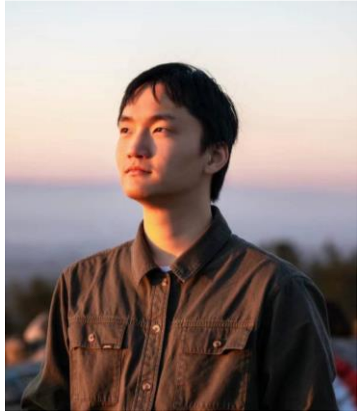
Sample-Based Matroid Prophet Inequalities




(me)

Hu Fu Pinyan Lu Zhihao Gavin Tang
(Shanghai University of Finance and Economics)
Hongxun Wu
(UC Berkeley)
Jinzhao Wu
(Yale)
Qianfan Zhang
(Princeton)

Single-item prophet inequality

- Given n independent distributions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$
- At each step $i = 1,2, \ldots, n$
  - Inspect $v_i \sim D_i$
  - Decide to accept/reject v_i immediately and irrevocably
- Goal: maximize the (expected) accepted value
  - While a *prophet* can get $\mathbf{E}[\max_i v_i]$

Images from "Slay the spire"

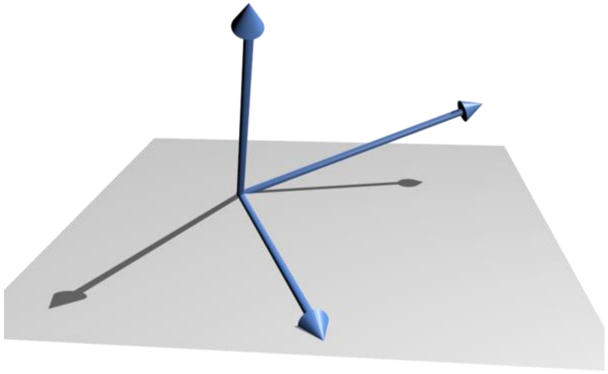
Single-item prophet inequality

- Given n independent distributions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$
- At each step $i = 1,2, \ldots, n$
  - Inspect $v_i \sim D_i$
  - Decide to accept/reject v_i immediately and irrevocably
- Goal: maximize the (expected) accepted value
  - While a *prophet* can get $\mathbf{E}[\max_i v_i]$

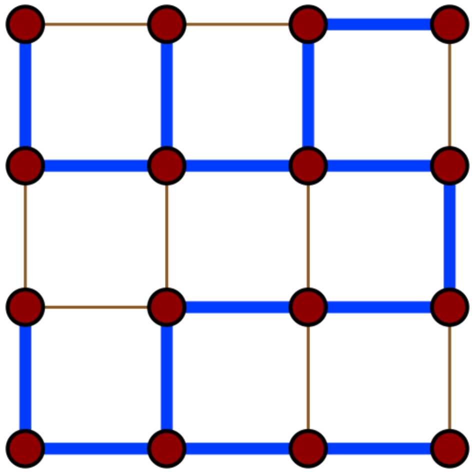
- [Krengel, Sucheston, Garling '78] There exists an algorithm with value $\geq \frac{1}{2} \mathbf{E}\left[\max_i v_i\right]$

Different feasibility constraints?

- At most k items
  - k-uniform matroids



- Linearly independent set of vectors
  - Linear matroids



- Set of edges that does not contain a cycle
  - Graphical matroids

Matroid prophet inequalities

- Given $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n$ **and a family of feasible sets $\boldsymbol{\mathcal{F}} \subseteq \mathbf{2}^{[n]}$** that is a matroid
- At each step $i = 1,2,\dots,n$
  - See $v_i \sim D_i$
  - Decide to accept/reject v_i
    - **Feasibility:** Must keep set of accepted items $A \in \mathcal{F}$
- Goal: maximize the (expected) **sum of** accepted values
  - While a *prophet* can get $\mathbf{E}[\max_{S \in \mathcal{F}} \sum_{i \in S} v_i]$
- [Kleinberg, Weinberg '12] There exists an algorithm with value $\geq \frac{1}{2} \mathbf{E}\left[\max_{S \in \mathcal{F}} \sum_{i \in S} v_i\right]$
- [Hajiaghayi, Kleinberg, Sandholm '07] [Chawla, Hartline, Malec, Sivan '10] Applications to auctions
  - But often only samples to distribution is available!

Sample-based prophet inequalities

- Given K independent samples from $\mathcal{D}_1, \mathcal{D}_2 \dots, \mathcal{D}_n$ each and $\mathcal{F} \subseteq 2^{[n]}$
- At each step $i = 1,2, \dots, n$
  - See $v_i \sim D_i$
  - Decide to accept/reject v_i subject to feasibility constraint $\mathcal{F}$
- Goal: maximize the (expected) sum of accepted values

- [Azar, Kleinberg, Weinberg '14] [Caramanis, Dütting, Faw, Fusco, Lazos, Leonardi, Papadigenopoulos, Pountourakis, Reiffenhäuser '22] [Kaplan, Naori, Raz '24] [Correa, Dütting, Fischer, Schewior '19] [Rubinstein, Wang, Weinberg '20] [Kaplan, Naori, Raz '20] [Correa, Dütting, Fischer, Schewior, Ziliotto '21] [Correa, Cristi, Epstein, Soto '23] [Correa, Cristi, Epstein, Soto '20] [Cristi, Ziliotto '24] ...

Sample-based prophet inequalities

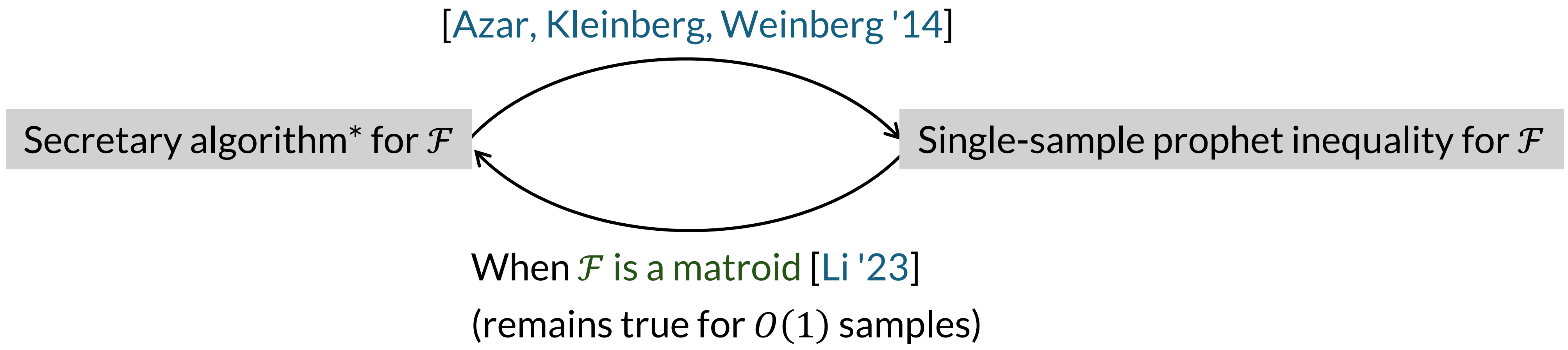
- Single-sample
  - Single-item prophet inequalities [Rubinstein, Wang, Weinberg ,20]
  - Special matroids [Azar, Kleinberg, Weinberg '14] [Caramanis, Dütting, Faw, Fusco, Lazos, Leonardi, Papadigenopoulos, Pountourakis, Reiffenhäuser '22]
  - Matchings [Caramanis, Dütting, Faw, Fusco, Lazos, Leonardi, Papadigenopoulos, Pountourakis, Reiffenhäuser '22] [Kaplan, Naori, Raz ’24]

Sample-based matroid prophet inequalities

- <u>Secretary problem</u>: values are chosen *adversarially*, but items arrive in a *uniformly random* order
  - **Long-standing open question:** $O(1)$-competitive secretary algorithm for general matroids

    ⬆

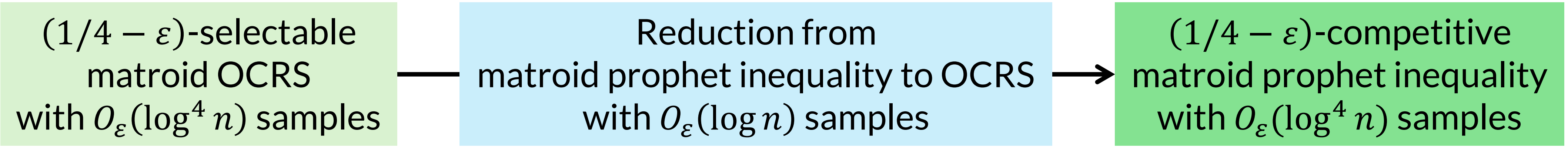
    $O(1)$-competitive matroid prophet inequality with $O(1)$ samples



Sample-based matroid prophet inequalities:
Existing results

		This work		
Ratio	$\frac{1}{O(\log\log rank)}$	$\frac{1}{4}-\varepsilon$	$\frac{1}{4}$	$\frac{1}{2}$
#(samples)	1	$O_\varepsilon(\log^4 n)$	$\text{poly}(n)$	∞
	Reduce to *Secretary Problem*	*Quantile-based* threshold strategy + OCRS	Reduce to *Online Contention Resolution Schemes*	*Expectation-based* threshold strategy
	[Azar, Kleinberg, Weinberg '14] [Lachish '14] [Feldman, Svensson, Zenklusen '14]		[Feldman, Svensson, Zenklusen '16]	[Kleinberg, Weinberg '12]

Our framework



Threshold strategies for single-item

- Idea: accept when $v_i \geq \boldsymbol{T}$

- [Samuel-Cahn '84] Median-based $T = Median(\max_i v_i)$
  - Don't know how to generalize to matroid
- [Kleinberg, Weinberg '12] Expectation-based $T = \frac{1}{2} E[\max_i v_i]$
  - Lack of concentration

Median-based threshold for matroids

- $\tau_i(v_{-i})$: minimum weight for i to be in OPT
- Let i be **active** whenever $v_i \geq T_i = \mathrm{Median}(\tau_i)$
  - Easy to estimate by $O(\log n)$ samples

- Ideally: accept every active items

Issue: Feasibility

The active set might not be feasible:

$$\{i\colon v_i \geq T_i\} \notin \mathcal{F}$$

→ Rounding via OCRS

Issue: Optimality

The active set must approximate the optimal set, e.g.,

$$\mathbf{E}\left[\sum_{i\in[n]} v_i \cdot \mathbf{1}[v_i \geq T_i]\right] > \frac{1}{2}\,\mathbf{E}\left[\sum_{i\in \mathrm{OPT}} v_i\right]$$

→ Proof via a global argument

Optimality:
local argument fails

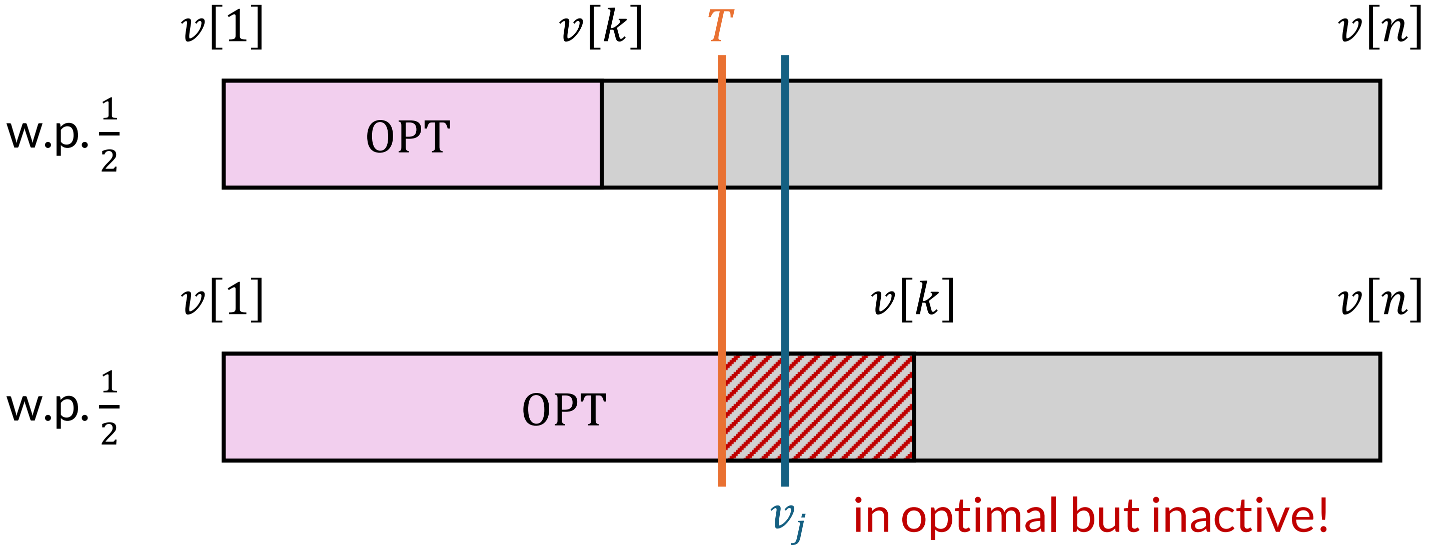
Issue: Optimality
The active set must approximate the optimal set, e.g.,

$$\mathbf{E}[\sum(Active\ items)] > \frac{1}{2}\mathbf{E}[\sum(\mathrm{OPT})]$$

Ideally:

$$E\left[v_j \cdot 1[v_j \geq T]\right] \geq \frac{1}{2} \cdot E\left[v_j \cdot 1[v_j \in OPT]\right]$$

- In k-uniform matroid, $\tau_i = v[k]$ (kth largest value)
  - A single threshold $T = \mathrm{Median}(v[k])$ for all i



- Consider an item j with deterministic value $v_j = T - \varepsilon$
  - $v_j \in \mathrm{OPT}$ w.p. $\frac{1}{2}$ $\rightarrow$ contribution to $\mathbf{E}[\sum(\mathrm{OPT})]$ is $\frac{T-\varepsilon}{2}$, while never active

Optimality for general matroids

- Benchmark: maximum basis B of thresholds $T_i = \mathrm{Median}(\tau_i)$

$$\mathbf{E}[\sum(\mathrm{OPT} \cap Active)] \geq \frac{1}{2}\mathrm{E}[\sum(B)] > \mathbf{E}[\sum(\mathrm{OPT} \cap Inactive)]$$

  - Key tool: Weighted Strong Basis Exchange Lemma [Buchbinder, Feldman, Garg ‘19]

- Actual strategy: a single median → multiple quantiles

Issue: Feasibility
The active set might not be feasible:
$\{i:\ v_i \geq T_i\} \notin \mathcal{F}$

Feasibility:
via Online Contention Resolution Schemes

- Each item i is active independently w.p. p_i
  - **Fact:** $\vec{\boldsymbol{p}} \in (1+\varepsilon) \cdot \mathcal{P}_{\mathcal{F}}$
- At each step $1,2, \ldots, n$:
  - See whether i is active or not
  - **If active**, decide to accept/reject i subject to feasibility constraint $\mathcal{F}$
- Goal: maximize *selectability* α such that

$$\Pr[i \text{ accepted}] \geq \alpha \cdot p_i \qquad \forall i \in [n]$$

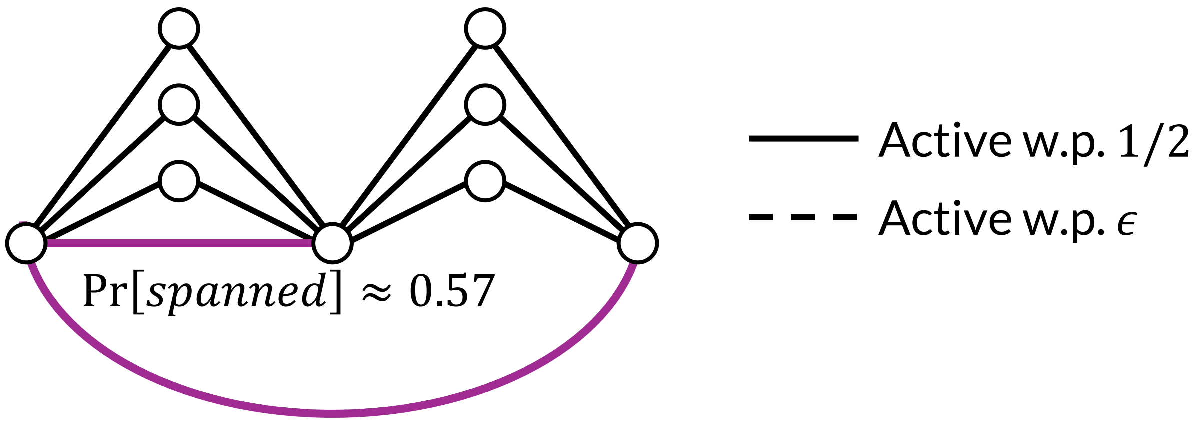
- [Feldman, Svensson, Zenklusen '16] There exists a $\frac{1}{4}$-selectable OCRS for matroids.

$$(\text{Active set}) \xrightarrow{\text{losing factor of } 4} (\text{Feasible set})$$

Constructing OCRS

- [Feldman, Svensson, Zenklusen '16] There exists a $\frac{1}{4}$-selectable OCRS for matroids
- Core subroutine `Protect`(c)
  - $S \leftarrow \emptyset$
  - `While` $(\exists\, i$ such that $\Pr[i \text{ spanned by } \{active\ items\} \cup S] > c)$ `do`
    - $S \leftarrow S \cup \{i\}$
  - `Return` S
- Example: a graphical matroid

`Protect`(0.5)



$\Pr[spanned] \approx \mathbf{0.57}$

Constructing OCRS from samples

- $\widehat{\texttt{Protect}}(c)$: $\text{poly}(n)$ queries in the form

$$\Pr[\boldsymbol{i} \text{ spanned by } \{active\ items\} \cup \boldsymbol{S}] \overset{?}{>} c$$

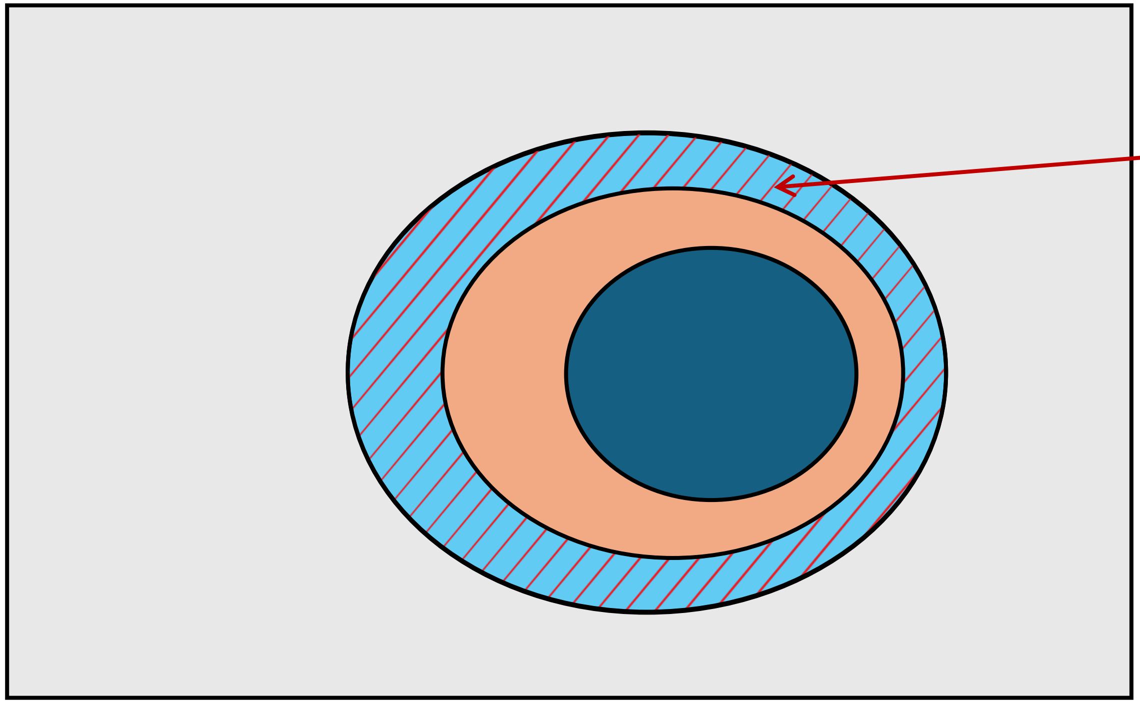
- Can be estimated efficiently!
  - By Chernoff bound + union bound, $O(\log n)$ samples suffices (?)

- No - because queries are **highly adaptive**
  - When true answer lies in $[c - \varepsilon, c + \varepsilon]$, whether $S \leftarrow S \cup \{i\}$ depends on the error
  - Which further determines later queries!

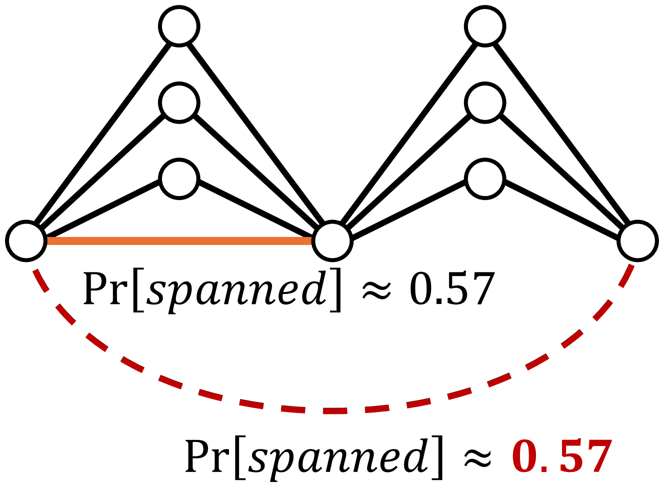
A sandwiching lemma

- **Lemma**: For any constants c, δ, with high probability

$$\mathtt{Protect}(c+\delta) \subseteq \widehat{\mathtt{Protect}}(c) \subseteq \mathtt{Protect}(c-\delta)$$



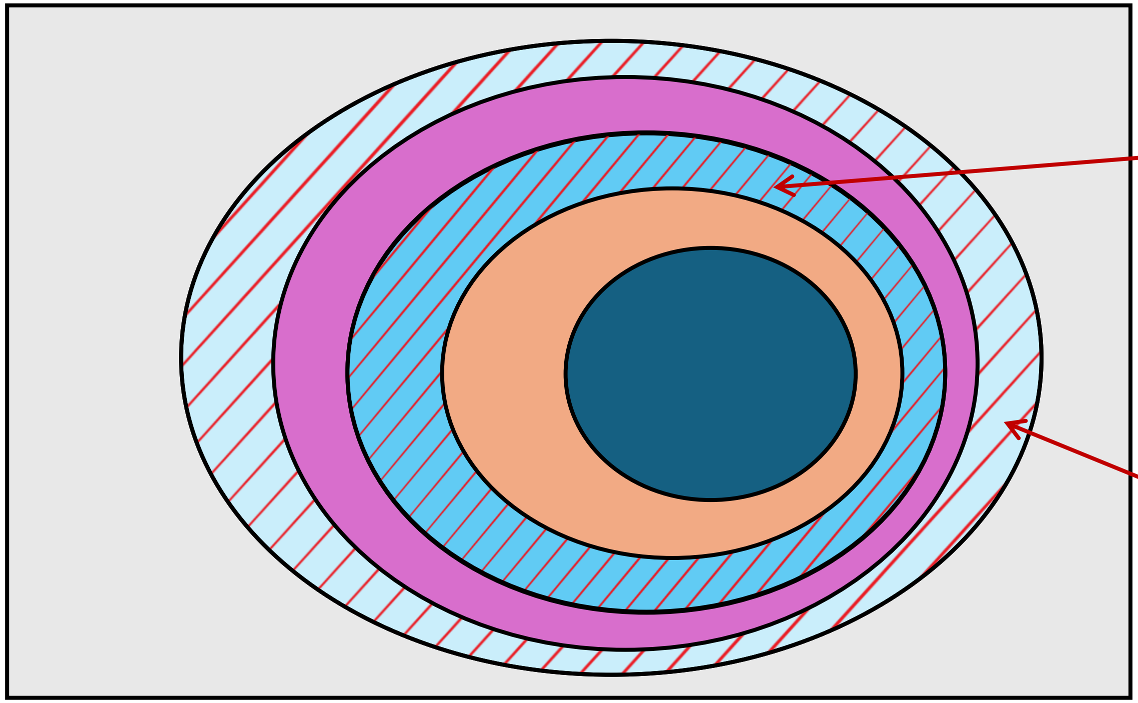
(Possibly) lack of protection against $\widehat{\mathtt{Protect}}(c) \setminus \mathtt{Protect}(c+\delta)$



More sandwichings

- **Lemma**: For any constants c, δ, with high probability

$$\texttt{Protect}(c+\delta) \subseteq \widehat{\texttt{Protect}}(c) \subseteq \texttt{Protect}(c-\delta)$$
$$\subseteq \widehat{\texttt{Protect}}(c-2\delta) \subseteq \texttt{Protect}(c-3\delta)$$



(Possibly) lack of protection against
$\widehat{\texttt{Protect}}(c) \setminus \texttt{Protect}(c+\delta)$

Disjoint!

(Possibly) lack of protection against
$\widehat{\texttt{Protect}}(c-2\delta) \setminus \texttt{Protect}(c-\delta)$

Randomization against adaptivity

- Randomize over 2 parameters → every item fails w.p. at most $\frac{1}{2}$
- Randomize over $\log n$ parameters → every item fails w.p. at most $\frac{1}{\log n}$
    - $O(\log^4 n)$ samples in total

Summary

- Our result:

$(1/4-\varepsilon)$-selectable matroid OCRS with $O_\varepsilon(\log^4 n)$ samples	—	Reduction from matroid prophet inequality to OCRS with $O_\varepsilon(\log n)$ samples	→	$(1/4-\varepsilon)$-competitive matroid prophet inequality with $O_\varepsilon(\log^4 n)$ samples

- Open questions:
  - Matroid OCRS with fewer samples
  - Matroid OCRS better selectability
  - Matroid prophet inequality with fewer samples
    - $O(1)$ samples ⇒ matroid secretary conjecture [Li '23]

Thank you!