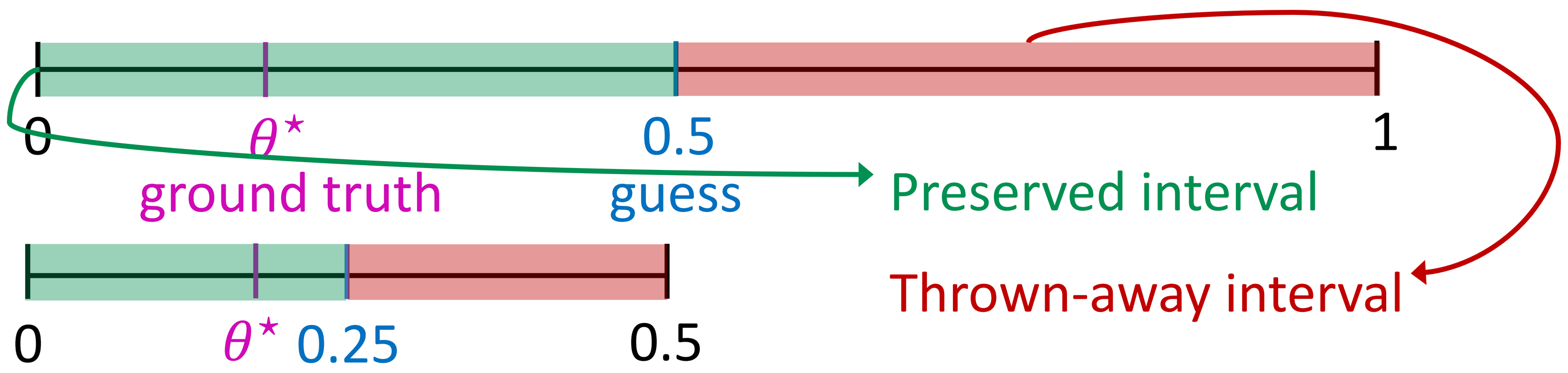
Corrupted Multidimensional Binary Search: Learning in the Presence of Irrational Agents

Akshay Krishnamurthy (MSR NYC), Thodoris Lykouris (MSR NYC), Chara Podimata (Harvard)

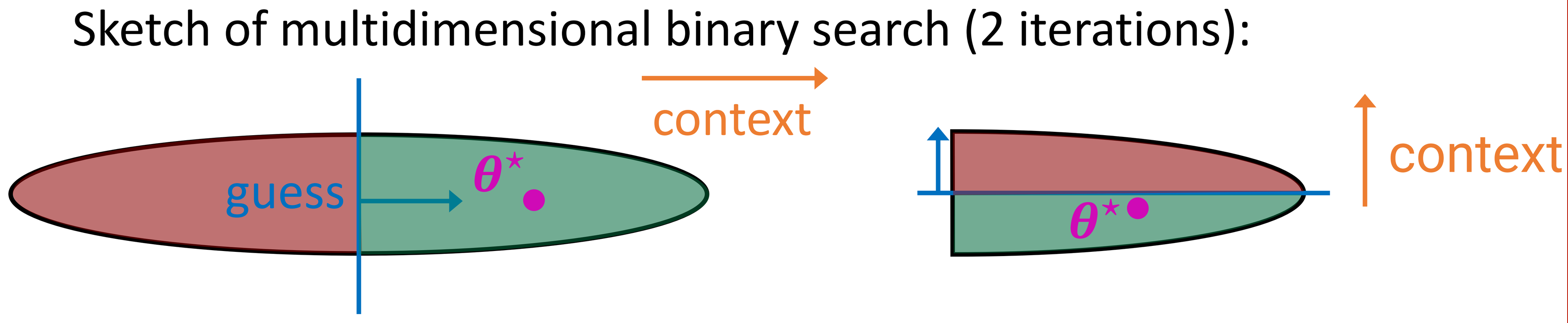
Main Question

Multidimensional binary search: core primitive in a lot of online learning problems (e.g., dynamic pricing)

Sketch of single-dimensional binary search (2 iterations):



Sketch of multidimensional binary search (2 iterations):



Core assumption in Algorithmic Game Theory: responses of agents are fully rational (i.e., agents' responses are consistent with $\boldsymbol{\theta}^\star$).

In the language of dynamic pricing: $\boldsymbol{\theta}^\star$: value vector, context: features of items sold, guess: posted price, sold item or not

How do we design learning algorithms (for multidimensional binary search) that are robust to the presence of some irrational agents?

"irrational" responses: inconsistent with $\boldsymbol{\theta}^\star$

→ Even few irrational responses, can have catastrophic effects in standard online learning algorithms.

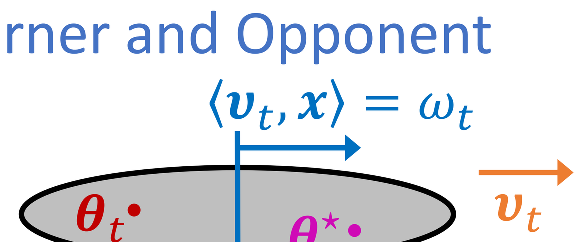
Model

$\boldsymbol{\theta}^\star \in \mathbb{R}^d$: $\|\boldsymbol{\theta}^\star\| \leq 1$: ground truth (unknown to the learner)

$\varepsilon > 0$: target accuracy

Repeated Interaction Protocol Between Learner and Opponent

For round $t \in [T]$:



1. Opponent chooses context $\boldsymbol{v}_t \in \mathbb{R}^d$.
2. Learner observes $\boldsymbol{v}_t$.
3. Opponent corrupts ($\boldsymbol{\theta}_t =$ arbitrary) or not ($\boldsymbol{\theta}_t = \boldsymbol{\theta}^\star$).
   → happens at most C times
4. Learner queries $\omega_t \in \mathbb{R}$ & observes feedback:
   $y_t = \text{sign}(\langle \boldsymbol{v}_t, \boldsymbol{\theta}_t \rangle - \omega_t) \in \{-1, 1\}$.
5. Learner incurs loss: $\ell(\omega_t, \boldsymbol{\theta}_t) = 1\{|\omega_t - \langle \boldsymbol{v}_t, \boldsymbol{\theta}_t \rangle| > \varepsilon\}$.

Learner's Goals

(1) Minimize: $R(T) = \sum_{t=1}^{T} \ell(\omega_t, \boldsymbol{\theta}_t)$

Same for **standard** multidimensional binary search, except $\boldsymbol{\theta}_t = \boldsymbol{\theta}^\star, \forall t$

(2.1) $R(T)$ degrading gracefully to C

(2.2) $R(T)$: agnostic C

[Lobel, Paes Leme, Vladu, '17] (**without corruptions**): $R(T) = O(d \log(d/\varepsilon))$

Applications for learning in the presence of irrational agents:

Contextual dynamic pricing with linear buyers, Stackelberg Security Games, contextual dynamic pricing with Lipschitz buyers

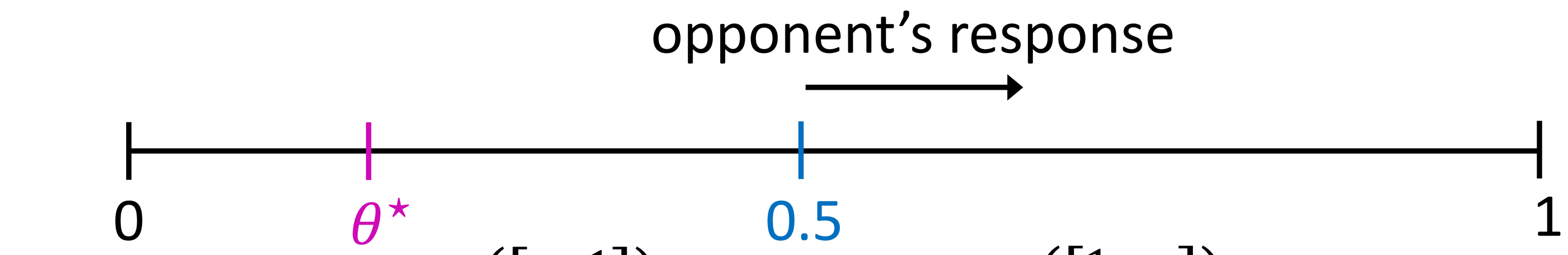
Main Result

CorPV.A runs in **quasipolynomial** time, is **agnostic in C** and achieves **regret**: $R(T) = O(d^3(C + \log T) \log T \log(d/\varepsilon))$

Key Idea: Undesirability Levels

Known C

High level: undesirability expresses the likelihood of where $\boldsymbol{\theta}^\star$ lies



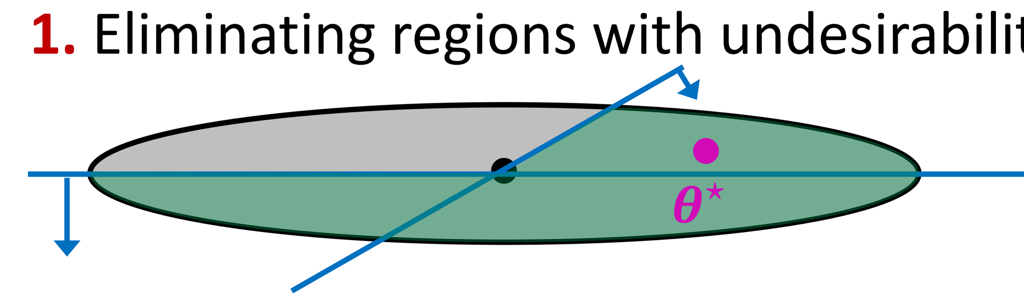
Undesirability: $u\left(\left[0, \frac{1}{2}\right]\right) = 1$ $\quad u\left(\left[\frac{1}{2}, 1\right]\right) = 0$

After $2C + 1$ rounds, $u\left(\left[0, \frac{1}{2}\right]\right) \leq C$ but $u\left(\left[\frac{1}{2}, 1\right]\right) \geq C + 1$

→ Generalize the undesirability levels idea for higher dimensions.

Challenges and Solutions

1. Eliminating regions with undesirability $\geq C + 1$ → non-convexity

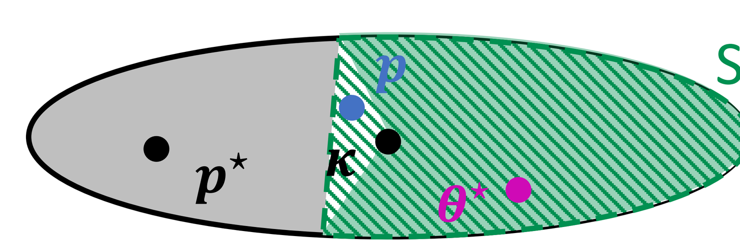


Solution: Always make hyperplane cuts

2. Existence of hyperplane cut with undesirability $\geq C + 1$.

Known C

Solution: After $2dC(d + 1) + 1$ contexts, there always exists a hyperplane with undesirability at least $C + 1$!



Shaded region ($\mathcal{P}(C)$): points with $u \leq C$

Patterned region: convex hull of $\mathcal{P}(C)$

[Carathéodory's Theorem] Every point $\boldsymbol{p}$ in the convex hull of $\mathcal{P}(C)$ can be written as the convex combination of ***at most $d + 1$ points*** in $\mathcal{P}(C)$.

→ $u(\boldsymbol{p}) \leq C(d + 1)$

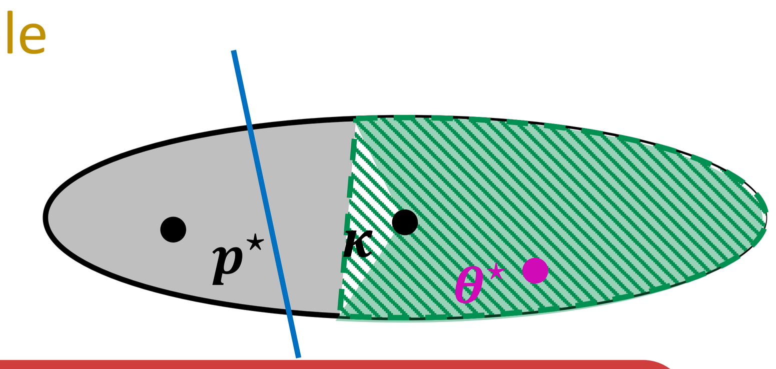
→ For separating $\mathcal{P}(C)$, suffices to have point: $u(\boldsymbol{p}^\star) = C(d + 1) + 1$

→ [Landmarks.] Set of $2d$ points such that at each round one of them gets undesirability point. + pigeonhole

3. Computation of hyperplane cut.

Known C

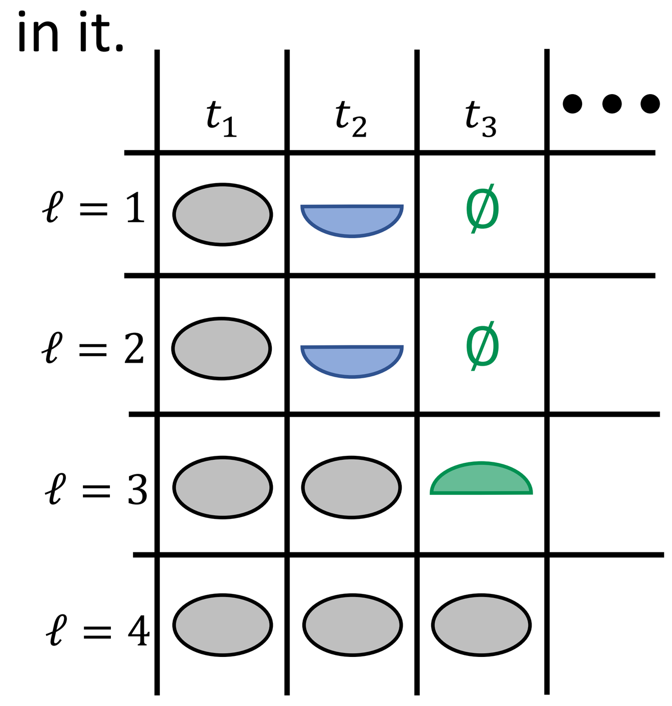
Solution: Building a dataset for perceptron with injected margin.



CorPV.K regret: $O\left(d^2(C + 1)d \log \frac{d}{\varepsilon}\right)$ and expected runtime: $O\left((d^2 C)^C \cdot poly(d \log d/\epsilon), C)\right)$

4. Known C and runtime exponential in it.

Solution: Extension of the multi-layering race technique of [Lykouris et al., '18] for continuous spaces.



References

[Lobel, Paes Leme, Vladu, '17]. Multidimensional Binary Search for Contextual Decision Making. EC'17.

[Lykouris, Mirrokni, Paes Leme, '18]. Stochastic Bandits Robust to Adversarial Corruptions. STOC'18.