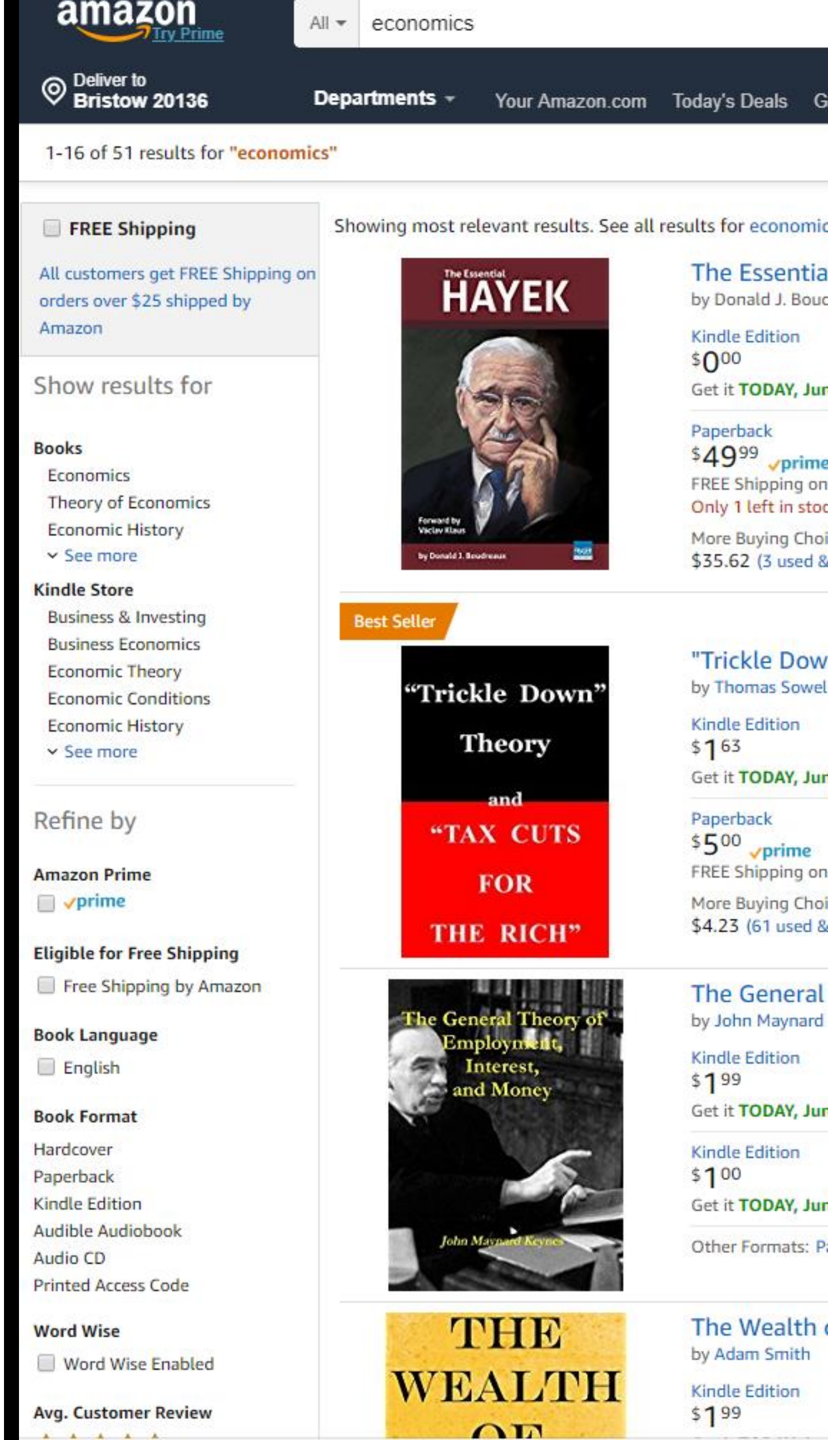
A General Theory of Sample Complexity for Multi-Item Revenue Maximization

Ellen Vitercik
Computer Science Department, CMU

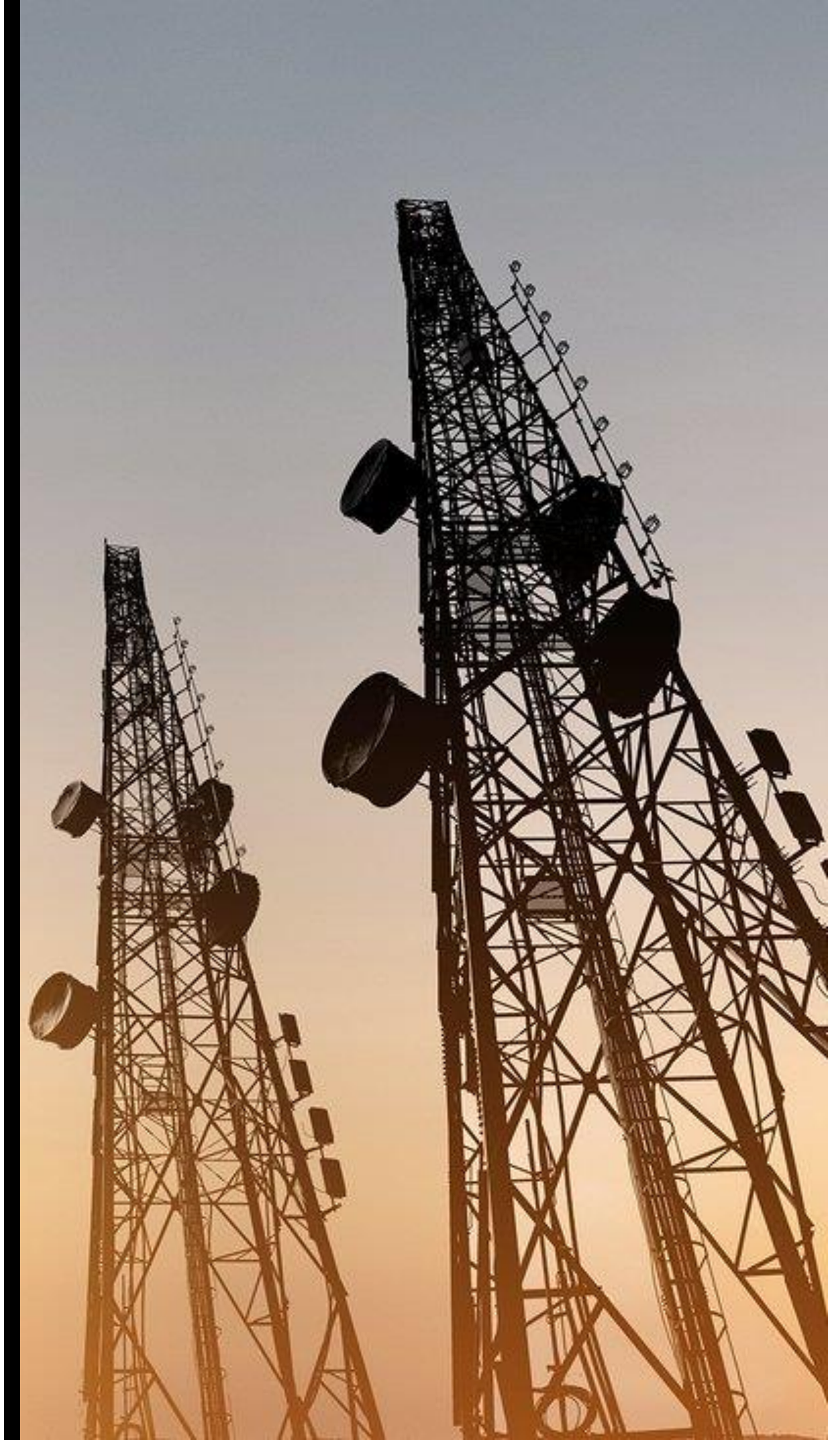
Joint work with Nina Balcan and Tuomas Sandholm

China Theory Week 2018

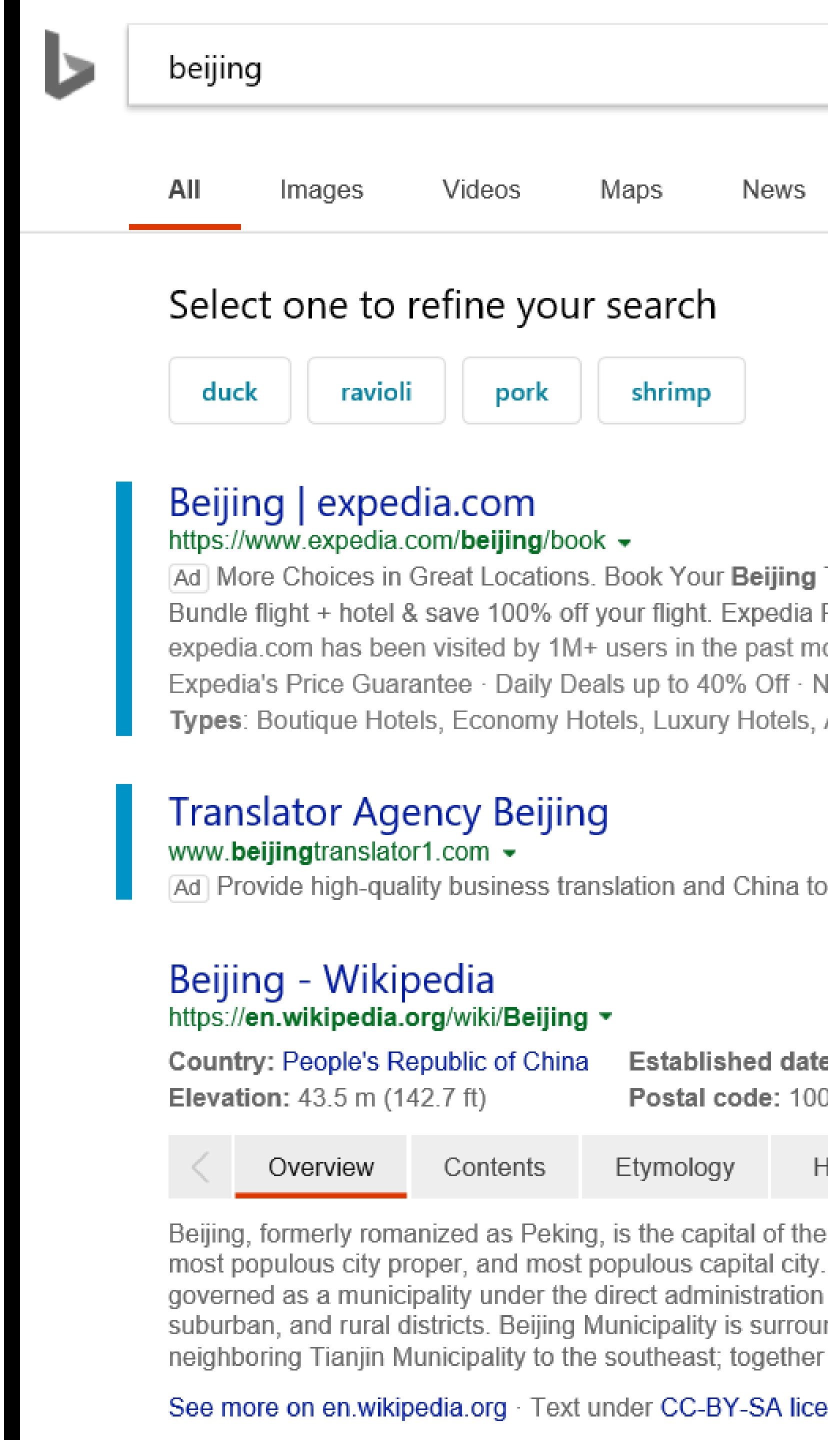
Amazon's profit swells to $1.6 billion
[NY Times '18]



Bidding in government auction of airwaves reaches $34 billion
[NYTimes ‘14]



	Ad revenue in 2016	Total revenue in 2016
Google	$79 billion	$89.46 billion
Facebook	$27 billion	$27.64 billion



Common misconception: There's only one way to hold an auction.

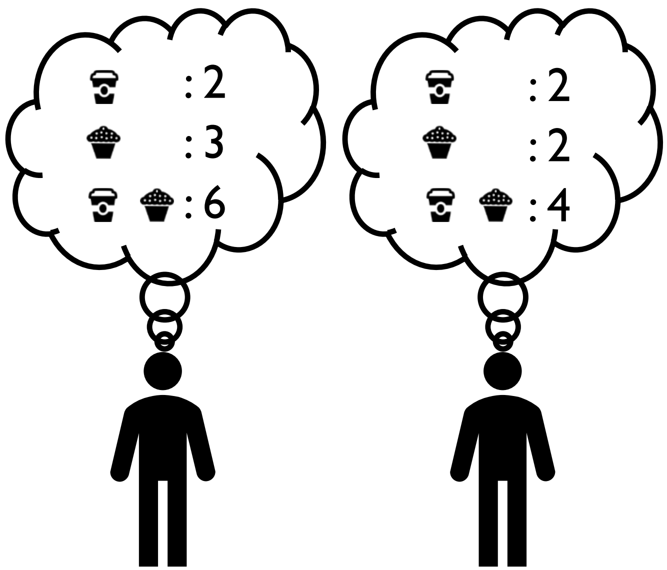
There are **infinitely**-many ways to hold an auction.

Mechanism design for sales settings

There is a set of **items** for sale and a set of **buyers**.

At a high level, a mechanism dictates:

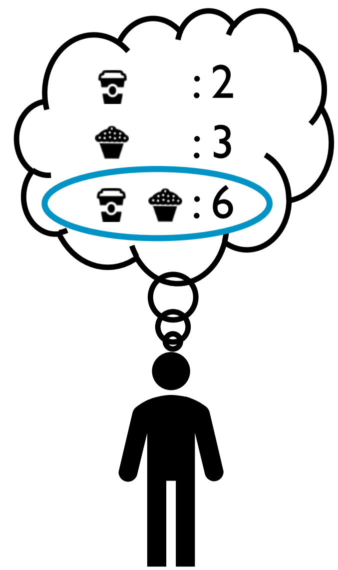
1. Which buyers receive which items.
2. What they pay.



Mechanism design example:
Posted price mechanisms

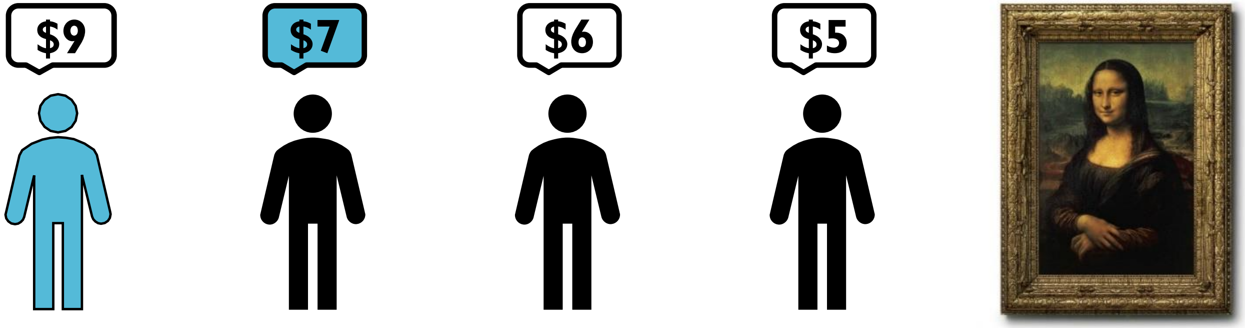
Set price per item.

Buyers buy the items maximizing their **utility** (value for items minus price).



Mechanism design example:
Second-price auction

The highest bidder wins and pays the second highest bid.



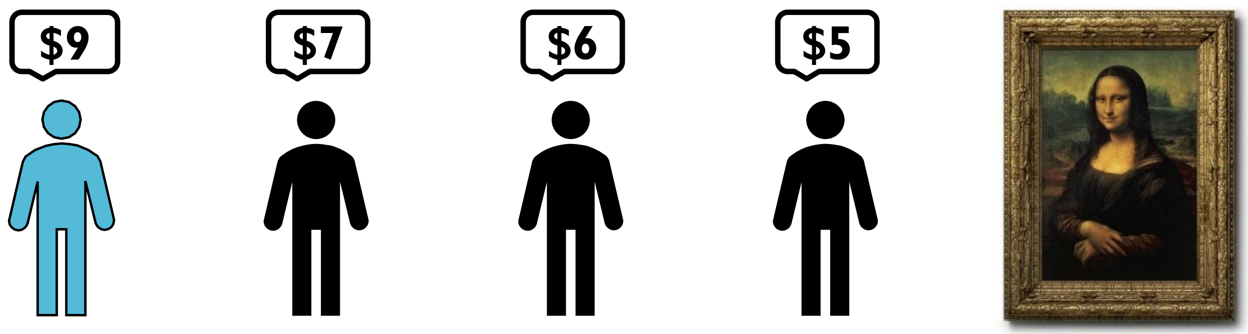
Mechanism design example:
Second-price auction with a reserve

Auctioneer sets reserve price p.

Highest bidder wins if bid $\geq p$. Pays maximum of second highest bid and p.

Reserve price: \$8 ➡ Revenue = **\$8**

Reserve price: \$6 ➡ Revenue = **\$7**



How to choose the reserve price?

This talk:

How can we use machine learning to design auctions with high revenue?

Booming area of economics and computer science

E.g., Likhodedov and Sandholm, AAAI'04, AAAI'05; Balcan, Blum, Hartline, and Mansour, FOCS'05; Elkind, SODA'07; Dhangwatnotai, Roughgarden, and Yan, EC'10; Mohri and Medina, ICML'14; Cole and Roughgarden STOC'14; Morgenstern and Roughgarden, COLT'16; Cai and Daskalakis FOCS'17; …

Helps overcome traditional, manual approaches to mechanism design

The revenue-maximizing auction is not known even when there are just two buyers and two items!

Outline

1. Introduction
2. **Background**
3. Machine learning for mechanism design
4. Conclusion

Notation

There are m items and n buyers.

Each buyer i has a value $v_i(b) \in \mathbb{R}$ for each bundle $b \subseteq [m]$.

Let $\boldsymbol{v}_i = \big(v_i(b_1), \dots, v_i(b_{2^m})\big)$ for all $b_1, \dots, b_{2^m} \subseteq [m]$.

Example

Items = {  ,  }

v_i ($\emptyset$) = 0 v_i () = 2 v_i () = 3 v_i ( , ) = 6

$\boldsymbol{v}_i =$ [v_i ($\emptyset$), v_i (), v_i (), v_i ( , )]

Notation

There are m items and n buyers.

Each buyer i has a value $v_i(b) \in \mathbb{R}$ for each bundle $b \subseteq [m]$.

Let $\boldsymbol{v}_i = \left(v_i(b_1), \dots, v_i(b_{2^m})\right)$ for all $b_1, \dots, b_{2^m} \subseteq [m]$.

Example

Items = {  ,  }

v_i ($\emptyset$) = 0

v_i () = 2

v_i () = 3

v_i ( , ) = 6

$\boldsymbol{v}_i = [\ 0\ ,\ 2\ ,\ 3\ ,\ 6\ ]$

Classical mechanism design

Standard assumption

A buyer's valuations are defined by a probability **distribution** over all the possible valuations she might have for all bundles of goods.
The mechanism designer knows this distribution.

Example

$(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \sim \mathcal{D}$, where $\boldsymbol{v}_i = \left[v_i(\emptyset), v_i(☕), v_i(🧁), v_i(☕\ 🧁) \right]$

Where does this information come from?

Outline

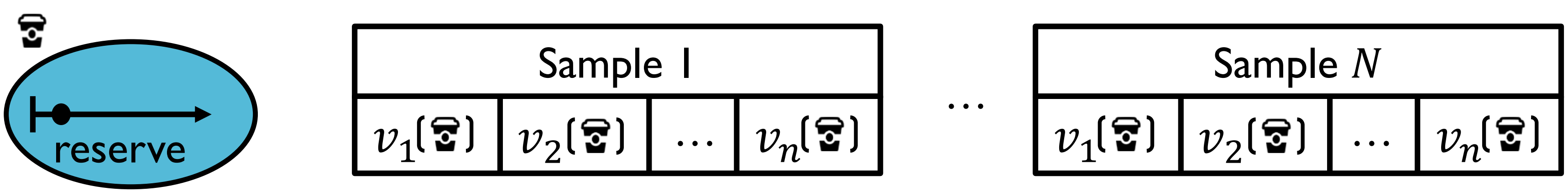
1. Introduction
2. Background
3. **Machine learning for mechanism design**
4. Conclusion

Mechanism design as a learning problem

Goal: Given mechanism family $\mathcal{M}$ and set of buyers' values sampled from unknown distribution $\mathcal{D}$, find mechanism with high expected revenue

- **Large family $\mathcal{M}$ of parametrized mechanisms**
  (E.g., 2nd-price auctions w/ reserves or posted price mechanisms)
- **Set of buyers' values sampled from unknown distribution $\mathcal{D}$**

2nd price auctions with reserves:



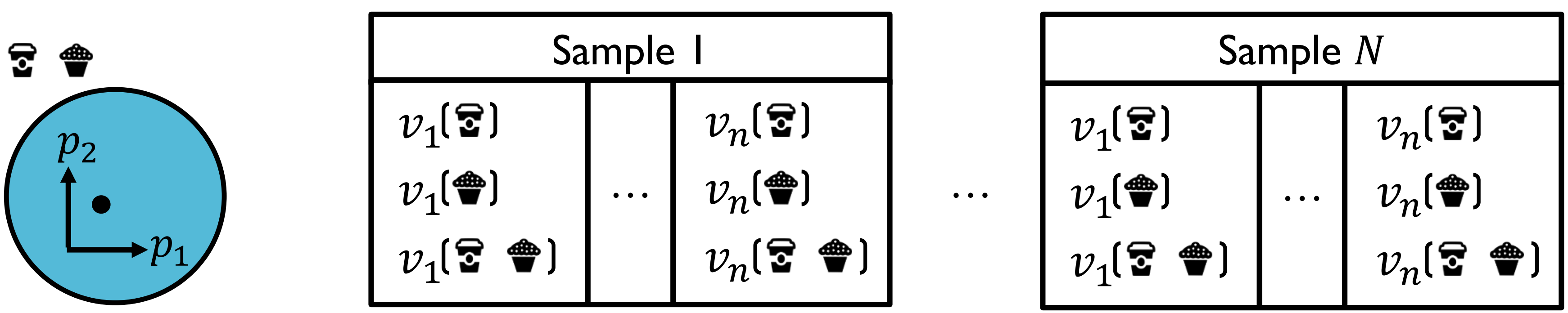
Mechanism design as a learning problem

Goal: Given mechanism family $\mathcal{M}$ and set of buyers' values sampled from unknown distribution $\mathcal{D}$, find mechanism with high expected revenue

- **Large family $\mathcal{M}$ of parametrized mechanisms**

  (E.g., 2nd-price auctions w/ reserves or posted price mechanisms)
- **Set of buyers' values sampled from unknown distribution $\mathcal{D}$**

Posted price mechanisms:



Mechanism design as a learning problem

Goal: Given mechanism family $\mathcal{M}$ and set of buyers' values sampled from unknown distribution $\mathcal{D}$, find mechanism with high expected revenue

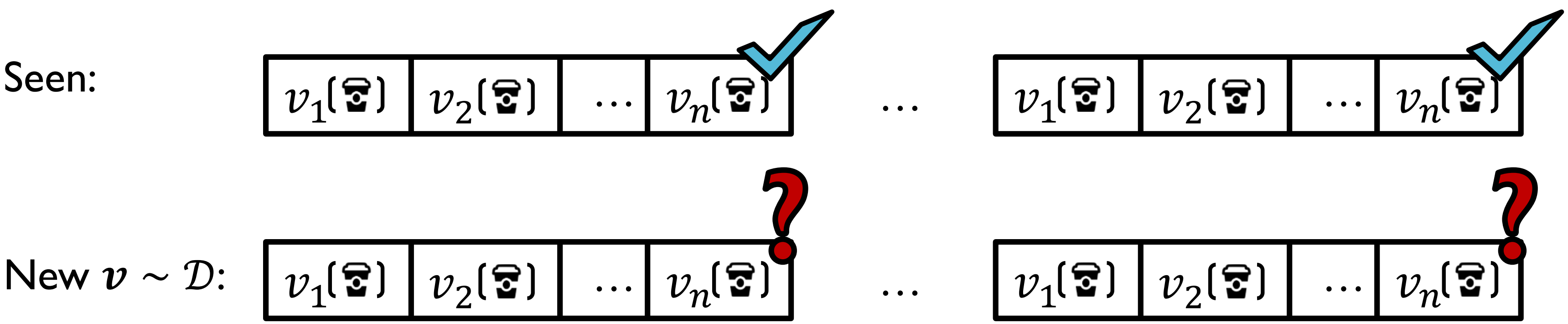
Approach: Find $\widehat{M}$ (nearly) optimal mechanism over the set of samples.

Mechanism design as a learning problem

Goal: Given mechanism family $\mathcal{M}$ and set of buyers' values sampled from unknown distribution $\mathcal{D}$, find mechanism with high expected revenue

Approach: Find $\widehat{M}$ (nearly) optimal mechanism over the set of samples.

Key question: Will $\widehat{M}$ have high expected revenue?



Will $\widehat{M}$ have high revenue over $\mathcal{D}$?

Mechanism design as a learning problem

Goal: Given mechanism family $\mathcal{M}$ and set of buyers' values sampled from unknown distribution $\mathcal{D}$, find mechanism with high expected revenue

Approach: Find $\hat{M}$ (nearly) optimal mechanism over the set of samples

Key question: Will $\hat{M}$ have high expected revenue?

Technical tool: uniform convergence

For any mechanism in class $\mathcal{M}$, average revenue over samples close to its expected revenue

Implies $\hat{M}$ has high expected revenue

Mechanism design as a learning problem

Goal: Given mechanism family $\mathcal{M}$ and set of buyers' values sampled from unknown distribution $\mathcal{D}$, find mechanism with high expected revenue

Approach: Find $\widehat{M}$ (nearly) optimal mechanism over the set of samples

Key question: Will $\widehat{M}$ have high expected revenue?

Technical tool: uniform convergence

Learning theory: $N = \tilde{O}(\mathrm{dim}(\mathcal{M})\ /\epsilon^2)$ samples suffice for ϵ-close

Challenge: Analyze dim($\mathcal{M}$) for complex combinatorial, modular mechanisms

Mechanism design as a learning problem

Goal: Given mechanism family $\mathcal{M}$ and set of buyers' values sampled from unknown distribution $\mathcal{D}$, find mechanism with high expected revenue

Learning theory: $N = \tilde{O}(\dim(\mathcal{M})\,/\epsilon^2)$ samples suffice for ϵ-close

Our results:
General way to bound $\dim(\mathcal{M})$ for any mechanism class satisfying **key structural property**: revenue is piecewise linear function of class's parameters

Many applications to multi-item, multi-buyer scenarios

Second-price auctions with reserves, posted price mechanisms, two-part tariffs, parameterized VCG mechanisms, etc.

Outline

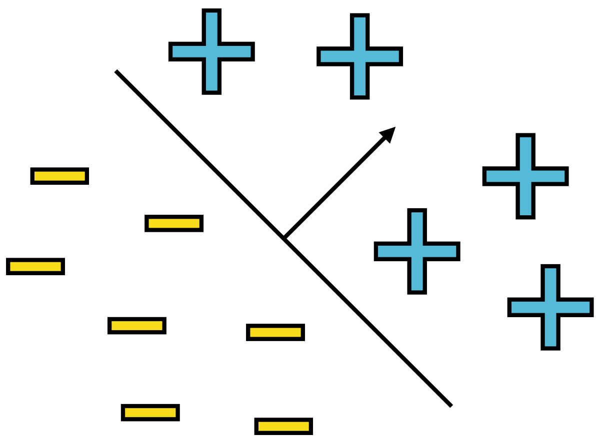
1. Introduction
2. Background
3. Machine learning for mechanism design
    a. **Learning theory tools**
    b. Simple example
    c. General theory
    d. Applications of general theory
4. Conclusion

VC dimension

Complexity measure characterizing the sample complexity of **binary-valued** function classes

(Classes of functions $h : \mathcal{X} \rightarrow \{-1,1\}$*)*

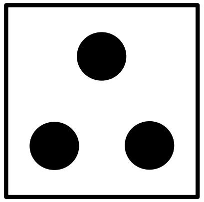
E.g., linear separators



VC dimension

VC-dimension of a function class $\mathcal{H} = \{h : \mathcal{X} \to \{-1,1\}\}$ is the cardinality of the largest set $\mathcal{S} \subseteq \mathcal{X}$ that can be labeled in all $2^{|\mathcal{S}|}$ ways by functions in $\mathcal{H}$.

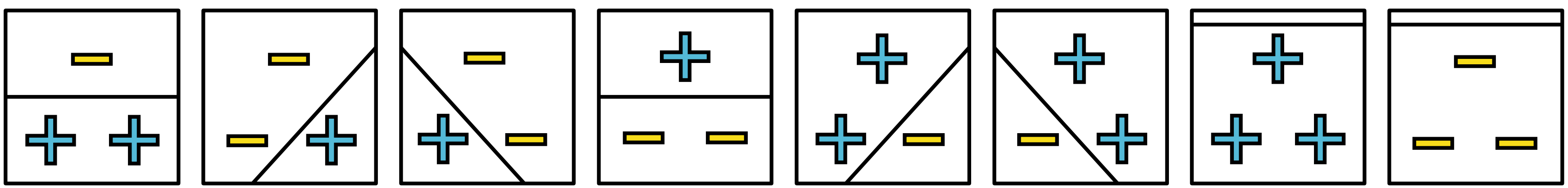
Example: $\mathcal{H}$ = Linear separators in $\mathbb{R}^2$. VCdim($\mathcal{H}$) ≥ 3.



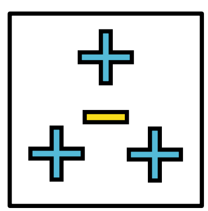
VC dimension

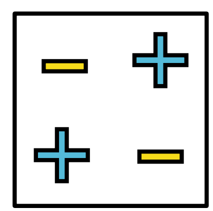
VC-dimension of a function class $\mathcal{H} = \{h : \mathcal{X} \to \{-1,1\}\}$ is the cardinality of the largest set $\mathcal{S} \subseteq \mathcal{X}$ that can be labeled in all $2^{|\mathcal{S}|}$ ways by functions in $\mathcal{H}$.

Example: $\mathcal{H}$ = Linear separators in $\mathbb{R}^2$.   VCdim($\mathcal{H}$) ≥ 3.



VCdim($\mathcal{H}$) ≤ 3.





VCdim({Linear separators in $\mathbb{R}^d$}) $= d + 1$.

VC dimension

VC-dimension of a function class $\mathcal{H} = \{h : \mathcal{X} \to \{-1,1\}\}$ is the cardinality of the largest set $\mathcal{S} \subseteq \mathcal{X}$ that can be labeled in all $2^{|\mathcal{S}|}$ ways by functions in $\mathcal{H}$.

Theorem [Vapnik and Chervonenkis, '71]

For any $\epsilon \in (0,1)$ and any distribution $\mathcal{D}$ over $\mathcal{X}$, with high probability over the draw of $N = \widetilde{\Theta}\left(\frac{\mathbf{VCdim}(\mathcal{H})}{\epsilon^2}\right)$ samples $\{x_1, \ldots, x_N\} \sim \mathcal{D}^N$, for all $h \in \mathcal{H}$,

$$\left| \mathbb{E}_{x\sim\mathcal{D}}[h(x)] - \frac{1}{N}\sum_{i=1}^{N} h(x_i) \right| \leq \epsilon.$$

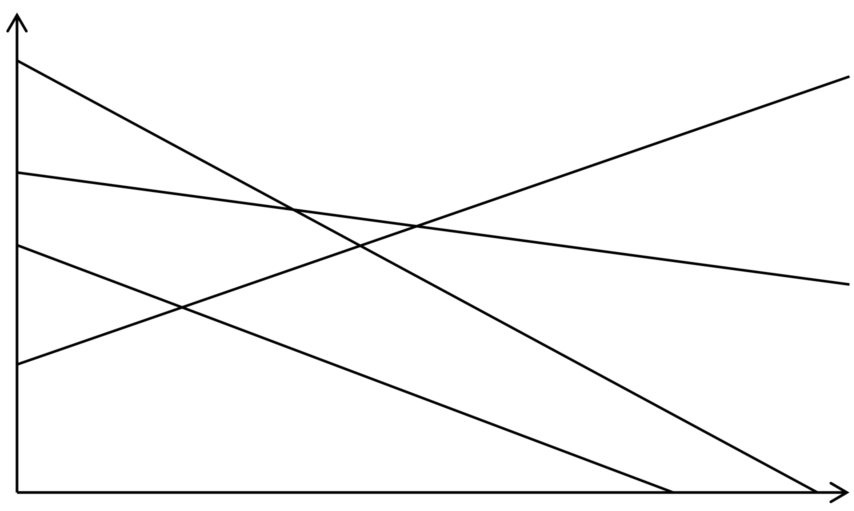
What about real-valued functions?

Pseudo-dimension

Complexity measure characterizing the sample complexity of **real-valued** function classes

(Classes of functions $f : \mathcal{X} \rightarrow [0,1]$*)*

E.g., affine functions

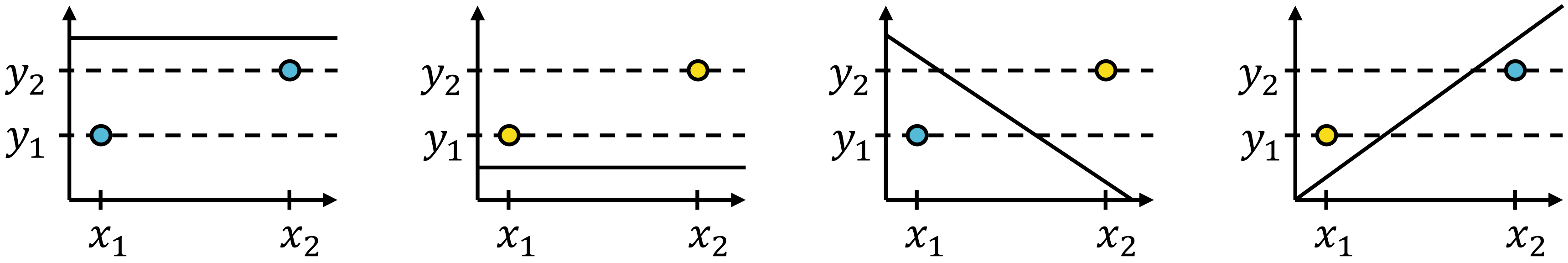


Pseudo-dimension

The **pseudo-dimension** of a class $\mathcal{F} = \{f : \mathcal{X} \to [0,1]\}$ is the cardinality of the largest set $\mathcal{S} = \{x_1, \dots, x_N\} \subseteq \mathcal{X}$ s.t. for some thresholds $y_1, \dots, y_N \in \mathbb{R}$, all 2^N above/below binary patterns can be achieved by functions $f \in \mathcal{F}$.

Example: $\mathcal{F}$ = Affine functions in $\mathbb{R}$.

$\mathrm{Pdim}(\mathcal{F}) \geq 2$.



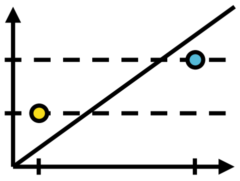
Pseudo-dimension

The **pseudo-dimension** of a class $\mathcal{F} = \{f : \mathcal{X} \to [0,1]\}$ is the cardinality of the largest set $\mathcal{S} = \{x_1, \dots, x_N\} \subseteq \mathcal{X}$ s.t. for some thresholds $y_1, \dots, y_N \in \mathbb{R}$, all 2^N above/below binary patterns can be achieved by functions $f \in \mathcal{F}$.

Theorem [Pollard, 1984]

For any $\epsilon \in (0,1)$ and any distribution $\mathcal{D}$ over $\mathcal{X}$, with high probability over the draw of $N = \widetilde{\Theta}\left(\frac{\mathbf{Pdim}(\mathcal{F})}{\epsilon^2}\right)$ samples $\{x_1, \dots, x_N\} \sim \mathcal{D}^N$, for all $f \in \mathcal{F}$,

$$\left| \mathbb{E}_{x \sim \mathcal{D}}[f(x)] - \frac{1}{N} \sum_{i=1}^{N} f(x_i) \right| \le \epsilon.$$



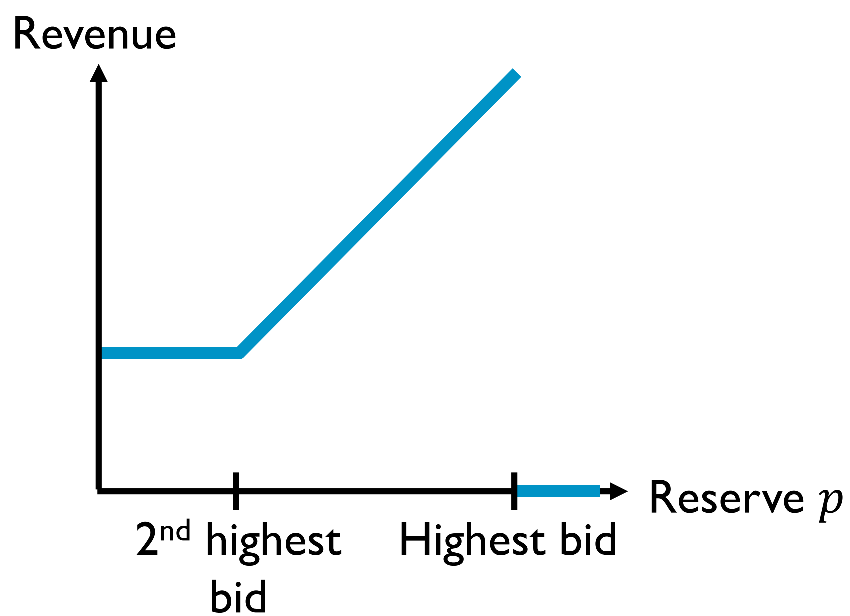
Outline

1. Introduction
2. Background
3. Machine learning for mechanism design
   a. Learning theory tools
   b. **Simple example**
   c. General theory
   d. Applications of general theory
4. Conclusion

Example:
P-dim of 2nd-price auctions with reserves

2nd-price auction with a reserve

- Auctioneer sets reserve price p
- Highest bidder wins if bid $\geq p$. Pays maximum of second highest bid and p



Claim
For a fixed set of bids, revenue is a piecewise linear function of the reserve.

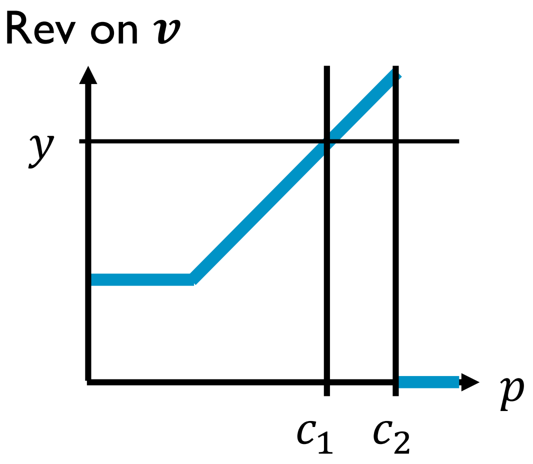
Example: P-dim of 2nd-price auctions with reserves

Theorem [Mohri-Medina'14; Morgenstern-Roughgarden'16; Balcan-Sandholm-**V**.'18]

$\mathcal{M} = \{\text{rev}_p :=$ revenue of 2nd-price auction with reserve $p\}$. Pdim$(\mathcal{M}) \leq 2$.

Key idea: Consider some valuation vector $\boldsymbol{v}$ and revenue-threshold y.

- Ranging p from 0 to ∞, will be (at most) two cutoff values c_1, c_2 where revenue goes from "below" to "above" to "below"
- With N examples, look at all $2N$ cutoff values
- All p in same interval between consecutive cutoff values will give same binary pattern
- So, at most $2N + 1$ binary patterns
- Pseudo-dimension is max N s.t. all 2^N binary above/below patterns are achievable
  - Need $2^N \leq 2N + 1$, so $N \leq 2$



Outline

1. Introduction
2. Background
3. Machine learning for mechanism design
   a. Learning theory tools
   b. Simple example
   c. **General theory**
   d. Applications of general theory
4. Conclusion

Bounding pseudo-dim of mechanism classes

Theorem

Suppose:

1. The mechanism class $\mathcal{M}$ is parameterized by vectors $\boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}$

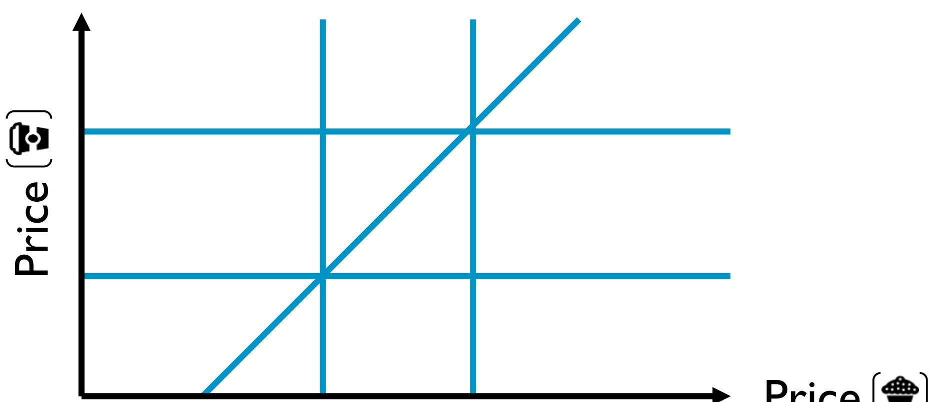
For example, $\boldsymbol{p}$ = (price (), price ())

Bounding pseudo-dim of mechanism classes

Theorem

Suppose:

1. The mechanism class $\mathcal{M}$ is parameterized by vectors $\boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}$
2. For every set $\boldsymbol{v}$ of buyers' values, a set of $\leq \boldsymbol{t}$ hyperplanes partition $\mathbb{R}^{\boldsymbol{d}}$ such that in every cell of this partition, $\text{revenue}_{\boldsymbol{v}}(\boldsymbol{p})$ is linear



In this example, $d = 2$ and $t = 5$.

Bounding pseudo-dim of mechanism classes

Theorem

Suppose:

1. The mechanism class $\mathcal{M}$ is parameterized by vectors $\boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}$
2. For every set $\boldsymbol{v}$ of buyers' values, a set of $\leq \boldsymbol{t}$ hyperplanes partition $\mathbb{R}^{\boldsymbol{d}}$ such that in every cell of this partition, $\text{revenue}_{\boldsymbol{v}}(\boldsymbol{p})$ is linear

Then $\text{Pdim}(\mathcal{M}) = O(\boldsymbol{d} \log(\boldsymbol{d}\boldsymbol{t}))$.

Bounding pseudo-dim of mechanism classes

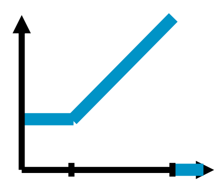
Corollary

Suppose:

1. The mechanism class $\mathcal{M}$ is parameterized by vectors $\boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}$
2. For every set $\boldsymbol{v}$ of buyers' values, a set of $\leq \boldsymbol{t}$ hyperplanes partition $\mathbb{R}^{\boldsymbol{d}}$ such that in every cell of this partition, $\text{revenue}_{\boldsymbol{v}}(\boldsymbol{p})$ is linear

For any $\epsilon \in (0,1)$, with high probability over the draw of $N = \tilde{\Theta}\left(\frac{\boldsymbol{d} \log(\boldsymbol{dt})}{\epsilon^2}\right)$ samples $\mathcal{S} = \{\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(N)}\} \sim \mathcal{D}^N$, for all mechanisms in $\mathcal{M}$:

$$|\text{average revenue over } \mathcal{S} - \text{expected revenue}| \leq \epsilon.$$



High-level learning theory bit

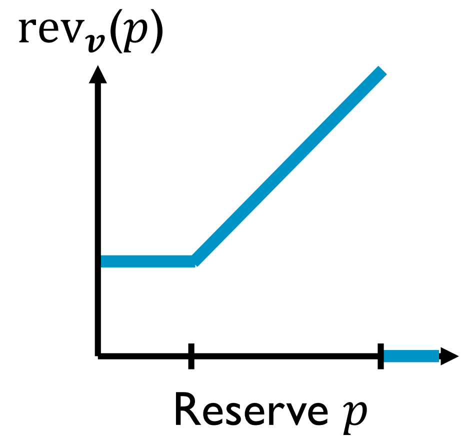
(Informal) **Theorem**
$\boldsymbol{d}$-dim. parameter space, $\boldsymbol{t}$ hyperplanes splitting parameters into linear pieces $\Rightarrow$ Pdim$(\mathcal{M}) = O(\boldsymbol{d}\log(\boldsymbol{dt}))$

Want to prove that for any mechanism parameters $\boldsymbol{p}$:

$$\frac{1}{|\mathcal{S}|}\sum_{\boldsymbol{v}\in\mathcal{S}} \mathrm{rev}_{\boldsymbol{p}}(\boldsymbol{v}) \text{ close to } \mathbb{E}[\mathrm{rev}_{\boldsymbol{p}}(\boldsymbol{v})]$$

Function class we analyze pseudo-dimension of:
$\{\mathrm{rev}_{\boldsymbol{p}} : \text{parameters } \boldsymbol{p} \in \mathbb{R}^{\boldsymbol{d}}\}$

Proof takes advantage of structure exhibited by **dual class** $\{\mathrm{rev}_{\boldsymbol{v}} : \text{buyer values } \boldsymbol{v}\}$



$$\mathrm{rev}_{\boldsymbol{v}}(\boldsymbol{p}) = \mathrm{rev}_{\boldsymbol{p}}(\boldsymbol{v})$$

Outline

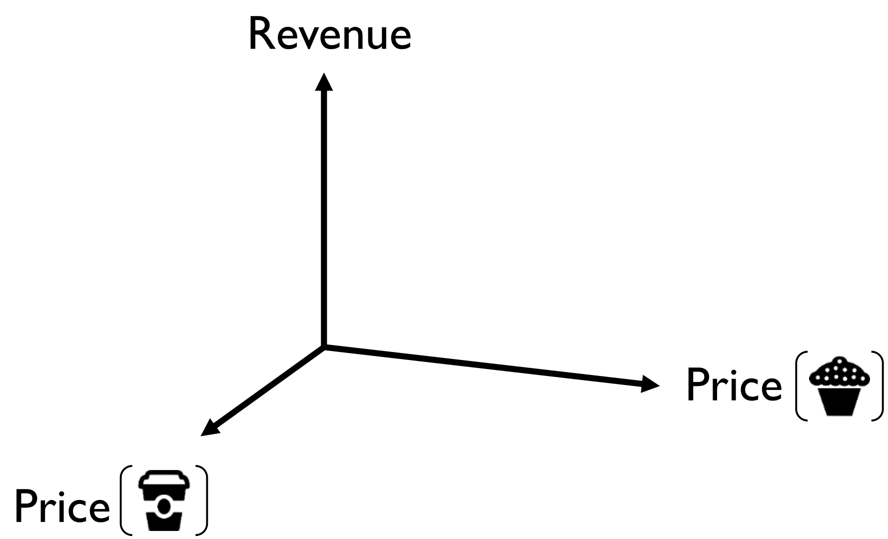
1. Introduction
2. Background
3. Machine learning for mechanism design
    a. Learning theory tools
    b. Simple example
    c. General theory
    d. **Applications of general theory**
4. Conclusion

Pseudo-dimension of posted price mechanisms

$\mathcal{M}$ = multi-item, multi-buyer posted price mechanisms

- Price per item.
- Fixed, arbitrary ordering on buyers.

1. First buyer in ordering arrives. Buys bundle of goods maximizing his utility.
2. Second buyer arrives. Buys bundle of remaining goods maximizing his utility.
3. Etc.



[E.g., Feldman, Gravin, Lucier, SODA'15; Babaioff, Immorlica, Lucier, Weinberg, FOCS'14; Cai Devanur, Weinberg, STOC'16]

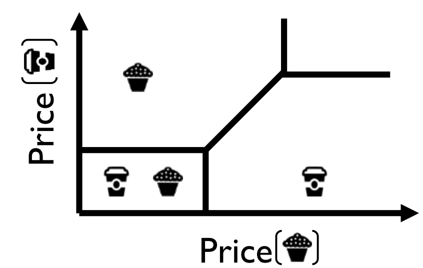
Pseudo-dimension of posted price mechanisms

Theorem

Pdim($\mathcal{M}$) $= O(d\log(dt))$ with $d = (\#\text{ dimensions}) = (\#\text{ items})$ and $t = (\#\text{ hyperplanes}) = (\#\text{ buyers}) \cdot \binom{2^{(\#\text{ items})}}{2}$.

Proof. For **every buyer** and **every pair of bundles**, decision boundary (determining where buyer prefers one bundle over another) is a hyperplane

- $(\#\text{ bundles}) = 2^{(\#\text{ items})}$, so $(\#\text{ buyers})\binom{2^{(\#\text{ items})}}{2}$ hyperplanes create partition where across all prices in a single region, all buyers' preference orderings are fixed
- When preference ordering fixed, bundles they buy are fixed. So revenue is linear function of items the buy



Our main applications

- Match or improve over the best-known guarantees for many those classes previously studied.
- Prove bounds for classes not yet studied from a learning perspective.

Mechanism class	**Sample complexity studied before?**
Randomized mechanisms (lotteries)	N/A
Multi-part tariffs and other non-linear pricing mechanisms	N/A
Posted price mechanisms	E.g., Morgenstern and Roughgarden, '16; Syrgkanis '17
Affine maximizer auctions	Balcan, Sandholm, and **V.**, '16
Second price auctions with reserves	E.g., Devanur et al., '16; Morgenstern and Roughgarden, '16

Outline

1. Introduction
2. Background
3. Machine learning for mechanism design
4. **Conclusion**

Discussion and open directions

- General way to analyze $\dim(\mathcal{M})$ for any class $\mathcal{M}$ of mechanisms whose revenue is a piecewise linear function of the class's parameters
- Many applications to multi-item, multi-buyer scenarios
  - Second-price auctions with reserves, posted price mechanisms, two-part tariffs, parameterized VCG mechanisms, etc.

Open questions

- Algorithmic aspects to data-driven mechanism design
- Other data-driven mechanism design applications beyond selling and/or revenue maximization

Thanks!

Questions?