# Contextual Search in the Presence of Irrational Agents

Chara Podimata, Harvard University

*Joint work with Akshay Krishnamurthy (MSR NYC), Thodoris Lykouris (MSR NYC → MIT), and Robert Schapire (MSR NYC)*

# Contextual/Feature-Based Pricing

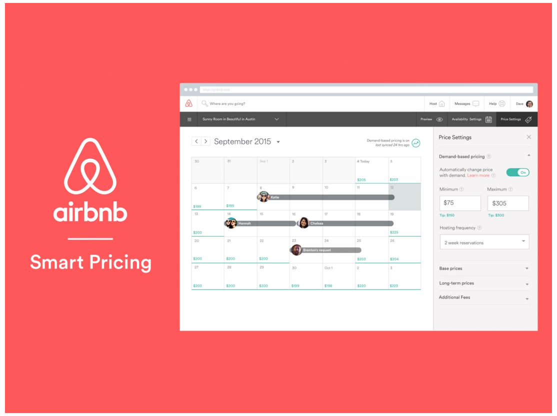
<u>Repeat:</u>

Nature chooses $d$-dimensional context $x_t$

1. Product arrives with *features (location, #rooms, etc)*
2. Airbnb suggests *a price*
3. User *purchases or not*
4. Airbnb *observes feedback (if purchase occurred)*
5. Airbnb *collects revenue*

Learner chooses price $p_t \in \mathbb{R}$

- Exists common feature value $\theta^\star \in \mathbb{R}^d$.
- Agent buys **only if** context valuation above price: $\langle x_t, \theta^\star \rangle \geq p_t$




Learner's feedback: $1\{\langle x_t, \theta^\star \rangle \geq p_t\}$

Learner's reward: $p_t \cdot 1\{\langle x_t, \theta^\star \rangle \geq p_t\}$

Set prices to collect as much revenue as possible.

$Regret$ = Total revenue for benchmark price – Total revenue collected

# Behavioral Model Core Assumption

## Full Rationality

- Exists common feature value $\theta^\star \in \mathbb{R}^d$.
- Agent buys **only if** their context valuation above price: $\langle x_t, \theta^\star \rangle \geq p_t$

!

But *sometimes* agents deviate from the prescribed behavioral model (irrationality)

e.g., agents may not buy, even if price is below their valuation, due to their psychological state

## Problem

Model misspecifications can cause arbitrary failures to standard algorithms

## Main Question

How do we design contextual search algorithms that are robust to the presence of some irrational agents?

# Contextual Pricing for Fully Rational Agents

For round $t = 1 \ldots T$:

1. Nature chooses observable context $x_t \in \mathbb{R}^d$ with $||x_t||_2 \leq 1$

[Cohen, Lobel, Paes Leme, EC16/MS19]
[Lobel, Paes Leme, Vladu, EC17/OR18]
[Paes Leme, Schneider, FOCS18]
[Liu, Paes Leme, Schneider, SODA21]

2. Learner selects *a price* $p_t$
3. Agent makes purchase if $\langle x_t, \theta^\star \rangle \geq p_t$
4. Learner observes if purchase occurred *(one-sided feedback)*
5. Learner collects revenue if purchase occurred

$$Regret = O(d \log T)$$

Benchmark pricing policy: $p_t^\star = \langle x_t, \theta^\star \rangle$

- For **revenue** objective, [Liu, Paes Leme, Schneider, SODA21] improved to: $O(d \log \log T)$
- For $d = 1$, optimal solution due to [Kleinberg and Leighton, FOCS03]

# Contextual Pricing for **Adversarially Irrational** Agents

Total corruption: $C = \sum_t c_t$

For round $t = 1 \dots T$:

1. Nature chooses observable context $x_t \in \mathbb{R}^d$ with $||x_t||_2 \leq 1$

   **and corruption $\boldsymbol{c_t} \in \{\mathbf{0}, \mathbf{1}\}$**

2. Learner selects *a price* $p_t$
3. Agent makes purchase if $\langle x_t, \theta^\star \rangle \geq p_t$ **if $\boldsymbol{c_t} = \mathbf{0}$**

   **or arbitrarily if $\boldsymbol{c_t} = \mathbf{1}$**

4. Learner observes if purchase occurred *(one-sided feedback)*
5. Learner collects revenue if purchase occurred

Extends the corruptions model studied in:
[Lykouris, Mirrokni, Paes Leme STOC18]
[Gupta, Koren, Talwar COLT19]
[Zimmert, Seldin JMLR21]

**Desiderata for our Algorithms**

✓ Graceful degradation of regret with $C$
✓ No knowledge of $C$ assumed (*agnostic*)

Total corruption: $C = \sum_t c_t$
Number of features: $d$

## Main Result

Robust contextual search algorithm agnostic to $C$ achieving
$Regret \leq C \cdot d^3 \cdot (\log T)^3$ (with high probability)

[Ulam 76], [Rivest, Meyer, Kleitman, Winklmann, Spencer, JCSS80], [Karp and Kleinberg, SODA07]
For Ulam's game ($d = 1$) tight bound of $\log T + C \log C + C \log \log T$ [RMKWS, JCSS80]

Extends classical Ulam's game (aka twenty questions with a liar) in 2 ways:

1. Multidimensional
2. Agnostic to $C$

# The Non-Robust Algorithm

[Lobel, Paes Leme, Vladu, EC17/OR18]

## Multidimensional Binary Search

- Maintain active knowledge set with feasible values for $\theta^\star$.
- Set price $\langle \kappa_1, x_1 \rangle$ where $\kappa_1$ centroid of knowledge set.
- Eliminate inconsistent side of knowledge set.

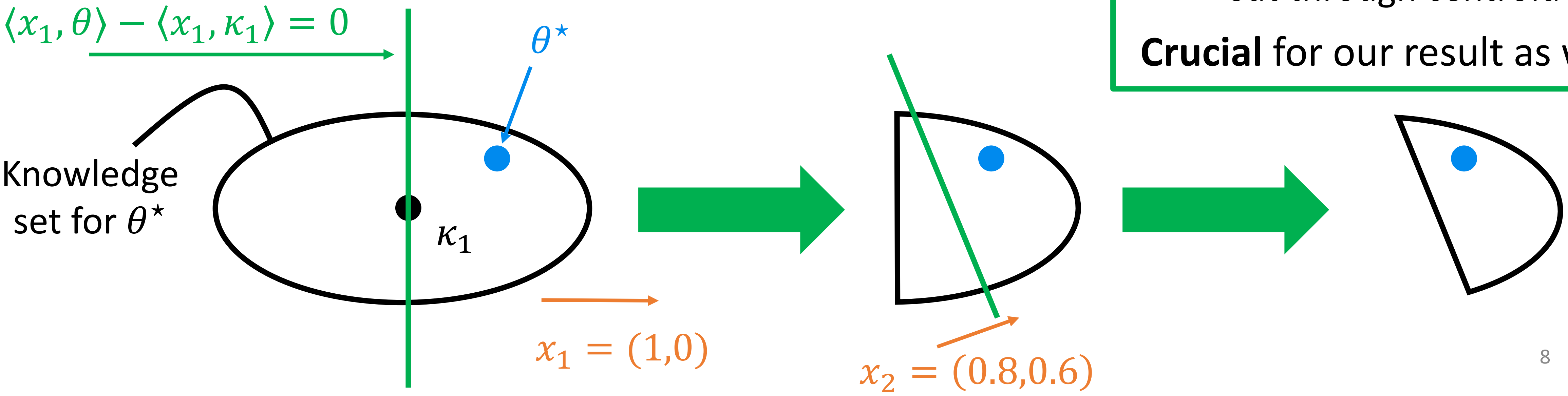
**Important properties of cut**

1. Never eliminate $\theta^\star$
   - Retain all parameters consistent with feedback
2. Volumetric progress
   - Cut through centroid

**Crucial** for our result as well

Cut: hyperplane perpendicular to context

$$\boxed{Regret \leq d \cdot \log T}$$



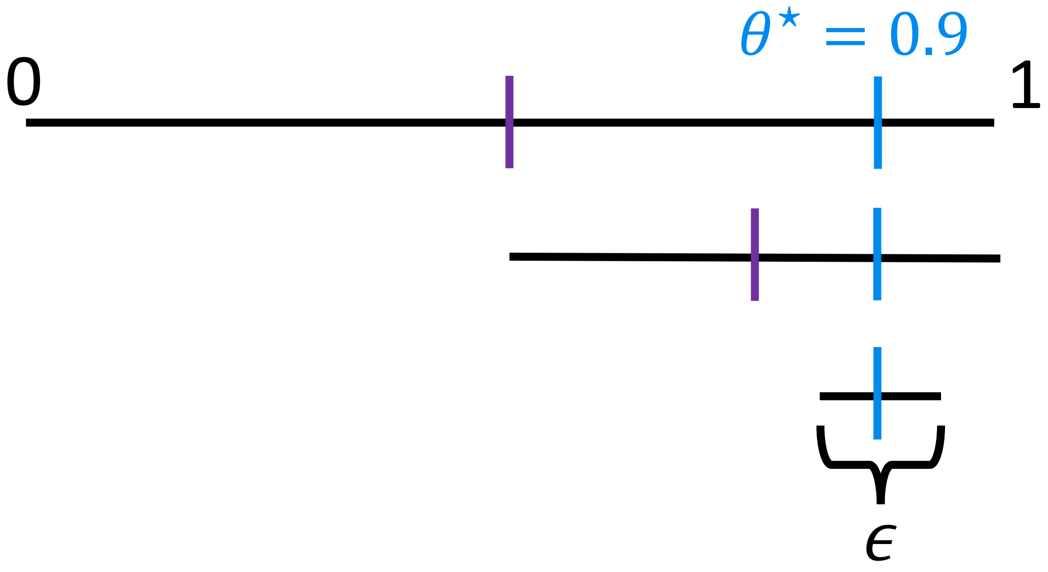
# Issue with Adversarial Irrationality: Ulam's Game

**Challenge**

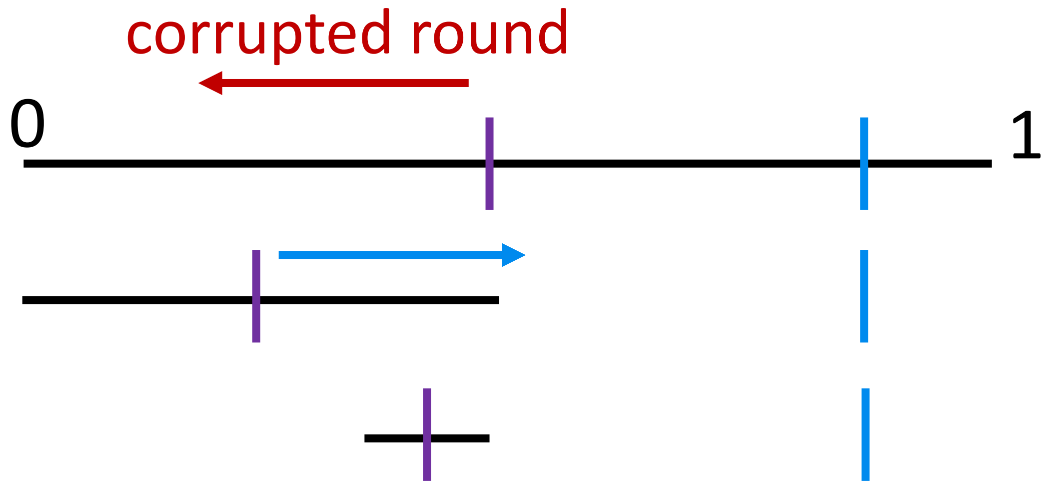
How can you fix it (in high dimensions) when you know that $\bar{c}$ responses are corrupted?

<u>Binary search ($d = 1$)</u>

- I fix a number $\theta^\star \in [0,1]$.
- How many queries do you need to find it with accuracy $\epsilon$?




… when *one response* is corrupted?

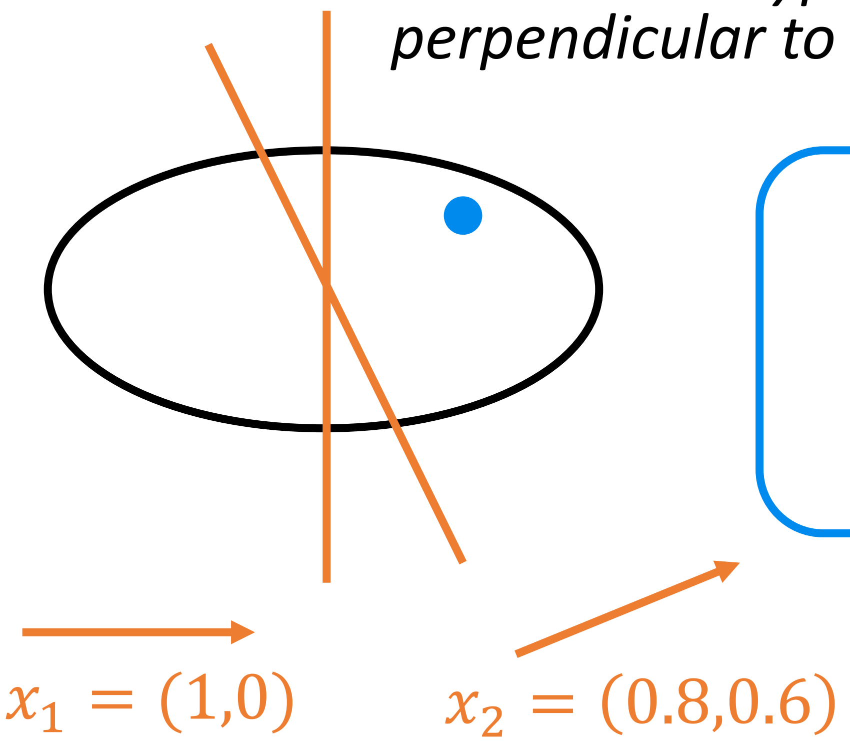


# Technical Crux: Robust Volumetric Progress

**Challenge 1**

We cannot repeat the same query (contexts are different at different rounds).

*Context cut: hyperplane perpendicular to context*




**Idea 1**

Keep "penalty" for each parameter & make cut once a context cut fully retains protected region on one side

- Protected region: all parameters with "penalty" $\leq \bar{c}$

# Technical Crux: Robust Volumetric Progress

**Challenge 2**

We may never have a context cut with the protected region fully on one side.

**Counterexample:** Even with infinite contexts and $\bar{c} = 1$, no such context cut.

# Technical Crux: Robust Volumetric Progress

**Challenge 2**

We may never have a context cut with the protected region fully on one side.

**Counterexample:** Even with infinite contexts and $\bar{c} = \mathbf{1}$, no such context cut.

**Idea 2**

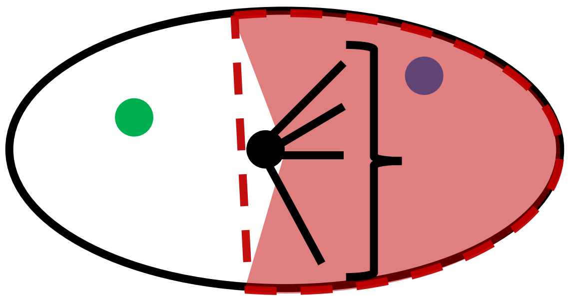
Combine context cuts to compute a "valid cut".

**Important properties of valid cut**

1. Never eliminate $\theta^\star$
   - Fully retains protected region on one side.
2. Volumetric progress
   - Cross **close** to centroid

**Idea 3**

Show that $2d \cdot (d+1) \cdot \bar{c} + 1$ context cuts have enough information to compute such a valid cut **(Caratheodory's theorem)**



# Technical Crux: Robust Volumetric Progress

**Challenge 2**

We may never have a context cut with the protected region fully on one side.

**Idea 2**

Combine context cuts to compute a "valid cut".

**Idea 3**

Show that $2d \cdot (d+1) \cdot \bar{c} + 1$ context cuts have enough information to compute such a valid cut **(Caratheodory's theorem)**

**Counterexample:** Even with infinite contexts and $\bar{\boldsymbol{c}} = \mathbf{1}$, no such context cut.

**Important properties of valid cut**

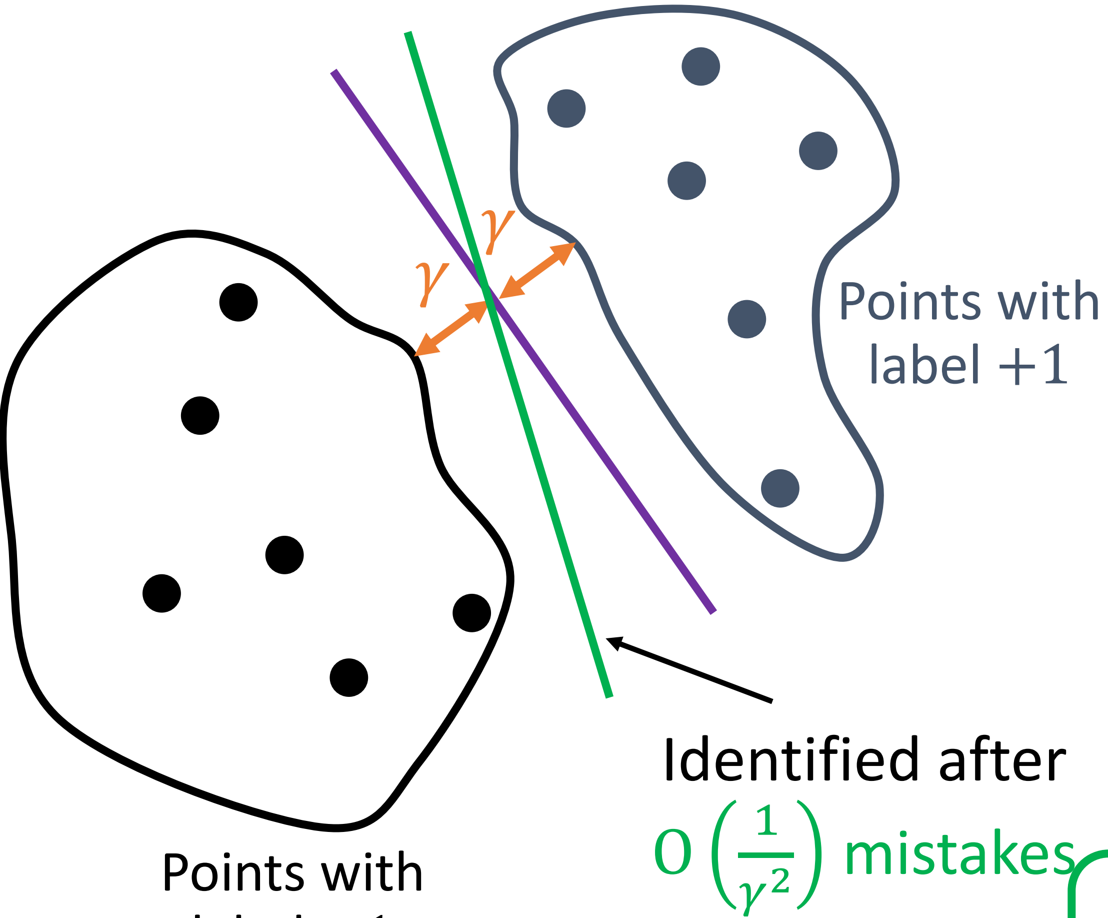
1. Never eliminate $\theta^\star$
   - Fully retains protected region on one side.
2. Volumetric progress
   - Cross **close** to centroid

**Idea 4**

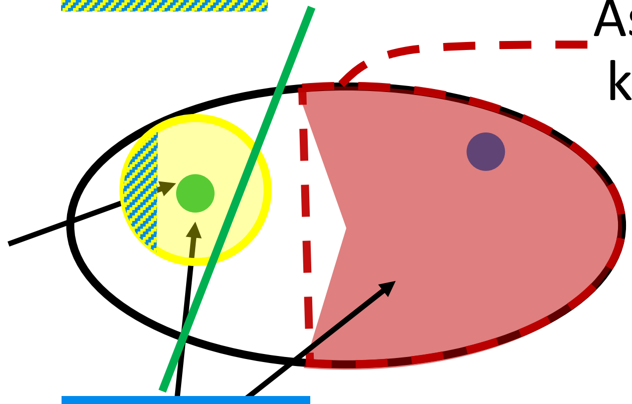
Use Perceptron as a **witness** for finding a valid cut.

# Perceptron as a Witness

<u>Standard Perceptron</u>




- Sample randomly from ball
- With constant probability, you will sample point from pattern.




**Idea**

If we knew them exactly + they had margin → Perceptron → done

- If dataset correct → Perceptron mistakes $O\left(\frac{1}{\gamma^2}\right)$
- Else, resample

# Unknown $C$: Layering to the Rescue!

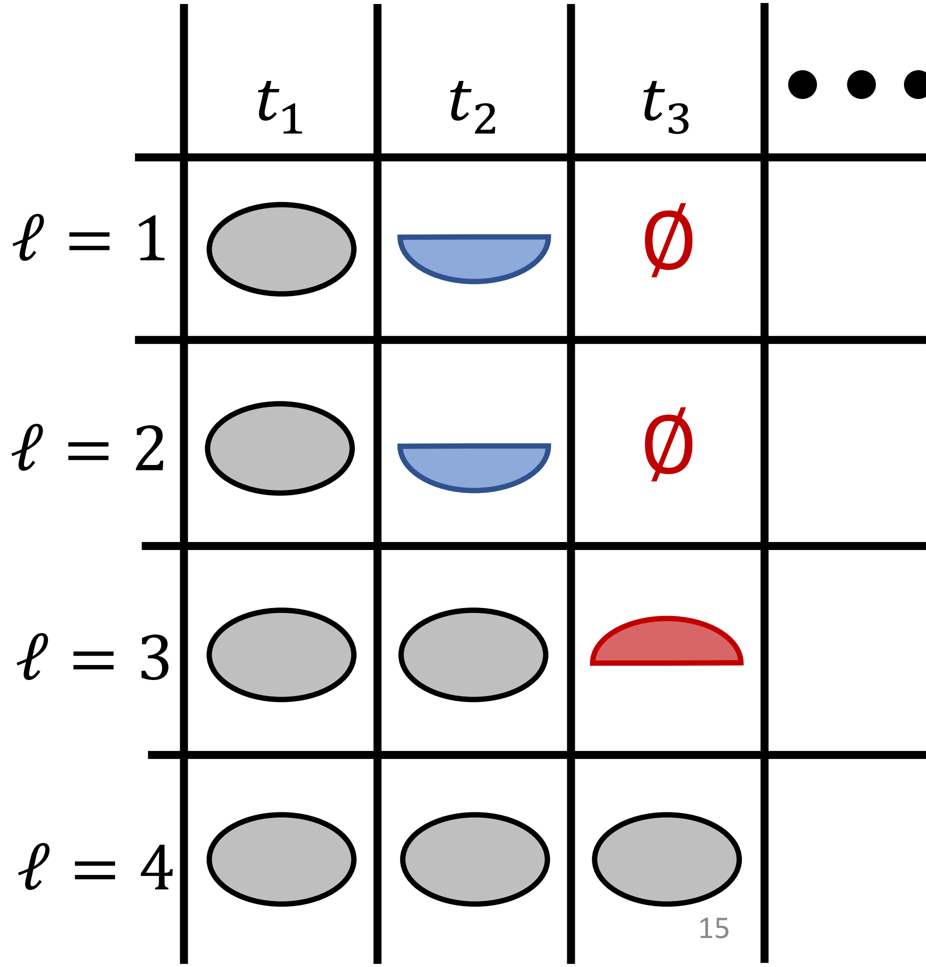
Extension of the multi-layering race technique of [Lykouris, Mirrokni, Paes Leme, STOC18] for continuous action spaces.

**Subsampling**

- $\log T$ layers
- Layer $i$ corresponds to corruption level: $C = 2^i$
- Sample a layer $\ell$ to play, with probability $2^{-\ell}$

**Global eliminations**

- Maintain consistent knowledge sets
- Below layers: more robust → they should inform above layers.



# Summary

## Theorem

Feature-based pricing with *adversarial irrationality*: $\boldsymbol{Regret} \leq \boldsymbol{C} \cdot \boldsymbol{d}^{\mathbf{3}} \cdot (\log \boldsymbol{T})^{\mathbf{3}}$

## Technical crux

1. **Cannot repeat same query:** *Work with protected region*
2. **Context cuts do not suffice:** *Combine context cuts to find a valid cut*
3. **Number of context cuts needed:** $2d \cdot (d+1) \cdot \bar{c} + 1$ *context cuts (Caratheodory's theorem)*
4. **Computation of valid cut:** *Novel use of Perceptron algorithm as "witness"*

Our multi-layering race runs parallel versions with $\bar{\boldsymbol{c}} = \log T$ .

## Also in the paper:

1. **Bounded rational agents**
2. **Gradient-descent based algorithm for contextual search**


Thank You!