THE UNIVERSITY OF CALIFORNIA UCLA

香港中文大學
The Chinese University of Hong Kong

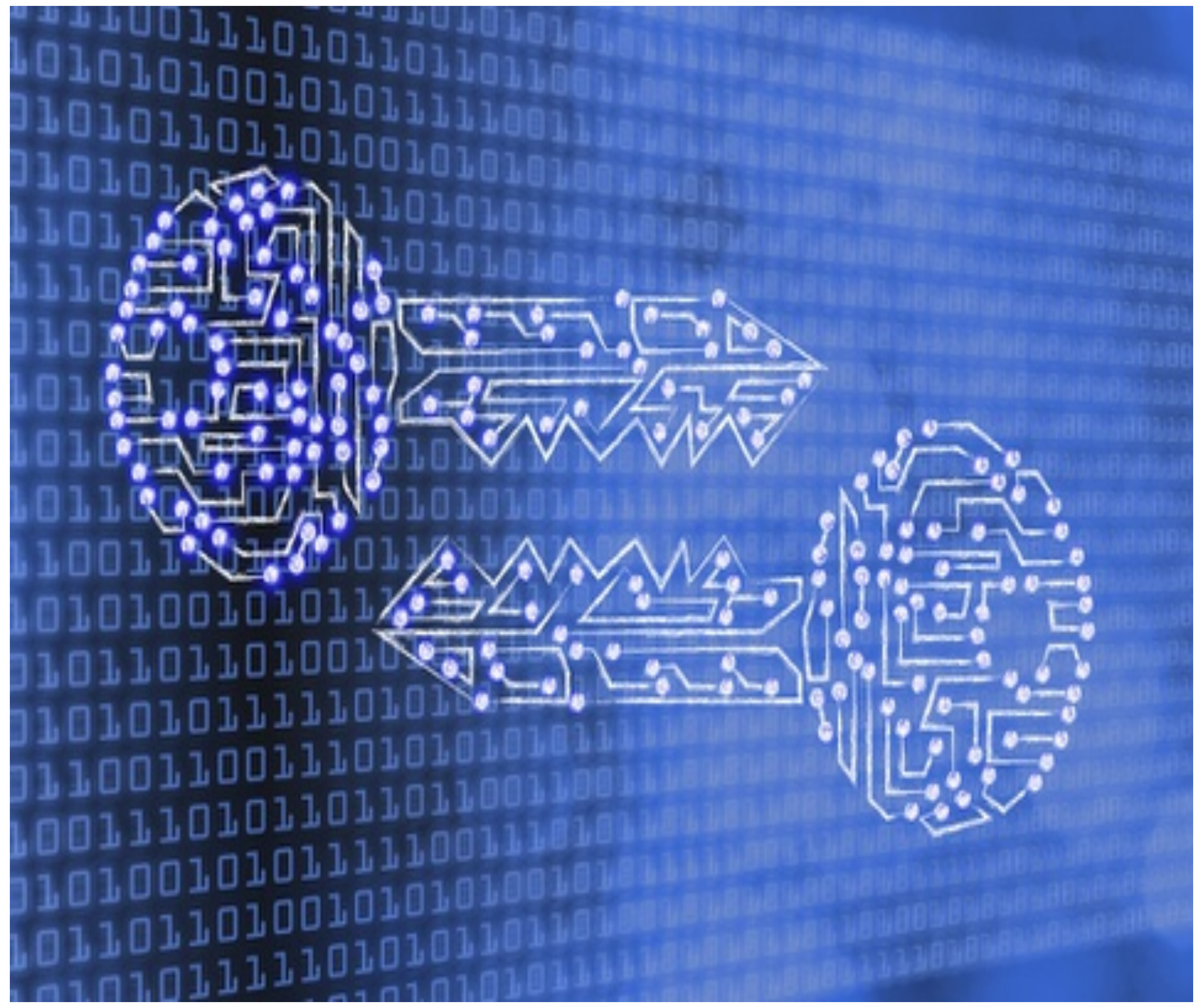
On Task Aware Compression: Common Information Dimension and Contextual Bandit Learning

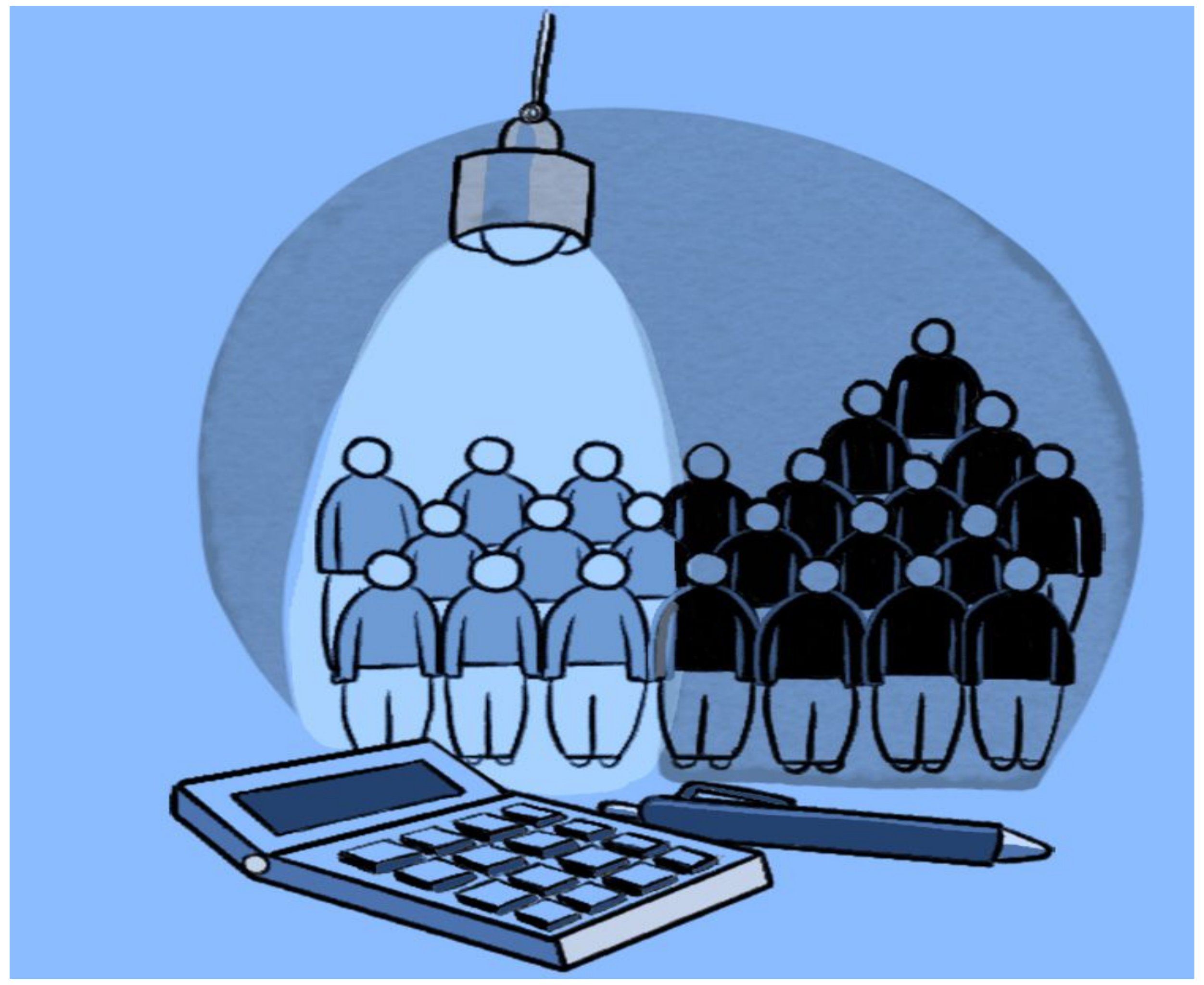
Osama A. Hanna, Christina Fragouli

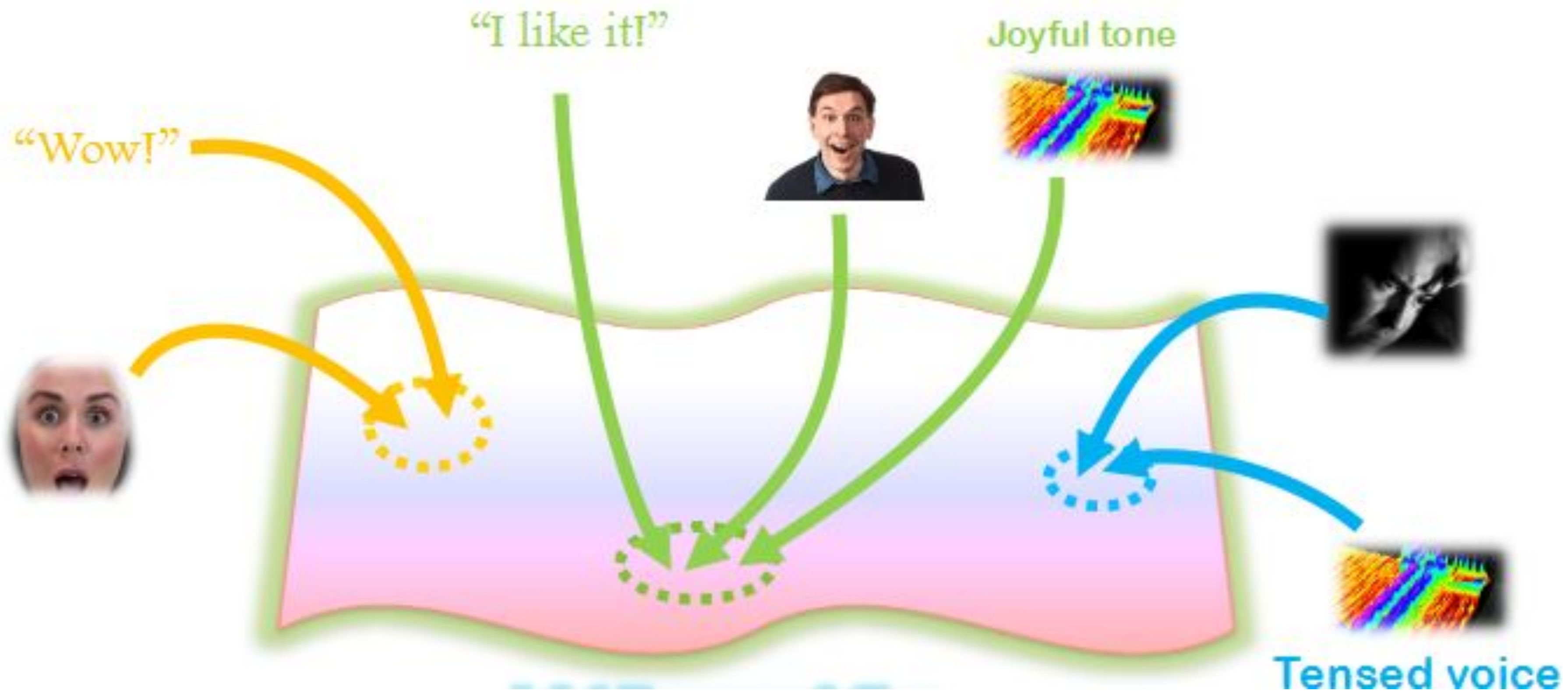
01
Common Information Dimension

Hanna, Osama, Xinlin Li, Suhas Diggavi, and Christina Fragouli. "Common Information Dimension." ISIT 2023.

Applications







Key generation in Cryptography

Hypothesis testing

Multi-modal representation learning

Common Information: Wyner

$$C_{\text{Wyner}}(X_1, X_2) := \min_{P_W P_{X_1|W} P_{X_2|W} : P_{X_1 X_2} = \pi_{X_1 X_2}} I(X_1, X_2; W)$$

X_1, X_2 : random vectors (sources)

W : common randomness

A. Wyner, "The common information of two dependent random variables," IEEE Transactions on Information Theory, vol. 21, no. 2, pp. 163–179, 1975.

Common Information: Wyner

$$C_{\text{Wyner}}(X_1, X_2) := \min_{P_W P_{X_1|W} P_{X_2|W} : P_{X_1 X_2} = \pi_{X_1 X_2}} I(X_1, X_2; W)$$

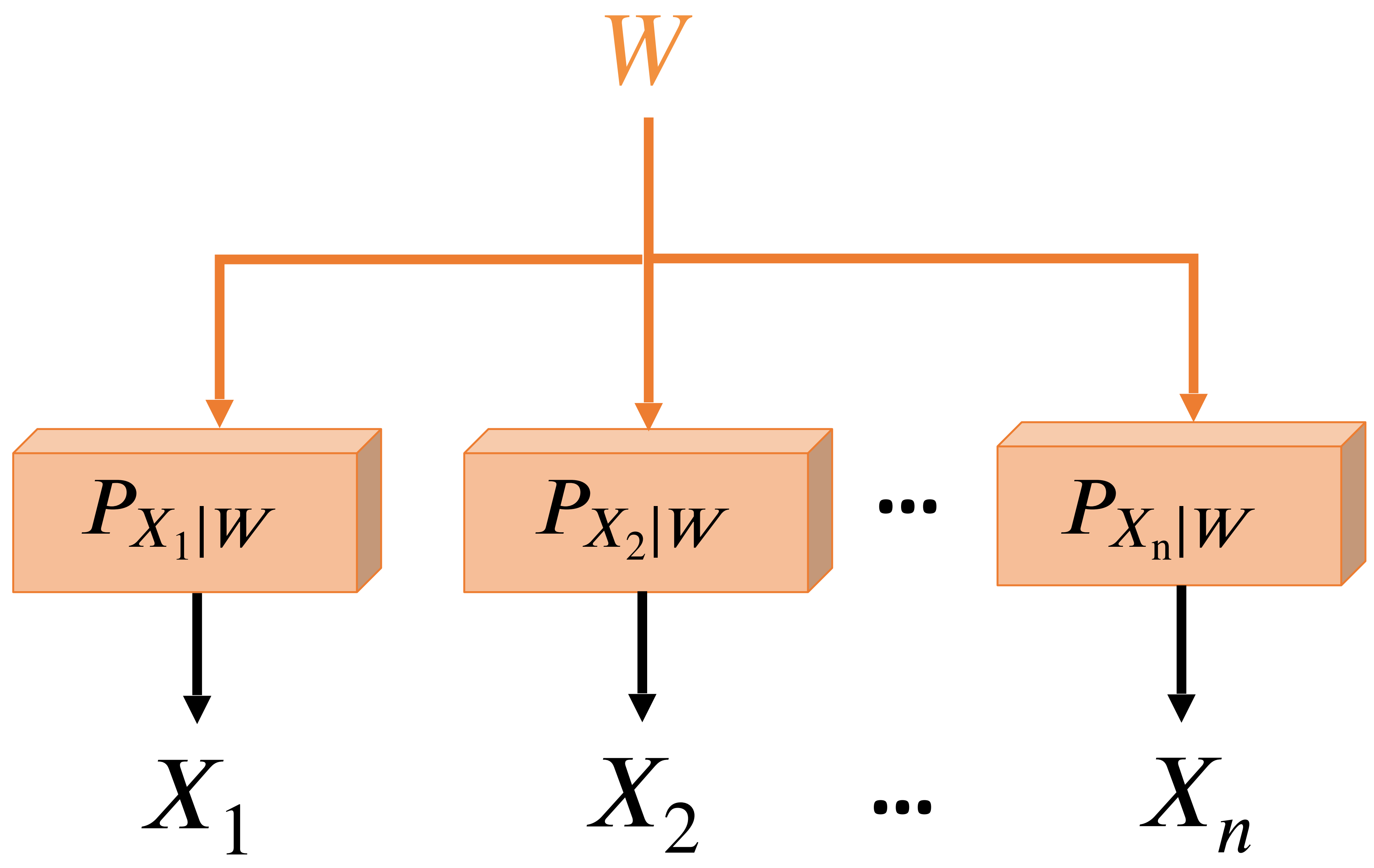
- Can be generalized to n sources

A. Wyner, "The common information of two dependent random variables," IEEE Transactions on Information Theory, vol. 21, no. 2, pp. 163–179, 1975.

Common Information: Wyner

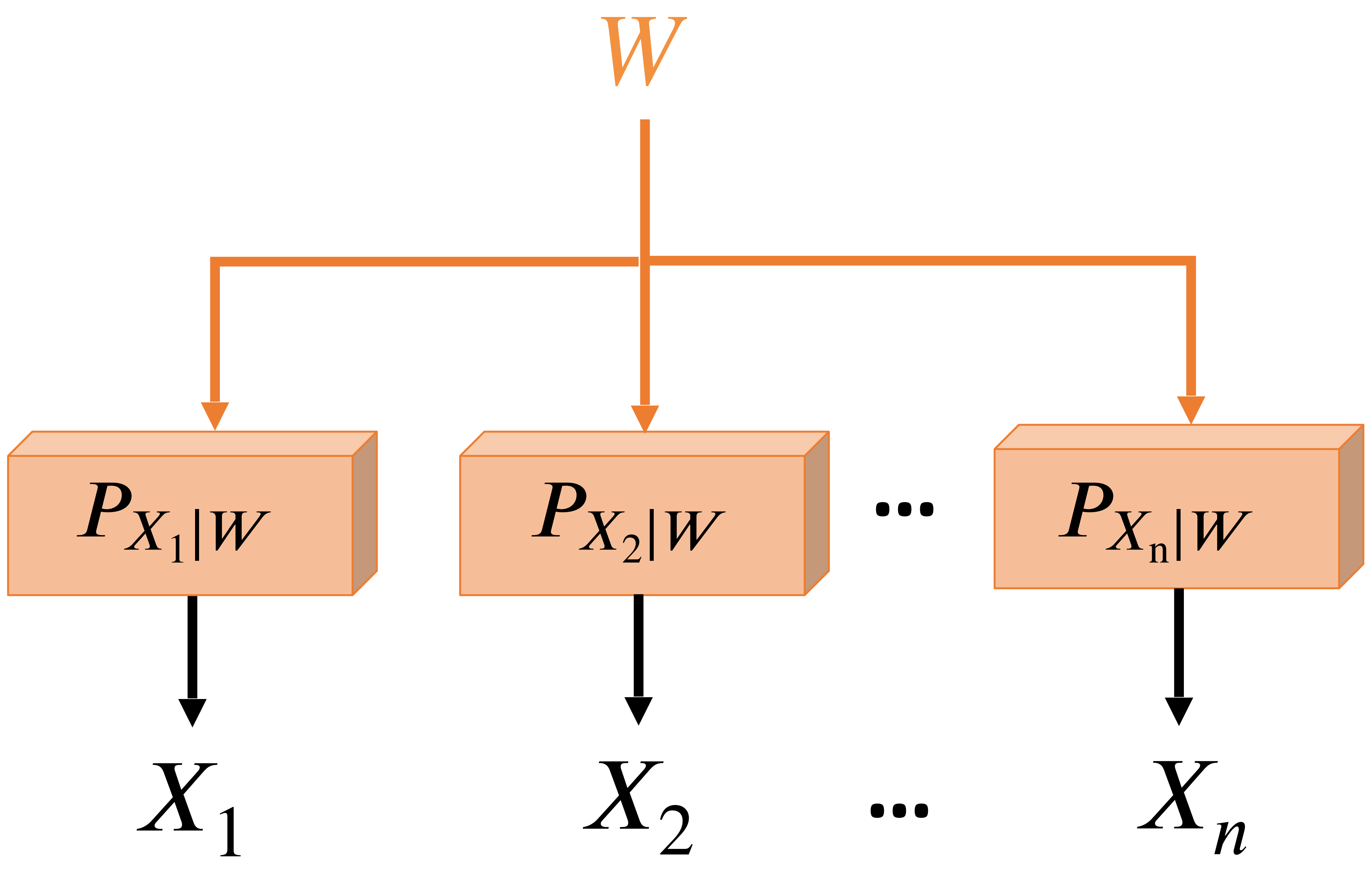
$$C_{\text{Wyner}}(X_1, X_2) := \min_{P_W P_{X_1|W} P_{X_2|W} : P_{X_1 X_2} = \pi_{X_1 X_2}} I(X_1, X_2; W)$$

- Can be generalized to n sources
- Multiple interpretations, e.g.,
  distributed simulation



A. Wyner, "The common information of two dependent random variables," IEEE Transactions on Information Theory, vol. 21, no. 2, pp. 163–179, 1975.

Common Information: Common Entropy

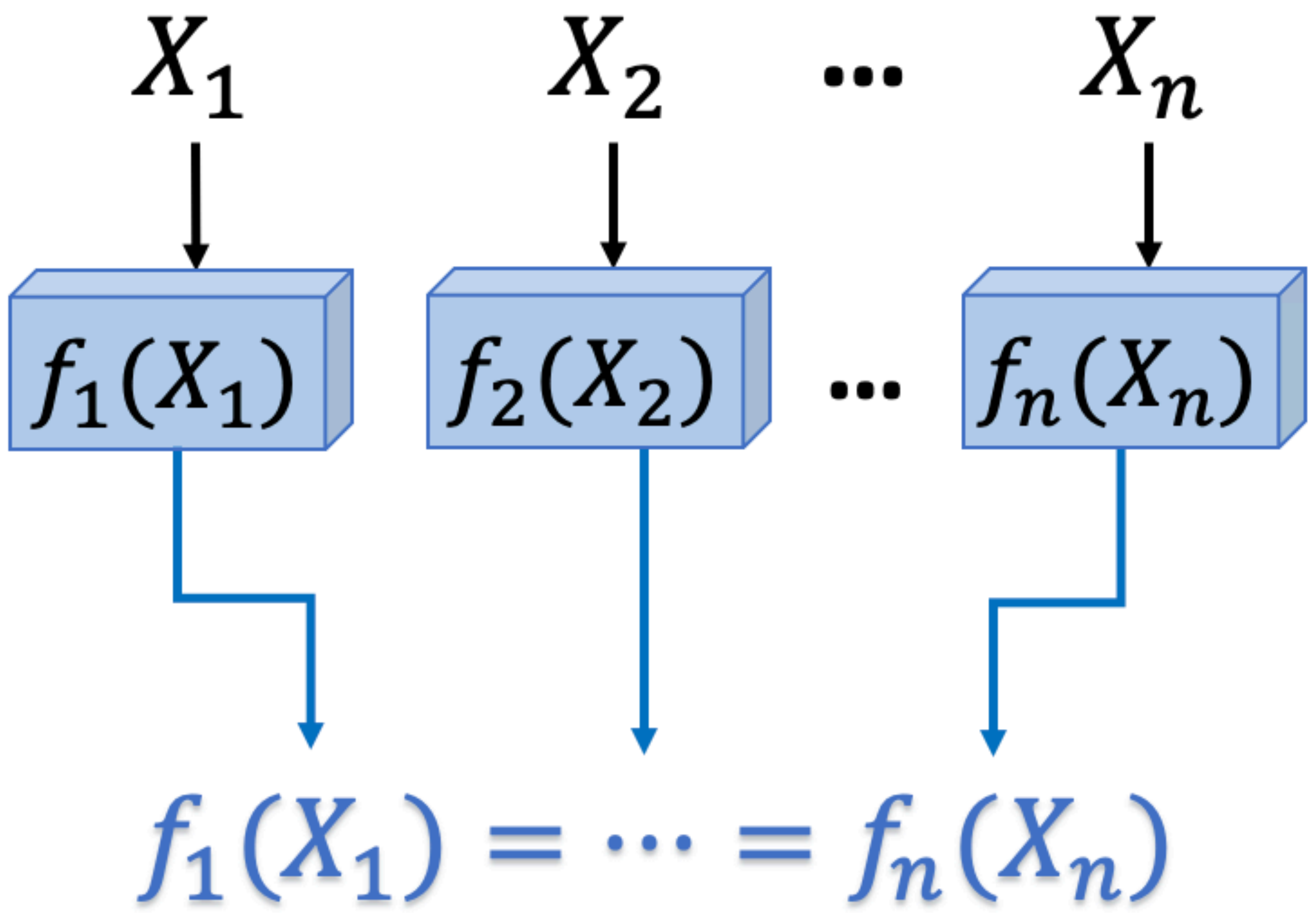


- one shot
- exact distributed simulation

$$C_{\text{Exact}}(X_1, X_2) := \min_{P_W P_{X_1|W} P_{X_2|W} : P_{X_1 X_2} = \pi_{X_1 X_2}} H(W)$$

G. R. Kumar, C. T. Li, and A. El Gamal, "Exact common information," in 2014 IEEE International Symposium on Information Theory. IEEE, 2014, pp. 161–165.

Common Information: Gacs-Korner



- distributed randomness extraction

$$C_{\mathrm{GK}}(X_1, X_2) := \max_{f,g: f(X_1)=g(X_2)} H(f(X_1))$$

P. Gács and J. Körner, "Common information is far less than mutual information," Problems of Control and Information Theory, vol. 2, no. 2, pp. 149–162, 1973.

Common Information Can be Infinite?

- $X_1, X_2 \in \mathbb{R}$, $X_1 \sim \mathcal{N}(0,1)$ and $X_1 = X_2$ almost surely
  - $C(X_1, X_2) = \infty$

Common Information Can be Infinite?

- $X_1, X_2 \in \mathbb{R}$, $X_1 \sim \mathcal{N}(0,1)$ and $X_1 = X_2$ almost surely
  - $C(X_1, X_2) = \infty$

- What if $X_1, X_2 \in \mathbb{R}^{100}$?
  - $C(X_1, X_2) = \infty$

Common Information Can be Infinite?

- $X_1, X_2 \in \mathbb{R}$, $X_1 \sim \mathcal{N}(0,1)$ and $X_1 = X_2$ almost surely
  - $C(X_1, X_2) = \infty$

- What if $X_1, X_2 \in \mathbb{R}^{100}$?
  - $C(X_1, X_2) = \infty$

How to measure the different complexities for the above cases?

Common Information Dimension (CID)

(First Attempt):

$$d(X_1, X_2) = \min\{d_W \mid W \in \mathscr{W}\}$$

$$\mathscr{W} = \{W \in \mathbb{R}^{d_W} \mid \exists g : (X_1, X_2) \mapsto W, \quad X_1 \perp\!\!\!\perp X_2 \mid W\}$$

X_1, X_2 : random vectors

W : common randomness (vector)

d_W : #coordinates

$X_1 \perp\!\!\!\perp X_2 \mid W$: X_1, X_2 conditionally independent given W

Common Information Dimension (CID)

(First Attempt):

$$d(X_1, X_2) = \min\{d_W \mid W \in \mathscr{W}\}$$

$$\mathscr{W} = \{W \in \mathbb{R}^{d_W} \mid \exists g : (X_1, X_2) \mapsto W, \quad X_1 \perp\!\!\!\perp X_2 \mid W\}$$

X_1, X_2, W : vectors

d_W : #coordinates

$X_1 \perp\!\!\!\perp X_2 \mid W$: X_1, X_2 conditionally independent given W

Issue: $\exists f : \mathbb{R} \leftrightarrow \mathbb{R}^n,\ d(X_1, X_2) = 1$ ⚠️

Common Information Dimension (CID)

$$d_{\mathscr{F}}(X_1, X_2) = \min\{d_W \,|\, W \in \mathscr{W}_{\mathscr{F}}\}$$

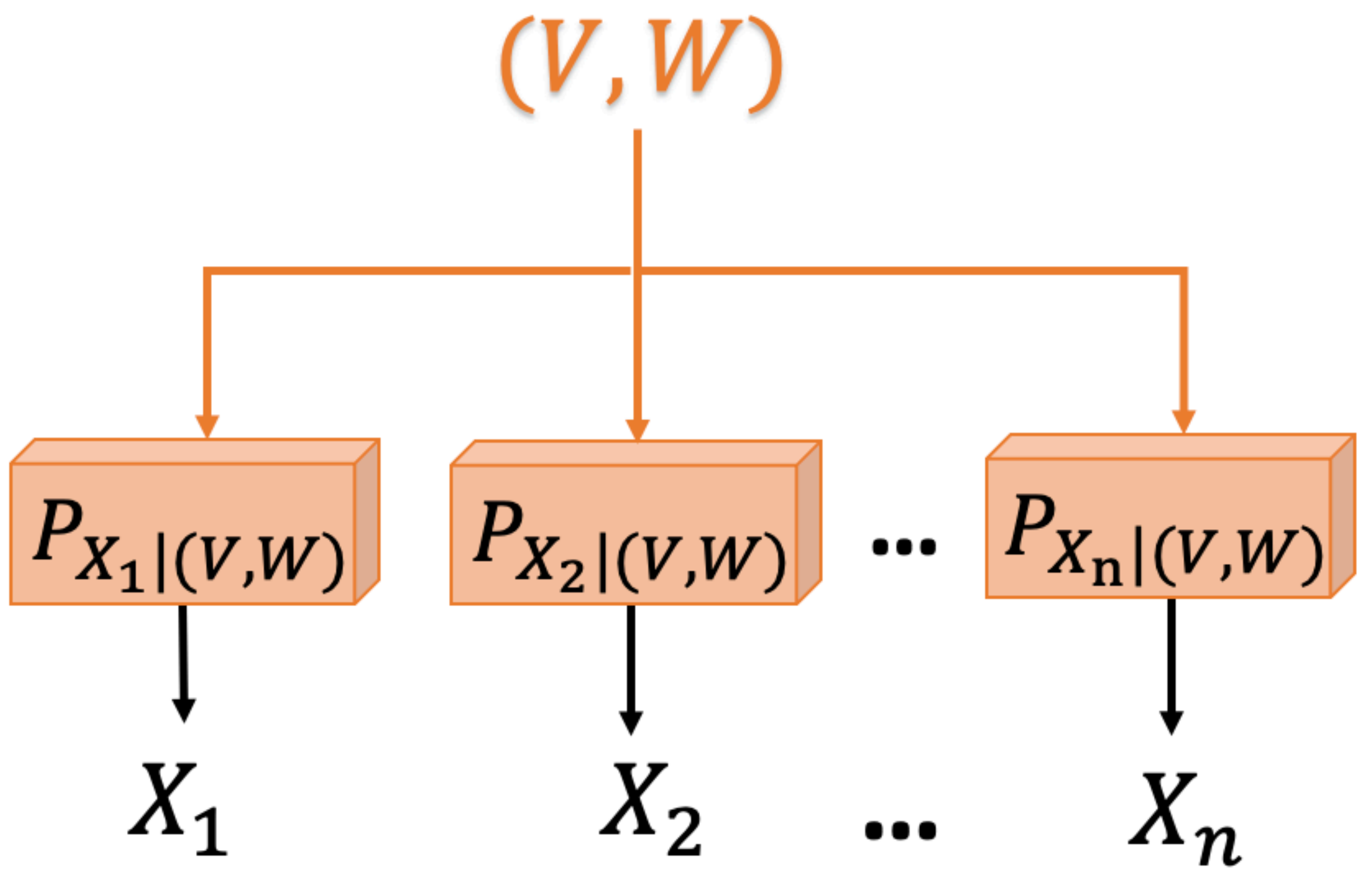
$$\mathscr{W}_{\mathscr{F}} = \{W |\ \exists g : (X_1, X_2) \mapsto W, \quad X_1 \perp\!\!\!\perp X_2 \,|\, W, \quad g \in \mathscr{F}\}$$

X_1, X_2, W : vectors

d_W : #coordinates

$X_1 \perp\!\!\!\perp X_2 \,|\, W$: X_1, X_2 conditionally independent given W

Common Information Dimension (CID)



Definition (CID):

$$d_{\mathscr{F}}(X_1, X_2) = \min\{d_W \,|\, W \in \mathscr{W}_{\mathscr{F}}\}$$

$$\mathscr{W}_{\mathscr{F}} = \{W \,|\, \exists V, g : (X_1, X_2) \mapsto W, \quad X_1 \perp\!\!\!\perp X_2 \,|\, (V, W), \quad g \in \mathscr{F}, \quad H(V) < \infty\}$$

d_W : #coordinates

Renyi Common Information Dimension (RCID)

Rényi Dimension:

$$d^R(W) = \lim_{m \to \infty} \frac{H(\langle W \rangle_m)}{\log m}$$

W : vector, $\langle W_i \rangle_m = \frac{\lfloor m W_i \rfloor}{m}$

Renyi Common Information Dimension (RCID)

Rényi Dimension:

$$d^R(W) = \lim_{m\to\infty} \frac{H(\langle W\rangle_m)}{\log m}$$

$W:$ vector, $\langle W_i\rangle_m = \dfrac{\lfloor mW_i\rfloor}{m}$

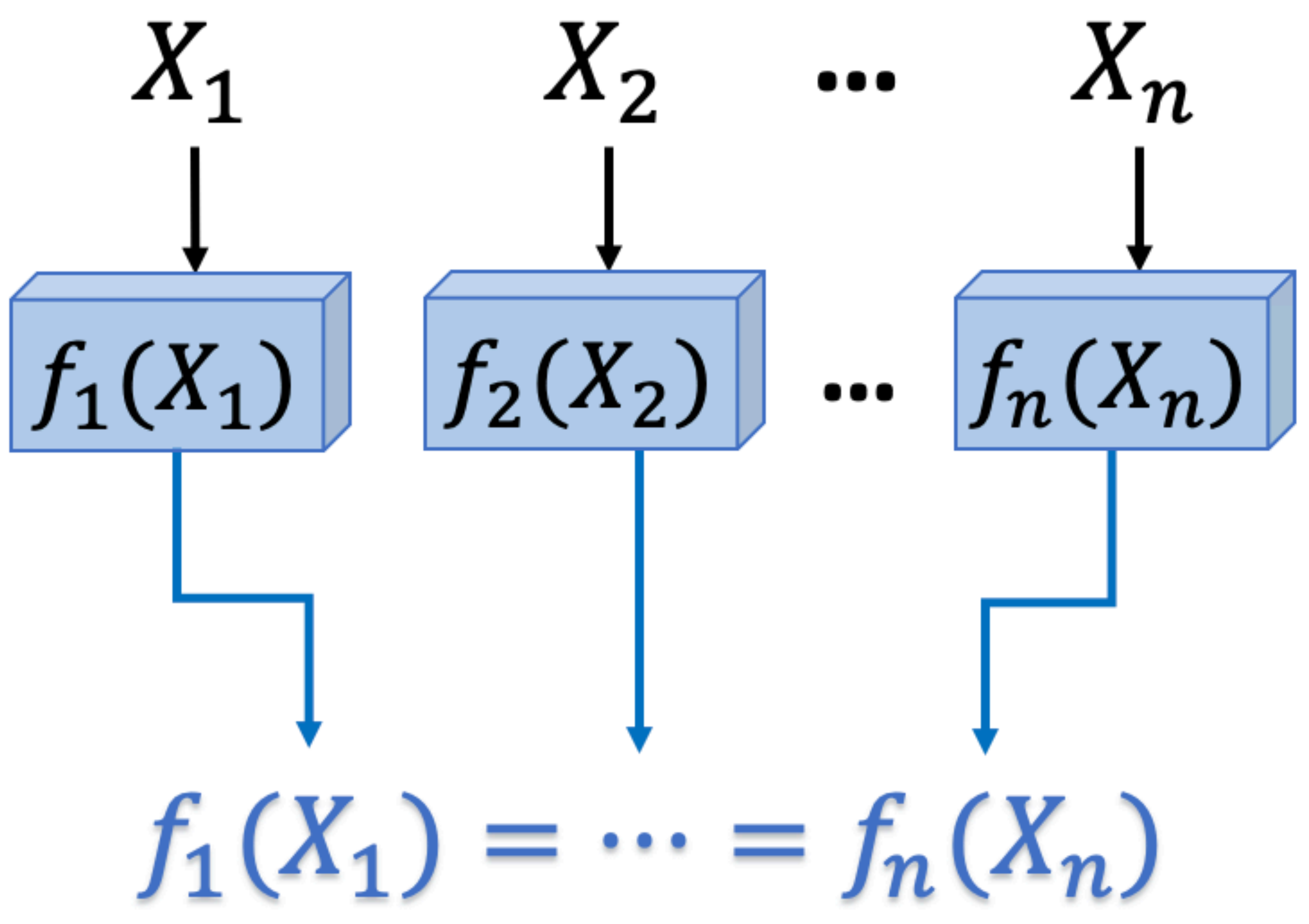
Definition (RCID):

$$d_{\mathscr{F}}(X_1, X_2) = \min\{d^R(W) \,|\, W \in \mathscr{W}_{\mathscr{F}}\}$$

Gacs-Korner Common Information Dimension (GKCID)

Definition (GKCID):

$$d_{\mathscr{F}}^{GK}(X_1, X_2) = \sup_{W=f_1(X_1)=f_2(X_2), f_i \in \mathscr{F}} d^R(W)$$



Gacs-Korner Common Information Dimension (GKCID)

Definition (GKCID):

$$d_{\mathscr{F}}^{GK}(X_1, X_2) = \sup_{W=f_1(X_1)=f_2(X_2), f_i \in \mathscr{F}} d^R(W)$$

Question: How to compute CID, RCID, GKCID?

CID for Gaussian Sources

Assumptions

- $X_1, X_2 \sim \mathcal{N}(\mu, \Sigma_{X_1,X_2})$
- $\mathcal{F} = \{f : \mathbb{R}^{d_{X_1}+d_{X_2}} \to d_W \,|\, f(X_1, X_2) = A[X_1 \;\; X_2]^\top$ for some matrix $A\}$

$X_1, X_2, W :$ vectors

CID for Two Gaussian Sources

Theorem:

If $[X_1, X_2]$ is a jointly Gaussian random vector, $\mathscr{F}$ is the class of linear function

$$d_{\mathscr{F}}(X_1, X_2) = \text{rank}(\Sigma_{X_1}) + \text{rank}(\Sigma_{X_2}) - \text{rank}(\Sigma_{X_1, X_2})$$

X_1, X_2 : vectors

$$d_{\mathscr{F}}(X_1, X_2) = \min\{d_W \mid W \in \mathscr{W}_{\mathscr{F}}\}$$

CID for Two Gaussian Sources: proof sketch

- WLOG assume $\Sigma_{X_1}, \Sigma_{X_2}$ are full rank

- $a^\top X_1 + b^\top X_2 = 0$ almost surely $\iff$ $[a^\top b^\top]\Sigma_{X_1,X_2} = 0$

- Find the null space of Σ_{X_1,X_2}, namely $N = [N_{X_1}\ N_{X_2}]$ with $N\Sigma_{X_1,X_2} = 0$
  - $N_{X_1}X_1 = -N_{X_2}X_2$ almost surely

CID for Two Gaussian Sources: proof sketch

Achievability

- Σ_{X_1,X_2} is full rank $\implies \quad d_{\mathcal{F}}(X_1, X_2) = 0$
- Recall: $N_1 = -N_2$

$N_1 = N_{X_1}X_1, N_2 = N_{X_2}X_2$

- Conditioned on $W = N_1$, $[X_1, X_2]$ effectively has full rank covariance matrix
- CID $\leq d_{N_1}$

C. T. Li and A. El Gamal, "Distributed simulation of continuous random variables," IEEE Transactions on Information Theory, vol. 63, no. 10, pp. 6329–6343, 2017.

CID for Two Gaussian Sources: proof sketch

Converse

- N_1 is a deterministic function of every $(V, W) : X_1 \perp\!\!\!\perp X_2 \,|\, (V, W)$
- N_1 can be obtained from W by a linear transformation

CID for Two Gaussian Sources: proof sketch

Converse

- N_1 is a deterministic function of every $(V, W) : X_1 \perp\!\!\!\perp X_2 \,|\, (V, W)$
- N_1 can be obtained from W by a linear transformation
- N_1 has full rank covariance matrix
- $d_W \geq$ #rows of N_{X_1}

RCID, GKCID for Two Gaussian Sources

Theorem:

If $[X_1, X_2]$ is a jointly Gaussian random vector, $\mathscr{F}$ is the class of linear function

$$d_{\mathscr{F}}(X_1, X_2) = d^R_{\mathscr{F}}(X_1, X_2) = d^{GK}_{\mathscr{F}}(X_1, X_2)$$

X_1, X_2 : vectors

$$d^R_{\mathscr{F}}(X_1, X_2) = \min\{d^R(W) \,|\, W \in \mathscr{W}_{\mathscr{F}}\}$$

$$d^{GK}_{\mathscr{F}}(X_1, X_2) = \sup_{W=f_1(X_1)=f_2(X_2), f_i \in \mathscr{F}} d^R(W)$$

CID for N Gaussian Sources

Theorem:

If $[X_1, \cdots, X_n]$ is a jointly Gaussian random vector, $\mathscr{F}$ is class of linear function

$$d_{\mathscr{F}}(X_1, \cdots, X_n) = \sum_{i=1}^{n} \text{rank}(\Sigma_{-i}) - (n-1)\text{rank}(\Sigma)$$

$$d_{\mathscr{F}}(X_1, \cdots, X_n) = d^{R}_{\mathscr{F}}(X_1, \cdots, X_n) \geq d^{GK}_{\mathscr{F}}(X_1, \cdots, X_n)$$

CID for N Gaussian Sources: proof sketch

Achievability

Converse

CID for N Gaussian Sources: proof sketch

Achievability

- Find $Z = [Z_1, \cdots, Z_n]$ s.t. conditioned on Z, the covariance matrix of X is effectively full rank
  - Intuitively: Z_i captures the information that X_i contains about $X_{i+1}, \cdots, X_n$ which $X_1, \cdots, X_{i-1}$ do not contain

Converse

CID for N Gaussian Sources: proof sketch

Achievability

- Find $Z = [Z_1, \cdots, Z_n]$ s.t. conditioned on Z, the covariance matrix of X is effectively full rank
  - Intuitively: Z_i captures the information that X_i contains about $X_{i+1}, \cdots, X_n$ which $X_1, \cdots, X_{i-1}$ do not contain

Converse

- Z is deterministic function of every $(V, W) : X_1 \perp\!\!\!\perp \cdots \perp\!\!\!\perp X_n \,|\, (V, W)$

CID for N Gaussian Sources

Theorem:

If $[X_1, \cdots, X_n]$ is a jointly Gaussian random vector, $\mathscr{F}$ is class of linear function

$$d_{\mathscr{F}}(X_1, \cdots, X_n) = \sum_{i=1}^{n} \text{rank}(\Sigma_{-i}) - (n-1)\text{rank}(\Sigma)$$

$$d_{\mathscr{F}}(X_1, \cdots, X_n) = d_{\mathscr{F}}^{R}(X_1, \cdots, X_n) \geq d_{\mathscr{F}}^{GK}(X_1, \cdots, X_n)$$

Example with GKCID < CID

- $X_1, X_2, X_3 \sim \mathcal{N}(0,1)$ and $X_1 = X_2$ almost surely
- $X_3 \perp\!\!\!\perp (X_1, X_2)$

Example with GKCID < CID

- $X_1, X_2, X_3 \sim \mathcal{N}(0,1)$ and $X_1 = X_2$ almost surely
- $X_3 \perp\!\!\!\perp (X_1, X_2)$
- GKCID $= 0$ while CID $= 1$

Future Work

How to compute CID, RCID, GKCID for general distributions and more general classes of function?

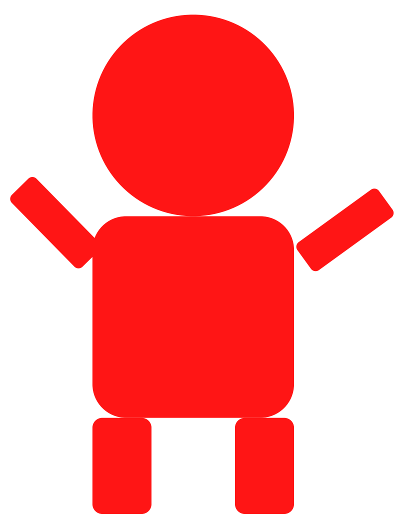
Any Questions?

02
Contextual Bandit Learning

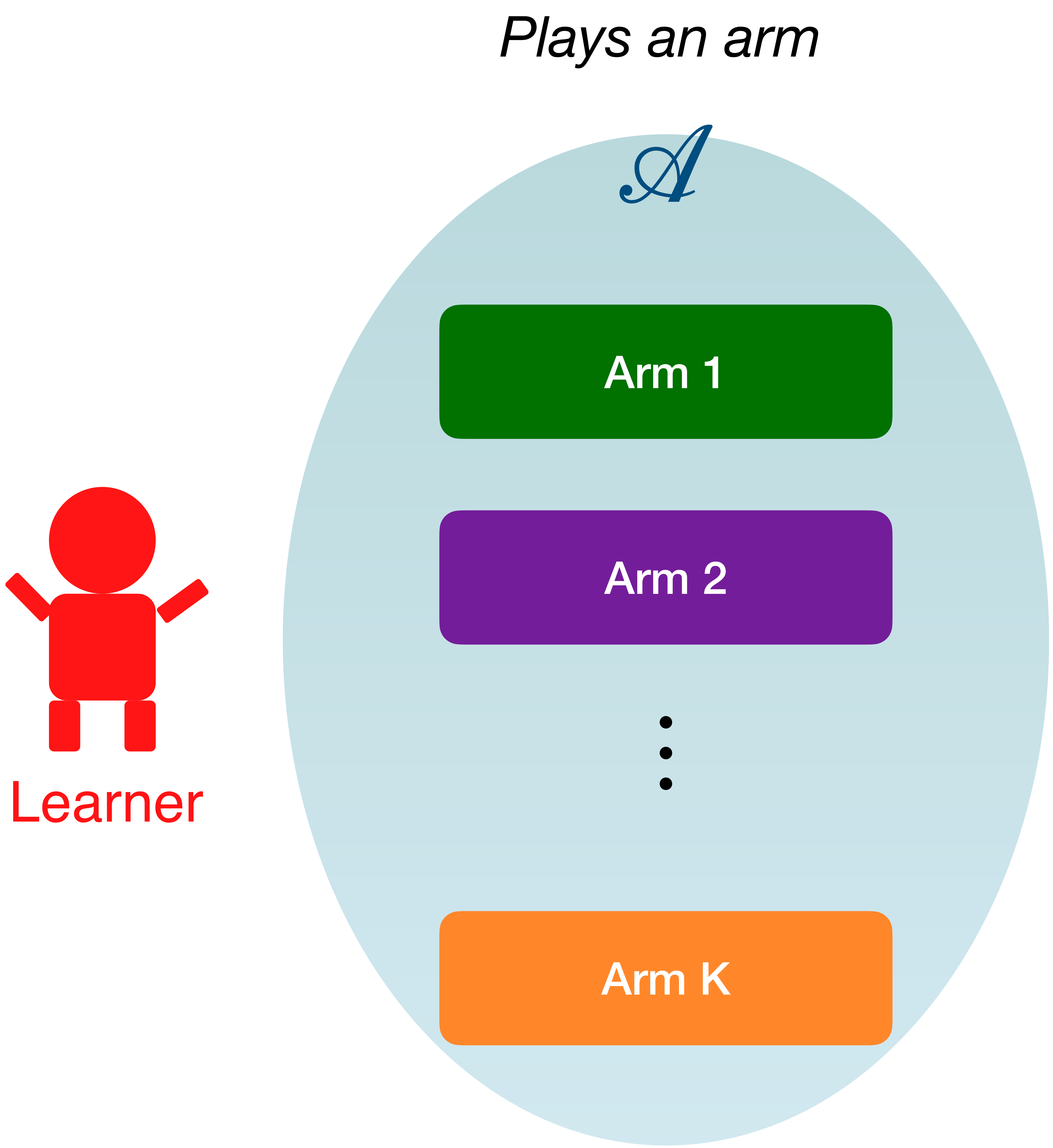
Hanna, Osama, Lin F. Yang, and Christina Fragouli. "Contexts can be Cheap: Solving Stochastic Contextual Bandits with Linear Bandit Algorithms." COLT 2023.

Multi Arm Bandits

Plays an arm



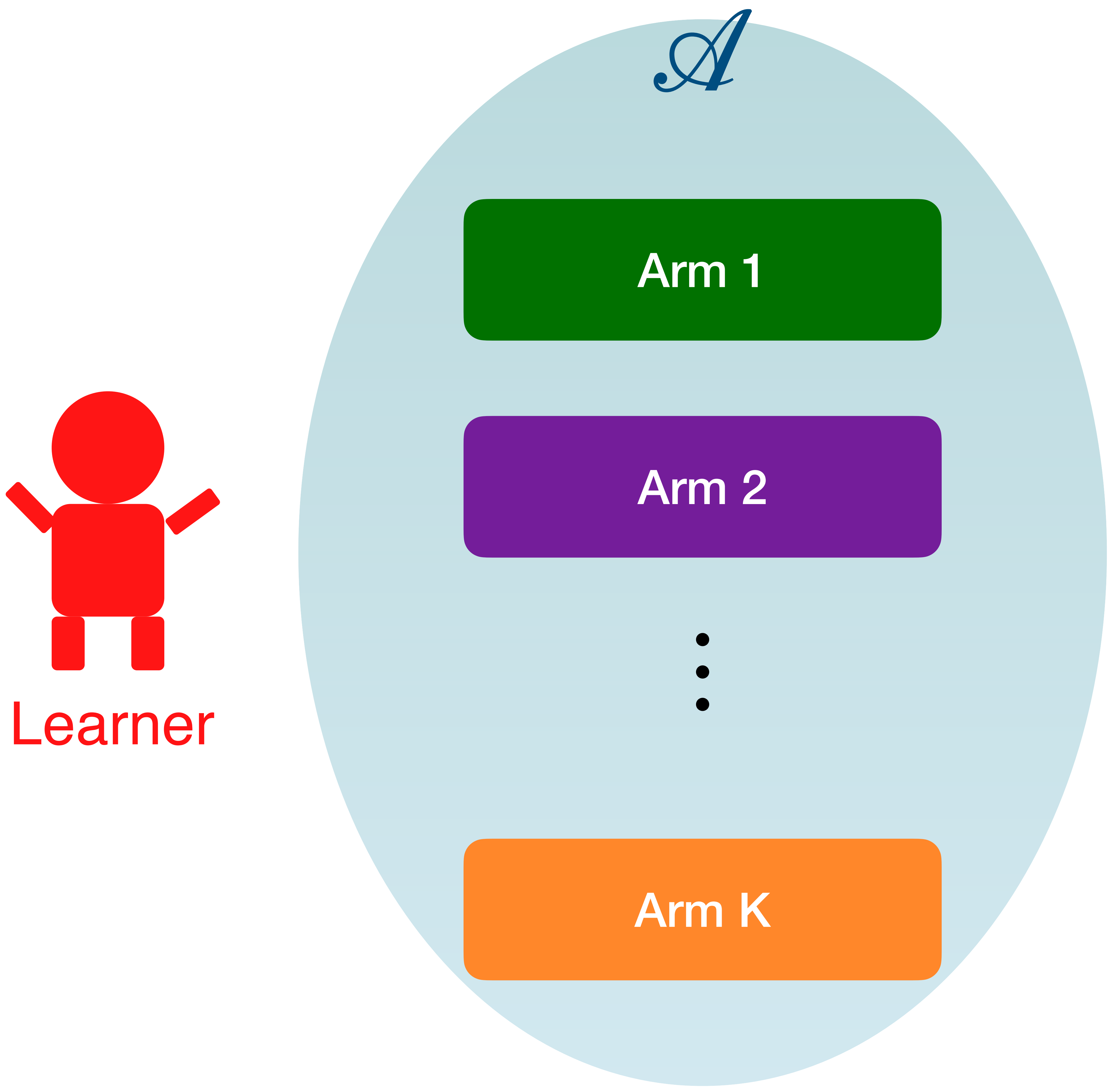
Multi Arm Bandits



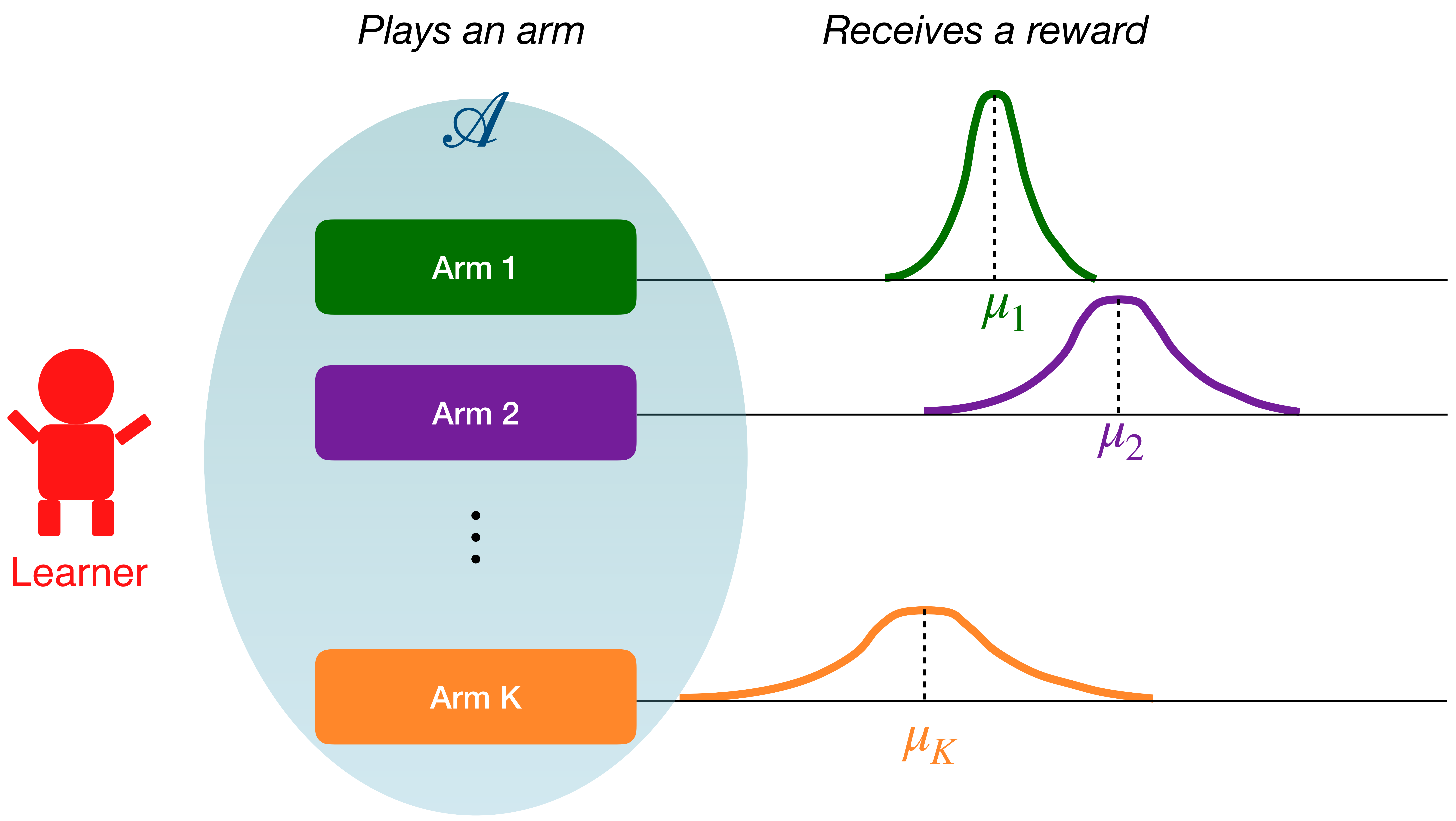
Multi Arm Bandits

Plays an arm

Receives a reward

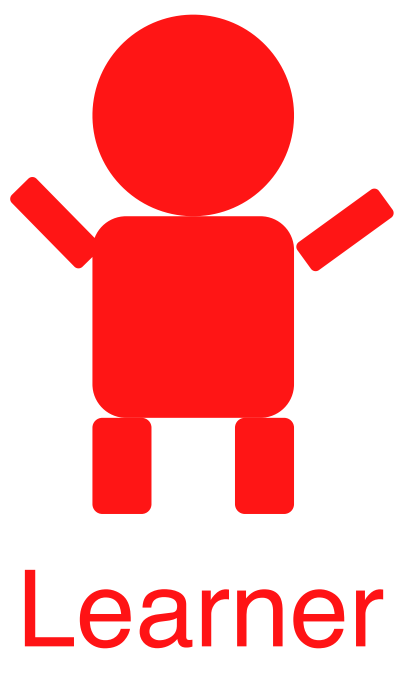


Multi Arm Bandits



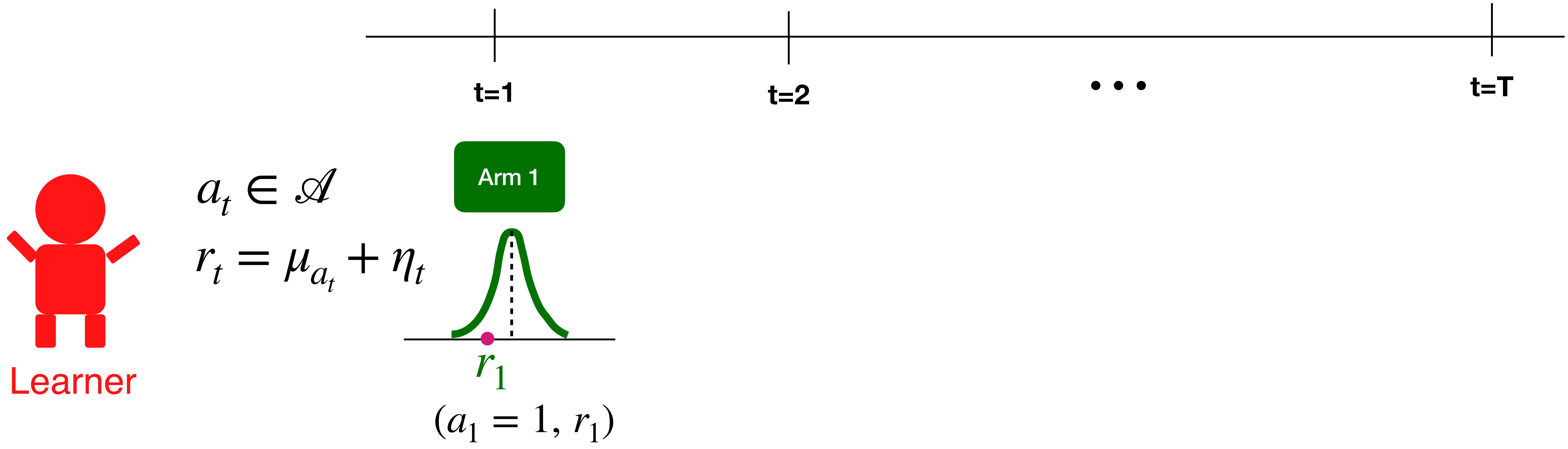
Multi Arm Bandits

Find which arm, among a set of choices, will provide on average the best reward



Multi Arm Bandits

Find which arm, among a set of choices, will provide on average the best reward

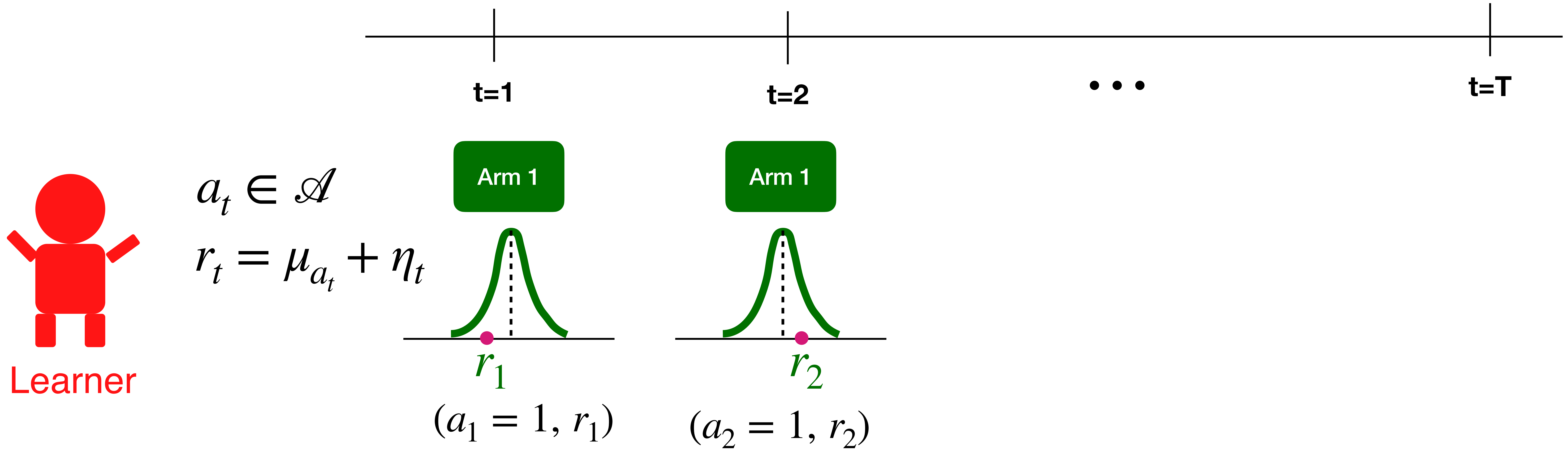


$$a_t \in \mathscr{A}$$

$$r_t = \mu_{a_t} + \eta_t$$

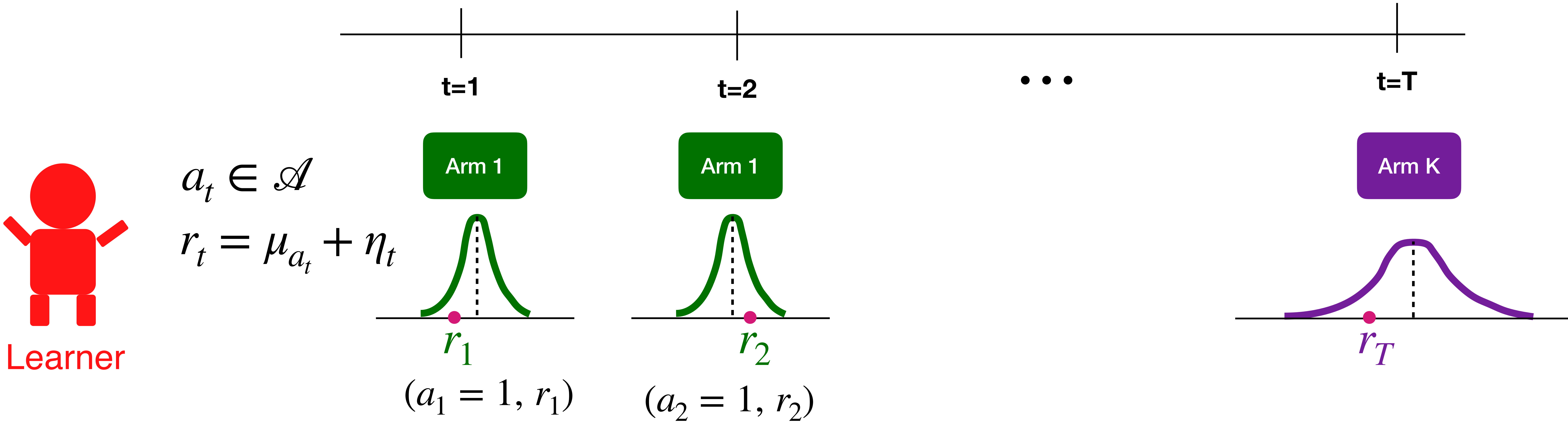
Multi Arm Bandits

Find which arm, among a set of choices, will provide on average the best reward



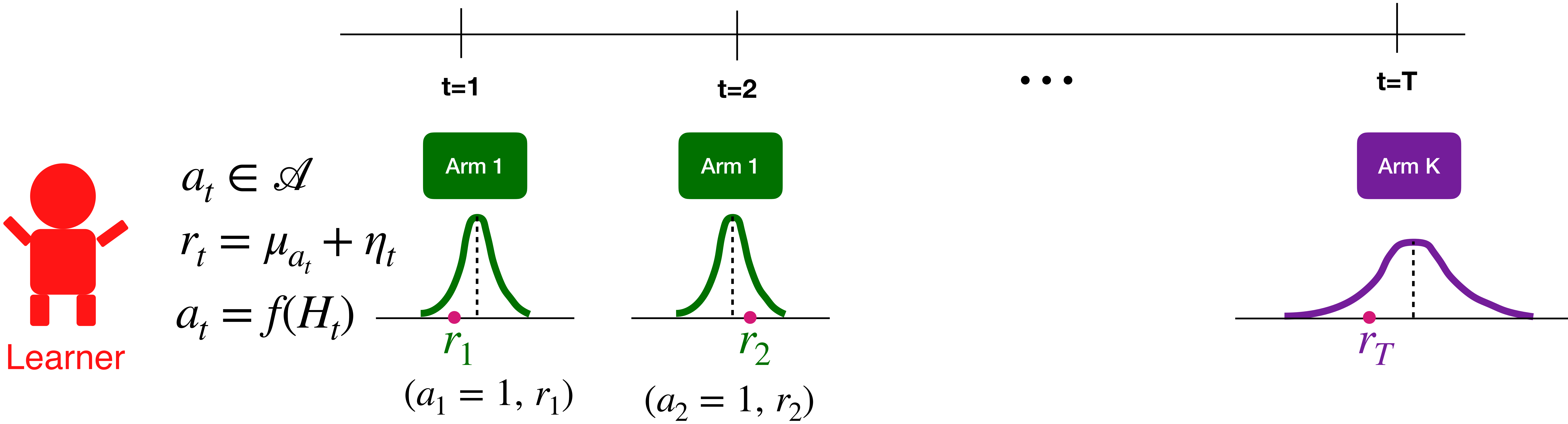
Multi Arm Bandits

Find which arm, among a set of choices, will provide on average the best reward



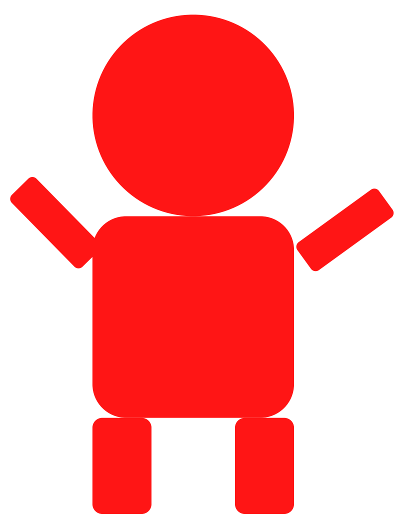
Multi Arm Bandits

Find which arm, among a set of choices, will provide on average the best reward



Multi Arm Bandits

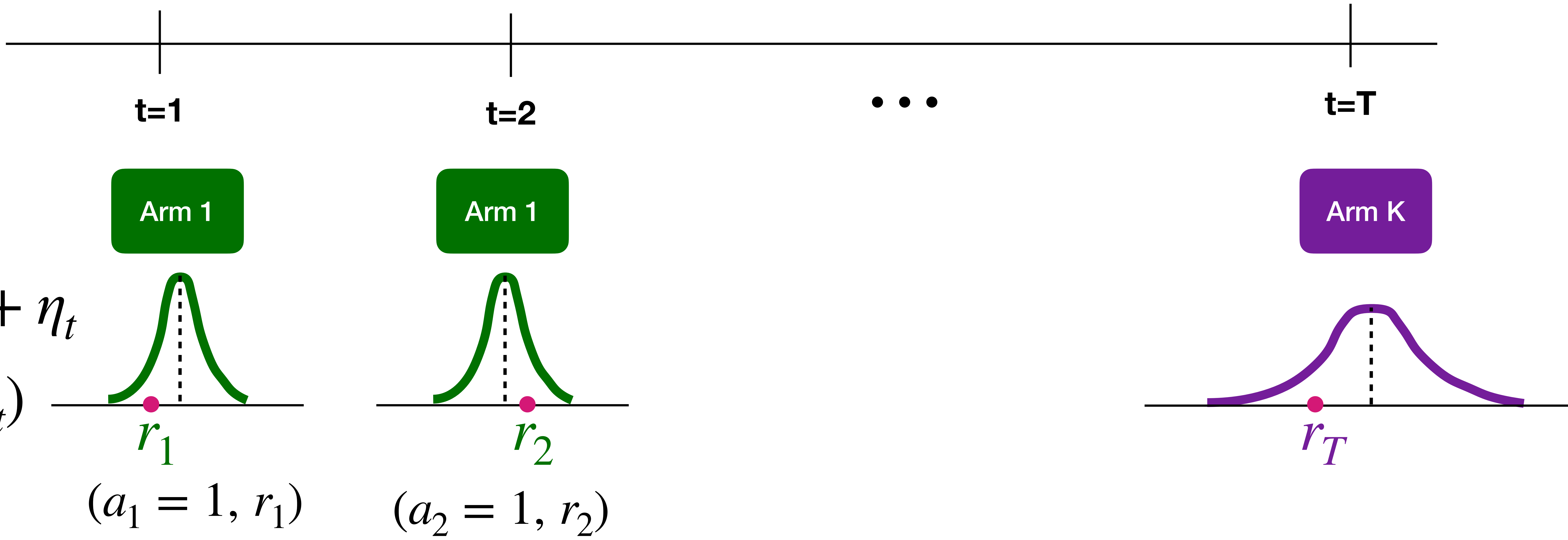
Find which arm, among a set of choices, will provide on average the best reward



$$a_t \in \mathscr{A}$$

$$r_t = \mu_{a_t} + \eta_t$$

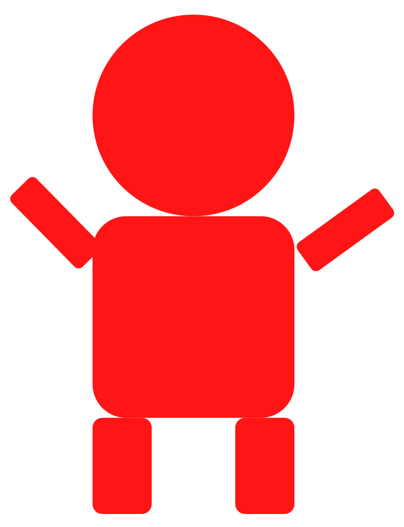
$$a_t = f(H_t)$$



Average Regret: $$R_T = \sum_{t=1}^{T} (\max_{a \in \mathscr{A}} \mu_a - \mu_{a_t})$$

Linear Bandits

Reward is a linear function of an unknown coefficient vector $\theta_\star$



$$a_t \in \mathcal{A} \subseteq \mathbb{R}^d$$

$$\mu_{a_t} = \langle a_t, \theta_\star \rangle, a_t \in \mathbb{R}^d$$

$$r_t = \mu_{a_t} + \eta_t$$

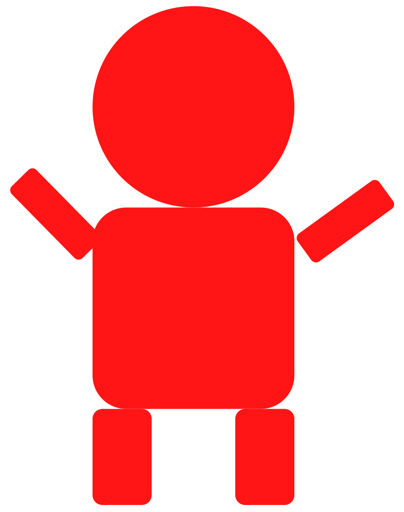
Regret: $$R_T = \sum_{t=1}^{T} \left(\max_{a \in \mathcal{A}} \langle a, \theta_\star \rangle - \langle a_t, \theta_\star \rangle\right)$$

Contextual Linear Bandits

Reward is a linear function of an unknown coefficient vector $\theta_\star$

Context for recommender systems:

gender, age, geographical location, past behavior,….



Learner

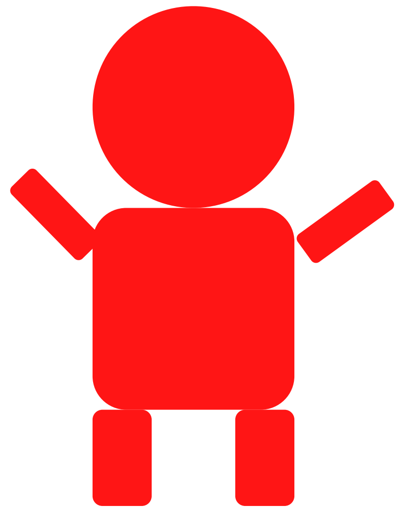
$a_t \in \mathcal{A}$

$r_t = \langle a_t, \theta_\star \rangle + \eta_t$

Regret: $$R_T = \sum_{t=1}^{T} \left(\max_{a \in \mathcal{A}} \langle a, \theta_\star \rangle - \langle a_t, \theta_\star \rangle \right)$$

Contextual Linear Bandits

Reward is a linear function of an unknown coefficient vector $\theta_\star$



Learner

$a_t \in \mathcal{A}$

known c_t, ϕ

$r_t = \langle \phi(c_t, a_t), \theta_\star \rangle + \eta_t$

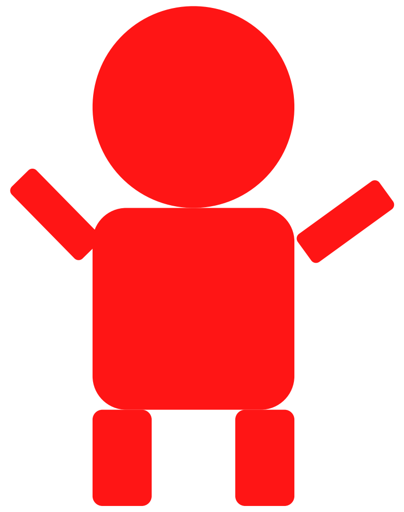
Context for recommender systems:

gender, age, geographical location, past behavior,….

Regret: $$R_T = \sum_{t=1}^{T} \left(\max_{a \in \mathcal{A}} \langle \phi(c_t, a), \theta_\star \rangle - \langle \phi(c_t, a_t), \theta_\star \rangle \right)$$

Contextual Linear Bandits

Reward is a linear function of an unknown coefficient vector $\theta_\star$



Learner

$$a_t \in \mathcal{A}_t$$

$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$

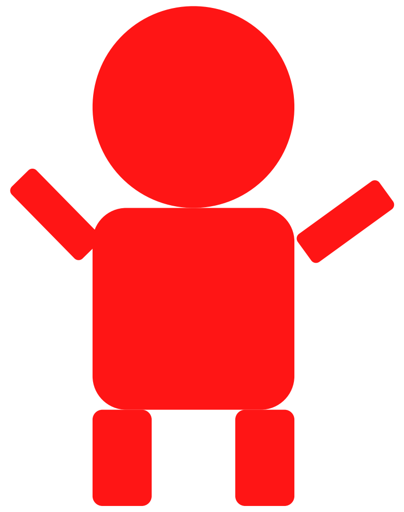
Context for recommender systems:

gender, age, geographical location, past behavior,….

Regret: $$R_T = \sum_{t=1}^{T} \left(\max_{a \in \mathcal{A}_t} \langle a, \theta_\star \rangle - \langle a_t, \theta_\star \rangle \right)$$

Challenge

Solving contextual linear bandits can be harder than solving linear bandits



Learner

$$a_t \in \mathcal{A}_t$$

$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$

Context for recommender systems:

gender, age, geographical location, past behavior,….

Regret: $$R_T = \sum_{t=1}^{T} \left(\max_{a \in \mathcal{A}_t} \langle a, \theta_\star \rangle - \langle a_t, \theta_\star \rangle \right)$$

Comparing Literature Results

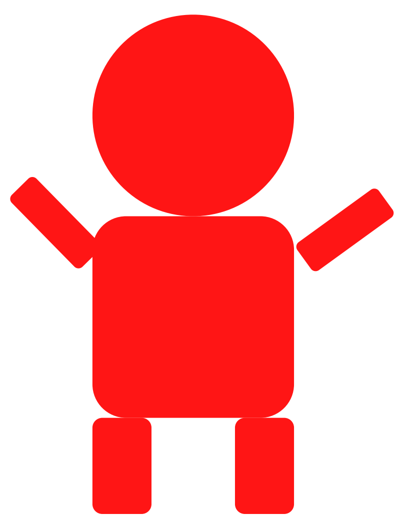
	Linear	Contextual
Basic setup	$O(d\sqrt{T\log T})$ w.h.p.	$O(d\sqrt{T}\log T)$ w.h.p.
Batched Algorithms	$O(d\sqrt{T\log T}\log\log T)$ w.h.p.	$O(d\sqrt{T\log d\log T}\log\log T)$ exp.
Adversarial corruption	$\tilde{O}(d\sqrt{T}+d^{1.5}C)$ w.h.p.	$\tilde{O}(d^{4.5}\sqrt{T}+d^{4}C)$ w.h.p.

Regret bound B_T in exp.: $\mathbb{E}[R_T] \leq B_T$
W.h.p.: $R_T \leq B_T$ w.p. at least $1 - 1/T$

C : amount of corruption

Building Intuition

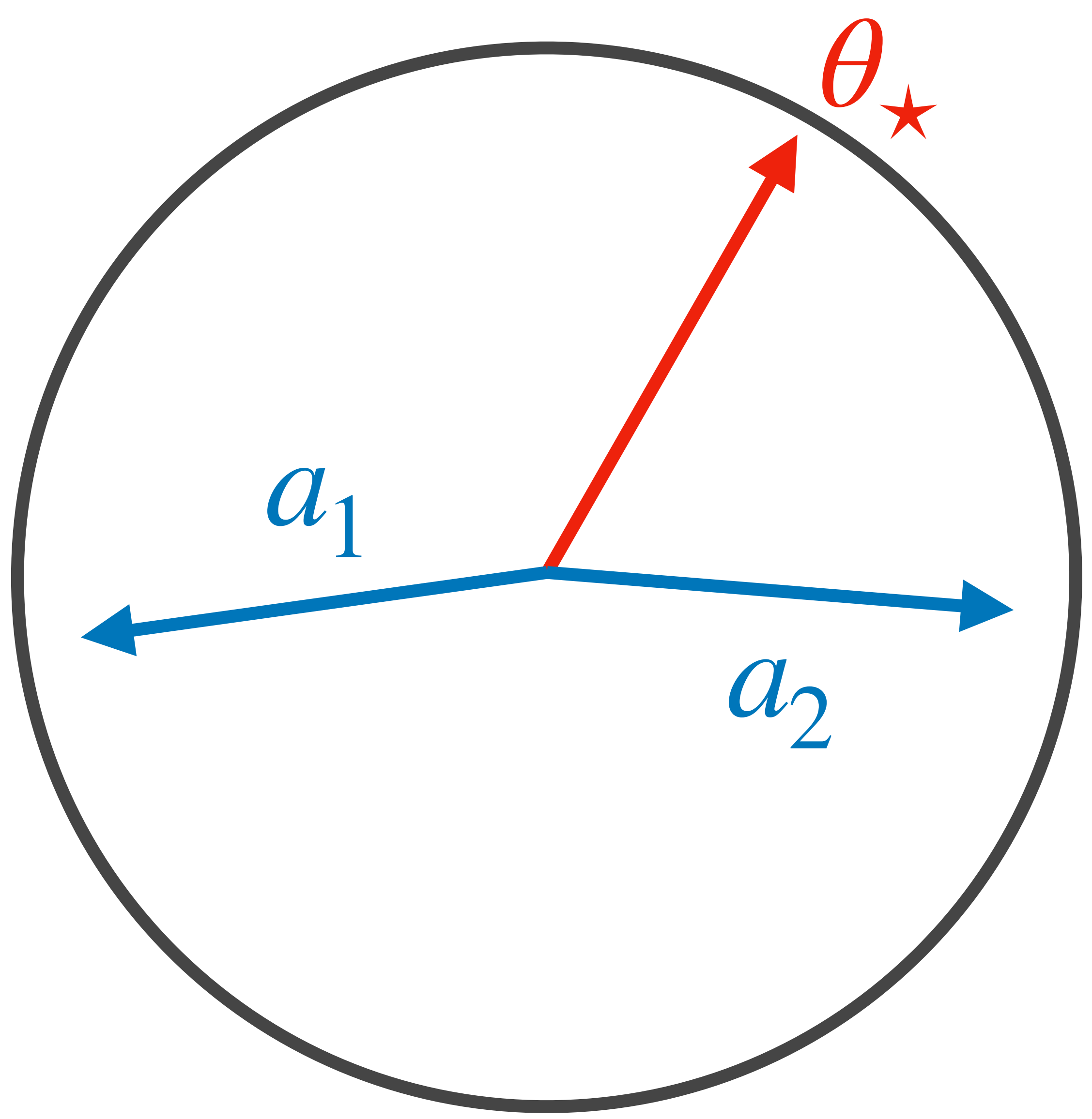
Linear Bandits goal: estimate optimal coefficient vector $\theta_\star$



Learner

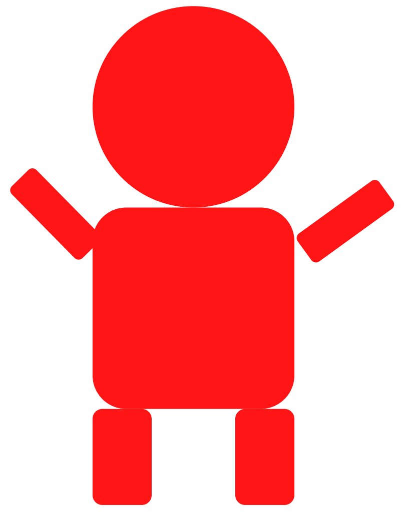
$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$

$$\alpha_t \in \mathscr{A}$$



Building Intuition

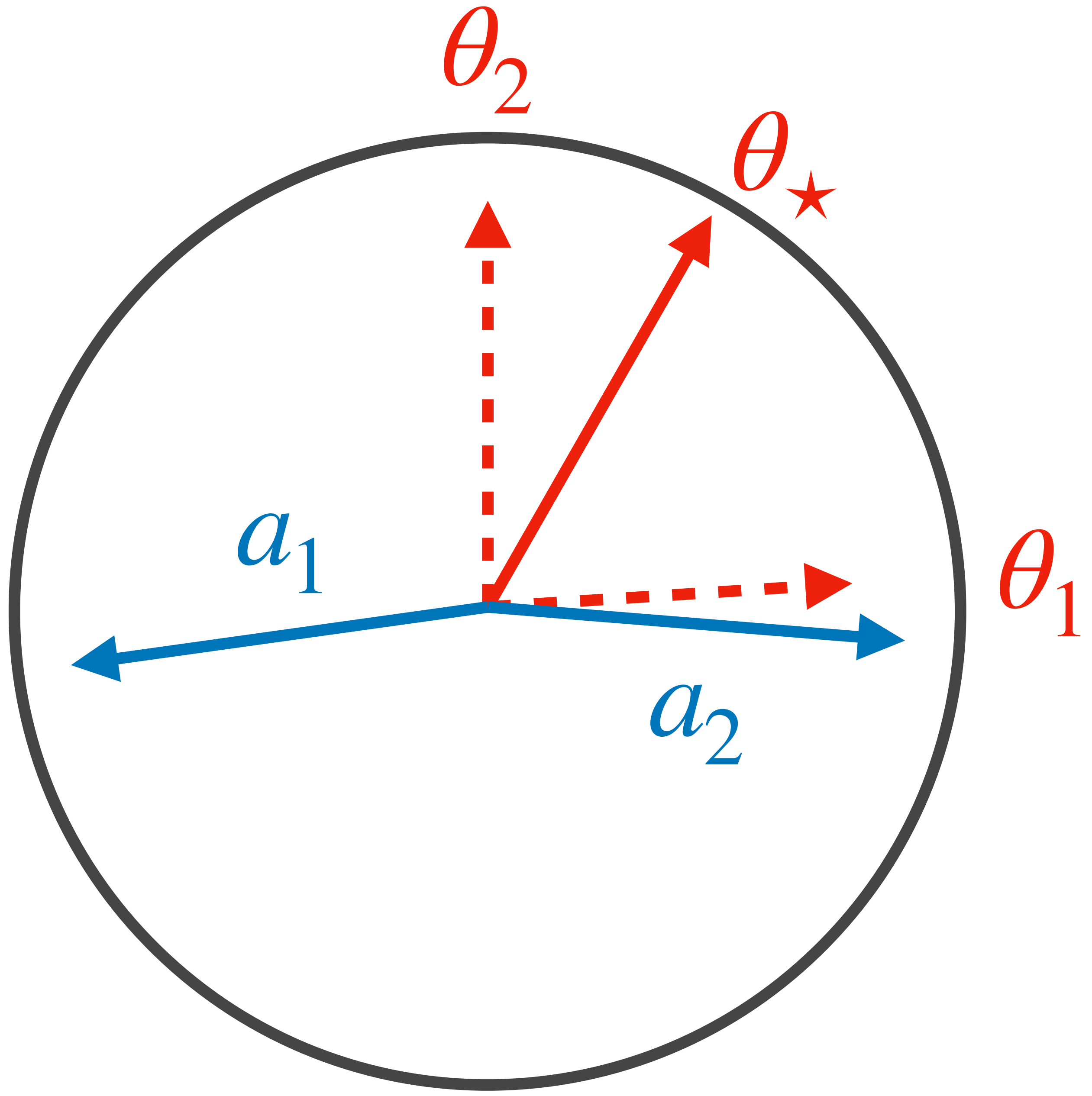
Linear Bandits goal: estimate optimal coefficient vector $\theta_\star$



Learner

$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$

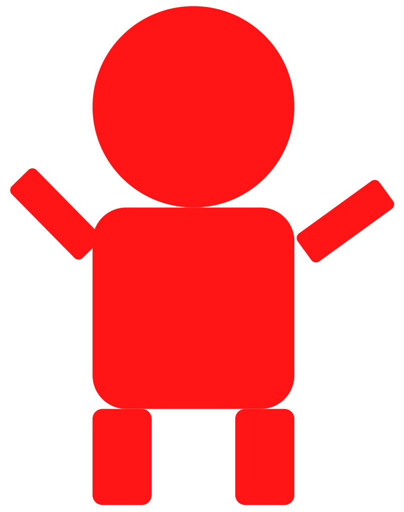
$$\alpha_t \in \mathscr{A}$$



Estimate $\theta_\star$ along actions directions

Building Intuition

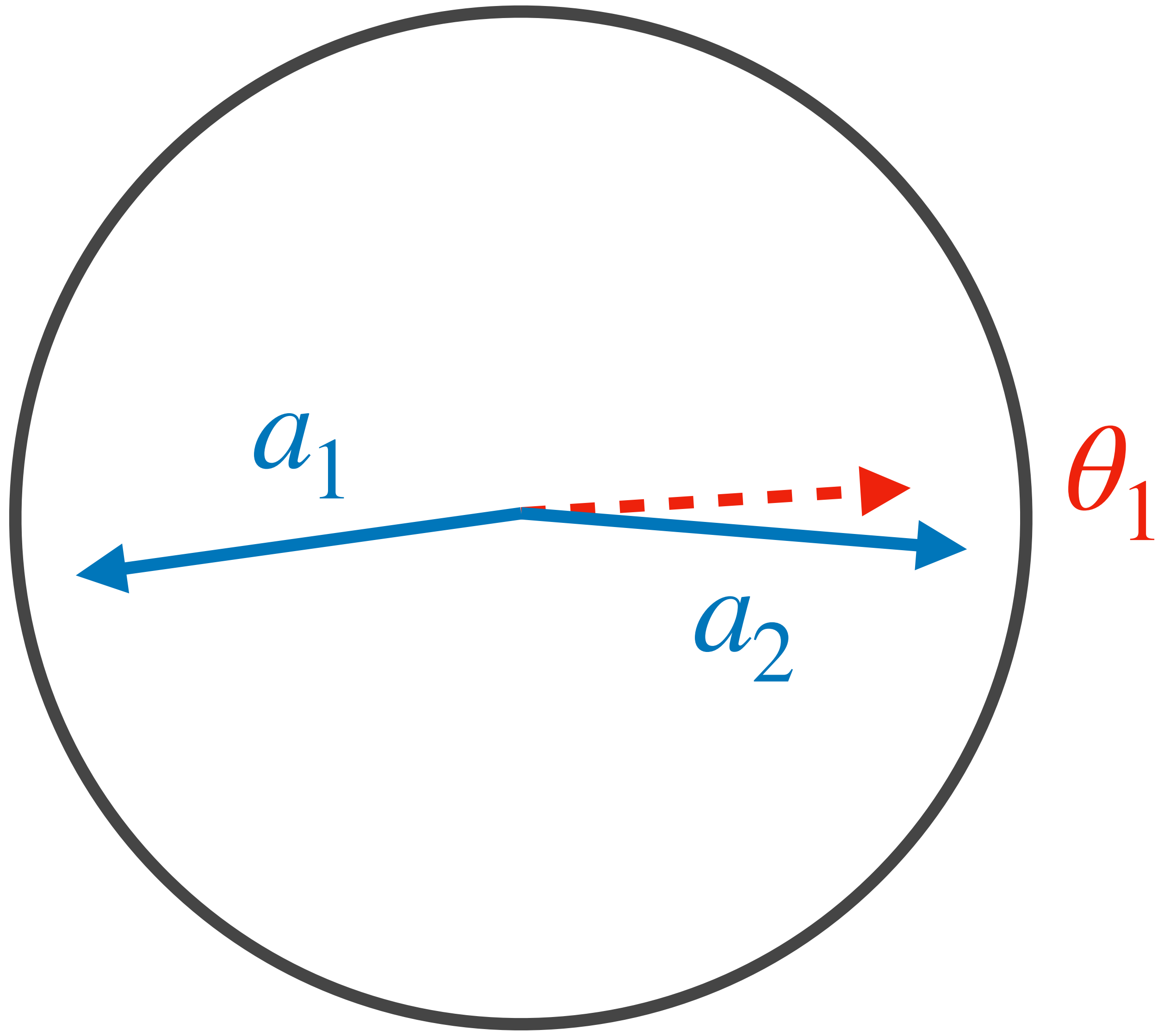
Linear Bandits goal: estimate optimal coefficient vector $\theta_\star$



Learner

$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$

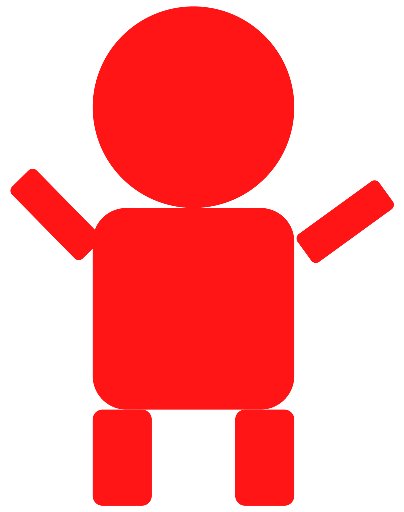
$$\alpha_t \in \mathcal{A}$$



Estimate $\theta_\star$ along actions directions

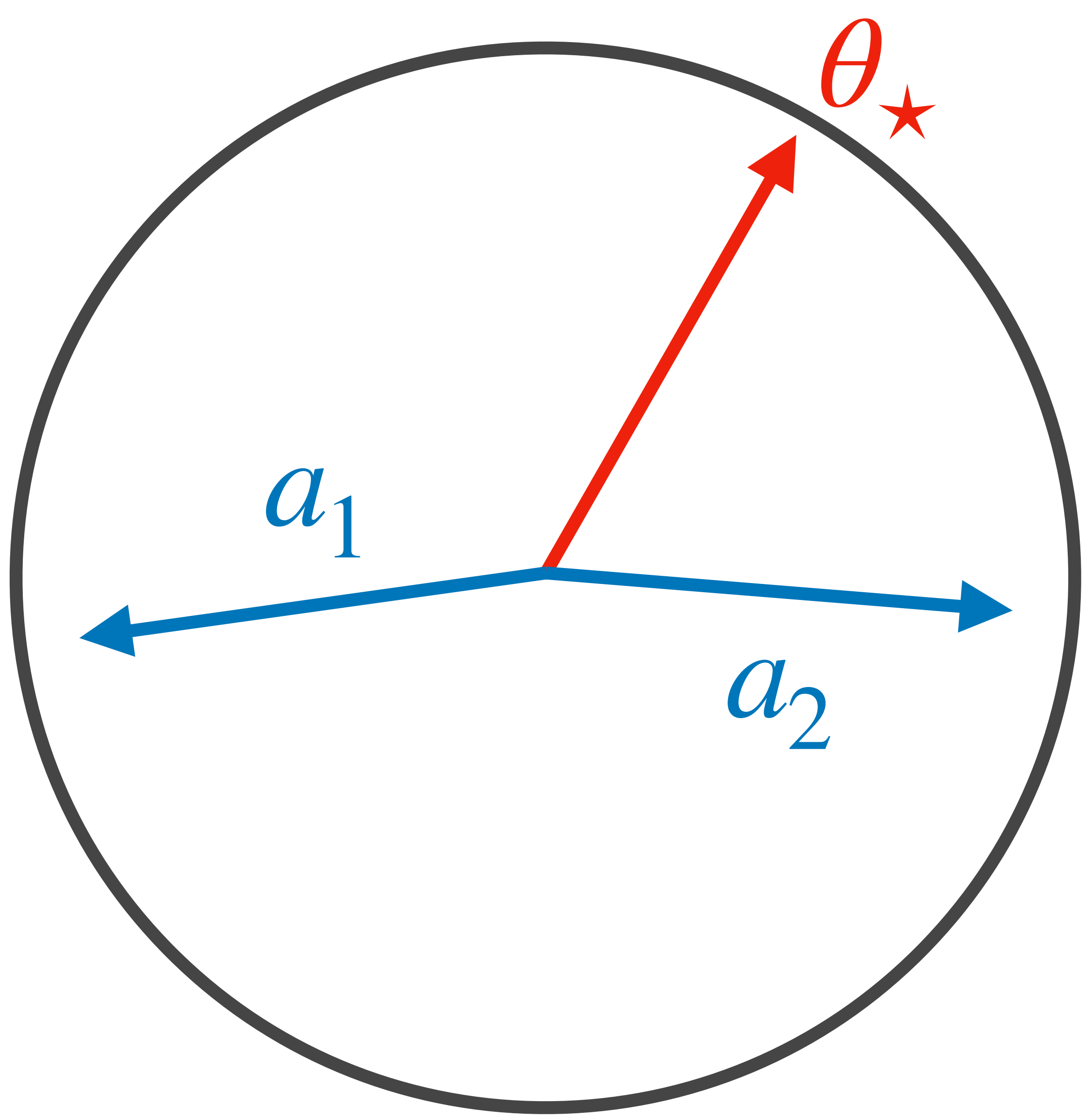
Building Intuition

Contextual Linear Bandits: directions change



Learner

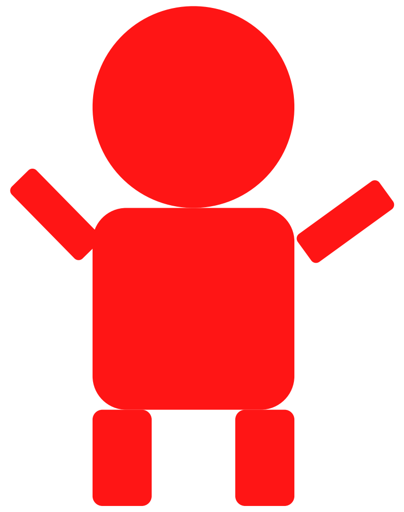
$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$
$$\alpha_t \in \mathcal{A}_t$$



$t = 1$

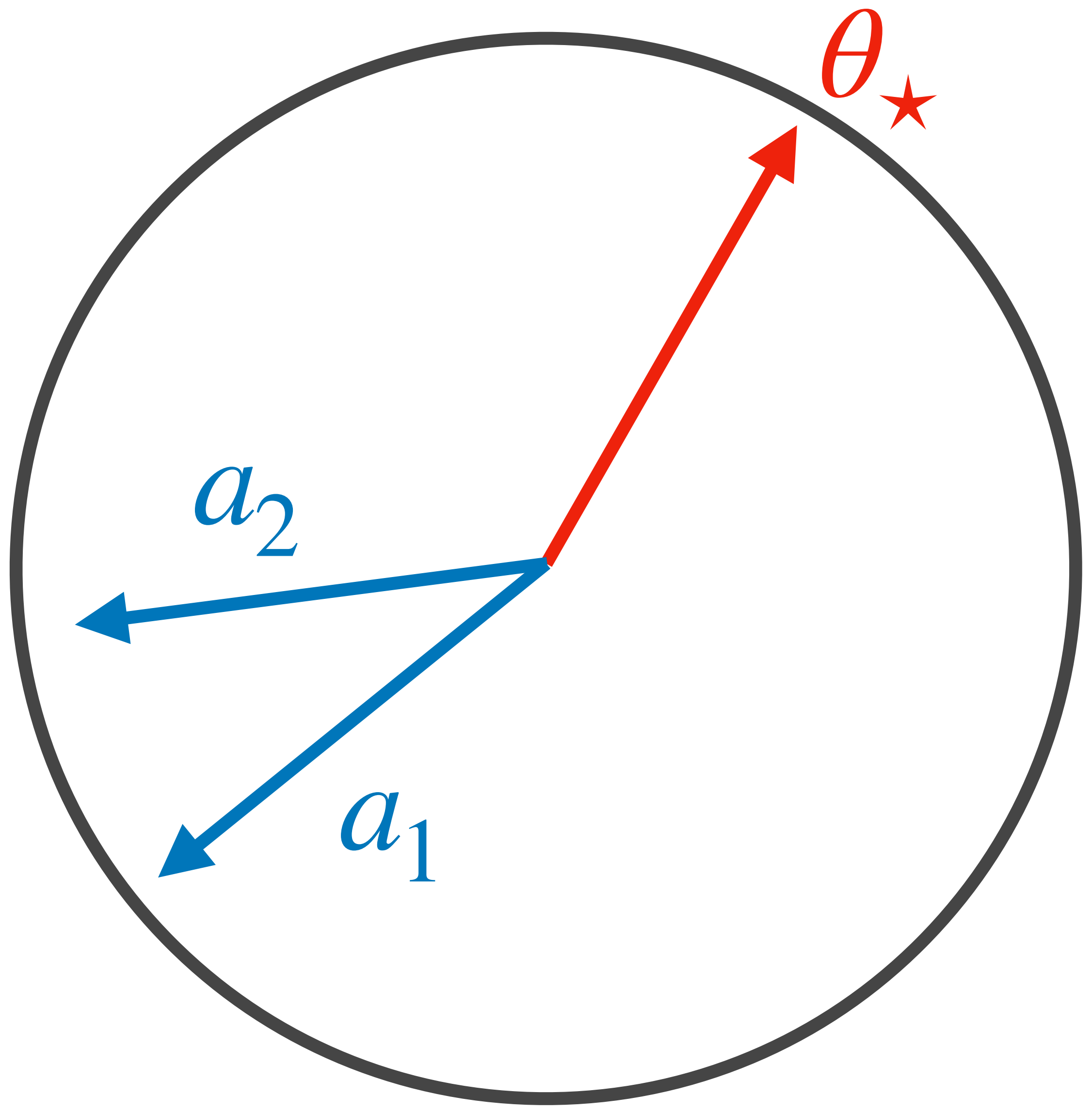
Building Intuition

Contextual Linear Bandits: directions change



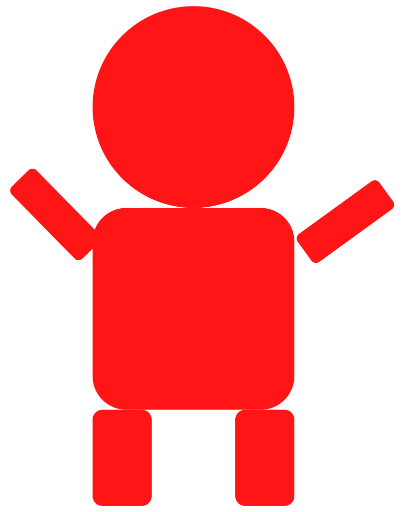
Learner

$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$
$$\alpha_t \in \mathcal{A}_t$$



$t = 2$

Main result



Learner

~~$\mathscr{A}_t$~~

$\mathscr{A}$

If the context is generated from a distribution,
we can reduce Contextual Linear Bandits
to Linear Bandits

Our results

Theorem 1:

For any contextual linear bandit instance I with **known** context distribution $\mathscr{D}$, there exists (constructively) a linear bandit instance L with the same action dimension, and any algorithm solving L solves I with the same worst-case regret bound as L.

Our results

Theorem 2:

For any contextual linear bandit instance I with **unknown** context distribution $\mathscr{D}$, there exist (constructively) $\log T$ linear bandit instances $L_1, \ldots, L_{\log T}$ with $\tilde{O}(1/\sqrt{T_i})$ misspecification, and any algorithm solving $L_1, \ldots, L_{\log T}$ solves I with the same worst-case regret bound.

Reduction

Use a Linear Bandit Algorithm to learn the optimal $\theta_\star$ *for the Contextual Bandit*

Instead of solving

$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$

$$R_T = \sum_{t=1}^{T} \max_{a \in \mathcal{A}_t} \langle a, \theta_\star \rangle - \langle a_t, \theta_\star \rangle$$

- $a_t \in \mathcal{A}_t$: context changes with t

Reduce to

$$r_t = \langle a_t, \theta_\star \rangle + \eta'_t$$

$$R_T = \sum_{t=1}^{T} \max_{a \in \mathcal{X}} \langle a, \theta_\star \rangle - \langle a_t, \theta_\star \rangle$$

- $a_t \in \mathcal{X}$: fixed $\forall t$

Reduction

We will use any standard LB algorithm

*(say **Alg**) to approximate $\theta_\star$ with action set $\mathscr{X}$*

Instead of solving

$$r_t = \langle a_t, \theta_\star \rangle + \eta_t$$

$$R_T = \sum_{t=1}^{T} \max_{a \in \mathscr{A}_t} \langle a, \theta_\star \rangle - \langle a_t, \theta_\star \rangle$$

- $a_t \in \mathscr{A}_t$: context changes with t

Reduce to

$$r_t = \langle a_t, \theta_\star \rangle + \eta_t'$$

$$R_T = \sum_{t=1}^{T} \max_{a \in \mathscr{X}} \langle a, \theta_\star \rangle - \langle a_t, \theta_\star \rangle$$

- $a_t \in \mathscr{X}$: fixed $\forall t$

How to create the set of actions $\mathscr{X}$

Reduction for known distribution

$$g(\theta) = \mathbb{E}_{\mathscr{A}_t \sim \mathscr{D}}[\arg\max_{a \in \mathscr{A}_t} \langle a, \theta \rangle \mid \mathscr{A}_t] \ \forall \theta \in \Theta$$

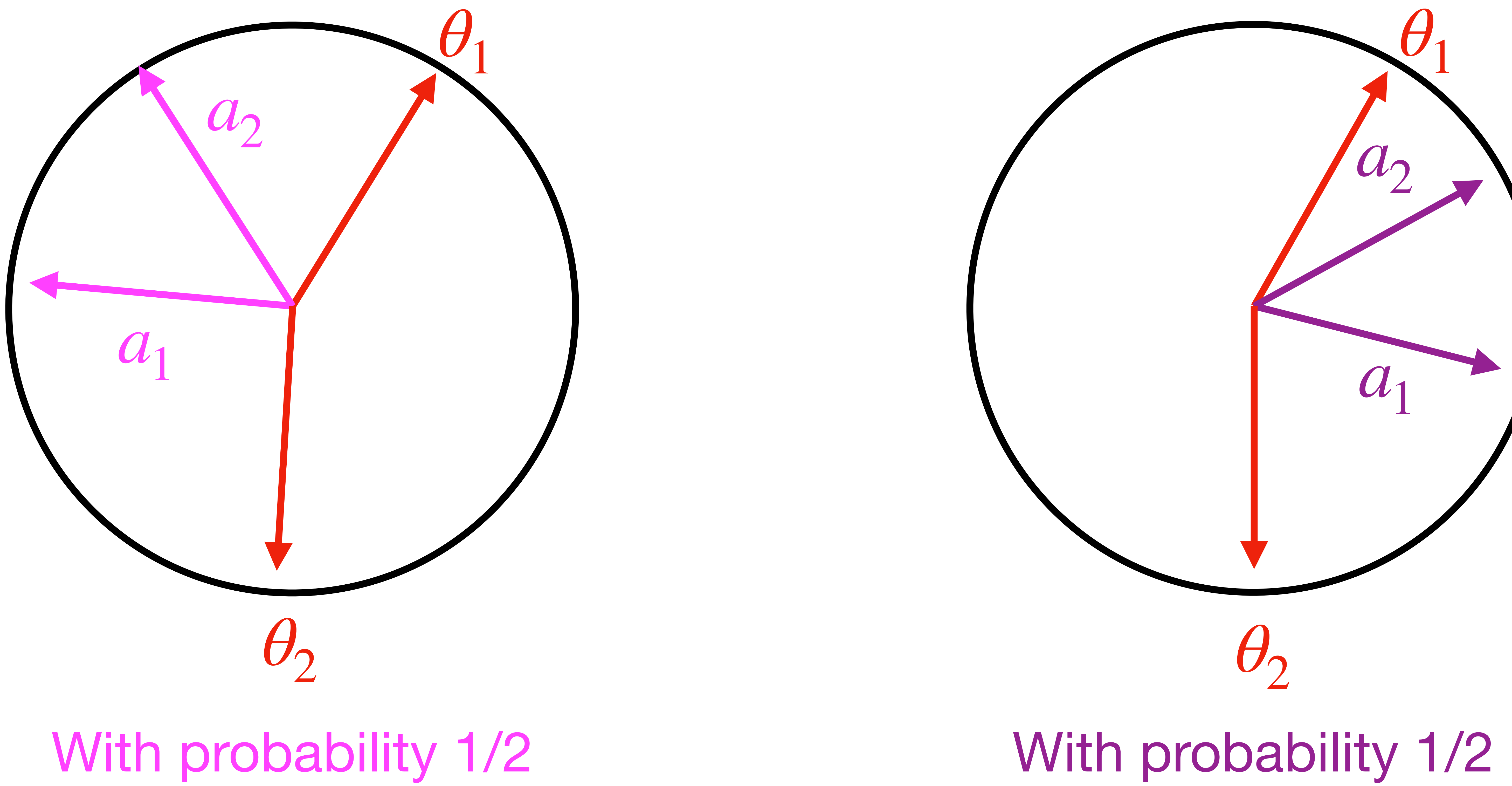
$$\mathscr{X} = \{g(\theta) \mid \theta \in \Theta\}$$

Example

$$g(\theta) = \mathbb{E}_{\mathcal{A}_t \sim \mathcal{D}}[\arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle \mid \mathcal{A}_t] \; \forall \theta \in \Theta$$

Assume two sets of actions

$$\Theta = \{\theta_1, \theta_2\}$$

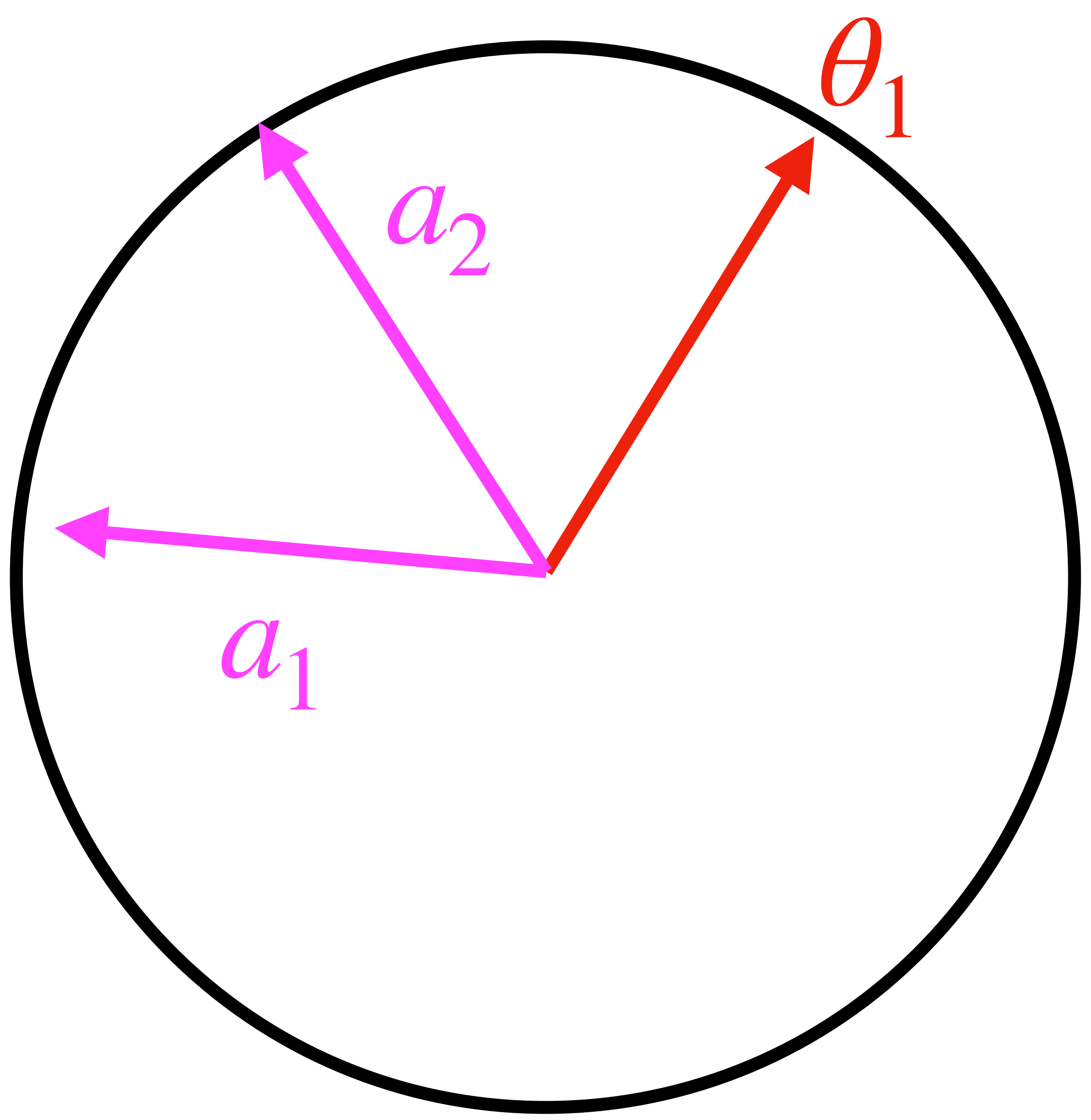


Example

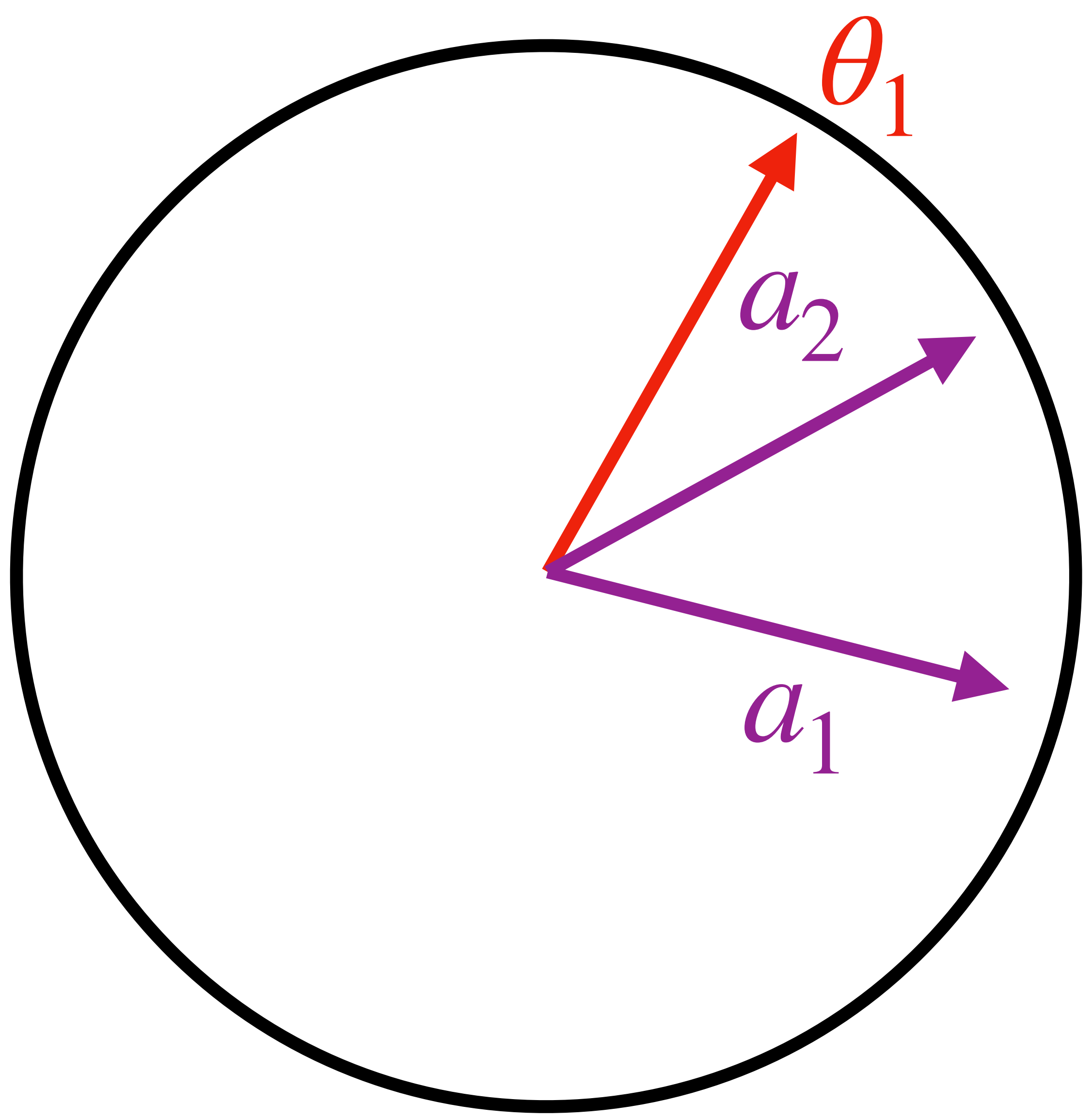
Assume two sets of actions

$$g(\theta) = \mathbb{E}_{\mathcal{A}_t \sim \mathcal{D}}[\arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle \mid \mathcal{A}_t] \ \forall \theta \in \Theta$$

$$\Theta = \{\theta_1, \theta_2\}$$



With probability 1/2



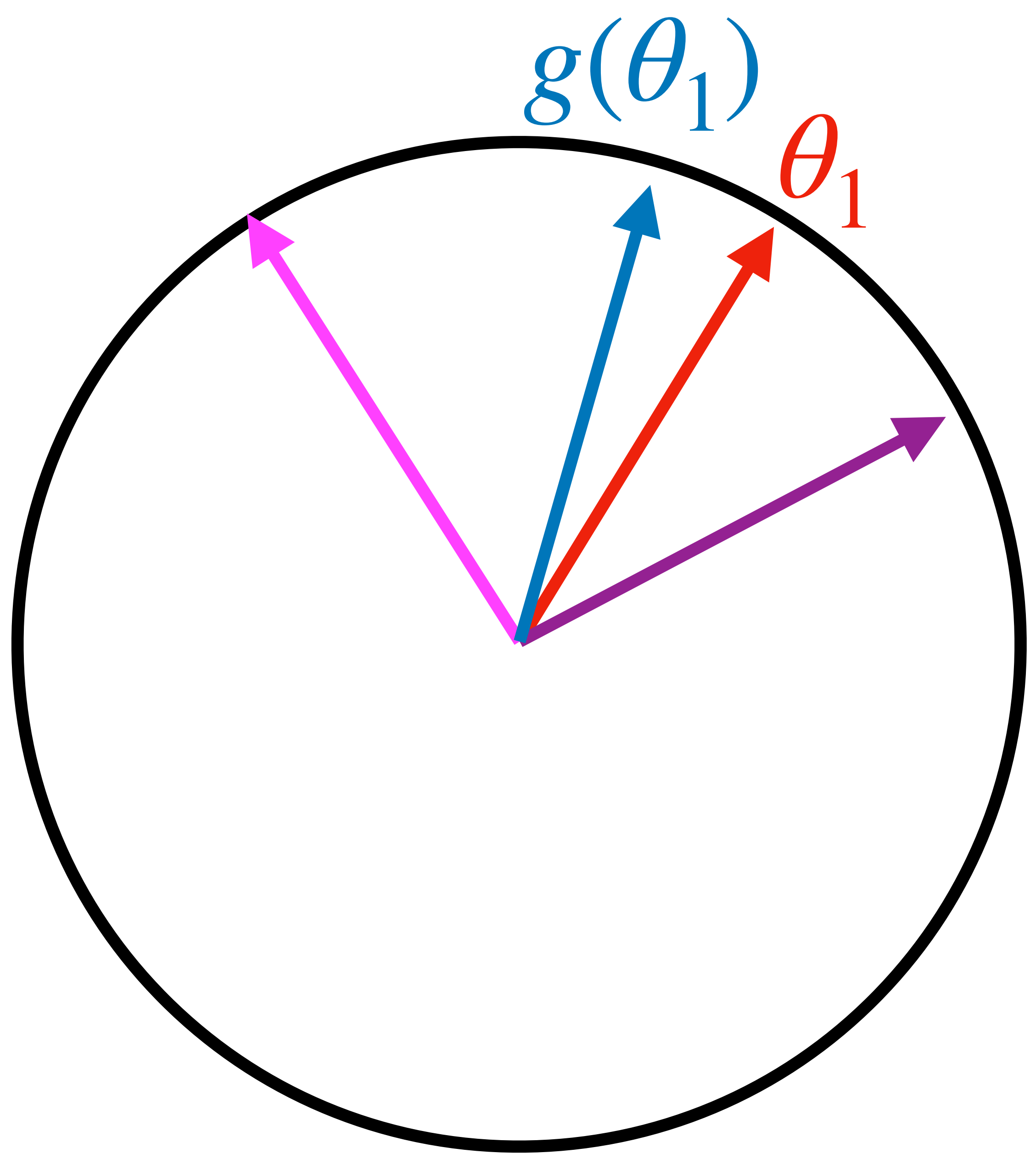
With probability 1/2

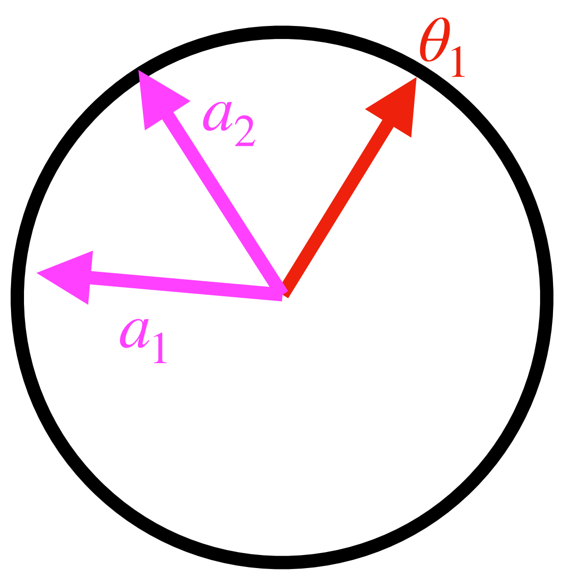
Example

$$g(\theta) = \mathbb{E}_{\mathscr{A}_t \sim \mathscr{D}}[\arg\max_{a \in \mathscr{A}_t} \langle a, \theta \rangle \mid \mathscr{A}_t] \; \forall \theta \in \Theta$$

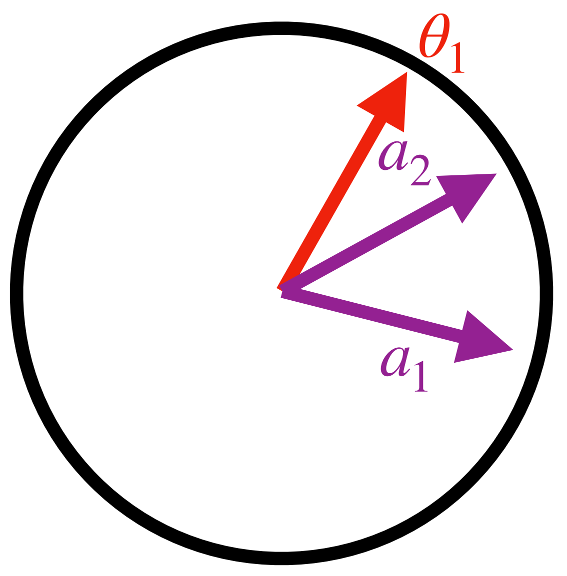
Assume two sets of actions

$$\Theta = \{\theta_1, \theta_2\}$$





With probability 1/2



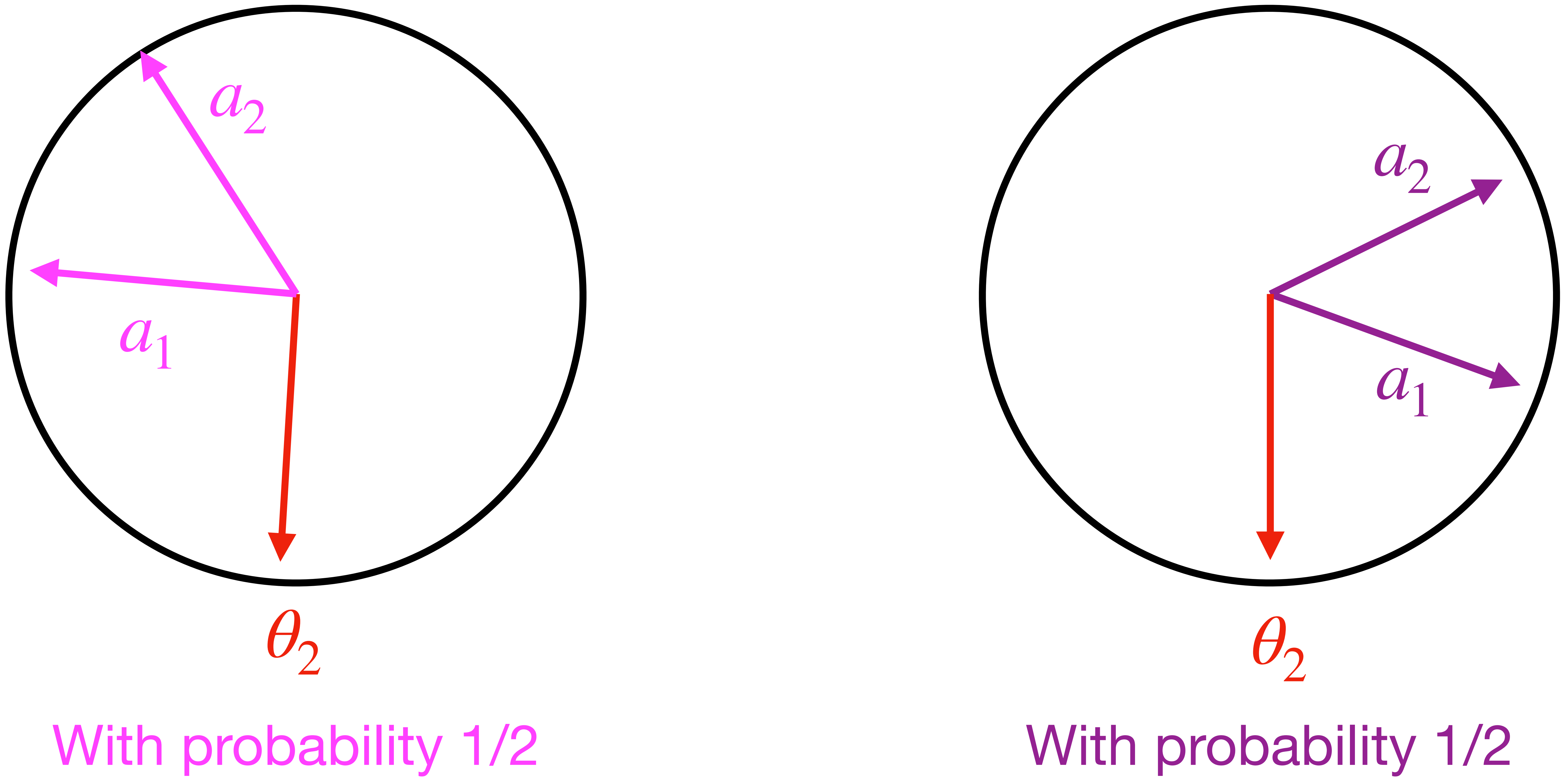
With probability 1/2

Example

$$g(\theta) = \mathbb{E}_{\mathcal{A}_t \sim \mathcal{D}}[\arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle \mid \mathcal{A}_t] \ \forall \theta \in \Theta$$

Assume two sets of actions

$$\Theta = \{\theta_1, \theta_2\}$$

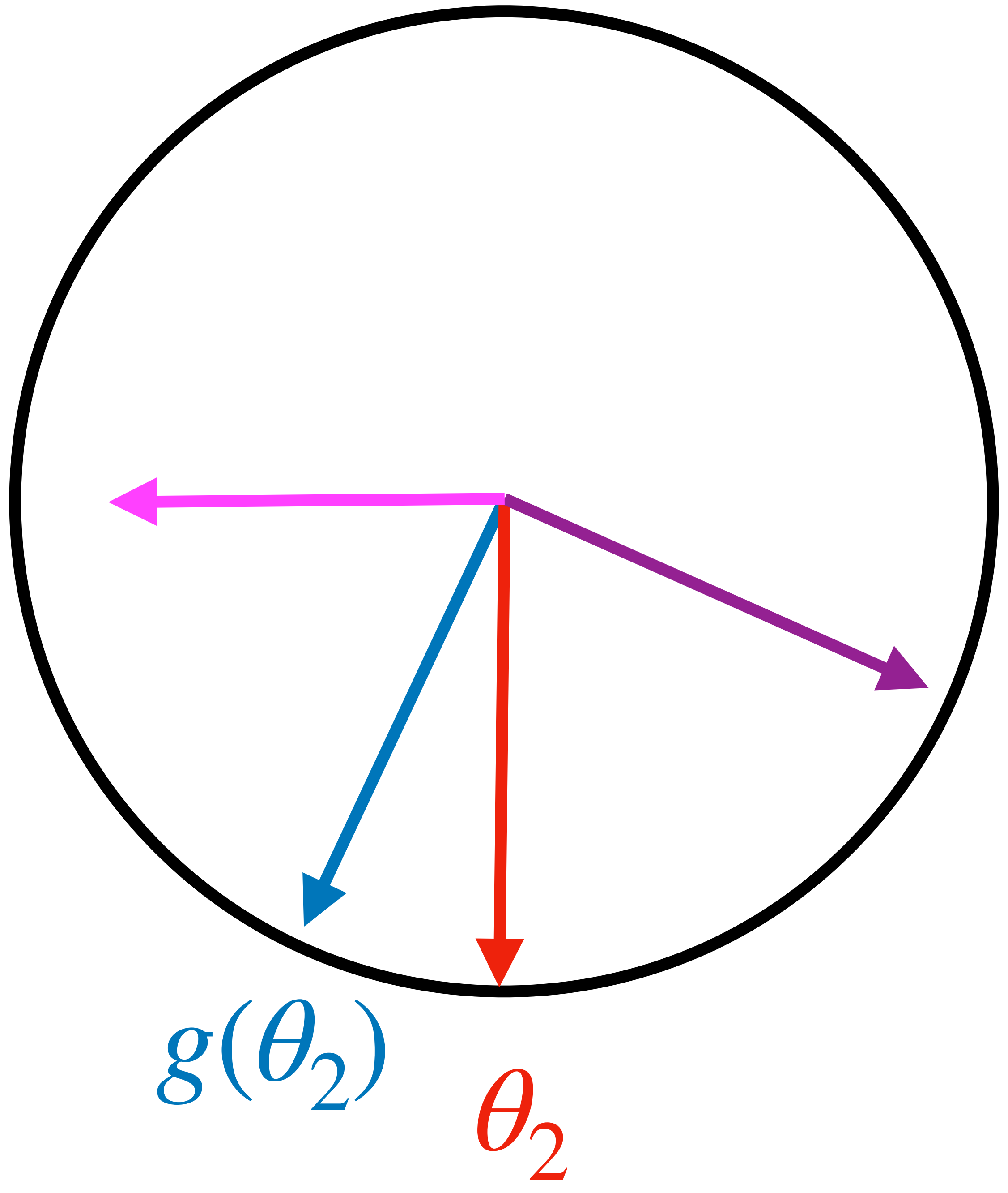


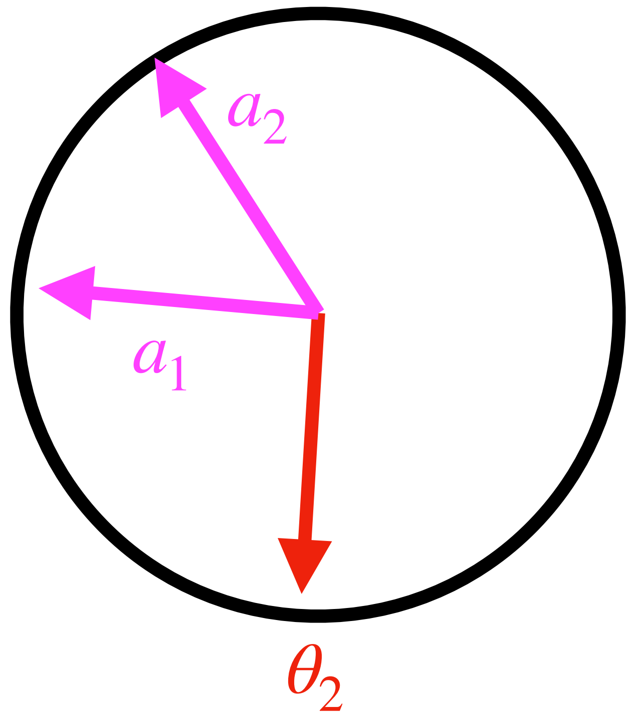
Example

Assume two sets of actions

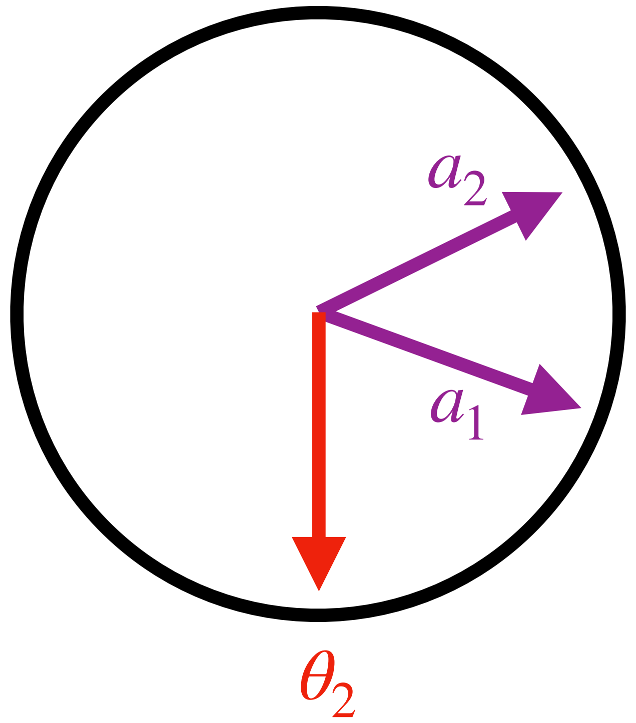
$$g(\theta) = \mathbb{E}_{\mathcal{A}_t \sim \mathcal{D}}[\arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle \mid \mathcal{A}_t] \ \forall \theta \in \Theta$$

$$\Theta = \{\theta_1, \theta_2\}$$





With probability 1/2



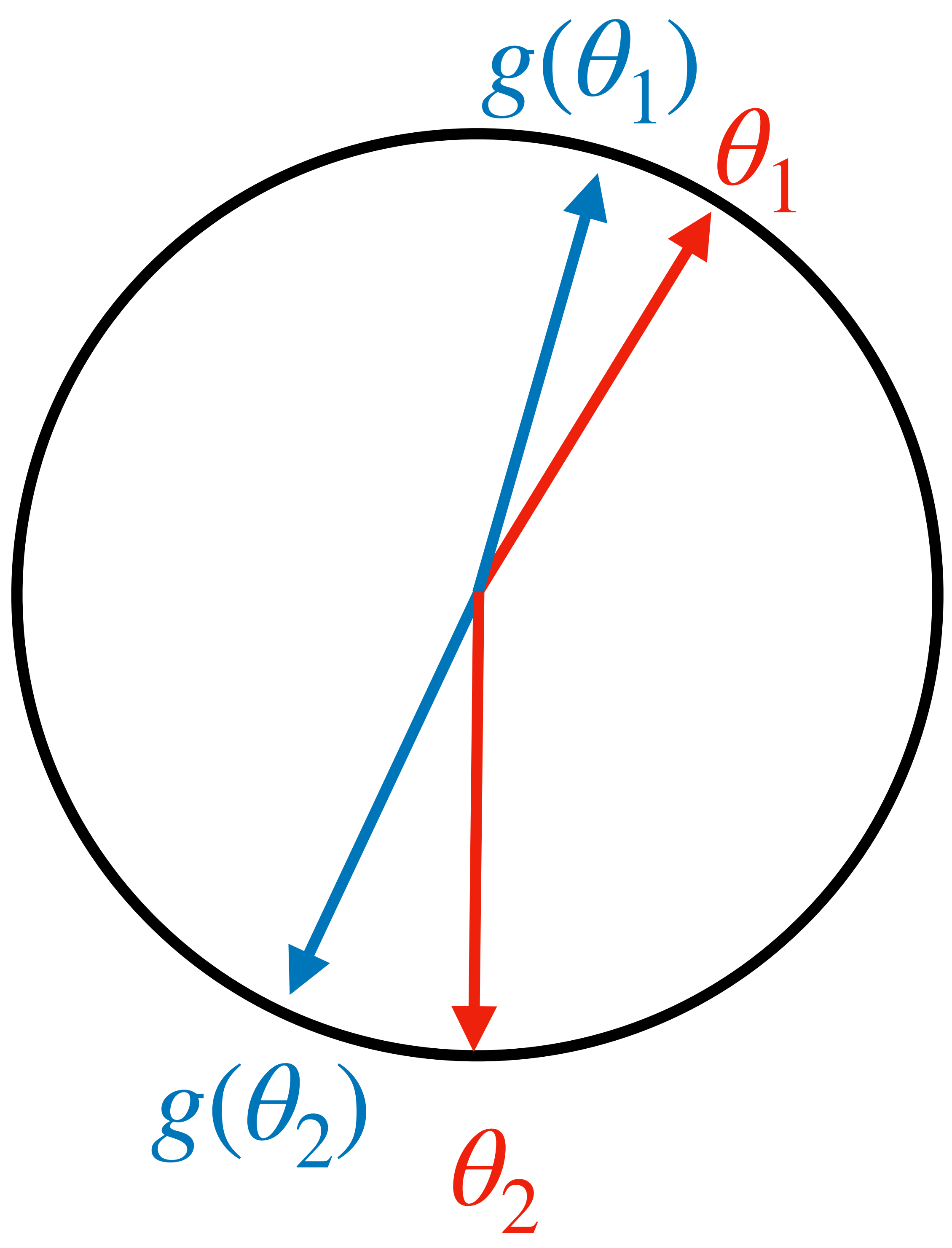
With probability 1/2

Example

Assume two sets of actions

$$g(\theta) = \mathbb{E}_{\mathcal{A}_t \sim \mathcal{D}}[\arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle \mid \mathcal{A}_t] \ \forall \theta \in \Theta$$

$$\Theta = \{\theta_1, \theta_2\}$$



$$\mathcal{X} = \{g(\theta_1), g(\theta_2)\}$$

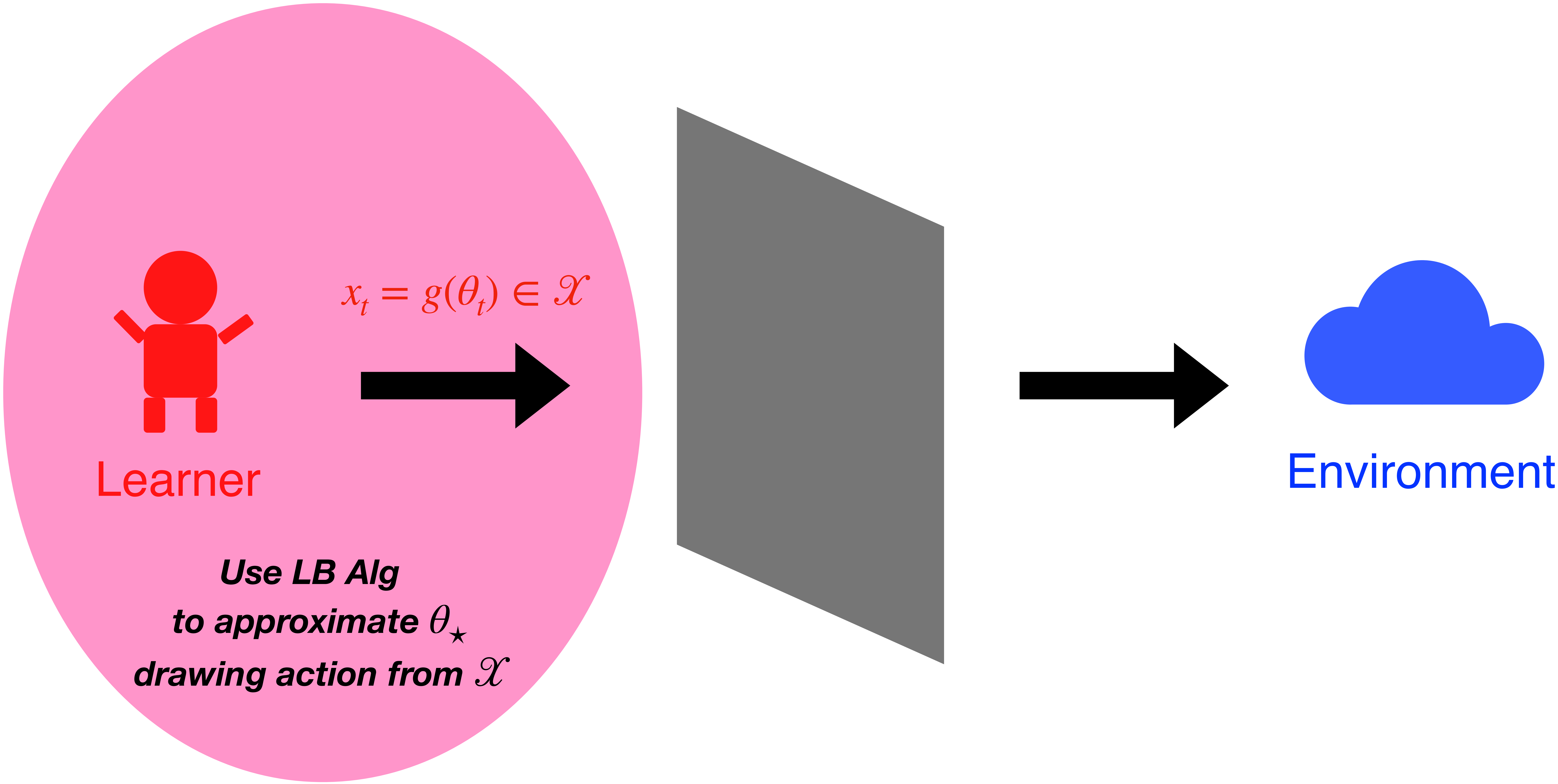
Reduction

Known distribution

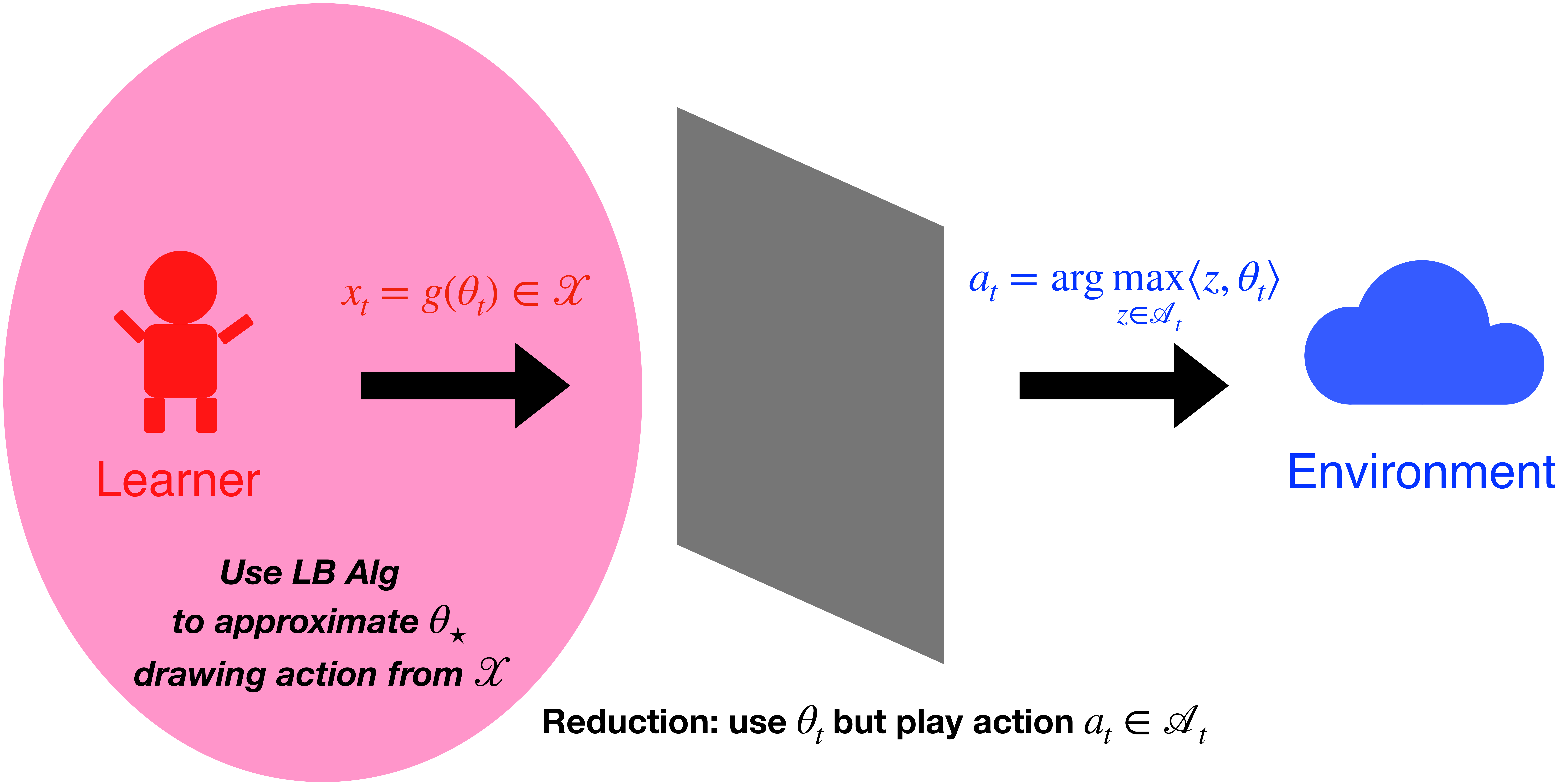
$$g(\theta) = \mathbb{E}_{\mathcal{A}_t \sim \mathcal{D}}[\arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle \mid \mathcal{A}_t] \quad \forall \theta \in \Theta$$

$$\mathcal{X} = \{g(\theta) \mid \theta \in \Theta\}$$

Reduction

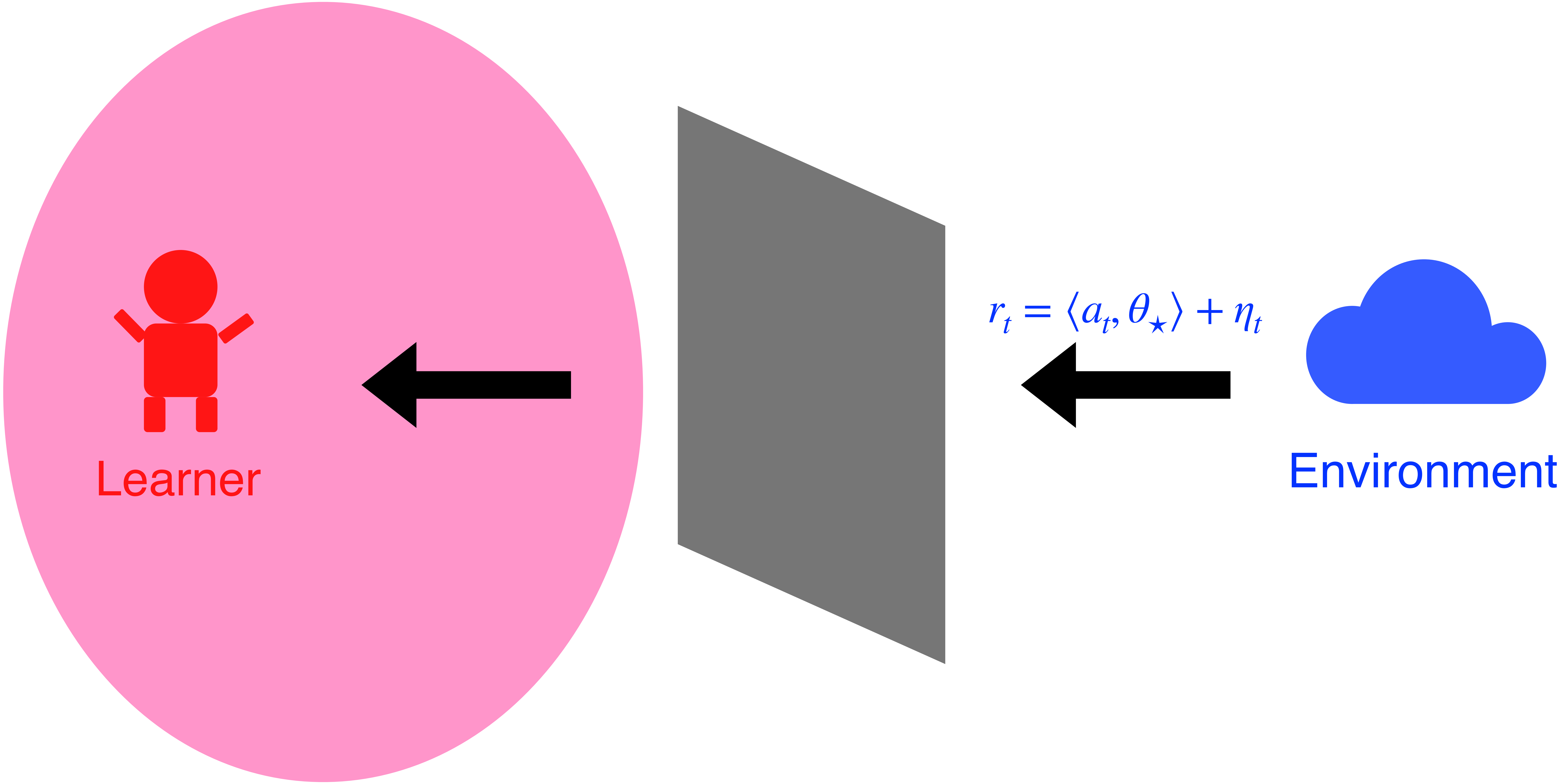

$x_t = g(\theta_t) \in \mathscr{X}$
Learner
Use LB Alg
to approximate $\theta_\star$
drawing action from $\mathscr{X}$
Environment

Reduction

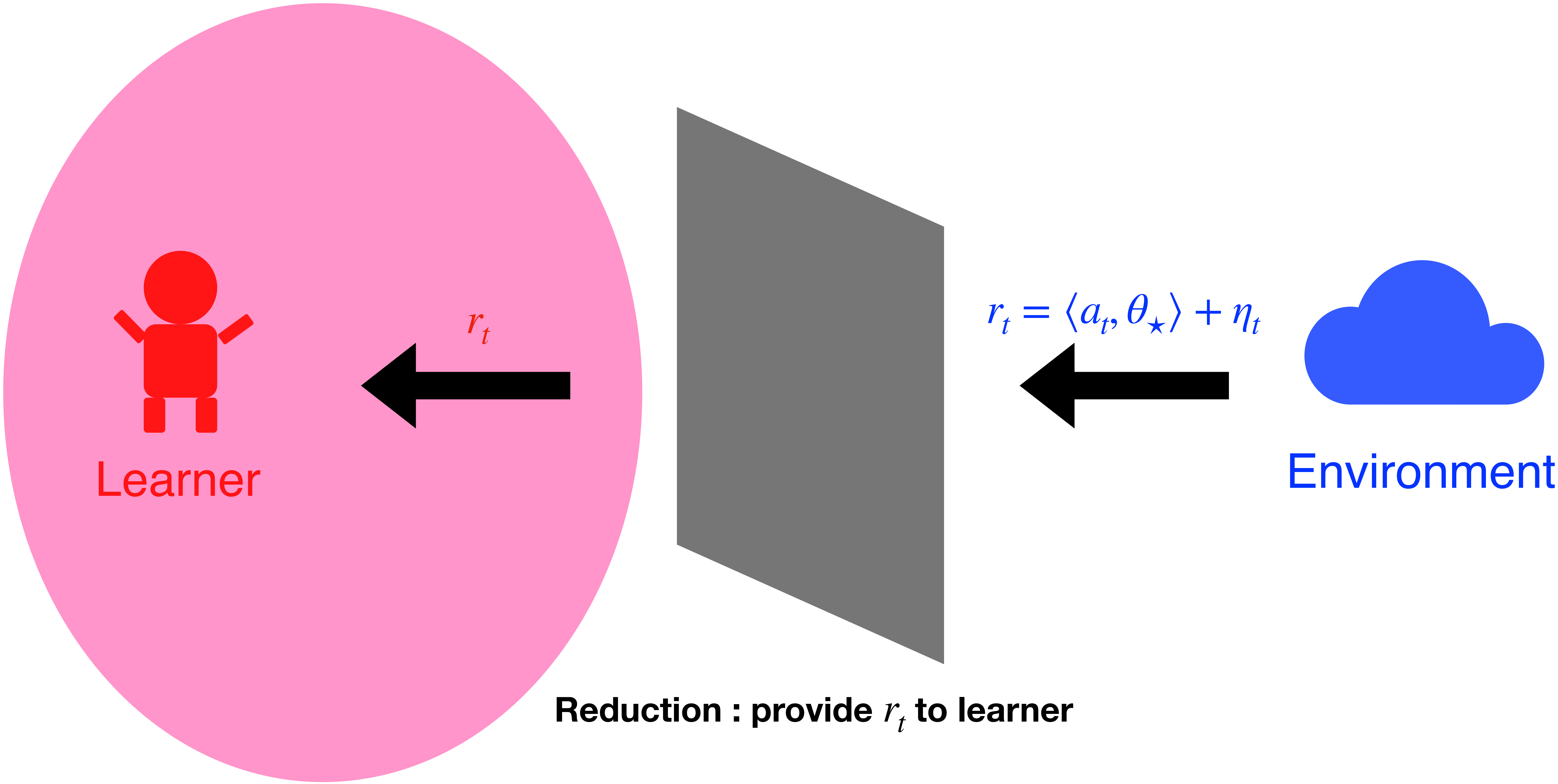


Reduction: use θ_t but play action $a_t \in \mathcal{A}_t$

Reduction

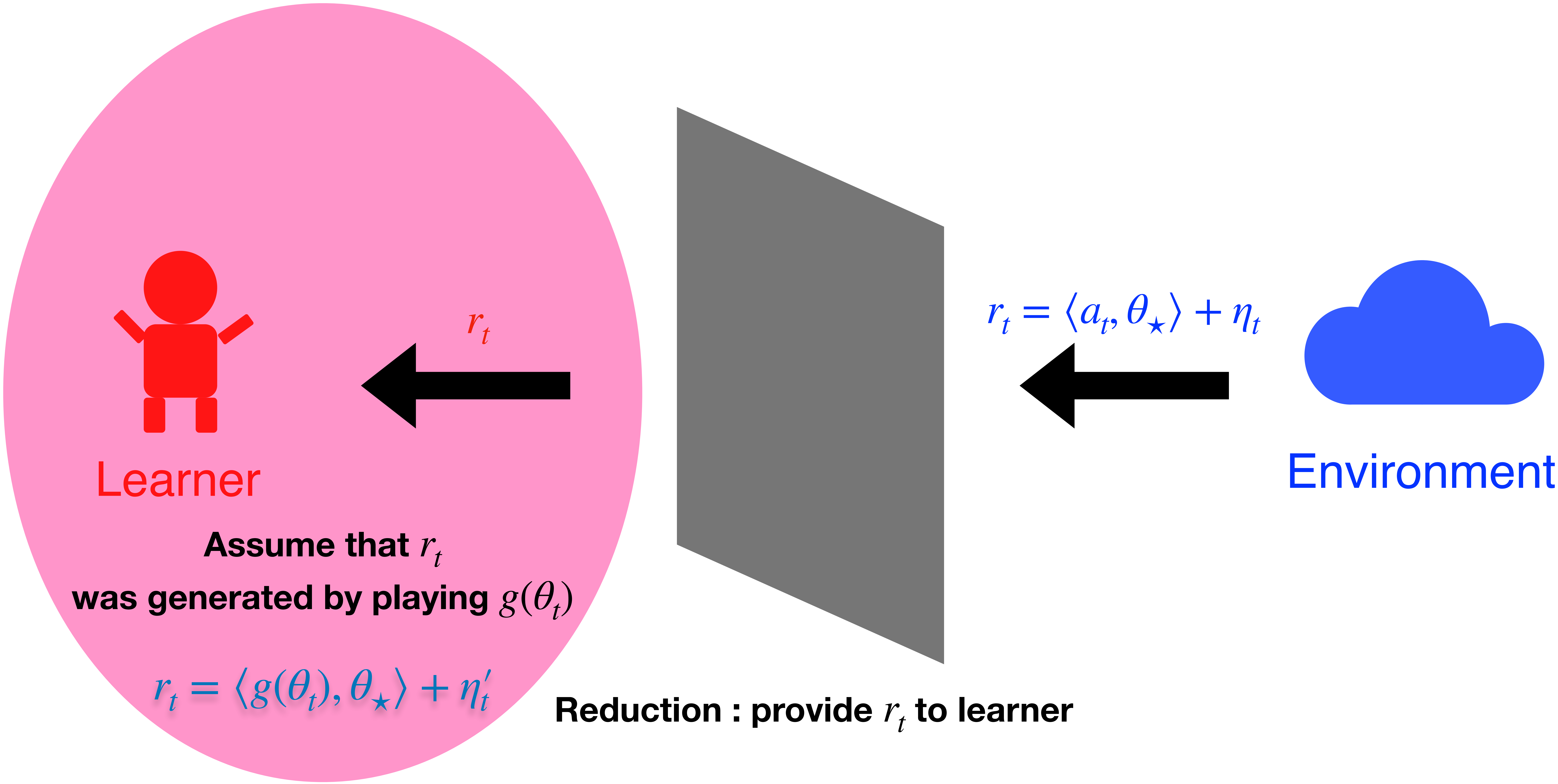


Reduction



Reduction : provide r_t to learner

Reduction



r_t
$r_t = \langle a_t, \theta_\star \rangle + \eta_t$
Learner
Environment
Assume that r_t
was generated by playing $g(\theta_t)$
$r_t = \langle g(\theta_t), \theta_\star \rangle + \eta'_t$
Reduction : provide r_t to learner

Main proof idea

Theorem:

$$|R_T^L - R_T^I| = O(\sqrt{T \log T}) \text{ w.h.p.}$$

1) ***Reward indeed can be expressed as:***

$$r_t = \langle g(\theta_t), \theta_\star \rangle + \eta_t'$$

Main proof idea

Theorem:

$$|R_T^L - R_T^I| = O(\sqrt{T \log T}) \text{ w.h.p.}$$

1) ***Reward indeed can be expressed as:***

$$r_t = \langle g(\theta_t), \theta_\star \rangle + \eta_t'$$

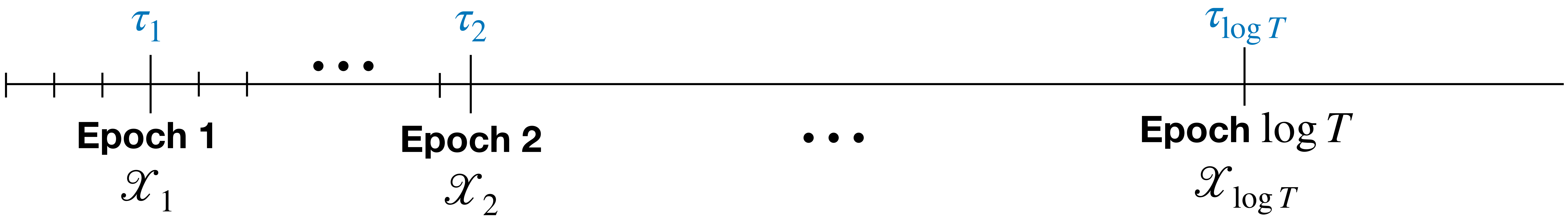
2) ***Difference between regrets of the two instances is not large:***

$$\arg\max_{x \in \mathcal{X}} \langle x, \theta_\star \rangle = g(\theta_\star) \quad \Longrightarrow \quad |R_T^L - R_T^I| = O(\sqrt{T \log T}) \text{ w.h.p.}$$

Unknown Distribution: empirically estimate X

$$g_m(\theta) = \frac{1}{\tau_m} \sum_{t=1}^{\tau_m} \arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle$$

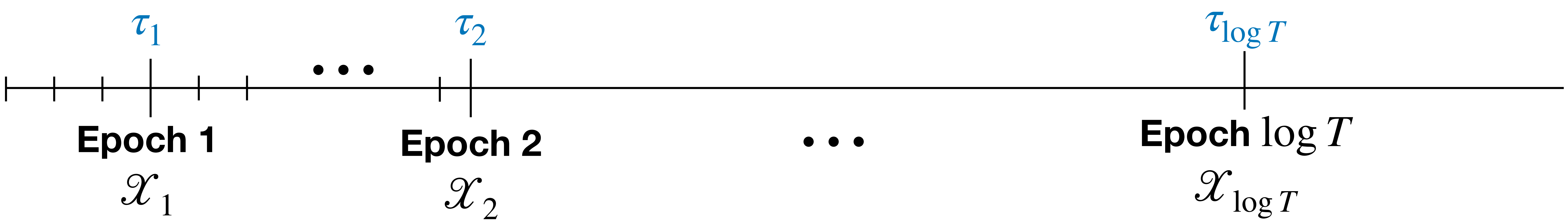
$$\tau_m = e^m, m = 1, \cdots, \log T$$



Unknown Distribution: empirically estimate X

$$g_m(\theta) = \frac{1}{\tau_m} \sum_{t=1}^{\tau_m} \arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle$$

$$\tau_m = e^m, m = 1, \cdots, \log T$$



$$\mu_{\theta_t} \neq \langle g_m(\theta_t), \theta_\star \rangle$$ ⚠️

Misspecified instance

Misspecified linear bandits

$$\mu_a = \langle a, \theta_\star \rangle + f(a),$$
$$|f(a)| \leq \epsilon$$

Main proof idea

Theorem:

$$|R_T^L - R_T^I| = O(\sqrt{T \log T}) \text{ w.h.p.}$$

1) ***Misspecification is small:***

$$\|g(\theta) - g_m(\theta)\|_2 = O(\sqrt{d/\tau}) \quad (\forall \theta \in \Theta?)$$

Main proof idea

Theorem:

$$|R_T^L - R_T^I| = O(\sqrt{T \log T}) \text{ w.h.p.}$$

1) ***Misspecification is small:***

$$\|g(\theta) - g_m(\theta)\|_2 = O(\sqrt{d/\tau}) \ \forall \theta \in \Theta_D$$

$$\forall \theta \in \Theta \ \exists \theta' \in \Theta_D : \|\theta - \theta'\|_2 \leq 1/T$$

2) $\theta_\star \in \Theta_D$?

g is not smooth

Main proof idea

Theorem:

$$|R_T^L - R_T^I| = O(\sqrt{T \log T}) \text{ w.h.p.}$$

1) ***Misspecification is small:***

$$\|g(\theta) - g_m(\theta)\|_2 = O(\sqrt{d/\tau}) \ \forall \theta \in \Theta_D$$

$$\forall \theta \in \Theta \ \exists \theta' \in \Theta_D : \|\theta - \theta'\|_2 \leq 1/T$$

2) ***Discretized set contains a good action***

g is smooth in a neighborhood of $\theta_\star$

Results

Literature

$$R_t = O(d\sqrt{T}\log T) \text{ w.h.p.}$$

$$R_t = O(d\sqrt{T\log T}\, poly(\log\log T)) \text{ exp.}$$

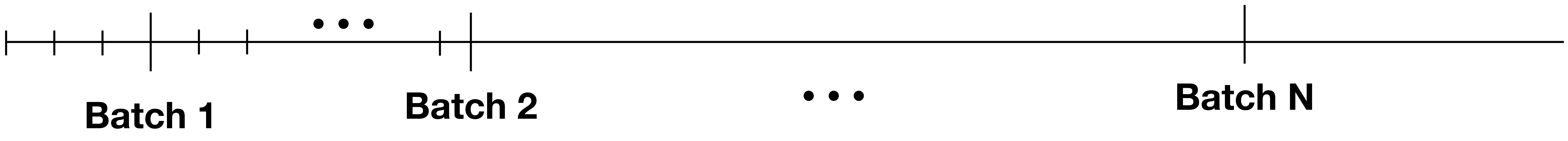
Ours

$$R_t = O(d\sqrt{T\log T}) \text{ w.h.p.}$$

Abbasi-Yadkori, Yasin, Dávid Pál, and Csaba Szepesvári. "Improved algorithms for linear stochastic bandits." Advances in neural information processing systems 24 (2011).

Li, Yingkai, et al. "Tight regret bounds for infinite-armed linear contextual bandits." International Conference on Artificial Intelligence and Statistics. PMLR, 2021.

Results: batch learning



Limited number of policy switches at preselected time instances

Results: batch learning

Literature

$R_t = O(d\sqrt{T \log d \log T} poly(\log\log T))$ exp.

#batches $= O(\log\log T)$

For contexts generated from a distribution

Ours

$R_t = O(d\sqrt{T \log T} \log\log T)$ w.h.p.

#batches $= O(\log\log T)$

Ruan, Yufei, Jiaqi Yang, and Yuan Zhou. "Linear bandits with limited adaptivity and learning distributional optimal design." Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing. 2021.

Results: misspecified setting

Literature

$R_t = O(d\sqrt{T}\log T + \epsilon\sqrt{d}T)$ in exp.

ϵ : amount of misspecification

Ours

$R_t = O(d\sqrt{T\log T} + \epsilon\sqrt{d}T\log T)$ w.h.p.

Foster, Dylan J., et al. "Adapting to misspecification in contextual bandits." Advances in Neural Information Processing Systems 33 (2020): 11478-11489.

Results: adversarial corruption

Results: adversarial corruption

Literature

$$R_t = \tilde{O}(d^{4.5}\sqrt{T} + d^4 C) \text{ w.h.p.}$$

C : amount of corruption

Ours

$$R_t = \tilde{O}(d\sqrt{T} + d^{3/2} C) \text{ w.h.p.}$$

Foster, Dylan J., et al. "Adapting to misspecification in contextual bandits." Advances in Neural Information Processing Systems 33 (2020): 11478-11489.

Results: s-sparse θ

$$\theta_\star = [\mu_1, \cdots, \mu_s, 0, 0 \ldots, 0]^T$$

Results: s-sparse θ

Literature

$$R_t = O(\sqrt{dsT} \log T) \text{ w.h.p.}$$

Ours

$$R_t = O(\sqrt{dsT \log T}) \text{ w.h.p.}$$

Abbasi-Yadkori, Yasin, David Pal, and Csaba Szepesvari. "Online-to-confidence-set conversions and application to sparse stochastic bandits." Artificial Intelligence and Statistics. PMLR, 2012.

Results: batch learning with s-sparse θ

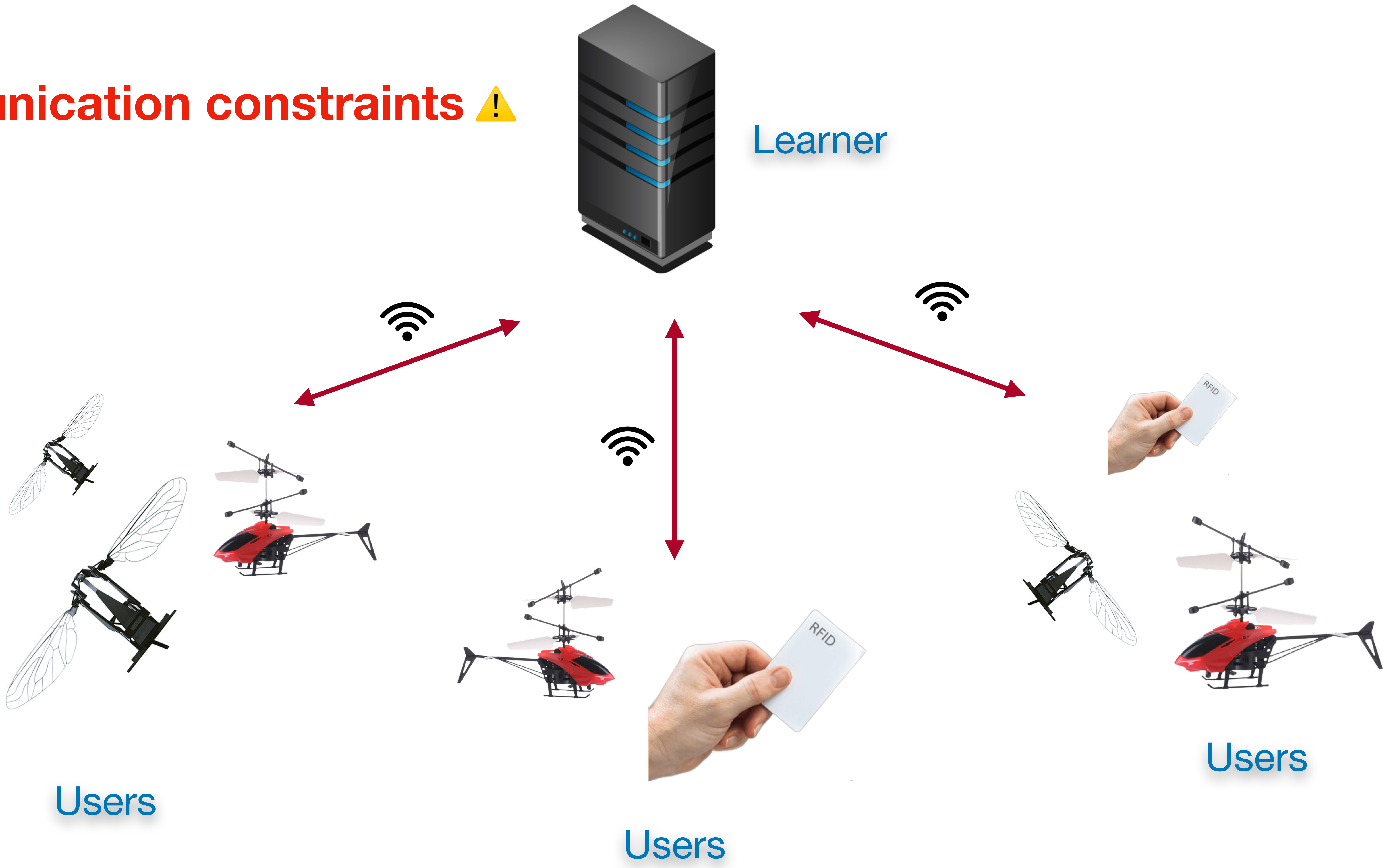
Ours

$$R_t = O(\sqrt{dsT \log T} \log \log T) \text{ w.h.p.}$$

$$\#\text{batches} = O(\log \log T)$$

Distributed contextual bandits

communication constraints ⚠️



Distributed contextual bandits

Reward compression:

Context compression:

Hanna, Osama A., Lin Yang, and Christina Fragouli. "Solving multi-arm bandit using a few bits of communication." International Conference on Artificial Intelligence and Statistics. PMLR, 2022.

Hanna, Osama, Lin Yang, and Christina Fragouli. "Learning from Distributed Users in Contextual Linear Bandits Without Sharing the Context." Advances in Neural Information Processing Systems 35 (2022): 11049-11062.

Distributed contextual bandits

Reward compression:

≈ 3 bits are enough

Context compression:

Hanna, Osama A., Lin Yang, and Christina Fragouli. "Solving multi-arm bandit using a few bits of communication." International Conference on Artificial Intelligence and Statistics. PMLR, 2022.

Hanna, Osama, Lin Yang, and Christina Fragouli. "Learning from Distributed Users in Contextual Linear Bandits Without Sharing the Context." Advances in Neural Information Processing Systems 35 (2022): 11049-11062.

Distributed contextual bandits

Reward compression:

≈ 3 bits are enough

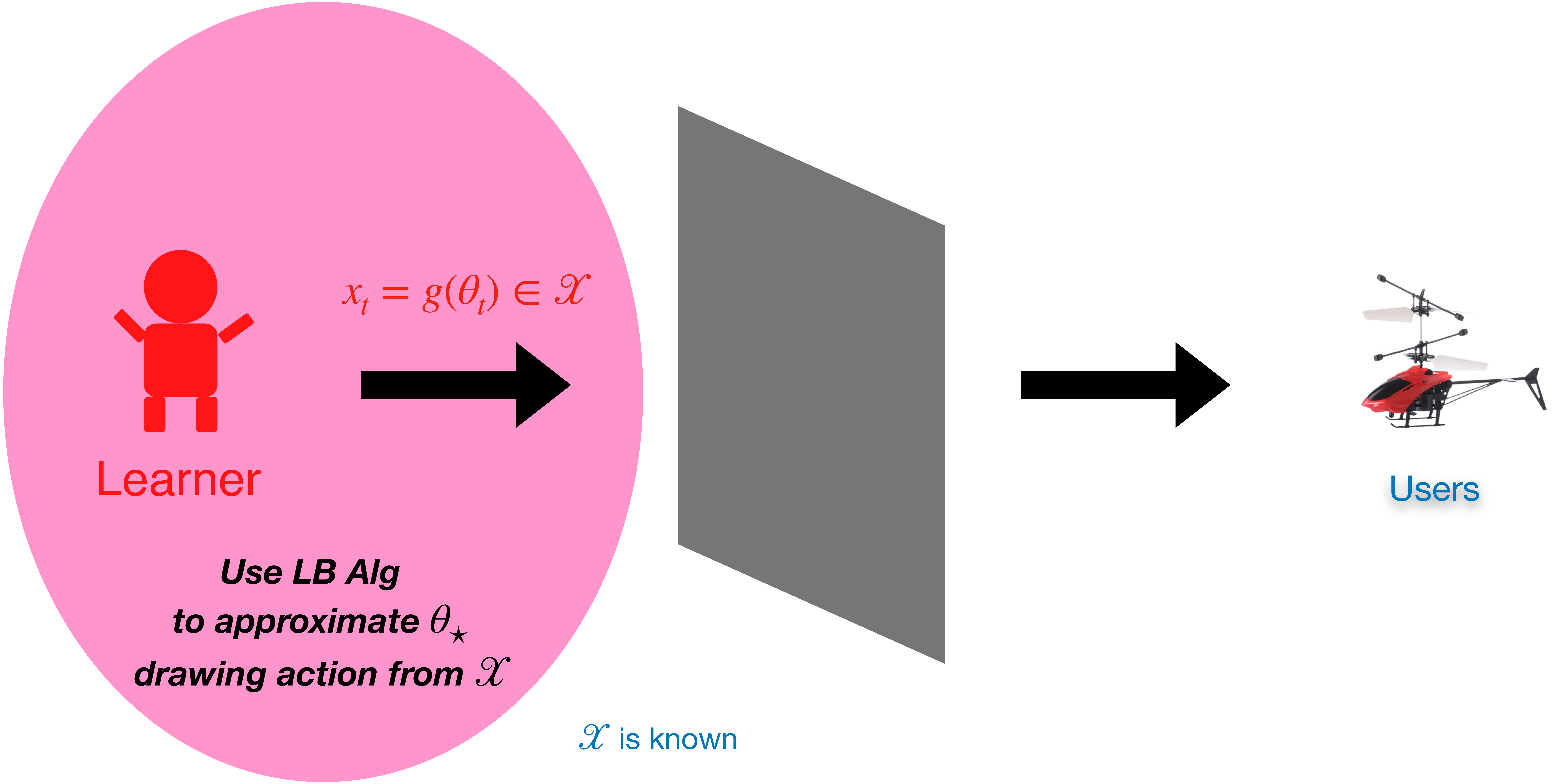
Context compression:

No need to share if context distribution is known!

Hanna, Osama A., Lin Yang, and Christina Fragouli. "Solving multi-arm bandit using a few bits of communication." International Conference on Artificial Intelligence and Statistics. PMLR, 2022.

Hanna, Osama, Lin Yang, and Christina Fragouli. "Learning from Distributed Users in Contextual Linear Bandits Without Sharing the Context." Advances in Neural Information Processing Systems 35 (2022): 11049-11062.

Reduction

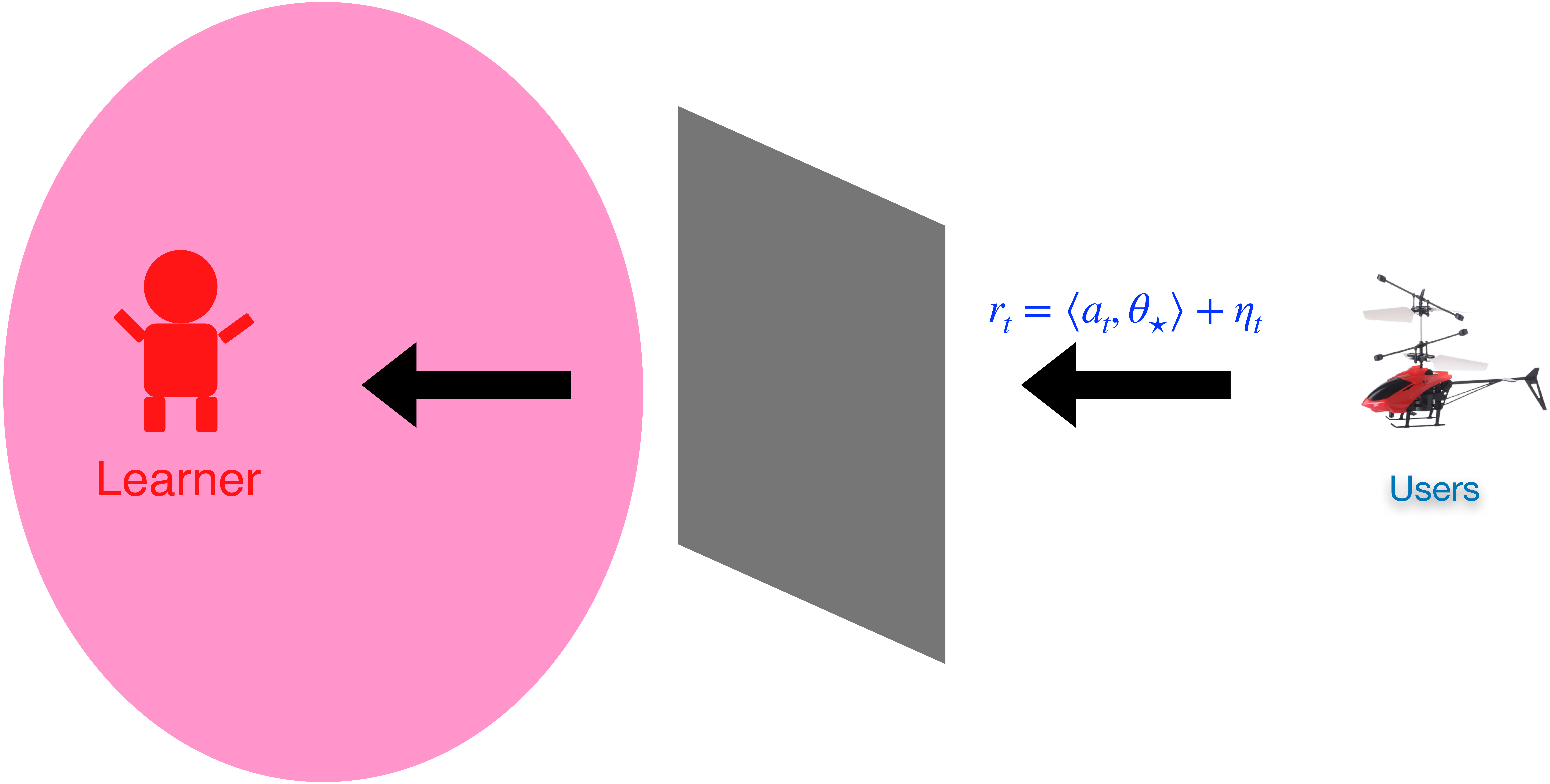


$x_t = g(\theta_t) \in \mathscr{X}$
Learner
Use LB Alg
to approximate $\theta_\star$
drawing action from $\mathscr{X}$
$\mathscr{X}$ is known
Users

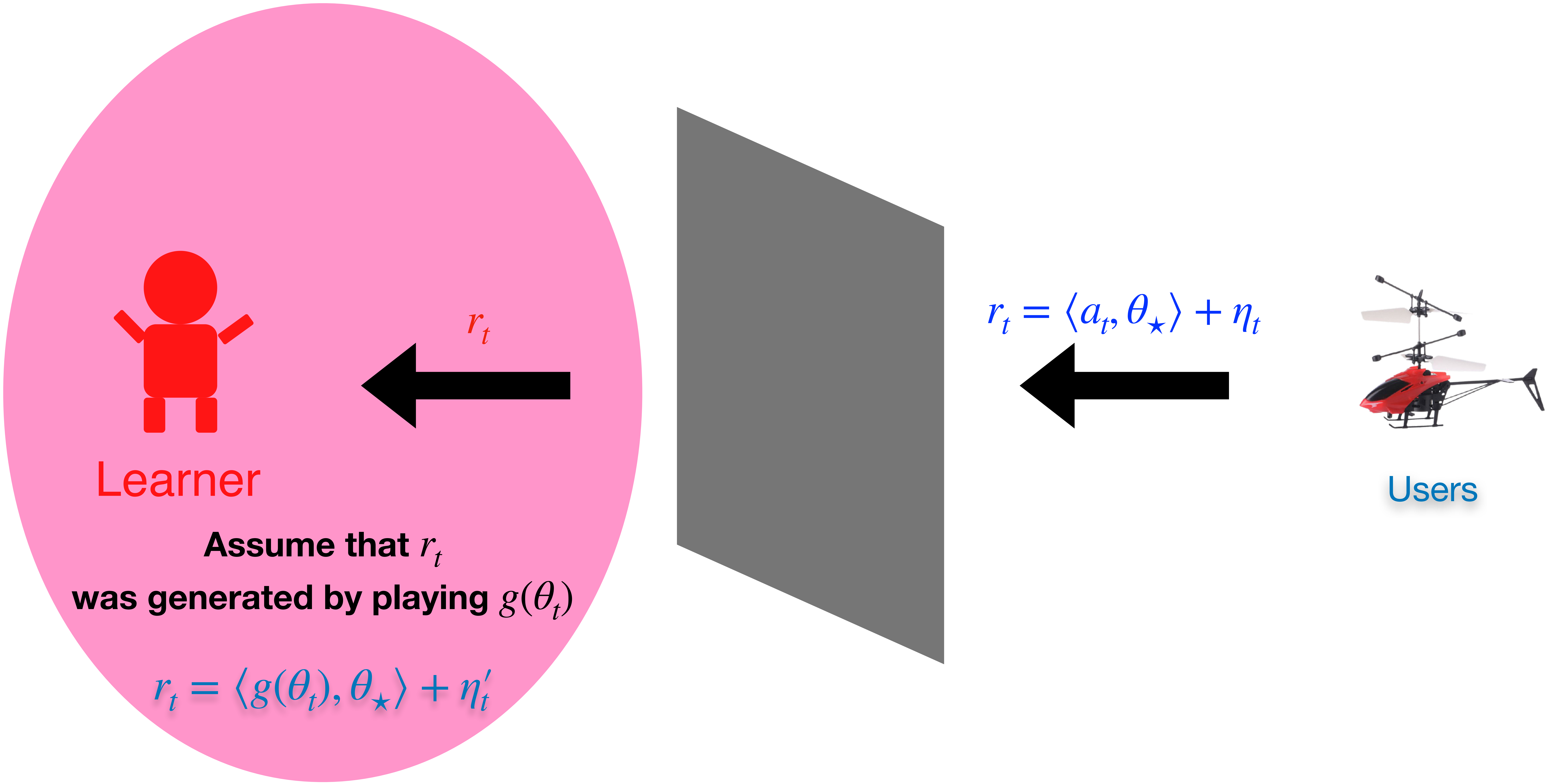
Reduction

$x_t = g(\theta_t) \in \mathcal{X}$
Learner
Use LB Alg
to approximate $\theta_\star$
drawing action from $\mathcal{X}$
$\mathcal{X}$ is known
θ_t
Users
Observe context $\mathcal{A}_t$ and use θ_t to play
$a_t = \arg\max_{z \in \mathcal{A}_t} \langle z, \theta_t \rangle$

Reduction



Reduction



r_t
Learner
Assume that r_t
was generated by playing $g(\theta_t)$
$r_t = \langle g(\theta_t), \theta_\star \rangle + \eta'_t$
$r_t = \langle a_t, \theta_\star \rangle + \eta_t$
Users

Complexity

$$\mathscr{X}_m = \{g_m(\theta)\,|\,\theta \in \Theta_D\}$$ ⚠️

Complexity

$$\mathcal{X}_m = \{g_m(\theta) \,|\, \theta \in \Theta_D\}$$

- Linear optimization and linear regression oracles are **sufficient**!

Ito, Shinji, et al. "Oracle-efficient algorithms for online linear optimization with bandit feedback." Advances in Neural Information Processing Systems 32 (2019).

Complexity

$$\mathcal{X}_m = \{g_m(\theta) \,|\, \theta \in \Theta_D\}$$

- Linear optimization and linear regression oracles are **sufficient**!
- Given $\arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle$, we can solve $\arg\max_{x \in \mathcal{X}_m} \langle x, \theta \rangle$

Ito, Shinji, et al. "Oracle-efficient algorithms for online linear optimization with bandit feedback." Advances in Neural Information Processing Systems 32 (2019).

ANY QUESTIONS?