

On Random Sampling Auctions for Digital Goods ^{*†}

Saeed Alaei [‡]

Azarakhsh Malekian [§]

Aravind Srinivasan [¶]

Abstract

In the context of auctions for digital goods, an interesting random sampling auction has been proposed by Goldberg, Hartline, and Wright [2001]. This auction has been analyzed by Feige, Flaxman, Hartline, and Kleinberg [2005], who have shown that it obtains in expectation at least 1/15 fraction of the optimal revenue – which is substantially better than the previously proven constant bounds but still far from the conjectured lower bound of 1/4. In this paper, we prove that the aforementioned random sampling auction obtains at least 1/4 fraction of the optimal revenue for a large class of instances where the number of bids above (or equal to) the optimal sale price is at least 6. We also show that this auction obtains at least 1/4.68 fraction of the optimal revenue for the small class of remaining instances, thus leaving a negligible gap between the lower and upper bound. We employ a mix of probabilistic techniques and dynamic programming to compute these bounds.

1 Introduction

In recent years, there has been a considerable amount of work in algorithmic mechanism design. Most of this work can be divided into two categories based on their assumption about prior: (i) Bayesian, and (ii) prior free. Bayesian mechanism design is based on exploiting the knowledge of the prior to optimize the expected performance, whereas prior free mechanism design is aimed at optimizing the worst case performance. Random sampling auction is perhaps the most popular revenue maximization technique in prior free mechanism design, yet an accurate analysis of its performance has proven difficult even in the simplest applications.

This paper focuses on analyzing the performance of the random sampling auction proposed by Goldberg et al. [2001], known as the “Random Sampling Optimal Price (RSOP)” auction. The basic problem can be described as follows. A seller has unlimited supply of a good (e.g., a digital good) ¹ which he is going to sell to unit demand bidders through the following auction: bids are partitioned into two sets uniformly at random; then the optimal (revenue maximizing) sale price is computed for each set, and offered as the sale price to the opposite set. The expected revenue of RSOP is then compared against the optimal revenue of selling two or more copies at a uniform price.

Most of our analysis is based on the following approach: we develop a lower bound on the performance of RSOP that depends on the level of *balancedness* of the partitions, but independent of the bid values; we then take the expectation of this lower bound over the varying level of balancedness to obtain a general lower bound on the performance of RSOP. This approach is in contrast to the previous work based on showing that a certain level of balancedness is met with a reasonable probability, which inevitably requires a tradeoff between how strong the balancedness condition is versus how likely it holds.

^{*}A preliminary version appeared in the 10th ACM Conference on Electronic Commerce, 2009.

[†]Id: rsop.tex 61 2013-07-14 23:38:03Z saeed

[‡]Department of Computer Science, Cornell University, Ithaca, NY 14850. saeed@cs.cornell.edu. Supported in part by NSF Award CNS-0720528.

[§]EECS department, Massachusetts Institute of Technology, Cambridge MA 02139. malekian@mit.edu. Supported in part by NSF Award CCF-0728839.

[¶]Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. srin@cs.umd.edu. Supported in part by NSF ITR Award CNS-0426683, NSF Award CNS-0626636, and NSF Award CNS 1010789.

¹If there is a fixed production cost per copy, the auction can still be used by simply subtracting the production cost from every bid.

Related work. The random sampling optimal price (RSOP) auction was proposed by Goldberg and Hartline [2001], but the problems was first studied by Goldberg et al. [2001]. The revenue of RSOP has been shown to be close to optimal for many classes of interesting inputs by Segal [2003], and Balcan et al. [2005]. There has also been a fair amount of work analyzing the performance of RSOP. Goldberg et al. [2006] showed that RSOP obtains a constant fraction of the optimal revenue, and conjectured the constant to be $1/4$; note that the conjecture is tight for an instance with 2 bidders with distinct bids. A better analysis was proposed by Feige et al. [2005] which proved the constant to be at least $1/15$.

Improving the lower bound on the expected revenue of RSOP is important because RSOP is a natural and popular mechanism which is easily implementable and adaptable to various settings such as double auctions [Baliga and Vohra, 2003], online limited supply auctions [Hajiaghayi et al., 2004], combinatorial auctions [Balcan et al., 2005, Goldberg and Hartline, 2001], and the money burning auction Hartline and Roughgarden [2008]). Indeed the results of this paper have been used in analysis of other auctions such as the random sampling based auction of Devanur and Hartline [2009] for limited and online supply.

Results. The following is a summary of our main results.

I. **Improved lower bounds:** We prove that the ratio of the expected revenue of RSOP to its benchmark is:

- at least $1/4.68$ (e.g., Theorem 1, and Theorem 2), improving the previous lower-bound of $1/15$ due to Feige et al. [2005];
- at least $1/4$, if there are at least 6 bids above (or equal to) the sale price.
- at least $1/3.53$, as the number of bids above (or equal to) the sale price approaches infinity.

Our analysis suggests that the worst case performance of RSOP is attained when there are only two bidders with distinct bids.

II. **Upper bound:** We show that there exist instances where the expected revenue of RSOP is still less than $1/2.65$ of its benchmark, even when the number of bids above the optimal sale price approaches infinity.

III. **Combinatorial approach:** We also present a combinatorial lower bound on the performance of RSOP for a special case when each non-zero bid can take one of only two possible values.

2 Preliminaries

We consider auctioning a digital good to n bidders whose bids are represented by the vector $\mathbf{v} = (v_1, \dots, v_n)$ which, without loss of generality, is sorted in decreasing order.

Definition 1 (RSOP). *The random sampling optimal price auction partitions the bidders into two sets \mathbf{A} and \mathbf{B} uniformly at random², computes the optimal sale price of each set, and offers it as the sale price to the opposite set.*

Definition 2 (OPT). *The optimal revenue from single price sale to at least two bidders is*

$$\text{OPT} = \max_{j \geq 2} jv_j. \tag{1}$$

See Goldberg et al. [2006] for motivation of the definition of OPT and why it requires selling to at least two bidders.

²I.e., each bidder independently goes to one of \mathbf{A} or \mathbf{B} with probability $\frac{1}{2}$. Each bidder is represented by her index in the sorted ordering of bids.

Assumptions. Without loss of generality, we assume there are infinitely many 0 bids, i.e., $v_j = 0$ for all $j > n$; consequently (\mathbf{A}, \mathbf{B}) is a partitioning of \mathbb{N} . The previous assumption allows us to make our analysis independent of n . Also, without loss of generality, we assume $1 \in \mathbf{B}$, i.e., the bidder with the highest bid is in \mathbf{B} .³

Throughout most of our analysis we ignore the revenue of RSOP from bidders in \mathbf{A} because in the pathological case where v_1 is too large (e.g., $v_1 > \text{OPT}$), the optimal sale price for \mathbf{B} is equal to v_1 which yields no revenue when offered to \mathbf{A} .

Notation. Throughout this paper we adopt the following convention: vectors are represented by bold letters; random variables are represented by italic capital letters if single dimensional, and bold capital letters if multidimensional (such as random sets); sets are represented by roman capital letters; and events are represented by calligraphy capital letters.

We will use $\mathbf{E}[\text{RSOP}]$ to denote the expected revenue of RSOP for an implicit bid vector \mathbf{v} , where the expectation is taken over all random partitions (\mathbf{A}, \mathbf{B}) ; however we sometimes specify an explicit bid vector by writing $\mathbf{E}[\text{RSOP}(\mathbf{v})]$ or $\text{OPT}(\mathbf{v})$.

We use λ to denote the index of the optimal sale price which sells to at least two bidders, i.e.,

$$\lambda \in \arg \max_{j \geq 2} j v_j \quad (2)$$

For every $j \in \mathbb{N}$, we define

$$S_j = |\mathbf{A} \cap \{1 \cdots j\}|, \quad (3)$$

$$Z_j = \frac{|\mathbf{B} \cap \{1 \cdots j\}|}{|\mathbf{A} \cap \{1 \cdots j\}|} = \frac{j - S_j}{S_j}, \quad (4)$$

$$Z = \min(\{Z_j\}_{j \in \mathbb{N}}, 1). \quad (5)$$

S_j , Z_j , and Z are random variables which depend only on how the bids are partitioned, but not on the actual value of the bids. S_j denotes how many bidders, out of the j highest bidders, are in \mathbf{A} . Z_j denotes the ratio of the revenue from \mathbf{B} to the revenue from \mathbf{A} when v_j is offered as the sale price to everyone. Z denotes the minimum of 1 and the worst case ratio of the revenue from \mathbf{B} to the revenue from \mathbf{A} when everyone is offered the same sale price.

For every $T \subset \mathbb{N}$ and $\alpha, \alpha' \in [0, 1]$, we define the following events:

$$\mathcal{E}_\alpha^T = \left\{ \max_{j \in T} \frac{S_j}{j} \leq \alpha \right\}, \quad (6)$$

$$\mathcal{E}_{(\alpha', \alpha]}^T = \left\{ \alpha' < \max_{j \in T} \frac{S_j}{j} \leq \alpha \right\} = \mathcal{E}_\alpha^T \setminus \mathcal{E}_{\alpha'}^T \quad (7)$$

We omit T if $T = \mathbb{N}$, i.e., $\mathcal{E}_\alpha = \mathcal{E}_\alpha^{\mathbb{N}}$ and $\mathcal{E}_{(\alpha', \alpha]} = \mathcal{E}_{(\alpha', \alpha]}^{\mathbb{N}}$. Intuitively, \mathcal{E}_α^T is the event that for every $j \in T$ at most α fraction of the j highest bidders fall in \mathbf{A} . Figure 1 illustrates an example of \mathcal{E}_α and $\mathcal{E}_{(\alpha', \alpha]}$.

Finally, for any random variable X and any event \mathcal{E} , we use $\widehat{\mathbf{E}}[X \mid \mathcal{E}]$ to denote the expectation of X conditioned on event \mathcal{E} normalized by the probability of \mathcal{E} , i.e.,

$$\widehat{\mathbf{E}}[X \mid \mathcal{E}] = \mathbf{E}[X \mid \mathcal{E}] \Pr[\mathcal{E}]. \quad (8)$$

We will use the following proposition extensively throughout this paper.

Proposition 1. *For any random variable X and any two events $\mathcal{E}, \mathcal{E}'$,*

- if $\mathcal{E}' \subseteq \mathcal{E}$, then $\widehat{\mathbf{E}}[X \mid \mathcal{E} \setminus \mathcal{E}'] = \widehat{\mathbf{E}}[X \mid \mathcal{E}] - \widehat{\mathbf{E}}[X \mid \mathcal{E}']$;
- if $\mathcal{E} \cap \mathcal{E}' = \emptyset$, then $\widehat{\mathbf{E}}[X \mid \mathcal{E} \cup \mathcal{E}'] = \widehat{\mathbf{E}}[X \mid \mathcal{E}] + \widehat{\mathbf{E}}[X \mid \mathcal{E}']$

³Otherwise we can rename \mathbf{A} and \mathbf{B} .

The following lemmas are used multiple times throughout the rest of this paper.

Lemma 1. For any $T, T' \subset \mathbb{N}$ and $\alpha \in [0, 1]$, the two events \mathcal{E}_α^T and $\mathcal{E}_\alpha^{T'}$ are positively correlated, i.e., $\Pr[\mathcal{E}_\alpha^T \cap \mathcal{E}_\alpha^{T'}] \geq \Pr[\mathcal{E}_\alpha^T] \Pr[\mathcal{E}_\alpha^{T'}]$ (alternatively $\Pr[\mathcal{E}_\alpha^T \cap \bar{\mathcal{E}}_\alpha^{T'}] \leq \Pr[\mathcal{E}_\alpha^T] \Pr[\bar{\mathcal{E}}_\alpha^{T'}]$).

Proof. The claim follows directly from the FKG inequality [Fortuin et al., 1971] and can be found in Appendix B. \square

Lemma 2. For any $\alpha \in (0, 1)$ and $j \in \mathbb{N}$,

$$\text{if } \alpha \geq 0.5, \text{ then } \Pr[\mathcal{E}_\alpha^{\{j\}}] \geq 1 - (r_\alpha)^j, \quad \text{where } r_\alpha = \frac{1}{2\alpha^\alpha(1-\alpha)^{1-\alpha}}, \quad (9)$$

$$\text{if } \alpha \leq 0.5 - 1/j, \text{ then } \Pr[\mathcal{E}_\alpha^{\{j\}}] \leq (r_{(\alpha+1/j)})^{j-1} \quad \text{where } r_\alpha \text{ is the same as above.} \quad (10)$$

Proof. The claim follows from a direct application of Chernoff–Hoeffding bound and can be found in Appendix B. \square

3 The Basic Lower Bound

In this section we prove that RSOP obtains at least 1/4 fraction of the optimal revenue for all input instances where $\lambda > 10$. In the next section, we prove a lower bound on the revenue of RSOP when $\lambda \leq 10$ using a more sophisticated analysis based on the ideas of the current section. The following theorem summarizes the main result of this section.

Theorem 1. $\mathbf{E}[\text{RSOP}] \geq \frac{1}{4} \text{OPT}$ for all $\lambda > 10$. Furthermore, $\mathbf{E}[\text{RSOP}] \geq \frac{1}{3.53} \text{OPT}$ for all $\lambda > 5000$. Table 1 lists the actual lower bounds obtained for various values of λ .

The outline of this section is as follows. First, we present a lower bound on $\mathbf{E}[\text{RSOP}]$ as a function of λ . Recall that expectation is taken over all random partitions (\mathbf{A}, \mathbf{B}) for a fixed set of bids (and thus a fixed λ). Our proposed lower bound depends only on λ and not the actual value of the bids. We present a dynamic program for numerically computing the lower bound for any fixed λ . By computing the lower bound on $\mathbf{E}[\text{RSOP}]$ for all $\lambda \in \{11 \dots 5000\}$ we confirm that it is indeed greater than $\frac{1}{4} \text{OPT}$. We then prove a lower bound of $\frac{1}{3.53} \text{OPT}$ on $\mathbf{E}[\text{RSOP}]$ for all $\lambda > 5000$.

The following lemma provides a lower bound on $\mathbf{E}[\text{RSOP}]$ as a function of λ .

Lemma 3. $\mathbf{E}[\text{RSOP}] \geq \mathbf{E}[\frac{S_\lambda}{\lambda} Z] \text{OPT}$.

Proof. Let v_{λ_A} be the optimal price for \mathbf{A} which RSOP offers to bidders in \mathbf{B} ; observe that $S_{\lambda_A} v_{\lambda_A} \geq S_j v_j$ for all $j \in \mathbb{N}$. The revenue of RSOP is at least the revenue it obtains from \mathbf{B} , therefore

$$\begin{aligned} \text{RSOP} &\geq (\lambda_A - S_{\lambda_A}) v_{\lambda_A} && \text{because at least } \lambda_A - S_{\lambda_A} \text{ bids in } \mathbf{B} \text{ are above or equal to } v_{\lambda_A} \\ &= Z_{\lambda_A} S_{\lambda_A} v_{\lambda_A} && \text{by definition of } Z_{\lambda_A} \text{ in (4)} \\ &\geq Z S_\lambda v_\lambda && \text{because } Z_{\lambda_A} \geq Z \text{ and } S_{\lambda_A} v_{\lambda_A} \geq S_\lambda v_\lambda \\ &= Z \frac{S_\lambda}{\lambda} \text{OPT} && \text{because } \text{OPT} = \lambda v_\lambda. \end{aligned}$$

Consequently, $\mathbf{E}[\text{RSOP}] \geq \mathbf{E}[\frac{S_\lambda}{\lambda} Z] \text{OPT}$ which proves the claim. \square

Notice that the lower bound provided by the above lemma only depends on λ and does not directly depend on the exact value of the bids.

3.1 Small λ

We start by proving the first part of Theorem 1, i.e., that $\mathbf{E}[\text{RSOP}] \geq \frac{1}{4} \text{OPT}$ for all $10 < \lambda \leq 5000$.

Recall that $\mathbf{E}[\text{RSOP}] \geq \mathbf{E}[\frac{S_\lambda}{\lambda} Z] \text{OPT}$ by Lemma 3. Ideally, we would like to approximate $\mathbf{E}[\frac{S_\lambda}{\lambda} Z]$ by $\mathbf{E}[\frac{S_\lambda}{\lambda}] \mathbf{E}[Z]$, however $\frac{S_\lambda}{\lambda}$ and Z are negatively correlated. To work around this obstacle we will decompose $\mathbf{E}[\frac{S_\lambda}{\lambda} Z]$ over a set of small and disjoint events such that, conditioned on each such event, Z can be approximated closely by a constant. The events are defined as follows. We partition the interval $[0, 1]$ to small disjoint intervals by picking m points $0.5 < \alpha_1 < \dots < \alpha_m < 1$. For each interval $(\alpha_{i-1}, \alpha_i]$ we consider the event $\mathcal{E}_{(\alpha_{i-1}, \alpha_i]}$. Recall that $\mathcal{E}_{(\alpha_{i-1}, \alpha_i]}$ is the event that $(\max_j \frac{S_j}{j}) \in (\alpha_{i-1}, \alpha_i]$ (see Figure 1). Conditioned on $\mathcal{E}_{(\alpha_{i-1}, \alpha_i]}$, it is easy to see that $Z \in [\frac{1-\alpha_i}{\alpha_i}, \frac{1-\alpha_{i-1}}{\alpha_{i-1}})$, and therefore we can obtain a good lower bound by substituting Z with $\frac{1-\alpha_i}{\alpha_i}$. Notice that there is no use in picking α_i from $[0, 0.5]$ because for any $\alpha \in [0, 0.5]$, $\Pr[\mathcal{E}_\alpha] = 0$ and therefore, for any bounded random variable X , we get $\widehat{\mathbf{E}}[X | \mathcal{E}_\alpha] = \mathbf{E}[X | \mathcal{E}_\alpha] \Pr[\mathcal{E}_\alpha] = 0$. Also notice that there is no use in considering the event $\mathcal{E}_{(\alpha_m, 1]}$ because we can only guarantee a trivial lower bound of 0 for Z under $\mathcal{E}_{(\alpha_m, 1]}$.

Lemma 4. *Consider any increasing sequence $\alpha_1, \dots, \alpha_m \in (0.5, 1)$ and let $\alpha_{m+1} = 1$. The following inequality holds for any non-negative random variable X .*

$$\mathbf{E}[XZ] \geq \sum_{i=1}^m \left(\frac{1}{\alpha_i} - \frac{1}{\alpha_{i+1}} \right) \mathbf{E}[X | \mathcal{E}_{\alpha_i}]. \quad (11)$$

Proof. Define $\alpha_0 = 0$. We decompose $\mathbf{E}[XZ]$ over the set of disjoint events $\mathcal{E}_{(\alpha_0, \alpha_1]}, \dots, \mathcal{E}_{(\alpha_{m-1}, \alpha_m]}$ as follows.

$$\begin{aligned} \mathbf{E}[XZ] &\geq \sum_{i=1}^m \widehat{\mathbf{E}}[XZ | \mathcal{E}_{(\alpha_{i-1}, \alpha_i]}] && \text{by law of total expectation} \\ &\geq \sum_{i=1}^m \widehat{\mathbf{E}}\left[X \frac{1-\alpha_i}{\alpha_i} \mid \mathcal{E}_{(\alpha_{i-1}, \alpha_i]}\right] && \text{because } Z \geq \frac{1-\alpha_i}{\alpha_i} \text{ conditioned on } \mathcal{E}_{(\alpha_{i-1}, \alpha_i]} \\ &= \sum_{i=1}^m \frac{1-\alpha_i}{\alpha_i} \left(\widehat{\mathbf{E}}[X | \mathcal{E}_{\alpha_i}] - \widehat{\mathbf{E}}[X | \mathcal{E}_{\alpha_{i-1}}] \right) && \text{by Proposition 1 given that } \mathcal{E}_{(\alpha_{i-1}, \alpha_i]} = \mathcal{E}_{\alpha_i} \setminus \mathcal{E}_{\alpha_{i-1}} \\ &= \sum_{i=1}^m \left(\frac{1}{\alpha_i} - \frac{1}{\alpha_{i+1}} \right) \widehat{\mathbf{E}}[X | \mathcal{E}_{\alpha_i}] && \text{by rearranging the terms.} \end{aligned}$$

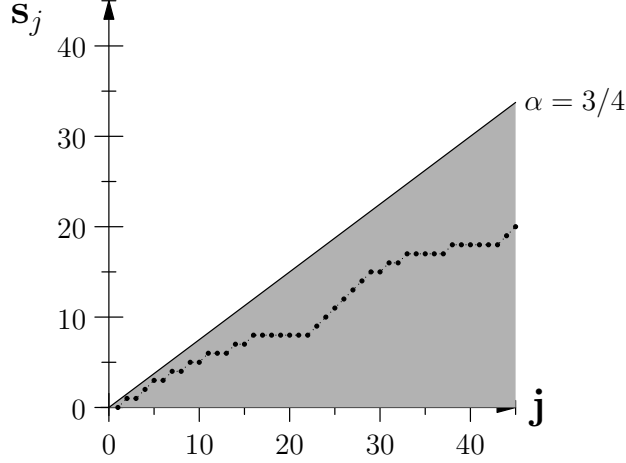
Note that in the last step we have used the fact that $\widehat{\mathbf{E}}[X | \mathcal{E}_{\alpha_0}] = 0$ (because $\Pr[\mathcal{E}_{\alpha_0}] = 0$). \square

The choice of m and $\alpha_1, \dots, \alpha_m$ in Lemma 4 greatly affects the value of the lower bound. Generally speaking, increasing m improves the lower bound but at the cost of more computation.

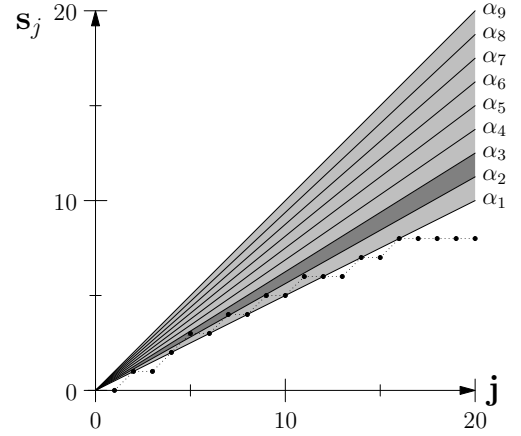
In order to use Lemma 4 effectively, we need to be able to compute $\mathbf{E}[\frac{S_\lambda}{\lambda} | \mathcal{E}_{\alpha_i}]$ for each α_i . However the events \mathcal{E}_{α_i} are hard to deal with computationally. The next two lemmas show that $\mathbf{E}[\frac{S_\lambda}{\lambda} | \mathcal{E}_{\alpha_i}]$ can be bounded below and thus approximated by $\mathbf{E}[\frac{S_\lambda}{\lambda} | \mathcal{E}_{\alpha_i}^{\{1 \dots \ell\}}] - \epsilon$ where ϵ approaches 0 exponentially fast as a function of ℓ .

Lemma 5. *For any random variable $X \in [0, 1]$, any $\alpha \in (0.5, 1]$, and $\ell \in \mathbb{N}$ the following holds:*

$$\widehat{\mathbf{E}}[X | \mathcal{E}_\alpha] \geq \widehat{\mathbf{E}}[X | \mathcal{E}_\alpha^{\{1 \dots \ell\}}] - \epsilon \quad \text{where} \quad \epsilon = \Pr[\mathcal{E}_\alpha^{\{1 \dots \ell\}}] \left(1 - \Pr[\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}] \right) \quad (12)$$



(a) \mathcal{E}_α denotes the event that the points (j, S_j) lie below the line $y = \alpha x$ (the gray area). This figure shows the plot of (j, S_j) for an instance of random partitioning in which the event \mathcal{E}_α with $\alpha = \frac{3}{4}$ has occurred.



(b) $\mathcal{E}_{(\alpha_{i-1}, \alpha_i]}$ denotes the event that the point (j, S_j) with the highest ratio of $\frac{S_j}{j}$ lies on or below the line $y = \alpha_{i-1}x$ and above the line $y = \alpha_i x$. This figure shows the plot of (j, S_j) for an instance of random partitioning in which the event $\mathcal{E}_{(\alpha_2, \alpha_3]}$ has occurred (indicated by the dark gray region).

Figure 1

Proof. Observe that $\mathcal{E}_\alpha = \mathcal{E}_\alpha^{\{1 \dots \ell\}} \setminus (\mathcal{E}_\alpha^{\{1 \dots \ell\}} \cap \overline{\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}})$ ⁴, therefore

$$\begin{aligned}
\widehat{\mathbf{E}} [X \mid \mathcal{E}_\alpha] &= \widehat{\mathbf{E}} [X \mid \mathcal{E}_\alpha^{\{1 \dots \ell\}}] - \widehat{\mathbf{E}} [X \mid \mathcal{E}_\alpha^{\{1 \dots \ell\}} \cap \overline{\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}}] && \text{by Proposition 1} \\
&\geq \widehat{\mathbf{E}} [X \mid \mathcal{E}_\alpha^{\{1 \dots \ell\}}] - \mathbf{Pr} [\mathcal{E}_\alpha^{\{1 \dots \ell\}} \cap \overline{\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}}] && \text{because } X \in [0, 1] \\
&\geq \widehat{\mathbf{E}} [X \mid \mathcal{E}_\alpha^{\{1 \dots \ell\}}] - \mathbf{Pr} [\mathcal{E}_\alpha^{\{1 \dots \ell\}}] (1 - \mathbf{Pr} [\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}]) && \text{by Lemma 1}
\end{aligned}$$

□

The following lemma allows us to compute an upper bound on the ϵ of the previous lemma.

Lemma 6. For any $\alpha \in (0.5, 1]$ and any $\ell, \ell' \in \mathbb{N}$ such that $\ell \leq \ell'$, the following holds:

$$\mathbf{Pr} [\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}] \geq \left(1 - \frac{(r_\alpha)^{\ell'+1}}{1 - r_\alpha}\right) \prod_{j=\ell+1}^{\ell'} (1 - (r_\alpha)^j) \quad \text{where } r_\alpha \text{ is defined in (9)} \quad (13)$$

Proof.

$$\begin{aligned}
\mathbf{Pr} [\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}] &= \mathbf{Pr} \left[\bigcap_{j=\ell+1}^{\infty} \mathcal{E}_\alpha^{\{j\}} \right] \geq \mathbf{Pr} \left[\bigcap_{j=\ell'+1}^{\infty} \mathcal{E}_\alpha^{\{j\}} \right] \prod_{j=\ell+1}^{\ell'} \mathbf{Pr} [\mathcal{E}_\alpha^{\{j\}}] && \text{by Lemma 1} \\
&\geq \left(1 - \sum_{j=\ell'+1}^{\infty} \mathbf{Pr} [\overline{\mathcal{E}_\alpha^{\{j\}}}] \right) \prod_{j=\ell+1}^{\ell'} \mathbf{Pr} [\mathcal{E}_\alpha^{\{j\}}] && \text{by union bound} \\
&\geq \left(1 - \frac{(r_\alpha)^{\ell'+1}}{1 - r_\alpha}\right) \prod_{j=\ell+1}^{\ell'} (1 - (r_\alpha)^j) && \text{by Lemma 2}
\end{aligned}$$

□

⁴ $\overline{\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}}$ denotes the complement of $\mathcal{E}_\alpha^{\{\ell+1 \dots \infty\}}$.

Observe that in the special case of the above lemma in which $\ell = \ell'$, the right hand side of (13) approaches 1 exponentially fast as a function of ℓ which implies that ϵ in (12) approaches 0 exponentially fast as a function of ℓ . Choosing $\ell' > \ell$ only improves the bound.

The next lemma provides a recurrence relation which can be used to compute the exact value of $\mathbf{E}[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_\alpha^{\{1 \dots \ell\}}]$ and $\Pr[\mathcal{E}_\alpha^{\{1 \dots \ell\}}]$ in time $O(\ell^2)$.

Lemma 7. *For any $\ell \in \mathbb{N}$ and $\alpha \in [0, 1]$, the exact value of $\widehat{\mathbf{E}}[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_\alpha^{\{1 \dots \ell\}}]$ and $\Pr[\mathcal{E}_\alpha^{\{1 \dots \ell\}}]$ can be computed using the following recurrence in which $\mathcal{E}_{\alpha,k}^{\{1 \dots \ell\}} = \mathcal{E}_\alpha^{\{1 \dots \ell\}} \cap \{S_\ell = k\}$ is the event that $\mathcal{E}_\alpha^{\{1 \dots \ell\}}$ happens and $S_\ell = k$.*

$$\Pr[\mathcal{E}_{\alpha,k}^{\{1 \dots \ell\}}] = \begin{cases} \frac{1}{2} \Pr[\mathcal{E}_{\alpha,k-1}^{\{1 \dots \ell-1\}}] + \frac{1}{2} \Pr[\mathcal{E}_{\alpha,k}^{\{1 \dots \ell-1\}}] & \ell > 1, k \leq \alpha\ell \\ 1 & \ell = 1, k = 0 \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

$$\widehat{\mathbf{E}}\left[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_{\alpha,k}^{\{1 \dots \ell\}}\right] = \begin{cases} \frac{1}{2} \widehat{\mathbf{E}}\left[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_{\alpha,k-1}^{\{1 \dots \ell-1\}}\right] + \frac{1}{2} \widehat{\mathbf{E}}\left[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_{\alpha,k}^{\{1 \dots \ell-1\}}\right] & \ell > \lambda, k \leq \alpha\ell \\ \frac{k}{\lambda} \Pr[\mathcal{E}_{\alpha,k}^{\{1 \dots \ell\}}] & \ell = \lambda \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

$$\Pr[\mathcal{E}_\alpha^{\{1 \dots \ell\}}] = \sum_{k=0}^{\ell} \Pr[\mathcal{E}_{\alpha,k}^{\{1 \dots \ell\}}] \quad (16)$$

$$\widehat{\mathbf{E}}\left[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_\alpha^{\{1 \dots \ell\}}\right] = \sum_{k=0}^{\ell} \widehat{\mathbf{E}}\left[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_{\alpha,k}^{\{1 \dots \ell\}}\right] \quad (17)$$

Proof. Let \mathcal{A}_ℓ denote the event that $\ell \in \mathbf{A}$. First consider (14): if $\ell > 1$ and $k \leq \alpha\ell$, then $\mathcal{E}_{\alpha,k}^{\{1 \dots \ell\}}$ can be decomposed as two disjoint events $\mathcal{E}_{\alpha,k-1}^{\{1 \dots \ell-1\}} \cap \mathcal{A}_\ell$ and $\mathcal{E}_{\alpha,k}^{\{1 \dots \ell-1\}} \cap \overline{\mathcal{A}}_\ell$, therefore its probability is the sum of the probabilities of those two event; note that $\mathcal{E}_{\alpha,k}^{\{1 \dots \ell-1\}}$ and \mathcal{A}_ℓ are independent for any ℓ and k and $\Pr[\mathcal{A}_\ell] = \frac{1}{2}$; furthermore the base of the recursion is $\Pr[\mathcal{E}_{\alpha,0}^{\{1\}}] = 1$ because by our assumption $\mathcal{A}_1 = 0$ (i.e., the highest bid is always in \mathbf{B}). The same argument implies the correctness of (15) for the case of $\ell > \lambda$. Furthermore, $\mathcal{E}_{\alpha,k}^{\{1 \dots \lambda\}}$ by its definition implies $S_\lambda = k$ which implies the correctness of (15) for the case of $\ell = \lambda$. Finally (16) and (17) follow trivially from the law of total probability and the law of total expectation. \square

Proof of Theorem 1 for small λ (i.e., $10 < \lambda \leq 5000$). We show how to numerically compute a lower bound on $\mathbf{E}[\text{RSOP}]$ for any fixed λ . Let $m = 100$ and $\alpha_i = 0.5 + \frac{i}{m+1}$ for each $i \in [m]$. Observe that

$$\begin{aligned} \mathbf{E}[\text{RSOP}] &\geq \mathbf{E}\left[\frac{S_\lambda}{\lambda} Z\right] \text{OPT} && \text{by Lemma 3} \\ &\geq \sum_{i=1}^m \left(\frac{1}{\alpha_i} - \frac{1}{\alpha_{i+1}}\right) \mathbf{E}\left[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_{\alpha_i}\right] \text{OPT} && \text{by Lemma 4} \end{aligned}$$

We then compute a lower bound for each term $\mathbf{E}[\frac{S_\lambda}{\lambda} \mid \mathcal{E}_{\alpha_i}]$ (using Lemma 5, Lemma 6, and Lemma 7 with $\ell = 5000$ and $\ell' = 100000$), and substitute them in the last inequality above to obtain a lower bound on $\mathbf{E}[\text{RSOP}]$. The computed numerical values of our lower bound are listed in Table 1 for various choices of λ . We have confirmed that $\mathbf{E}[\text{RSOP}] \geq \frac{1}{4} \text{OPT}$ for all $\lambda \in \{11 \dots 5000\}$. \square

3.2 Large λ

We now prove the second part of Theorem 1, i.e., $\mathbf{E}[\text{RSOP}] \geq \frac{1}{3.53} \text{OPT}$ for all $\lambda > 5000$.

Recall that $\mathbf{E}[\text{RSOP}] \geq \mathbf{E}[\frac{S_\lambda}{\lambda} Z] \text{OPT}$ by Lemma 3. Also recall that $\frac{S_\lambda}{\lambda}$ and Z are negatively correlated, thus $\mathbf{E}[\frac{S_\lambda}{\lambda}] \mathbf{E}[Z]$ does not yield a lower bound on $\mathbf{E}[\frac{S_\lambda}{\lambda} Z]$. Nevertheless, the correlation decreases as λ increases which suggests that for sufficiently large λ we can separate the two terms. In other words, when λ is large (i.e., $\lambda > 5000$), the two random variables $\frac{S_\lambda}{\lambda}$ and Z are almost independent and so the expected

value of their product is very close to the product of their expected values. Also for a large λ the value of $\frac{S_\lambda}{\lambda}$ is very close to $\frac{1}{2}$ so $\mathbf{E}[\frac{S_\lambda}{\lambda}Z]$ is close to $\frac{1}{2}\mathbf{E}[Z]$. We formalize this argument in the following lemma.

Lemma 8. For any $\alpha \in [0, 1]$:

$$\mathbf{E}\left[\frac{S_\lambda}{\lambda}Z\right] \geq \alpha\left(\mathbf{E}[Z] - \mathbf{Pr}\left[\mathcal{E}_\alpha^{\{\lambda\}}\right]\right) \quad (18)$$

Proof.

$$\begin{aligned} \mathbf{E}\left[\frac{S_\lambda}{\lambda}Z\right] &= \widehat{\mathbf{E}}\left[\frac{S_\lambda}{\lambda}Z \mid \bar{\mathcal{E}}_\alpha^{\{\lambda\}}\right] + \widehat{\mathbf{E}}\left[\frac{S_\lambda}{\lambda}Z \mid \mathcal{E}_\alpha^{\{\lambda\}}\right] \\ &\geq \alpha \widehat{\mathbf{E}}\left[Z \mid \bar{\mathcal{E}}_\alpha^{\{\lambda\}}\right] && \text{because } \frac{S_\lambda}{\lambda} > \alpha \text{ conditioned on } \bar{\mathcal{E}}_\alpha^{\{\lambda\}} \\ &= \alpha\left(\mathbf{E}[Z] - \widehat{\mathbf{E}}\left[Z \mid \mathcal{E}_\alpha^{\{\lambda\}}\right]\right) \\ &\geq \alpha\left(\mathbf{E}[Z] - \mathbf{Pr}\left[\mathcal{E}_\alpha^{\{\lambda\}}\right]\right) && \text{because } Z \leq 1 \text{ by its definition, i.e., equation (5).} \end{aligned}$$

□

Recall that we can compute an upper bound on $\mathbf{Pr}[\mathcal{E}_\alpha^{\{\lambda\}}]$ using Lemma 2. Also observe that, for any fixed $\alpha \in (0, 0.5)$, $\mathbf{Pr}[\mathcal{E}_\alpha^{\{\lambda\}}]$ approaches 0 exponentially fast as a function of λ as $\lambda \rightarrow \infty$. The only remaining task is to compute a good lower bound on $\mathbf{E}[Z]$.

Lemma 9. $\mathbf{E}[Z] \geq 0.61$.

Proof. Let $\ell = 60000$, $m = 100$, and $\alpha_i = 0.5 + \frac{i}{m+1}$ for each $i \in [m]$. By applying Lemma 4 and plugging $X = 1$ we get

$$\begin{aligned} \mathbf{E}[Z] &\geq \sum_{i=1}^m \left(\frac{1}{\alpha_i} - \frac{1}{\alpha_{i+1}}\right) \mathbf{Pr}[\mathcal{E}_{\alpha_i}] \\ &\geq \sum_{i=1}^m \left(\frac{1}{\alpha_i} - \frac{1}{\alpha_{i+1}}\right) \mathbf{Pr}[\mathcal{E}_{\alpha_i}^{\{1 \dots \ell\}}] \mathbf{Pr}[\mathcal{E}_{\alpha_i}^{\{\ell+1 \dots \infty\}}] && \text{by Lemma 1} \end{aligned}$$

We then use Lemma 7 to compute $\mathbf{Pr}[\mathcal{E}_{\alpha_i}^{\{1 \dots \ell\}}]$, and use Lemma 6 with $\ell' = 100000$ to compute a lower bound on $\mathbf{Pr}[\mathcal{E}_{\alpha_i}^{\{\ell+1 \dots \infty\}}]$; by substituting the computed numerical values in the above inequality we get $\mathbf{E}[Z] \geq 0.61$. □

It is worth mentioning that by using a similar method we have computed an upper bound of $\mathbf{E}[Z] \leq 0.63$ which indicates that our analysis is almost tight. ⁵

Proof of Theorem 1 for large λ (i.e., $\lambda > 5000$). Let $\alpha = 0.48$. Then

$$\begin{aligned} \mathbf{E}[\text{RSOP}] &\geq \mathbf{E}\left[\frac{S_\lambda}{\lambda}Z\right] \text{OPT} && \text{by Lemma 3} \\ &\geq \alpha\left(\mathbf{E}[Z] - \mathbf{Pr}\left[\mathcal{E}_\alpha^{\{\lambda\}}\right]\right) \text{OPT} && \text{by Lemma 8.} \end{aligned}$$

Using Lemma 2 we get $\mathbf{Pr}[\mathcal{E}_\alpha^{\{\lambda\}}] \leq 0.0183$ for all $\lambda > 5000$; furthermore $\mathbf{E}[Z] \geq 0.61$ by Lemma 9; substituting them in the above inequality we get $\mathbf{E}[\text{RSOP}] \geq 1/3.53 \text{OPT}$. □

⁵Note that $1 - 1/e \simeq 0.6321$ which is slightly greater than the upper bound of $\mathbf{E}[Z] \leq 0.63$.

4 The Exhaustive Search Lower-Bound

In this section we propose an exhaustive search approach which yields an improved lower bound for RSOP for small values of λ (i.e., $\lambda \leq 10$). The following theorem summarizes the main result of this section.

Theorem 2. $\mathbf{E}[\text{RSOP}] \geq \frac{1}{4.68} \text{OPT}$ for $\lambda \geq 2$ and $\mathbf{E}[\text{RSOP}] \geq \frac{1}{4} \text{OPT}$ for $\lambda \geq 6$. Table 2 lists the actual lower bounds obtained for various choices of λ .

The basic lower bound of $\mathbf{E}[\text{RSOP}] \geq \mathbf{E}[\frac{S_\lambda}{\lambda}] \text{OPT}$ which we used in Section 3 does not yield a good enough bound when λ is small, mainly because

- (I) $\frac{S_\lambda}{\lambda}$ and Z are negatively correlated and their correlation is much stronger when λ is small,
- (II) the bidder with the highest bid is always in \mathbf{B} , so $\mathbf{E}[\frac{S_\lambda}{\lambda}]$ approaches $\frac{1}{4}$ as λ goes down to 2.

Therefore, for $\lambda = 2$, $\mathbf{E}[\frac{S_\lambda}{\lambda}] < \mathbf{E}[\frac{S_\lambda}{\lambda}] \mathbf{E}[Z] < 0.25 \times 0.63 \approx \frac{1}{6.55}$. In fact, the lower bounds of Table 1 are quite close to the exact value of $\mathbf{E}[\frac{S_\lambda}{\lambda}]$, which suggests for small values of λ we need a different approach.

We now provide a high level description of the approach of this section. Without loss of generality we assume $\text{OPT} = 1$ which immediately implies $v_j \in [0, \frac{1}{j}]$ for all j . In addition to fixing the index of the optimal price, λ , we fix the index of the second optimal price of a higher index, λ' , and also fix its corresponding revenue, OPT' , i.e.,

$$\lambda' \in \arg \max_{j > \lambda} j v_j \quad (19)$$

$$\text{OPT}' = \max_{j > \lambda} j v_j \quad (20)$$

We divide the range of possible values for each of the d highest bids (for an appropriate choice of d) as follows. For each $j \in [d]$ we uniformly divide the interval $[0, \frac{1}{j}]$ to η subintervals for an appropriate choice of η and restrict v_j to one of them, say $[\underline{v}_j, \bar{v}_j]$; similarly we uniformly divide the interval $[0, 1]$ to η' subintervals for an appropriate choice of η' and restrict OPT' to one of them, say $[\underline{w}', \bar{w}']$. Let $\text{RSOPMinExpect}(\lambda, \mathbf{v}, \bar{\mathbf{v}}, \underline{w}', \bar{w}')$ denote the minimum expected revenue of $\mathbf{E}[\text{RSOP}]$ where the minimum is taken over all instances in which $\text{OPT}' \in [\underline{w}', \bar{w}']$, and $v_j \in [\underline{v}_j, \bar{v}_j]$ (for all $j \in [d]$), and λ is the index of the optimal price. We compute a lower bound on the expected revenue of RSOP by computing a lower bound on $\text{RSOPMinExpect}(\lambda, \mathbf{v}, \bar{\mathbf{v}}, \underline{w}', \bar{w}')$ for each combination of intervals and taking the minimum over all such combinations. Formally, we have

$$\mathbf{E}[\text{RSOP}] \geq \min_{\substack{\forall j \in [d]: (\underline{v}_j, \bar{v}_j) \in \{(\frac{i-1}{\eta}, \frac{i}{\eta}) \mid i \in [\eta]\} \\ (\underline{w}', \bar{w}') \in \{(\frac{i-1}{\eta'}, \frac{i}{\eta'}) \mid i \in [\eta']\}}} \text{RSOPMinExpect}(\lambda, \mathbf{v}, \bar{\mathbf{v}}, \underline{w}', \bar{w}'). \quad (21)$$

Note that some combinations of intervals might be inconsistent/infeasible. For example it is not possible to have both $[\underline{v}_2, \bar{v}_2] = [\frac{0}{10}, \frac{1}{10}]$ and $[\underline{v}_3, \bar{v}_3] = [\frac{4}{15}, \frac{5}{15}]$ because that would imply $v_3 > v_2$. We simply define RSOPMinExpect to be ∞ if a combination is infeasible.

Computing a lower bound on $\text{RSOPMinExpect}(\lambda, \mathbf{v}, \bar{\mathbf{v}}, \underline{w}', \bar{w}')$. In the rest of this section we show how to compute a lower bound on the expected revenue of RSOP given the assumption that $\text{OPT}' \in [\underline{w}', \bar{w}']$ and $v_j \in [\underline{v}_j, \bar{v}_j]$ for all $j \in [d]$. The high level idea is to enumerate all possible partitions of the first d bids, define an event for each such partition, decompose $\mathbf{E}[\text{RSOP}]$ over those events, and compute a lower bound conditioned on each such event.

We start with a few definitions. For every $\mathbf{T} \subset [d]$, we define the event

$$\mathcal{A}_{\mathbf{T}} = \left\{ \mathbf{A} \cap \{1 \cdots d\} = \mathbf{T} \right\}. \quad (22)$$

Intuitively, $\mathcal{A}_{\mathbf{T}}$ is the event that, among the highest d bidders, the subset that fall in \mathbf{A} is exactly \mathbf{T} . Observe that under the event $\mathcal{A}_{\mathbf{T}}$, for every $j \in [d]$, both S_j and Z_j are constants; we will denote those constants respectively by

$$s_j^{\mathbf{T}} = (S_j \mid \mathcal{A}_{\mathbf{T}}) = |\{1 \cdots j\} \cap \mathbf{T}|, \quad (23)$$

$$z_j^{\mathbf{T}} = (Z_j \mid \mathcal{A}_{\mathbf{T}}) = \frac{|\{1 \cdots j\} \setminus \mathbf{T}|}{|\{1 \cdots j\} \cap \mathbf{T}|}. \quad (24)$$

Our approach is to decompose $\mathbf{E}[\text{RSOP}]$ over the set of disjoint events $\{\mathcal{A}_T\}_{T \subseteq \{2 \dots d\}}$ and then decompose $\widehat{\mathbf{E}}[\text{RSOP} \mid \mathcal{A}_T]$ further over the set of disjoint events $\{\mathcal{E}_{(\alpha_{i-1}, \alpha_i)}^{\{d+1 \dots \infty\}}\}_{i \in [m]}$ for some choice of $0.5 < \alpha_1 < \dots < \alpha_m < 1$ (this decomposition is similar to Section 3). Formally, we have

$$\mathbf{E}[\text{RSOP}] = \sum_{T \subseteq \{2 \dots d\}} \sum_{i=1}^m \widehat{\mathbf{E}} \left[\text{RSOP} \mid \mathcal{A}_T \cap \mathcal{E}_{(\alpha_{i-1}, \alpha_i)}^{\{d+1 \dots \infty\}} \right]. \quad (25)$$

Next we show how to compute a lower bound on $\widehat{\mathbf{E}}[\text{RSOP} \mid \mathcal{A}_T \cap \mathcal{E}_{(\alpha_{i-1}, \alpha_i)}^{\{d+1 \dots \infty\}}]$. We first define the following constants for all $T \subseteq \{2 \dots d\}$ and $\alpha \in (0, 1]$.

$$r^T = \max_{j \in [d]} s_j^T v_j, \quad (26)$$

$$\overline{\mathbf{A}}^T = \{j \in T \mid s_j^T v_j \geq r^T\}, \quad (27)$$

$$\rho_\alpha^T = \begin{cases} \min \left\{ z_j^T \mid j \in \overline{\mathbf{A}}^T \right\} & \text{if } r^T > \alpha \overline{w}' \\ \min \left(\left\{ z_j^T \mid j \in \overline{\mathbf{A}}^T \right\}, \frac{1-\alpha}{\alpha} \right) & \text{otherwise} \end{cases}. \quad (28)$$

Under each event $\mathcal{A}_T \cap \mathcal{E}_{(\alpha_{i-1}, \alpha_i)}^{\{d+1 \dots \infty\}}$ appearing in (25), the above definitions can be interpreted as follows. r^T is a lower bound on the maximum revenue from \mathbf{A} by offering one of the highest d bids as the sale price. $\overline{\mathbf{A}}^T$ contains all of the d highest bidders except those whose bid cannot possibly be an optimal sale price for \mathbf{A} . ρ_α^T is a lower bound on the ratio of the revenue from \mathbf{B} to the revenue from \mathbf{A} when everyone is offered the optimal sale price of \mathbf{A} .

Lemma 10. *For any $d \geq 2$, any $T \subseteq \{2 \dots d\}$, and any $0.5 < \alpha' < \alpha \leq 1$,*

$$\widehat{\mathbf{E}} \left[\text{RSOP} \mid \mathcal{A}_T \cap \mathcal{E}_{(\alpha', \alpha)}^{\{d+1 \dots \infty\}} \right] \geq \widehat{\mathbf{E}} \left[\max \left(r^T, \frac{S_{\lambda'} w'}{\lambda'} \right) \rho_\alpha^T \mid \mathcal{A}_T \cap \mathcal{E}_{(\alpha', \alpha)}^{\{d+1 \dots \infty\}} \right]. \quad (29)$$

Proof. Let $v_{\lambda_{\mathbf{A}}}$ be the optimal price for \mathbf{A} which RSOP offers to bidders in \mathbf{B} . Observe that

$$\text{RSOP} \geq (\lambda_{\mathbf{A}} - S_{\lambda_{\mathbf{A}}}) v_{\lambda_{\mathbf{A}}} = S_{\lambda_{\mathbf{A}}} v_{\lambda_{\mathbf{A}}} Z_{\lambda_{\mathbf{A}}}.$$

Under event $\mathcal{A}_T \cap \mathcal{E}_{(\alpha', \alpha)}^{\{d+1 \dots \infty\}}$, we show that $S_{\lambda_{\mathbf{A}}} v_{\lambda_{\mathbf{A}}} \geq \max \left(r^T, \frac{S_{\lambda'} w'}{\lambda'} \right)$ (i.e., a lower bound on the optimal revenue from \mathbf{A}) and also show that $Z_{\lambda_{\mathbf{A}}} \geq \rho_\alpha^T$, which combined with the above inequality imply the statement of the lemma.

- First we prove $S_{\lambda_{\mathbf{A}}} v_{\lambda_{\mathbf{A}}} \geq \max \left(r^T, \frac{S_{\lambda'} w'}{\lambda'} \right)$. Notice that $S_{\lambda_{\mathbf{A}}} v_{\lambda_{\mathbf{A}}}$ is the optimal revenue of \mathbf{A} which must be at least r^T ; furthermore, the optimal revenue of \mathbf{A} is no less than the revenue of selling to \mathbf{A} at price $v_{\lambda'}$ which is at least $\frac{S_{\lambda'} w'}{\lambda'}$.
- Next we prove $Z_{\lambda_{\mathbf{A}}} \geq \rho_\alpha^T$. The inequality follows immediately by considering the following two possibilities:
 - (I) If $\lambda_{\mathbf{A}} \leq d$, then $\lambda_{\mathbf{A}}$ must be in $\overline{\mathbf{A}}^T$, because for any $j \in \{1 \dots d\} \setminus \overline{\mathbf{A}}^T$ selling to \mathbf{A} at price v_j generates a revenue which is less than r^T , therefore v_j cannot be the optimal price for \mathbf{A} .
 - (II) If $\lambda_{\mathbf{A}} > d$, then it must be $r^T \leq \alpha \overline{w}'$, otherwise the revenue of selling to \mathbf{A} at price $v_{\lambda_{\mathbf{A}}}$ would be less than r^T which contradicts its optimality.⁶ Furthermore, under event $\mathcal{E}_{(\alpha', \alpha)}^{\{d+1 \dots \infty\}}$, we have $Z_j \geq \frac{1-\alpha}{\alpha}$ for all $j > d$ which implies $Z_{\lambda_{\mathbf{A}}} \geq \frac{1-\alpha}{\alpha}$ too.

□

⁶The revenue of selling to \mathbf{A} at price $v_{\lambda_{\mathbf{A}}}$ is $S_{\lambda_{\mathbf{A}}} v_{\lambda_{\mathbf{A}}}$ which is at most $\alpha \lambda_{\mathbf{A}} v_{\lambda_{\mathbf{A}}} = \alpha \text{OPT}'$ under event $\mathcal{E}_{(\alpha', \alpha)}^{\{d+1 \dots \infty\}}$.

Lemma 11. For any increasing sequence $\alpha_1, \dots, \alpha_m \in (0.5, 1)$ the following inequality holds (assume $\alpha_{m+1} = 1$).

$$\mathbf{E}[\text{RSOP}] \geq \sum_{T \subseteq \{2 \dots d\}} \sum_{i=1}^m (\rho_{\alpha_i}^T - \rho_{\alpha_{i+1}}^T) \widehat{\mathbf{E}} \left[\max \left(r^T, \frac{S_{\lambda'}}{\lambda'} w' \right) \mid \mathcal{A}_T \cap \mathcal{E}_{\alpha_i}^{\{d+1 \dots \infty\}} \right] \quad (30)$$

Proof. The claim follows by applying Lemma 10 to equation (25), then decomposing each event $\mathcal{E}_{(\alpha_{i-1}, \alpha_i]}^{\{d+1 \dots \infty\}}$ as $\mathcal{E}_{\alpha_i}^{\{d+1 \dots \infty\}} \setminus \mathcal{E}_{\alpha_{i-1}}^{\{d+1 \dots \infty\}}$ and applying Proposition 1, and then rearranging the terms. \square

Next we sketch the proof of the main theorem of this section.

Proof of Theorem 2. We use (21) with $d = 11$, $\eta = 3$ and $\eta' = 100$ together with Lemma 11 with $m = 100$ and $\alpha_i = 0.5 + \frac{i}{m+1}$ for each $i \in [m]$. To compute an accurate approximation (lower bound) on each term $\widehat{\mathbf{E}}[\max(r^T, \frac{S_{\lambda'}}{\lambda'} w') \mid \mathcal{A}_T \cap \mathcal{E}_{\alpha_i}^{\{d+1 \dots \infty\}}]$, we use a combination of dynamic programming and tail bounds similar to those of Lemma 5, Lemma 6, Lemma 7, Lemma 8, and Lemma 9 (observe that $S_{\lambda'}$ is the only random variable in this term). However doing so naively requires computing a lower bound on as many as $\eta^{d-1} \eta' 2^{d-1} m$ such terms.⁷ Instead, we pre-compute $\widehat{\mathbf{E}}[\max(c, S_{\lambda'} c') \mid \{S_d = a\} \cap \mathcal{E}_{\alpha}^{\{d+1 \dots \infty\}}]$ for all $c, c' \in \{\frac{0}{\eta'}, \dots, \frac{\eta'}{\eta'}\}$, all $a \in \{0 \dots d\}$, and all $\alpha \in \{\alpha_1, \dots, \alpha_m\}$; and then we approximate $\widehat{\mathbf{E}}[\max(r^T, \frac{S_{\lambda'}}{\lambda'} w') \mid \mathcal{A}_T \cap \mathcal{E}_{\alpha_i}^{\{d+1 \dots \infty\}}]$ with $\widehat{\mathbf{E}}[\max(c, S_{\lambda'} c') \mid \{S_d = a\} \cap \mathcal{E}_{\alpha}^{\{d+1 \dots \infty\}}]$ where c, c' are the result of rounding r^T and $\frac{w'}{\lambda'}$ down to the nearest integer multiples of $\frac{1}{\eta'}$ respectively and $a = |T|$ and $\alpha = \alpha_i$.⁸ Notice that we only need to pre-compute $(\eta' + 1)^2 dm$. Table 2 lists the lower bound obtained for each $\lambda \in \{1 \dots 10\}$. We should also mention that we refine each configuration of intervals by cutting off infeasible regions of each interval prior to any further computation⁹. \square

5 An Upper Bound on Performance of RSOP

It is easy to see that $\mathbf{E}[\text{RSOP}] = \frac{1}{4} \text{OPT}$ for any instance with exactly two non-zero bids, yielding a straightforward upper bound of $\frac{1}{4} \text{OPT}$ on the performance of RSOP. However in such instances we always have $\lambda = 2$. That raises the question of whether the performance of RSOP approaches optimality asymptotically as $\lambda \rightarrow \infty$. We give a negative answer to the previous question by exhibiting a family of instances for which $\mathbf{E}[\text{RSOP}]$ is no more than $\frac{1}{2.65} \text{OPT}$ as $\lambda \rightarrow \infty$.

Theorem 3. For any $\lambda \geq 2$ there exists an input instance where there are λ bids above or equal to the optimal sale price and such that $\mathbf{E}[\text{RSOP}] < \frac{1}{2.65} \text{OPT}$.

Next, we define the family of *equal revenue* instances for which Feige et al. [2005] proved a lower bound of $\frac{1}{4} \text{OPT}$ on the performance of RSOP. We prove the above theorem by deriving an upper bound on the performance of RSOP on such instances.

Definition 3 (Equal Revenue Instance). An instance of bids is called an equal revenue instance if choosing any of the bids as the sale price yields the same revenue. The equal revenue instance with n non-zero distinct bids is unique (up to scaling) and given by the bid vector $\mathbf{q}^{(n)} = (q_1^{(n)}, q_2^{(n)}, \dots)$, where

$$q_j^{(n)} = \begin{cases} \frac{1}{j} & j \leq n \\ 0 & \text{otherwise} \end{cases}$$

Proposition 2. For any equal revenue instance, RSOP offers the worst price to each of the sets \mathbf{A} and \mathbf{B} . In other words, the optimal price of each set generates the least revenue when offered to the opposite set (i.e., less revenue than offering any of the other non-zero bids as the sale price).

⁷Because there are $\eta^{d-1} \eta'$ possible combinations of intervals in (21) and 2^{d-1} events of the form \mathcal{A}_T and m events of the form $\mathcal{E}_{\alpha_i}^{\{d+1 \dots \infty\}}$.

⁸It is easy to see that $\widehat{\mathbf{E}}[\max(c, \frac{S_{\lambda'}}{\lambda'} c') \mid \mathcal{A}_T \cap \mathcal{E}_{\alpha}^{\{d+1 \dots \infty\}}] = \widehat{\mathbf{E}}[\max(c, \frac{S_{\lambda'}}{\lambda'} c') \mid \{S_d = |T|\} \cap \mathcal{E}_{\alpha}^{\{d+1 \dots \infty\}}]$.

⁹For example if $v_j < v_{j+1}$, we set $v_j \leftarrow v_{j+1}$.

Proof. It follows immediately from the fact that offering any of the bids as the sale price for both sets generates a total revenue that is equal to OPT. So the price that generates highest revenue for \mathbf{A} also generates lowest revenue for \mathbf{B} and vice versa. \square

Proposition 2 suggests that, for any given λ , an equal revenue instance might actually be the worst case instance for RSOP among all instances with the same λ ; however based on computer simulation that seems not to be true at least for small values of λ .

To prove Theorem 3, we need to show the expected revenue of RSOP is no more than $\frac{1}{2.65}$ OPT for any equal revenue instance with distinct bids. However a direct analysis of the performance of RSOP for all such instances is not easy. Instead we define a modified variant of RSOP whose performance is easy to analyze, and whose revenue is close to the revenue of RSOP (e.g., asymptotically equal as $\lambda \rightarrow \infty$).

Definition 4 (RSOP*). *The modified random sampling optimal price auction behaves exactly the same way as RSOP (see Definition 1), except if all of the non-zero bids fall in the same set, it offers them the lowest non-zero bid as the sale price (instead of 0).*

Note that truthful reporting of bids is not a dominant strategy for the bidders in RSOP*, however it is only used in the analysis and not as an actual auction. Next we show that the revenue of RSOP is asymptotically equal to the revenue of RSOP*.

Lemma 12. $\mathbf{E}[\text{RSOP}] \leq \mathbf{E}[\text{RSOP}^*] \leq \mathbf{E}[\text{RSOP}] + (\frac{1}{2})^{n-1} \text{OPT}$, with the second inequality being met with equality for equal revenue instances.

Proof. Recall that RSOP* behaves exactly like RSOP except when all the n bids fall in the same set which happens with probability $(\frac{1}{2})^{n-1}$, in which case RSOP* still generates a revenue of at most OPT (exactly OPT if it is an equal revenue instance), while RSOP generates zero revenue. \square

Lemma 13. $\mathbf{E}[\text{RSOP}^*(\mathbf{q}^{(n)})]$ is a monotonically decreasing function of n .

Proof. Consider an arbitrary partition (\mathbf{A}, \mathbf{B}) .¹⁰ We prove the following equation

$$\text{RSOP}^*(\mathbf{q}^{(n)}) = \min \left(1, \left\{ \frac{j - S_j}{j} \mid j \in \mathbf{A} \cap \{1 \cdots n\} \right\} \right).$$

Increasing n only extends the range over which the minimum is taken which immediately implies the claim of the lemma. Next we prove the above equation.

Observe that $\text{RSOP}^*(\mathbf{q}^{(n)})$ is exactly equal to the revenue RSOP* obtains from \mathbf{B} because no revenue is obtained from \mathbf{A} which is offered the optimal price of \mathbf{B} which is 1. If $\mathbf{A} \cap \{1 \cdots n\} \neq \emptyset$, then the optimal price of \mathbf{A} is $\frac{1}{\lambda_{\mathbf{A}}}$ where $\lambda_{\mathbf{A}} \in \arg \max_{j \in \mathbf{A} \cap \{1 \cdots n\}} \frac{S_j}{j}$ and the revenue obtained by offering that price to \mathbf{B} is exactly $\min_{j \in \mathbf{A} \cap \{1 \cdots n\}} \frac{j - S_j}{j}$; otherwise if $\mathbf{A} \cap \{1 \cdots n\} = \emptyset$, the price offered to \mathbf{B} , by definition of RSOP*, is exactly $\frac{1}{n}$ which yields a revenue of 1. That completes the proof. \square

The following is obtained by direct calculation using a computer.

Proposition 3. $\mathbf{E}[\text{RSOP}^*(\mathbf{q}^{(400)})] = 0.377208 \pm 10^{-6}$.

We now prove the main theorem of this section.

Proof of Theorem 3. To prove the theorem for any λ we exhibit a bid vector \mathbf{v} with λ bids above the optimal sale price such that $\mathbf{E}[\text{RSOP}(\mathbf{v})] < \frac{1}{2.65} \text{OPT}(\mathbf{v})$. Let $n = \max(\lambda, 400)$, then

$$\begin{aligned} \mathbf{E} \left[\text{RSOP}(\mathbf{q}^{(n)}) \right] &\leq \mathbf{E} \left[\text{RSOP}^*(\mathbf{q}^{(n)}) \right] && \text{by Lemma 12} \\ &\leq \mathbf{E} \left[\text{RSOP}^*(\mathbf{q}^{(400)}) \right] && \text{by Lemma 13} \\ &< 0.377209 && \text{by Proposition 3} \\ &< \frac{1}{2.65} \text{OPT}(\mathbf{q}^{(n)}) && \text{because } \text{OPT}(\mathbf{q}^{(n)}) = 1 \end{aligned}$$

¹⁰Recall that by our definition there are infinitely many zero bids so (\mathbf{A}, \mathbf{B}) is a partitioning of \mathbb{N} .

Observe that the optimal sale price for $\mathbf{q}^{(n)}$ is not unique. Let \mathbf{v} be the same as $\mathbf{q}^{(n)}$ everywhere except $v_\lambda = q_\lambda^{(n)} + \epsilon$ for a small $\epsilon \in (0, \frac{1}{\lambda^2})$. Observe that v_λ is now the unique optimal sale price for \mathbf{v} . It is easy to see that $\lim_{\epsilon \rightarrow 0} \mathbf{E}[\text{RSOP}(\mathbf{v})] = \mathbf{E}[\text{RSOP}(\mathbf{q}^{(n)})]$ and $\lim_{\epsilon \rightarrow 0} \text{OPT}(\mathbf{v}) = \text{OPT}(\mathbf{q}^{(n)}) = 1$, so for a small enough ϵ , we get $\mathbf{E}[\text{RSOP}(\mathbf{v})] < \frac{1}{2.65} \text{OPT}(\mathbf{v})$ which completes the proof. \square

6 A Combinatorial Lower Bound

In this section we present a combinatorial approach for obtaining a lower bound on the expected revenue of RSOP for equal revenue instance where each non-zero bid is either 1 or h (for some fixed $h \in \mathbb{N}$). We hope the ideas we present in the section help develop a more general combinatorial approach in the future for proving lower bounds on mechanisms based on random sampling.

Observe that in an equal revenue instance where non-zero bids are either h or 1, if there are k bids of value h , there must be $k(h-1)$ bids of value 1. Throughout the rest of this section we assume h is an implicit constant. The following theorem summarizes the main result of this section.

Theorem 4. *For any equal revenue instance where each non-zero bid is either h or 1,*

$$\mathbf{E}[\text{RSOP}] \geq \left(\frac{1}{2} + \frac{1}{2h} - \frac{1}{2^{kh-1}} \right) \text{OPT} \quad (31)$$

where k is the number of bids of value h .

Observe that in the above theorem the worst case of the lower bound is when $k = 1$ and $h = 2$ for which the lower bound becomes $\text{OPT}/4$. Notice that the lower bound approaches $\text{OPT}/2$ quickly as either k or h increases.

Definition 5. $Q^{(k)}$ denotes the multi-set of bids corresponding to an equal revenue instance with k bids of value h and $k(h-1)$ bids of value 1.

For the rest of this section we assume that \mathbf{A} and \mathbf{B} are multi-sets containing the actual bids in each side of the partition, as opposed to the previous sections where we assumed \mathbf{A} and \mathbf{B} contained the indices of those bids. Furthermore, for any multi-set of bids such as \mathbf{I} , we use the notation $\mathbf{E}[\text{RSOP}(\mathbf{I})]$, $\mathbf{E}[\text{RSOP}^*(\mathbf{I})]$ and $\text{OPT}(\mathbf{I})$ to denote the respective quantity being computed on bids explicitly specified by \mathbf{I} . We also make no assumption about which of \mathbf{A} or \mathbf{B} gets the highest bid, unless explicitly stated otherwise.

We start by proving a lower bound on the expected revenue of RSOP^* (see Definition 4) for equal revenue instances where each non-zero bid is either h or 1. We then extend the lower bound to RSOP . Recall that RSOP^* behaves exactly the same way as RSOP , except if all non-zero bids fall in the same set, RSOP^* offers them the lowest non-zero bid as the sale price (instead of 0).

Lemma 14. *For any $k \in \mathbb{N}$,*

$$\mathbf{E}[\text{RSOP}^*(Q^{(k)})] \geq \frac{k(h+1)}{2} = \frac{1}{2} \text{OPT}(Q^{(k)}) + \frac{k}{2}. \quad (32)$$

Proof. We prove the claim by induction on k .

We first prove the base case which is $k = 1$. The single bid of value h is the highest bid. Without loss of generality assume that the h bid is in \mathbf{B} . Observe that the optimal price of \mathbf{B} is h which is also the price offered to \mathbf{A} , so no revenue is obtained from \mathbf{A} . Furthermore the optimal price of \mathbf{A} is 1 which is also the price offered to \mathbf{B} . Each bid of value 1 falls in \mathbf{B} with probability $1/2$, so $\mathbf{E}[\text{RSOP}^*(Q^{(1)})] = 1 + \frac{h-1}{2}$ which proves the base of the induction.

We now prove the induction step. For any two multi-sets of bids such as \mathbf{T} and \mathbf{U} , let $\text{Rev}^*(\mathbf{T}, \mathbf{U})$ denote the revenue obtained from \mathbf{T} by computing the optimal sale price for \mathbf{U} (let the optimal price be 1 if $\mathbf{U} = \emptyset$) and offering that price to \mathbf{T} ; also let $\text{Rev}(\mathbf{T}, p)$ denote the revenue obtained by offering price p to \mathbf{T} . Let (\mathbf{A}, \mathbf{B}) be a random partition of $Q^{(1)}$, and let $(\mathbf{A}', \mathbf{B}')$ be a random partition of $Q^{(k-1)}$. Observe that

$(\mathbf{A} \cup \mathbf{A}', \mathbf{B} \cup \mathbf{B}')$ is a random partition of $Q^{(k)}$. The induction step follows from the following inequities.

$$\begin{aligned}
\mathbf{E} \left[\text{RSOP}^*(Q^{(k)}) \right] &= \mathbf{E} \left[\text{Rev}^*(\mathbf{A} \cup \mathbf{A}', \mathbf{B} \cup \mathbf{B}') + \text{Rev}^*(\mathbf{B} \cup \mathbf{B}', \mathbf{A} \cup \mathbf{A}') \right] \\
&\geq \mathbf{E} \left[\text{Rev}^*(\mathbf{A}, \mathbf{B}) + \text{Rev}^*(\mathbf{A}', \mathbf{B}') + \text{Rev}^*(\mathbf{B}, \mathbf{A}) + \text{Rev}^*(\mathbf{B}', \mathbf{A}') \right] \quad \text{to be proven} \\
&= \mathbf{E} \left[\text{RSOP}^*(Q^{(1)}) \right] + \mathbf{E} \left[\text{RSOP}^*(Q^{(k-1)}) \right] \\
&\geq \frac{h+1}{2} + \frac{(k-1)(h+1)}{2} > \frac{k(h+1)}{2} \quad \text{by the induction hypothesis}
\end{aligned}$$

We shall prove $\text{Rev}^*(\mathbf{B} \cup \mathbf{B}', \mathbf{A} \cup \mathbf{A}') > \text{Rev}^*(\mathbf{B}, \mathbf{A}) + \text{Rev}^*(\mathbf{B}', \mathbf{A}')$ and by symmetry we can argue $\text{Rev}^*(\mathbf{A} \cup \mathbf{A}', \mathbf{B} \cup \mathbf{B}') > \text{Rev}^*(\mathbf{A}, \mathbf{B}) + \text{Rev}^*(\mathbf{A}', \mathbf{B}')$ which completes the proof. Let p, p' and p'' denote the optimal price of \mathbf{A}, \mathbf{A}' and $\mathbf{A} \cup \mathbf{A}'$ respectively as computed by RSOP^* (i.e., the optimal price for an empty set would be 1). We argue that

$$\begin{aligned}
\text{Rev}^*(\mathbf{B} \cup \mathbf{B}', \mathbf{A} \cup \mathbf{A}') &= \text{Rev}(\mathbf{B}, p'') + \text{Rev}(\mathbf{B}', p'') \\
&\geq \text{Rev}(\mathbf{B}, p) + \text{Rev}(\mathbf{B}', p') \quad \text{explained below} \\
&= \text{Rev}^*(\mathbf{B}, \mathbf{A}) + \text{Rev}^*(\mathbf{B}', \mathbf{A}').
\end{aligned}$$

Observe that both $\mathbf{A} \cup \mathbf{B}$ and $\mathbf{A}' \cup \mathbf{B}'$ are equal revenue instances and by Proposition 2 in any equal revenue instance the price that is optimal for one side generates the least revenue for the opposite side so $\text{Rev}(\mathbf{B}, p'') \geq \text{Rev}(\mathbf{B}, p)$ and $\text{Rev}(\mathbf{B}', p'') \geq \text{Rev}(\mathbf{B}', p')$. □

Proof of Theorem 4. Recall that the only situation where RSOP^* and RSOP behave differently is when either \mathbf{A} or \mathbf{B} is empty which happens with probability $1/2^{kh-1}$, therefore

$$\begin{aligned}
\mathbf{E} \left[\text{RSOP}(Q^{(k)}) \right] &= \mathbf{E} \left[\text{RSOP}^*(Q^{(k)}) \right] - \frac{1}{2^{kh-1}} \text{OPT} \\
&\geq \frac{k(h+1)}{2} - \frac{1}{2^{kh-1}} \text{OPT} \quad \text{by Lemma 14} \\
&= \left(\frac{1}{2} + \frac{1}{2h} - \frac{1}{2^{kh-1}} \right) \text{OPT} \quad \text{because } \text{OPT} = kh
\end{aligned}$$

That completes the proof. □

7 Acknowledgment

We would like to thank Jason Hartline for several valuable discussions. We also thank the anonymous referees for their helpful and detailed comments.

References

- M.-F. Balcan, A. Blum, J. D. Hartline, and Y. Mansour. Mechanism design via machine learning. In *FOCS*, pages 605–614, 2005.
- S. Baliga and R. Vohra. Market research and market design. *Advances in Theoretical Economics*, 3(1): 1059–1059, 2003.
- N. R. Devanur and J. D. Hartline. Limited and online supply and the bayesian foundations of prior-free mechanism design. In *ACM Conference on Electronic Commerce*, pages 41–50, 2009.
- U. Feige, A. Flaxman, J. D. Hartline, and R. D. Kleinberg. On the competitive ratio of the random sampling auction. In *WINE*, pages 878–886, 2005.
- C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics*, 22(2):89–103, 1971.
- A. V. Goldberg and J. D. Hartline. Competitive auctions for multiple digital goods. In *ESA '01: Proceedings of the 9th Annual European Symposium on Algorithms*, pages 416–427, London, UK, 2001. Springer-Verlag. ISBN 3-540-42493-8.
- A. V. Goldberg, J. D. Hartline, and A. Wright. Competitive auctions and digital goods. In *SODA '01: Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, pages 735–744, Philadelphia, PA, USA, 2001. Society for Industrial and Applied Mathematics. ISBN 0-89871-490-7.
- A. V. Goldberg, J. D. Hartline, A. R. Karlin, M. Saks, and A. Wright. Competitive auctions. *Games and Economic Behavior*, 55(2):242–269, May 2006.
- M. T. Hajiaghayi, R. D. Kleinberg, and D. C. Parkes. Adaptive limited-supply online auctions. In *ACM Conference on Electronic Commerce*, pages 71–80, 2004.
- J. D. Hartline and T. Roughgarden. Optimal mechanism design and money burning. *CoRR*, abs/0804.2097, 2008.
- W. Hoeffding. Probability inequalities for sums of bounded random variables. *American Statistical Association Journal*, 58:13–30, 1963.
- I. Segal. Optimal pricing mechanisms with unknown demand. *American Economic Review*, 93(3):509–529, June 2003.

A Results

λ	$\mathbf{E}[\text{RSOP}]/\text{OPT}$	$\text{OPT}/\mathbf{E}[\text{RSOP}]$
2	0.125148	7.99
3	0.166930	5.99
4	0.192439	5.20
5	0.209222	4.78
6	0.221407	4.52
7	0.230605	4.34
8	0.237862	4.20
9	0.243764	4.10
10	0.248647	4.02
11	0.252774	3.96
15	0.264398	3.78
20	0.273005	3.66
30	0.282297	3.54
50	0.290384	3.44
100	0.296993	3.37
200	0.300549	3.33
300	0.301784	3.31
500	0.302792	3.30
1000	0.303560	3.29
1500	0.303818	3.29
2000	0.303949	3.29

Table 1: Computed numerical values for the basic lower-bound of $\mathbf{E}[\text{RSOP}] \geq \mathbf{E}[\frac{S_\lambda}{\lambda} Z] \text{OPT}$.

λ	$\mathbf{E}[\text{RSOP}]/\text{OPT}$	$\text{OPT}/\mathbf{E}[\text{RSOP}]$
2	0.2138	4.68
3	0.2178	4.59
4	0.238	4.20
5	0.243	4.11
6	0.2503	3.99
7	0.2545	3.93
8	0.2602	3.84
9	0.2627	3.81
10	0.2669	3.75

Table 2: Computed numerical values for the exhaustive-search lower-bound

B Proofs

Theorem 5 (Chernoff–Hoeffding1963). *For (i.i.d.) random variables $X_1, X_2, \dots, X_\ell \in \{0, 1\}$ with $\mathbf{E}[X_i] = p$, the following inequality holds for all $\varepsilon \in (0, 1 - p)$:*

$$\Pr \left[\frac{1}{\ell} \sum X_i \geq p + \varepsilon \right] \leq \left(\left(\frac{p}{p + \varepsilon} \right)^{p + \varepsilon} \left(\frac{1 - p}{1 - p - \varepsilon} \right)^{1 - p - \varepsilon} \right)^\ell \quad (33)$$

Lemma (2). *For any $\alpha \in (0, 1)$ and $j \in \mathbb{N}$,*

$$\begin{aligned} \text{if } \alpha \geq 0.5, \text{ then } & \Pr \left[\mathcal{E}_\alpha^{\{j\}} \right] \geq 1 - (r_\alpha)^j, & \text{where } r_\alpha &= \frac{1}{2\alpha^\alpha(1-\alpha)^{1-\alpha}} \\ \text{if } \alpha \leq 0.5 - 1/j, \text{ then } & \Pr \left[\mathcal{E}_\alpha^{\{j\}} \right] \leq (r_{(\alpha+1/j)})^{j-1} & \text{where } r_\alpha & \text{ is the same as above.} \end{aligned}$$

Proof. Let A_j be an indicator random variable which is 1 if $j \in \mathbf{A}$, and 0 otherwise.

The first inequality of the lemma follows immediately from Theorem 5 by setting $X_j = A_j$, $\ell = j$, $p = 0.5$, and $\varepsilon = \alpha - 0.5$ which yields an upper bound on $\Pr[\overline{\mathcal{E}}_\alpha^{\{j\}}]$ and thus a lower bound on $\Pr[\mathcal{E}_\alpha^{\{j\}}]$. Note that $A_1 = 0$ with probability 1, however that only decreases the probability on the left hand side of (33) so it still holds.

To prove the second inequality, we proceed as follows.

$$\begin{aligned}
\Pr \left[\mathcal{E}_\alpha^{\{j\}} \right] &= \Pr \left[\frac{S_j}{j} \leq \alpha \right] = \Pr \left[\frac{\sum_{k=1}^j \bar{A}_k}{j} > 1 - \alpha \right] \\
&= \Pr \left[\frac{\sum_{k=2}^j \bar{A}_k}{j} > 1 - \alpha - \frac{1}{j} \right] && \text{because } \bar{A}_1 = 1 \text{ always.} \\
&\leq \Pr \left[\frac{\sum_{k=2}^j \bar{A}_k}{j-1} > 1 - \alpha - \frac{1}{j} \right].
\end{aligned}$$

The second inequality of the lemma now follows immediately from Theorem 5 by setting $X_j = \bar{A}_{j-1}$, $\ell = j-1$, $p = 0.5$, and $\epsilon = 0.5 - \alpha - \frac{1}{j}$. Note that $r_{1-\alpha-\frac{1}{j}} = r_{\alpha+\frac{1}{j}}$. \square

Theorem 6 (Fortuin, Kasteleyn, and Ginibre [1971]). *Let L be a finite distributive lattice, and $\mu : L \rightarrow \mathbb{R}_+$ be a function that satisfies*

$$\mu(x \wedge y)\mu(x \vee y) \geq \mu(x)\mu(y), \quad \text{for all } x, y \in L. \quad (34)$$

Then for any two functions $f, g : L \rightarrow \mathbb{R}_+$ which are either both increasing, or both decreasing, the following inequality holds.

$$\left(\sum_{x \in L} f(x)g(x)\mu(x) \right) \left(\sum_{x \in L} \mu(x) \right) \geq \left(\sum_{x \in L} f(x)\mu(x) \right) \left(\sum_{x \in L} g(x)\mu(x) \right) \quad (35)$$

Lemma (1). *For any $T, T' \subset \mathbb{N}$ and $\alpha \in [0, 1]$, the two events \mathcal{E}_α^T and $\mathcal{E}_\alpha^{T'}$ are positively correlated, i.e., $\Pr[\mathcal{E}_\alpha^T \cap \mathcal{E}_\alpha^{T'}] \geq \Pr[\mathcal{E}_\alpha^T] \Pr[\mathcal{E}_\alpha^{T'}]$.*

Proof. For every $n \in \mathbb{N}$, define $T_n = T \cap \{1 \cdots n\}$; similarly define $T'_n, \mathbf{A}_n, \mathbf{B}_n$, etc.

We start by proving $\Pr[\mathcal{E}_\alpha^{T_n} \cap \mathcal{E}_\alpha^{T'_n}] \geq \Pr[\mathcal{E}_\alpha^{T_n}] \Pr[\mathcal{E}_\alpha^{T'_n}]$ for every $n \in \mathbb{N}$. Let L_n be a distributive lattice whose elements are the subsets of $\{2 \cdots n\}$ and whose meet/join operators correspond to taking intersection/union. For all $A \in L_n$ let $\mu(A) = 1/2^{n-1}$. Define $\mathcal{E}_\alpha^{T_n}(A)$ to be an indicator function which is defined for each $A \in L_n$ as

$$\mathcal{E}_\alpha^{T_n}(A) = \begin{cases} 1 & \text{if } |A \cap \{1 \cdots j\}| \leq \alpha j \text{ for all } j \in T_n \\ 0 & \text{otherwise} \end{cases}.$$

By invoking Theorem 6 on lattice L_n and substituting $f(x)$ and $g(x)$ with $\mathcal{E}_\alpha^{T_n}(A)$ and $\mathcal{E}_\alpha^{T'_n}(A)$ respectively we get the following inequality.

$$\left(\sum_{A \subseteq \{2 \cdots n\}} \frac{\mathcal{E}_\alpha^{T_n}(A)\mathcal{E}_\alpha^{T'_n}(A)}{2^{n-1}} \right) \geq \left(\sum_{A \subseteq \{2 \cdots n\}} \frac{\mathcal{E}_\alpha^{T_n}(A)}{2^{n-1}} \right) \left(\sum_{A \subseteq \{2 \cdots n\}} \frac{\mathcal{E}_\alpha^{T'_n}(A)}{2^{n-1}} \right)$$

Observe that the left hand side of the above inequality is exactly $\mathbf{E}_A[\mathcal{E}_\alpha^{T_n}(\mathbf{A})\mathcal{E}_\alpha^{T'_n}(\mathbf{A})] = \Pr[\mathcal{E}_\alpha^{T_n} \cap \mathcal{E}_\alpha^{T'_n}]$ while its right hand side is exactly $\mathbf{E}_A[\mathcal{E}_\alpha^{T_n}(\mathbf{A})] \mathbf{E}_A[\mathcal{E}_\alpha^{T'_n}(\mathbf{A})] = \Pr[\mathcal{E}_\alpha^{T_n}] \Pr[\mathcal{E}_\alpha^{T'_n}]$, so we have proved that $\Pr[\mathcal{E}_\alpha^{T_n} \cap \mathcal{E}_\alpha^{T'_n}] \geq \Pr[\mathcal{E}_\alpha^{T_n}] \Pr[\mathcal{E}_\alpha^{T'_n}]$ for every $n \in \mathbb{N}$.

We now prove the infinite case. For every $n \in \mathbb{N}$, define $\ell_n = \Pr[\mathcal{E}_\alpha^T \cap \mathcal{E}_\alpha^{T'_n}]$, $r_n = \Pr[\mathcal{E}_\alpha^T] \Pr[\mathcal{E}_\alpha^{T'_n}]$, and $d_n = \ell_n - r_n$. Observe that d_n is an infinite sequence which is bounded in $[0, 1]$, so by invoking BolzanoWeierstrass theorem we argue that it has an infinite converging subsequence, i.e., there exists an infinite sequence of indices $n_1 < n_2 < \cdots$ and $d^* \in [0, 1]$ such that $\lim_{j \rightarrow \infty} d_{n_j} = d^*$. On the other hand both ℓ_{n_j} and r_{n_j} are decreasing sequences which are bounded below by 0 so they both converge, therefore

$$\Pr \left[\mathcal{E}_\alpha^T \cap \mathcal{E}_\alpha^{T'} \right] - \Pr \left[\mathcal{E}_\alpha^T \right] \Pr \left[\mathcal{E}_\alpha^{T'} \right] = \lim_{j \rightarrow \infty} \ell_{n_j} - \lim_{j \rightarrow \infty} r_{n_j} = \lim_{j \rightarrow \infty} d_{n_j} = d^* \geq 0$$

which proves the claim of the lemma. \square