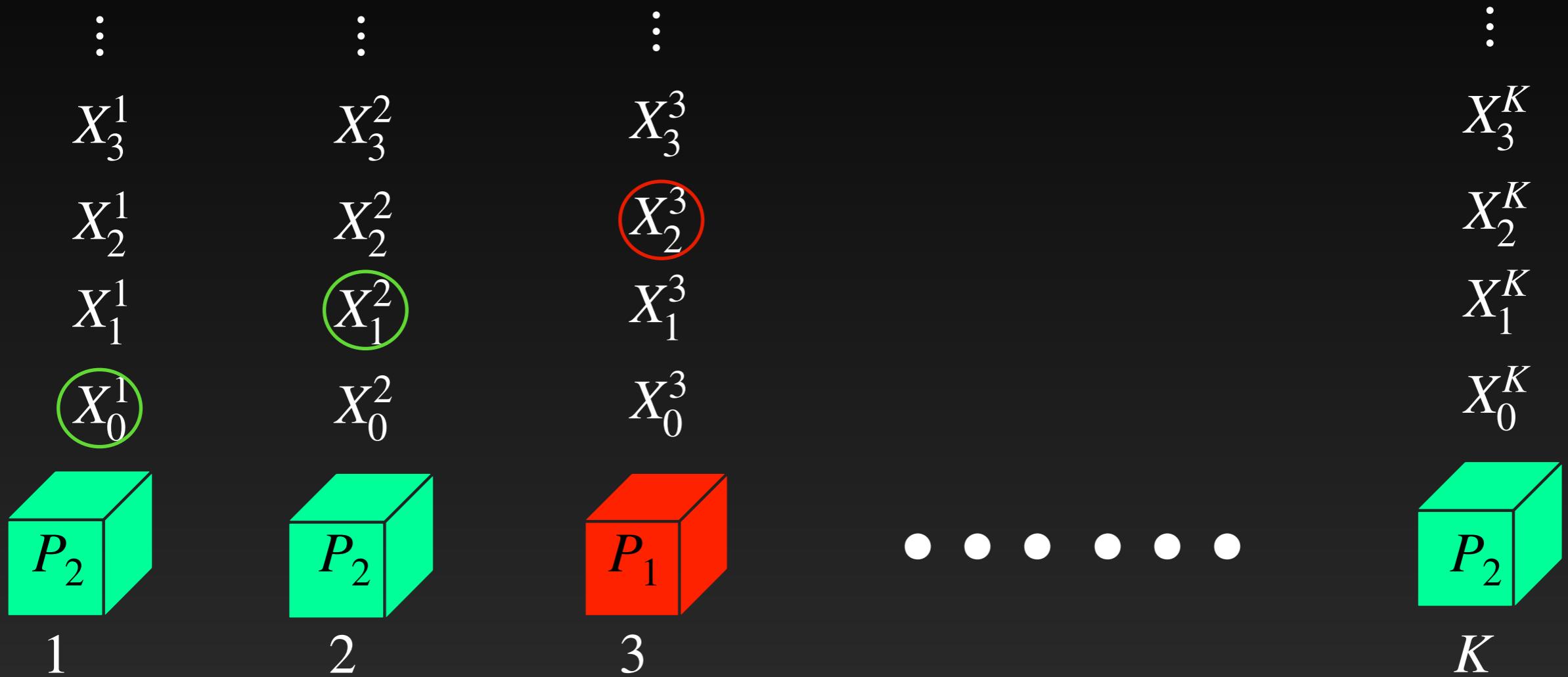
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1					

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

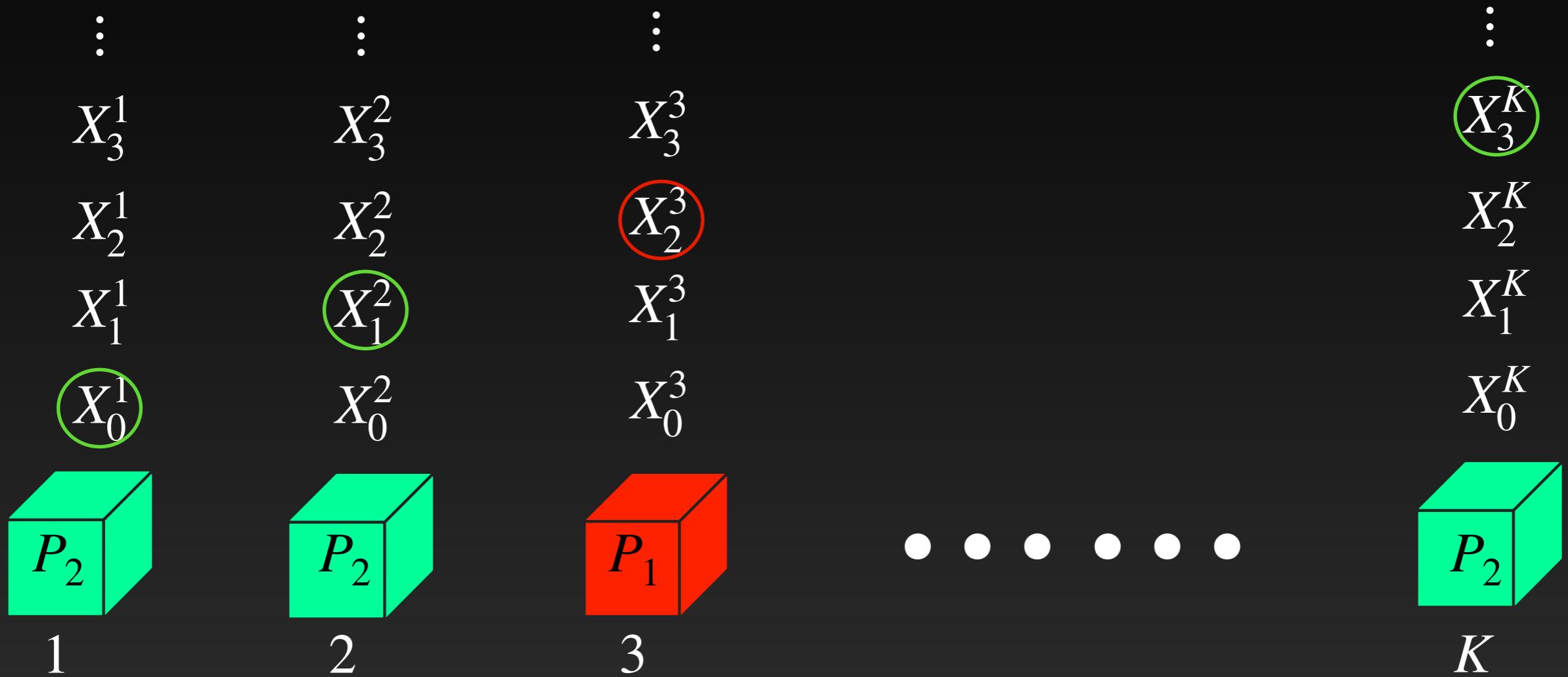
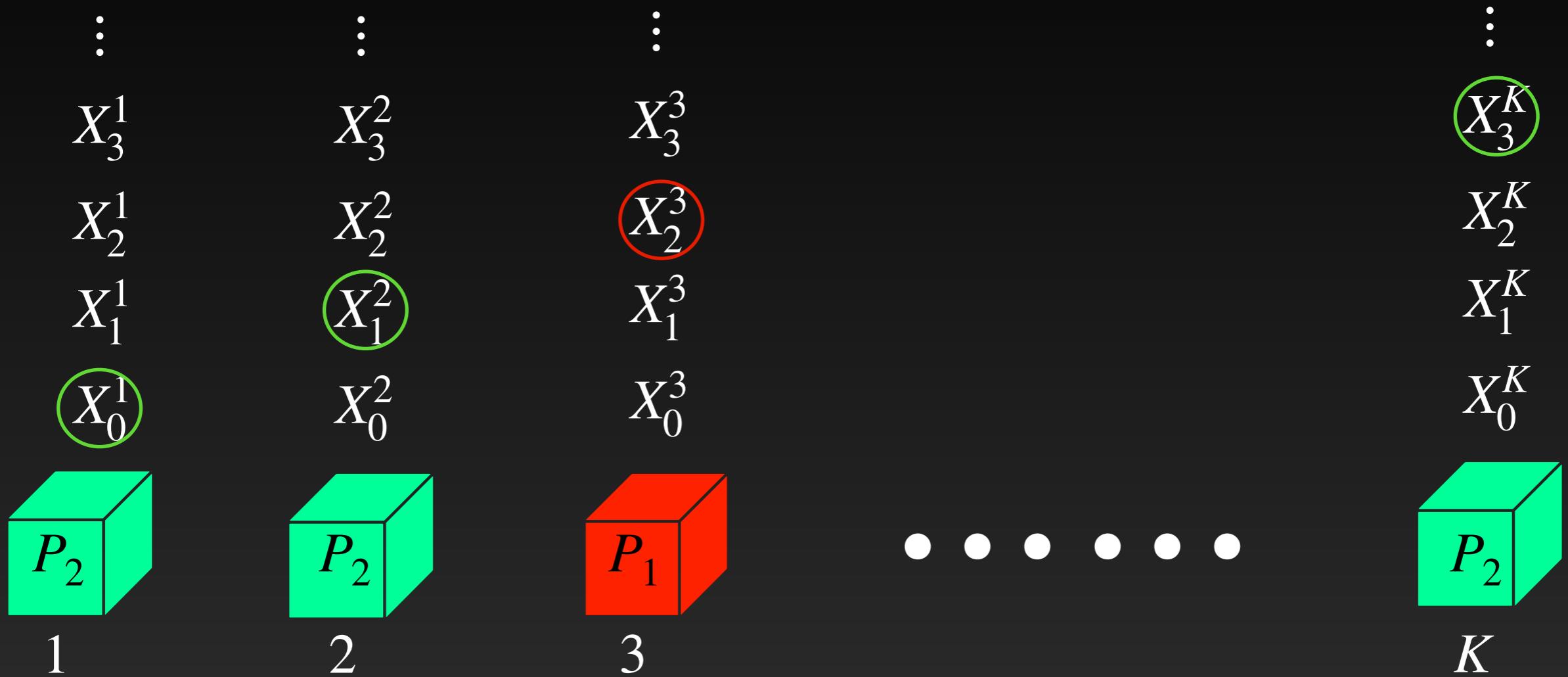


Table showing the sequence of observations for each arm over time steps $t = 0, 1, 2, 3, 4, 5$. The table has columns for t (0 to 5), Arm (1 to K), and Obs (X_0^1 to X_3^K). Red circles highlight $X_0^1, X_1^2, \text{ and } X_2^3$, while green circles highlight X_3^K .

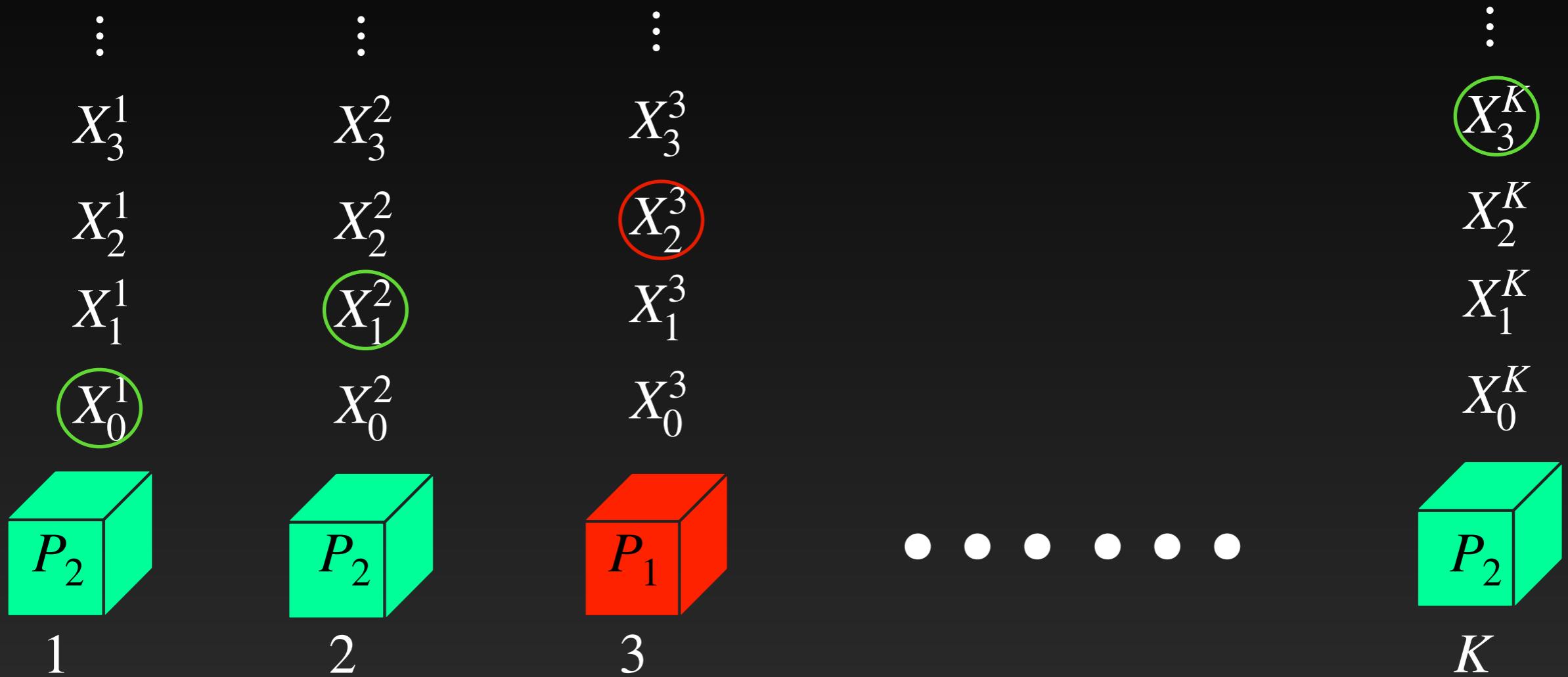
t	0	1	2	3	4	5
Arm	1	2	3	K	3	K
Obs	X_0^1	X_1^2	$\circled{X_2^3}$	$\circled{X_3^K}$		

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

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

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

Exact in the asymptotic limit
as **error** vanishes

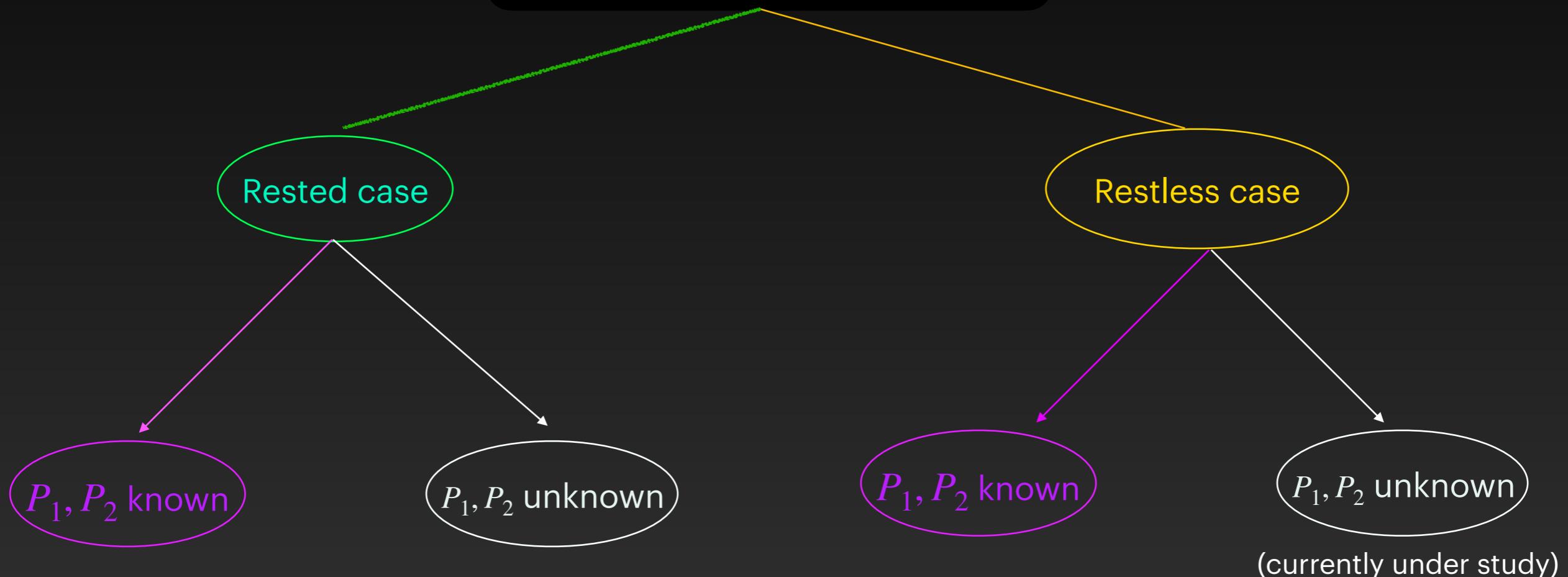
$$\text{Time to identify the anomaly} \approx \log \frac{1}{\text{Error in reporting the anomaly correctly}} = \frac{1}{D^*}$$

YES!

We characterise D^* explicitly

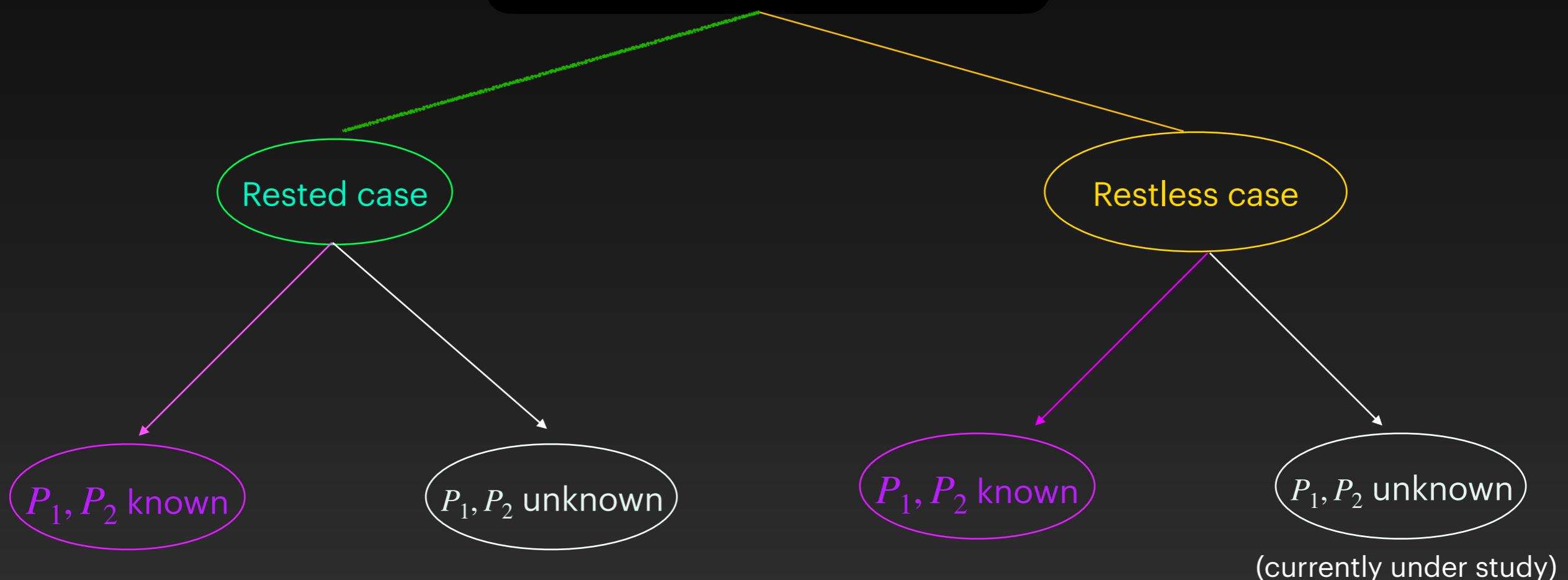
Organisation

Odd arm identification with
Markov observations



Organisation

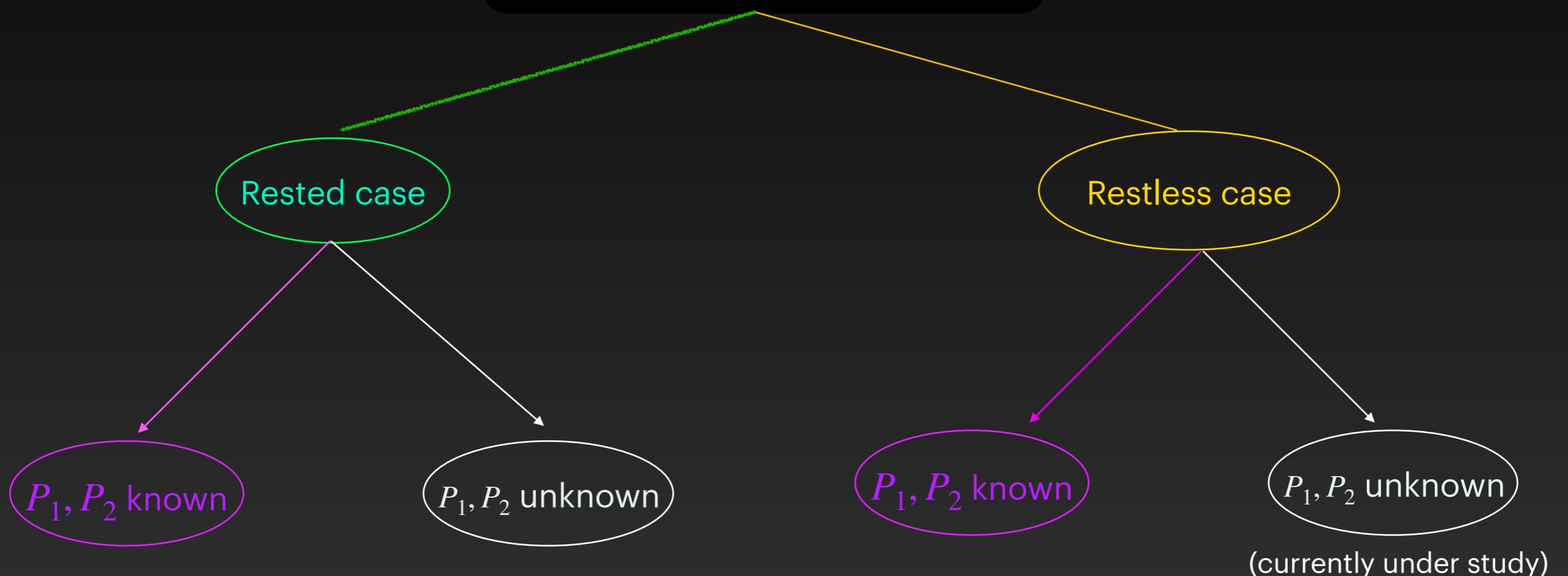
Odd arm identification with
Markov observations



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

Organisation

Odd arm identification with
Markov observations



$$\log \frac{1}{\text{Error in reporting the anomaly correctly}} \cdot \frac{1}{D^*} \stackrel{\substack{\text{Data processing inequality} \\ \text{Asymptotic lower bound}}}{\lesssim}$$

Time to identify
the anomaly

$$\stackrel{\substack{\text{Policy design} \\ \text{Asymptotic upper bound}}}{\lesssim}$$

$$\log \frac{1}{\text{Error in reporting the anomaly correctly}} \cdot \frac{1}{D^*}$$

Key Ideas and Challenges

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

When P_1 and P_2 are known:

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

Key Ideas and Challenges

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

When P_1 and P_2 are known:

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

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:

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

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:

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

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 $\eta > 0$

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:

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

Supremum is attained for some λ_h^*

Keep an estimate $\hat{h}(t)$ of h at time t , and sample according to

$$\frac{\eta}{K} + (1 - \eta) \lambda_{\hat{h}(t)}^*(\cdot)$$

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

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:

$$D_{h,P_1,P_2}^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \dots \dots \dots$$

Key Ideas and Challenges

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

When P_1 and P_2 are unknown:

$$D_{h,P_1,P_2}^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \dots \dots \dots$$

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:

$$D_{h,P_1,P_2}^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \dots \dots \dots$$

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

Keep estimates $(\hat{h}(t), \hat{P}_1^t, \hat{P}_2^t)$ of (h, P_1, P_2) at time t , and sample according to $\lambda_{\hat{h}(t), \hat{P}_1^t, \hat{P}_2^t}^*$
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:

$$D_{h,P_1,P_2}^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \dots \dots \dots$$

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

Keep estimates $(\hat{h}(t), \hat{P}_1^t, \hat{P}_2^t)$ of (h, P_1, P_2) at time t , and sample according to $\lambda_{\hat{h}(t), \hat{P}_1^t, \hat{P}_2^t}^*$ 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:

$$D_{h,P_1,P_2}^* = \sup_{\lambda(\cdot)} \min_{h' \neq h} \dots \dots \dots$$

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

Keep estimates $(\hat{h}(t), \hat{P}_1^t, \hat{P}_2^t)$ of (h, P_1, P_2) at time t , and sample according to $\lambda_{\hat{h}(t), \hat{P}_1^t, \hat{P}_2^t}^*$ with η -trembling

$$(\hat{h}(t), \hat{P}_1^t, \hat{P}_2^t) \rightarrow (h, P_1, P_2) \text{ as } t \rightarrow \infty$$

Principle of certainty equivalence

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:

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

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:

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

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:

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

Not clear if the supremum is attained

Supremum can be approached δ – 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:

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

Not clear if the supremum is attained

Supremum can be approached δ – 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!