Limited and Online Supply and the Bayesian Foundations of Prior-free Mechanism Design

Nikhil R. Devanur
Microsoft Research
One Microsoft Way
Redmond, WA
nikdev@microsoft.com

Jason D. Hartline*
Electrical Engineering and Computer Science
Northwestern University
Evanston, IL
hartline@eecs.northwestern.edu

ABSTRACT

We study auctions for selling a limited supply of a single commodity in the case where the supply is known in advance and the case it is unknown and must be instead allocated in an online fashion. The latter variant was proposed by Mahdian and Saberi [12] as a model of an important phenomena in auctions for selling Internet advertising: advertising impressions must be allocated as they arrive and the total quantity available is unknown in advance. We describe the Bayesian optimal mechanism for these variants and extend the random sampling auction of Goldberg et al. [8] to address the prior-free case.

Categories and Subject Descriptors

F.2.m [**ANALYSIS OF ALGORITHMS AND PROBLEM COMPLEXITY**]: Miscellaneous

General Terms

Algorithms, Economics

Keywords

Online, Auctions, Prior-free, Mechanism design, Limited supply

1. INTRODUCTION

Consider a prior-free mechanism designer looking for a mechanism with good profit. Does limited supply pose an additional challenge over unlimited supply? Does online supply pose a challenge over offline supply? In attempt to answer the first question Fiat et al. [6] gave a simple approximation-preserving reduction from limited to unlimited supply auctions. *Their answer: no.* In attempt to answer the second question Mahdian and Saberi [12] solved an online pricing problem and with it adopted the auction from [8] to solve the online case, though, at significant loss in performance. *Their answer: yes.* We reconsider both of these questions through a prior-free methodology that is rigorously grounded in the Bayesian mechanism design theory. *Our answers are the opposite!*

Consider the same problems but from the perspective of a Bayesian designer. Suppose the agents' valuations are drawn i.i.d. from a known distribution. What is the optimal mechanism for the three cases: unlimited supply, (offline) limited supply, and online limited supply? For unlimited supply, the optimal mechanism is simply to post the price that maximizes the price times the probability that an agent would buy (recall, all agents are identically distributed). Notice that this price optimization (i.e., revenue curve) may not be concave.

For limited supply and a concave revenue curve, Myerson showed that the optimal mechanism is the Vickrey auction with a reservation price set to the same price as in the unlimited supply case [13]. In the non-concave case Myerson gives a technique that obtains the same revenue as if the revenue curve was the concave hull of the actual revenue curve. To do this the auction applies a distribution dependent weakly monotone transformation of agents' valuations to get *ironed virtual valuations* and the optimal auction allocates to the agents with the highest virtual valuations (breaking ties arbitrarily or randomly, but not by the agents' valuations). Notice that transformation is distribution dependent but not supply dependent. Furthermore, the breaking of valuation space into equivalence classes (of equal virtual valuation) makes the limited supply problem conceptually different from the unlimited supply problem.

For online supply, the designer can make the same transformation from valuations to virtual valuations and then assign the units as they arrive to the agents with the highest virtual valuations (again breaking ties in virtual valuations arbitrarily or randomly, but not by the agents' valuations). Proceeding thusly, the online Bayesian designer will make the exact same allocation as the offline Bayesian designer. By *revenue equivalence* [13], the payment rule is implied by the allocation rule and the two scenarios are equivalent. Thus, our online designer faces no complication rising from the online nature of the supply.

Our Bayesian designer must conclude: *unlimited supply is conceptually easier than (offline) limited supply, but there is*

*Supported in part by NSF CAREER Award CCF-0846113.

EC'09, July 6–10, 2009, Stanford, California, USA.

no conceptual difficulty in online supply over offline supply. The primary focus of this paper is in deriving the same result in the prior-free case (which is the opposite conclusion of Fiat et al. [6] and Mahdian and Saberi [12]).

In prior-free mechanism design the performance of a mechanism is compared to a (distribution independent) *performance benchmark*. The mechanism that performs best relative to, i.e., minimizes the maximum ratio to, this benchmark is the *prior-free optimal* mechanism (for the given benchmark). Hartline and Roughgarden [11] propose as a benchmark the performance of the *optimal Bayesian optimal* mechanism. This benchmark has the strong consequence that a mechanism that approximates it is guaranteed to simultaneously approximate, for all i.i.d. distributions, the performance of the Bayesian optimal mechanism for that distribution. For unlimited supply where the Bayesian optimal mechanism is a posted price, this benchmark is simply the *optimal posted-price* profit. Coincidentally this is exactly the benchmark chosen for prior-free unlimited supply auctions (modulo one small technical detail: the restriction to prices for which there are at least two units sold) [6]. Fiat et al. generalized the unlimited supply benchmark to limited supply settings in a natural way by considering the *optimal posted-price profit selling at least two units and at most the full supply.* Though natural, this definition does not coincide with the benchmark of the optimal Bayesian optimal profit; in fact, it can be up to but at most a factor of two off (as we prove). Precisely for this reason the lossless reduction from the limited to unlimited supply proposed in [6] fails. Thus, Fiat et al.'s results can be interpreted as giving four and eight approximations for the unlimited and limited supply problems, respectively (against the optimal Bayesian optimal benchmark). Hartline and McGrew's 3.25–approximation for the unlimited supply problem [10] gives a 6.5–approximation to the limited supply problem.

These prior-free auctions cannot be easily adapted to the case of online supply because the number of units they allocate is a discontinuous function of the supply. Yet, as we already described, the Bayesian optimal auction can easily be adapted to online supply. Thus it seems like these prior-free mechanisms are doing something that is intuitively wrong. To rectify this, we consider the limited supply generalization of the *random sampling optimal price (RSOP)* auction [8]: we randomly partition the agents, then for each partition we sell half the units using the Bayesian optimal mechanism for the distribution given by the agents in the opposite partition. Notice that because this mechanism is derived from a Bayesian optimal auction for limited supply, it inherits the property that it does not need to know the supply limit in advance.

We extend the analysis approach of Alaei et al. [1] for RSOP to this limited supply auction. Our analysis is necessarily more complicated, and we lose slightly in the approximation factor we are able to prove. We prove a 25-approximation which is the best known factor for the online case. Furthermore, we have no reason to believe that the approximation factor is not four as is conjectured for RSOP. Notice that a bound of four would be better than the best known (offline) limited-supply approximation factor of 6.5.

Related Work. Prior-free optimal mechanism design was initiated by Goldberg et al.'s design of the random sampling optimal price (RSOP) auction for selling a single commodity in unlimited supply [8]. The analysis framework was refined by Fiat et al. who showed that RSOP is a constant approximation in worst case relative to a natural benchmark, *the optimal posted price profit selling at least two units* [6]. Feige et al. refine the analysis to improve the constant approximation factor to 15 [5], which was later improved to 4.68 by Alaei et al. [1].

Fiat et al. extend the analysis framework to the case of limited supply in a natural way; though in hindsight this generalization is not as well motivated as the original unlimited supply framework. Hartline and Roughgarden give a methodology for prior-free mechanism design that is based on Bayesian mechanism design that suggests an alternative and well motivated approach to limited supply [11]. We reexamine limited supply profit maximization from this new perspective. (Note that [11] does not consider the objective of profit maximization and their mechanism is simpler to analyze because they do not allow the possibility of online supply.)

Mahdian and Saberi [12] adapt the limited supply auctions from Goldberg et al. [8] to the case where the supply arrives online and must be allocated immediately while payments may be determined offline. They consider a non-game-theoretic online pricing problem where the seller is constrained to sell at a single price to all winners, and gave an algorithm for it with a constant competitive ratio [12]. Devanur and Chakraborty gave another algorithm for the same online pricing problem that improved the competitive ratio to a factor of two [4].

It is crucial in our problem of online supply that the mechanism is free to defer agent payments until all supply has been realized. As shown by Babaioff et al. [2] when payments must be calculated online, incentive constraints must incorporate the agent's beliefs over the supply or significant loss in performance is inevitable. Deferred payments are reasonable for the motivating application of advertisement auctions, since the advertisers are typically charged at the end of the billing cycle.

2. PRELIMINARIES

We consider the problem of a monopolistic seller attempting to maximize their profit when selling k indivisible units of a single item to n unit-demand agents. We consider ex post incentive compatible and individually rational auctions for solving this problem. In such a mechanism each agent has a (weakly) dominant strategy of participating in the auction and reporting their true valuation as their bid. Such an auction selects at most k winners and demands a payment from each. From agent i's perspective, v_i is their valuation upon winning, x_i is an indicator for their winning a unit or not, and p_i is their payment. Feasibility for k units requires that $\sum_i x_i \leq k$.

We consider two variants of the limited supply problem. In the standard variant the supply limitation k is known in advance. We refer to this as the *offline* variant. In the *online* variant the supply is not known in advance, instead the mechanism is given units to allocate one at a time. Upon receiving a unit the mechanism must either choose an agent to whom to allocate it or to throw it away (the designer has free disposal). Eventually, perhaps adversarially, the designer is told that there are no more units. At this point the designer calculates payments for agents who received units. Notice that with both online and offline supply the mechanism is single-round and sealed-bid with respect to the agents: the agents bid, units are allocated online or offline, payments are determined.

Formally, an auction maps the *valuation profile* $\mathbf{v} = (v_1, \ldots, v_n)$ into an *allocation* $\mathbf{x}(\mathbf{v})$ and *payments* $\mathbf{p}(\mathbf{v})$. It is often useful to look at the probability of allocation and expected payment of an agent with a specific value (randomization taken over agent valuations and coins flipped by a randomized auction protocol); define $p_i(v_i) = \mathbf{E}_{\mathbf{v}_{-i}}[p_i(v_i, \mathbf{v}_{-i})]$ and $x_i(v_i) = \mathbf{E}_{\mathbf{v}_{-i}}[x_i(v_i, \mathbf{v}_{-i})]$ where $\mathbf{v}_{-i} = (v_1, \ldots, v_{i-1}, ?, v_{i+1}, \ldots, v_n)$. The lemma below that characterizes incentive compatible auctions allows us to ignore the payment rule (it is unique) and focus on choosing a good allocation rule.

LEMMA 1. *[13] Any incentive compatible auction in which losers pay nothing satisfies (for all agents i):*

1. *allocation monotonicity: $x_i(v_i)$ is non-decreasing in v_i.*
2. *payment identity: $p_i(v_i) = v_i x_i(v_i) - \int_0^{v_i} x_i(z)dz$.*

Our results for the prior-free designer are best discussed in the context of the following standard results for the Bayesian designer [13] (for a survey see [9]). Assume the valuations of the agents are drawn independently and identically from a distribution with *distribution function* F and *density function* f. For this setting the *Bayesian optimal auction* specifies a distribution-dependent partitioning valuation space into intervals of equal "priority" and allocates the units to the (at most) k agent with the highest positive priority with ties broken randomly. Clearly such an allocation rule satisfies the monotonicity condition of Lemma 1. The payments can be calculated simply using the payment identity. *Notice that it does not matter whether the supply is online or offline.*

DEFINITION 2. *The* Bayesian optimal *auction for distribution F is $\mathcal{M}_F$.*

Our Bayesian designer obtains an auction which is optimal in an absolute sense: for valuations from the given distribution, the expected performance of his chosen (i.e., the Bayesian optimal) auction is at least that of any other auction. Our prior-free designer is not endowed with prior knowledge of the distribution. For any auction our prior-free designer considers, there is always a distribution for which some other auction strictly out performs the considered auction. Thus, the prior-free auction design literature has turned to a relative notion of optimality that bears close resemblance to the *competitive analysis of online algorithms*. Here an auction is good if its maximum (worst case over inputs) ratio with a *benchmark performance* is small. The *prior-free optimal auction* for a given benchmark is the one with the smallest worst-case ratio. This framework for prior-free design is given in [7] (for a survey see [9]).

DEFINITION 3. *A* performance benchmark *maps valuation profile* $\mathbf{v}$ *and supply limit* k *to a target performance, notated, e.g.,* $\mathcal{G}(k, \mathbf{v})$*. Where* k *or* $\mathbf{v}$ *is implicit in the context use short and notations* $\mathcal{G}(k)$ *and* $\mathcal{G}(\mathbf{v})$ *respectively. An* auction *maps a valuation profile and supply limit to an expected revenue, notated, e.g.,* $\mathcal{A}(k, \mathbf{v})$.

DEFINITION 4. *The* prior-free optimal auction *for benchmark* $\mathcal{G}$ *is auction* $\mathcal{A}$ *that minimizes*

$$\min_{\mathcal{A}} \max_{\mathbf{v}} \frac{\mathcal{G}(\mathbf{v})}{\mathcal{A}(\mathbf{v})}$$

Some care must be taken in choosing performance benchmarks for the prior-free results to be economically meaningful. Hartline and Roughgarden give a general theory for meaningful benchmarks: choose as a benchmark the performance of the *optimal Bayesian optimal* auction.

DEFINITION 5. *The* optimal Bayesian optimal *benchmark is* $\mathcal{G}(\mathbf{v}) = \sup_F \mathcal{M}_F(\mathbf{v})$.

An prior-free approximation to this benchmark is a very strong result as stated by the following fact.

FACT 6. *[11] A prior-free auction that β-approximates $\mathcal{G}$ on any input $\mathbf{v}$ also β-approximates the Bayesian optimal auction on any i.i.d. distribution F.*

In unlimited supply settings, the optimal Bayesian optimal benchmark can be characterized succinctly. Let $v_{(i)}$ denote the ith largest agent valuation. For unlimited supply (i.e., $k = n$) the optimal-Bayesian-optimal benchmark coincides precisely with the optimal-posted-price benchmark, $\mathcal{F}(n, \mathbf{v}) = \max_i iv_{(i)}$, proposed by [6]. For limited supply the optimal-Bayesian-optimal benchmark does not generally coincide with the limited supply benchmark $\mathcal{F}(k, \mathbf{v}) = \max_{i \le k} iv_{(i)}$ from [6].

For technical reasons described by [7] we must adjust the benchmark to ignore the case where there is a single agent with an extreme high valuation. This restriction gives the unlimited supply benchmark of $\mathcal{F}^{(2)}(\mathbf{v}) = \max_{i \ge 2} iv_{(i)}$ and the general benchmark of $\mathcal{G}^{(2)}(\mathbf{v}) = \sup_F \mathcal{M}_F^2(\mathbf{v})$. where $\mathcal{M}_F^{(2)}$ is the Bayesian optimal auction for distribution F constrained to offer prices that are at most $v_{(2)}$. Note that $\mathcal{G}^{(2)}$ and $\mathcal{G}$ coincide when it is optimal to sell at least two units.

We focus on the analysis of a generalization of Goldberg et al.'s [8] *Random Sampling Optimal Price (RSOP)* auction to limited supply. This generalization is based on Baliga and Vohra [3] approach to prior-free mechanisms (referred to as the *Random Sampling Empirical Myerson (RSEM)* auction in Hartline and Karlin's survey [9]). An important construct in the definition of this auction is the *empirical distribution* for a valuation profile. This is simply the distribution F with $F(z)$ equal to the faction of agents in $\mathbf{v}$ with $v_i < z$.

DEFINITION 7. *The k-unit* Random Sampling Empirical Myerson (RSEM) *auction works as follows:*

1. *Randomly partition the agents into two sets, $\mathbf{v}^A$ and $\mathbf{v}^B$.*
2. *Calculate the* empirical distributions *for each set, F^A and F^B.*
3. *Run $\mathcal{M}_{F^A}(\mathbf{v}^B)$ and $\mathcal{M}_{F^B}(\mathbf{v}^A)$ with $k/2$ units each.*

Our main theorem is that RSEM is a 25-approximation to $\mathcal{G}^{(2)}$. *Notice that it does not matter whether the supply is online or offline.*

3. ANALYSIS OF THE RANDOM SAMPLING AUCTION.

In this section we give our analysis of the random sampling auction RSEM. We do so first with a more detailed discussion of Bayesian optimal auctions for continuous distributions. This discussion will enable deeper understanding of our benchmark $\mathcal{G}$ which is the supremum over such auctions. With this understanding we will show how a lemma from [1] proves the approximation bound.

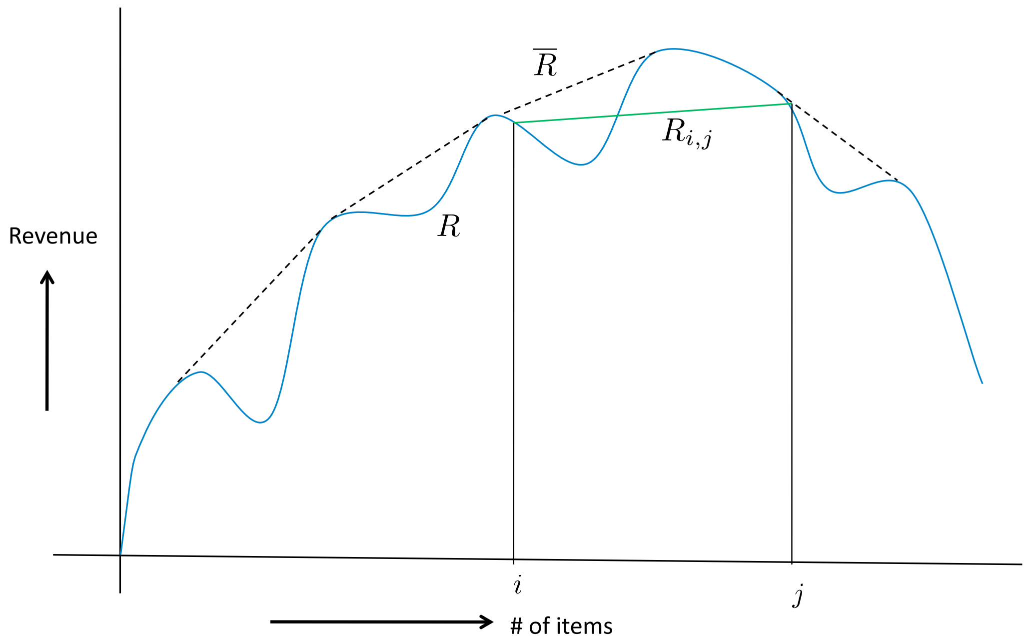

Figure 1: Revenue curves: $R(\cdot)$, $R_{i,j}(\cdot)$, and $\bar{R}(\cdot)$.

3.1 Bayesian optimal auctions, revisited.

Recall our setting for Bayesian optimal auctions where the agent valuations are distributed i.i.d. from distribution F. Consider the case all k units are sold. In such a case, any particular agent's ex ante probability of winning is $\frac{k}{n}$.

Bayesian optimal auctions can be intuitively understood by considering the revenue obtainable as a function of the probability that the agent wins. To allocate with probability $\frac{k}{n}$ to an agent with value drawn from distribution F, we could simply use a posted price of $F^{-1}(1-\frac{k}{n})$, which the agent would accept with probability $\frac{k}{n}$. Our expected revenue is thus

$$R(k) = \tfrac{k}{n} F^{-1}(1 - \tfrac{k}{n}).$$

Notice that this revenue function may not be concave in k (as depicted by Figure 1).[1]

Another way we could allocate with probability $\frac{k}{n}$ is to pick any i and j such that $i < k < j \le n$, and allocate to the agent always if their valuation is at least price $F^{-1}(1-\frac{i}{n})$ and with probability $(k-i)/(j-i)$ if their valuation is between prices $F^{-1}(1-\frac{j}{n})$ and $F^{-1}(1-\frac{i}{n})$. By the payment identity of Lemma 1, the revenue of such an allocation strategy is

$$R_{i,j}(k) = \frac{(k-i)R(i) + (j-k)R(j)}{j-i},$$

i.e., $R_{i,j}(k)$ is the function that is equal to $R(k)$ for $k < i$ and $k > j$, and for $k \in (i,j)$ it is the line connecting $R(i)$ to $R(j)$ (See Figure 1). Notice that when $R(k)$ is non-concave this random allocation can give more revenue than the aforementioned posted price. It is clear that the optimal revenue possible form this kind of approach is obtained from considering the concave hull of $R(k)$ which we denote by $\bar{R}(k)$.

Intuitively in the above example the agent has the same priority when their value is anywhere in the interval between $F^{-1}(1-\frac{i}{n})$ and $F^{-1}(1-\frac{j}{n})$. To achieve the optimal revenue possible, we break valuation space into intervals of equal priority that correspond with the line segments of $\bar{R}(\cdot)$. For instance, the standard approach in the literature is to consider the derivative (a.k.a., slope) of $\bar{R}(\cdot)$ as the priority. (Since $\bar{R}(\cdot)$ is concave, priority is a weakly monotone function of valuation; this particular priority function is known in the literature as the *ironed virtual valuation* of the agent [13].)

3.2 The optimal Bayesian optimal benchmark

Now consider repeating the above discussion but with the empirical distribution F for valuation profile $\mathbf{v}$. We could sell k units at price $v_{(k)}$. Denote this revenue by

$$R(k) = k v_{(k)}.$$

Again, $\bar{R}(k)$ is the concave hull of $R(k)$.

We could then pick an i and j with $i < k < j \le n$ and make the following offer to this particular set of agents with profile $\mathbf{v}$: high-valued agents with values in $[v_{(i)}, \infty)$ win for sure and low-valued agents with values in $[v_{(j)}, v_{(i)})$ win with probability given by the ratio between the number of units left and the number of agents on the interval (i.e., by *lottery*). The payments for such an allocation rule are given by Lemma 1: low-valued agents pay $v_{(j)}$ and high-valued agents pay $v_{(i)} - \frac{(v_{(i)}-v_{(j)})(k-i+1)}{j-i+1}$. Notice that to calculate the high-valued agent payments one must note that a high-

[1] As pointed out by Baliga and Vohra [3], random sampling based auctions cannot assume concave revenue functions as these revenue functions are given the empirical distribution of a sample of the bidders.

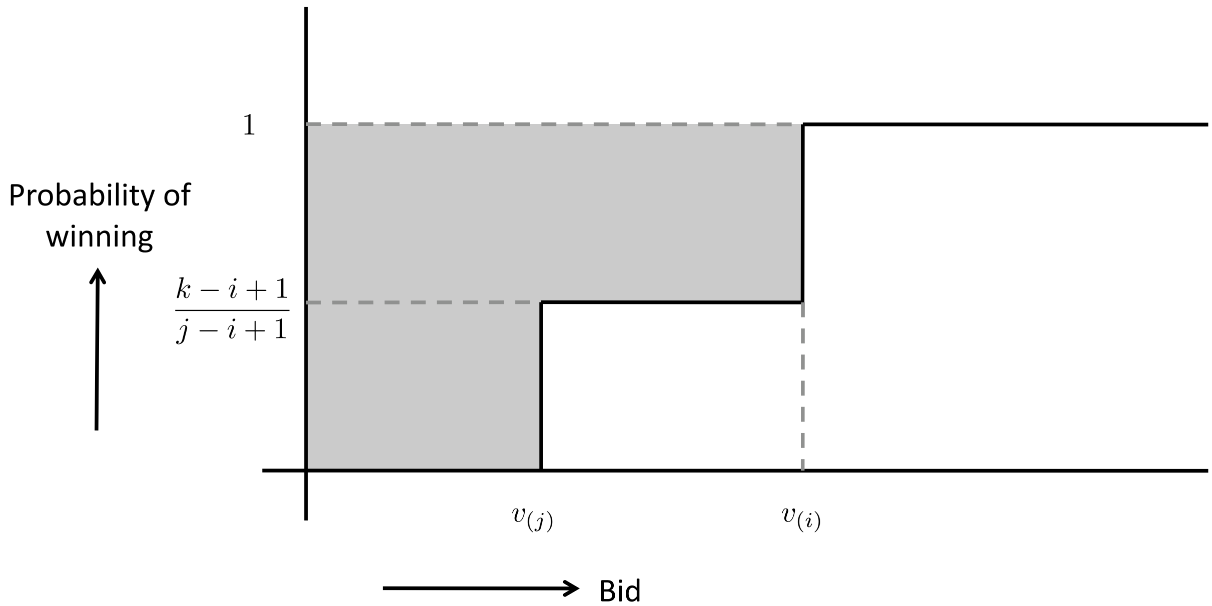

Figure 2: The allocation function $x_i(v_i)$ for $\mathbf{v}_{-i}$ fixed. The area of the shaded region is the payment $p_i(v_i)$ specified by Lemma 1.

valued agent might try to bid as if they were low-valued, this would result in one fewer units being claimed by high-valued agents which means there are $k-i+1$ units to be randomly allocated among the remaining $j-i+1$ agents (see Figure 2). Let $\hat{R}_{i,j}(k)$ denote the total resulting revenue from this approach. Notice that $\hat{R}_{i,j}(k) \neq R_{i,j}(k)$.

As described in the preceding section the Bayesian optimal auction divides valuation space into intervals of equal priority. For the purpose of maximizing profit one would only want to use intervals bounded by agent valuations, i.e., of the form $[v_{(j)}, v_{(i)})$, as this would have all winners paying the maximum possible. This and the above discussion gives intuition for the following lemma.

Lemma 8. *[11]* $\mathcal{G}(k, \mathbf{v}) = \sup_{i,j} \hat{R}_{i,j}(k, \mathbf{v})$ *and* $\mathcal{G}^{(2)}(k, \mathbf{v}) = \sup_{i,j\geq 2} \hat{R}_{i,j}(k, \mathbf{v})$.

$\mathcal{G}$ will be tough for us to work with so we use the following definition which (as the theorem below shows) provides a good upper bound.

Definition 9. $\bar{\mathcal{F}}(k, \mathbf{v}) = \max_{i\leq k} \bar{R}(i, \mathbf{v})$ *and* $\bar{\mathcal{F}}^{(2)}(k, \mathbf{v}) = \max_{2\leq i\leq k} \bar{R}(i, \mathbf{v})$.

The following theorem relates $\mathcal{G}$, $\mathcal{F}$, and $\bar{\mathcal{F}}$. This will be used later in our proof to relate conceptually easier bounds in terms of $\bar{\mathcal{F}}$ with our desired bound in terms of $\mathcal{G}$. Also, the argument $\mathbf{v}$ in the above benchmarks will be dropped when it is clear from the context.

Theorem 10. *For all valuations,* $\mathbf{v}$*,*

1. $\mathcal{F}(k, \mathbf{v}) \leq \mathcal{G}(k, \mathbf{v}) \leq \bar{\mathcal{F}}(k, \mathbf{v})$. *Similarly* $\mathcal{F}^{(2)}(k, \mathbf{v}) \leq \mathcal{G}^{(2)}(k, \mathbf{v}) \leq \bar{\mathcal{F}}^{(2)}(k, \mathbf{v})$.
2. $\bar{\mathcal{F}}(k, \mathbf{v}) \leq 2\mathcal{F}(k, \mathbf{v})$. *There exists* $\mathbf{v}$ *such that* $\mathcal{G}(k, \mathbf{v}) = (2 - \frac{1}{k})\mathcal{F}(k, \mathbf{v})$.
3. $\bar{\mathcal{F}}(k, \mathbf{v}) \leq \min\{\frac{4}{3}, 1 + \frac{1}{k}\}\mathcal{G}(k, \mathbf{v})$.

Proof. We prove each part separately.

1. The first part of the theorem follows almost immediately from the definitions and Lemma 8.

2. For the second part, recall that $\bar{R}$ is concave. If $\bar{R}$ attains its maximum before the supply runs out, that is $\bar{\mathcal{F}}(k) = R(l)$ for some $l \leq k$, then the lemma follows trivially since $\bar{\mathcal{F}}(k) = \mathcal{F}(k)$. So suppose not. Then $\bar{R}$ is monotonically non-decreasing in the interval $[1, k]$ and hence $\bar{\mathcal{F}}(k) = \bar{R}(k)$. Suppose $\bar{R}(k)$ is on the line joining $R(i)$ and $R(j)$, where $i < k < j$. Then

$$\begin{aligned}\bar{R}(k) &= \tfrac{1}{j-i}\left((k-i)R(j) + (j-k)R(i)\right)\\ &= R(i) + \tfrac{k-i}{j-i}(R(j) - R(i))\\ &\qquad \text{(The 2nd term is segment AB in Figure 3)}\\ &\leq R(i) + \tfrac{k}{j}R(j) \text{ (segment CD in Figure 3)}\\ &\leq R(i) + R(k) \leq 2\mathcal{F}(k).\end{aligned}$$

The first inequality on the last line follows since

$$\tfrac{k}{j}R(j) = \tfrac{k}{j}jv_{(j)} = kv_{(j)} \leq kv_{(k)} = R(k).$$

The above bound is almost tight. Let $v_1 = k$, $v_2 = v_3 = \cdots = v_n = 1$. Then $\mathcal{F}(k) = k$, where as $\mathcal{G}(k) = \hat{R}_{1,k}(k) = k + (k-1)(1 - \frac{k}{n})$ which tends to $2k-1$ as n tends to infinity.

3. For the third part, again, the interesting case is when $\bar{\mathcal{F}}(k) = \bar{R}(k)$, and

$$\bar{R}(k) = \tfrac{1}{j-i}\left((k-i)R(j) + (j-k)R(i)\right).$$

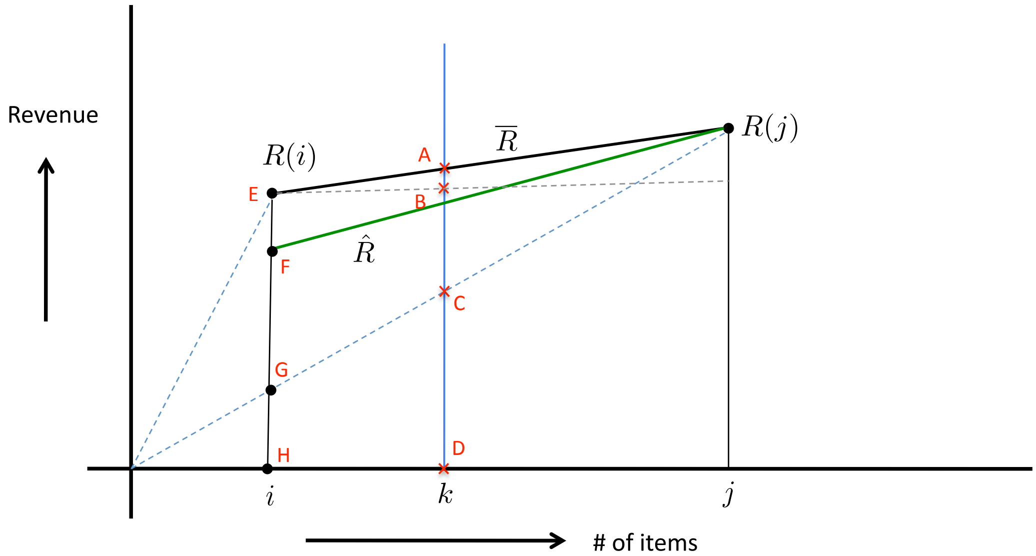

Figure 3: Illustration for proof of Theorem 10.

Let $p = v_{(i)}$ and $q = v_{(j)}$ and define

$$\hat{R}(k) = \hat{R}_{i,j}(k) = (k-i)q + i\left(p - \frac{(p-q)(k-i+1)}{j-i+1}\right).$$

In the interval $[i, j]$, $\hat{R}$ is a linear function of k with $\hat{R}(i) = i\left(p - \frac{p-q}{j-i+1}\right)$ and slope $\frac{q(j+1)-ip}{j-i+1}$. In comparison, in the same interval, $\bar{R}$ is a linear function of k with $\bar{R}(i) = ip$ and slope $\frac{qj-ip}{j-i}$. $\bar{R}(i) > \hat{R}(i)$ where as the slope of $\hat{R}$ is higher than $\bar{R}$. Hence the minimum ratio of $\hat{R}(k)$ to $\bar{R}(k)$ occurs at $k = i$, and $\bar{R}(i) - \hat{R}(i) = \frac{i(p-q)}{j-i+1}$ which is clearly less than $\frac{ip}{4}$ if $j - i \geq 3$. (In Figure 3, ratio of segment EF to EG is 1:j-i+1.) Also, $j-i$ is at least 2, since $i < k < j$. So we are only left with the case when $j - i = 2$. In this case we use the fact that $R(j) \geq R(i)$, which follows from the definition of $\bar{R}$. This implies $(i+2)q \geq ip$, which in turn implies that $\frac{i(p-q)}{j-i+1}$ is less than $\frac{ip}{4}$ by simple algebra. In any case, we have that $\bar{R}(i) \leq \frac{4}{3}\hat{R}(i)$ and in turn $\bar{\mathcal{F}}(k) \leq \frac{4}{3}\mathcal{G}(k)$.

Again, using the fact that $R(j) \geq R(i)$, we get that $\frac{i(p-q)}{j-i+1}$ is less than $\frac{q(j-i)}{j-i+1} \leq q$. Further note that $qk \leq \mathcal{G}(k)$ since qk is the revenue of the posted price auction with price q. Thus $\bar{R}(i) - \hat{R}(i)$, and in fact, $\bar{R}(k) - \hat{R}(k)$ is less than $\mathcal{G}(k)/k$ for all $k \in [i, j]$. Thus, $\bar{\mathcal{F}}(k) \leq \left(1 + \frac{1}{k}\right)\mathcal{G}(k)$.

□

Remark. It immediately follows from the above theorem that any algorithm that is an α-approximation to $\mathcal{F}$ is a 2α-approximation to $\bar{\mathcal{F}}$. Currently the best known auction is from Hartline and McGrew [10] and achieves an approximation factor of 3.25 to $\mathcal{F}$. Thus, it is a 6.5 approximation to $\bar{\mathcal{F}}$. We do not manage to prove that RSEM is better than this; however, there is no reason to believe that it is not.

3.3 The Myerson auction for empirical distributions

The random sampling auction we wish to analyze runs the Myerson auction on a partition of the agents with distribution given empirically by the opposite partition. We now explicitly describe what $\mathcal{M}_F$ does when F is the empirical distribution for valuation profile $\mathbf{v}$. Recall that Myerson's auction for a distribution simply allocates the k units to the agents with highest priority breaking ties in priority randomly. Of course, the key property of Myerson's auction is that it partitions valuation space into intervals of equal priority in the optimal way. From our preceding discussion this partitioning is given by looking at the revenue function $R(\cdot)$ and its concave hull $\bar{R}(\cdot)$. Intervals where theses functions are not equal form a priority class. Specifically, let $i_1 \leq \cdots \leq i_T$ index the T agents i satisfying $R(i) = \bar{R}(i)$ on the rising slope of $\bar{R}(\cdot)$ (i.e., $i_T = \arg\max_i R(i)$). The equal-priority intervals are

$$[v_{(i_T)}, v_{(i_{T-1})}], (v_{(i_{T-1})}, v_{(i_{T-2})}], \ldots, (v_{(2)}, v_{(1)}], (v_{(1)}, \infty).$$

Agents with values below the smallest interval, i.e., strictly less than $v_{(i_T)}$ are rejected.

The set of bidders is partitioned into sets A and B. Let the bids in A be $v^A_{(1)} \geq v^A_{(2)} \geq \ldots$ and so on. Let F^A be the empirical distribution from bidders in A and consider running the Myerson auction with this distribution on the bidders in B, i.e., $\mathcal{M}_{F^A}(\mathbf{v}^B)$. Myerson allocates as follows. Let $i^A_1 < i^A_2 < \cdots < i^A_T$ be the indices of agents on the concave hull of the revenue curve for F^A (as described in the preceding paragraph). Let $p^A_t = v^A_{(i^A_t)}$. The equal priority intervals are

$$[p^A_T, p^A_{T-1}], (p^A_{T-1}, p^A_{T-2}], \ldots, (p^A_2, p_1], (p^A_1, \infty).$$

The bidders in B are allocated odd-numbered items. Let

$j_t^B = |\{l \in B : v^B_{(l)} \geq p_t^A\}|$. Now if $\frac{k}{2} \in [j_t^B, j_{t+1}^B)$, then the top j_t^B always get the item and pay $p_t^A - \frac{p_t^A - p_{t+1}^A}{j_{t+1}^B - j_t^B + 1}$. Bidders with values in the interval $(p_{t+1}^A, p_t^A]$ win remaining units at price p_{t+1}^A with ties broken randomly.

3.4 The performance of the random sampling auction

Our analysis of the performance of the RSEM auction uses a lemma from the recent paper of [1] that analyses the RSOP auction for unlimited supply. They prove a lemma about the expectation of a certain random variable, let's call it X, that is used to lower bound the revenue of the RSOP auction. We show that the same quantity X can be used to bound the revenue of the RSEM auction for limited supply as well. While the fact that X lower bounds the revenue of the RSOP auction for unlimited supply is more or less straightforward, the analogous statement for the RSEM auction for limited supply is more complicated because the revenue of lotteries is more complicated that the revenue of posted price auctions.

We now define the random variable X. Without loss of generality, let the highest bidder be in B. Let $s_i = \left|\{j \in A : v_j \geq v_{(i)}\}\right|$ be the number of bidders in A among the top i bidders. Note that the distribution of s_i is independent of the actual values, and has the same distribution as the following discrete random walk on integers: for each i, s_i is either s_{i-1}, or $s_{i-1}+1$, with probability half each, with $s_1 = 0$. Let $z = \min_i (i - s_i)/s_i$. Suppose that the optimum single price auction sells λ units, i.e. $\mathcal{F}^{(2)} = \lambda v_{(\lambda)}$. Then

$$X = z \cdot \frac{s_\lambda}{\lambda}.$$

Note that for every possible value of λ we get a different random variable X. The main lemma that we need from [1] lower bounds $\mathbf{E}[X]$ for all possible values of λ.

Lemma 11. *[1] For all positive integers λ, $\mathbf{E}[X] \geq \frac{1}{4.68}$.*

As is usual in the analysis of random sampling auctions, we relate the revenue of the RSEM auction from side B to some function of the values from side A. Here, that function is the *concave hull* of the revenue curve from side A. Let R^A be the revenue curve restricted to the bidders in A; more precisely $R^A(i) = iv^A_{(i)}$. Let $\bar{R}^A$ be its concave hull. However, for $i > i_T^A$, we define $\bar{R}^A(i) = \bar{R}^A(i_T^A)$. Note that this makes $\bar{R}^A$ a monotonically non-decreasing concave function. Let RSEM(k) denote the profit of the RSEM Auction with k items.

Lemma 12. $\text{RSEM}(k) \geq z\bar{R}^A(\frac{k}{2})/\min\{\frac{4}{3}, 1 + \frac{1}{k}\}$.

Proof. Recall that the highest bidder is in B. Define $\text{RSEM}^B(i)$ to be the revenue obtained by the RSEM auction from side B, as a function of the number of units allocated to B. In the proof we will only consider the revenue obtained from side B, i.e., we will use that $\text{RSEM}(k) \geq \text{RSEM}^B(k)$. This is tight in the worst case, because if $v_{(1)}$ is high enough, then the optimal Myerson auction for F^B (the empirical distribution for B) is to run a Vickrey Auction with reserve price equal to $v_{(1)}$. This auction would fetch no revenue from side A.

Suppose $\frac{k}{2} \in [j_t^B, j_{t+1}^B)$. Consider the following function defined on the interval $[j_t^B, j_{t+1}^B)$: let $L(i)$ be the linear function such that $L(j_t^B)$ is equal to $j_t^B p_t^A$ and $L(j_{t+1}^B) = j_{t+1}^B p_{t+1}^A$. As in the proof of part 3 of Theorem 10, in the interval $[j_t^B, j_{t+1}^B)$, the function $\text{RSEM}^B(i)$ is within a factor $\min\{\frac{4}{3}, 1 + \frac{1}{k}\}$ of $L(i)$. From the definition of z, we get that $j_t^B \geq zi_t^A$. These two facts point to the conclusion we need. However, the proof needs more argument because of the following:

- It could be that $L(j_{t+1}^B) < L(j_t^B)$. We used the fact that this does not happen in the proof of Theorem 10.
- It could be that $\frac{k}{2} \notin [i_t^A, i_{t+1}^A)$. Thus we cannot directly compare $L(\frac{k}{2})$ to $\bar{R}^A(\frac{k}{2})$.

We get around these difficulties by comparing L to the following function instead. Let $\underline{R}^A(zi_t^A) = zi_t^A p_t^A$. Extend $\underline{R}^A$ to all $i \in [1, zi_T^A]$ by a linear interpolation. For $i > zi_T^A$, let $\underline{R}^A(i) = \underline{R}^A(zi_T^A)$. We note the following easy observations without proof (see Figure 4):

- For all i, $\underline{R}^A(zi) = z\bar{R}^A(i)$.
- $\underline{R}^A$ is an increasing concave function.
- For all $i \in [j_t^B, j_{t+1}^B]$, $L(i) \geq \underline{R}^A(i)$.

We now argue that $L(\frac{k}{2}) \geq z\bar{R}^A(\frac{k}{2})$ via the following sequence of inequalities; this sequence of inequalities corresponds to the points $A \geq B \geq C = zD$ in Figure 4.

$$L(\frac{k}{2}) \geq \underline{R}^A(\frac{k}{2}) \geq \underline{R}^A(z\frac{k}{2}) = z\bar{R}^A(\frac{k}{2}).$$

Therefore it is enough to show that

$$\text{RSEM}^B(\tfrac{k}{2}) \geq L(\tfrac{k}{2})/\min\{\tfrac{4}{3}, 1 + \tfrac{1}{k}\}.$$

The proof of this is identical[2] to that in the proof of part 3 of Theorem 10, unless $L(j_{t+1}^B) < L(j_t^B)$. In case $L(j_{t+1}^B) < L(j_t^B)$, we use the observation that the revenue of the auction actually decreases in the interval $[j_t^B, j_{t+1}^B)$ and is in fact at least $L(j_{t+1}^B)$. The conclusion holds from the following sequence of inequalities, each of which is easy to see given the discussion so far.

$$\text{RSEM}^B(\tfrac{k}{2}) \geq L(j_{t+1}^B) \geq \underline{R}^A(j_{t+1}^B) \geq \underline{R}^A(\tfrac{k}{2}) \geq z\bar{R}^A(\tfrac{k}{2}).$$

This completes the proof of the lemma. □

Finally, we relate $\bar{R}^A(k/2)$ to the benchmark $\mathcal{F}^{(2)}(k)$ as in the following lemma.

Lemma 13. $\bar{R}^A(k/2) \geq \frac{s_\lambda}{2\lambda} \cdot \mathcal{F}^{(2)}(k)$.

Proof. In unlimited supply, if the optimum single price sells λ units, i.e. $\mathcal{F}^{(2)} = \lambda v_{(\lambda)}$, then one can offer $v_{(\lambda)}$ to the bidders in A and get a revenue of $v_{(\lambda)} s_\lambda$, so one would conclude that

$$\mathcal{F}^A \geq v_{(\lambda)} s_\lambda = \frac{s_\lambda}{\lambda} \cdot \mathcal{F}^{(2)}.$$

However the same argument does not work in case of limited supply, because we split the supply also in half for each side. Thus offering $v_{(\lambda)}$ to the bidders in A is guaranteed to get a revenue of $v_{(\lambda)} \cdot \min\{s_\lambda, \frac{k}{2}\}$. Since $s_\lambda \leq k$, we can use that $\min\{s_\lambda, \frac{k}{2}\} \geq \frac{s_\lambda}{2}$ to get the required conclusion for the lemma. □

[2]We also have to consider an additional case, when $j_{t+1}^B - j_t^B = 1$, which is similar to the case $j_{t+1}^B - j_t^B = 2$.

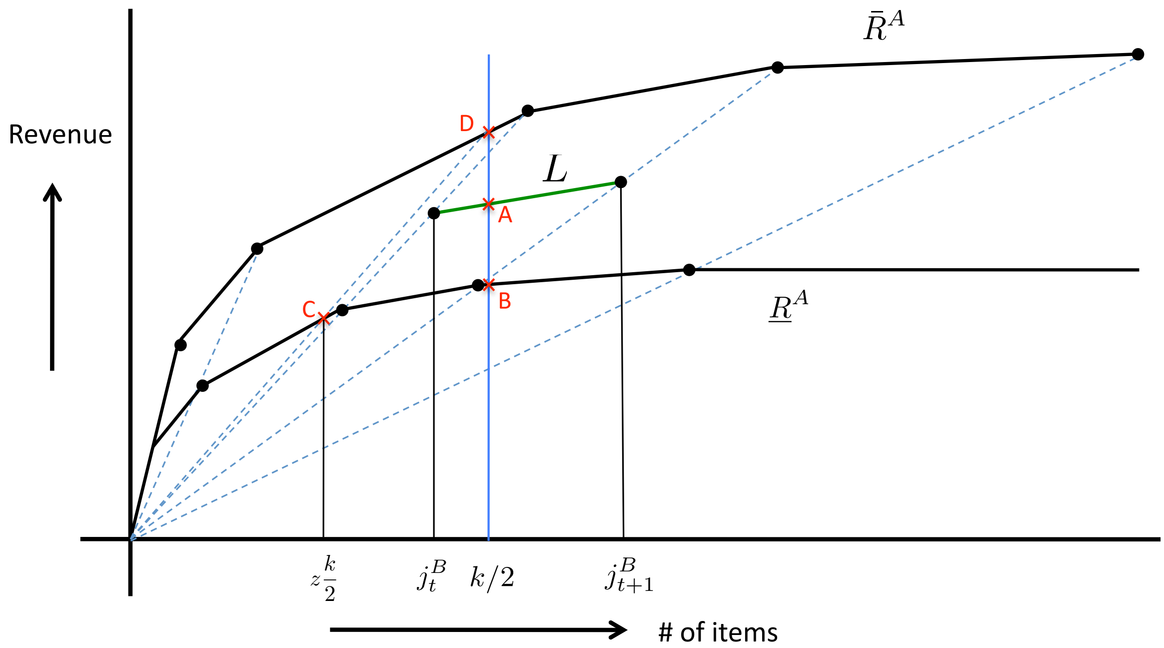

Figure 4: Revenue curves illustrated for the proof of Lemma 12.

Our main theorem now follows easily from Lemmas 11, 12, 13 and Theorem 10.

THEOREM 14. *RSEM is at most a* $\min\{8/3, 2(1+\frac{1}{k})\} \cdot 4.68$*-approximation with respect to* $\mathcal{F}^{(2)}$ *and at most a* $\min\{16/3, 4(1+\frac{1}{k})\} \cdot 4.68$*-approximation with respect to* $\bar{\mathcal{F}}^{(2)}$*.* $\frac{16}{3} \cdot 4.68 \approx 25$*.*

Remark. In using Lemma 11 from [1] as a blackbox, we lose a factor of 2 in the proof of Lemma 13. It is natural to wonder if the random variable X is defined as $z \cdot \min\{\frac{s_\lambda}{\lambda}, \frac{1}{2}\}$, then what is the lower bound on $\mathbf{E}[X]$? Naively, the bound is $1/(2 \cdot 4.68)$ (as used in the present analysis). An improvement in this bound would improve the approximation factor in our analysis.

4. CONCLUSIONS

In this paper we show that the standard prior-free mechanism design approach to random sampling gives an approximately optimal auction for multi-unit single-item settings of limited supply. Furthermore, the supply limitation may be determined online and this does not affect the auction protocol nor its performance. We obtain a bound on the approximation factor of 25 which implies that the auction (RSEM) is the best known auction for the online supply problem. Furthermore we have no reason to believe that the actual approximation factor does not in fact match the conjectured 4-approximation for the unlimited supply case. Thus, the random sampling auction could potentially outperform the best known auction for (offline) limited supply [10] which is a 6.5 approximation.

Open Question. What is RSEM's approximation factor to $\mathcal{G}^{(2)}$?

One direction for future study is in whether the benchmark $\mathcal{G}(k)$ has a *profit extractor*, i.e., is there an (incentive compatible and individually rational) auction that when given target profit R can extract profit R on any valuation profile $\mathbf{v}$ with $R \leq \mathcal{G}(k, \mathbf{v})$. This question is especially interesting because the best prior-free auctions for unlimited supply are based on profit extractors for $\mathcal{F}$. As we argued, $\mathcal{F}$ is the wrong benchmark for limited supply. The challenge in developing a profit extractor for $\mathcal{G}(k)$ is that the benchmark is inherently two dimensional where as all known profit extractors work by searching linearly in a single dimension.

Open Question. Is there a profit extractor for $\mathcal{G}(k)$?

5. REFERENCES

[1] S. Alaei, A. Malekian, and A. Srinivasan. On random sampling auctions for digital goods. In *Proc. 11th ACM Conf. on Electronic Commerce*, 2009.

[2] M. Babaioff, L. Blumrosen, and A. Roth. Auctions with online supply. Working paper.

[3] S. Baliga and R. Vohra. Market research and market design. *Advances in Theoretical Economics*, 3, 2003.

[4] S. Chakraborty and N. Devanur. An Online Multi-unit Auction with Improved Competitive Ratio. *ArXiv e-prints*, January 2009.

[5] U. Feige, A. Flaxman, J. Hartline, and R. Kleinberg. On the Competitive Ratio of the Random Sampling Auction. In *Proc. 1st Workshop on Internet and Network Economics*, pages 878–886, 2005.

[6] A. Fiat, A. Goldberg, J. Hartline, and A. Karlin. Generalized competitive auctions. In *Proc. 34th ACM*

Symp. on Theory of Computing, pages 72–81. ACM Press, 2002.

[7] A. V. Goldberg, J. D. Hartline, A. Karlin, M. Saks, and A. Wright. Competitive auctions. *Games and Economic Behavior*, 55:242–269, 2006.

[8] A. V. Goldberg, J. D. Hartline, and A. Wright. Competitive auctions and digital goods. In *Proc. 12th ACM Symp. on Discrete Algorithms*, pages 735–744. ACM/SIAM, 2001.

[9] J. Hartline and A. Karlin. Profit maximization in mechanism design. In N. Nisan, T. Roughgarden, É. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*, chapter 13, pages 331–362. Cambridge University Press, 2007.

[10] J. Hartline and R. McGrew. From optimal limited to unlimited supply auctions. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, pages 175–182, 2005.

[11] J. Hartline and T. Roughgarden. Optimal mechanism design and money burning. In *Proc. 39th ACM Symp. on Theory of Computing*, 2008.

[12] M. Mahdian and A. Saberi. Multi-unit auctions with unknown supply. In *Proc. 8th ACM Conf. on Electronic Commerce*, 2006.

[13] R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981.