Bounding the Competition Complexity via Dual Flows, Discretizations, and Symmetries

Emily Ryu
Advisor: Prof. Matt Weinberg
PACM Reader: Prof. Mark Braverman

April 23, 2021

Abstract

Designing a multi-item auction to extract optimal revenue is very difficult and often practically infeasible, but by recruiting additional buyers to compete for the items, a seller can run a simple auction (such as running a separate second-price auction for each item) and still extract greater revenue than the optimal mechanism without extra buyers. The number of additional bidders necessary such that selling the items separately (to additional bidders) guarantees greater expected revenue than the optimal mechanism (without additional bidders) is termed the *competition complexity*.

Seminal work by Bulow and Klemperer showed that perhaps surprisingly, only one additional buyer is needed in the single-item setting. But with even two independent items, things are more complex; previous work has shown that the competition complexity for n buyers with additive values for two items drawn independently and identically distributed from the equal revenue distribution is $\Omega(\log n)$ and $\mathcal{O}(\sqrt{n})$. We seek to close the gap between the currently established upper and lower bounds by exploring a number of different techniques to make this bound tight in n.

Contents

1 Introduction

Consider a monopolist seller with m distinct items, facing n distinct buyers. The seller decides to run an *auction*, a process in which the buyers compete to win the items by submitting bids. This setting is very common in the real world: for example, eBay, Google, and the NYSE all use auctions to set prices for items, ads, and stocks. In all of these cases, the seller's goal is naturally to maximize their revenue. We define an *auction mechanism* as taking as input a set of bids and outputting an allocation of the items awarded to and a price charged to each bidder. So, what mechanism should the seller use to maximize revenue?

In the single-item setting, the revenue-optimal auction has been found by [Mye81], and is a simple and straightforward mechanism. However, multi-item auctions are significantly more complicated. Intuitively, it might seem that we can just treat a multi-item auction as selling the items separately since they are independent, but this turns out to be suboptimal, even with only 1 buyer and 2 independent items from a simple distribution [HN17]. From the buyer's perspective, the items are independent and there is no interaction between them, but from the seller's perspective, the existence of additional items enriches their strategy space and allows them to price options for which the buyer's value has lower variance (such as bundling the items together), thus extracting more of the buyer's value as revenue.

Further, the optimal multi-item auction may be randomized rather than deterministic [DDT12] or may offer the buyer an uncountable number of distinct menu options [DDT17]. There are also instances of *revenue monotonicity*, in which a distribution D^+ stochastically dominates D ($F(x) \geq F^+(x)$ for all x), yet the optimal revenue for D strictly exceeds the optimal revenue for D^+; that is, a "strictly better" distribution achieves "strictly worse" revenue [HR15]. Given these challenges, a seller might hope that a simple auction such as selling the items separately or bundling them together is at least approximately-optimal, but neither one of these mechanisms provides a constant-factor approximation; the approximation gets worse as the number of items increases [HN17].

Clearly, optimal multi-item actions are difficult and impractical to understand, so the seller may settle for just using a simple mechanism. However, they do not want to lose out on revenue from doing this, so they may decide to recruit additional buyers in hopes that increased competition for the items will drive prices up. This now prompts the question: how many additional bidders are needed so that the seller can use a simple mechanism such as selling separately, but still guarantee at least the optimal revenue (without the additional bidders)? We term this amount the *competition complexity*.

1.1 Roadmap

We begin in Section 2 with a discussion of the core theoretical concepts in auction mechanism design. Sections 3 through 5 each cover a different approach we explored to bound the competition complexity, each beginning with additional background to familiarize the reader with the necessary theory, then detailing our results from applying the technique. We conclude in Section 6 with notes on potential future directions of investigation.

2 Background and preliminaries

For a comprehensive treatment of auction theory and the equal revenue distribution, we refer the reader to [Bar20]. Here, we cover only the background theory directly used in this project.

2.1 Auction format

To begin, we assume that the seller only has the ability to specify the parameters of the auction before it is run. Once the auction begins, the seller simply executes the auction as specified beforehand, and does not actively participate or change the outcome with their actions.

We label the bidders as elements of the set $[n] = \{1, \dots, n\}$ and the items as elements of the set $[m] = \{1, \dots, m\}$. In the setup phase before the auction, each bidder i draws v_{ij}, their true *private valuation* for each item j, from distribution $\mathcal{D}_{ij}$. The seller does not know the value vector $\vec{v}$ (hence the term "private values"), but does know each $\mathcal{D}_{ij}$ (one can imagine that the seller can obtain information about the distribution of values in the general population through market research, but cannot obtain the true values of each individual bidder). The true valuations are *additive*, so that from a certain bidder's point of view, their value for a set of items is exactly the sum of their values for each item in the set.

In the auction, each bidder simultaneously submits a single *sealed bid* to the seller, meaning that only that bidder and the seller see the bid. The seller thus receives a reported profile of bids, and then uses some mechanism to allocate the prices and charge prices accordingly. We can formally define an *auction mechanism* as follows:

Definition 1 (Auction mechanism)**.** *An auction mechanism takes as input a set of bids and outputs an allocation vector and a price vector. Let $X_i(\vec{v})$ and $P_i(\vec{v})$ be the allocation and price vectors, respectively, for bidder i when value vector $\vec{v} = (\vec{v_1}, \dots, \vec{v_n})$ (the concatenation of each individual bidder's value vector) is input. Then, we can formally represent the mechanism as $M = (X, P)$, where $X = (X_1(\vec{v}), \dots, X_n(\vec{v}))$, $P = (P_1(\vec{v}), \dots, P_n(\vec{v}))$.*

Given the output allocation vector X (which is a probability distribution for each item), the seller then samples the allocation distribution to obtain the final allocation of items to bidders. Throughout this paper, we use the terms *auction* and *mechanism* interchangeably.

2.2 Notation

We establish the following definitions and notation for use throughout this paper:

- $\pi_i(\vec{v}_i) \coloneqq \mathbb{E}_{\vec{v}_{-i} \leftarrow \mathcal{D}_{-i}}[X_i(\vec{v})]$, the *expected allocation* (over all other bidders reporting their true values) for bidder i when bidding $\vec{v}_i$, also referred to as the *interim allocation*.
- $\pi_{ij}(\vec{v}_i) \coloneqq \mathbb{E}_{\vec{v}_{-i} \leftarrow \mathcal{D}_{-i}}[X_{ij}(\vec{v})]$, the *expected probability* of bidder i winning item j when bidding $\vec{v}_i$, also referred to as the *interim probability*.
- $q_i(\vec{v}_i) \coloneqq \mathbb{E}_{\vec{v}_{-i} \leftarrow \mathcal{D}_{-i}}[P_i(\vec{v})]$, the *expected payment* of bidder i when bidding $\vec{v}_i$.
- $U(\vec{v}_i, M_i(\vec{v})) \coloneqq \vec{v}_i \cdot X_i(\vec{v}) - P_i(\vec{v})$, the *utility* of bidder i when the complete bidder profile is $\vec{v}$. Note that values for items are *additive*, and utilities are *quasi-linear* (subtracting price from value).
- $\vec{v}_i \cdot \pi_i(\vec{v}_i) - q_i(\vec{v}_i)$, the *expected utility* of bidder i when bidding $\vec{v}_i$.
- $\textsc{Rev}(\mathcal{D})$, the *optimal achievable revenue* for bidders whose values are drawn from $\mathcal{D} = \times_{i=1}^{n} \times_{j=1}^{m} \mathcal{D}_{ij}$.
- $\textsc{Rev}^M(\mathcal{D})$, the *expected revenue achieved by mechanism M* for distribution $\mathcal{D}$ when buyers bid truthfully.
- $\textsc{SRev}(\mathcal{D})$, the optimal achievable revenue by any mechanism that *sells separately* for distribution $\mathcal{D}$.
- $\textsc{ARev}(\mathcal{D}) \coloneqq \textsc{Rev}(\mathcal{D}) - \textsc{SRev}(\mathcal{D})$, the additional revenue that the optimal mechanism provides over selling separately to bidders from distribution $\mathcal{D}$. We term this the *adjusted revenue*.
- $\textsc{Val}(\mathcal{D}) \coloneqq \mathbb{E}_{\vec{v} \leftarrow \mathcal{D}}\left[\sum_j \max_i\{v_{ij}\}\right]$, the *expected optimal welfare* for distribution $\mathcal{D}$ (obtained by giving each item to the bidder with highest value).
- If all the buyers are i.i.d., *i.e.* for all i, $\mathcal{D}_i = \mathcal{D}$ for some distribution $\mathcal{D}$, then $\textsc{Rev}_n(\mathcal{D})$ denotes the optimal achievable revenue for n bidders, each of whom has values drawn from $\mathcal{D}$. $\textsc{Rev}_n^M(\mathcal{D})$, $\textsc{SRev}_n(\mathcal{D})$, $\textsc{ARev}_n(\mathcal{D})$, and $\textsc{Val}_n(\mathcal{D})$ are defined analogously.
- $\mathcal{ER}$, the *equal revenue distribution*, defined by the CDF $F(x) = 0$ for $x < 1$, $F(x) = 1 - \frac{1}{x}$ for $x \geq 1$.
- $\mathcal{ER}_t$, the equal revenue distribution *truncated at t*, defined by drawing x from $\mathcal{ER}$ and replacing x with t if and only if $x > t$.
- $\mathcal{ER}_{t,d}$, the equal revenue distribution *truncated at t and discretized* according to some discretization (see Section 4 for further discussion).

- $c(n,m) \coloneqq \min\{k \,|\, \text{SRev}_{n+k}(\mathcal{ER}^m) \geq \text{Rev}_n(\mathcal{ER}^m)\}$, the *competition complexity* for n additive buyers with values for m items each drawn from $\mathcal{ER}$. This is the minimum number of additional buyers k such that selling separately to $n+k$ bidders guarantees at least as much revenue as the optimal revenue for n bidders.

2.3 Properties of auctions

Naturally, buyers are incentived to maximize their expected payoff, regardless of whether or not they bid their true valuations $\vec{v}_i$. However, in order to reason about expected revenue, the seller would like the buyers to bid their true values (otherwise, attempting to reason about non-truthful bids rather than true values may be challenging, or possibly not even a well-posed question). This motivates the following definition of Bayesian incentive compatibility, which states that for each buyers, if all other buyers bid truthfully, it is in their best interest to also bid truthfully.

Definition 2 (Bayesian incentive compatible)**.** *A mechanism $M = (X,P)$ is Bayesian incentive compatible (BIC) if for all bidders i, for all $\vec{v}_i, \vec{v}_i'$,*

$$\vec{v}_i \cdot \pi_i(\vec{v}_i) - q_i(\vec{v}_i) \geq \vec{v}_i \cdot \pi_i(\vec{v}_i') - q_i(\vec{v}_i').$$

A BIC auction achieves a Bayesian Nash equilibrium in which every buyer bids their true valuation.

We also define two additional properties that are crucial to ensuring that all agents (the seller and all the bidders) participate willingly in the auction.

Definition 3 (Individually rational)**.** *A mechanism $M = (X,P)$ is ex interim individually rational (IR) if for all $i, \vec{v}_i$,*

$$\vec{v}_i \cdot \pi_i(\vec{v}_i) - q_i(\vec{v}_i) \geq 0.$$

That is, the expected utility is positive for all bidders. An auction that is ex interim IR prohibits the seller from setting arbitrarily high prices, because then the bidders can just choose to not participate at all.

Definition 4 (No positive transfers)**.** *A mechanism $M = (X,P)$ is no positive transfers (NPT) if for all $i, \vec{v}_i$,*

$$q_i(\vec{v}_i) \geq 0.$$

That is, the seller never expects to pay the bidders. Otherwise, the seller could just choose not to run the auction at all.

In our work, we are concerned with the optimal achievable revenue from a BIC, IR, NPT auction.

2.4 Related work

We first establish the connection between revenue and competition complexity that will inform our analysis.

Proposition 5 ([Bar20])**.** *The revenue obtained from selling m items separately to n buyers, with each value drawn from $\mathcal{ER}$, is mn. That is,*

$$\text{SRev}_n(\mathcal{ER}^m) = mn.$$

In particular, if we consider the competition complexity when $m = 2$, we have

$$\begin{aligned} \textsc{SRev}_{n+k}(\mathcal{ER}^2) &\geq \textsc{Rev}_n(\mathcal{ER}^2) \\ 2n + 2k &\geq \textsc{Rev}_n(\mathcal{ER}^2) \\ 2k &\geq \textsc{Rev}_n(\mathcal{ER}^2) - 2n = \textsc{Rev}_n(\mathcal{ER}^2) - \textsc{SRev}_n(\mathcal{ER}^2) \\ \implies c(n,2) &= \frac{\textsc{ARev}_n(\mathcal{ER}^2)}{2}. \end{aligned}$$

That is, the competition complexity and the adjusted revenue have the same order of growth in n. Thus, much of the previous work on bounding the competition complexity has framed the problem in terms of bounding the adjusted revenue (note that while the optimal *mechanism* is generally challenging to understand, we may still be able to reason about the optimal *revenue*, and consequently the adjusted revenue). We continue using this framework throughout our work.

Currently, it is known that the competition complexity is $c(n,2) = \Omega(\log n)$ and $c(n,2) = \mathcal{O}(\sqrt{n})$. In particular, the best known bounds are:

Theorem 6 (Upper bound, due to [BW18])**.** *The competition complexity of n bidders with additive values over m independent items is at most $n(\log(1 + m/n) + 2)$ (tight when $n \leq m$) and also at most $9\sqrt{nm}$ (tight when $n \geq m$).*

Theorem 7 (Lower bound, due to [BW18])**.** *The competition complexity of n bidders with additive values over $m = 2$ i.i.d. items from $\mathcal{ER}^2$ is at least $\log n/20$.*

3 Dual flows

In this section, we seek to utilize the duality-based framework for Bayesian mechanism design developed by [CDW16]. A thorough exposition of the theory of dual flows is included in [Bar20], so here we provide only a brief overview of the necessary background.

3.1 Introduction to Lagrangian duality

Definition 8 (Lagrangian relaxation)**.** *Consider a linear program (LP) $\mathcal{L}$ of the form:*

$$\max \sum_i c_i x_i,$$

subject to the constraints

$$\begin{aligned} \sum_i A_{ji} x_i &\leq b_j, \quad \forall j \\ x_i &\geq 0, \quad \forall i. \end{aligned}$$

Let S be a subset of the constraints. For all $j \in S$, let $\lambda_j \geq 0$ be given, and let $\vec{\lambda}$ be the vector of all $\{\lambda_j\}_{j \in S}$. Then, a Lagrangian relaxation of the above LP for the subset S and Lagrangian multipliers $\vec{\lambda}$ is the following (which we call $\mathcal{L}_S^{\vec{\lambda}}$):

$$\max \sum_i c_i x_i + \sum_{j \in S} \lambda_j \cdot \left(b_j - \sum_i A_{ji} x_i \right),$$

subject to the constraints

$$\sum_i A_{ji} x_i \le b_j, \quad \forall j \notin S$$

$$x_i \ge 0, \quad \forall i.$$

Lemma 9. *For any linear program $\mathcal{L}$, all subsets of constraints S, and all non-negative Lagrangian multipliers $\vec{\lambda}$ for constraints in S, the optimal value of $\mathcal{L}_S^{\vec{\lambda}}$ is at least the optimal value of $\mathcal{L}$.*

Now, consider the revenue-maximizing LP for a BIC, IR, NPT multi-item auction with n bidders with valuation functions drawn from $\mathcal{D} = \times_i \mathcal{D}_i$:

Variables:

- $x_{i,S}(\vec{v})$, the probability that bidder i receives set S on input $\vec{v}$ (a complete bidder profile).
- $x_{\vec{S}}(\vec{v})$, the probability of selecting the partition $\vec{S}$ of the items on input $\vec{v}$.
- $\pi_{i,S}(v_i)$, the interim probability that bidder i receives set S when reporting $v_i(\cdot)$ (*i.e.*, the probability in expectation over all other bidders bidding their true types $\vec{v}_{-i}$).
- $p_i(v_i)$, the interim price paid by bidder i when reporting $v_i(\cdot)$ (in expectation over all other bidders bidding their true types).

Constraints:

- $\pi_{i,S}(v_i) = \sum_{\vec{v}_{-i}} f_{-i}(\vec{v}_{-i}) \cdot x_{i,S}(v_i; \vec{v}_{-i})$ for all i, v_i, S (guarantees that interim probabilities are computed correctly).
- $x_{i,S}(\vec{v}) = \sum_{\vec{S}|S_i=S} x_{\vec{S}}(\vec{v})$ for all $i, S, \vec{v}$ (guarantees that $x_{i,S}$ is computed correctly).
- $\sum_{\vec{S}} x_{\vec{S}}(\vec{v}) = 1$, $x_{\vec{S}}(\vec{v}) \ge 0$ for all $\vec{S}, \vec{v}$ (guarantees that probabilities of selections are non-negative and sum to 1).
- $\sum_S v_i(S) \cdot \pi_{i,S}(v_i) - p_i(v_i) \ge \sum_S v_i(S) \cdot \pi_{i,S}(v_i') - p_i(v_i')$ for all i, v_i, v_i' (BIC constraints).
- $\pi_{i,S}(\varnothing) = p_i(\varnothing) = 0$ for all i, S (interim probabilities and payments are 0 for any bidder who chooses not to participate).

Objective: Maximize the expected revenue; that is,

$$\max \sum_i \sum_{v_i} f_i(v_i) \cdot p_i(v_i).$$

We now construct a Lagrangian relaxation by putting the Lagrangian multiplier $\lambda_i(v_i, v_i')$ on the BIC constraint for a bidder of type v_i misreporting v_i'. By Lemma 9, a feasible solution to this relaxed LP upper bounds the optimal achievable revenue. We refer to a profile of Lagrangian multipliers $\vec{\lambda}$ as *useful* if it results in a finite upper bound, and it can be shown that:

Lemma 10. *A profile of Lagrangian multipliers $\vec{\lambda}$ is useful if and only if*

$$\lambda_i(v_i, v_i') = f_i(v_i) + \sum_{v_i'} \lambda_i(v_i', v_i)$$

for all i, v_i.

We can interpret this condition as follows: for all i, there is a *flow network* with a super source and a super sink (where the sink is $\varnothing$, the type corresponding to not participating). There is an internal node for each $v_i \in \operatorname{supp}(\mathcal{D}_i)$. There is flow from the source to each v_i of value $f_i(v_i)$, and flow from each v_i to v_i' of $\lambda(v_i, v_i')$. Then $\vec{\lambda}$ is useful if and only if this picture is a flow; that is, the flow leaving a point v_i is equal to the flow entering v_i plus the density at v_i.

Next, define

$$\begin{aligned}\Phi_{i,S}^{\vec{\lambda}}(\vec{v}_i) &:= \frac{\sum_{v_i'} \lambda_i(v_i, v_i') v_i(S) - \lambda_i(v_i', v_i) v_i'(S)}{f_i(v_i)} \\ &= v_i(S) - \frac{\sum_{v_i'} (v_i'(S) - v_i(S)) \cdot \lambda_i(v_i', v_i)}{f_i(v_i)},\end{aligned}$$

which we can interpret as the *virtual valuation* function of bidder i when their real valuation function is $v_i(\cdot)$.

Lemma 11 ([CDW16, BW18])**.** *The expected revenue of the optimal mechanism does not exceed the sum of the expected maximum virtual value for each item. That is,*

$$\textsc{Rev}_n(\mathcal{ER}^m) \leq \sum_{j=1}^{m} \mathbb{E}_{\vec{v} \leftarrow (\mathcal{ER}^m)^n} \left[\max_{i \in [n]} \{\Phi_{ij}^{\lambda}(\vec{v}_i)\} \right].$$

Thus, our approach is to attempt to find a profile of Lagrangian multipliers (satisfying the flow constraints, so that the profile is useful) such that the upper bound furnished by Lemma 11 is logarithmic (or at least improves upon the current upper bound).

Further, [Bar20] describes a framework for analyzing flow in a continuous setting (since $\mathcal{ER}^2$ is a continuous distribution) and a quantile transform of values. The latter allows us to reduce analysis for two items to flows over the unit square, originating at $(q_1, q_2) = (1, 1)$ and terminating at $(q_1, q_2) = (0, 0)$. To do this, we let q_i be the quantile corresponding to v_i (for $i = 1, 2$) and note that for $\mathcal{ER}$, we have $q_i = 1 - \frac{1}{v_i}$, $v_i = \frac{1}{1 - q_i}$. The key property of note is that q_i is uniformly distributed on the interval $[0, 1]$.

Previous analysis of linear flows yielded only $\mathcal{O}(\sqrt{n})$ bounds on competition complexity, so we now attempt to analyze nonlinear flows. We refer the reader to [Bar20] for precise definitions of all formalisms used, and simply adopt the notation used therein without further explanation.

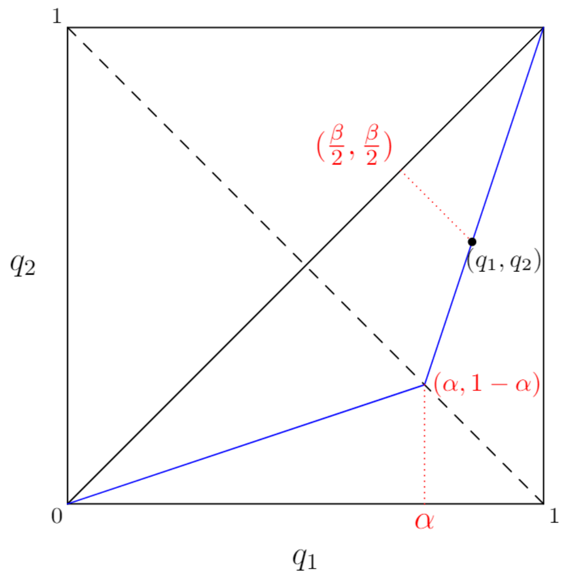
3.2 Piecewise linear flow

Our first nonlinear flow network was a simple piecewise linear flow, attempting to simultaneously capture properties of two different linear flows analyzed by [Bar20].

Proposition 12. *The symmetric piecewise linear flow originating from $(1, 1)$ and terminating at $(0, 0)$ does not improve upon the $\mathcal{O}(\sqrt{n})$ competition complexity bound.*

Proof. Since the proposed lines of flow are not smooth along the boundary $q_1 + q_2 = 1$, we consider the portions of the unit square above and below the boundary separately. We first consider the amount of flow accumulated in the top half.

For a given point (q_1, q_2) with $q_1 + q_2 > 1$ (and WLOG $q_1 > q_2$), let $(\alpha, 1-\alpha)$ be the point where the line of flow intersects the boundary $q_1 + q_2 = 1$, and let $(\frac{\beta}{2}, \frac{\beta}{2})$ be the foot of the perpendicular from (q_1, q_2) to the line $q_1 = q_2$. Clearly we have $\alpha \in [\frac{1}{2}, 1], \beta \in [1, 2]$, and α and β are uniformly distributed over their respective integrals. Further, each flow line is parametrized by a unique value of α, and for fixed α, each point on the flow line is parametrized by a unique value of β.



Using coordinate geometry and algebra we can see that we have

$$\alpha = \frac{q_2 - 1}{q_1 + q_2 - 2} \qquad \beta = q_1 + q_2$$
$$q_1 = 2\alpha + \beta - \alpha\beta - 1 \qquad q_2 = \alpha\beta - 2\alpha + 1$$
$$\frac{\partial q_1}{\partial \alpha} = 2 - \beta \qquad \frac{\partial q_2}{\partial \alpha} = \beta - 2$$
$$\frac{\partial q_1}{\partial \beta} = 1 - \alpha \qquad \frac{\partial q_2}{\partial \beta} = \alpha.$$

Then we can compute

$$h(\alpha, \beta) = \left| \frac{\partial q_1}{\partial \alpha} \cdot \frac{\partial q_2}{\partial \beta} - \frac{\partial q_2}{\partial \alpha} \cdot \frac{\partial q_1}{\partial \beta} \right| = 2 - \beta,$$

so $\frac{\partial \lambda}{\partial \beta} = -h = \beta - 2 \implies \lambda(\alpha, \beta) = \frac{\beta^2}{2} - 2\beta + c$ for some constant c. Recall that since $\beta \in [1, 2]$, we choose the maximum value of c such that $\lambda(\alpha, 2)$ is nonpositive. This gives $c = 2$, so we have

$$\lambda(\alpha, \beta) = \frac{\beta^2}{2} - 2\beta + 2 = \frac{(\beta - 2)^2}{2}.$$

Finally, we compute

$$
\begin{aligned}
\Phi_1 &= \frac{1}{1-q_1} - \frac{\frac{\partial q_1}{\partial \beta} \cdot \lambda(\alpha,\beta)}{h(\alpha,\beta)\cdot(1-q_1)^2} \\
&= \frac{1}{2(\alpha-1)(\beta-2)}, \\
\Phi_2 &= \frac{1}{1-q_2} - \frac{\frac{\partial q_2}{\partial \beta} \cdot \lambda(\alpha,\beta)}{h(\alpha,\beta)\cdot(1-q_2)^2} \\
&= \frac{1}{2\alpha(2-\beta)}.
\end{aligned}
$$

Note that both virtual values are always positive in this region.

To better understand the competition complexity that results from this flow, we simulated the average adjusted revenue as a function of n. As discussed in [Bar20], the adjusted revenue comes from flow lines strictly in the interior of the unit square (bidders with quantile values on the boundary of the square will contribute $2n$ to the total revenue), so we now take $\alpha \in [\frac{1}{2}, 1), \beta \in [1, 2)$. For $n \in [100, 500]$, we generated n random draws of α and β, computed the resulting virtual values Φ_1 and Φ_2, and then calculated the sum of the maximum Φ_1 and the maximum Φ_2.

When numerically evaluated, we found that the flow accumulated above the boundary appears to be $\Theta(\sqrt{n})$. Then, regardless of what happens below the boundary, the total amount of flow accumulated over the entire square must be $\Omega(\sqrt{n})$, so this profile cannot possibly result in an upper bound on $\text{AREV}_n(\mathcal{ER}^2)$ that is $\mathcal{O}(\log n)$. $\square$

3.3 General smooth nonlinear flows

To avoid the issue of having to analyze different regions of the unit square separately, we next attempted to construct a more general nonlinear flow network originating from $(1,1)$ and terminating at $(0,0)$, consisting of *smooth* flow lines. We will let these flow lines have the form $y(x)$ and again parametrize each flow line by α, where $(\alpha, 1-\alpha)$ denotes the intersection with $x+y=1$. We require the following properties:

1. y' increases smoothly from $y'(0) = f(\alpha)$ to $y'(\alpha) = 1$ along the smooth function g_α (g is a family of functions, each parametrized by α):
$$y'(x) = f(\alpha) + g_\alpha\left(\tfrac{x}{a}\right), \qquad g_\alpha(0) = 0,\, g_\alpha(1) = 1 - f(\alpha).$$
2. Boundary conditions of slope: $\lim_{\alpha \to \frac{1}{2}^+} f(\alpha) = 1$, $\lim_{\alpha \to 1^-} f(\alpha) = 0$.
3. f is decreasing and continuous in α, and $f(\alpha) < \frac{1-\alpha}{\alpha} \leq 1$.
4. $\int_0^\alpha y'(x)dx = 1 - \alpha$.

However, we claim that this is impossible.

Proposition 13. *Such a function $y(x)$ satisfying properties 1-4 does not exist.*

Proof. Substituting property 1 into property 4 gives

$$\begin{aligned}
\int_0^\alpha y'(x)dx &= \int_0^\alpha \left(f(\alpha) + g_\alpha(\tfrac{x}{\alpha})\right)dx \\
1-\alpha &= \alpha f(\alpha) + \int_0^\alpha g_\alpha(\tfrac{x}{\alpha})dx = \alpha f(\alpha) + \int_0^1 \alpha g_\alpha(x)dx \\
\frac{1}{\alpha} - 1 &= f(\alpha) + \int_0^1 g_\alpha(x)dx \\
f(\alpha) &= \frac{1}{\alpha} - 1 - \int_0^1 g_\alpha(x)dx.
\end{aligned}$$

Now, define

$$k_\alpha := \frac{\int_0^1 g_\alpha(x)dx}{1-f(\alpha)} \in [0,1].$$

Observe that $k_{1/2} = k_1 = 0$, and we have

$$f(\alpha) = \frac{1}{\alpha} - 1 - k_\alpha(1-f(\alpha)) \implies f(\alpha) = 1 + \frac{1}{\alpha} + \frac{2}{k_\alpha - 1}.$$

We seek a k_α so that $f(\alpha)$ is decreasing from $f\left(\frac{1}{2}\right) = 1$ to $f(1) = 0$. Clearly $k_\alpha = 0$ works. Now suppose k_α is not identically 0, so it attains maximum value $k_{\alpha'} = \varepsilon > 0$ for some $\alpha' \in (0,1)$ (if α' is not unique, take the largest α'). We consider $f(\alpha' + \delta)$ for $\delta > 0$:

$$\begin{aligned}
f(\alpha'+\delta) &\le f(\alpha') \\
1 + \frac{1}{\alpha'+\delta} + \frac{2}{k_{\alpha'+\delta} - 1} &\le 1 + \frac{1}{\alpha'} + \frac{2}{\varepsilon - 1} \\
\frac{2}{1-k_{\alpha'+\delta}} &\ge \frac{1}{\alpha'+\delta} - \frac{1}{\alpha'} + \frac{2}{1-\varepsilon} = \frac{2\alpha'(\alpha'+\delta) - \delta(1-\varepsilon)}{\alpha'(\alpha'+\delta)(1-\varepsilon)} \\
k_{\alpha'+\delta} &\ge 1 - \frac{2\alpha'(\alpha'+\delta)(1-\varepsilon)}{2\alpha'(\alpha'+\delta) - \delta(1-\varepsilon)} = \frac{2\alpha'(\alpha'+\delta)\varepsilon - \delta(1-\varepsilon)}{2\alpha'(\alpha'+\delta) - \delta(1-\varepsilon)}.
\end{aligned}$$

We also have $k_{\alpha'+\delta} < \varepsilon$, so this gives

$$\begin{aligned}
\frac{2\alpha'(\alpha'+\delta)\varepsilon - \delta(1-\varepsilon)}{2\alpha'(\alpha'+\delta) - \delta(1-\varepsilon)} &< \varepsilon \\
\delta\varepsilon(1-\varepsilon) &< \delta(1-\varepsilon),
\end{aligned}$$

which holds as long as $\varepsilon < 1$. But we already know $k_\alpha \le 1$, so this does not provide us with new information.

On the other hand, take α' to be the smallest value such that $k_{\alpha'} = \varepsilon$ (the max attained by k_α), and consider $f(\alpha' - \delta)$ for $\delta > 0$:

$$
\begin{aligned}
f(\alpha' - \delta) &\geq f(\alpha') \\
1 + \frac{1}{\alpha' - \delta} + \frac{2}{k_{\alpha'-\delta} - 1} &\geq 1 + \frac{1}{\alpha'} + \frac{2}{\varepsilon - 1} \\
\frac{2}{1 - k_{\alpha'-\delta}} &\leq \frac{1}{\alpha' - \delta} - \frac{1}{\alpha'} + \frac{2}{1-\varepsilon} = \frac{2\alpha'(\alpha'-\delta) + \delta(1-\varepsilon)}{\alpha'(\alpha'-\delta)(1-\varepsilon)} \\
k_{\alpha'-\delta} &\leq 1 - \frac{2\alpha'(\alpha'-\delta)(1-\varepsilon)}{2\alpha'(\alpha'-\delta)+\delta(1-\varepsilon)} = \frac{2\alpha'(\alpha'-\delta)\varepsilon + \delta(1-\varepsilon)}{2\alpha'(\alpha'-\delta)+\delta(1-\varepsilon)}.
\end{aligned}
$$

We also have $k_{\alpha'-\delta} \geq 0$, so we require

$$
\frac{2\alpha'(\alpha'-\delta)\varepsilon + \delta(1-\varepsilon)}{2\alpha'(\alpha'-\delta)+\delta(1-\varepsilon)} \geq 0.
$$

Since $\varepsilon < 1$, the numerator of the above expression is smaller than the denominator. We can verify that for $\delta > \frac{2\alpha'^2}{2\alpha'+\varepsilon-1}$, we have $2\alpha'(\alpha'-\delta)+\delta(1-\varepsilon) < 0$, thus $2\alpha'(\alpha'-\delta)\varepsilon + \delta(1-\varepsilon) < 0$, and $\frac{2\alpha'(\alpha'-\delta)\varepsilon+\delta(1-\varepsilon)}{2\alpha'(\alpha'-\delta)+\delta(1-\varepsilon)} \geq 0$ holds.

For $\delta < \frac{2\alpha'^2}{2\alpha'+\varepsilon-1}$, we have $2\alpha'(\alpha'-\delta)+\delta(1-\varepsilon) > 0$, so we now require

$$
\begin{aligned}
2\alpha'(\alpha'-\delta)\varepsilon + \delta(1-\varepsilon) &> 0 \\
2\alpha'^2\varepsilon &> (2\alpha'\varepsilon + \varepsilon - 1)\delta
\end{aligned}
$$

If $\varepsilon > \frac{1}{1+2\alpha'}$, then $\delta < \frac{2\alpha'^2\varepsilon}{2\alpha'\varepsilon+\varepsilon-1}$. But $\frac{2\alpha'^2\varepsilon}{2\alpha'\varepsilon+\varepsilon-1} \leq \frac{2\alpha'^2}{2\alpha'\varepsilon+\varepsilon-1}$, so we have a contradiction.

Otherwise $\varepsilon < \frac{1}{1+2\alpha'}$, which implies $\delta > \frac{2\alpha'^2\varepsilon}{1-2\alpha'\varepsilon-\varepsilon}$. But this means that for $\delta < \frac{2\alpha'^2\varepsilon}{1-2\alpha'\varepsilon-\varepsilon}$, the numerator is negative while the denominator is positive, so we end up with $k_{\alpha'-\delta} < 0$, which is a contradiction.

Thus, taking δ to be sufficiently small gives us a contradiction for $k_{\alpha'-\delta}$, so we conclude that such $y(x)$ cannot exist. □

While other nonlinear flow lines are likely to exist, the above shows that they cannot simultaneously satisfy basic symmetry and smoothness constraints. Without insisting on both of these constraints, we are unlikely to be able to analyze the resulting flows by hand (and even satisfying properties 1-4 is not a guarantee of straightforward analysis), so continuing to study nonlinear flows does not appear to be a promising direction. Instead, we move forwards with other forms of computational analysis.

4 Discretizations and computational analysis

Our primary tool for analyzing the optimal revenue via computation was RoaSolver [Shu19], a software program that, given the distribution of bidder values for each item, constructs the linear program corresponding to the BIC, IR, and NPT constraints and computes the revenue-optimal auction. The goal was to better understand whether the competition complexity behaves more like $\Theta(\log n)$, matching the current lower bound, or like $\Theta(\sqrt{n})$, matching the current upper bound.

The major limitation of the software is that it only takes discrete probability distributions as input, so it is impossible to draw values perfectly from the equal revenue distribution. Further, RoaSolver runs in size polynomial in s^2, where s is the size of the support of the distribution, and in practice can only be used up to $s = 14$. Therefore, effectively converting the unbounded, continuous $\mathcal{ER}$ distribution into a truncated, discretized distribution $\mathcal{ER}_{t,d}$ is a non-trivial, but important task.

4.1 Initial discretization

We initially discretized along the first 14 powers of 2, resulting in the distribution shown in Table 1. The probabilities of drawing v, $\mathbb{P}(v)$, were calculated to ensure that $\mathcal{ER}_{t,d}$ was stochastically dominated by $\mathcal{ER}$ while also being as close of an approximation as possible, by using

$$\begin{aligned}\mathbb{P}_{v\leftarrow\mathcal{ER}_{t,d}}\left(v = 2^k\right) &= \begin{cases}\mathbb{P}_{v\leftarrow\mathcal{ER}}\left(v \in [2^k, 2^{k+1})\right) & \text{for } k \in [0,12] \\ \mathbb{P}_{v\leftarrow\mathcal{ER}}\left(v \geq 2^k\right) & \text{for } k = 13\end{cases} \\ &= \begin{cases}\frac{1}{2^k} - \frac{1}{2^{k+1}} & \text{for } k \in [0,12] \\ \frac{1}{2^k} & \text{for } k = 13\end{cases}.\end{aligned}$$

v	1	2	4	8	16	32	64	128	256	512	1024	2048	4096	8192
$\mathbb{P}(v)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$	$\frac{1}{256}$	$\frac{1}{512}$	$\frac{1}{1024}$	$\frac{1}{2048}$	$\frac{1}{4096}$	$\frac{1}{8192}$	$\frac{1}{8192}$

Table 1: Initial $\mathcal{ER}$ discretization, $s = 14$.

We used RoaSolver to compute the optimal revenue for an auction with $m = 2$ items and $n \in [1, 100]$ bidders, and then plotted $\text{AR}\textsc{ev}_n(\mathcal{ER}^2_{t,d})$ as a function of n, as shown in Figure 1. The solution output by RoaSolver appears logarithmic for small values of n (for $n \in [1, 36]$, the adjusted revenue had squared error ≈ 0.162 to the closest fit square root function, but squared error ≈ 0.002 to the closest fit logarithmic function [Bar20]).

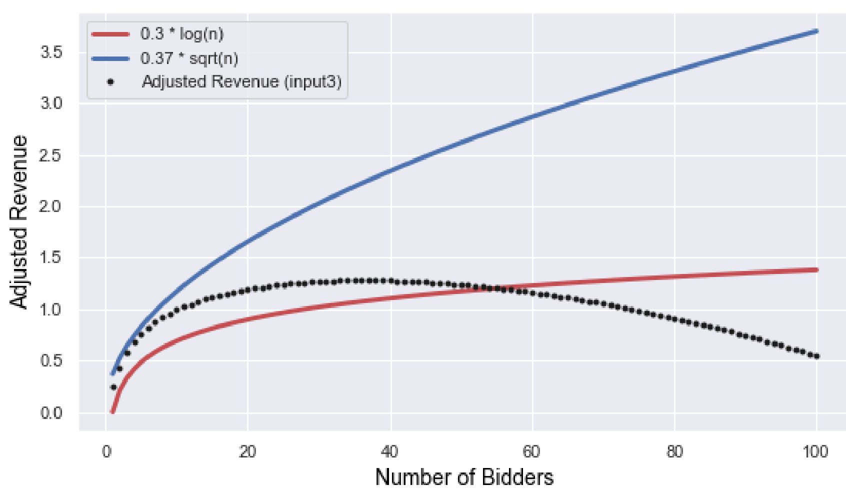


Figure 1: $\text{AR}\textsc{ev}_n(\mathcal{ER}^2_{t,d})$ for $m = 2$ items and $n \in [1, 100]$ bidders, computed by RoaSolver using the distribution detailed in Table 1.

However, the adjusted revenue begins to decrease around $n = 36$. We rationalized this as being due to the truncation of the distribution, where we can see that when we draw $2n$ values from $\mathcal{ER}$, the probability of having at least one value above a truncation cutoff T is $1 - \left(1 - \frac{1}{T}\right)^{2n}$. By a union bound this is bounded above by $\frac{2n}{T}$, and for the values of n and T we investigated here, this bound is close to tight. Thus as n grows, we lose more and more revenue from truncation, meaning that the discretization is no longer a good simulation of the true equal revenue curve. But for small values of n, we expect the discretization to be a fairly good simulation, giving us some degree of confidence that the competition complexity is $\Theta(\log n)$.

4.2 Discretization optimizations

We made various modifications to the initial discretization as described below and plotted the adjusted revenue for each discretization. For each of the following discretizations, $\mathbb{P}(v)$ is reported as the *unnormalized* probability of drawing value v (so the normalized probability is obtained by dividing by the sum of all the unnormalized probabilities.) Probabilities were computed using the same method described in Section 3.1 to ensure stochastic dominance. We first analyzed each discretization separately, and then plotted the best adjusted revenue across all tested discretizations for each n, shown in Figure 9.

By increasing the largest point in the discretization to a very high value of 2^{23} as in Table 2, we were able to minimize the effects of truncation for n in the interval $[1, 100]$, so that the computed adjusted revenue was monotonically increasing in n. This experiment suggests that a single extremely high value in the support is sufficient to avoid significant revenue loss due to truncation.

v	1	2	4	8	16	32	64
$\mathbb{P}(v)$	4194304	2097152	1048576	524288	262144	131072	65536
v	128	256	512	1024	2048	4096	8388608
$\mathbb{P}(v)$	32768	16384	8192	4096	2048	2047	1

Table 2: Modification of initial $\mathcal{ER}$ discretization, with the largest point changed from 2^{13} to 2^{23}, $s = 14$.

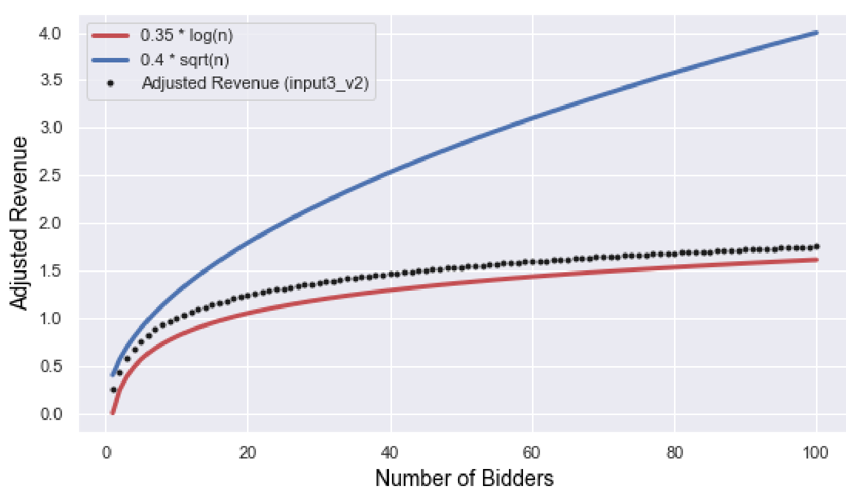


Figure 2: $\mathrm{AREV}_n(\mathcal{ER}^2_{t,d})$ for $m = 2$ items and $n \in [1, 100]$ bidders, using the distribution detailed in Table 2.

We discretized low values extremely finely at the expense of mid-range and high values, but attempted to compensate by taking larger high values and an extremely high truncation cutoff (Table 3). However, this resulted in decreasing adjusted revenue starting at a low value of $n \approx 20$. This can be understood using the same concept as the loss due to truncation: given the limited support size, putting *too much* weight on low values means that we do not accurately discretize mid-range and high values, from which we lose more and more revenue as n increases. The inaccuracy in Table 3 becomes significant at a lower point in the distribution than in Table 1, explaining the earlier onset of decreasing adjusted revenue.

v	1	1.333	1.5	2	3	4	10
$\mathbb{P}(v)$	7500000	2500000	5000000	5000000	2500000	4500000	1500000
v	20	100	200	1000	2000	10000	10000000
$\mathbb{P}(v)$	1200000	150000	120000	15000	12000	2997	3

Table 3: Particularly fine discretization of lower values and higher truncation cutoff, $s = 14$.

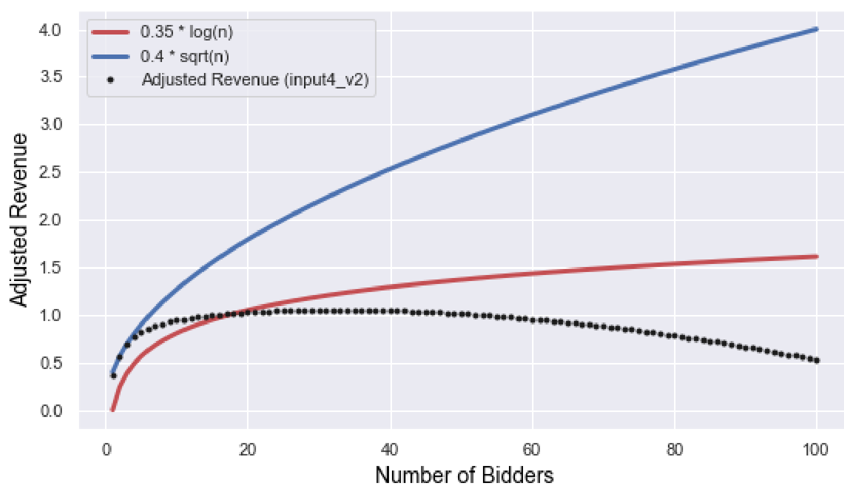


Figure 3: $\text{AREV}_n(\mathcal{ER}^2_{t,d})$ for $m = 2$ items and $n \in [1, 100]$ bidders, using the distribution detailed in Table 3.

The discretization in Table 4, an "intermediate" discretization between Table 2 and Table 3 also performed well. As shown in Figure 9, this discretization performed better for small values of n ($n \leq 14$), while Table 2 performed better for $n \geq 15$. This suggests that when n is small, it is important to finely discretize low values (as it is likely that all bidders will have small values, so we want to accurately capture their values in this range rather than rounding down to a few small values), but for larger n we must discretize mid-range values somewhat more carefully.

v	1	2	3	4	5	10	20
$\mathbb{P}(v)$	7500000	2500000	1250000	750000	1500000	750000	450000
v	50	100	200	500	1000	2500	5000000
$\mathbb{P}(v)$	150000	75000	45000	15000	9000	5997	3

Table 4: Lower values discretized more finely than Table 2 but less extremely than Table 3, $s = 14$.

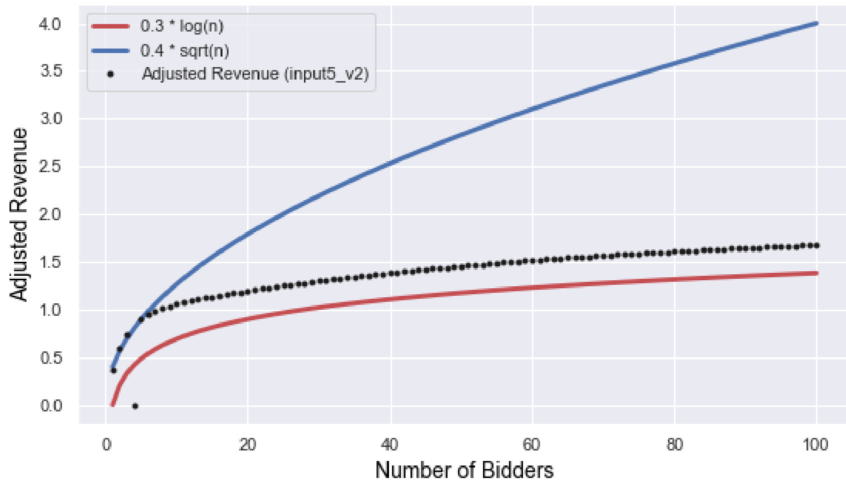


Figure 4: $\text{AREV}_n(\mathcal{ER}^2_{t,d})$ for $m = 2$ items and $n \in [1, 100]$ bidders, using the distribution detailed in Table 4.

Removing the second highest point to get Table 5 did not result in significant revenue loss, even with $s = 13$. We also see that the second highest point in the support does not have to be very large (*e.g.* 1000 suffices).

v	1	2	3	4	5	10	20
$\mathbb{P}(v)$	7500000	2500000	1250000	750000	1500000	750000	450000
v	50	100	200	500	1000	5000000	
$\mathbb{P}(v)$	150000	75000	45000	15000	14997	3	

Table 5: Modification of discretization in Table 4 with second highest point removed, $s = 13$.

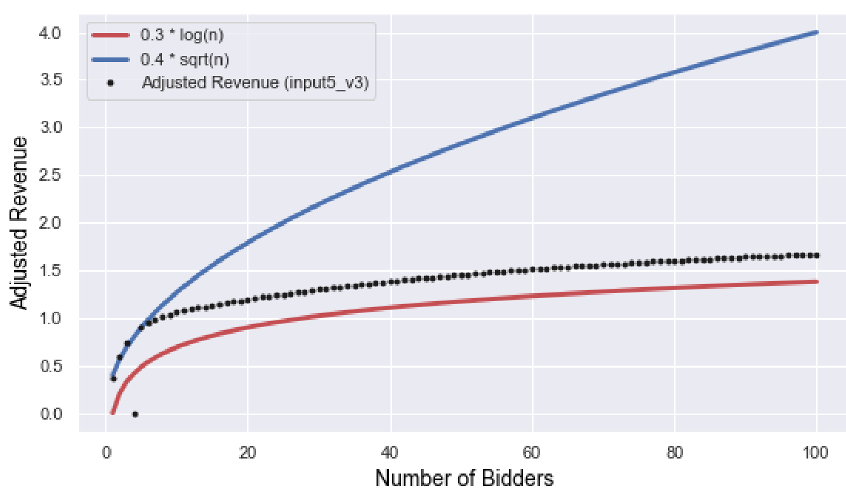


Figure 5: $\text{AREV}_n(\mathcal{ER}^2_{t,d})$ for $m = 2$ items and $n \in [1, 100]$ bidders, using the distribution detailed in Table 5.

When we removed 5, a relatively low point, from the support (Table 6), adjusted revenue suffered noticeably for all n and became approximately constant very quickly (at $n \approx 7$). This indicates that when attempting to reduce the size of the support, maintaining fine discretization of the low values remains important.

v	1	2	3	4	10	20
$\mathbb{P}(v)$	7500000	2500000	1250000	2250000	750000	450000
v	50	100	200	500	1000	5000000
$\mathbb{P}(v)$	150000	75000	45000	15000	14997	3

Table 6: Modification of discretization in Table 5 with low value 5 removed, $s = 12$.

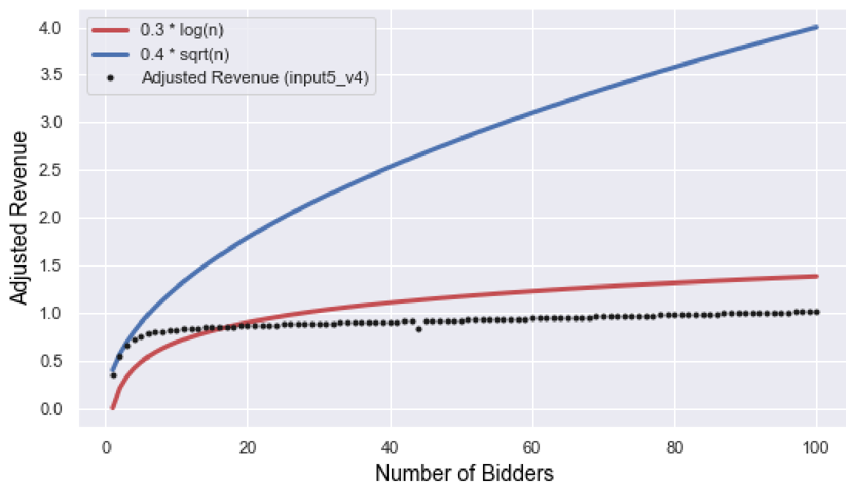


Figure 6: $\mathrm{AR{\small EV}}_n(\mathcal{ER}^2_{t,d})$ for $m = 2$ items and $n \in [1, 100]$ bidders, using the distribution detailed in Table 6.

We modified Table 4 to discretize mid-range values less finely so that we could include additional high values in the support. However, the resulting Table 7 performed less well for all n (with the exception of $n = 4$), suggesting that in the $n \in [1, 100]$ range, the low and mid-range values are more important to discretize well than the high values.

v	1	2	3	4	5	10	25
$\mathbb{P}(v)$	15000000	5000000	2500000	1500000	3000000	1800000	900000
v	100	250	1000	2500	10000	25000	10000000
$\mathbb{P}(v)$	180000	90000	18000	9000	1800	1197	3

Table 7: Modification of Table 4 with larger, less finely discretized higher values, $s = 14$.

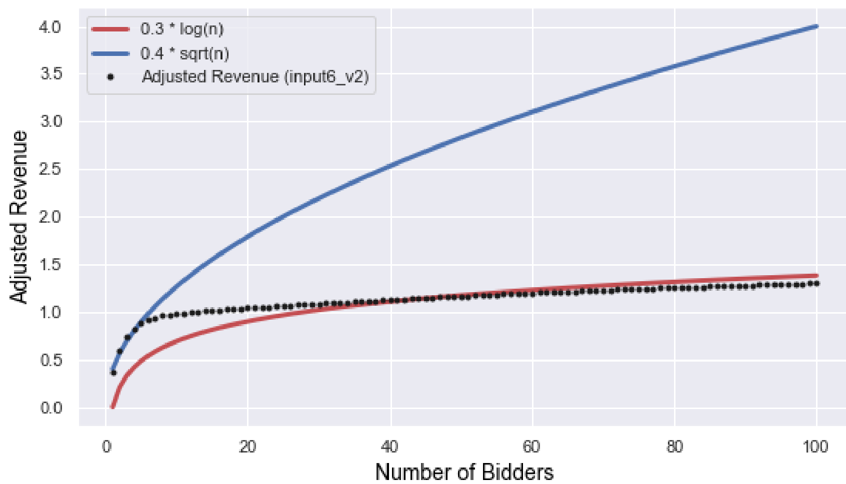


Figure 7: $\text{AR}\textsc{ev}_n(\mathcal{ER}^2_{t,d})$ for $m = 2$ items and $n \in [1, 100]$ bidders, using the distribution detailed in Table 7.

Finally, we created Table 8 an overall coarse distribution with generally high values. This resulted in relatively low adjusted revenue for all n (although the extreme high point prevented a dropoff in revenue due to truncation loss), further supporting our conclusions that for $n \in [1, 100]$, it is more important to discretize low values, and adding more high values does not improve adjusted revenue significantly.

v	1	1.5	2	3	6	24
$\mathbb{P}(v)$	7500000	2500000	1250000	2250000	750000	450000
v	1209600	604800	604800	604800	453600	120960
$\mathbb{P}(v)$	25200	4320	630	80	9	1

Table 8: Discretization using factorials (giving a coarse discretization with very high values), $s = 12$.

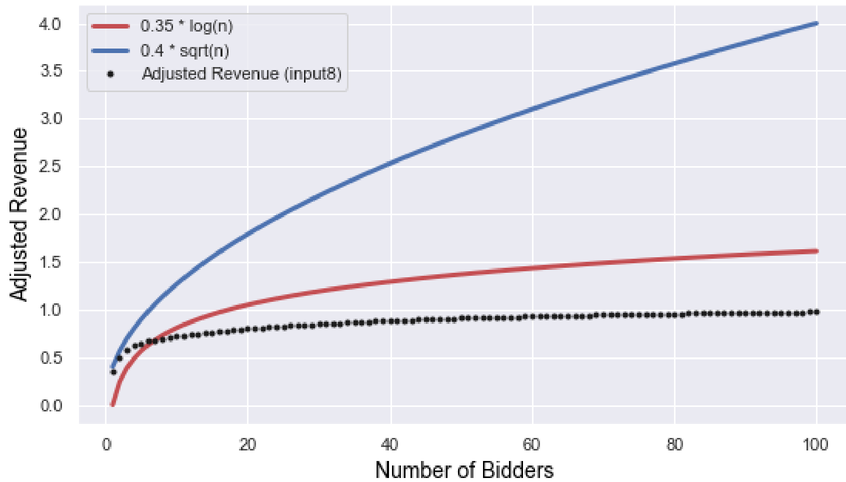


Figure 8: $\text{AREV}_n(\mathcal{ER}^2_{t,d})$ for $m = 2$ items and $n \in [1, 100]$ bidders, using the distribution detailed in Table 8.

In the masterplot (Figure 9), we can see that the best adjusted revenue for $n \in [1, 100]$ appears logarithmic. The adjusted revenue had squared error ≈ 0.453 to the closest fit square root function ($y = 0.115\sqrt{x} + 0.685$), but squared error only ≈ 0.026 to the closest fit logarithmic function ($y = 0.297 \log x + 0.373$).

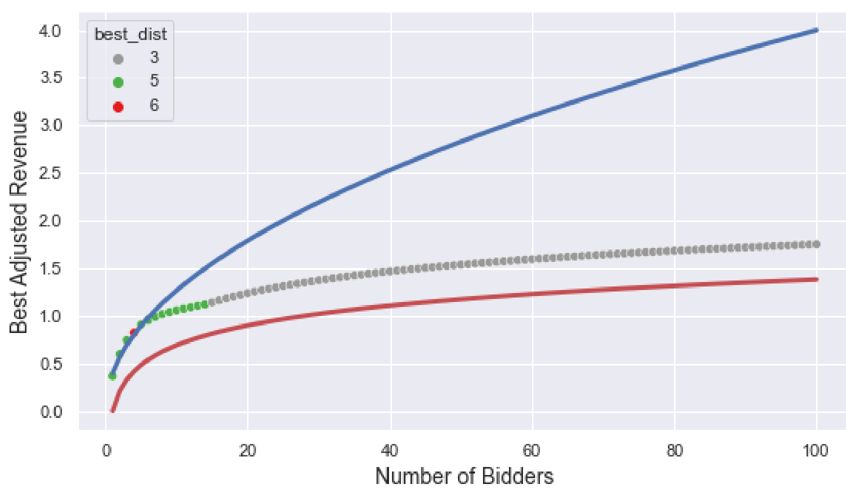


Figure 9: Best adjusted revenue across the discretizations tested, for $m = 2$ items and $n \in [1, 100]$ bidders. Here, label 3 is the discretization in Table 2, label 5 is the discretization in Table 4, and label 6 is the discretization in Table 7.

4.3 Discrete to continuous revenue bounds

With a few well-performing discretizations in hand and the issue of truncation resolved, we then sought to understand how much revenue we lose from the process of discretization, since $\mathcal{ER}_t$ stochastically dominates $\mathcal{ER}_{t,d}$. Our key tool is a bound on the revenue from a product distribution $\mathcal{D}^+$ that stochastically dominates another product distribution $\mathcal{D}$ in terms of $\textsc{Rev}(\mathcal{D})$ and the expected difference between $\mathcal{D}^+$ and $\mathcal{D}$, via a reduction from a mechanism for $\mathcal{D}$ to a mechanism for $\mathcal{D}^+$ due to [RW15].

In this section, we use the following notation (slightly different from the notation used throughout the rest of this paper): let $\phi_j(v)$ denote the (random) allocation awarded to bidder j when reporting type v; we also abuse notation by letting $v(\phi)$ denote the expected utility (over the randomness in the mechanism and the other bidders reporting their true types), that a bidder with type v gains from a random allocation ϕ. Let $p_j(v)$ denote the expected price paid by bidder j when reporting type v (over the same randomness). Further, $\delta_j(\cdot)$ denotes the random function $v_j^+(\cdot) - v_j(\cdot)$ when couples v^+ and v are sampled jointly from $\mathcal{D}^+$ and $\mathcal{D}$. We again abuse notation and refer to δ_j as the distribution over $\delta_j(\cdot)$ as well (allowing us to write terms like $\textsc{Val}(\delta)$). Lastly, for v drawn from $\mathcal{D}^+$, we denote v's couple from $\mathcal{D}$ by v'.

The reduction proceeds in two phases. Informally, the first phase is a *surrogate sale*, in which buyers from $\mathcal{D}^+$ pay to be represented in the auction by a *surrogate* from $\mathcal{D}$. The *replicas* for a bidder j are bidder j's "competition" for buying the surrogates, which allows the seller to decide how to price the surrogates appropriately. The second phase is the *surrogate competition*, in which the surrogates participate in the auction. Finally, each bidder (from $\mathcal{D}^+$) is awarded their surrogate's allocation and pays the price charged to their surrogate, plus the price of their surrogate from the surrogate sale.

Phase 1, Surrogate Sale:

(1) Let M be any BIC mechanism for buyers from $\mathcal{D}$. Multiply all prices charged by M by $(1-\varepsilon)$ and call the new mechanism M^ε. Interpret the ε fraction of prices given back as *rebates*.

(2) For each bidder j, create $r-1$ *replicas* sampled i.i.d. from $\mathcal{D}_j^+$ and r *surrogates* sampled i.i.d. from $\mathcal{D}_j$. Let $r \to \infty$.

(3) Ask each bidder to report their value $v_j(\cdot)$.

(4) Create a weighted bipartite graph with bidder j and the $r-1$ replicas on the left and the r surrogates on the right. The weight of an edge between a replica (or bidder j) with type $r_j(\cdot)$ and surrogate of type $s_j(\cdot)$ is the utility of r_j for the expected outcome of M^ε when reporting s_j, which is $r_j(\phi_j^\varepsilon(s_j)) - p_j^\varepsilon(s_j)$.

(5) Compute a maximum perfect matching and VCG prices in this bipartite graph; henceforth refer to it as the *VCG matching*. If a replica (or bidder j) is unmatched in the VCG matching (for instance if all edges incident to some replica have negative weight), add an edge to a random unmatched surrogate. The surrogate selected for bidder j is their match.

Phase 2, Surrogate Competition

(1) Let s_j denote the surrogate chosen to represent bidder j in Phase 1, and let $\vec{s}$ denote the entire profile of surrogates (the ones matched to real buyers). Have the surrogates participate in M^ε.

(2) If bidder j was matched to their surrogate through VCG, charge them the VCG price and award them $M_j^\varepsilon(\vec{s})$ (recall that this auction outcome consists of an allocation and a price; the price is added onto the VCG price). If bidder j was matched to a random surrogate after VCG, award them nothing and charge them nothing.

Theorem 14 (Implicit in [RW15])**.** *Let M' denote the mechanism obtained by the above reduction, starting from any BIC mechanism M for buyers from $\mathcal{D}$. Then M' is BIC for bidders from $\mathcal{D}^+$, and for any $\varepsilon \in (0,1)$, we have*

$$\textsc{Rev}^{M'}(\mathcal{D}^+) \geq (1-\varepsilon)\left(\textsc{Rev}^{M}(\mathcal{D}) - \frac{1}{\varepsilon}\mathbb{E}\left[\frac{1}{r}\sum_j \sum_{r_j} \delta_j(\phi_j^M(s_j))\right]\right).$$

Here, our major contribution is applying Theorem 14 to $\mathcal{D}^+ = \left(\mathcal{ER}_t^2\right)^n$, $\mathcal{D} = \left(\mathcal{ER}_{t,d}^2\right)^n$, and M the mechanism output by RoaSolver. In particular, we explicitly evaluate the "error term" (the expected value term subtracted from $\textsc{Rev}^M(\mathcal{D})$) in terms of the parameters of $\mathcal{D}$.

Lemma 15 (Lower bound on $\textsc{Rev}_n(\mathcal{ER}_t^2)$)**.** *Let $\mathcal{ER}_{t,d}$ have support $\{x_1 = 1, x_2, \dots, x_s = t\}$. Then for any $\varepsilon \in (0,1)$, we have*

$$\textsc{Rev}_n(\mathcal{ER}_t^2) \geq (1-\varepsilon)\left(\textsc{Rev}_n(\mathcal{ER}_{t,d}^2) - \frac{2n}{\varepsilon}\sum_{k=1}^{s-1}\sum_{\ell=1}^{s-1}\phi_1(x_k, x_\ell)\left(\frac{1}{x_\ell} - \frac{1}{x_{\ell+1}}\right)\log\frac{x_{k+1}}{x_k}\right).$$

Proof. Our goal is to find an explicit formula for the expectation $\mathbb{E}\left[\frac{1}{r}\sum_j\sum_{r_j}\delta_j(\phi_j^M(s_j))\right]$, where the expectation is taken over the sampling of the j bidders with values from $\mathcal{ER}_t^2$ and the randomness in the mechanism. Observe that we have

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{r}\sum_j\sum_{r_j}\delta_j(\phi_j^M(s_j))\right] &= \mathbb{E}\left[\sum_{j=1}^n \frac{1}{r}\sum_{r_j}\delta_j(\phi_j^M(s_j))\right] \\
&= \mathbb{E}\left[\sum_{j=1}^n \frac{1}{r}\sum_{r_j}\delta_j(\phi_j^M(r_j'))\right] \\
&= \mathbb{E}\left[\sum_{j=1}^n \mathbb{E}_{\vec{v}\leftarrow\mathcal{ER}_t^2}\left[\delta_j(\phi_j^M(\vec{v}'))\right]\right] \\
&= n\mathbb{E}_{\vec{v}\leftarrow\mathcal{ER}_t^2}\left[\delta_j(\phi_j^M(\vec{v}'))\right].
\end{aligned}$$

Here, the first line is just a rearrangement, the second line holds because as $r \to \infty$, s_j concentrates around the type that is "close" to r_j, which is r_j' [RW15], the third line follows from the definition of expectation over replicas as we take $r \to \infty$, and the last line follows by linearity of expectation since the n buyers are i.i.d. (and is just written with respect to some bidder j).

Since values are additive and the items are i.i.d., we have

$$\begin{aligned}
\mathbb{E}_{\vec{v}\leftarrow\mathcal{ER}_t^2}\left[\delta_j(\phi_j^M(\vec{v}'))\right] &= \mathbb{E}_{\vec{v}\leftarrow\mathcal{ER}_t^2}\left[(\vec{v}-\vec{v}')\cdot\phi_j^M(\vec{v}')\right] \\
&= \int_{v_1=1}^t\int_{v_2=1}^t (\vec{v}-\vec{v}')\cdot\phi_j^M(v_1', v_2')\mathbb{P}(v_1,v_2)dv_1dv_2 \\
&= 2\int_{v_1=1}^t\int_{v_2=1}^t (v_1 - v_1')\phi_1(v_1', v_2')\frac{1}{v_1^2v_2^2}dv_1dv_2,
\end{aligned}$$

where $\phi_1(\vec{v})$ now denotes the probability that a bidder wins item 1 when bidding $\vec{v}'$ in mechanism M (the solution output by RoaSolver). Now, note that $v' = x_k$ for $v \in [x_k, x_{k+1})$ and that $\phi_1(x_k, x_\ell)$ is just a constant, so we have

$$\begin{aligned}\int_{v_1=x_k}^{x_{k+1}}\int_{v_2=x_\ell}^{x_{\ell+1}}(v_1-v_1')\phi_1(v_1',v_2')\frac{1}{v_1^2v_2^2}dv_1dv_2 &= \int_{v_1=x_k}^{x_{k+1}}\int_{v_2=x_\ell}^{x_{\ell+1}}(v_1-x_k)\phi_1(x_k,x_\ell)\frac{1}{v_1^2v_2^2}dv_1dv_2\\ &= \phi_1(x_k,x_\ell)\int_{v_1=x_k}^{x_{k+1}}\int_{v_2=x_\ell}^{x_{\ell+1}}\left(\frac{1}{v_1v_2^2}-\frac{x_k}{v_1^2v_2^2}\right)dv_1dv_2\\ &= \phi_1(x_k,x_\ell)\int_{v_1=x_k}^{x_{k+1}}\left(\frac{1}{v_1x_\ell}-\frac{1}{v_1x_{\ell+1}}\right)dv_1\\ &= \phi_1(x_k,x_\ell)\left(\frac{1}{x_\ell}-\frac{1}{x_{\ell+1}}\right)\log\frac{x_{k+1}}{x_k}.\end{aligned}$$

Finally, summing over $1\le k,\ell<s$ gives the desired integral. Putting everything together, we get

$$\mathbb{E}\left[\frac{1}{r}\sum_j\sum_{r_j}\delta_j(\phi_j(s_j))\right]=2n\sum_{k=1}^{s-1}\sum_{\ell=1}^{s-1}\phi_1(x_k,x_\ell)\left(\frac{1}{x_\ell}-\frac{1}{x_{\ell+1}}\right)\log\frac{x_{k+1}}{x_k},$$

and the result follows from observing that $\textsc{Rev}_n(\mathcal{ER}^2_{t,d})=\textsc{Rev}_n^M(\mathcal{ER}^2_{t,d})$ since RoaSolver computes the optimal auction, and also $\textsc{Rev}_n(\mathcal{ER}^2_t)\ge\textsc{Rev}_n^{M'}(\mathcal{ER}^2_t)$. □

Unfortunately, for the discretizations we tested, the lower bound obtained by optimizing over $\varepsilon\in(0,1)$ for every n did not improve upon the $\log n/20$ lower bound. Discretizations with an extremely high point in the support (such as in Table 5) gave strictly negative lower bounds on $\textsc{Rev}_n(\mathcal{ER}^2_t)$, while discretizations without an extremely high point (such as the initial discretization, Table 1) gave a lower bound of approximately 0. This can be understood by observing that the error term in Lemma 15 is multiplied by a factor of $\frac{1}{\varepsilon}$, so in order to obtain meaningful bounds, we need the error term to be very small. But recall that the error term is essentially a function of the expected difference between $\mathcal{ER}^2_t$ and $\mathcal{ER}^2_{t,d}$, so we need the discretization to match the continuous distribution very closely. So far, this does not seem feasible with a support size of only $s=14$. But if we are able to bypass the limitation imposed by RoaSolver's runtime and perform these experiments with a much finer discretization, this framework may be able to provide a more meaningful lower bound on $\textsc{Rev}_n(\mathcal{ER}^2_t)$, and consequently on $\textsc{ARev}_n(\mathcal{ER}^2)$.

Nevertheless, we pressed ahead with applying Theorem 14 to also formulate an upper bound on $\textsc{Rev}_n(\mathcal{ER}^2_t)$. The key observation is that stochastic dominance is not used in the reduction at all (it is only required for later results in [RW15]), so we can simply reverse the roles of $\mathcal{ER}^2_{t,d}$ and $\mathcal{ER}^2_t$ as $\mathcal{D}^+$ and $\mathcal{D}$.

Lemma 16 (Upper bound on $\textsc{Rev}_n(\mathcal{ER}^2_t)$)**.** *Let $\mathcal{ER}_{t,d}$ have support $\{x_1=1,x_2,\dots,x_s=t\}$. Then for any $\varepsilon\in(0,1)$, we have*

$$\textsc{Rev}_n(\mathcal{ER}^2_t)\le\frac{1}{1-\varepsilon}\textsc{Rev}_n(\mathcal{ER}^2_{t,d})+\frac{2n}{\varepsilon}\sum_{k=1}^{s-1}\sum_{\ell=1}^{s-1}\left(\frac{1}{x_\ell}-\frac{1}{x_{\ell+1}}\right)\log\frac{x_{k+1}}{x_k}.$$

Proof. Now, M is the optimal mechanism for buyers from $\mathcal{ER}^2_t$ and M' is the mechanism obtained from the reduction, for buyers from $\mathcal{ER}^2_{t,d}$. Then $\textsc{Rev}_n^M(\mathcal{ER}^2_t)=\textsc{Rev}_n(\mathcal{ER}^2_t)$ and $\textsc{Rev}_n^{M'}(\mathcal{ER}^2_{t,d})\le\textsc{Rev}_n(\mathcal{ER}^2_{t,d})$

(which we recall is the revenue of the solution output by RoaSolver), so we have

$$\begin{aligned}
\textsc{Rev}_n^{M'}(\mathcal{ER}^2_{t,d}) &\geq (1-\varepsilon)\left(\textsc{Rev}_n^{M}(\mathcal{ER}^2_t) - \frac{1}{\varepsilon}\mathbb{E}\left[\frac{1}{r}\sum_j\sum_{r_j}\delta_j(\phi_j^M(s_j))\right]\right)\\
\textsc{Rev}_n(\mathcal{ER}^2_{t,d}) &\geq (1-\varepsilon)\left(\textsc{Rev}_n(\mathcal{ER}^2_t) - \frac{1}{\varepsilon}\mathbb{E}\left[\frac{1}{r}\sum_j\sum_{r_j}\delta_j(\phi_j^M(s_j))\right]\right)\\
\textsc{Rev}_n(\mathcal{ER}^2_t) &\leq \frac{1}{1-\varepsilon}\textsc{Rev}_n(\mathcal{ER}^2_{t,d}) + \frac{1}{\varepsilon}\mathbb{E}\left[\frac{1}{r}\sum_j\sum_{r_j}\delta_j(\phi_j^M(s_j))\right].
\end{aligned}$$

Unfortunately, we now do not have an explicit allocation ϕ^M since we do not know the optimal mechanism for M, and the best we can do is use the bound $\phi_1(\vec{v}) \leq 1$. Using this bound and following calculations in the same style as Lemma 15 furnishes the desired result. ☐

Again, the discretizations we experimented with resulted in linear upper bounds for $\textsc{ARev}_n(\mathcal{ER}^2)$, therefore not improving upon the currently known $9\sqrt{2n}$ upper bound. However, we remain hopeful that a finer discretization will allow success of this framework in proving tighter bounds.

5 Symmetries

This section closely follows the main ideas from [DW11], with special attention paid to the fact that in our particular setting, we have both item and bidder symmetries. In general, it is known that we can write a revenue-optimizing linear program of size polynomial in $|\mathrm{supp}(\mathcal{D})|$ (see [DW11], Appendix B, for an explicit specification of the full LP). However, $\mathrm{supp}(\mathcal{D})$ may be infinite, and when it is finite it is usually exponential in both the number of bidders and the number of buyers (since we specify a value from the distribution for each item, for every bidder).

We seek to use symmetries in the distribution $\mathcal{D}$ to reason about symmetries in the optimal mechanism M, and use this to reduce the number of variables needed to fully describe the structure of M in the LP. Then, with a smaller LP, we will be able to run RoaSolver with finer discretizations and hopefully utilize the framework from Section 4 to prove tighter bounds. We begin with some relevant definitions and notation.

5.1 Preliminaries and notation

Recall that the seller has m items and faces n buyers. Denote by v_{ij} bidder i's value for item j, so that $\vec{v}_i := (v_{ij})_{j\in[m]}$ is bidder i's *type*, and $\vec{v} = (\vec{v}_i)_{i\in[n]} = (v_{ij})_{i\in[n],j\in[m]}$ is the *complete bidder profile*. Let $\pi_{ij}(\vec{v}_i)$ be the *interim probability* that bidder i of type $\vec{v}_i$ is awarded item j, $q_i(\vec{v}_i)$ be the *expected price* charged to bidder i of type $\vec{v}_i$, and $U(\vec{v}_i, M_i(\vec{w}))$ be the *utility* of a bidder with type $\vec{v}_i$ for the expected outcome $M_i(\vec{w})$.

Denote by S_n, S_m the symmetric groups over the sets $[n] := \{1,\cdots,n\}$ and $[m] := \{1,\ldots,m\}$ respectively, and for $\sigma = (\sigma_1,\sigma_2) \in S_n \times S_m$, assume that σ maps $(i,j) \in [n]\times[m]$ to $\sigma(i,j) := (\sigma_1(i),\sigma_2(j))$. For a complete bidder profile $\vec{v}$, we define $\sigma(\vec{v})$ to be the value vector $\vec{w}$ such that $w_{\sigma(i,j)} = v_{i,j}$ for all i,j. For a value distribution $\mathcal{D}$, we define $\sigma(\mathcal{D})$ to be the distribution that first draws $\vec{v}$ from $\mathcal{D}$ and then outputs $\sigma(\vec{v})$.

We can also formally define the notion of symmetries for distributions and mechanisms.

Definition 17 (Symmetry in a distribution). *We say that a distribution $\mathcal{D}$ has symmetry $\sigma \in S_n \times S_m$ if, for all $\vec{v} \in \mathbb{R}^{n \times m}$, $\mathbb{P}_{\mathcal{D}}(\vec{v}) = \mathbb{P}_{\mathcal{D}}(\sigma(\vec{v}))$.*

Definition 18 (Symmetry in a mechanism). *We say that a mechanism M respects symmetry $\sigma \in S_n \times S_m$ if, for all $\vec{v} \in \mathbb{R}^{n \times m}$, $M(\sigma(\vec{v})) = \sigma(M(\vec{v}))$.*

Definition 19 (Permutation of a mechanism). *For any permutation $\sigma \in S_n \times S_m$ and any mechanism M, define $\sigma(M)$, the permutation of M, as $[\sigma(M)](\vec{v}) = \sigma(M(\sigma^{-1}(\vec{v})))$.*

5.2 Existence of an optimal symmetric mechanism

Our key result is an analog of Nash's theorem for symmetric games, defined for randomized mechanisms. Informally, we show that every symmetric distribution $\mathcal{D}$ has a symmetric BIC optimal mechanism M, corresponding to a symmetric Bayes-Nash equilibrium in which every bidder reports their true value.

Theorem 20. *For all $\mathcal{D}$, any BIC mechanism M can be symmetrized into a BIC mechanism M' such that for all $\sigma \in S_n \times S_m$, if $\mathcal{D}$ has symmetry σ, then M' respects σ, and $\text{REV}^M(\mathcal{D}) = \text{REV}^{M'}(\mathcal{D})$.*

We break the proof into three smaller lemmas.

Lemma 21. *If M is a BIC mechanism, then for any $\sigma \in S_n \times S_m$ that $\mathcal{D}$ has, the mechanism $\sigma(M)$ is also BIC, and $\text{REV}^M(\mathcal{D}) = \text{REV}^{\sigma(M)}(\mathcal{D})$.*

Proof of Lemma 21. Observe that by definition of $\sigma(M)$, on input $\vec{v}$ (a complete bidder profile), bidder i is allocated the permutation σ of the lottery offered to bidder $\sigma^{-1}(i)$ by M on input $\sigma^{-1}(\vec{v})$, and charged the price charged to bidder $\sigma^{-1}(i)$ by M on input $\sigma^{-1}(\vec{v})$. Then, for all $i, \vec{v}_i$, we have

$$\mathbb{E}_{\vec{v}_{-i} \sim \mathcal{D}_{-i}}\left[U(\vec{v}_i, [\sigma(M)]_i(\vec{v}))\right] = \mathbb{E}_{\vec{v}_{-i} \sim \mathcal{D}_{-i}}\left[U(\sigma^{-1}(\vec{v}_i), M_{\sigma^{-1}(i)}(\sigma^{-1}(\vec{v})))\right].$$

Since M is BIC, we also have, for all $\vec{v}_i, \vec{w}_i$,

$$\begin{aligned}\mathbb{E}_{\vec{v}_{-i} \sim \mathcal{D}_{-i}}\left[U(\sigma^{-1}(\vec{v}_i), M_{\sigma^{-1}(i)}(\sigma^{-1}(\vec{v})))\right] &\geq \mathbb{E}_{\vec{v}_{-i} \sim \mathcal{D}_{-i}}\left[U(\sigma^{-1}(\vec{v}_i), M_{\sigma^{-1}(i)}(\sigma^{-1}(\vec{w}_i; \vec{v}_{-i})))\right] \\ &= \mathbb{E}_{\vec{v}_{-i} \sim \mathcal{D}_{-i}}\left[U(\vec{v}_i, [\sigma(M)]_i(\vec{w}_i; \vec{v}_{-i}))\right].\end{aligned}$$

Putting these together, we see that

$$\mathbb{E}_{\vec{v}_{-i} \sim \mathcal{D}_{-i}}\left[U(\vec{v}_i, [\sigma(M)]_i(\vec{v}))\right] \geq \mathbb{E}_{\vec{v}_{-i} \sim \mathcal{D}_{-i}}\left[U(\vec{v}_i, [\sigma(M)]_i(\vec{w}_i; \vec{v}_{-i}))\right]$$

for all $i, \vec{v}_i, \vec{w}_i$, so $\sigma(M)$ is BIC.

Now, since all bidders in $\sigma(M)$ play truthfully, the expected revenue of $\sigma(M)$ on complete bidder profile $\vec{v}$

is exactly $\textsc{Rev}^M(\sigma^{-1}(\vec{v}))$, so we can compute

$$\begin{aligned}
\textsc{Rev}^{\sigma(M)}(\mathcal{D}) &= \mathbb{E}_{\vec{v}\in \mathrm{supp}(D)}\left[\textsc{Rev}^M(\sigma^{-1}(\vec{v}))\right] \\
&= \sum_{\vec{v}\in\mathrm{supp}(\mathcal{D})} \textsc{Rev}^M(\sigma^{-1}(\vec{v}))\cdot \mathbb{P}_{\mathcal{D}}(\vec{v}) \\
&= \sum_{\vec{v}\in\mathrm{supp}(\mathcal{D})} \textsc{Rev}^M(\sigma^{-1}(\vec{v}))\cdot \mathbb{P}_{\mathcal{D}}(\sigma^{-1}(\vec{v})) \\
&= \sum_{\sigma^{-1}(\vec{v})\in\mathrm{supp}(\mathcal{D})} \textsc{Rev}^M(\vec{v})\cdot \mathbb{P}_{\mathcal{D}}(\vec{v}) \\
&= \sum_{\vec{v}\in\mathrm{supp}(\mathcal{D})} \textsc{Rev}^M(\vec{v})\cdot \mathbb{P}_{\mathcal{D}}(\vec{v}) \\
&= \textsc{Rev}^M(\mathcal{D}),
\end{aligned}$$

where we moved from the second to the third line using the definition of symmetry in a distribution, and from the fourth to the fifth line using the fact that permutations are bijective mappings. ☐

Lemma 22. *Let $\mathcal{G}$ denote any distribution over elements of $S_n \times S_m$. For a BIC mechanism M, let $\mathcal{G}(M)$ denote the randomized mechanism that samples an element σ from $\mathcal{G}$ and then uses the mechanism $\sigma(M)$. Then, for all $\mathcal{G}$, $\mathcal{G}(M)$ is BIC. Further, if $\mathcal{G}$ samples only σ such that $\mathcal{D}$ has symmetry σ, then $\textsc{Rev}^M(\mathcal{D}) = \textsc{Rev}^{\mathcal{G}(M)}(\mathcal{D})$.*

Proof of Lemma 22. By Lemma 21, each $\sigma(M)$ is a BIC mechanism, so randomly sampling from a set of BIC mechanisms also results in a BIC mechanism (a bidder cannot possibly improve their utility in any outcome by deviating from reporting their true value), proving the first part of the lemma. For the second part, we can compute

$$\begin{aligned}
\textsc{Rev}^{\mathcal{G}(M)}(\mathcal{D}) &= \sum_{\sigma\in\mathrm{supp}(\mathcal{G})} \textsc{Rev}^{\sigma(M)}(\mathcal{D})\cdot \mathbb{P}_{\mathcal{G}}(\sigma) \\
&= \sum_{\sigma\in\mathrm{supp}(\mathcal{G})} \textsc{Rev}^{M}(\mathcal{D})\cdot \mathbb{P}_{\mathcal{G}}(\sigma) \\
&= \textsc{Rev}^{M}(\mathcal{D}) \sum_{\sigma\in\mathrm{supp}(\mathcal{G})} \mathbb{P}_{\mathcal{G}}(\sigma) \\
&= \textsc{Rev}^{M}(\mathcal{D}),
\end{aligned}$$

where we moved from the second line to the third line by a direct application of Lemma 21. ☐

Lemma 23. *Let $\mathcal{G}$ sample a permutation uniformly at random from a subgroup S of $S_n \times S_m$. Then $\mathcal{G}(M)$ respects every permutation in S.*

Proof of Lemma 23. For any $\vec{v}$, by definition we have $\mathcal{G}(M)(\vec{v}) = \frac{1}{|S|}\sum_{\rho\in S}\rho(M(\rho^{-1}(\vec{v})))$. Now, since S is a subgroup (and thus closed under multiplication and inverses), for any $\sigma \in S$, we have

$$
\begin{aligned}
\mathcal{G}(M)(\sigma(\vec{v})) &= \frac{1}{|S|} \sum_{\sigma\rho\in S} (\sigma\rho)(M((\sigma\rho)^{-1}(\sigma(\vec{v})))) \\
&= \frac{1}{|S|} \sum_{\rho\in S} \sigma\rho(M(\rho^{-1}(\vec{v}))) \\
&= \sigma\left(\frac{1}{|S|} \sum_{\rho\in S} \rho(M(\rho^{-1}(\vec{v})))\right) \\
&= \sigma(\mathcal{G}(M))(\vec{v}).
\end{aligned}
$$

That is, $\mathcal{G}(M)$ respects every $\sigma \in S$. □

Proof of Theorem 20. If $\mathcal{D}$ has symmetries σ and τ, then applying the definition of symmetry twice gives

$$\mathbb{P}_{\mathcal{D}}(\sigma\tau(\vec{v})) = \mathbb{P}_{\mathcal{D}}(\tau(\vec{v})) = \mathbb{P}_{\mathcal{D}}(\vec{v}),$$

so $\mathcal{D}$ has symmetry $\sigma\tau$. Also, applying the definition of symmetry using vector $\sigma^{-1}(\vec{v})$ gives

$$\mathbb{P}_{\mathcal{D}}(\sigma^{-1}(\vec{v})) = \mathbb{P}_{\mathcal{D}}(\vec{v}),$$

so $\mathcal{D}$ also has symmetry σ^{-1}. Thus the set of symmetries of $\mathcal{D}$ is a subgroup, and we can apply the three lemmas to obtain a symmetric BIC mechanism M' that respects all symmetries of $\mathcal{D}$ and obtains the same revenue as M. □

In our setting, the bidders are i.i.d., so there exists a BIC revenue-optimal mechanism with $\pi_{ij}(\vec{v}) = \pi_{i'j}(\vec{v})$ for all bidders i, i', bidder types $\vec{v}_i = \vec{v}_i{}' = \vec{v}$, and items j. In other words, any two bidders who report the same type receive the same allocation, so in the symmetric LP we only need to keep track of allocation and price variables for bidder 1 rather than for all n bidders. Additionally, the items are i.i.d., so $\pi_{ij}(\vec{v}_i) = \pi_{i\sigma(j)}(\sigma(\vec{v}_i))$ for all bidders i, items j, and item permutations $\sigma \in S_m$. In other words, relabeling the items according to permutation σ does not change the optimal mechanism.

5.3 Strong monotonicity of a BIC mechanism

Theorem 24. *If $\mathcal{D}$ is item-symmetric, then every item-symmetric BIC mechanism is* strongly monotone*:*

$$\text{for all bidders } i \text{ and items } j, j'\text{: } v_{ij} \geq v_{ij'} \implies \pi_{ij}(\vec{v}_i) \geq \pi_{ij'}(\vec{v}_i).$$

Proof. Suppose for the sake of contradiction that M is item-symmetric and BIC but not strongly monotone. Then there exists some bidder type $\vec{v}_i{}^*$ and items $j \neq j'$ such that

$$v^*_{ij} < v^*_{ij'}, \quad \pi_{ij}(\vec{v}_i{}^*) > \pi_{ij'}(\vec{v}_i{}^*).$$

Let σ denote the transposition $(jj') \in S_n$. By item symmetry of $\mathcal{D}$, for complete bidder profiles $\vec{w}$, we have $\mathbb{P}(\vec{w}|\vec{w}_i = \vec{v}_i{}^*) = \mathbb{P}(\sigma(\vec{w})|\vec{w}_i = \sigma(\vec{v}_i{}^*))$. By item symmetry of M, we have $M(\sigma(\vec{w})) = \sigma(M(\vec{w}))$ for all $\vec{w}$. Putting these together, we have

$$\pi_{ij}(\vec{v}_i{}^*) = \pi_{ij'}(\sigma(\vec{v}_i{}^*)), \quad \pi_{ij'}(\vec{v}_i{}^*) = \pi_{ij}(\sigma(\vec{v}_i{}^*)).$$

Then, if bidder i reports $\sigma(\vec{v_i}^*)$ instead of $\vec{v_i}^*$, their value for their allocation changes by

$$
\begin{aligned}
&\left(\pi_{ij}(\sigma(\vec{v_i}^*))v^*_{ij} + \pi_{ij'}(\sigma(\vec{v_i}^*))v^*_{ij'}\right) - \left(\pi_{ij}(\vec{v_i}^*)v^*_{ij} + \pi_{ij'}(\vec{v_i}^*)v^*_{ij'}\right) \\
&= \pi_{ij'}(\vec{v_i}^*)v^*_{ij} + \pi_{ij}(\vec{v_i}^*)v^*_{ij'} - \pi_{ij}(\vec{v_i}^*)v^*_{ij} - \pi_{ij'}(\vec{v_i}^*)v^*_{ij'} \\
&= \left(\pi_{ij}(\vec{v_i}^*) - \pi_{ij'}(\vec{v_i}^*)\right)\left(v^*_{ij'} - v^*_{ij}\right) \\
&> 0.
\end{aligned}
$$

That is, bidder i strictly increases their expected value (and pays the same price), so their expected utility strictly increases by misreporting. However, this contradicts the fact that M is BIC. $\square$

That is, if bidder i likes item j more than item j', their expected probability of getting item j is higher. This is a useful property for simplifying the LP because it implies that a bidder of type $\vec{v_i}$ with $v_{ij} \geq v_{ij'}$ has no incentive to misreport any $\vec{w_i}$ with $w_{ij} < w_{ij'}$. For the two-item setting, this means that if $v_{i1} > v_{i2}$, we only need to consider BIC constraints for bidder types $\vec{w_i}$ with $w_{i1} \geq w_{i2}$.

5.4 Symmetric LP formulation

The bidder symmetry optimization has already been implemented in RoaSolver, but the item symmetry optimization has not yet been implemented because detecting item symmetry in the general case is not a simple problem [Shu19]. For our particular setting, we have added the command-line execution flag `-sym` to ensure symmetric output, but this implementation starts from the longer LP (utilizing reduced forms and bidder symmetry, but not item symmetry) and *adds* equality constraints to enforce symmetry, rather than formulating the fully symmetric LP with *fewer* variables and constraints. Implementing the symmetric LP (both in our particular setting, and in the general case) is a possible direction for future development.

Here, we explicitly formulate the symmetric LP for two symmetric items and n bidders. We build upon [DW11] by also formulating the symmetric LP using reduced forms, which further reduces the size of the LP. Since the bidders are i.i.d., we drop the subscript i when referring to the allocation given to or price charged to a bidder when reporting type $\vec{v_i}$. Finally, let E denote the subset of bidder types where item 1 is the favorite item; that is, $E := \{\vec{v} \in \mathcal{ER}^2_{t,d} \,|\, v_1 \geq v_2\}$, and s be the size of the support of the discretization $\mathcal{ER}_{t,d}$. Note that $|E| = \binom{s}{2} + s = \frac{s(s+1)}{2}$.

The value in parentheses at the end of each line is an upper bound on the number of such variables/constraints.

Variables:

- $\pi_1(\vec{v})$, for all $\vec{v} \in E$: the expected probability (over all other bidders bidding their true types) that a bidder gets item 1 when reporting type $\vec{v}$ ($|E|$).
- $\pi_2(\vec{v})$, for all $\vec{v} \in E$: the expected probability (over all other bidders bidding their true types) that a bidder gets item 2 when reporting type $\vec{v}$ ($|E|$).
- $q(\vec{v})$, for all $\vec{v} \in E$: the price paid by a bidder when reporting type $\vec{v}$ ($|E|$).

Constraints:

- $0 \leq \pi_j(\vec{v}) \leq 1$, for all $\vec{v} \in E, j \in \{1, 2\}$: valid probabilities ($4|E|$).

- $v_1\pi_1(\vec{v}) + v_2\pi_2(\vec{v}) - q(\vec{v}) \geq 0$, for all $\vec{v} = (v_1, v_2) \in E$: ex-interim IR ($|E|$).
- $v_1\pi_1(\vec{v}) + v_2\pi_2(\vec{v}) - q(\vec{v}) \geq v_1\pi_1(\vec{v}') + v_2\pi_2(\vec{v}') - q(\vec{v}')$, for all $\vec{v}, \vec{v}' \in E$: BIC ($|E|^2$).
- $\pi_1(\vec{v}) \geq \pi_2(\vec{v})$, for all $\vec{v} \in E$: strong monotonicity ($|E|$).

Separation oracle:

$\mathfrak{S}$, so that if interim allocation rule π satisfies all other constraints and $\mathfrak{S}(\pi) = \ell$, we add the broken constraint ℓ to the linear program and iterate. The separation oracle $\mathfrak{S}$ ensures that π is feasible; that is, that there is an ex-post allocation rule inducing π.

Using Border's theorem, [CDW17] shows that feasibility holds if and only if all the *Border constraints* are satisfied. We now reformulate the Border constraints using the smaller set of symmetric variables in our LP. First define the sets

$$S_j(x) \coloneqq \left\{ \vec{v} \in \mathcal{ER}^2_{t,d} \;\middle|\; \pi_j(\vec{v}) \sum_{\substack{\vec{v}' \in \mathcal{ER}^2_{t,d} \\ \pi_j(\vec{v}) \geq \pi_j(\vec{v}')}} \mathbb{P}(\vec{v}') \geq x \right\}$$

for all items $j \in \{1, 2\}$, as a function of $x \in \mathbb{R}$. Then, for all $x \in \mathbb{R}$, we require

$$n \sum_{\vec{v} \in S_j(x)} \pi_j(\vec{v}) \cdot \mathbb{P}_{\mathcal{ER}^2_{t,d}}(\vec{v}) \leq 1 - \left(1 - \sum_{\vec{v} \in S_j(x)} \mathbb{P}_{\mathcal{ER}^2_{t,d}}(\vec{v})\right)^n,$$

where the left hand side represents the expected probability that item j is awarded to a bidder with type in $S_j(x)$, and the right hand side represents the probability that at least one of the n bidders has type in $S_j(x)$ (when phrased in this way, it is clear that all the Border constraints must be satisfied for feasibility to hold, but Border's theorem states that satisfying all the Border constraints is also *sufficient* for feasibility).

We now exploit item symmetry by observing that if $\vec{v} = (v_1, v_2) \notin E$, then $\sigma(\vec{v}) = (v_2, v_1) \in E$, and

$$\pi_j(\vec{v}) = \pi_{j'}(\sigma(\vec{v})), \quad \mathbb{P}_{\mathcal{ER}^2_{t,d}}(\vec{v}) = \mathbb{P}_{\mathcal{ER}^2_{t,d}}(\sigma(\vec{v})),$$

so we can write all the Border constraints using only the variables for $\vec{v} \in E$.

Further, note that by the discrete nature of the sets $S_j(x)$, it suffices to check the Border constraints only for the threshold values of x in the set

$$X_j = \left\{ \pi_j(\vec{v}) \sum_{\substack{\vec{v}' \in \mathcal{ER}^2_{t,d} \\ \pi_j(\vec{v}) \geq \pi_j(\vec{v}')}} \mathbb{P}(\vec{v}') \;\middle|\; \vec{v} \in \mathcal{ER}^2_{t,d} \right\}.$$

For every j, $|X_j| \leq s$. Finally, the mechanism is symmetric, so we only need to run the separation oracle for item 1. Thus the separation oracle adds at most s constraints to the LP.

Objective:

$$\max n \cdot \sum_{\vec{v} \in E} \mathbb{P}\left(\bigcup_{\sigma \in S_2} \sigma(\vec{v})\right) \cdot q(\vec{v})$$

Putting everything together, we have a revenue-optimizing LP in size polynomial in $|E|$, which is an improvement on the current RoaSolver LP of size polynomial in $|\text{supp}(\mathcal{ER}_{t,d})|^2$. Since $|E| \approx \frac{s^2}{2}$, implementing this symmetric LP can improve the runtime of RoaSolver by at least a factor of 2.

6 Conclusions

In summary, the gap between the logarithmic and square root dependency of the competition complexity on n remains challenging to close. Our simulations using discretizations of the equal revenue distribution give us a fair amount of confidence that the competition complexity is $\Theta(\log n)$, but we do recognize that our simulations are limited by the (in)accuracy of discretizing. Taking this difference into account, we have found explicit upper and lower bounds on the (continuous) competition complexity as a function of the (discrete) optimal revenue computed by RoaSolver. Although this did not improve upon the currently known bounds on competition complexity for the discretizations we tested, this overall framework can ideally applied to finer discretizations to prove tighter bounds.

One way to enable the use of finer discretizations is to optimize the runtime of RoaSolver to allow for larger support sizes. This can be done by exploiting item and bidder symmetries to reduce the size of the LP representing the mechanism. We have formulated the fully symmetric LP and separation oracle here, but have not yet implemented the detection and use of item symmetries in the RoaSolver codebase; this is a possible area for future development.

Presuming that the competition complexity is indeed $\Theta(\log n)$, yet another possible approach relies on the direct construction of an approximately-optimal mechanism from an optimal mechanism for bidders from $\mathcal{ER}^2$. Informally, an outline of the reduction provided by [KMS+19] is:

(1) Ignore buyers with *two* high values (above some cutoff H).

(2) Simplify expensive menu options costing at least $E > 2H$ (possibly higher), noting that a buyer must have at least one high value to purchase an expensive option.

(3) Trim the inexpensive options to reduce menu complexity while preserving supremum probability p of awarding each item.

We seek to lose only an additive $\mathcal{O}(\frac{\log n}{n})$ fraction of revenue per bidder in each step of the reduction. Then, we can interpret the resulting mechanism M' as allocating a p fraction of each item to buyers with two low values (and optimizing revenue over this), and then just selling separately the remaining $1-p$ fraction of each item (which will end up getting purchased by the buyers with one high value) to get revenue $2n(1-p)$. If we can show that M' still achieves total expected revenue $2n$, then we will be able to conclude that $\text{AR}\textsc{ev}_n(\mathcal{ER}^2) = \mathcal{O}(\log n)$, thus furnishing the desired logarithmic upper bound on competition complexity.

References

[Bar20] Rebecca Barber. Bounding the competition complexity for additive buyers over two independent items, 2020.

[BW18] Hedyeh Beyhaghi and S. Matthew Weinberg. Optimal (and benchmark-optimal) competition complexity for additive buyers over independent items. *CoRR*, abs/1812.01794, 2018.

[CDW16] Yang Cai, Nikhil R. Devanur, and S. Matthew Weinberg. A duality based unified approach to bayesian mechanism design. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, page 926–939, New York, NY, USA, 2016. Association for Computing Machinery.

[CDW17] Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. A constructive approach to reduced-form auctions with applications to multi-item mechanism design, 2017.

[DDT12] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. The complexity of optimal mechanism design. *CoRR*, abs/1211.1703, 2012.

[DDT17] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. Strong duality for a multiple-good monopolist. *Econometrica*, 85(3):735–767, 2017.

[DW11] Constantinos Daskalakis and S. Matthew Weinberg. On optimal multi-dimensional mechanism design, 2011.

[HN17] Sergiu Hart and Noam Nisan. Approximate revenue maximization with multiple items. *Journal of Economic Theory*, 172(C):313–347, 2017.

[HR15] Sergiu Hart and Philip J. Reny. Maximal revenue with multiple goods: Nonmonotonicity and other observations. *Theoretical Economics*, 10(3):893–922, 2015.

[KMS+19] Pravesh Kothari, Divyarthi Mohan, Ariel Schvartzman, Sahil Singla, and S. Matthew Weinberg. Approximation schemes for a buyer with independent items via symmetries. *CoRR*, abs/1905.05231, 2019.

[Mye81] Roger B. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1):58–73, February 1981.

[RW15] Aviad Rubinstein and S. Matthew Weinberg. Simple mechanisms for a combinatorial buyer and applications to revenue monotonicity. *CoRR*, abs/1501.07637, 2015.

[Shu19] Mel Y. Shu. Computation of optimal auctions, 2019.