AGT
+ Data Science

Jamie Morgenstern, University of Pennsylvania
Vasilis Syrgkanis, Microsoft Research

AGT + DS 1: Sample complexity of auction design

Jamie Morgenstern, University of Pennsylvania

Tutorial Structure:

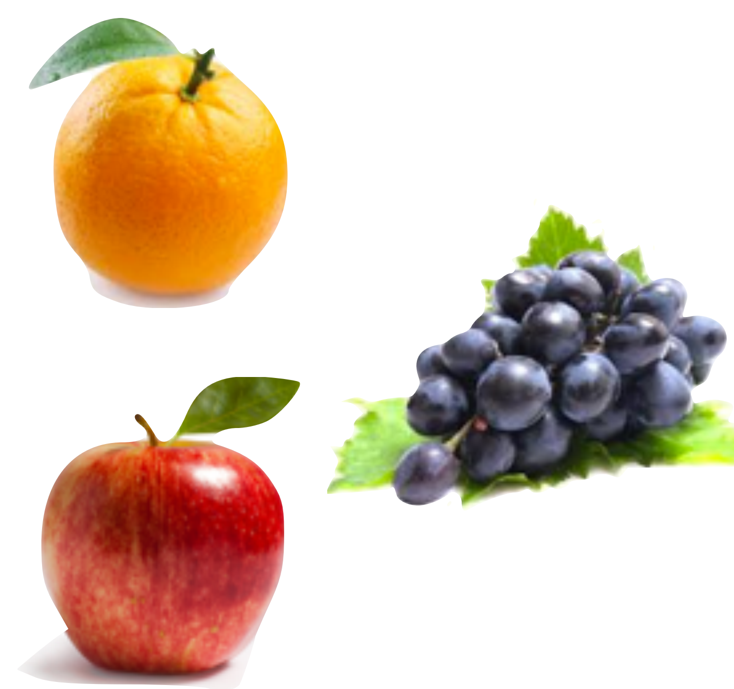
9-10:30 The Sample Complexity of Auction Design [Jamie]

10:30-11 Coffee break

11:00-12:30 Econometrics for Games
[Vasilis]

1:30-5:10 today and 8:30-10:30 tomorrow:
Workshop On Interface Between Algorithmic Game Theory And Data Science

How much do we need to know about buyers to sell to them (nearly) optimally, and how should we sell with limited information about buyers?



Main take-away

"True" Lost Revenue of auction class C on distribution D designed from data

=

Representation error: how much revenue best $c \in C$ loses on true distribution D

+

Generalization error: difference btwn $c \in C$'s revenue on data vs. on D

“Perfect Prior” Model

n buyers, $i \in [n]$ has value $v_i \sim \mathcal{D}_i$

Knowing $\mathcal{D}_1, \ldots, \mathcal{D}_n$, seller picks an auction

Auction’s (expected) revenue measured on

$$(v_1, \ldots, v_n) \sim \mathcal{D}_1 \times \ldots \times \mathcal{D}_n$$

Exact knowledge of prior distributions

[M’81]

“Prior Free” Model

n buyers, $i \in [n]$ has value $v_i \sim \mathcal{D}_i$

Not knowing $\mathcal{D}_1, \ldots, \mathcal{D}_n$, seller picks auction

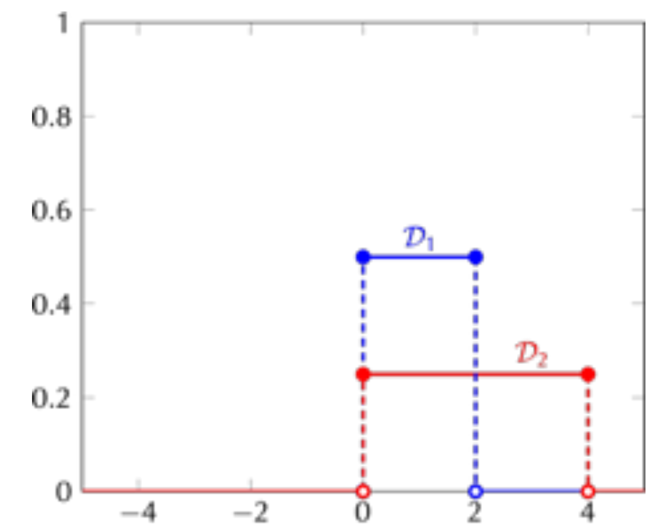
Auction’s (expected) revenue measured on

$$(v_1, \ldots, v_n) \sim \mathcal{D}_1 \times \ldots \times \mathcal{D}_n$$

No knowledge of prior distributions

[GHW ’01, HR’08, BBHM’08*]

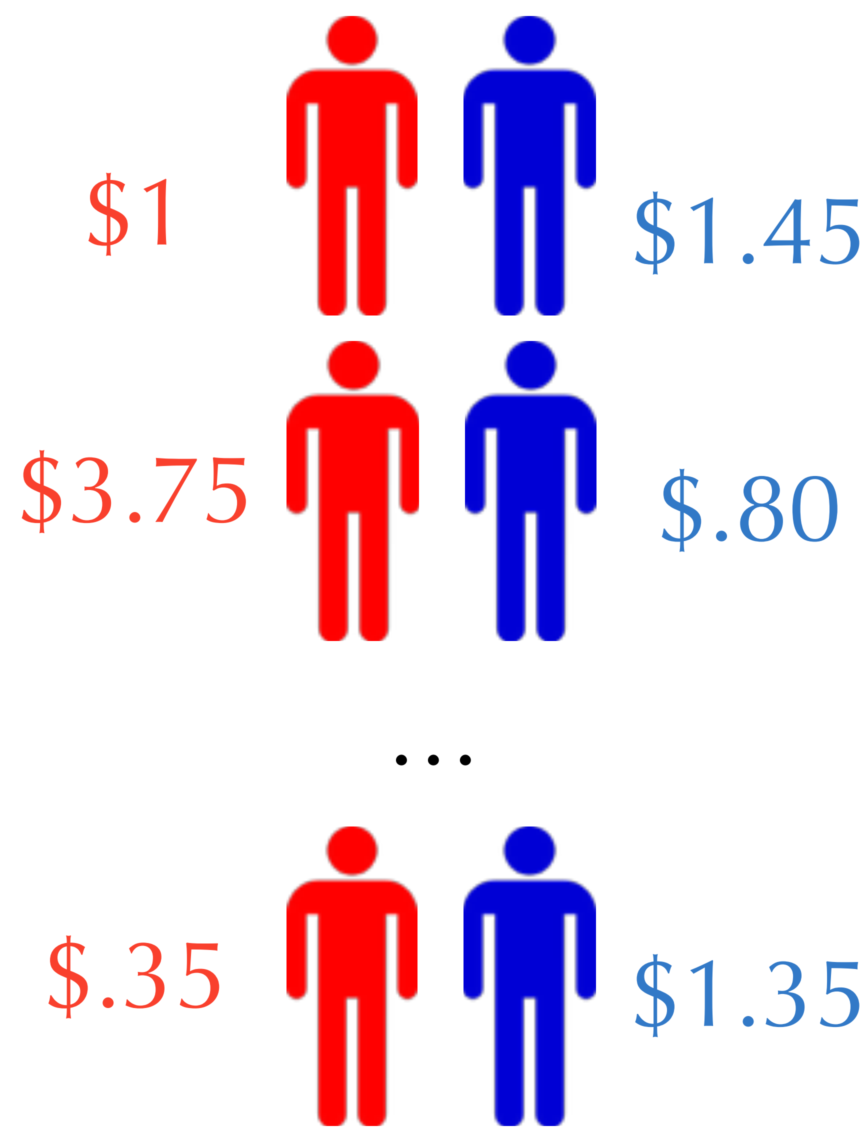
Sampling Model (this talk)



Auction(m samples ~ (D_1, … D_n))

Revenue(Auction(m samples), (D_1, … D_n))

Sample access to prior distributions



[DRY'10, CR'14,…]

Total Revenue

Revenue(Auction(samples), D)

= Revenue(Auction(samples))

Representation error

+ (Revenue(Auction(samples), D) - Revenue(Auction(samples)))

Generalization error

Worse revenue, better generalization

Unrealistic, detail-dependent

More Revenue

Less info

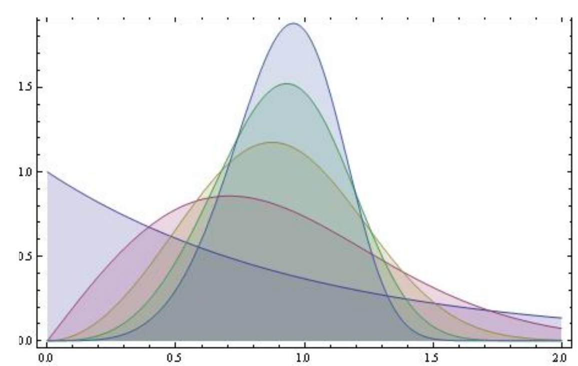
Prior Free

Samples from prior

Perfect prior

Perfect info

Tradeoff between optimality on sample and future performance



$v_1 \sim D_1$

$v_2 \sim D_2$

$v_3 \sim D_3$

General technique:

1. What does OPT do, assuming perfect prior?
2. Can OPT be learned from poly samples?
3. If not, what's a good apx for OPT which is "simpler"?

Outline

- **Sample complexity definitions**
- Single Parameter (Single Item)
  - IID
  - Non-IID
    - Regular
    - Irregular
  - Open Qs in this area
- Multi-parameter (* If time)

Learning Theory

(binary setting)

$$C = \{c : X \to \{\ ,\ \}\}$$

m samples $S = \{(x, y) : x \in X, y \in \{\ ,\ \}\}, (x, y) \sim D$

ERM Learner:

Given S, C, pick best $c \in C$ for S

When does ERM do well on D?

Depends on C and S together

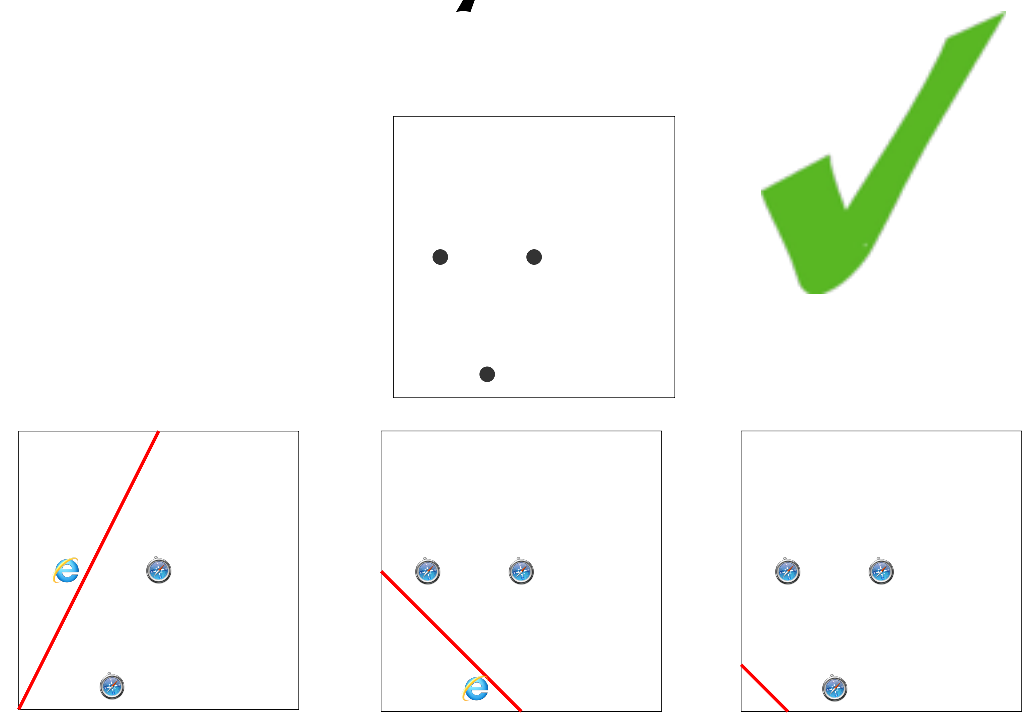
Learning Theory

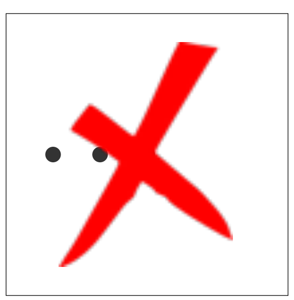
C can **shatter** S if

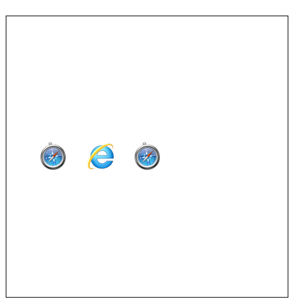
any $S' \subseteq S$ has a $c_{S'} \in C$ s.t.

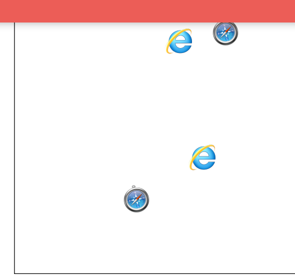
$c_{S'}(x) =$ for $x \in S'$

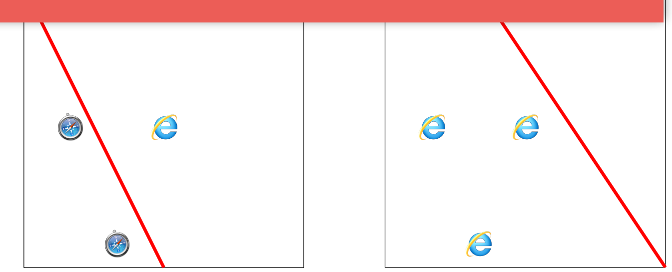
and $c_{S'}(x) =$ for $x \notin S'$

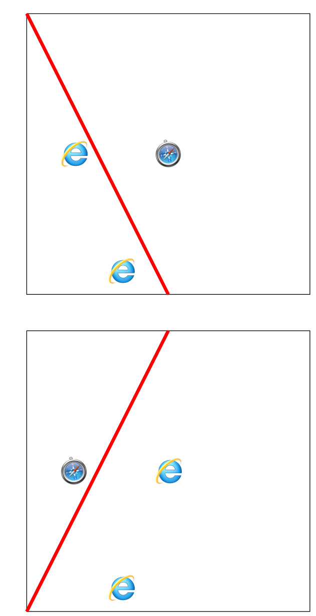












$$VC(C) = \max_{S \text{ that } C \text{ can shatter}} |S|$$

Learning Theory

ERM Learner:

Given S, C, pick best $c \in C$ for S

S is big enough

Thm: Need $|S| \geq \tilde{\Theta}\left(\frac{\nu c(C)}{\epsilon^2}\right)$ for ERM to pick $c \in C$ which has error $\leq \epsilon + \mathrm{OPT}(C)$

C must contain a good classifier

Learning Theory

(real-valued/auction setting)

$C = \{c : X \to [0, H]\}$

m samples $S = \{v^t \in [0, H]^n\}_t \sim D$

$c(v^t)$ is c's *revenue* on v^t

ERM Learner:

Given S, pick $c \in C$ which has highest revenue for S

When does ERM do well on D?

Shattering

C can **shatter** S if there is some $(r^1, \ldots, r^m) \in \mathbb{R}^m$, $m = |S|$, s.t

any $S' \subseteq S$ has a $c_{S'} \in C$ s.t.

$c_{S'}(x^t) \geq r^t$ for $x^t \in S'$

and $c_{S'}(x^t) < r^t$ for $x^t \notin S'$

$$PD(C) = \max_{S \text{ that } C \text{ can shatter}} |S|$$

Learning Theory

ERM Learner:

Given S, C, pick best $c \in C$ for S

Thm: Need $|S| \geq \tilde{\Theta}\left(\frac{H^2 VC(C)}{\epsilon^2}\right)$ for ERM to pick

Thm: If $|S| \geq \tilde{\Omega}\left(\frac{H^2 PD(C)}{\epsilon^2}\right)$, ERM will pick

$c \in C$ which has error $\leq \epsilon + \mathrm{OPT}(C)$

$c \in C$ which has revenue $\geq \mathrm{OPT}(C) - \epsilon$

Thm 2: If $|S| \geq \tilde{\Omega}\left(\frac{H^2 PD(C)}{\epsilon^2}\right)$,

every $c \in C$ has $c(S) \in \mathbb{E}_{v \sim D}[c(v)] \pm \epsilon$.

uniform convergence over C

Outline

- A few words/notations about sample complexity
- Single Parameter (Single Item)
  - IID
  - Non-IID
  - Computational
  - Open Qs in this area
- Multi-parameter
- Online questions

General technique:

1. What does OPT do?
2. Can OPT be learned from poly samples?
3. If not, what's a good apx for OPT which is "simpler"?

In Single Parameter Settings:

Always do Myerson
(or some approximation)

…. because Myerson is optimal

Single Parameter Setting

n buyers, with $v_i \sim \mathcal{D}_i$, $v_i \in \mathbb{R}$

Feasibility space $X \subseteq 2^n$

Auction: $(a, p) : \mathbb{R}^n \to X \times \mathbb{R}^n$

(an allocation and payment rule)

Independent

$\forall v_1, \ldots, v_n, \forall i, v_i',$

So, we will (mostly) only talk about allocation rules,
since they define truthful payments

$\Leftarrow \quad a$ monotone, p charges winner(s) min winning bid

The Myerson Auction: maximizes revenue in this setting

$$a_{\mathcal{M}}(v_1, \ldots, v_n) = \text{argmax}_{x \in X} \sum_{i \in X} \phi_i(v_i)$$

Depends delicately on prior distributions

$$v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}$$

"virtual value"

MHR: when $\frac{1-F_i(v_i)}{f_i(v_i)}$ is non-increasing

Regularity: when ϕ_i is non-decreasing

Extra tricks (ironing) needed for irregular

Assume $v \in [1, H]$ for irregular settings

The Myerson Auction: maximizes revenue in this setting

$$a_{\mathcal{M}}(v_1, \ldots, v_n) = \text{argmax}_{x \in X} \sum_{i \in X} \phi_i(v_i)$$

More generally, for any auction $\mathcal{A}$

$$\mathbb{E}_{v \sim \mathcal{D}}\left[\text{Rev}(\mathcal{A}(v))\right] = \mathbb{E}_{v \sim \mathcal{D}}\left[a_{\mathcal{A},i}(v)\phi_i(v_i)\right]$$

$$v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}$$

"virtual value"

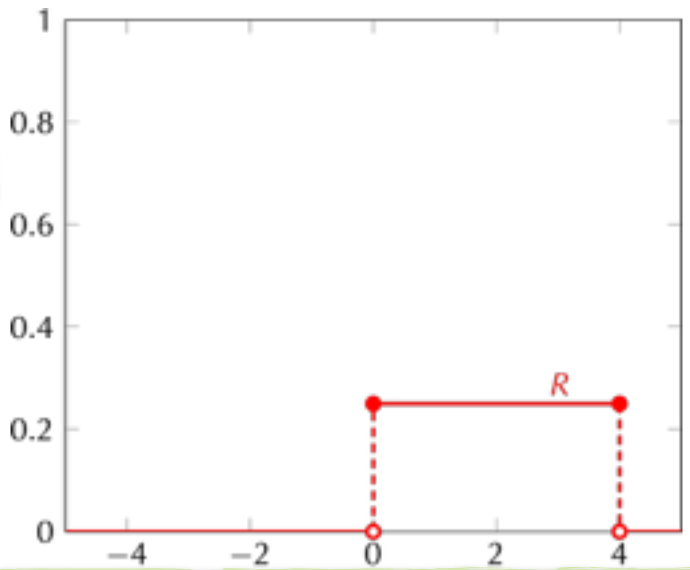
Outline

- A few words/notations about sample complexity
- Single Parameter (Single Item)
  - **Regular, IID**
    - **Why the "obvious" approach might overfit**
  - Regular, Non-IID
  - Irregular, Non-IID
  - Open Qs in this area
- Multi-parameter
- Online questions

n regular IID Buyers

The set of *possible* allocations pretty large…

… but Myerson is pretty simple.



Opt for n regular IID Buyers

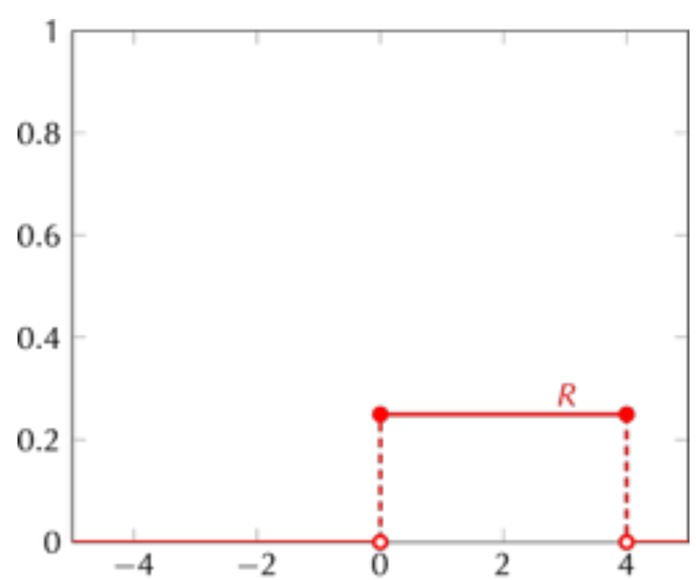
$$a(v_1,\ldots,v_n) = \mathrm{argmax}_{x\in X}\sum_{i\in X}\phi_i(v_i)$$

$$= \mathrm{argmax}_{x\in X}\sum_i \phi(v_i)$$

$$= \begin{cases} \emptyset & \boxed{\max_i \phi(v_i) \le 0} \\ \mathrm{argmax}_i v_i & \text{o/w} \end{cases}$$

Only part to "learn", a single parameter: smallest v* for which ϕ(v*) > 0

v* acts as a reserve price



n regular IID Buyers from samples

$$a(v_1, \ldots, v_n) = \text{argmax}_{x \in X} \sum_{i \in X} \phi_i(v_i)$$

$$= \text{argmax}_{x \in X} \sum_i \phi(v_i)$$

$$= \begin{cases} \emptyset & \boxed{\max_i \phi(v_i) \leq 0} \\ \text{argmax}_i v_i & \text{o/w} \end{cases}$$

Only part to "learn", a single parameter: smallest v* for which $\phi(v^*) > 0$

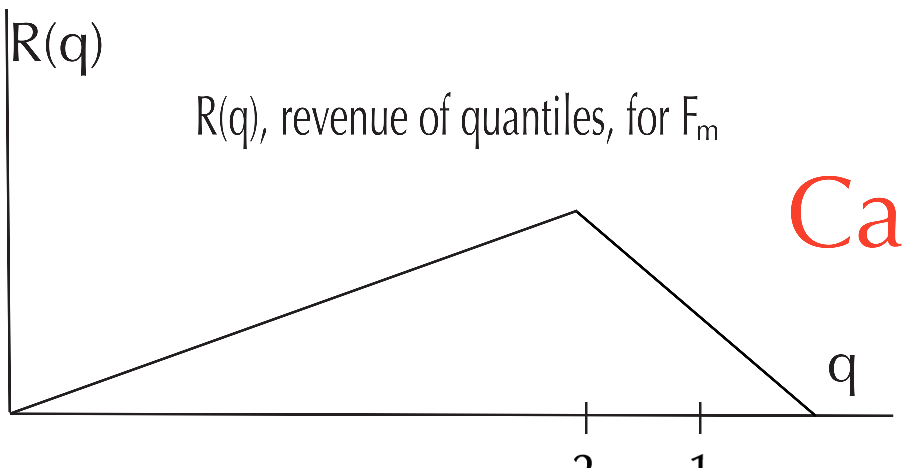
v* acts as a reserve price

How to do this with samples?

A: pick v* w. highest revenue on samples!

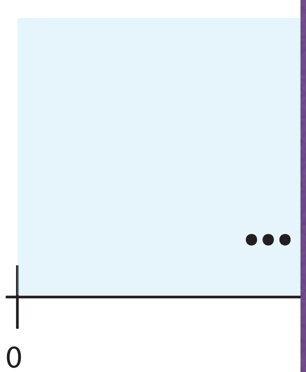
Why this might not work

“straight up” empirical Myerson might overfit.



Can see too many large samples of m…

w.p. $\Omega(1)$:



So, even for a single regular distribution, empirical Myerson “overfits”!

(No distribution-independent 1-ε-approximation)

[* Due to DRY’10]

Fixing the overfitting

$$a(v_1, \ldots, v_n) = \text{argmax}_{x \in X} \textstyle\sum_{i \in X} \phi_i(v_i)$$

$$= \text{argmax}_{x \in X} \textstyle\sum_i \phi(v_i)$$

Only part to "learn", a single parameter: smallest v* for which ϕ(v*) > 0

$$= \begin{cases} \emptyset & \boxed{\max_i \phi(v_i) \leq 0} \\ \text{argmax}_i v_i & \text{o/w} \end{cases}$$

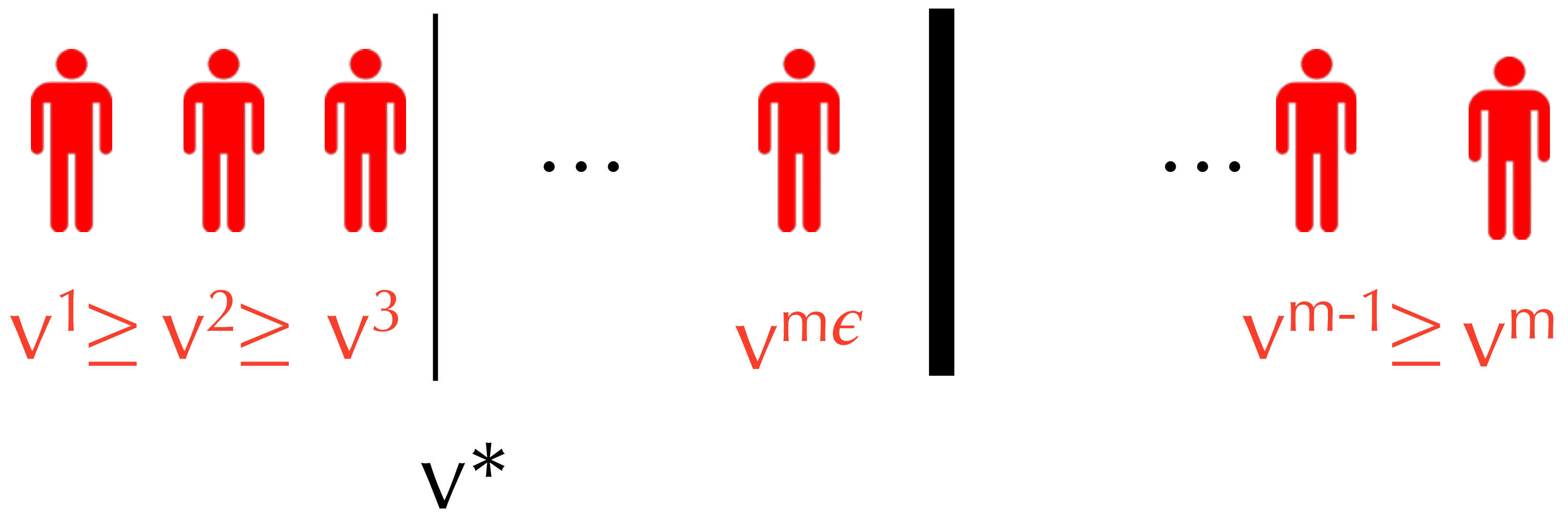
v* acts as a reserve price

How to do this with samples?

A′: pick v* w. highest revenue on samples

s.t v* < $m\epsilon$-th highest sampled value

Theorem:
For IID regular bidders



Theorem:

If $m = \tilde{\Omega}\left(\text{poly}\left(\frac{1}{\epsilon}\right)\right)$ and the distribution is regular then this $1 - \epsilon$ approximates OPT.

[DRY'10, HMR'15]

Main take-away: IID Regular

"True" Lost Revenue of
auction class C
on distribution D

=

Representation error:
At most (1-ϵ) for
guarded reserve

+

Generalization
error:
(1-ϵ) for guarded
reserve

Outline

- A few words/notations about sample complexity
- Single Parameter (Single Item)
  - IID, Regular
  - **Non-IID, Regular**
  - Non-IID, Irregular
  - Open Qs in this area
- Multi-parameter (Time permitting)

Myerson for n non-iid regular buyers

$$a(v_1, \ldots, v_n) = \text{argmax}_{x \in X} \sum_{i \in X} \phi_i(v_i)$$

Need not be well-behaved

So, in general, no simple form

$$\phi_i(v_i) = v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}$$

Myerson for n non-iid buyers from samples?

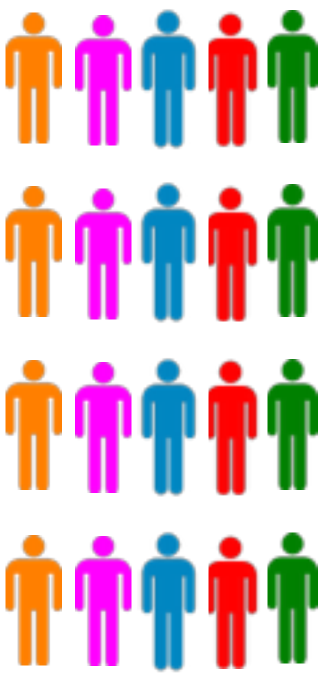
$$a(v_1, \ldots, v_n) = \text{argmax}_{x \in X} \sum_{i \in X} \phi_i(v_i)$$

$$= v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}$$

How to do this with samples?

Something that "looks like" Myerson on samples,

truncating $m\epsilon$ highest sampled values for each dist.

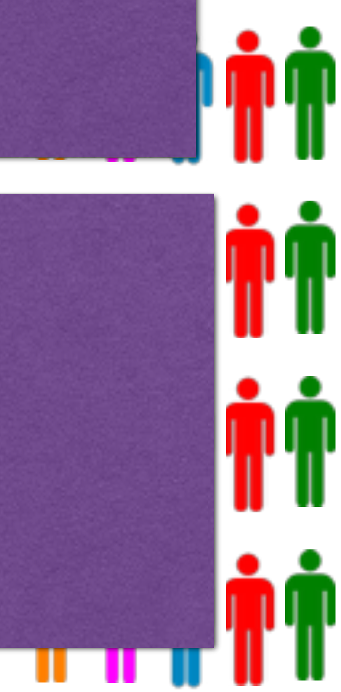


[CR'14]

Theorem:
For non-IID regular bidders

Something that “looks like” Myerson on samples, truncating $m\epsilon$ highest sampled values for each dist.

Theorem: [CR’14]
If $m = \tilde{\Omega}\left(\text{poly}(n, \frac{1}{\epsilon})\right)$ and the distribution is regular then this $1 - \epsilon$ approximates OPT.



Stylized analysis, less “portable” to different settings.

Main take-away: non-IID regular setting

"True" Lost Revenue of auction class C on distribution D designed from data

=

Representation error: (1-ϵ) for truncating empirical distributions

+

Generalization error: at most (1-ϵ) for running empirical Myerson

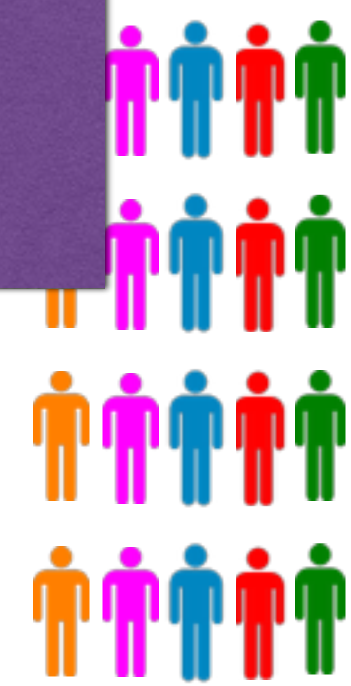
Outline

- A few words/notations about sample complexity
- Single Parameter (Single Item)
  - IID, Regular
  - Non-IID, Regular
  - **Non-IID, Irregular**
  - Open Qs in this area
- Multi-parameter (Time permitting)

For n non-iid buyers

CR'14 → poly(n, $^{1}/_{\varepsilon}$) sample complexity
(upper and lower bounds)
for single item, (MHR or regular) distributions

Uses "standard" ML techniques:
applies to more general settings, tighter bounds,
NOT computationally efficient



DHP'16 → improved regular upper bound,
general downwards-closed single-parameter

General technique:

1. What does OPT do?
2. Can OPT be learned from poly samples?
3. If not, what's a good apx for OPT which is "simpler"?

Myerson for n non-iid irrregular buyers

$$a(v_1, \ldots, v_n) = \text{argmax}_{x \in X} \sum_{i \in X} \bar{\phi}_i(v_i)$$

"Ironed" virtual value

Need not be well-behaved

So, in general, no simple form

Proof Technique

Design set of auctions C :

C has a $(1-\varepsilon)$-opt auction for all distributions

Show C is "simple" :
For all $f \in C$

polynomial sample's empirical revenue (f) $\approx$ true revenue (f)

Based on poly sample S, $f \in C$ w. best revenue on S

will be $(1-\varepsilon)$-opt for true distribution

[MR'15]

Attempt 1

C =
Set of Myerson auctions for all distributions
(perhaps truncated?)

Use pseudo-dimension to show polynomial uniform convergence of C

Based on poly sample S, f∈C w. best revenue on S

will be (1- Ɛ)-opt for true distribution

[MR'15]

Where this goes wrong

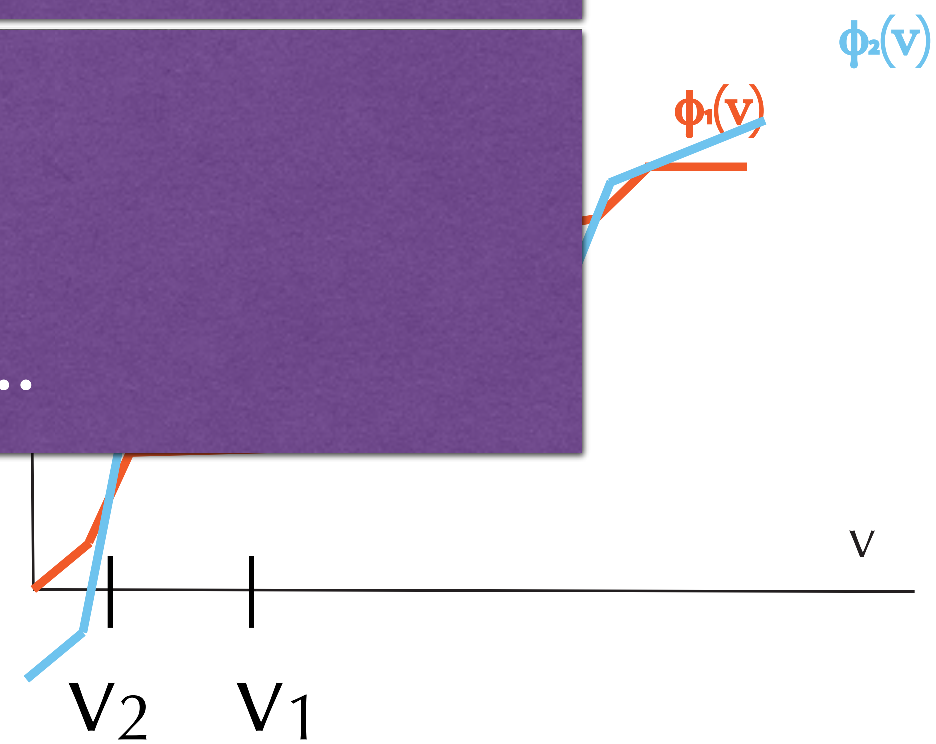
Myerson's Class is "complicated"

(in a formal sense)

Myerson's Class is Complicated

The set of all allocation rules is highly unconstrained….

Myerson's Class
1. Has infinite pseudo-dimension
2. Doesn't have finite-sample uniform convergence guarantees…



Attempt 2

C =
Set of approximate Myerson auctions for all distributions

Use pseudo-dimension to show polynomial uniform convergence of C

Based on poly sample S, f∈C w. best revenue on S will be (1- ϵ)-opt for true distribution

[MR'15]

Apx optimal auctions for n irregular iid bidders

Want:

A set of auctions with

1. more constrained allocation rules than Myerson
2. Still contains an auction for each distribution:
   1. agrees with Myerson's allocation mostly?
   2. Or only disagrees with Myerson when doing so loses very little revenue

The Myerson Auction: maximizes revenue in this setting

$$a_{\mathcal{M}}(v_1, \ldots, v_n) = \operatorname{argmax}_{x \in X} \sum_{i \in X} \bar{\phi}_i(v_i)$$

More generally, for any auction $\mathcal{A}$

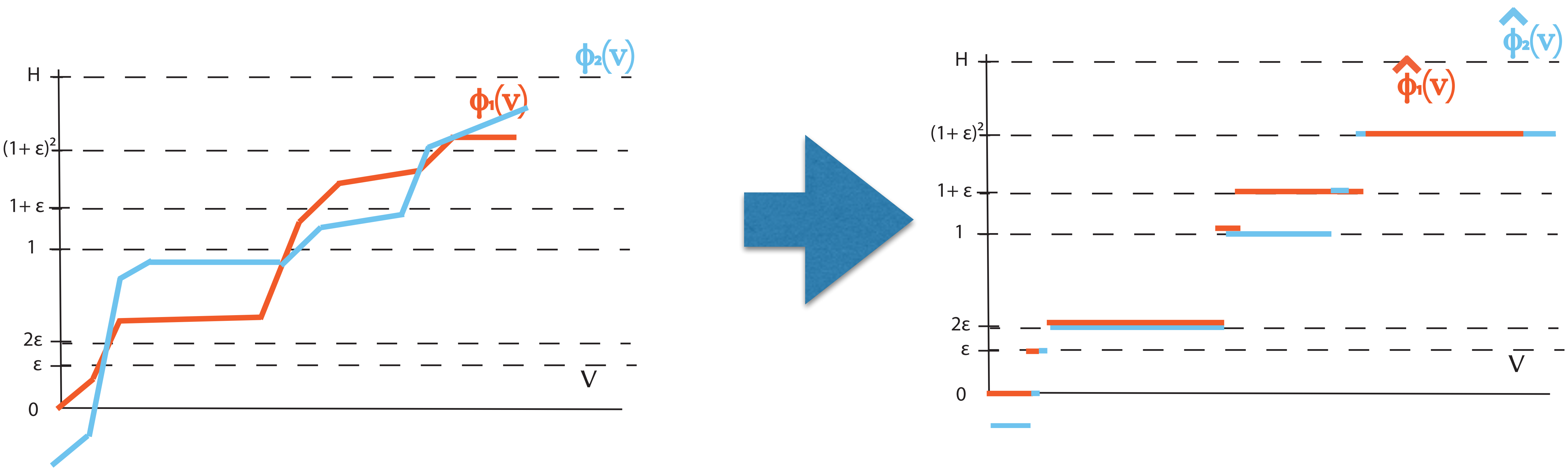
$$\mathbb{E}_{v \sim \mathcal{D}}\left[\mathrm{Rev}(\mathcal{A}(v))\right]$$

$$=$$

$$\mathbb{E}_{v \sim \mathcal{D}}\left[a_{\mathcal{A},i}(v)\bar{\phi}_i(v_i)\right]$$

For n non-iid buyers

So… what about estimating virtual value curves to some reasonable precision?



Class of Apx OPT Auctions C_B

One Auction A

For each $i \in [n]$, choose $\hat{\phi}_i^\epsilon$:

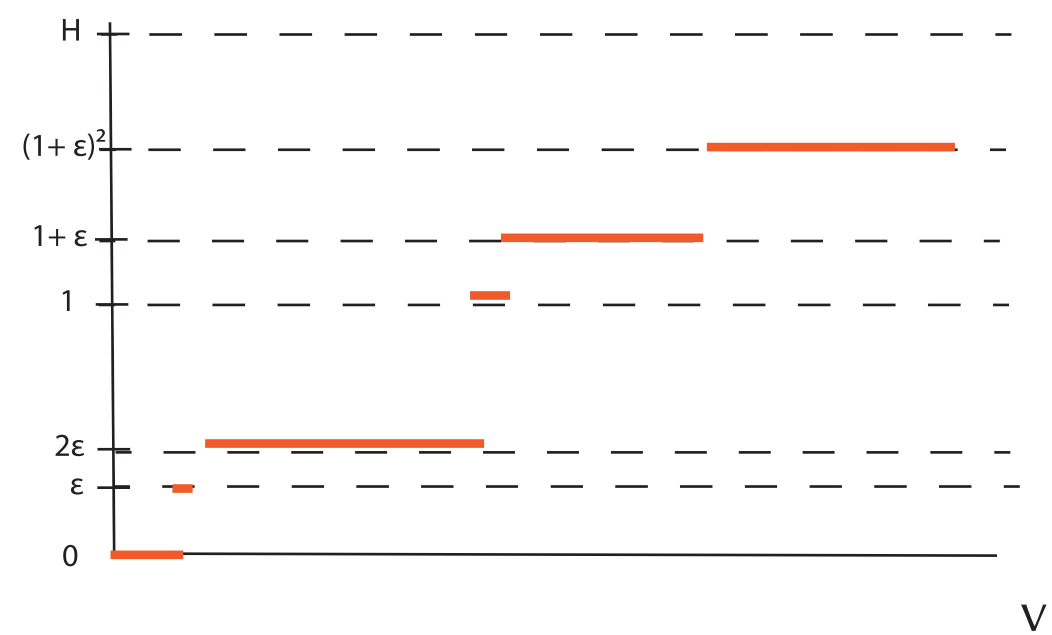
monotone

piecewise constant fn

w. values $\in \{-1, 0, \epsilon, 2\epsilon, \ldots, 1, (1+\epsilon), (1+\epsilon)^2, \ldots H\}$

Run Myerson w.r.t $\hat{\phi}_1^\epsilon, \ldots, \hat{\phi}_n^\epsilon$

"Parameters": locations of jumps



Why does this class contain an (1-ϵ)-optimal auction?

Fix distribution, find auction in class which is (1-ϵ)-optimal .

Fix D_1, … D_n. Find auction in C_B which is apx optimal

One Auction A

For each buyer i

$$\text{Let } \hat{\phi}_i^{\epsilon}(v_i) = \begin{cases} -1 & \bar{\phi}_i(v_i) < 0 \\ 0 & 0 \leq \bar{\phi}_i(v_i) < \epsilon \\ t\epsilon & t\epsilon \leq \bar{\phi}_i(v_i) < (t+1)\epsilon \leq 1 \\ (1+\epsilon)^t & 1 \leq (1+\epsilon)^t \leq \bar{\phi}_i(v_i) < (1+\epsilon)^{t+1} \end{cases}$$

What is the expected revenue of this auction?

Hint: Compare to true Myerson

Never get negative v.v.

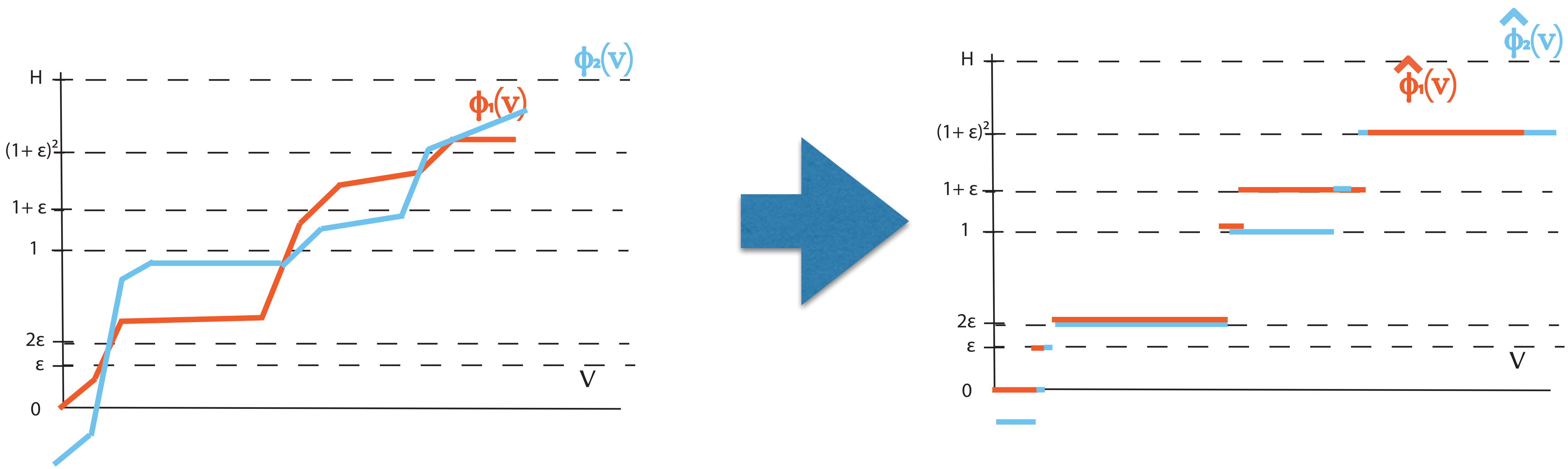
ϵ additive apx

(1-ϵ) mult apx

So, will ultimately be a(1-ϵ) mult apx to ironed virtual values!

💪

For n non-iid buyers



Attempt 2

C =
Set of approximate Myerson auctions for all distributions

Use pseudo-dimension to show polynomial uniform convergence of C

Based on poly sample S, f∈C w. best revenue on S

will be (1- Ɛ)-opt for true distribution

[MR'15]

Why is this class learnable?

Let $B = \frac{1}{\epsilon} + \log_{1+\epsilon}(H)$

$C_B = \{\text{auctions below}\}$

One Auction A

For each $i \in [n]$, choose $\hat{\phi}_i^\epsilon$:

monotone

piecewise constant fn

w. values $\in \{-1, 0, \epsilon, 2\epsilon, \ldots, 1, (1+\epsilon), (1+\epsilon)^2, \ldots H\}$

Run Myerson w.r.t $\hat{\phi}_1^\epsilon, \ldots, \hat{\phi}_n^\epsilon$

Bound the pseudo-dimension of C_B…

Pseudo-Dimension

C can **shatter** S if there is some $(r^1, \ldots, r^m) \in \mathbb{R}^m, m = |S|$, s.t

any $S' \subseteq S$ has a $c_{S'} \in C$ s.t.

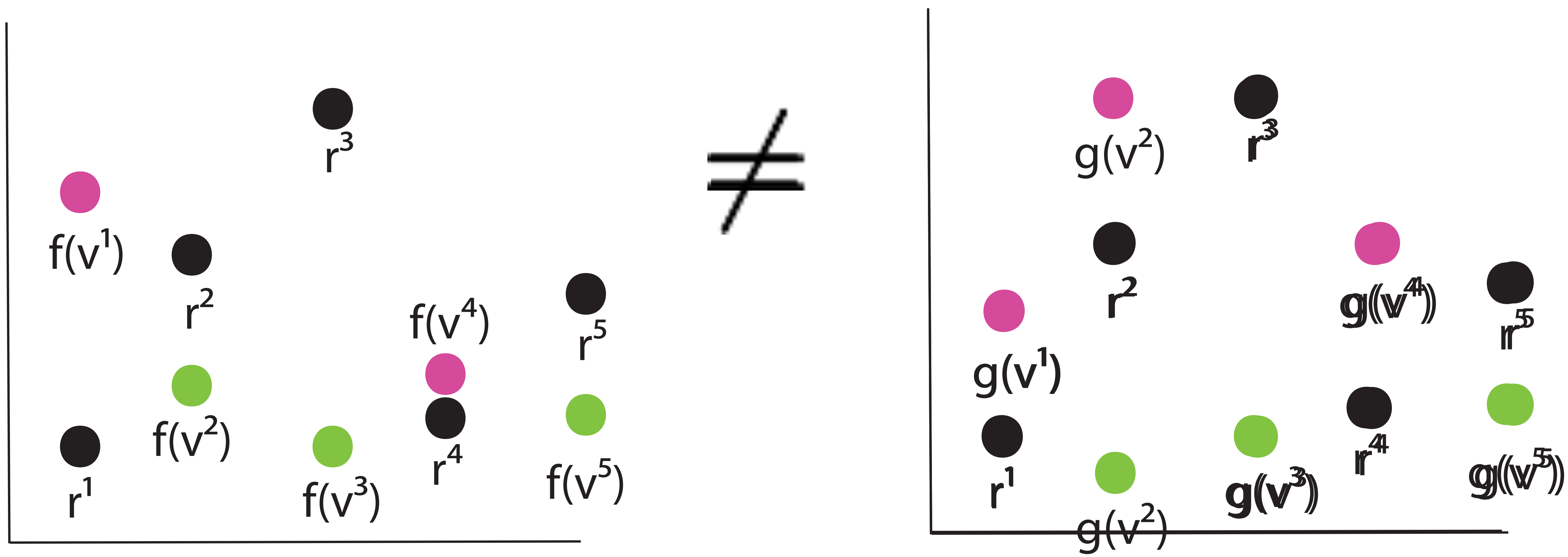
$c_{S'}(v^t) \geq r^t$ for $v^t \in S'$

$c_{S'}(v^t) < r^t$ for $v^t \notin S'$

$$PD(C) = \max_{S \text{ that } C \text{ can shatter}} |S|$$

Upper bound on PD(C_B)

Upper bound M, # of distinct "sign patterns" for S:



Since $M \geq 2^{|S|}$ if S is shatterable, $\rightarrow \log M \geq |S| = m$

Upper bounding
of sign patterns

Fix sample $S = \{v^1, \ldots, v^m\}, v^t = (v^t_1, \ldots, v^t_n)$

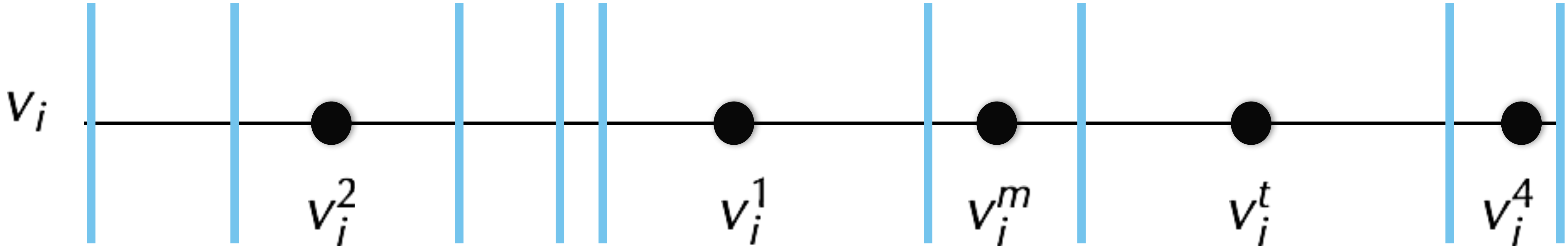
And some $r = (r^1, \ldots, r^m)$

Separate C_B into $C_{B,1}, \ldots, C_{B,k}$

$$k = O\left(\binom{B+m}{B}^n\right)$$

s.t. ordering of break points and values is fixed $\in C_{B,j}$

for all $i \in [n]$

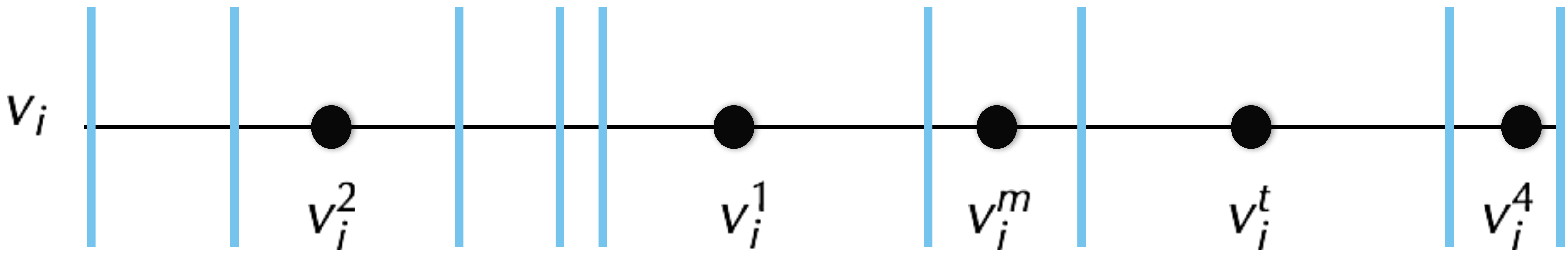


Bounding the pseudo-dimension

Then, in a fixed $C_{B,j}$,

Revenue a Bn-dimensional linear function of break points

(which can induce $\Theta(m^{Bn})$ sign patterns)



Bounding the pseudo-dimension

Separate C_B into $C_{B,1}, \ldots, C_{B,k}$ $\qquad k = O\left(\binom{B+m}{B}^n\right)$

Revenue a Bn-dimensional linear function of break points

(which can induce $\Theta(m^{Bn})$ sign patterns)

Then in total $M \leq \binom{B+m}{B}^n \cdot m^{Bn}$ sign patterns

If $M \geq 2^m$, $m = \tilde{O}(Bn)$

So pseudo-dimension is $\tilde{O}(Bn)$

Theorem:
For non-IID irregular bidders

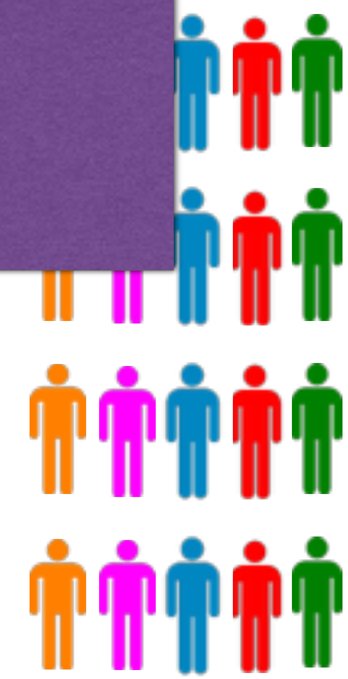
Theorem: [MR'15]

If $m = \tilde{\Omega}\left(\frac{H^2}{\epsilon^3} n\right)$

then ERM over Myerson's class

w. discretized virtual values

is $1-\epsilon$-approximately optimal.



Main take-away: non-IID regular setting

"True" Lost Revenue of auction class C on distribution D designed from data

=

Representation error: (1-ϵ) for using only estimated virtual values

+

Generalization error: at most (1-ϵ) from uniform convergence

Outline

- Learning Theory Basics
- Single Parameter (Single Item)
  - Regular, IID
  - Regular, Non-IID
  - Irregular, Non-IID
  - **Related Work/Open Qs in this area**
- Multi-parameter

Related Work

General sampling for mechanism design
- BBHM ’05: **Reducing mechanism design to algorithm design via machine learning.**

Finite Support
- Elkind ’07: **Designing and learning optimal finite support auctions**

IID, MHR and Regular
- DRY’10: **Revenue Maximization with a Single Sample**
- HMR’15: **Making the Most of Your Samples**

IID, irregular
- SR’16 [This EC!]: **Ironing in the Dark**

Non-IID
- CR’14 (MHR + Regular): **The Sample Complexity of Revenue Maximization**
- MR’15 (also MHR): **The Pseudo-Dimension of Near-Optimal Auctions**
- DHR’16 (Regular): **The Sample Complexity of Auctions with Side Information**

Technical open questions

Do irregular iid settings need poly(n) samples for computationally efficient algorithms?

What is a computationally efficient algorithm for non-iid irregular settings w. polynomial sample complexity?

Are there separations in sample complexity information theoretically vs. computationally?

Close various gaps…
- Does regular single parameter s.c. depend on n?

Open-ended open questions

In what contexts is it better to “mix” two distributions and draw twice from an irregular distribution for sample complexity?

What other properties of distributions might decrease the sample complexity of learning nearly optimal auctions?

Related Work: Single Parameter

General sampling for mechanism design

- BBHM '05: **Reducing mechanism design to algorithm design via machine learning.**

Finite Support

- Elkind '07: **Designing and learning optimal finite support auctions**

IID, MHR and Regular

- DRY'10: **Revenue Maximization with a Single Sample**
- HMR'15: **Making the Most of Your Samples**

IID, irregular

- SR'16 [This EC!]: **Ironing in the Dark**

Non-IID

- CR'14 (MHR + Regular): **The Sample Complexity of Revenue Maximization**
- MR'15 (also MHR): **The Pseudo-Dimension of Near-Optimal Auctions**
- DHR'16 (Regular): **The Sample Complexity of Auctions with Side Information**

Grand Slide of Single Buyer/IID results. Red are non-computational.

Feasibility	Single Buyer MHR	Single Buyer α-strongly regular	Single Buyer regular	Single Buyer bounded	IID MHR	IID regular	IID bounded
Single Item	$\tilde{\Theta}\left(\frac{1}{\epsilon^{3/2}}\right)$ [HMR'15]	$\tilde{\Omega}\left(p(\frac{1}{\epsilon}, n)^{p(\alpha)}\right)$ [CR'15]	$\tilde{\Theta}\left(\frac{1}{\epsilon^3}\right)$ [HMR'15] [DRY'10]	$\tilde{\Theta}\left(\frac{H}{\epsilon^2}\right)$ [HMR'15]	$\tilde{O}\left(\frac{1}{\epsilon^2}\right)$ [DRY'10] $\tilde{\Omega}\left(\frac{1}{\epsilon^{3/2}}\right)$ [HMR'15]	$\tilde{\Theta}\left(\frac{1}{\epsilon^3}\right)$ [HMR'15] [DRY'10]	$\tilde{O}\left(\frac{n^2H^2}{\epsilon^2}\right)$ [SR'16] $\tilde{\Omega}\left(\frac{1}{\epsilon^2}\right)$ [CGM'13] $\tilde{\Omega}(H)$ [HMR'15]
Matroid					$\tilde{O}\left(\frac{1}{\epsilon^3}\right)$ [DRY'10] $\tilde{\Omega}\left(\frac{1}{\epsilon^{3/2}}\right)$ [HMR'15]	$\tilde{\Omega}\left(\frac{1}{\epsilon^3}\right)$ [HMR'15] $\tilde{O}\left(\frac{n}{\epsilon^4}\right)$ [DMP'16]	$\tilde{O}\left(\frac{n^2H^2}{\epsilon^2}\right)$ [SR'16] $\tilde{\Omega}\left(\frac{1}{\epsilon^2}\right)$ [CGM'13] $\tilde{\Omega}(H)$ [HMR'15]
Downward closed					$\tilde{O}\left(\frac{n^2}{\epsilon^4}\right)$ [DHP'16] $\tilde{\Omega}\left(\frac{1}{\epsilon^{3/2}}\right)$ [HMR'15]	$\tilde{\Omega}\left(\frac{1}{\epsilon^3}\right)$ [HMR'15] $\tilde{O}\left(\frac{n^2}{\epsilon^4}\right)$ [DMP'16]	*$\tilde{O}\left(\frac{n^3H^3}{\epsilon^3}\right)$ [MR'15]* $\tilde{\Omega}\left(\frac{1}{\epsilon^2}\right)$ [CGM'13] $\tilde{\Omega}(H)$ [HMR'15]

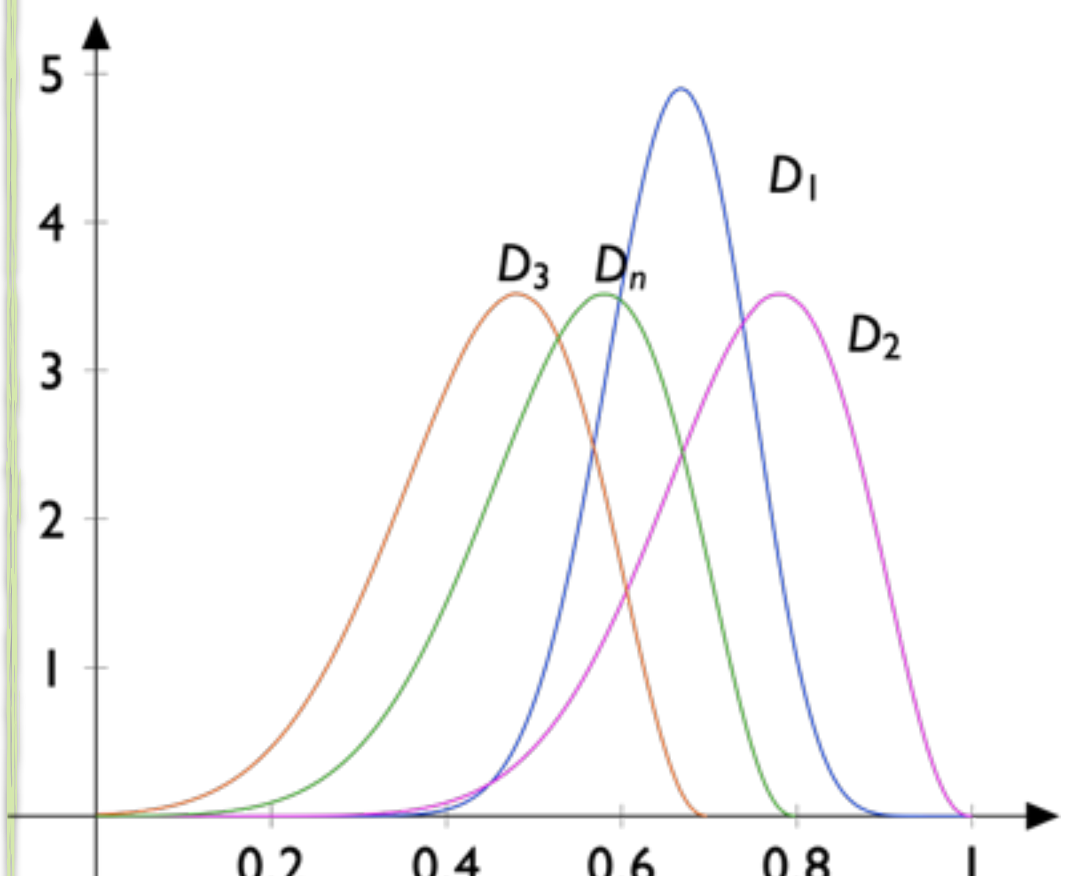
Grand Slide of All results. Red are non-computational.

Feasibility	Non-IID, MHR	Non-IID, α-strongly Regular	Non-IID, Regular	Non-IID, Bounded
Single Item	$\tilde{O}\left(\frac{n}{\epsilon^3}\right)$ [MR'15] $\tilde{O}\left(\frac{n^{10}}{\epsilon^7}\right)$ [CR'14] $\tilde{\Omega}\left(\frac{n}{\epsilon^{1/2}}\right)$ [CR'14]	$\tilde{\Omega}\left(\mathrm{p}(\frac{1}{\epsilon}, n)^{\mathrm{p}(\alpha)}\right)$ [CR'15] $\tilde{O}\left(\frac{n}{\epsilon^4}\right)$ [DHP'16]	$\Omega\left(\max\left(\frac{n}{\epsilon}, \frac{1}{\epsilon^3}\right)\right)$ [CR'14] [HMR'15] $\tilde{O}\left(\frac{n}{\epsilon^4}\right)$ [DHP'16]	$\tilde{O}\left(\frac{H^2 n}{\epsilon^3}\right)$ [MR'15] $\tilde{\Omega}\left(\frac{1}{\epsilon^2}\right)$ [CGM'13] $\tilde{\Omega}(H)$ [HMR'15]
Matroid	$\tilde{O}\left(\frac{n}{\epsilon^3}\right)$ [MR'15] $\tilde{\Omega}\left(\frac{n}{\epsilon^{1/2}}\right)$ [CR'14]	$\tilde{\Omega}\left(\mathrm{p}(\frac{1}{\epsilon}, n)^{\mathrm{p}(\alpha)}\right)$ [CR'15] $\tilde{O}\left(\frac{n}{\epsilon^4}\right)$ [DHP'16]	$\Omega\left(\max\left(\frac{n}{\epsilon}, \frac{1}{\epsilon^3}\right)\right)$ [CR'14] [HMR'15] $\tilde{O}\left(\frac{n}{\epsilon^4}\right)$ [DHP'16]	$\tilde{O}\left(\frac{H^2 n}{\epsilon^3}\right)$ [MR'15] $\tilde{\Omega}\left(\frac{1}{\epsilon^2}\right)$ [CGM'13] $\tilde{\Omega}(H)$ [HMR'15]
Downward closed	$\tilde{\Omega}\left(\frac{n}{\epsilon^{1/2}}\right)$ [CR'14]	$\tilde{\Omega}\left(\mathrm{p}(\frac{1}{\epsilon}, n)^{\mathrm{p}(\alpha)}\right)$ [CR'15] $\tilde{O}\left(\frac{n^2}{\epsilon^4}\right)$ [DHP'16]	$\Omega\left(\max\left(\frac{n}{\epsilon}, \frac{1}{\epsilon^3}\right)\right)$ [CR'14] [HMR'15] $\tilde{O}\left(\frac{n^2}{\epsilon^4}\right)$ [DHP'16]	$\tilde{\Omega}\left(\frac{1}{\epsilon^2}\right)$ [CGM'13] $\tilde{\Omega}(H)$ [HMR'15]
General	$\tilde{\Omega}\left(\frac{n}{\epsilon^{1/2}}\right)$ [CR'14]	$\tilde{\Omega}\left(\mathrm{p}(\frac{1}{\epsilon}, n)^{\mathrm{p}(\alpha)}\right)$ [CR'15]	$\Omega\left(\max\left(\frac{n}{\epsilon}, \frac{1}{\epsilon^3}\right)\right)$ [CR'14] [HMR'15]	$\tilde{O}\left(\frac{H^3 n^5}{\epsilon^3}\right)$ (additive) [MR'15] $\tilde{\Omega}\left(\frac{1}{\epsilon^2}\right)$ [CGM'13] $\tilde{\Omega}(H)$ [HMR'15]

Outline

- Learning Theory Basics
- Single Parameter (Single Item)
  - Regular, IID
  - Regular, Non-IID
  - Irregular, Non-IID
  - **Related Work/Open Qs in this area**
- Multi-parameter

Setting:
Selling to Combinatorial bidders



$v_1 \sim D_1$

$v_2 \sim D_2$

$v_3 \sim D_3$

$\ldots$

$v_n \sim D_n$

k

$v_i : 2^k \to [0, H]$

Multiparameter… less well understood on both sides.

A few pointers to papers on sample complexity

- Balcan, Blum, Hartline, Mansour '05
- Balcan, Devanur, Hartline, Talwar '07
- Agrawal, Wang, Ye '14
- Devanur and Hayes '09
- Dughmi, Han, Nissan '14
- Morgenstern and Roughgarden '16
- Goldner, Karlin '16

Thanks!