Introduction to Auction Design via Machine Learning

Ellen Vitercik

December 4, 2017
CMU 10-715
Advanced Introduction to Machine Learning

Ad auctions contribute a huge portion of large internet companies' revenue.

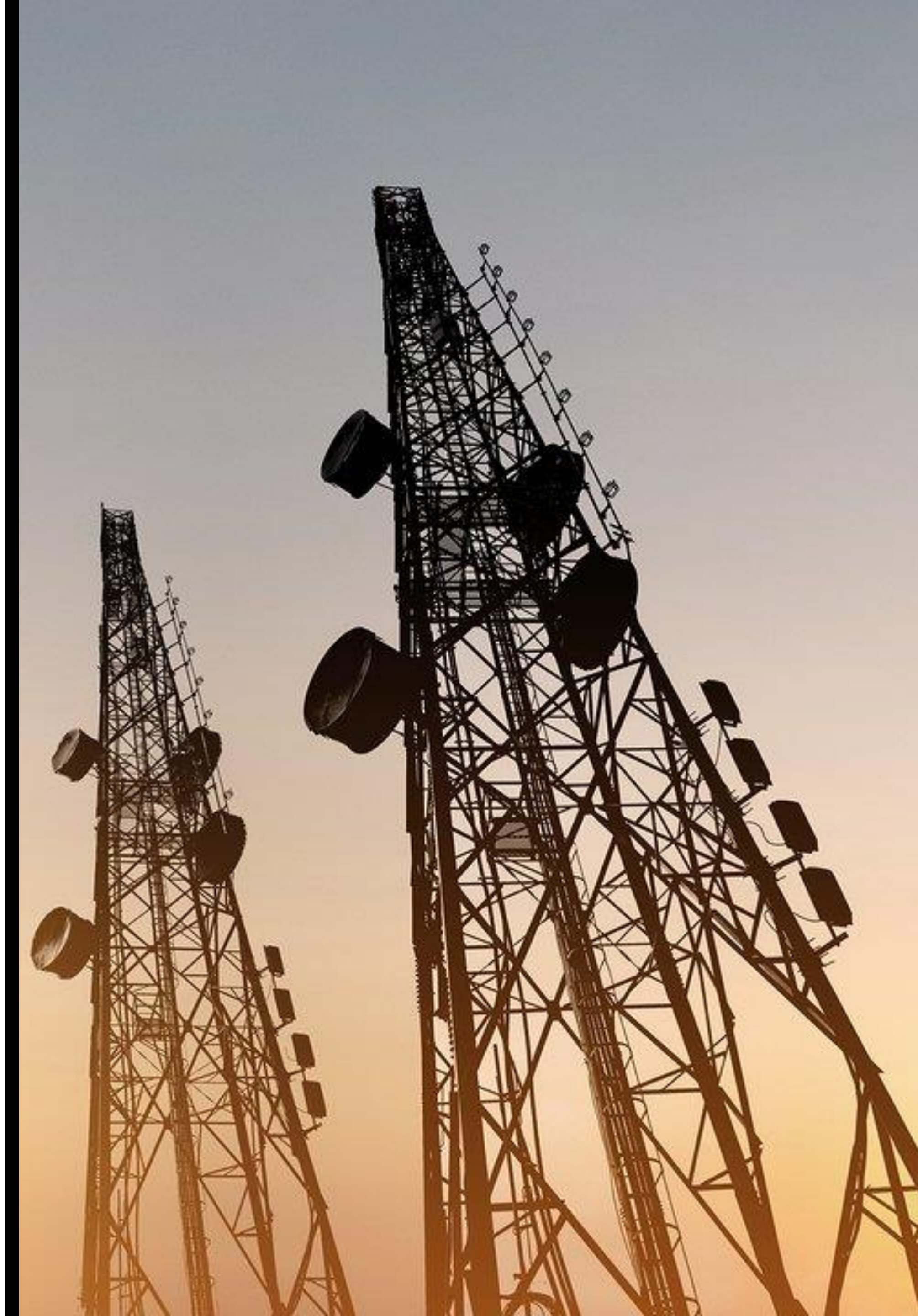
	Ad revenue in 2016	Total revenue in 2016
Google	$79 billion	$89.46 billion
Facebook	$27 billion	$27.64 billion

Leonardo da Vinci painting sells for $450.3m, shattering auction highs
[NY Times ‘17]

Diamond fetches $33.7m at Christie's auction in Geneva
[BBC '17]

The FCC spectrum auction is sending $10 billion to broadcasters

[NiemanLab ‘17]



Very-large-scale generalized combinatorial multi-attribute auctions: Lessons from conducting $60 billion of sourcing

[Sandholm '13]

How can we design revenue-maximizing auctions using machine learning?

1. Introduction
2. Auction design background
3. Revenue-maximizing auctions
4. Learning-theoretic auction design
5. Tools
6. Case study: Second price auctions with reserve prices
7. Related work

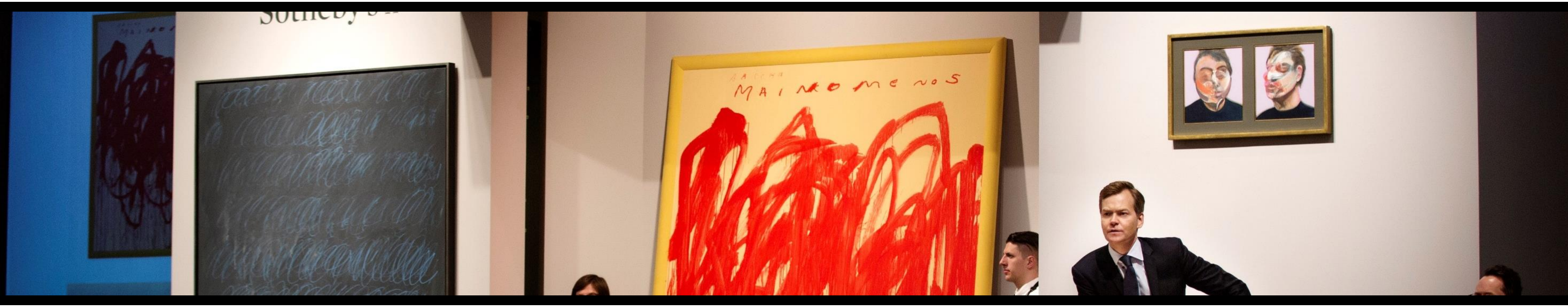
Common misconception:

There's only one way to hold an auction.

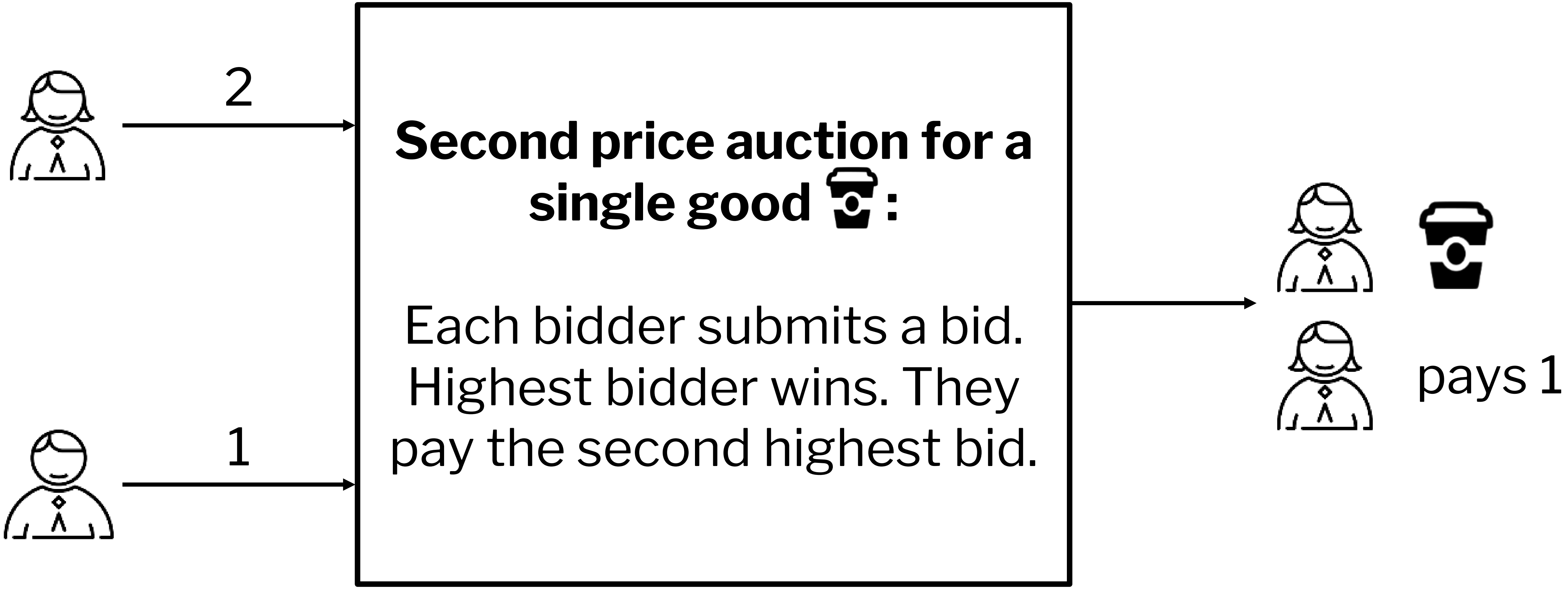
There are **infinitely** many ways to design auctions.

Ascending price auction: Auctioneer calls out the opening bid. Bidders call out bids until there is no new bid within a period of time. Last bidder to bid wins the item and pays their bid.

Ascending price auction: Auctioneer calls out the opening bid. Bidders call out bids until there is no new bid within a period of time. Last bidder to bid wins the item and pays their bid.



Online auctions can only take a **fraction of a second**, so we need faster auction formats.


2
1
Second price auction for a single good :
Each bidder submits a bid. Highest bidder wins. They pay the second highest bid.
pays 1

The second price auction is **incentive compatible** (every bidder will maximize their **utility** by bidding truthfully).

$$\boxed{\left(\text{value}\left(\ \right) - \text{payment}\right) \cdot \mathbf{1}\left(\text{wins item}\right)}$$

Why not bid above value ( ) ?

If winner, will stay winner and price won't change.

If loser, might become winner, but will have to pay a price larger than value ( ), which will result in negative utility.

The second price auction is **incentive compatible** (every bidder will maximize their **utility** by bidding truthfully).

(value() — payment) $\cdot\, \mathbf{1}$(wins item)

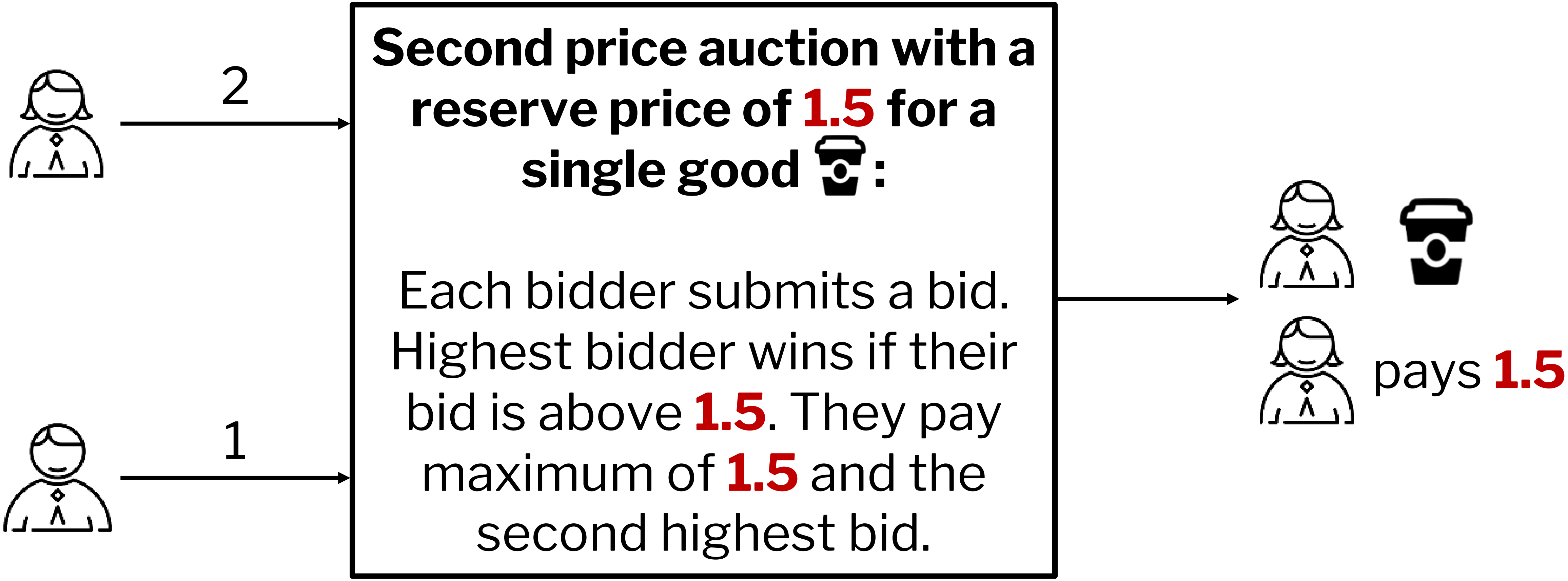
Why not bid below value()?

If winner, might become loser, and will shift from non-negative to zero utility.

If loser, will still be loser, so utility will still be zero.

Second price auction with a reserve price $r \in \mathbb{R}$ for a single good :

Each bidder submits a bid. Highest bidder wins if their bid is above r. They pay maximum of r and the second highest bid.



Optimizing reserve prices can lead to higher revenue.

But how should we choose the reserve price?

1. Introduction
2. Auction design background
3. Revenue-maximizing auctions
4. Learning-theoretic auction design
5. Tools
6. Case study: Second price auctions with reserve prices
7. Related work

In 1981, Roger Myerson discovered the **revenue-maximizing** single-item, multi-bidder auction.

He won the 2007 Nobel prize in economics for his work on mechanism design.

R. Myerson. Optimal auction design. Mathematics of Operations Research, 6(1):58–73, 1981.

Standard assumption

A buyer's value for a bundle of goods is defined by a probability **distribution** over all the possible valuations she might have for that bundle.

value() ~ Uniform(1, 3)

Standard assumption

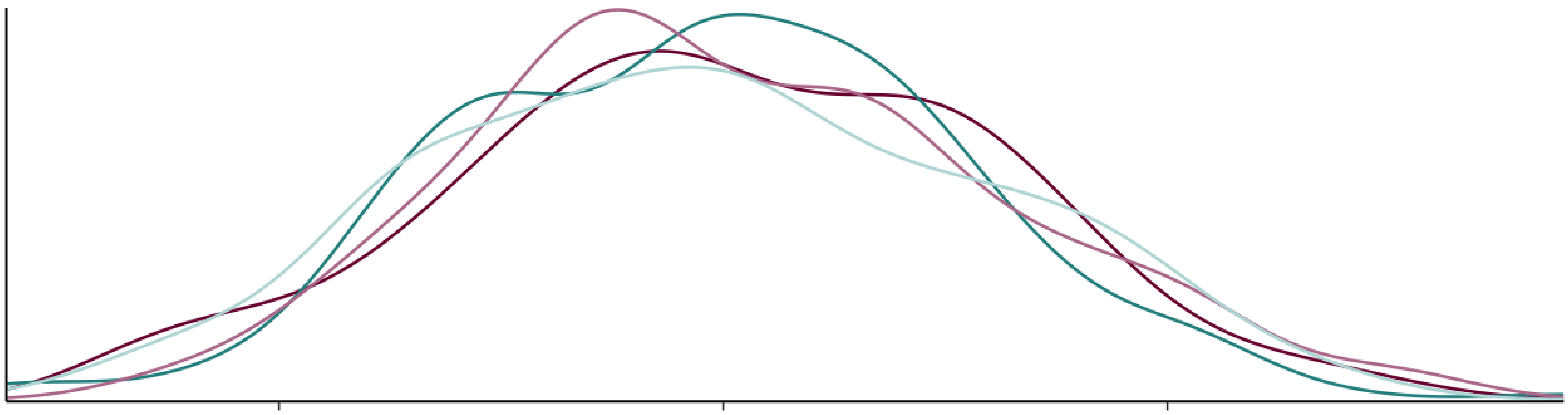
A buyer's value for a bundle of goods is defined by a probability **distribution** over all the possible valuations she might have for that bundle.

For simplicity, suppose all bidders have the same valuation distribution, with PDF f, CDF F, and support in [0, 1].

Myerson's optimal auction

Let $\phi(x) = x - \frac{1-F(x)}{f(x)}$. Run a **second price auction** with a reserve price of $\phi^{-1}(0)$.

Myerson also analyzed the case where the bidders' values are **not identically distributed**.



The revenue-maximizing **multi-item** auction is **unknown**!

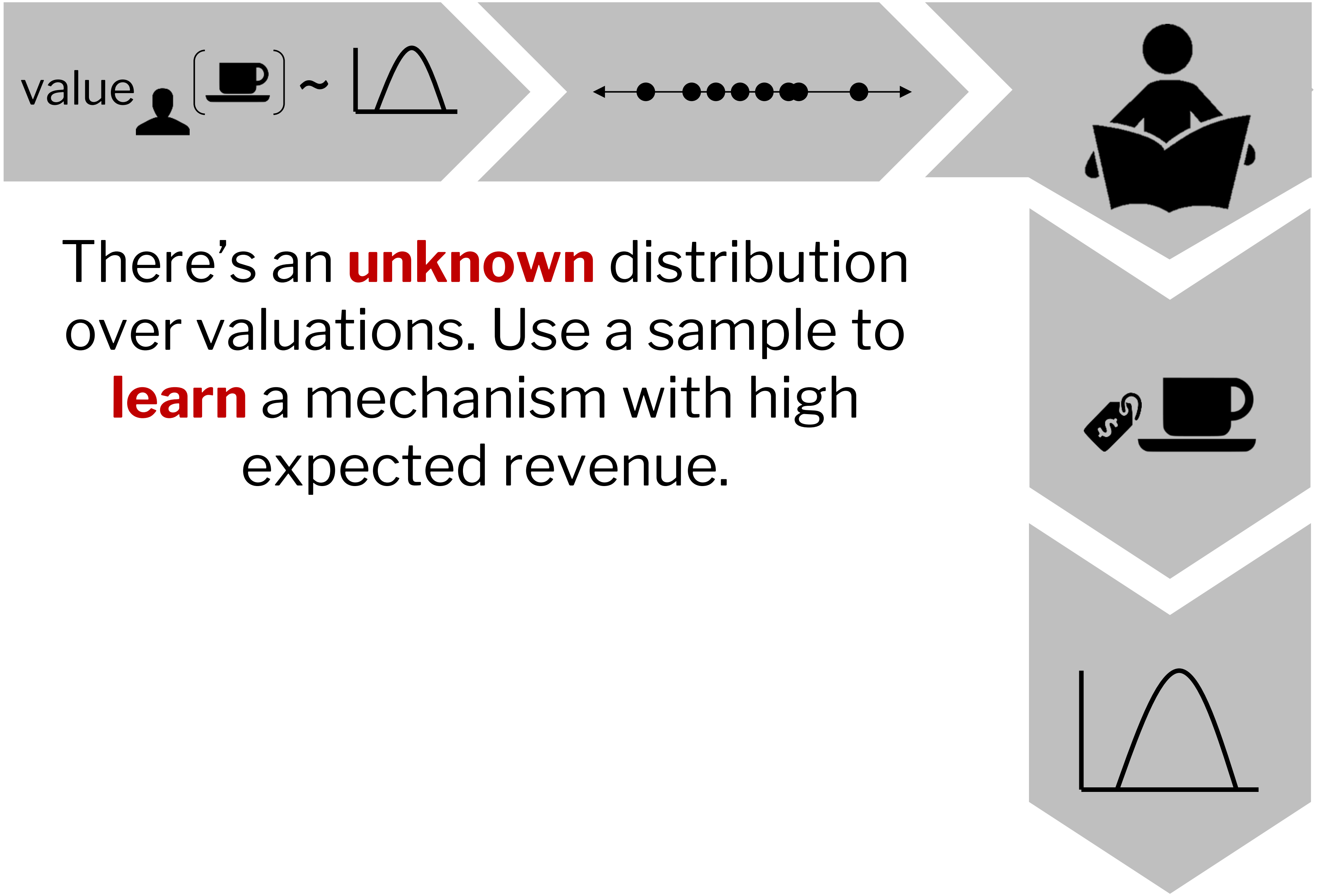
?

Myerson's optimal auction
Let $\phi(x) = x - \frac{1-F(x)}{f(x)}$. Run a **second price auction** with a reserve price of $\phi^{-1}(0)$.

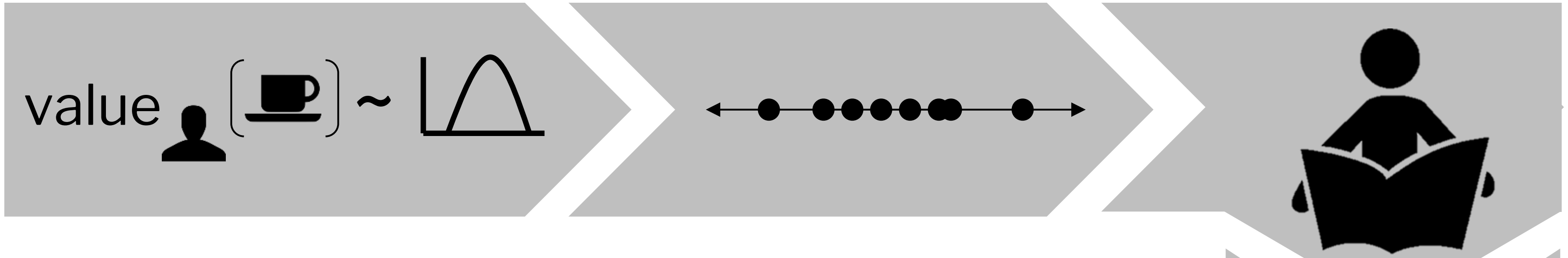
This requires that we **know** the distribution over bidders' values.

- ❌ Where does this knowledge come from?
- ❌ In multi-item cases, the support of the distribution might be doubly exponential!

1. Introduction
2. Auction design background
3. Revenue-maximizing auctions
4. Learning-theoretic auction design
5. Tools
6. Case study: Second price auctions with reserve prices
7. Related work

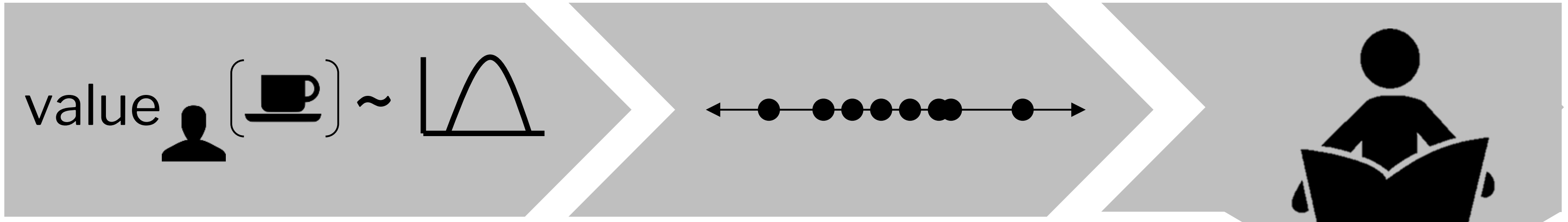


There's an **unknown** distribution over valuations. Use a sample to **learn** a mechanism with high expected revenue.



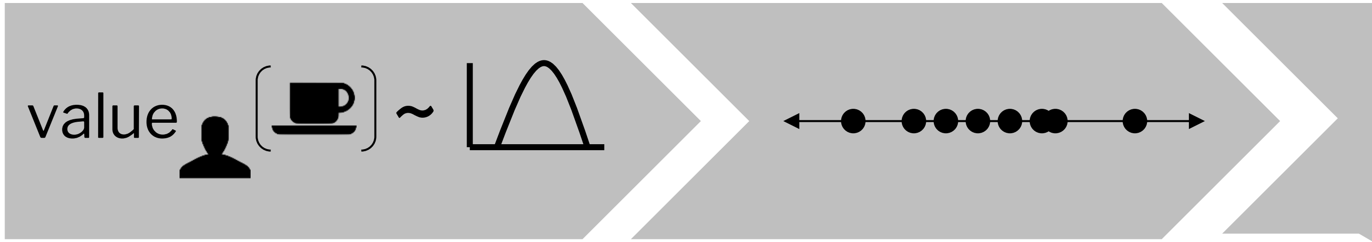
Research papers:

[Likhodedov and Sandholm ‘04, ‘05, Elkind ’07, Dhangwatnotai et al. ’10, Medina and Mohri ‘14, Cole and Roughgarden ’14, Huang et al. ’15, Morgenstern and Roughgarden ’15, ‘16, Roughgarden and Schrijvers ’16, Devanur et al. ’16, Balcan, Sandholm, and **V** ‘16, ’17, Hartline and Taggart ’17, Syrgkanis ‘17, Alon et al. ’17, Cai and Daskalakis ’17, Gonczarowski and Nisan ’17, Medina and Vassilvitskii ’17, Dütting et al. ‘17]



Classes:

Nina Balcan	Connections between Learning, Game Theory, and Optimization	Georgia Tech	Fall 2010
Constantinos Daskalakis and Vasilis Syrgkanis	Algorithmic Game Theory and Data Science	MIT	Spring 2017



Workshops and Tutorials:

2016	EC Tutorial on Algorithmic Game Theory and Data Science
2015, 2016, 2017	EC Workshop on Algorithmic Game Theory and Data Science
2017	NIPS Workshop on Learning in Presence of Strategic Behavior
2017	Dagstuhl Workshop on Game Theory Meets Computational Learning Theory

Fix an auction class $\mathcal{M}$

$\Downarrow$

Look at the samples

$\Downarrow$

Optimize over $\mathcal{M}$

$\Downarrow$

Field auction

Central question: Is the resulting auction nearly optimal?

Fix an auction class $\mathcal{M}$

$\Downarrow$

Look at the samples

$\Downarrow$

Optimize over $\mathcal{M}$

$\Downarrow$

Field auction

Central question: Does the resulting auction approximately maximize revenue in expectation over the unknown distribution?

Central question: Does the resulting auction approximately maximize revenue in expectation over the unknown distribution?

Depends on the class $\mathcal{M}$.

Let M^* be the optimal auction and let $\widehat{M} \in \mathcal{M}$ be the auction returned by the learning algorithm.

$$\underbrace{\mathbb{E}[\mathrm{Rev}(M^*)] - \mathbb{E}\left[\mathrm{Rev}\left(\widehat{M}\right)\right]}_{\text{Total revenue loss}}$$

Central question: Does the resulting auction approximately maximize revenue in expectation over the unknown distribution?

Depends on the class $\mathcal{M}$.

Let M^* be the optimal auction and let $\widehat{M} \in \mathcal{M}$ be the auction returned by the learning algorithm.

$$\mathbb{E}[\mathrm{Rev}(M^*)] - \mathbb{E}\left[\mathrm{Rev}\left(\widehat{M}\right)\right]$$
$$= \underbrace{\mathbb{E}[\mathrm{Rev}(M^*)] - \max_{M \in \mathcal{M}} \mathbb{E}[\mathrm{Rev}(M)]}_{\text{Approximation loss}} + \underbrace{\max_{M \in \mathcal{M}} \mathbb{E}[\mathrm{Rev}(M)] - \mathbb{E}\left[\mathrm{Rev}\left(\widehat{M}\right)\right]}_{\text{Estimation loss}}$$

Central question: Does the resulting auction approximately maximize revenue in expectation over the unknown distribution?

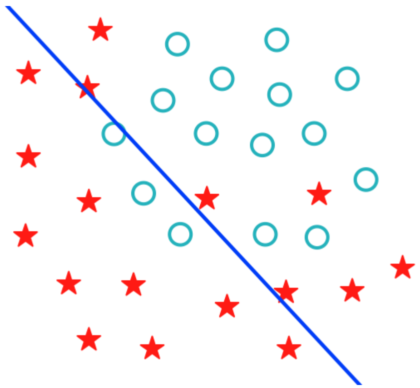
Depends on the class $\mathcal{M}$.

Simple class ⟷ Complex class

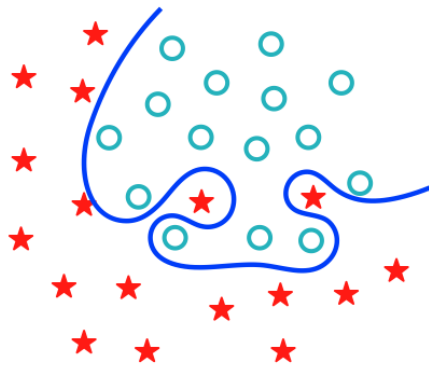
Bad approximation loss
Good estimation loss

Good approximation loss
Bad estimation loss

These are the same problems we face in machine learning.



Simple class



Complex class

Bad approximation loss
Good estimation loss

Good approximation loss
Bad estimation loss

Say $\mathcal{M}$ is the class of second price auctions with reserve prices.

Myerson's optimal auction for i.i.d. bidders
Let $\phi(x) = x - \frac{1-F(x)}{f(x)}$. Run a second price auction with a reserve price of $\phi^{-1}(0)$.

$\mathcal{M}$ has zero **approximation loss** when the bidders' values are i.i.d.!

This talk: What is the **estimation loss** of $\mathcal{M}$?

1. Introduction
2. Auction design background
3. Revenue-maximizing auctions
4. Learning-theoretic auction design
5. Tools
6. Case study: Second price auctions with reserve prices
7. Related work

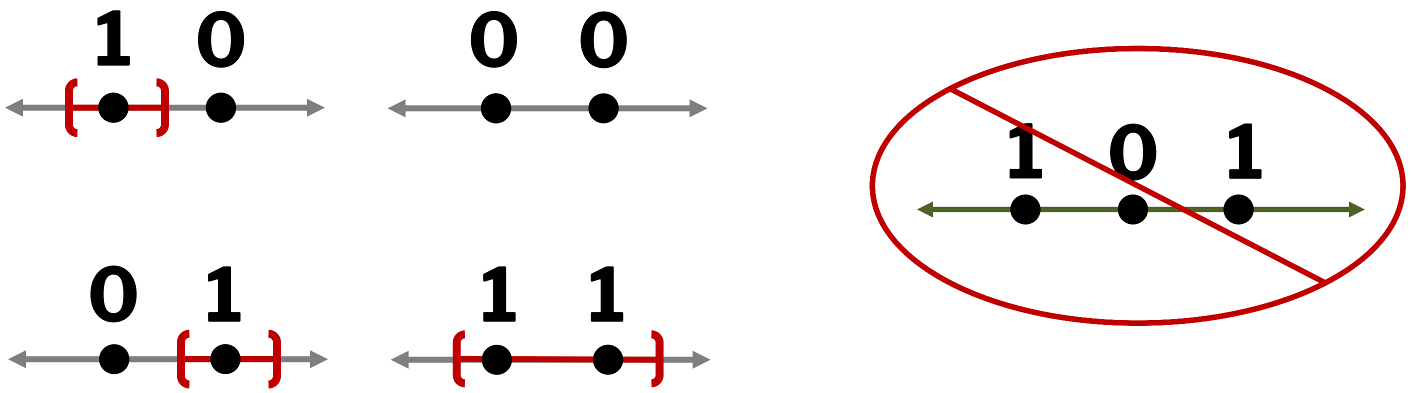
VC dimension:

A complexity measure for binary-valued functions ($h: \mathcal{X} \to \{0,1\}$)

Example.

For $a, b \in \mathbb{R}, a < b$, let $h_{a,b}(x) = \begin{cases} 1, & x \in (a,b) \\ 0, & x \notin (a,b) \end{cases}$

Let $\mathcal{H}_{\text{int}} = \{h_{a,b}: a, b \in \mathbb{R}, a < b\}$.

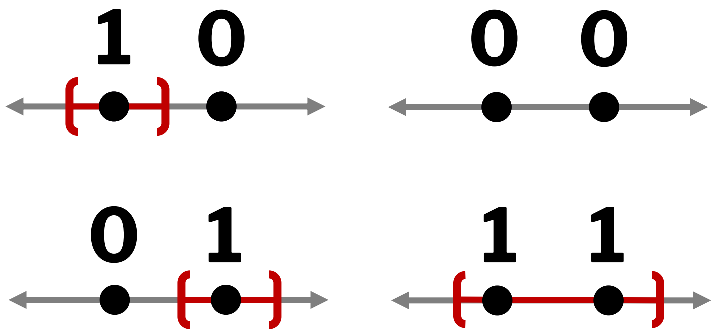


No set $S \subset \mathbb{R}$ of size 3 can be labeled in all 2^3 ways by $\mathcal{H}_{\text{int}}$.

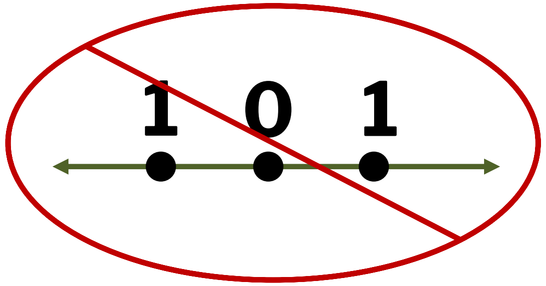
The **VC dimension** of a class $\mathcal{H}$ is the size of the largest set S that can be labeled in all $2^{|S|}$ ways by functions in $\mathcal{H}$.

VC dimension
A set $S = \{x_1, \ldots, x_m\}$ is *shattered* by $\mathcal{H}$ if for all $\boldsymbol{b} \in \{0,1\}^{\lvert S \rvert}$, there exists a function $h \in \mathcal{H}$ such that for all $i \in [m]$, $h(x_i) = b_i$. The VC dimension of $\mathcal{H}$ is the size of the largest set S that can be shattered by $\mathcal{H}$.

The **VC dimension** of a class $\mathcal{H}$ is the size of the largest set S that can be labeled in all $2^{|S|}$ ways by functions in $\mathcal{H}$.

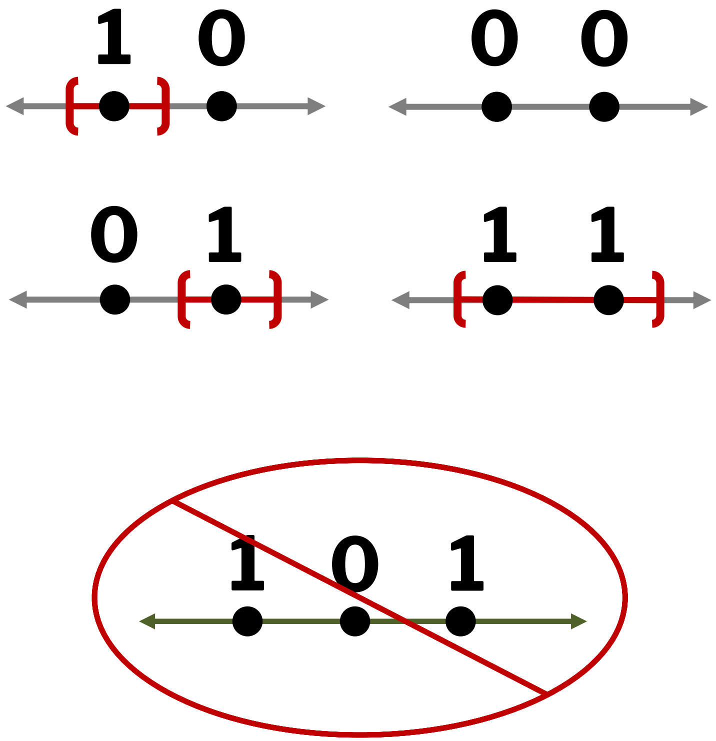


The VC dimension of $\mathcal{H}_{\mathrm{int}}$ is at least 2.



The VC dimension of $\mathcal{H}_{\mathrm{int}}$ is at most 2.

The **VC dimension** of a class $\mathcal{H}$ is the size of the largest set S that can be labeled in all $2^{|S|}$ ways by functions in $\mathcal{H}$.



The VC dimension of $\mathcal{H}_{\text{int}}$ is 2.

Pseudo-dimension:

A complexity measure for real-valued functions ($h: \mathcal{X} \to \mathbb{R}$)

Set $S = \{x_1, \dots, x_m\}$
Class of functions $\mathcal{H} = \{h: \mathcal{X} \to \mathbb{R}\}$

$\boldsymbol{z} = \left(z^{(1)}, \dots, z^{(m)}\right) \in \mathbb{R}^m$ **witnesses** the **shattering** of S by $\mathcal{H}$ if for all $\boldsymbol{b} \in \{0,1\}^m$, there exists $h \in \mathcal{H}$ such that $h(x_i) \leq z^{(i)}$ if $b_i = 0$ and $h(x_i) > z^{(i)}$ if $b_i = 1$

x_1

x_2

x_3

x_4

x_5

x_6

x_7

x_8

Set $S = \{x_1, \dots, x_m\}$

Class of functions $\mathcal{H} = \{h: \mathcal{X} \to \mathbb{R}\}$

$\boldsymbol{z} = \left(z^{(1)}, \dots, z^{(m)}\right) \in \mathbb{R}^m$ **witnesses** the **shattering** of S by $\mathcal{H}$ if for all $\boldsymbol{b} \in \{0,1\}^m$, there exists $h \in \mathcal{H}$ such that $h(x_i) \le z^{(i)}$ if $b_i = 0$ and $h(x_i) > z^{(i)}$ if $b_i = 1$

x_1 $z^{(1)}$

x_2 $z^{(2)}$

x_3 $z^{(3)}$

x_4 $z^{(4)}$

x_5 $z^{(5)}$

x_6 $z^{(6)}$

x_7 $z^{(7)}$

x_8 $z^{(8)}$

Set $S = \{x_1, \dots, x_m\}$
Class of functions $\mathcal{H} = \{h: \mathcal{X} \to \mathbb{R}\}$

$\boldsymbol{z} = \left(z^{(1)}, \dots, z^{(m)}\right) \in \mathbb{R}^m$ **witnesses** the **shattering** of S by $\mathcal{H}$ if for all $\boldsymbol{b} \in \{0,1\}^m$, there exists $h \in \mathcal{H}$ such that $h(x_i) \leq z^{(i)}$ if $b_i = 0$ and $h(x_i) > z^{(i)}$ if $b_i = 1$

x_1	$z^{(1)}$	0	
x_2	$z^{(2)}$	0	
x_3	$z^{(3)}$	1	
x_4	$z^{(4)}$	0	
x_5	$z^{(5)}$	1	
x_6	$z^{(6)}$	0	
x_7	$z^{(7)}$	1	
x_8	$z^{(8)}$	1	

Set $S = \{x_1, \ldots, x_m\}$

Class of functions $\mathcal{H} = \{h: \mathcal{X} \to \mathbb{R}\}$

$\boldsymbol{z} = \left(z^{(1)}, \ldots, z^{(m)}\right) \in \mathbb{R}^m$ **witnesses** the **shattering** of S by $\mathcal{H}$ if for all $\boldsymbol{b} \in \{0,1\}^m$, there exists $h \in \mathcal{H}$ such that $h(x_i) \leq z^{(i)}$ if $b_i = 0$ and $h(x_i) > z^{(i)}$ if $b_i = 1$

x_1	$h(x_1)$	$\leq$	$z^{(1)}$	0
x_2	$h(x_2)$	$\leq$	$z^{(2)}$	0
x_3	$h(x_3)$	$>$	$z^{(3)}$	1
x_4	$h(x_4)$	$\leq$	$z^{(4)}$	0
x_5	$h(x_5)$	$>$	$z^{(5)}$	1
x_6	$h(x_6)$	$\leq$	$z^{(6)}$	0
x_7	$h(x_7)$	$>$	$z^{(7)}$	1
x_8	$h(x_8)$	$>$	$z^{(8)}$	1

Set $S = \{x_1, \ldots, x_m\}$

Class of functions $\mathcal{H} = \{h: \mathcal{X} \to \mathbb{R}\}$

$\boldsymbol{z} = \left(z^{(1)}, \ldots, z^{(m)}\right) \in \mathbb{R}^m$ **witnesses** the **shattering** of S by $\mathcal{H}$ if for all $\boldsymbol{b} \in \{0,1\}^m$, there exists $h \in \mathcal{H}$ such that $h(x_i) \leq z^{(i)}$ if $b_i = 0$ and $h(x_i) > z^{(i)}$ if $b_i = 1$

The **pseudo-dimension** of $\mathcal{H}$ is the cardinality of the largest set that can be shattered by $\mathcal{H}$.

x_1	$h(x_1)$	$\leq$	$z^{(1)}$	0
x_2	$h(x_2)$	$\leq$	$z^{(2)}$	0
x_3	$h(x_3)$	$>$	$z^{(3)}$	1
x_4	$h(x_4)$	$\leq$	$z^{(4)}$	0
x_5	$h(x_5)$	$>$	$z^{(5)}$	1
x_6	$h(x_6)$	$\leq$	$z^{(6)}$	0
x_7	$h(x_7)$	$>$	$z^{(7)}$	1
x_8	$h(x_8)$	$>$	$z^{(8)}$	1

Set $S = \{x_1, \ldots, x_m\}$

Class of functions $\mathcal{H} = \{h : \mathcal{X} \to \mathbb{R}\}$

$\boldsymbol{z} = \left(z^{(1)}, \ldots, z^{(m)}\right) \in \mathbb{R}^m$ **witnesses** the **shattering** of S by $\mathcal{H}$ if for all $\boldsymbol{b} \in \{0,1\}^m$, there exists $h \in \mathcal{H}$ such that $h(x_i) \leq z^{(i)}$ if $b_i = 0$ and $h(x_i) > z^{(i)}$ if $b_i = 1$

The **pseudo-dimension** of $\mathcal{H}$ is the VC dimension of the set of "below the graph" functions $\left\{(x, z) \longmapsto \mathbf{1}_{h(x)-z>0} \mid h \in \mathcal{H}\right\}$.

Theorem [Haussler 1992]

Suppose $\mathcal{H} = \{h: \mathcal{X} \to [-U, U]\}$

For every $\epsilon, \delta > 0$, and distribution D, if $m = \tilde{O}\left(\left(\frac{U}{\epsilon}\right)^2 \cdot \mathrm{Pdim}(\mathcal{H})\right)$,
then with probability $1 - \delta$ over the draw of $S = \{x_1, \ldots, x_m\} \sim D^m$,
for all $h \in \mathcal{H}$,

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

Corollary [Haussler 1992]

Suppose $\mathcal{H} = \{h: \mathcal{X} \to [-U, U]\}$

Given a sample $S = \{x_1, \dots, x_m\}$, let $h_S = \max_{h \in \mathcal{H}} \sum_{i=1}^{m} h(x_i)$.

For all $\epsilon, \delta > 0$, $m = \tilde{O}\left(\left(\frac{U}{\epsilon}\right)^2 \text{Pdim}(\mathcal{H})\right)$ samples are sufficient to ensure that w.p. $1 - \delta$ over the draw of $S = \{x_1, \dots, x_m\} \sim D^m$,

$$\max_{h \in \mathcal{H}} \mathbb{E}_{x \sim D}[h(x)] - \mathbb{E}_{x \sim D}[h_S(x)] \leq \epsilon$$

Central question: Does the resulting auction approximately maximize revenue in expectation over the unknown distribution?

Depends on the class $\mathcal{M}$.

Let M^* be the optimal auction and let $\widehat{M} \in \mathcal{M}$ be the auction returned by the learning algorithm.

$$\mathbb{E}[\mathrm{Rev}(M^*)] - \mathbb{E}\left[\mathrm{Rev}\left(\widehat{M}\right)\right]$$
$$= \underbrace{\mathbb{E}[\mathrm{Rev}(M^*)] - \max_{M \in \mathcal{M}} \mathbb{E}[\mathrm{Rev}(M)]}_{\text{Approximation loss}} + \underbrace{\max_{M \in \mathcal{M}} \mathbb{E}[\mathrm{Rev}(M)] - \mathbb{E}\left[\mathrm{Rev}\left(\widehat{M}\right)\right]}_{\text{Estimation loss}}$$

Corollary [Haussler 1992]

Suppose $\mathcal{H} = \{h: \mathcal{X} \to [-U, U]\}$

Given a sample $S = \{x_1, \dots, x_m\}$, let $h_S = \max_{h \in \mathcal{H}} \sum_{i=1}^{m} h(x_i)$.

For all $\epsilon, \delta > 0$, $m = \tilde{O}\left(\left(\frac{U}{\epsilon}\right)^2 \text{Pdim}(\mathcal{H})\right)$ samples are sufficient to ensure that w.p. $1 - \delta$ over the draw of $S = \{x_1, \dots, x_m\} \sim D^m$,

$$\max_{h \in \mathcal{H}} \mathbb{E}_{x \sim D}[h(x)] - \mathbb{E}_{x \sim D}[h_S(x)] \leq \epsilon$$

Pseudo-dimension allows us to bound estimation loss.

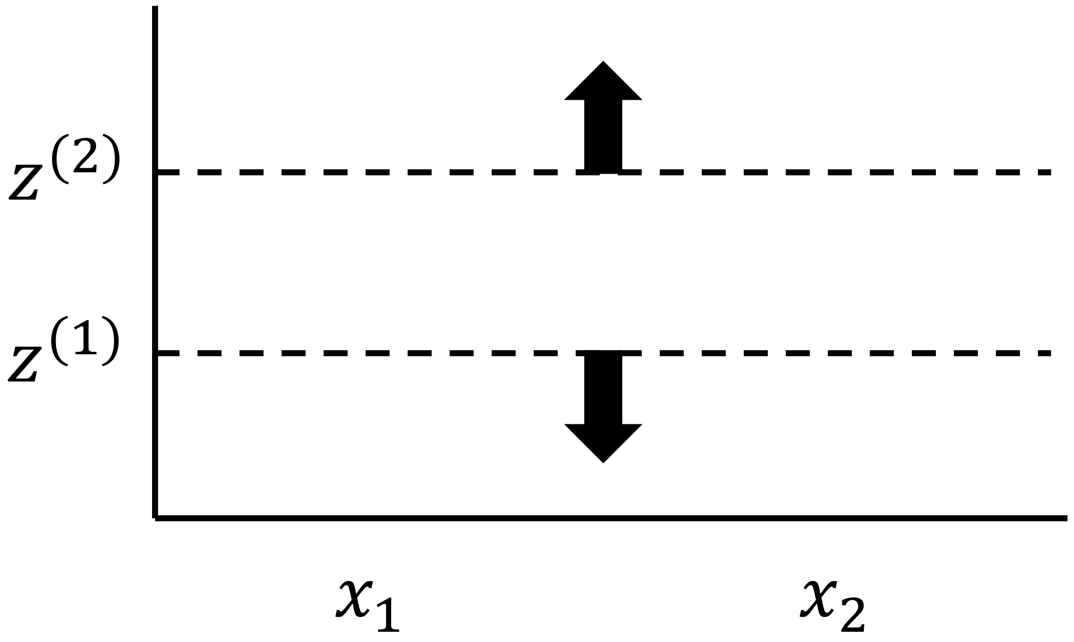
Example.

Let $\mathcal{H}$ be the set of all constant functions

$$(\mathcal{H} = \{h : \forall x \in \mathcal{X}, h(x) = c \text{ for some } c \in \mathbb{R}\})$$

Example.

For $S = \{x_1, x_2\} \subseteq \mathcal{X}$, let $z^{(1)}$ and $z^{(2)}$ be two potential witnesses. WLOG, suppose $z^{(1)} \leq z^{(2)}$. There is no constant function $h(x) = c$ where $z^{(2)} < h(x_2) = c = h(x_1) \leq z^{(1)}$. ⇒ S is **not shatterable**.



Example.

For $S = \{x_1, x_2\} \subseteq \mathcal{X}$, let $z^{(1)}$ and $z^{(2)}$ be two candidate witnesses. WLOG, suppose $z^{(1)} \leq z^{(2)}$. There is no constant function $h(x) = c$ where $z^{(2)} < h(x_2) = c = h(x_1) \leq z^{(1)}$. $\Rightarrow$ S is **not shatterable**.

Clearly, a **single point** is shatterable.

Therefore, the pseudo-dimension of $\mathcal{H}$ is 1.

1. Introduction
2. Auction design background
3. Revenue-maximizing auctions
4. Learning-theoretic auction design
5. Tools
6. Case study: Second price auctions with reserve prices
    a) Recap
    b) Result
7. Related work

Fix an auction class $\mathcal{M}$

$\Downarrow$

Look at the samples

$\Downarrow$

Optimize over $\mathcal{M}$

$\Downarrow$

Field auction

Central question: Does the resulting auction approximately maximize revenue in expectation over the unknown distribution?

Central question: Does the resulting auction approximately maximize revenue in expectation over the unknown distribution?

Depends on the class $\mathcal{M}$.

Let M^* be the optimal auction and let $\widehat{M} \in \mathcal{M}$ be the auction returned by the learning algorithm.

$$\mathbb{E}[\mathrm{Rev}(M^*)] - \mathbb{E}\left[\mathrm{Rev}\left(\widehat{M}\right)\right]$$
$$= \underbrace{\mathbb{E}[\mathrm{Rev}(M^*)] - \max_{M \in \mathcal{M}} \mathbb{E}[\mathrm{Rev}(M)]}_{\text{Approximation loss}} + \underbrace{\max_{M \in \mathcal{M}} \mathbb{E}[\mathrm{Rev}(M)] - \mathbb{E}\left[\mathrm{Rev}\left(\widehat{M}\right)\right]}_{\text{Estimation loss}}$$

Suppose $\mathcal{M}$ is the class of second price auctions with reserve prices.

Myerson's optimal auction for i.i.d. bidders
Let $\phi(x) = x - \frac{1-F(x)}{f(x)}$. Run a second price auction with a reserve price of $\phi^{-1}(0)$.

$\mathcal{M}$ has **zero approximation loss** when the bidders valuations are i.i.d.!

What is the estimation loss of $\mathcal{M}$?

1. Introduction
2. Auction design background
3. Revenue-maximizing auctions
4. Learning-theoretic auction design
5. Tools
6. Case study: Second price auctions with reserve prices
    a) Recap
    b) Result
7. Related work

Notation: $\boldsymbol{v} = (v_1, \ldots, v_n) \sim D$, where v_i is the value of bidder i

Let $v_{(1)} = \max\{v_1, \ldots, v_n\}$ and $v_{(2)} = \max\left\{\{v_1, \ldots, v_n\} \setminus v_{(1)}\right\}$

For a reserve price r, $\mathrm{rev}_r(\boldsymbol{v}) = \max\{r, v_{(2)}\} \cdot \mathbf{1}\left(v_{(1)} \geq r\right)$.

Define the class $\mathcal{H} = \{\mathrm{rev}_r(\boldsymbol{v}) : r \geq 0\}$.

We've seen that the **approximation loss** for learning over $\mathcal{H}$ is zero when the bidders have i.i.d. valuations.

Notation: $\boldsymbol{v} = (v_1, \dots, v_n) \sim D$, where v_i is the value of bidder i

Let $v_{(1)} = \max\{v_1, \dots, v_n\}$ and $v_{(2)} = \max\left\{\{v_1, \dots, v_n\} \setminus v_{(1)}\right\}$

For a reserve price r, $\mathrm{rev}_r(\boldsymbol{v}) = \max\{r, v_{(2)}\} \cdot \mathbf{1}\left(v_{(1)} \geq r\right)$.

Define the class $\mathcal{H} = \{\mathrm{rev}_r(\boldsymbol{v}) : r \geq 0\}$.

What is the **estimation loss** for learning over $\mathcal{H}$?

To answer this, we'll bound the **pseudo-dimension** of $\mathcal{H}$.

Theorem*
The pseudo-dimension of the class $\mathcal{H} = \{\mathrm{rev}_r(\boldsymbol{v}) : r \geq 0\}$ is 2.

*Special case of a theorem by Morgenstern and Roughgarden [2016]

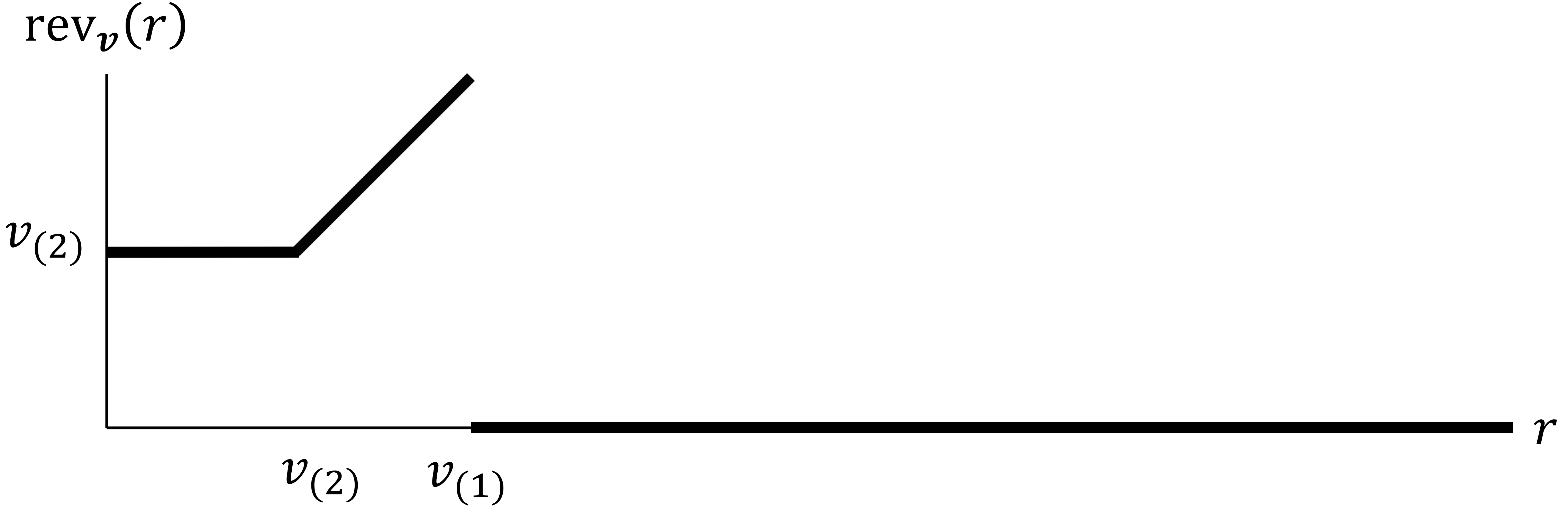
Proof.

Suppose $S = \{\boldsymbol{v}^{(1)}, \dots, \boldsymbol{v}^{(m)}\}$ is shatterable by $\mathcal{H} = \{\mathrm{rev}_r(\boldsymbol{v}) : r \geq 0\}$.

There is a **set of witnesses** $z^{(1)}, \dots, z^{(m)}$ such that for every $\boldsymbol{b} \in \{0,1\}^m$, there exists $r_{\boldsymbol{b}} \geq 0$ such that $\mathrm{rev}_{r_{\boldsymbol{b}}}(\boldsymbol{v}^{(i)}) \leq z^{(i)}$ if $b_i = 0$ and $\mathrm{rev}_{r_{\boldsymbol{b}}}(\boldsymbol{v}^{(i)}) > z^{(i)}$ if $b_i = 1$.

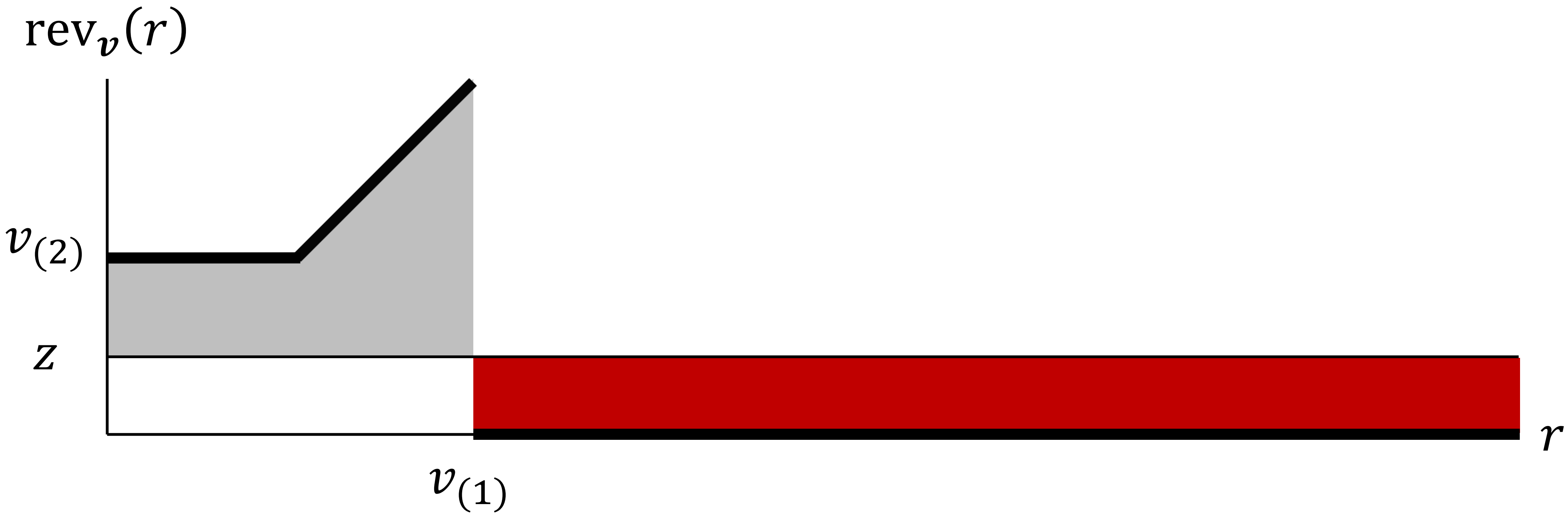
Proof.

Let $\boldsymbol{v} \in S$ and let z be its witness. Write $\mathrm{rev}_{\boldsymbol{v}}(r) = \mathrm{rev}_r(\boldsymbol{v})$.



Proof. Suppose $z < v_{(2)}$.

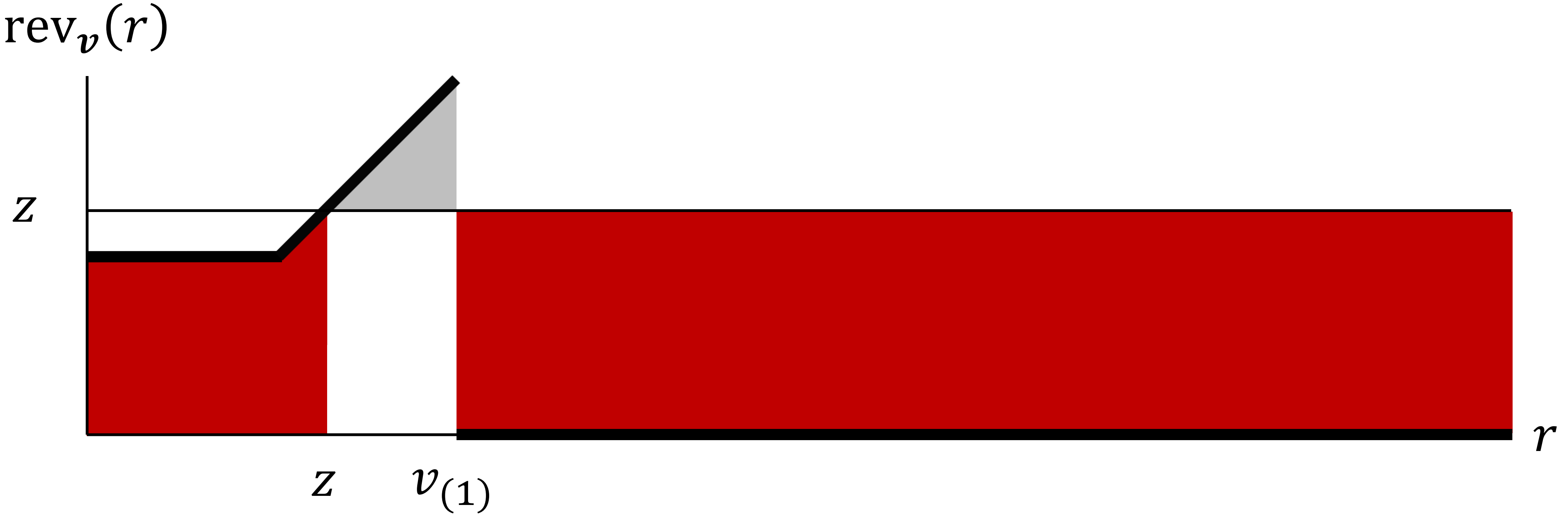
When $r \leq v_{(1)}$, $\mathrm{rev}_{\boldsymbol{v}}(r) > z$ and when $r > v_{(1)}$, $\mathrm{rev}_{\boldsymbol{v}}(r) \leq z$.



Proof. Suppose $z \geq v_{(2)}$.

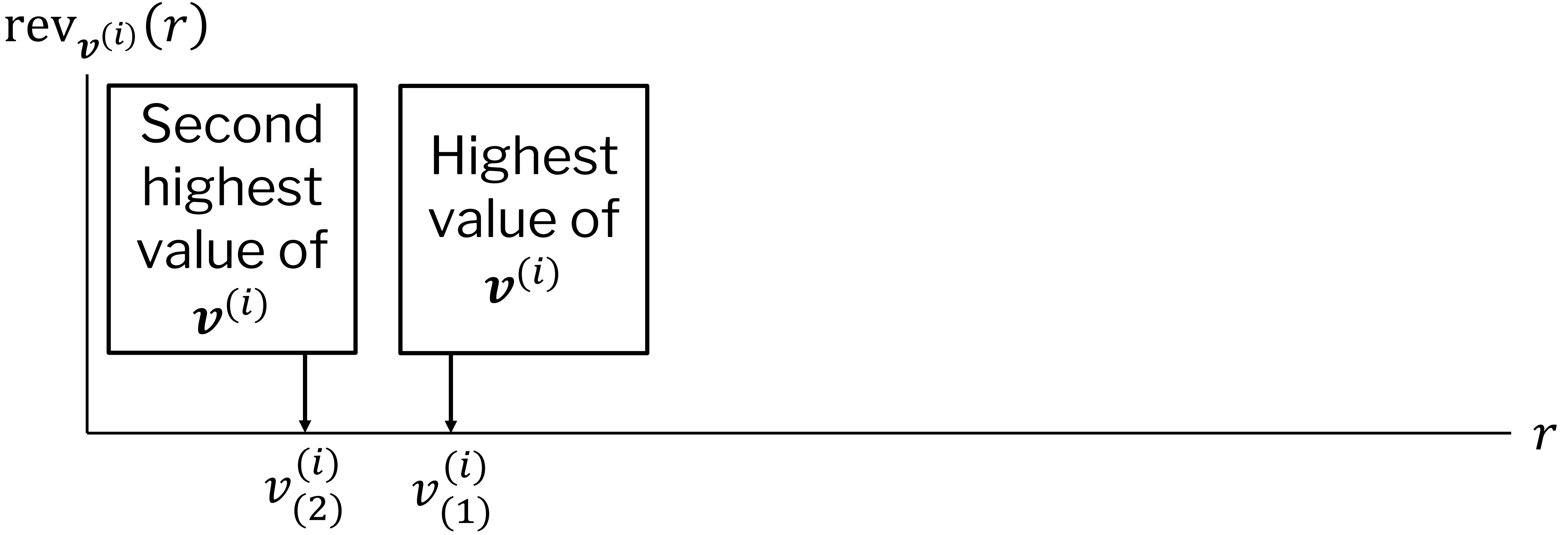
When $r \in [0, z]$ or $r \in \left[v_{(1)}, \infty\right]$, $\operatorname{rev}_{\boldsymbol{v}}(r) \leq z$.

When $r \in \left[z, v_{(1)}\right]$, $\operatorname{rev}_{\boldsymbol{v}}(r) > z$.



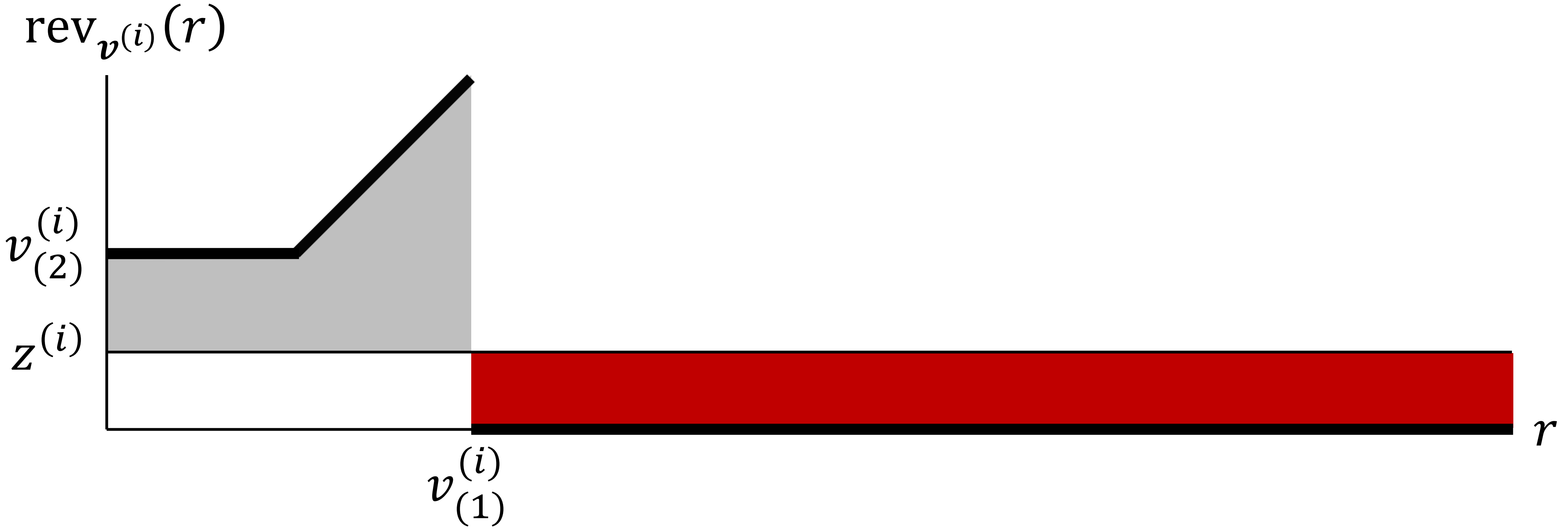
Proof.

Notation: Let $\boldsymbol{v}^{(i)} \in S$ and let $z^{(i)}$ be its witness.



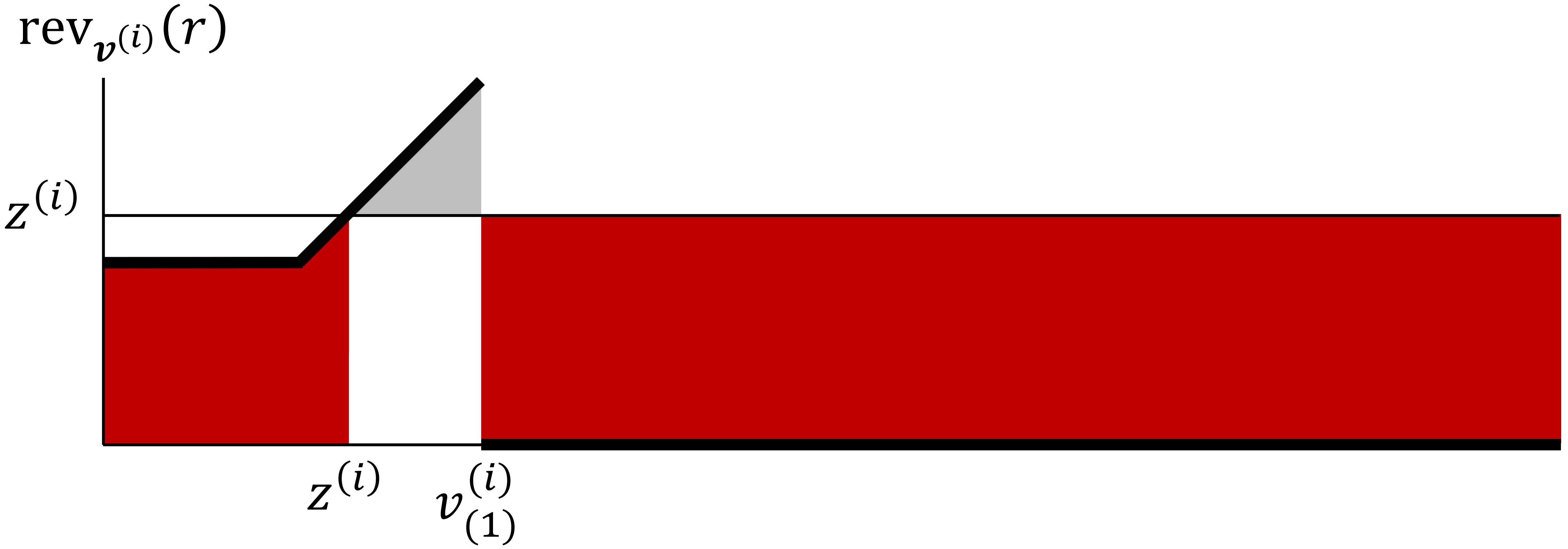
Proof.

Sort $\left\{z^{(1)}, v_{(1)}^{(1)}, \ldots, z^{(m)}, v_{(1)}^{(m)}\right\}$. For all $\boldsymbol{v}^{(i)} \in S$, between these sorted values, either $\mathrm{rev}_{\boldsymbol{v}^{(i)}}(r) > z^{(i)}$ or $\mathrm{rev}_{\boldsymbol{v}^{(i)}}(r) \leq z^{(i)}$.



Proof.

Sort $\left\{z^{(1)}, v_{(1)}^{(1)}, \ldots, z^{(m)}, v_{(1)}^{(m)}\right\}$. For all $\boldsymbol{v}^{(i)} \in S$, between these sorted values, either $\mathrm{rev}_{\boldsymbol{v}^{(i)}}(r) > z^{(i)}$ or $\mathrm{rev}_{\boldsymbol{v}^{(i)}}(r) \leq z^{(i)}$.



Proof.

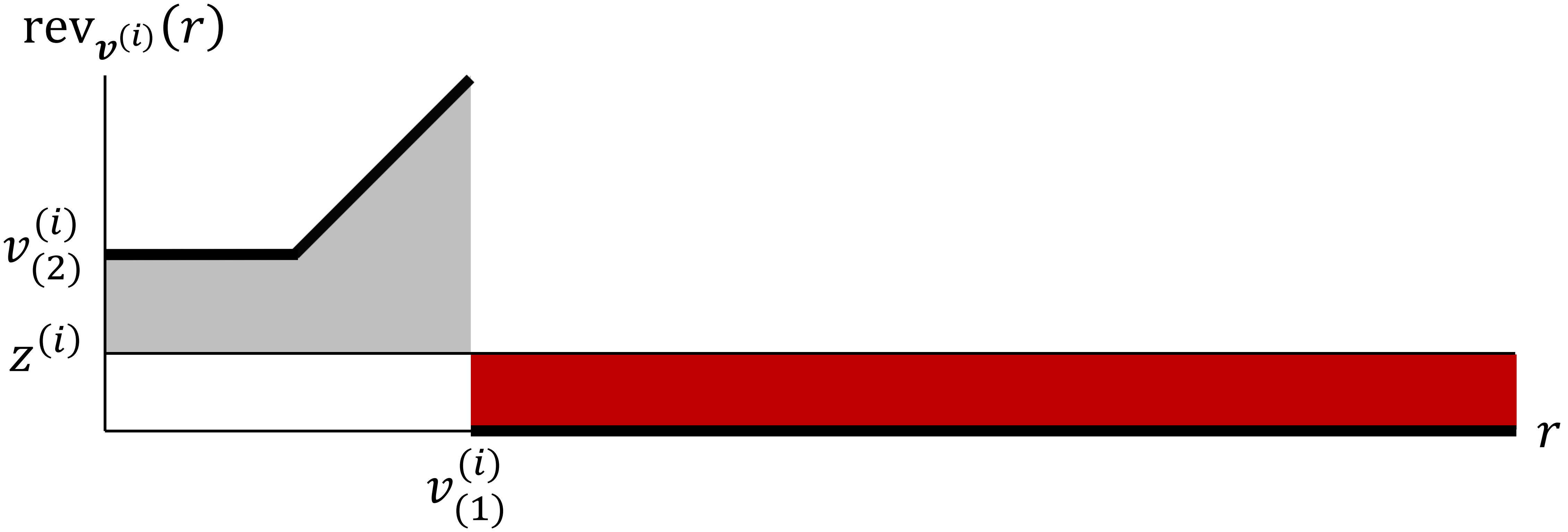
Recall:

Suppose $S = \{\boldsymbol{v}^{(1)}, \dots, \boldsymbol{v}^{(m)}\}$ is shatterable by $\mathcal{H} = \{\mathrm{rev}_r(\boldsymbol{v}) : r \geq 0\}$.

There is a set of witnesses $z^{(1)}, \dots, z^{(m)}$ such that for every $\boldsymbol{b} \in \{0,1\}^m$, there exists $r_{\boldsymbol{b}} \geq 0$ such that $\mathrm{rev}_{r_{\boldsymbol{b}}}\left(\boldsymbol{v}^{(i)}\right) \leq z^{(i)}$ if $b_i = 0$ and $\mathrm{rev}_{r_{\boldsymbol{b}}}\left(\boldsymbol{v}^{(i)}\right) > z^{(i)}$ if $b_i = 1$. Let $R^* = \{r_{\boldsymbol{b}} : \boldsymbol{b} \in \{0,1\}^m\}$.

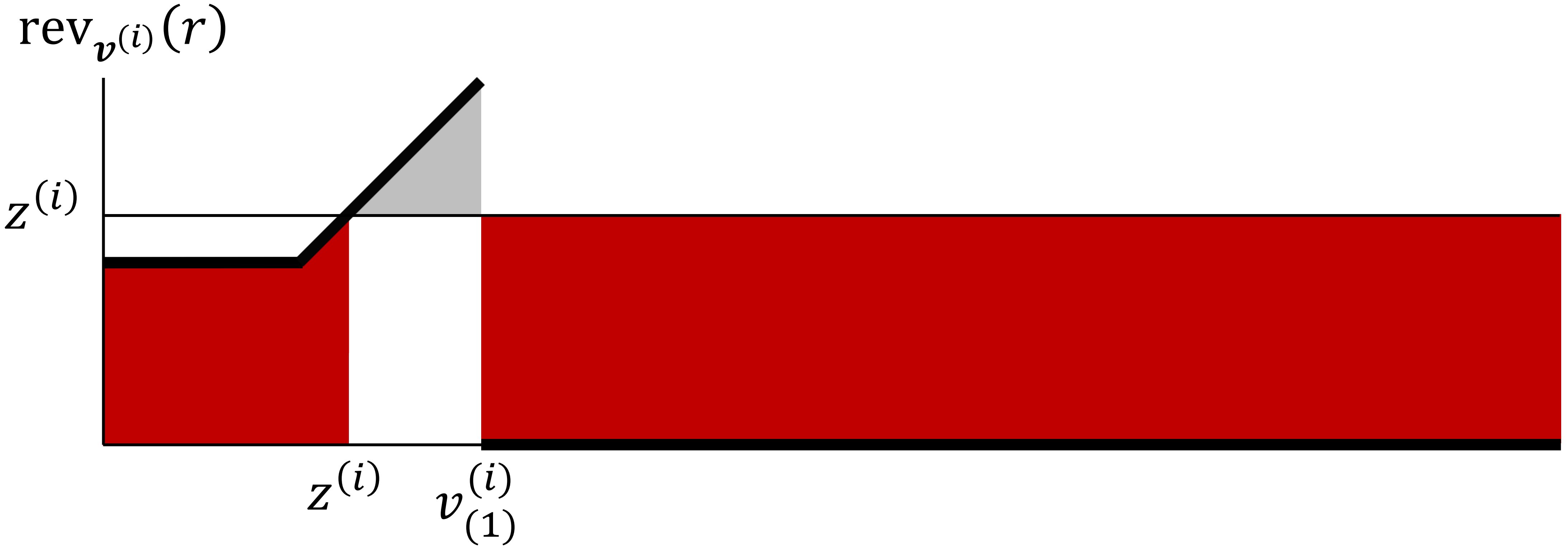
Proof.

Sort $\left\{z^{(1)}, v_{(1)}^{(1)}, \ldots, z^{(m)}, v_{(1)}^{(m)}\right\}$. For all $\boldsymbol{v}^{(i)} \in S$, between these sorted values, either $\mathrm{rev}_{\boldsymbol{v}^{(i)}}(r) > z^{(i)}$ or $\mathrm{rev}_{\boldsymbol{v}^{(i)}}(r) \leq z^{(i)}$.



Proof.

Sort $\left\{z^{(1)}, v_{(1)}^{(1)}, \ldots, z^{(m)}, v_{(1)}^{(m)}\right\}$. For all $\boldsymbol{v}^{(i)} \in S$, between these sorted values, either $\mathrm{rev}_{\boldsymbol{v}^{(i)}}(r) > z^{(i)}$ or $\mathrm{rev}_{\boldsymbol{v}^{(i)}}(r) \leq z^{(i)}$.



Proof.

Sort $\left\{z^{(1)}, v_{(1)}^{(1)}, \ldots, z^{(m)}, v_{(1)}^{(m)}\right\}$. For all $\boldsymbol{v}^{(i)} \in S$, between these sorted values, either $\operatorname{rev}_{\boldsymbol{v}^{(i)}}(r) > z^{(i)}$ or $\operatorname{rev}_{\boldsymbol{v}^{(i)}}(r) \leq z^{(i)}$.

At most one $r \in R^*$ is from each interval between sorted values.

Therefore, $2^m = |R^*| \leq 2m + 1$, so $m \leq 2$.

Theorem
The pseudo-dimension of the class $\mathcal{H} = \{\operatorname{rev}_r(\boldsymbol{v}) : r \geq 0\}$ is 2.

Corollary

Suppose the bidders' valuations are i.i.d with support in [0,1].
Let $\mathcal{H}$ be the class of second price auctions with reserves.

Approximation loss: 0

Estimation loss: $\tilde{O}\left(\frac{1}{|S|^2}\right)$

Corollary

Suppose the bidders' valuations are i.i.d with support in [0,1]. Let $\mathcal{H}$ be the class of second price auctions with reserves.

For a sample $S = \{\boldsymbol{v}^{(1)}, \dots, \boldsymbol{v}^{(m)}\} \sim D$, let r_S be the reserve price that maximizes $\sum_{i=1}^{m} \mathrm{rev}_r\left(\boldsymbol{v}^{(i)}\right)$.

$\tilde{O}\left(\frac{1}{\varepsilon^2}\right)$ samples are sufficient to ensure that with probability at least $1 - \delta$, $\mathbb{E}_{\boldsymbol{v}\sim D}\left[\mathrm{rev}_{r_S}(\boldsymbol{v})\right] \geq \max_{r\geq 0} \mathbb{E}_{\boldsymbol{v}\sim D}[\mathrm{rev}_r(\boldsymbol{v})] - \varepsilon$.

This is tight [Cesa-Bianchi et al. 2015]!

1. Introduction
2. Auction design background
3. Revenue-maximizing auctions
4. Learning-theoretic auction design
5. Tools
6. Case study: Second price auctions with reserve prices
7. Related work

Other single-item auction batch learning papers:

Elkind ’07, Dhangwatnotai et al. ’10, Medina and Mohri ‘14, Cole and Roughgarden ’14, Huang et al. ’15, Morgenstern and Roughgarden ’15, Roughgarden and Schrijvers ’16, Devanur et al. ’16, Hartline and Taggart ’17, Alon et al. ’17, Gonczarowski and Nisan ’17

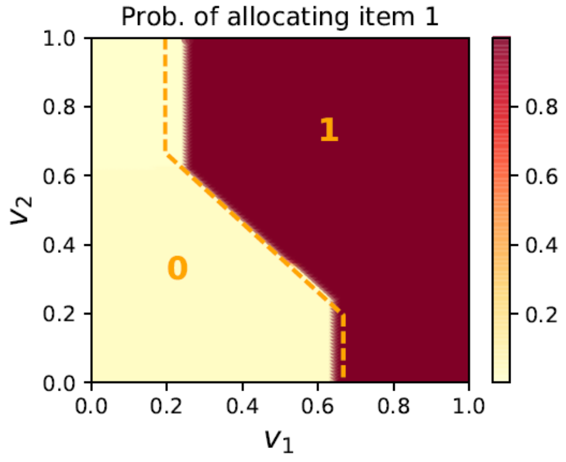
Multi-item auction batch learning:

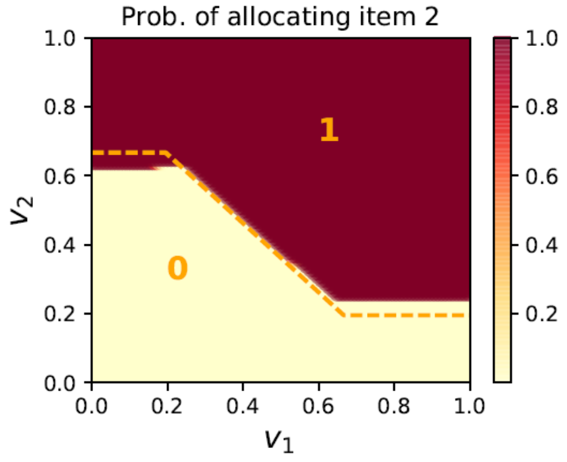
Likhodedov and Sandholm ‘04, ’05, Morgenstern and Roughgarden ‘16, Balcan, Sandholm, and **V** ‘16, ’17, Syrgkanis ’17, Cai and Daskalakis ’17, Medina and Vassilvitskii ’17

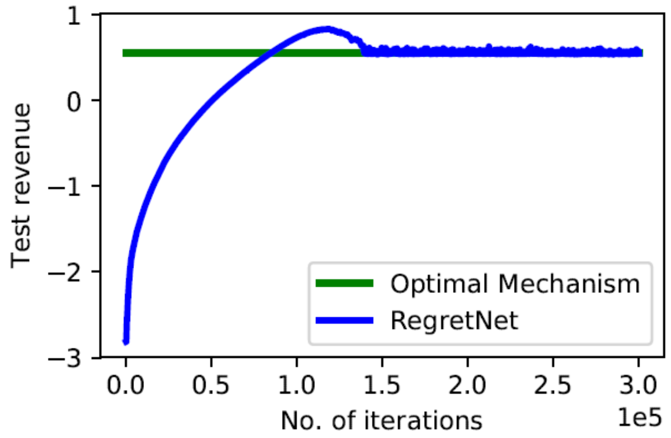
Online auction design:

Kleinberg and Leighton ‘03, Blum and Hartline ‘05, Cesa-Bianchi et al. ‘15, Dudík et al. ‘17, Balcan, Dick and **V** ‘17

Auction design via deep learning: Dütting et al. '17







Solid regions represent allocation probability learned by neural network when there's a single bidder with $v_1, v_2 \sim U[0,1]$. The optimal mechanism of Manelli and Vincent ['06] is described by the regions separated by the dashed orange lines.