Multi-Item Mechanisms: Revenue Maximization from Samples

Yannai A. Gonczarowski

Harvard

SLMath (MSRI) Summer School
June 22–23, 2023

Based upon (but all typos are my own):

Learning Simple Auctions, Jamie Morgenstern and Tim Roughgarden, 2016

The Sample Complexity of Up-to-ε Multi-Dimensional Revenue Maximization, Y.A.G. and S. Matthew Weinberg, 2021

Revenue Maximization with Samples as Input

- With the advent of internet-scale marketplaces, from Google/Bing/Facebook ad auctions through Amazon, now is a time for microeconomic theory to shine brightly!

Revenue Maximization with Samples as Input

- With the advent of internet-scale marketplaces, from Google/Bing/Facebook ad auctions through Amazon, now is a time for microeconomic theory to shine brightly!
- E.g., **ad auctions**: literature tells us how to maximize revenue (in some settings), if we **fully know** the market distribution.

Revenue Maximization with Samples as Input

- With the advent of internet-scale marketplaces, from Google/Bing/Facebook ad auctions through Amazon, now is a time for microeconomic theory to shine brightly!
- E.g., **ad auctions**: literature tells us how to maximize revenue (in some settings), if we **fully know** the market distribution.
- Even Google/Facebook/. . . , even with lots of past data and structural econ/io, do not know the market distribution, but at best only see many **samples** from it.

Revenue Maximization with Samples as Input

- With the advent of internet-scale marketplaces, from Google/Bing/Facebook ad auctions through Amazon, now is a time for microeconomic theory to shine brightly!
- E.g., **ad auctions**: literature tells us how to maximize revenue (in some settings), if we **fully know** the market distribution.
- Even Google/Facebook/. . . , even with lots of past data and structural econ/io, do not know the market distribution, but at best only see many **samples** from it.
- **Can a seller make do with an (internet-scale but) reasonable amount of samples?**

Revenue Maximization with Samples as Input

- With the advent of internet-scale marketplaces, from Google/Bing/Facebook ad auctions through Amazon, now is a time for microeconomic theory to shine brightly!
- E.g., **ad auctions**: literature tells us how to maximize revenue (in some settings), if we **fully know** the market distribution.
- Even Google/Facebook/. . . , even with lots of past data and structural econ/io, do not know the market distribution, but at best only see many **samples** from it.
- **Can a seller make do with an (internet-scale but) reasonable amount of samples?**
- Turn to statistical/machine learning for modeling inspiration.

Revenue Maximization with Samples as Input

- With the advent of internet-scale marketplaces, from Google/Bing/Facebook ad auctions through Amazon, now is a time for microeconomic theory to shine brightly!
- E.g., **ad auctions**: literature tells us how to maximize revenue (in some settings), if we **fully know** the market distribution.
- Even Google/Facebook/. . . , even with lots of past data and structural econ/io, do not know the market distribution, but at best only see many **samples** from it.
- **Can a seller make do with an (internet-scale but) reasonable amount of samples?**
- Turn to statistical/machine learning for modeling inspiration.
- **Sample complexity**: the number of samples required for learning a "good enough" auction.

Revenue Maximization with Samples as Input

- With the advent of internet-scale marketplaces, from Google/Bing/Facebook ad auctions through Amazon, now is a time for microeconomic theory to shine brightly!
- E.g., **ad auctions**: literature tells us how to maximize revenue (in some settings), if we **fully know** the market distribution.
- Even Google/Facebook/..., even with lots of past data and structural econ/io, do not know the market distribution, but at best only see many **samples** from it.
- **Can a seller make do with an (internet-scale but) reasonable amount of samples?**
- Turn to statistical/machine learning for modeling inspiration.
- **Sample complexity**: the number of samples required for learning a "good enough" auction.
  - CS: "polynomially" = many *vs.* "exponentially" = too many

Setting

- Standard setup: one **seller**, n **items** for sale, m (potential) **buyers**.
- Buyer i's **valuation** for item j is **independently** drawn from some distribution V_{ij} supported on $[0, H]$.
- Each buyer's valuation is **additive** across items.
- Seller wishes to find an **auction mechanism** that would yield good **revenue in expectation** over $\bigtimes_{i,j} V_{ij}$.

Setting

- Standard setup: one **seller**, n **items** for sale, m (potential) **buyers**.
- Buyer i's **valuation** for item j is **independently** drawn from some distribution V_{ij} supported on $[0, H]$.
- Each buyer's valuation is **additive** across items.
- Seller wishes to find an **auction mechanism** that would yield good **revenue in expectation** over $\bigtimes_{i,j} V_{ij}$.
    - Individually Rational (IR): $\forall i, \forall$valuations $v_{i1}, \ldots, v_{im}$, if i bids v_i, then : $v_i(\text{outcome}) - \text{payment}_i \geq 0$.

Setting

- Standard setup: one **seller**, n **items** for sale, m (potential) **buyers**.
- Buyer i's **valuation** for item j is **independently** drawn from some distribution V_{ij} supported on $[0, H]$.
- Each buyer's valuation is **additive** across items.
- Seller wishes to find an **auction mechanism** that would yield good **revenue in expectation** over $\bigtimes_{i,j} V_{ij}$.
  - Individually Rational (IR):
    $\forall i, \forall$valuations $v_{i1}, \ldots, v_{im}$, if i bids v_i, then : $v_i(\text{outcome}) - \text{payment}_i \geq 0$.
  - Bayesian Incentive Compatible (BIC):

$$\forall i, v_i, v_i' : \quad \mathbb{E}_{v_{-i} \sim V_{-i}} \left[v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i \right] \geq \mathbb{E}_{v_{-i} \sim V_{-i}} \left[v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i' \right].$$

Setting

- Standard setup: one **seller**, n **items** for sale, m (potential) **buyers**.
- Buyer i's **valuation** for item j is **independently** drawn from some distribution V_{ij} supported on $[0, H]$.
- Each buyer's valuation is **additive** across items.
- Seller wishes to find an **auction mechanism** that would yield good **revenue in expectation** over $\bigtimes_{i,j} V_{ij}$.
  - Individually Rational (IR):
    $\forall i, \forall$valuations $v_{i1}, \ldots, v_{im}$, if i bids v_i, then : $v_i(\text{outcome}) - \text{payment}_i \geq 0$.
  - Bayesian Incentive Compatible (BIC):
    $$\forall i, v_i, v_i' : \quad \mathbb{E}_{v_{-i} \sim V_{-i}} \left[v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i \right] \geq \mathbb{E}_{v_{-i} \sim V_{-i}} \left[v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i' \right].$$

Two standard settings:

- **Bayesian** revenue maximization: seller has **complete knowledge** of $\bigtimes V_{ij}$.

Setting

- Standard setup: one **seller**, n **items** for sale, m (potential) **buyers**.
- Buyer i's **valuation** for item j is **independently** drawn from some distribution V_{ij} supported on $[0, H]$.
- Each buyer's valuation is **additive** across items.
- Seller wishes to find an **auction mechanism** that would yield good **revenue in expectation** over $\bigtimes_{i,j} V_{ij}$.
  - Individually Rational (IR):
    $\forall i, \forall$valuations $v_{i1}, \ldots, v_{im}$, if i bids v_i, then : $v_i(\text{outcome}) - \text{payment}_i \geq 0$.
  - Bayesian Incentive Compatible (BIC):
    $$\forall i, v_i, v_i' : \quad \mathbb{E}_{v_{-i} \sim V_{-i}} \left[v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i \right] \geq \mathbb{E}_{v_{-i} \sim V_{-i}} \left[v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i' \right].$$

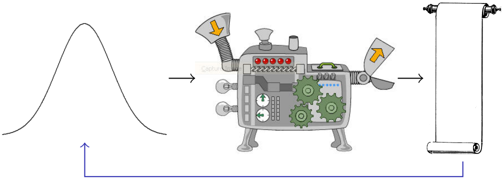
Two standard settings:

- **Bayesian** revenue maximization: seller has **complete knowledge** of $\bigtimes V_{ij}$.
- Revenue maximization **from samples**: seller has access to **polynomially many samples** from $\bigtimes V_{ij}$. ("PAC learning-like.")
  - Benchmark remains the optimal auction given $\bigtimes V_{ij}$.

Bayesian Auction (Mechanism) Design

- The seller is given
  a distribution from which the buyers' types (item valuations) are drawn, but does not know
  the realizations.

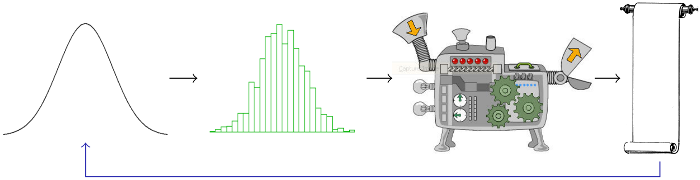
- The goal: find a **truthful** auction that **maximizes** the revenue of the seller, **in expectation** over this distribution.



~~Bayesian~~ Empirical Auction (Mechanism) Design

- The seller is given **polynomially** many samples from a distribution from which the buyers' types (item valuations) are drawn, but does not know the distribution or the realizations.

- The goal: find a **truthful** auction that **maximizes*** the revenue of the seller, **in expectation** over this **(unknown)** distribution.

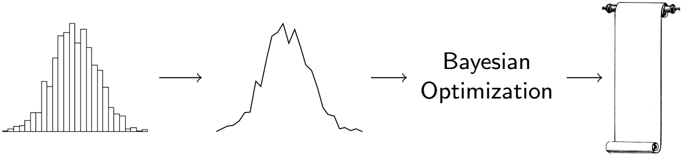


*Up to an additive ε, with high probability (**PAC learning**-like).

The Challenge in Empirical Auction Design

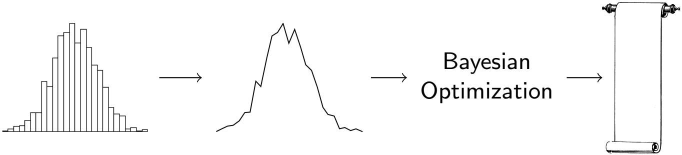
- Single-item Bayesian optimization is completely solved. (Myerson 1981)

The Challenge in Empirical Auction Design



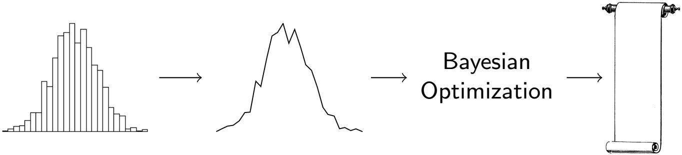
- Single-item Bayesian optimization is completely solved. (Myerson 1981)
- The naïve approach (at least when there is just one item): perform Bayesian optimization with respect to the empirical distribution over the samples.

The Challenge in Empirical Auction Design



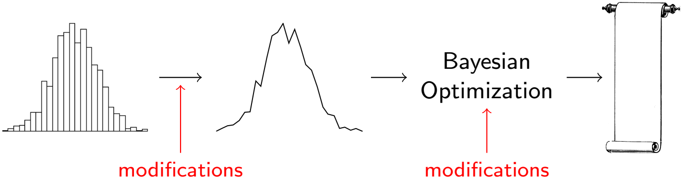
- Single-item Bayesian optimization is completely solved. (Myerson 1981)
- The naïve approach (at least when there is just one item): perform Bayesian optimization with respect to the empirical distribution over the samples.
- The potential problem: **overfitting**. A mechanism tailored for a slightly noisy version of the true underlying distribution can conceivably perform very poorly on the true distribution.

The Challenge in Empirical Auction Design



- Single-item Bayesian optimization is completely solved. (Myerson 1981)
- The naïve approach (at least when there is just one item): perform Bayesian optimization with respect to the empirical distribution over the samples.
- The potential problem: **overfitting**. A mechanism tailored for a slightly noisy version of the true underlying distribution can conceivably perform very poorly on the true distribution.

The Challenge in Empirical Auction Design



- Single-item Bayesian optimization is completely solved. (Myerson 1981)
- The naïve approach (at least when there is just one item): perform Bayesian optimization with respect to the empirical distribution over the samples.
- The potential problem: **overfitting**. A mechanism tailored for a slightly noisy version of the true underlying distribution can conceivably perform very poorly on the true distribution.

Learning-Theoretic Dimensionality

Theorem (Morgenstern and Roughgarden, 2016)

The pseudodimension of the set of all separate-item-pricing mechanisms (when viewed as functions from valuations to revenue) is $O(n \log n)$.

Learning-Theoretic Dimensionality

Theorem (Morgenstern and Roughgarden, 2016)

The pseudodimension of the set of all separate-item-pricing mechanisms (when viewed as functions from valuations to revenue) is $O(n \log n)$.

- Babaioff, Immorlica, Lucier, and Weinberg (2014) prove that the better of optimal separate item pricing and optimal bundle pricing attains at least a $1/6$ fraction of the optimal revenue. Therefore, the above theorem guarantees that $\text{poly}(n, H, \log 1/\delta)$ samples suffice for learning an up-to-$1/7$ optimal mechanism with probability $1 - \delta$.

Learning-Theoretic Dimensionality

Theorem (Morgenstern and Roughgarden, 2016)

The pseudodimension of the set of all separate-item-pricing mechanisms (when viewed as functions from valuations to revenue) is $O(n \log n)$.

- Babaioff, Immorlica, Lucier, and Weinberg (2014) prove that the better of optimal separate item pricing and optimal bundle pricing attains at least a $1/6$ fraction of the optimal revenue. Therefore, the above theorem guarantees that $\text{poly}(n, H, \log 1/\delta)$ samples suffice for learning an up-to-$1/7$ optimal mechanism with probability $1 - \delta$.
- Fix samples $(v^{(1)}, \ldots, v^{(s)})$ and fix thresholds $t_1, \ldots, t_s$. How big can we make s while keeping $(v^{(1)}, \ldots, v^{(s)})$ shatterable?

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0)

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1, 1, 1, 1)$

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \rightarrow (1,1,0,1)$

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1)$

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1)$

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1, 1, 1, 1) \to (1, 1, 0, 1) \to (1, 0, 0, 1) \to (0, 0, 0, 1) \to (0, 0, 0, 0)$.
- At most $s + 1$ such "sold/unsold labelings" for item i

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.
- At most $s+1$ such "sold/unsold labelings" for item $i \Rightarrow$ at most $(s+1)^n$ "which item(s) are sold/unsold labelings" altogether.

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.
- At most $s+1$ such "sold/unsold labelings" for item $i \Rightarrow$ at most $(s+1)^n$ "which item(s) are sold/unsold labelings" altogether.
- For a fixed "which item(s) are sold/unsold labeling" with sold/unsold labels $(\ell_i^1, \ldots, \ell_i^s)$ for each item i, the revenue from sample (valuations) $v^{(k)}$, for prices so labeled, is $\sum_{i=1}^n \ell_i^k p_i$.

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.
- At most $s+1$ such "sold/unsold labelings" for item i $\Rightarrow$ at most $(s+1)^n$ "which item(s) are sold/unsold labelings" altogether.
- For a fixed "which item(s) are sold/unsold labeling" with sold/unsold labels $(\ell_i^1, \ldots, \ell_i^s)$ for each item i, the revenue from sample (valuations) $v^{(k)}$, for prices so labeled, is $\sum_{i=1}^n \ell_i^k p_i$.
- Linear function of the prices (for each fixed "which item(s) are sold/unsold labeling") $\Rightarrow$ pseudodimension $n+1$.

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.
- At most $s+1$ such "sold/unsold labelings" for item i $\Rightarrow$ at most $(s+1)^n$ "which item(s) are sold/unsold labelings" altogether.
- For a fixed "which item(s) are sold/unsold labeling" with sold/unsold labels $(\ell_i^1,\ldots,\ell_i^s)$ for each item i, the revenue from sample (valuations) $v^{(k)}$, for prices so labeled, is $\sum_{i=1}^n \ell_i^k p_i$.
- Linear function of the prices (for each fixed "which item(s) are sold/unsold labeling") $\Rightarrow$ pseudodimension $n+1$.
  $\Rightarrow$ Even if all prices were so labeled, by the Perles–Sauer–Shelah Lemma, the overall number of distinct "which sample(s) have high/low revenue labelings" would be at most s^{n+1}.

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.
- At most $s+1$ such "sold/unsold labelings" for item $i \Rightarrow$ at most $(s+1)^n$ "which item(s) are sold/unsold labelings" altogether.
- For a fixed "which item(s) are sold/unsold labeling" with sold/unsold labels $(\ell_i^1, \ldots, \ell_i^s)$ for each item i, the revenue from sample (valuations) $v^{(k)}$, for prices so labeled, is $\sum_{i=1}^n \ell_i^k p_i$.
- Linear function of the prices (for each fixed "which item(s) are sold/unsold labeling") $\Rightarrow$ pseudodimension $n+1$.
  $\Rightarrow$ Even if all prices were so labeled, by the Perles–Sauer–Shelah Lemma, the overall number of distinct "which sample(s) have high/low revenue labelings" would be at most s^{n+1}.
- Overall number of distinct "which sample(s) have high/low revenue labelings" is at most $(s+1)^n \cdot s^{n+1}$.

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.
- At most $s+1$ such "sold/unsold labelings" for item i $\Rightarrow$ at most $(s+1)^n$ "which item(s) are sold/unsold labelings" altogether.
- For a fixed "which item(s) are sold/unsold labeling" with sold/unsold labels $(\ell_i^1, \ldots, \ell_i^s)$ for each item i, the revenue from sample (valuations) $v^{(k)}$, for prices so labeled, is $\sum_{i=1}^n \ell_i^k p_i$.
- Linear function of the prices (for each fixed "which item(s) are sold/unsold labeling") $\Rightarrow$ pseudodimension $n+1$.
  $\Rightarrow$ Even if all prices were so labeled, by the Perles–Sauer–Shelah Lemma, the overall number of distinct "which sample(s) have high/low revenue labelings" would be at most s^{n+1}.
- Overall number of distinct "which sample(s) have high/low revenue labelings" is at most $(s+1)^n \cdot s^{n+1}$.
- $(v^{(1)}, \ldots, v^{(s)})$ shatterable $\Rightarrow 2^s \leq (s+1)^n \cdot s^{n+1}$

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.
- At most $s+1$ such "sold/unsold labelings" for item $i \Rightarrow$ at most $(s+1)^n$ "which item(s) are sold/unsold labelings" altogether.
- For a fixed "which item(s) are sold/unsold labeling" with sold/unsold labels $(\ell_i^1, \ldots, \ell_i^s)$ for each item i, the revenue from sample (valuations) $v^{(k)}$, for prices so labeled, is $\sum_{i=1}^n \ell_i^k p_i$.
- Linear function of the prices (for each fixed "which item(s) are sold/unsold labeling") $\Rightarrow$ pseudodimension $n+1$.
  $\Rightarrow$ Even if all prices were so labeled, by the Perles–Sauer–Shelah Lemma, the overall number of distinct "which sample(s) have high/low revenue labelings" would be at most s^{n+1}.
- Overall number of distinct "which sample(s) have high/low revenue labelings" is at most $(s+1)^n \cdot s^{n+1}$.
- $(v^{(1)}, \ldots, v^{(s)})$ shatterable $\Rightarrow 2^s \le (s+1)^n \cdot s^{n+1} \Rightarrow s = O(n \log n)$.

Bounding the Pseudodimension

- Fix some item i, gradually increase p_i from 0 to ∞. Keep track of at which samples item i gets sold (1) or doesn't get sold (0), e.g. $(1,1,1,1) \to (1,1,0,1) \to (1,0,0,1) \to (0,0,0,1) \to (0,0,0,0)$.
- At most $s+1$ such "sold/unsold labelings" for item $i \Rightarrow$ at most $(s+1)^n$ "which item(s) are sold/unsold labelings" altogether.
- For a fixed "which item(s) are sold/unsold labeling" with sold/unsold labels $(\ell_i^1, \ldots, \ell_i^s)$ for each item i, the revenue from sample (valuations) $v^{(k)}$, for prices so labeled, is $\sum_{i=1}^n \ell_i^k p_i$.
- Linear function of the prices (for each fixed "which item(s) are sold/unsold labeling") $\Rightarrow$ pseudodimension $n+1$.
  $\Rightarrow$ Even if all prices were so labeled, by the Perles–Sauer–Shelah Lemma, the overall number of distinct "which sample(s) have high/low revenue labelings" would be at most s^{n+1}.
- Overall number of distinct "which sample(s) have high/low revenue labelings" is at most $(s+1)^n \cdot s^{n+1}$.
- $(v^{(1)}, \ldots, v^{(s)})$ shatterable $\Rightarrow 2^s \leq (s+1)^n \cdot s^{n+1} \Rightarrow s = O(n \log n)$.

□

From Bayesian Design to Empirical Design

- Revenue maximization from samples ubiquitously seen as the "next step" beyond Bayesian revenue maximization:
  1. Understand the **structure** of good-revenue auctions.
  2. **Low-dimensional** set of good auctions ⇒ no overfitting.

From Bayesian Design to Empirical Design

- Revenue maximization from samples ubiquitously seen as the "next step" beyond Bayesian revenue maximization:
  1. Understand the **structure** of good-revenue auctions.
  2. **Low-dimensional** set of good auctions ⇒ no overfitting.

	Bayesian	From Samples
Single-Item (and more generally, single-parameter)	Exact revenue maximization: Myerson'81	

From Bayesian Design to Empirical Design

- Revenue maximization from samples ubiquitously seen as the "next step" beyond Bayesian revenue maximization:
  1. Understand the **structure** of good-revenue auctions.
  2. **Low-dimensional** set of good auctions ⇒ no overfitting.

	Bayesian	From Samples
Single-Item (and more generally, single-parameter)	Exact revenue maximization: Myerson'81	Up-to-ε: CR'14, MR'15, DHP'16, HT'16, RS'16, **G**N'17

From Bayesian Design to Empirical Design

- Revenue maximization from samples ubiquitously seen as the "next step" beyond Bayesian revenue maximization:
  1. Understand the **structure** of good-revenue auctions.
  2. **Low-dimensional** set of good auctions ⇒ no overfitting.

	Bayesian	From Samples
Single-Item (and more generally, single-parameter)	Exact revenue maximization: Myerson'81	Up-to-ε: CR'14, MR'15, DHP'16, HT'16, RS'16, **G**N'17
Multi-Item (and more generally, multi-parameter)	Some percentage: CHK'07, CHMS'10, CMS'15, HN'12, BILW'14, RW'15, Yao'15, CDW'16, CM'16, CZ'17, HR'19	

From Bayesian Design to Empirical Design

- Revenue maximization from samples ubiquitously seen as the "next step" beyond Bayesian revenue maximization:
  1. Understand the **structure** of good-revenue auctions.
  2. **Low-dimensional** set of good auctions ⇒ no overfitting.

	Bayesian	From Samples
Single-Item (and more generally, single-parameter)	Exact revenue maximization: Myerson'81	Up-to-ε: CR'14, MR'15, DHP'16, HT'16, RS'16, **G**N'17
Multi-Item (and more generally, multi-parameter)	Some percentage: CHK'07, CHMS'10, CMS'15, HN'12, BILW'14, RW'15, Yao'15, CDW'16, CM'16, CZ'17, HR'19	Some percentage: MR'16, BSV'16, CD'17, S'17, BSV'18

From Bayesian Design to Empirical Design

- Revenue maximization from samples ubiquitously seen as the "next step" beyond Bayesian revenue maximization:
  1. Understand the **structure** of good-revenue auctions.
  2. **Low-dimensional** set of good auctions ⇒ no overfitting.

	Bayesian	From Samples
Single-Item (and more generally, single-parameter)	Exact revenue maximization: Myerson'81	Up-to-ε: CR'14, MR'15, DHP'16, HT'16, RS'16, **G**N'17
Multi-Item (and more generally, multi-parameter)	Some percentage: CHK'07, CHMS'10, CMS'15, HN'12, BILW'14, RW'15, Yao'15, CDW'16, CM'16, CZ'17, HR'19	Some percentage: MR'16, BSV'16, CD'17, S'17, BSV'18
	Up-to-ε: ??? (even for one buyer, two items)	

From Bayesian Design to Empirical Design

- Revenue maximization from samples ubiquitously seen as the "next step" beyond Bayesian revenue maximization:
  1. Understand the **structure** of good-revenue auctions.
  2. **Low-dimensional** set of good auctions ⇒ no overfitting.

	Bayesian	From Samples
Single-Item (and more generally, single-parameter)	Exact revenue maximization: Myerson'81	Up-to-ε: CR'14, MR'15, DHP'16, HT'16, RS'16, **G**N'17
Multi-Item (and more generally, multi-parameter)	Some percentage: CHK'07, CHMS'10, CMS'15, HN'12, BILW'14, RW'15, Yao'15, CDW'16, CM'16, CZ'17, HR'19	Some percentage: MR'16, BSV'16, CD'17, S'17, BSV'18
	Up-to-ε: ??? (even for one buyer, two items)	Up-to-ε: Why bother trying when Bayesian case still open?

From Bayesian Design to Empirical Design

- Revenue maximization from samples ubiquitously seen as the "next step" beyond Bayesian revenue maximization:
  1. Understand the **structure** of good-revenue auctions.
  2. **Low-dimensional** set of good auctions ⇒ no overfitting.

	Bayesian	From Samples
Single-Item (and more generally, single-parameter)	Exact revenue maximization: Myerson'81	Up-to-ε: CR'14, MR'15, DHP'16, HT'16, RS'16, **G**N'17
Multi-Item (and more generally, multi-parameter)	Some percentage: CHK'07, CHMS'10, CMS'15, HN'12, BILW'14, RW'15, Yao'15, CDW'16, CM'16, CZ'17, HR'19	Some percentage: MR'16, BSV'16, CD'17, S'17, BSV'18
	Up-to-ε: ??? (even for one buyer, two items)	Up-to-ε: Why bother trying when Bayesian case still open ✓

Sample Complexity: A **Nonparametric** Approach

Sample Complexity: A **Nonparametric** Approach

Notation

m **buyers**, n **items**,
independent **valuation** distributions supported on $[0, H]$.

Sample Complexity: A **Nonparametric** Approach

Notation

m **buyers**, n **items**,
independent **valuation** distributions supported on $[0, H]$.

Theorem (**G.** and Weinberg, 2021)

For every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and ε-BIC auction that maximizes revenue (among all such auctions) up to an additive ε is $\text{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$.

Sample Complexity: A **Nonparametric** Approach

Notation

m **buyers**, n **items**,
independent **valuation** distributions supported on $[0, H]$.

Theorem (**G.** and Weinberg, 2021)

For every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and ε-BIC auction that maximizes revenue (among all such auctions) up to an additive ε is $\mathrm{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$.

- BIC: $\forall i, v_i, v_i'$:

$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i] \geq$$
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i']$$

Sample Complexity: A **Nonparametric** Approach

Notation

m **buyers**, n **items**,
independent **valuation** distributions supported on $[0, H]$.

Theorem (**G.** and Weinberg, 2021)

For every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and ε-BIC auction that maximizes revenue (among all such auctions) up to an additive ε is $\mathrm{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$.

- ε-BIC: $\forall i, v_i, v_i'$:

$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i] \geq$$
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i'] - \varepsilon$$

Sample Complexity: A **Nonparametric** Approach

Notation

m **buyers**, n **items**,
independent **valuation** distributions supported on $[0, H]$.

Theorem (**G.** and Weinberg, 2021)

For every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and ε-BIC auction that maximizes revenue (among all such auctions) up to an additive ε is $\text{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$.

- ε-BIC: $\forall i, v_i, v_i'$:
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i] \geq$$
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i'] - \varepsilon$$
- Applies mutatis mutandis also for other incentive compatibility notions.

Sample Complexity: A **Nonparametric** Approach

Notation

m **buyers**, n **items**,
independent **valuation** distributions supported on $[0, H]$.

Theorem (**G.** and Weinberg, 2021)

For every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and ε-BIC auction that maximizes revenue (among all such auctions) up to an additive ε is $\mathrm{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$.

- ε-BIC: $\forall i, v_i, v_i'$:
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i] \geq \mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i'] - \varepsilon$$
- Applies mutatis mutandis also for other incentive compatibility notions.
- Computationally unbounded seller. (Information-theoretic result.)

Sample Complexity: A **Nonparametric** Approach

Notation

m **buyers**, n **items**,
independent **valuation** distributions supported on $[0, H]$.

Theorem (**G.** and Weinberg, 2021)

For every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and ε-BIC auction that maximizes revenue (among all such auctions) up to an additive ε is $\text{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$.

- ε-BIC: $\forall i, v_i, v_i'$:
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i] \geq$$
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i'] - \varepsilon$$
- Applies mutatis mutandis also for other incentive compatibility notions.
- Computationally unbounded seller. (Information-theoretic result.)
- Proof **assumes nothing** regarding structure/dimensionality of optimal-, or approximately optimal-, revenue auctions.

Sample Complexity: A **Nonparametric** Approach

Notation

m **buyers**, n **items**,
independent **valuation** distributions supported on $[0, H]$.

Theorem (**G.** and Weinberg, 2021)

For every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and ε-BIC auction that maximizes revenue (among all such auctions) up to an additive ε is $\text{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$.

- ε-BIC: $\forall i, v_i, v_i'$:
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i] \geq$$
$$\mathbb{E}_{v_{-i} \sim V_{-i}} [v_i(\text{outcome}) - \text{payment}_i \mid i \text{ bids } v_i'] - \varepsilon$$
- Applies mutatis mutandis also for other incentive compatibility notions.
- Computationally unbounded seller. (Information-theoretic result.)
- Proof **assumes nothing** regarding structure/dimensionality of optimal-, or approximately optimal-, revenue auctions.
- Holds even **far beyond** additive valuations.

Strengthened Results for Special Cases

Corollary (Single Buyer (Digital Goods), Many Items)

For $m = 1$ buyer (recall: also models selling digital goods), for every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and **IC** *mechanism that maximizes revenue (among all such mechanisms) up to an additive ε is* $\mathrm{poly}(n, H, 1/\varepsilon, \log 1/\delta)$.

Cf. DHN'14: fails for correlated distributions.

Strengthened Results for Special Cases

Corollary (Single Buyer (Digital Goods), Many Items)

For $m=1$ buyer (recall: also models selling digital goods), for every $\varepsilon, \delta > 0$, the sample complexity of learning, w.p. $1-\delta$, an IR and **IC** *mechanism that maximizes revenue (among all such mechanisms) up to an additive ε is* $\mathrm{poly}(n, H, 1/\varepsilon, \log 1/\delta)$.

Cf. DHN'14: fails for correlated distributions.

Corollary (Single Item, Many Buyers)

For $n=1$ item, for every $\varepsilon, \delta > 0$, the sample complexity of **efficiently** *learning, w.p. $1-\delta$, an IR and* **DSIC** *auction that maximizes revenue (among all IR and BIC/DSIC auctions) up to an additive ε is* $\mathrm{poly}(m, H, 1/\varepsilon, \log 1/\delta)$.

Cf. parametric approaches: even in "Myersonian" settings, generalizes slightly beyond previous "top-right table cell" results.

What Drives the Results of the Paper

- Assuming access to an **oracle** that can solve the Bayesian revenue maximization problem for **explicitly given** discrete distributions (but assuming nothing about the structure of its output!), we explicitly construct our learning algorithm.

What Drives the Results of the Paper

- Assuming access to an **oracle** that can solve the Bayesian revenue maximization problem for **explicitly given** discrete distributions (but assuming nothing about the structure of its output!), we explicitly construct our learning algorithm.
- In fact, we prove a general **black-box reduction/transformation**, converting "middle-column results" (Bayesian Auction Design) to "right-column results" (Empirical Auction Design):

What Drives the Results of the Paper

- Assuming access to an **oracle** that can solve the Bayesian revenue maximization problem for **explicitly given** discrete distributions (but assuming nothing about the structure of its output!), we explicitly construct our learning algorithm.
- In fact, we prove a general **black-box reduction/transformation**, converting "middle-column results" (Bayesian Auction Design) to "right-column results" (Empirical Auction Design):

Meta Theorem (**G.** and Weinberg, 2021)

For any percentage C:

If $\exists$ *algorithm for C% revenue maximization given an explicitly specified finite distribution,*

Then $\forall\, \varepsilon, \delta > 0$, $\exists$ *"as computationally efficient" algorithm for an ε less than C% revenue maximization w.p. $1-\delta$, given* $\text{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$ *samples from the underlying (not necessarily finite) distribution.*

- Latter loses ε in IC compared to former.
- But, for a single buyer (digital goods) OR a single item: no loss in IC.

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

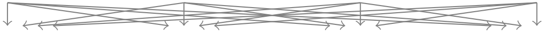
$$(w_{11}^1, \ldots, w_{mn}^1) \qquad (w_{11}^2, \ldots, w_{mn}^2) \qquad \cdots \qquad (w_{11}^S, \ldots, w_{mn}^S)$$

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

$$(w_{11}^{1}, \ldots, w_{mn}^{1}) \qquad (w_{11}^{2}, \ldots, w_{mn}^{2}) \qquad \cdots \qquad (w_{11}^{S}, \ldots, w_{mn}^{S})$$

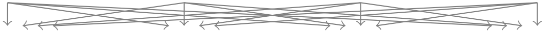


Learning-Algorithm Outline

Notation

m **buyers**, *n* **items**, *S* **samples**.

$$(w_{11}^1, \ldots, w_{mn}^1) \qquad (w_{11}^2, \ldots, w_{mn}^2) \qquad \cdots \qquad (w_{11}^S, \ldots, w_{mn}^S)$$



$$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\} \;\cdots\; W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$$

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

$(w_{11}^1, \ldots, w_{mn}^1)$ $\quad (w_{11}^2, \ldots, w_{mn}^2)$ $\quad \cdots \quad$ $(w_{11}^S, \ldots, w_{mn}^S)$

$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\} \;\cdots\; W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$

Learning-Algorithm Outline

Notation

m **buyers**, *n* **items**, *S* **samples**.

$$(w_{11}^1, \ldots, w_{mn}^1) \qquad (w_{11}^2, \ldots, w_{mn}^2) \qquad \cdots \qquad (w_{11}^S, \ldots, w_{mn}^S)$$

$$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\} \;\cdots\; W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$$

$$\bigtimes W_{ij}$$

Learning-Algorithm Outline

Notation

m **buyers**, *n* **items**, *S* **samples**.

$$(w^1_{11}, \ldots, w^1_{mn}) \qquad (w^2_{11}, \ldots, w^2_{mn}) \qquad \cdots \qquad (w^S_{11}, \ldots, w^S_{mn})$$

$$W_{11} = \mathrm{Unif}\{\lfloor w^1_{11} \rfloor_\varepsilon, \ldots, \lfloor w^S_{11} \rfloor_\varepsilon\} \ \cdots \ W_{mn} = \mathrm{Unif}\{\lfloor w^1_{mn} \rfloor_\varepsilon, \ldots, \lfloor w^S_{mn} \rfloor_\varepsilon\}$$

$$\downarrow$$

$$\bigtimes W_{ij}$$

$$\downarrow$$

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

$$(w_{11}^1, \ldots, w_{mn}^1) \quad (w_{11}^2, \ldots, w_{mn}^2) \quad \cdots \quad (w_{11}^S, \ldots, w_{mn}^S)$$

$$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\} \quad \cdots \quad W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$$

$$\bigtimes W_{ij}$$

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

$$(w^1_{11}, \ldots, w^1_{mn}) \qquad (w^2_{11}, \ldots, w^2_{mn}) \qquad \cdots \qquad (w^S_{11}, \ldots, w^S_{mn})$$

$$W_{11} = \mathrm{Unif}\{\lfloor w^1_{11} \rfloor_\varepsilon, \ldots, \lfloor w^S_{11} \rfloor_\varepsilon\} \quad \cdots \quad W_{mn} = \mathrm{Unif}\{\lfloor w^1_{mn} \rfloor_\varepsilon, \ldots, \lfloor w^S_{mn} \rfloor_\varepsilon\}$$

$$\downarrow$$

$$\bigtimes W_{ij}$$

$$\downarrow$$

$$\downarrow$$

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

$(w_{11}^1, \ldots, w_{mn}^1)$ $(w_{11}^2, \ldots, w_{mn}^2)$ $\cdots$ $(w_{11}^S, \ldots, w_{mn}^S)$

$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\}$ $\cdots$ $W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$

$\bigtimes W_{ij}$

μ

Learning-Algorithm Outline

Notation

m **buyers**, *n* **items**, *S* **samples**.

$(w_{11}^1, \ldots, w_{mn}^1)$ $(w_{11}^2, \ldots, w_{mn}^2)$ $\cdots$ $(w_{11}^S, \ldots, w_{mn}^S)$

$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\}$ $\cdots$ $W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$

$\bigtimes W_{ij}$

μ

modify auction to round its inputs to ε-grid

Learning-Algorithm Outline

Notation

m **buyers**, *n* **items**, *S* **samples**.

$(w_{11}^1, \ldots, w_{mn}^1)$ $(w_{11}^2, \ldots, w_{mn}^2)$ $\cdots$ $(w_{11}^S, \ldots, w_{mn}^S)$

$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\}$ $\cdots$ $W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$

$\downarrow$

$\bigtimes W_{ij}$

$\downarrow$

$\downarrow$

μ

modify auction to round its inputs to ε-grid

$\downarrow$

$\hat{\mu}$

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

$$(w_{11}^1, \ldots, w_{mn}^1) \quad (w_{11}^2, \ldots, w_{mn}^2) \quad \cdots \quad (w_{11}^S, \ldots, w_{mn}^S)$$

$$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\} \quad \cdots \quad W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$$

$$\times W_{ij}$$

Know **nothing** about the structure of the output of the oracle!

μ

modify auction to round its inputs to ε-grid

$\hat{\mu}$

Learning-Algorithm Outline

Notation

m **buyers**, n **items**, S **samples**.

$(w_{11}^1, \ldots, w_{mn}^1)$ $(w_{11}^2, \ldots, w_{mn}^2)$ $\cdots$ $(w_{11}^S, \ldots, w_{mn}^S)$

$W_{11} = \text{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\}$ $\cdots$ $W_{mn} = \text{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$

$\times W_{ij}$

Know **nothing** about the structure of the output of the oracle!

Optimal auctions have very high dimensionality! Why is there **no overfitting**?

μ

modify auction to round its inputs to ε-grid

$\hat{\mu}$

Learning-Algorithm Outline

Notation

m **buyers**, *n* **items**, *S* **samples**.

$(w_{11}^1, \ldots, w_{mn}^1)$ $(w_{11}^2, \ldots, w_{mn}^2)$ $\cdots$ $(w_{11}^S, \ldots, w_{mn}^S)$

$W_{11} = \mathrm{Unif}\{\lfloor w_{11}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{11}^S \rfloor_\varepsilon\}$ $\cdots$ $W_{mn} = \mathrm{Unif}\{\lfloor w_{mn}^1 \rfloor_\varepsilon, \ldots, \lfloor w_{mn}^S \rfloor_\varepsilon\}$

Deliberately introduced controlled model misspecification for statistical analysis to rule out overfitting

$\bigtimes W_{ij}$

Know **nothing** about the structure of the output of the oracle!

Optimal auctions have very high dimensionality! Why is there **no overfitting**?

μ

modify auction to round its inputs to ε-grid

$\hat{\mu}$

Algorithm Analysis

μ

$\times W_{ij}$

round inputs down

$\hat{\mu}$

Algorithm Analysis

μ $\times W_{ij}$

round inputs down

$\hat{\mu}$ $\times V_{ij}$

Algorithm Analysis

μ $\quad$ $\times W_{ij}$

round inputs down

round down to ε-grid

$\hat{\mu}$ $\quad$ $\times V_{ij}$

Algorithm Analysis

μ $\times W_{ij}$

round inputs down

$\times \lfloor V_{ij} \rfloor_\varepsilon$

round down to ε-grid

$\hat{\mu}$ $\times V_{ij}$

Algorithm Analysis

μ $\qquad$ $\times W_{ij}$

draw S samples

$\times \lfloor V_{ij} \rfloor_{\varepsilon}$

round inputs down $\qquad$ round down to ε-grid

$\hat{\mu}$ $\qquad$ $\times V_{ij}$

Algorithm Analysis

μ $\qquad$ $\times W_{ij}$

draw S samples

μ $\qquad$ $\times \lfloor V_{ij} \rfloor_{\varepsilon}$

round inputs down

round down to ε-grid

$\hat{\mu}$ $\qquad$ $\times V_{ij}$

Algorithm Analysis

μ —is BIC and optimal for→ $\times W_{ij}$

μ $\| $ μ

draw S samples

$\times \lfloor V_{ij} \rfloor_\varepsilon$

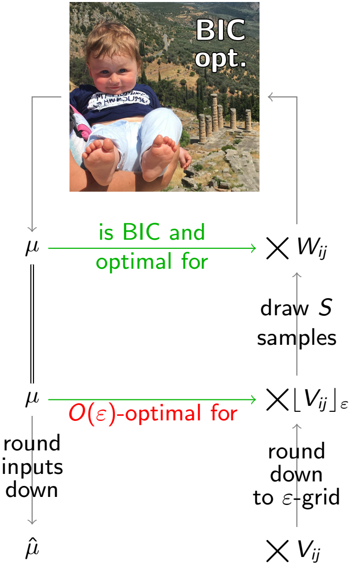
round inputs down

round down to ε-grid

$\hat{\mu}$

$\times V_{ij}$

Algorithm Analysis



Algorithm Analysis

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

μ — is BIC and optimal for → $\bigtimes W_{ij}$

draw S samples

μ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

round down to ε-grid

$\hat{\mu}$ $\qquad$ $\bigtimes V_{ij}$

Algorithm Analysis

μ —is BIC and optimal for→ $\bigtimes W_{ij}$

μ —$O(\varepsilon)$-optimal for→ $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

$\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$ —draw S samples→ $\bigtimes W_{ij}$

$\bigtimes V_{ij}$ —round down to ε-grid→ $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

μ —round inputs down→ $\hat{\mu}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.

Algorithm Analysis

μ — is BIC and optimal for → $\bigtimes W_{ij}$

μ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down: μ → $\hat{\mu}$

draw S samples: $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$ → $\bigtimes W_{ij}$

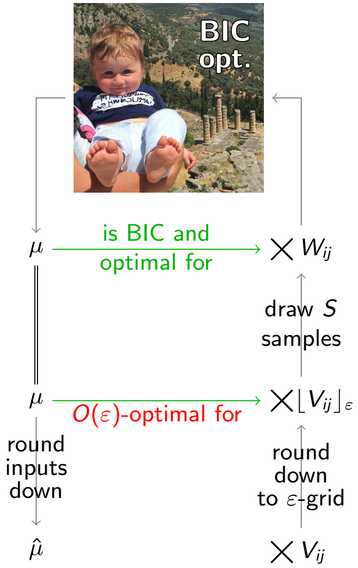
round down to ε-grid: $\bigtimes V_{ij}$ → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.

Algorithm Analysis



μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta}$

Algorithm Analysis

μ — is BIC and optimal for → $\bigtimes W_{ij}$

μ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

μ round inputs down → $\hat{\mu}$

$\bigtimes V_{ij}$ round down to ε-grid → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$ draw S samples → $\bigtimes W_{ij}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \text{poly}(m,n,{}^H/_\varepsilon) \cdot \log S + \log {}^1/_\delta$.

Algorithm Analysis

μ — is BIC and optimal for → $\bigtimes W_{ij}$

μ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down: μ → $\hat{\mu}$

draw S samples

round down to ε-grid: $\bigtimes V_{ij}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S...

Algorithm Analysis

μ is BIC and optimal for $\bigtimes W_{ij}$

draw S samples

μ $O(\varepsilon)$-optimal for $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

round down to ε-grid

$\hat{\mu}$ $\bigtimes V_{ij}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \text{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S… moderately.

Algorithm Analysis

μ is BIC and optimal for $\bigtimes W_{ij}$

draw S samples

μ $O(\varepsilon)$-optimal for $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

round down to ε-grid

$\hat{\mu}$ $\bigtimes V_{ij}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \text{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \text{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

μ is BIC and optimal for $\bigtimes W_{ij}$

draw S samples

$\lfloor \text{OPT} \rfloor_\varepsilon$ μ $O(\varepsilon)$-optimal for $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

round down to ε-grid

$\hat{\mu}$ $\bigtimes V_{ij}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \text{poly}(m, n, {}^{H}\!/\!{}_{\varepsilon}) \cdot \log S + \log {}^{1}\!/\!{}_{\delta}$.
- Grows with S... moderately.
- $S = \text{poly}(m, n, H, {}^{1}\!/\!{}_{\varepsilon}, \log {}^{1}\!/\!{}_{\delta})$ samples suffice!

Algorithm Analysis

μ is BIC and optimal for $\bigtimes W_{ij}$

draw S samples

$\lfloor \text{OPT} \rfloor_\varepsilon$ $\quad$ μ $O(\varepsilon)$-optimal for $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

round down to ε-grid

$\hat{\mu}$ $\quad$ $\bigtimes V_{ij}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \text{poly}(m,n,{}^H/_\varepsilon) \cdot \log S + \log {}^1/_\delta$.
- Grows with S... moderately.
- $S = \text{poly}(m,n,H,{}^1/_\varepsilon,\log {}^1/_\delta)$ samples suffice!

Algorithm Analysis

$\lfloor \mathrm{OPT} \rfloor_\varepsilon$ ‖ $\lfloor \mathrm{OPT} \rfloor_\varepsilon$

μ — is BIC and optimal for → $\bigtimes W_{ij}$

μ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down → $\hat{\mu}$

$\bigtimes V_{ij}$ — round down to ε-grid → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$ — draw S samples → $\bigtimes W_{ij}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

$\lfloor \text{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\bigtimes W_{ij}$

$\lfloor \text{OPT} \rfloor_\varepsilon$ μ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down: $\hat{\mu}$

draw S samples

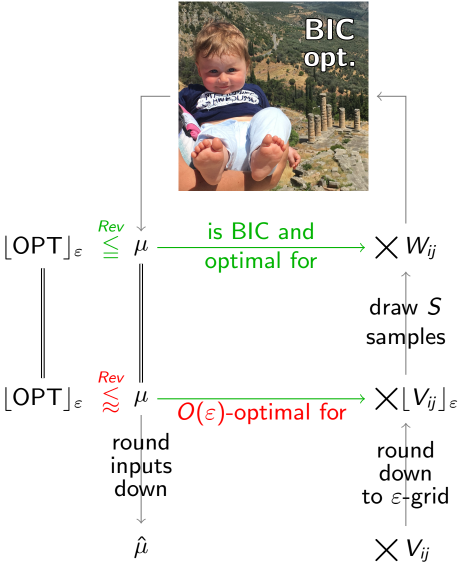
round down to ε-grid: $\bigtimes V_{ij}$

μ
attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \text{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \text{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis



μ
attains similar revenue

on $\times W_{ij}$ and $\times \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}\!/\!_{\varepsilon}) \cdot \log S + \log {}^{1}\!/\!_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}\!/\!_{\varepsilon},\log {}^{1}\!/\!_{\delta})$ samples suffice!

Algorithm Analysis

$\lfloor \text{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\bigtimes W_{ij}$

$\lfloor \text{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down

round down to ε-grid

$\hat{\mu}$

$\bigtimes V_{ij}$

μ, as well as $\lfloor \text{OPT} \rfloor_\varepsilon$, attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \text{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \text{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\times W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — $O(\varepsilon)$-optimal for → $\times \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down → $\hat{\mu}$

round down to ε-grid ← $\times V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue

on $\times W_{ij}$ and $\times \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}\!/\!_{\varepsilon}) \cdot \log S + \log {}^{1}\!/\!_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}\!/\!_{\varepsilon},\log {}^{1}\!/\!_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\bigtimes W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down → $\hat{\mu}$

round down to ε-grid ← $\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}+1$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}\!/\!_{\varepsilon}) \cdot \log S + \log {}^{1}\!/\!_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}\!/\!_{\varepsilon},\log {}^{1}\!/\!_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\bigtimes W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down → $\hat{\mu}$

round down to ε-grid ← $\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}}+1$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}+1}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}\!/\!_{\varepsilon}) \cdot \log S + \log {}^{1}\!/\!_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}\!/\!_{\varepsilon},\log {}^{1}\!/\!_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leqq} \mu$ — is BIC and optimal for → $\bigtimes W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down: $\hat{\mu}$

round down to ε-grid: $\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}+1}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}+1}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}+1}}{\delta} \approx \mathrm{poly}(m,n,H/\varepsilon) \cdot \log S + \log 1/\delta$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,1/\varepsilon,\log 1/\delta)$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leqq} \mu$ —is BIC and optimal for→ $\bigtimes W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ —$O(\varepsilon)$-optimal for→ $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down → $\hat{\mu}$

$\bigtimes V_{ij}$ — round down to ε-grid

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}+1}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}+1}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}+1}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}\!/\!_{\varepsilon}) \cdot \log S + \log {}^{1}\!/\!_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}\!/\!_{\varepsilon},\log {}^{1}\!/\!_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \text{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\times W_{ij}$

$\lfloor \text{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — $O(\varepsilon)$-optimal for → $\times \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down → $\hat{\mu}$

round down to ε-grid

$\times V_{ij}$

μ, as well as $\lfloor \text{OPT} \rfloor_\varepsilon$, attains similar revenue

on $\times W_{ij}$ and $\times \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}+1}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}+1}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}+1}}{\delta} \approx \text{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \text{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\bigtimes W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

μ — round inputs down → $\hat{\mu}$

$\bigtimes V_{ij}$ — round down to ε-grid → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$ — draw S samples → $\bigtimes W_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue

on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}+1}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}+1}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}+1}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ is BIC and optimal for $\times W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for $\times \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down: $\hat{\mu}$

round down to ε-grid: $\times V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form**

on $\times W_{ij}$ and $\times \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}+1}$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}+1}}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}+1}}{\delta} \approx \mathrm{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leqq} \mu$ is BIC and optimal for $\bigtimes W_{ij}$

draw S samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

round down to ε-grid

$\hat{\mu}$

$\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form** $\left\{\mathbb{E}_{v_{-i}}\left[v_i(o) - p_i \mid \text{bid } v_i'\right]\right\}_{i,v_i,v_i'}$ on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $(S+1)^{\frac{mnH}{\varepsilon}} + 1$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}} + 1}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}} + 1}{\delta} \approx \mathrm{poly}(m,n,{}^{H}\!/\!_{\varepsilon}) \cdot \log S + \log {}^{1}\!/\!_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}\!/\!_{\varepsilon},\log {}^{1}\!/\!_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leqq} \mu$ is BIC and optimal for $\bigtimes W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\approx} \mu$ is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down

$\hat{\mu}$

round down to ε-grid

$\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form** $\{\mathbb{E}_{v_{-i}}[v_i(o) - p_i \mid \text{bid } v_i']\}_{i,v_i,v_i'}$ on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right) \cdot \left(1+m\lceil \frac{H}{\varepsilon} \rceil^{2n}\right)$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{(S+1)^{\frac{mnH}{\varepsilon}}+1}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}+1}{\delta} \approx \mathrm{poly}(m,n,{}^{H}\!/\!_{\varepsilon}) \cdot \log S + \log {}^{1}\!/\!_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}\!/\!_{\varepsilon},\log {}^{1}\!/\!_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\times W_{ij}$

draw S samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for → $\times \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down: $\hat{\mu}$

round down to ε-grid: $\times V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form** $\left\{\mathbb{E}_{v_{-i}}\left[v_i(o) - p_i \mid \text{bid } v_i'\right]\right\}_{i,v_i,v_i'}$ on $\times W_{ij}$ and $\times \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil\frac{H}{\varepsilon}\right\rceil^{2n}\right)$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil\frac{H}{\varepsilon}\right\rceil^{2n}\right)}$.
- Need #samples $S \approx \log \frac{(S+1)^{\frac{mnH}{\varepsilon}}+1}{\delta} \approx \mathrm{poly}(m,n,{}^{H}/_{\varepsilon})\cdot\log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\times W_{ij}$

draw S samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for → $\times \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

$\hat{\mu}$

round down to ε-grid

$\times V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form** $\left\{\mathbb{E}_{v_{-i}}\left[v_i(o) - p_i \mid \text{bid } v_i'\right]\right\}_{i,v_i,v_i'}$ on $\times W_{ij}$ and $\times \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)}$.
- Need #samples $S \approx \log\frac{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)}{\delta} \approx \mathrm{poly}(m,n,{}^{H}/_{\varepsilon})\cdot\log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\bigtimes W_{ij}$

draw S samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

round down to ε-grid

$\hat{\mu}$ $\qquad$ $\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form** $\{\mathbb{E}_{v_{-i}}[v_i(o) - p_i \mid \text{bid } v_i']\}_{i,v_i,v_i'}$ on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right) \cdot \left(1+m\left\lceil \frac{H}{\varepsilon} \right\rceil^{2n}\right)$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right) \cdot \left(1+m\left\lceil \frac{H}{\varepsilon} \right\rceil^{2n}\right)}$.
- Need #samples $S \approx \log \frac{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right) \cdot \left(1+m\left\lceil \frac{H}{\varepsilon} \right\rceil^{2n}\right)}{\delta} \approx \mathrm{poly}(m,n,{}^{H}/_{\varepsilon}) \cdot \log S + \log {}^{1}/_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}/_{\varepsilon},\log {}^{1}/_{\delta})$ samples suffice!

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ — is BIC and optimal for → $\bigtimes W_{ij}$

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ — is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for → $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

draw S samples

round inputs down

round down to ε-grid

$\hat{\mu}$ — is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for → $\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form** $\left\{\mathbb{E}_{v_{-i}}\left[v_i(o) - p_i \mid \text{bid } v_i'\right]\right\}_{i,v_i,v_i'}$ on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)}$.
- Need #samples $S \approx \log\frac{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)}{\delta} \approx \mathrm{poly}(m,n,{}^H\!/\!_\varepsilon)\cdot\log S + \log {}^1\!/\!_\delta$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^1\!/\!_\varepsilon,\log {}^1\!/\!_\delta)$ samples suffice!

(A Peak) Behind the Scenes

- ε-BIC vs. BIC revenue maximization:

Theorem (Rubinstein and Weinberg, 2015; see also DW12)

Let $\mathcal{W}$ be any joint distribution over arbitrary valuations, where the valuations of different buyers are independent.
The maximum revenue attainable by any IR and ε-BIC auction for $\mathcal{W}$ is at most $2m\sqrt{nLH\varepsilon}$ greater than the maximum revenue attainable by any IR and BIC auction for $\mathcal{W}$.

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ is BIC and optimal for $\bigtimes W_{ij}$

$\|$ $\|$ draw S samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down; round down to ε-grid

$\hat{\mu}$ is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for $\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form** $\left\{ \mathbb{E}_{v_{-i}} \left[v_i(o) - p_i \mid \text{bid } v_i' \right] \right\}_{i, v_i, v_i'}$ on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $\left((S+1)^{\frac{mnH}{\varepsilon}} + 1 \right) \cdot \left(1 + m \left\lceil \frac{H}{\varepsilon} \right\rceil^{2n} \right)$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{\left((S+1)^{\frac{mnH}{\varepsilon}} + 1 \right) \cdot \left(1 + m \left\lceil \frac{H}{\varepsilon} \right\rceil^{2n} \right)}$.
- Need #samples $S \approx \log \frac{\left((S+1)^{\frac{mnH}{\varepsilon}} + 1 \right) \cdot \left(1 + m \left\lceil \frac{H}{\varepsilon} \right\rceil^{2n} \right)}{\delta} \approx \mathrm{poly}(m, n, H/\varepsilon) \cdot \log S + \log 1/\delta$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m, n, H, 1/\varepsilon, \log 1/\delta)$ samples suffice!

(A Peak) Behind the Scenes

- ε-BIC vs. BIC revenue maximization:

Theorem (Rubinstein and Weinberg, 2015; see also DW12)

Let $\mathcal{W}$ be any joint distribution over arbitrary valuations, where the valuations of different buyers are independent.
The maximum revenue attainable by any IR and ε-BIC auction for $\mathcal{W}$ is at most $2m\sqrt{nLH\varepsilon}$ greater than the maximum revenue attainable by any IR and BIC auction for $\mathcal{W}$.

- Chernoff-style concentration bound for product distributions:

Theorem (Babichenko et al., 2017; see also DHP16)

Let $V_1, \ldots, V_\ell$ be discrete distributions. Let $S \in \mathbb{N}$.
For every i, draw S independent samples from V_i, and let W_i be the uniform distribution over these samples.
For every $\varepsilon > 0$ and $f : \bigtimes_{i=1}^{\ell} \operatorname{supp} V_i \to [0, H]$, we have that

$$\Pr\Big(\big|\mathbb{E}_{\times_{i=1}^{\ell} W_i}[f] - \mathbb{E}_{\times_{i=1}^{\ell} V_i}[f]\big| > \varepsilon\Big) \leq \frac{4H}{\varepsilon} \exp\Big(-\frac{\varepsilon^2 S}{8H^2}\Big).$$

Algorithm Analysis

Unknown, yet well defined; independent of the samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\leq} \mu$ is BIC and optimal for $\bigtimes W_{ij}$

draw S samples

$\lfloor \mathrm{OPT} \rfloor_\varepsilon \overset{Rev}{\lessapprox} \mu$ is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$

round inputs down

round down to ε-grid

$\hat{\mu}$ is $O(\varepsilon)$-BIC and $O(\varepsilon)$-optimal for $\bigtimes V_{ij}$

μ, as well as $\lfloor \mathrm{OPT} \rfloor_\varepsilon$, attains similar revenue and has similar **reduced form** $\left\{\mathbb{E}_{v_{-i}}\left[v_i(o) - p_i \mid \text{bid } v_i'\right]\right\}_{i,v_i,v_i'}$ on $\bigtimes W_{ij}$ and $\bigtimes \lfloor V_{ij} \rfloor_\varepsilon$:

- Enough to take a union bound over $\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)$ random variables.
- Need convergence for a single variable w.p. $\geq 1 - \frac{\delta}{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)}$.
- Need #samples $S \approx \log \frac{\left((S+1)^{\frac{mnH}{\varepsilon}}+1\right)\cdot\left(1+m\left\lceil \frac{H}{\varepsilon}\right\rceil^{2n}\right)}{\delta} \approx \mathrm{poly}(m,n,{}^{H}\!/\!_{\varepsilon})\cdot \log S + \log {}^{1}\!/\!_{\delta}$.
- Grows with S... moderately.
- $S = \mathrm{poly}(m,n,H,{}^{1}\!/\!_{\varepsilon},\log {}^{1}\!/\!_{\delta})$ samples suffice!

Some Remarks on Model Misspecification

- One of the main proof ideas: relatively small (i.e., grows "moderately enough") number of possible oracle inputs.

Some Remarks on Model Misspecification

- One of the main proof ideas: relatively small (i.e., grows "moderately enough") number of possible oracle inputs.
    - Independent valuation distributions +
      Deliberately introduced a controlled amount of model misspecification.

Some Remarks on Model Misspecification

- One of the main proof ideas: relatively small (i.e., grows "moderately enough") number of possible oracle inputs.
  - Independent valuation distributions +
    Deliberately introduced a controlled amount of model misspecification.
  - A fairly moderate price: ε-IC.

Some Remarks on Model Misspecification

- One of the main proof ideas: relatively small (i.e., grows "moderately enough") number of possible oracle inputs.
  - Independent valuation distributions + Deliberately introduced a controlled amount of model misspecification.
  - A fairly moderate price: ε-IC.
- For $n = 1$ item: in fact **precise IC**, computational **efficiency**.

Some Remarks on Model Misspecification

- One of the main proof ideas: relatively small (i.e., grows "moderately enough") number of possible oracle inputs.
  - Independent valuation distributions + Deliberately introduced a controlled amount of model misspecification.
  - A fairly moderate price: ε-IC.
- For $n = 1$ item: in fact **precise IC**, computational **efficiency**.
- For $m = 1$ buyer, there is a known **"ε-IC to IC transformation"** that loses negligible (poly(ε)) revenue without requiring any knowledge of any distribution (Nisan, ∼'05; see also Madarász and Prat, 2017)

Some Remarks on Model Misspecification

- One of the main proof ideas: relatively small (i.e., grows "moderately enough") number of possible oracle inputs.
  - Independent valuation distributions +
    Deliberately introduced a controlled amount of model misspecification.
  - A fairly moderate price: ε-IC.
- For $n=1$ item: in fact **precise IC**, computational **efficiency**.
- For $m=1$ buyer, there is a known **"ε-IC to IC transformation"** that loses negligible (poly(ε)) revenue without requiring any knowledge of any distribution (Nisan, ∼'05; see also Madarász and Prat, 2017) ⇒ **overcome** the model misspecification to overall learn a **precisely IC** mechanism!

Some Remarks on Model Misspecification

- One of the main proof ideas: relatively small (i.e., grows "moderately enough") number of possible oracle inputs.
  - Independent valuation distributions + Deliberately introduced a controlled amount of model misspecification.
  - A fairly moderate price: ε-IC.
- For $n=1$ item: in fact **precise IC**, computational **efficiency**.
- For $m=1$ buyer, there is a known **"ε-IC to IC transformation"** that loses negligible (poly(ε)) revenue without requiring any knowledge of any distribution (Nisan, ∼'05; see also Madarász and Prat, 2017) ⇒ **overcome** the model misspecification to overall learn a **precisely IC** mechanism!
  - We do not know a "direct path" to learning an IC mechanism.

Some Remarks on Model Misspecification

- One of the main proof ideas: relatively small (i.e., grows "moderately enough") number of possible oracle inputs.
  - Independent valuation distributions + Deliberately introduced a controlled amount of model misspecification.
  - A fairly moderate price: ε-IC.

- For $n = 1$ item: in fact **precise IC**, computational **efficiency**.

- For $m = 1$ buyer, there is a known **"ε-IC to IC transformation"** that loses negligible (poly(ε)) revenue without requiring any knowledge of any distribution (Nisan, ∼'05; see also Madarász and Prat, 2017) ⇒ **overcome** the model misspecification to overall learn a **precisely IC** mechanism!
  - We do not know a "direct path" to learning an IC mechanism.

- Deliberate model misspecification as a **tool against overfitting.**

Conclusion & Further Research

- Main takeaway: **empirical revenue maximization not harder than Bayesian revenue maximization** in many settings: any result that holds given full information immediately implies a robust result from samples.
  - For many buyers AND many items, a fairly moderate price to pay: ε-IC.
  - Otherwise (see the paper for details): same incentive compatibility notion.

Conclusion & Further Research

- Main takeaway: **empirical revenue maximization not harder than Bayesian revenue maximization** in many settings: any result that holds given full information immediately implies a robust result from samples.
  - For many buyers AND many items, a fairly moderate price to pay: ε-IC.
  - Otherwise (see the paper for details): same incentive compatibility notion.

Open Question

Given an IR and ε-BIC auction for some product distribution, even in an additive multi-item setting, is it possible to transform it into an IR and (precisely) BIC auction with negligible ($\text{poly}(\varepsilon) \cdot \text{poly}(m, n, H)$) revenue loss using polynomially many samples from this product distribution?

Conclusion & Further Research

- Main takeaway: **empirical revenue maximization not harder than Bayesian revenue maximization** in many settings: any result that holds given full information immediately implies a robust result from samples.
  - For many buyers AND many items, a fairly moderate price to pay: ε-IC.
  - Otherwise (see the paper for details): same incentive compatibility notion.

Open Question

Given an IR and ε-BIC auction for some product distribution, even in an additive multi-item setting, is it possible to transform it into an IR and (precisely) BIC auction with negligible ($\mathrm{poly}(\varepsilon)\cdot\mathrm{poly}(m, n, H)$) revenue loss using polynomially many samples from this product distribution?

- Known ε-BIC-to-BIC reduction from samples (DHKN'17) requires a number of samples that is polynomial in the size of the type space = exponential in the number of items, but does **not** assume independence.

- Can independence come to the rescue?

Questions?