

The Vickrey Auction with a Single Duplicate Bidder Approximates the Optimal Revenue

Hu Fu* Christopher Liaw* Sikander Randhawa*

May 10, 2019

Bulow and Klemperer’s well-known result states that, in a single-item auction where the n bidders’ values are independently and identically drawn from a regular distribution, the Vickrey auction with one additional bidder (a duplicate) extracts at least as much revenue as the optimal auction without the duplicate. Hartline and Roughgarden, in their influential 2009 paper, removed the requirement that the distributions be identical, at the cost of allowing the Vickrey auction to recruit n duplicates, one from each distribution, and relaxing its revenue advantage to a 2-approximation.

In this work we restore Bulow and Klemperer’s number of duplicates in Hartline and Roughgarden’s more general setting with a worse approximation ratio. We show that recruiting a duplicate from one of the distributions suffices for the Vickrey auction to 10-approximate the optimal revenue. We also show that in a k -items unit demand auction, recruiting k duplicates suffices for the VCG auction to $O(1)$ -approximate the optimal revenue.

As another result, we tighten the analysis for Hartline and Roughgarden’s Vickrey auction with n duplicates for the case with two bidders in the auction. We show that in this case the Vickrey auction with two duplicates obtains at least $3/4$ of the optimal revenue. This is tight by meeting a lower bound by Hartline and Roughgarden. En route, we obtain a transparent analysis of their 2-approximation for n bidders, via a natural connection to Ronen’s lookahead auction.

*Department of Computer Science, University of British Columbia. {hufu, cvliaw, srand}@cs.ubc.ca

1 Introduction

Bulow and Klemperer’s theorem [3] is a fundamental result in the auction literature, drawing a connection between two basic auctions for selling a single item. The second price auction, also known as the Vickrey auction [18], lets the highest bidder win the item at the price of the second highest bid. It always sells to the highest bidder, thereby maximizing the market’s efficiency, requires no prior information on the bidders, and is easy to implement in practice. On the other hand, revenue optimal auctions, those that generate the most revenue for the auctioneer, are usually more intricate. They may sometimes not sell to anyone, or sell to a bidder who does not place the highest bid, and their rules in general depend on prior information on bidders’ values [14]. Bulow and Klemperer showed that, in the case where bidders’ values are drawn independently and identically from a distribution that satisfies a commonly assumed property known as regularity, the simple Vickrey auction with $n + 1$ bidders extracts no less revenue than the revenue optimal auction with n bidders, for any $n \geq 1$. This fundamental result makes an elegant comparison between the revenue impact brought by enhanced competition and that attributed to more precise market information and tailored auction design. It holds for a large family of distributions, as long as the bidders’ values are i.i.d. drawn.

Hartline and Roughgarden [11], in their seminal work that started a fruitful line of research on the revenue of simple auctions (as exemplified by the Vickrey auction), extended Bulow and Klemperer’s result, with approximation, when one removes the assumption of symmetry, i.e., when bidders’ values are drawn independently but not identically from regular distributions. This extension is approximate for two changes it makes in the comparison between the two auctions. First, the Vickrey auction, instead of recruiting only one additional bidder, is now allowed to recruit n additional bidders (called duplicates), one from each distribution. Second, the revenue of this Vickrey auction with n duplicates, instead of being lower bounded by the optimal revenue without duplicates, is shown to be at least half as much. Bidders with different value distributions are often interpreted as coming from different populations or possessing different characteristics. Under this interpretation, the Vickrey auction is shown to approximate the optimal revenue if one is allowed to recruit at least one additional bidder from each population. Hartline and Roughgarden also gave an example with two bidders where the Vickrey auction, even with two duplicate bidders, one from each distribution, extracts only $3/4$ of the revenue of the optimal auction without duplicates.

The two changes made in Hartline and Roughgarden’s result leave two gaps between the behavior of auctions with i.i.d. bidders and those with independent but non-identical ones. The first gap is a necessary one: Hartline and Roughgarden’s example shows that, absent symmetry among bidders, a constant fraction of revenue has to be lost by using the simpler Vickrey auction, even when all bidders are duplicated. The other gap has been an open question: *how many duplicate bidders* are needed for the Vickrey auction to give a constant approximation to the optimal revenue? Whereas for the i.i.d. case a single duplicate bidder suffices, it has remained unknown whether n duplicates are necessary when the value distributions are not identical.

1.1 Main Results

Vickrey auction with a single duplicate. In this work, we close this second gap. We show that, in a single-item auction with n bidders, whose values are drawn independently from possibly different regular distributions, the Vickrey auction with *one* additional bidder with value drawn from *one of the existing distributions* obtains at least $\frac{1}{10}$ of the optimal revenue extractable from the original bidders.¹ Therefore, to retain a constant fraction of the optimal revenue, the number

¹In fact, it extracts at least $\frac{1}{10}$ of the optimal *ex ante* revenue from the original bidders.

of duplicate bidders needed in the non-i.i.d. setting is in fact no more than in the i.i.d. setting.

Our analysis parts ways from the very beginning with that by Hartline and Roughgarden, who use the Myersonian “virtual surplus” as an (exact) upper bound on the optimal revenue without duplicates. Instead, we work with the looser bound given by the ex ante optimal revenue. The looseness of this bound is compensated by the ease with which it allows connections to be drawn between the optimal revenue and values (or prices) in bidders’ distributions, and all such quantities across bidders to be manipulated in a linear fashion. With this, we identify a set of bidders that, in total, contribute a constant fraction of the ex ante optimal revenue by using high prices (so high as to be comparable to the optimal revenue). Two observations are in order: (a) if a bidder accepts a high price with a constant probability, then duplicating this bidder alone suffices for the Vickrey auction to extract a constant fraction of the optimal revenue; and (b) if there is a high price such that the *sum* of probabilities with which bidders take the price is a constant, then the VCG auction even *without* duplicates is a constant approximation. A technical lemma (Lemma 3.2) shows that we must, in fact, be in one of these scenarios. The lemma makes use of regularity and the set of bidders we identified in the previous step.

Choosing the distribution to duplicate. Naturally, the distribution to be duplicated needs to be carefully chosen; for instance, if most bidders’ values are constantly 0 or negligible, duplicating an arbitrary or a random bidder helps little with the revenue. We show several forms of minimal information that the auctioneer could acquire in order to choose the distribution to duplicate: to guarantee a constant approximation achieved by the Vickrey auction with a duplicate, it suffices that the auctioneer knows (approximately) the value of each bidder at a specified quantile; that is, the auctioneer only needs to know, for each bidder i , what is roughly the value v_i such that with probability q bidder i bids above v_i , where q is a fixed probability, say, 0.5. Note that the approximation ratio may depend on how well the auctioneer can query the value of a bidder at the specified quantile q as well as the quantile itself. It also suffices if the auctioneer has a single sample from each distribution albeit with a different constant approximation (see Theorem 3.4). Our result may be seen therefore as describing a tradeoff between reducing the number of duplicate bidders and acquiring some minimal knowledge on the value distributions.

We remark that we do not claim that the best use of such prior information is to find a duplicate and run the Vickrey auction. In fact, with the said types of information, one may run other auctions for comparable revenue guarantees. For example, with a single sample from each distribution, one may use the maximum of these samples as an anonymous reserve in the Vickrey auction and guarantee a 4-approximation [11]. We see the main contribution of this work as conceptually closing the long-standing gap in the understanding on the number of duplicates needed for the Vickrey auction to be approximately revenue optimal. The Vickrey auction is fundamental and ubiquitous, and, partly owing to this, Bulow and Klemperer’s original result is a central link in the auction literature; our result extends it to non-i.i.d. settings in a spirit close to the original. The results on choosing the duplicate are natural consequences of our techniques used to prove the main result.

On the other hand, as we discuss at the end of the paper, the upper bounds we give are unlikely to be tight; for all that we know, it is possible that a tighter analysis could show the VCG auction with a duplicate to be a comparable or even better way to utilize the kind of partial information we discussed. We leave this intriguing question to future work.

k duplicates for k -items auctions with unit demands. We extend our result to k -items auctions with unit demands, where there are k identical items for sale and each bidder needs only

one of them. The generalization of the Vickrey auction here is the VCG auction, which sells to the highest k bidders at the price of the $(k + 1)$ -st highest bid. A simple example shows that one needs more duplicates to generate *any* revenue in the VCG auction: let there be one bidder whose value is 1 and all the other bidders have value 0. Here even if we duplicate all bidders, the VCG auction still has revenue 0, whereas the optimal revenue is 1. To address this, we need to be allowed to make *multiple* duplicates of the same bidder. The above example shows that k duplicates are necessary for the VCG auction to give any approximation. We show that this is tight: as long as the bidders' values are drawn independently from regular distributions, there exists a bidder such that one may add k duplicates of her to the VCG auction and guarantee a constant fraction of the optimal revenue without duplicates.

Hartline and Roughgarden [11] addressed the issue in the example in a different way: the VCG auction may still add only one duplicate for each bidder, but each original bidder and her duplicate can win at most one item. With this restriction, they show that the VCG auction with a duplicate for every bidder is again a 2-approximation. We extend our analysis to this scenario as well, and show that duplicating k bidders suffices for the VCG auction to be a constant approximation, albeit with a smaller constant. The following example shows the tightness of our result: let $k/3$ bidders have value 1 and all other bidders have value 0, then one must duplicate $\Omega(k)$ bidders to guarantee a constant approximation.

Tighter analysis of the Vickrey auction with n duplicates. Since Hartline and Roughgarden first gave the upper bound of 2 and lower bound of $\frac{4}{3}$ on the approximation factor of the Vickrey auction with n duplicates, it has remained an open question what this factor is in the worst case. We make progress towards resolving this by showing that for $n = 2$, $\frac{4}{3}$ is the tight bound. We first give an alternate proof of Hartline and Roughgarden's original 2-approximation. Our proof uses the geometry of the so-called revenue curves and lower bounds the revenue of the Vickrey auction with n duplicates by that of Ronen's *lookahead auction* [15], which is known to 2-approximate the optimal revenue. We then further exploit properties of the Vickrey auction and also switch to a *stronger* benchmark, the so-called *ex ante* optimal revenue. This allows us to identify a small family of distributions as scenarios; a thorough analysis of these reveals the tight ratio of $\frac{4}{3}$.

Related literature. From the large literature on “simple versus optimal auctions” that followed Hartline and Roughgarden [11], we point out a few that are closer to our work. Sivan and Syrgkanis [17] extended Hartline and Roughgarden's result to distributions that are convex combinations of regular ones (and the duplicate bidders are drawn from each component regular distribution). Roughgarden et al. [16] showed a Bulow-Klemperer type result for matching environments. Eden et al. [7] and Feldman et al. [8] studied, for auctions selling multiple *heterogenous* items, the number of duplicate bidders needed for the VCG auction's revenue to approximate the optimal without duplicates. The bidders' preferences are *i.i.d.* drawn, and the non-trivial number of duplicates needed is a consequence of multi-dimensional preferences rather than asymmetry among bidders.

Our results on the minimal information needed to choose the bidder to duplicate also place the work in the literature on prior-independent mechanisms [e.g. 2, 6, 9, 5]. In particular, we show in Theorem 3.4 that a *single sample* in our setting also suffices to choose the duplicate.

Organization of the paper In Section 2 we set up the model formally and give preliminaries. In Section 3 we show our result on duplicating few bidders, first for single-item auctions and then for k -items auctions with unit demands in Section 4. In Section 5 we show our tighter analysis for the auction with n duplicated bidders. We conclude with some discussion in Section 6.

2 Preliminaries

Single-item auctions. A single item is to be sold to n bidders. Each bidder i values the item at v_i , which is private to i and independently drawn from a publicly known distribution F_i . We use F_i interchangeably as both the distribution of bidder i 's value as well as the cumulative density function (cdf) of the distribution, i.e. $F_i(w) = \Pr[v_i \leq w]$. We will also use the notation $\mathbf{v} = (v_1, \dots, v_n)$ for the vector containing the values of all bidders and $\mathbf{v}_{-i} = (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$ for the vector *excluding* bidder i 's value.

A selling mechanism consists of an allocation rule $\mathbf{x}(\mathbf{v}) = (x_1(\mathbf{v}), \dots, x_n(\mathbf{v}))$, and a payment rule $\mathbf{p}(\mathbf{v}) = (p_1(\mathbf{v}), \dots, p_n(\mathbf{v}))$, where $x_i(\mathbf{v}) \in [0, 1]$ indicates the probability that bidder i receives the item when the bid vector is \mathbf{v} , and $p_i(\mathbf{v}) \in \mathbb{R}$ indicates the amount that bidder i must pay. The utility of bidder i is quasi-linear, given by $u_i = x_i(\mathbf{v})v_i - p_i(\mathbf{v})$. Note that we must enforce the condition $\sum_i x_i(\mathbf{v}) \leq 1$ for all \mathbf{v} as there is only one item for sale.

A mechanism is said to be *dominant strategy incentive compatible* (DSIC) if for all i , any value profile \mathbf{v} , it holds that for any deviation v'_i , we have

$$x_i(\mathbf{v})v_i - p_i(\mathbf{v}) \geq x_i(v'_i, \mathbf{v}_{-i})v_i - p_i(v'_i, \mathbf{v}_{-i}). \quad (1)$$

A mechanism is said to be *Bayesian incentive compatible* (BIC) if for any i , any value v_i and any deviation v'_i , we have

$$\mathbf{E}_{\mathbf{v}_{-i}} [x_i(v_i, \mathbf{v}_{-i})v_i - p_i(v_i, \mathbf{v}_{-i})] \geq \mathbf{E}_{\mathbf{v}_{-i}} [x_i(v'_i, \mathbf{v}_{-i})v_i - p_i(v_i, \mathbf{v}_{-i})]. \quad (2)$$

A mechanism is said to be *ex-post individually rational* (IR) if a bidder's utility is always nonnegative, i.e. for all i and \mathbf{v} , we have

$$x_i(\mathbf{v})v_i - p_i(\mathbf{v}) \geq 0. \quad (3)$$

As will be clear in the summary of Myerson's work below, for our purpose there is no need to make a distinction between DSIC and BIC. Without loss of generality, we focus in this work on DSIC and ex-post IR mechanisms. In such mechanisms, one may assume that bidders bid truthfully. The *revenue* of a mechanism is then $\mathbf{E}_{\mathbf{v}}[\sum_i p_i(\mathbf{v})]$.

Myerson's characterization of optimal mechanisms. Myerson [14] completely characterized the revenue of incentive compatible auctions:

- Lemma 2.1** (Myerson [14]).
1. *The payment rule of any BIC mechanism is uniquely determined by its allocation rule up to a constant. With this, a mechanism is BIC if and only if its allocation rule $x_i(v_i, \mathbf{v}_{-i})$ for each i is monotone nondecreasing in v_i for any \mathbf{v}_{-i} .*
 2. *For any BIC mechanism with allocation rule \mathbf{x} , its expected revenue, up to a constant, is given by $\mathbf{E}_{\mathbf{v}}[\sum_i x_i(v_i)\varphi_i(v_i)]$, where $\varphi_i(v_i) := v_i - \frac{1-F_i(v_i)}{f_i(v_i)}$ is called the virtual value of v_i .*
 3. *There exists a DSIC and ex-post IR mechanism that is revenue optimal amongst all BIC and ex-post IR mechanisms.*

Throughout the paper, we fix the constant so that the lowest type in the support has expected utility 0.

A distribution is *regular* if the virtual value $\varphi(v)$ is monotone nondecreasing in v . Given the characterization, it is immediate that, for regular distributions, the revenue-optimal auction allocates to the bidder with the highest virtual non-negative value; if all virtual values are negative, the item is not sold. We will refer to the revenue of Myerson's optimal auction as the *optimal revenue*.

Revenue curves. Bulow and Roberts [4] gave an influential reinterpretation of Myerson’s characterization, and we will use heavily this interpretation. Given a bidder whose value is drawn from distribution F , for every take-it-or-leave-it price p , she will buy with probability $1 - F(p)$, thereby generating revenue $p(1 - F(p))$. Let $q(p) := 1 - F(p)$ be the probability of selling at p , and call it the *quantile* of p . The plot of revenue $\text{Rev}(q) := q \cdot F^{-1}(1 - q)$ against $q \in [0, 1]$ is called the *revenue curve* of the distribution. With a slight overloading of notation, for a quantile q we denote by $v(q) := F^{-1}(1 - q)$ its corresponding value.

It is without loss of generality to assume that $\text{Rev}(0) = \text{Rev}(1) = 0$.² The most important fact about the revenue is that the derivative of $\text{Rev}(q)$ at quantile q is equal to the virtual value of $v(q)$. In this paper, we do not appeal to this connection other than the curve’s concavity for regular distributions.

The highest point of a revenue curve gives the maximum revenue extractable from a bidder by posting a take-it-or-leave-it price. This price is called the *monopoly reserve* for the bidder. For bidder i we denote this by r_i^* . When there is a single bidder in an auction, the revenue yielded by posting r^* is in fact the optimal revenue, by Myerson’s characterization.

The Vickrey Auction. In the Vickrey auction (a.k.a. the second price auction), the item is allocated to the bidder with the highest value, who is then charged the bid of the second highest bidder. Vickrey [18] showed that this auction is DSIC and ex-post IR.

Auctions with duplicate bidders and Bulow-Klemperer type results. In general, the revenue of the Vickrey auction can be far from the optimal revenue. Bulow and Klemperer [3] showed that, when bidders’ values are drawn i.i.d. from a regular distribution, the second price auction, by recruiting one additional bidder (called a *duplicate bidder*), could reverse the comparison:

Theorem 2.2 (Bulow and Klemperer [3]). *Let OPT be the optimal revenue for an n bidder, single-item auction where each bidder’s value is drawn independently from an identical regular distribution. Then the expected revenue of the second price auction with $n + 1$ independent bidders from the same distribution is at least OPT.*

Hartline and Roughgarden [11] extended this theorem in an approximate manner to settings where bidders’ values are independently but not identically drawn:

Theorem 2.3 (Hartline and Roughgarden [11]). *Let OPT be the optimal revenue extractable in an auction with n bidders where bidder i ’s value is drawn independently from a regular distribution F_i . The expected revenue of the second price auction with $2n$ bidders, where the values of bidder i and $n + i$ are independently drawn from F_i , is at least $\frac{1}{2}$ OPT.*

Hartline and Roughgarden also gave an example with $n = 2$ in which the Vickrey auction with $2n$ bidders extracts a revenue that is $\frac{3}{4}$ OPT. We recall this example in Section 5.

Lookahead auction. The lookahead auction, first devised and analyzed by Ronen [15], offers a technical tool to analyze auctions. At a bid profile \mathbf{v} , the lookahead auction first identifies the highest bidder, say i^* , then offers the item to i^* at the optimal posted price for i^* conditioned on \mathbf{v}_{-i^*} and the fact $v_{i^*} \geq \max_{j \neq i^*} v_j$.

Theorem 2.4 (Ronen [15]). *The lookahead auction is a DSIC, ex-post IR mechanism whose revenue 2-approximates the optimal revenue.*

²See, e.g. Appendix A of Fu et al. [9] for an argument.

Ex ante optimal revenue. Almost all approximation results we present in this paper are in fact with respect to the *ex ante optimal revenue*, a benchmark that is usually strictly stronger than the optimal revenue. Given regular distributions with revenue curves $\text{Rev}_1, \dots, \text{Rev}_n$, the ex ante optimal revenue is the value of the following program:

$$\begin{aligned} \max \quad & \sum_{i=1}^n \text{Rev}_i(q_i) \\ \text{s.t.} \quad & \sum_{i=1}^n q_i \leq 1. \end{aligned} \tag{4}$$

We denote by $\tilde{q}_1, \dots, \tilde{q}_n$ the optimal solution to this program. We refer the reader to Hartline [10] for an exposition that the ex ante optimal revenue upper bounds the optimal revenue.

***k*-Items Unit-Demand Auctions.** In a *k*-items unit-demand auction, there are *k* identical items to sell, and each bidder wants at most one item. Each bidder’s value for an item is, as before, drawn independently. All the notions defined above carry over to this setting, with minimal change as follows: the feasibility constraint on the allocation rule now becomes, for all \mathbf{v} , $\sum_i x_i(\mathbf{v}) \leq k$, and $0 \leq x_i(\mathbf{v}) \leq 1$, for all *i*. The generalization of the Vickrey auction in this setting is the VCG auction, which allocates the items to the *k* highest bidders, and then charges each winner the (*k* + 1)-st highest bid. In the program that defines the ex ante optimal revenue, the constraint now becomes $\sum_i q_i \leq k$, and $q_i \leq 1$ for all *i*. Myerson’s characterization (Lemma 2.1) still holds, and the revenue-optimal auction allocates the items to at most *k* bidders with the highest non-negative virtual values.

As we explained in the introduction, in order for VCG with duplicates to be approximately revenue optimal, one needs the further restriction that an original bidder and her duplicate should never both win an item. With this modification, Hartline and Roughgarden showed:

Theorem 2.5 (Hartline and Roughgarden [11]). *Let OPT be the optimal revenue extractable in a *k*-items auction with *n* unit-demand bidders whose values are independently drawn from regular, non-identical distributions. The expected revenue of the VCG auction with every bidder duplicated is at least half of OPT.*

In fact, Hartline and Roughgarden showed a more general result for matroid settings. As we discuss in the conclusion, we leave the question open whether our result for *k*-items auctions can be extended to general matroids.

3 The Vickrey Auction with a Single Duplicate Bidder

3.1 Warm-up: a loose bound for single-item auctions

We illustrate in this section the main ideas behind our result on single-duplicate auctions and prove a 40-approximation. In Section 3.2, we optimize the parameters and get an approximation ratio better than 10.

Let $\{\tilde{q}_i\}_{i \in [n]}$ be the optimal solution to the ex ante relaxation and let OPT denote the optimal value of the ex ante relaxation. In the case where $v_i(1/4) \geq \text{OPT}/2$ for some *i*, posting a price of $\text{OPT}/2$ to bidder *i* obtains a revenue of at least $\text{OPT}/8$. Hence, Bulow-Klemperer’s Theorem (Theorem 2.2) implies that running a second-price auction with only two copies of bidder *i* already obtains a revenue of at least $\text{OPT}/8$; introducing additional bidders only weakly increases that

revenue and therefore in this case the SPA with bidder i duplicated obtains a revenue of at least $\text{OPT}/8$.

Henceforth, we may assume $v_i(1/4) < \text{OPT}/2$ for all i . We show that in this case the SPA without duplicates in fact already obtains an $O(1)$ -approximation of the optimal revenue (and introducing bidders never hurts the revenue). Let S be $\{i : v_i(\tilde{q}_i) \geq \text{OPT}/2\}$. Intuitively, bidders not in S have low prices in the optimal ex ante solution, and altogether contribute at most a constant fraction in the ex ante optimal. Specifically,

$$\sum_{i \notin S} v_i(\tilde{q}_i) \tilde{q}_i \leq \frac{\text{OPT}}{2} \sum_{i \notin S} \tilde{q}_i \leq \frac{\text{OPT}}{2}$$

where the first inequality is because $v_i(\tilde{q}_i) \leq \text{OPT}/2$ for $i \notin S$ and the second inequality is because $\sum_{i \notin S} \tilde{q}_i \leq 1$. Hence,

$$\sum_{i \in S} \text{Rev}_i(\tilde{q}_i) = \sum_{i \in S} v_i(\tilde{q}_i) \tilde{q}_i \geq \text{OPT}/2. \quad (5)$$

For $i \in S$, define $q'_i = q_i(\text{OPT}/2)$ and notice that $\tilde{q}_i \leq q'_i < 1/4$. We show that the *sum* of q'_i 's is at least a constant, which in turn allows us to show that with constant probability, at least two bidders in S bid at least $\text{OPT}/2$.

By concavity of the revenue curve (see Claim A.1), we have $\text{Rev}_i(q'_i) \geq \frac{3}{4} \text{Rev}_i(\tilde{q}_i)$. Plugging this bound into Equation (5), we get that

$$\sum_{i \in S} v_i(q'_i) q'_i = \sum_{i \in S} \text{Rev}_i(q'_i) \geq \frac{3}{4} \sum_{i \in S} \text{Rev}_i(\tilde{q}_i) \geq \frac{3 \text{OPT}}{8}. \quad (6)$$

On the other hand, since $v_i(q'_i) = \text{OPT}/2$ (by definition of q'_i), we have

$$\sum_{i \in S} v_i(q'_i) q'_i = \frac{\text{OPT}}{2} \sum_{i \in S} q'_i. \quad (7)$$

Combining Equation (6) and Equation (7) gives

$$\sum_{i \in S} q'_i \geq \frac{3}{4}. \quad (8)$$

We now show that with probability $\Omega(1)$, at least *two* bidders bid at least $\text{OPT}/2$. Indeed, let X be the random variable that counts the number of bidders in S that bid at least $\text{OPT}/2$. Then $\Pr[X = 0] = \prod_{i \in S} (1 - q'_i) \leq \exp(-\sum_{i \in S} q'_i) \leq \exp(-3/4)$ and

$$\begin{aligned} \Pr[X = 1] &= \sum_{i \in S} q'_i \prod_{j \in S \setminus \{i\}} (1 - q'_j) \\ &\leq \sum_{i \in S} q'_i \exp\left(-\sum_{j \in S \setminus \{i\}} q'_j\right) \\ &\leq \sum_{i \in S} q'_i \exp\left(1/4 - \sum_{j \in S} q'_j\right) && \text{(because } 1/4 - q'_i \geq 0\text{)} \\ &= \exp(1/4) \cdot \left(\sum_{i \in S} q'_i\right) \cdot \exp\left(-\sum_{i \in S} q'_i\right) \\ &\leq \exp(-3/4), \end{aligned}$$

where the last inequality is because the function $f(x) = x \exp(-x)$ is maximized at $x = 1$.³ To conclude, we have $\Pr[X \geq 2] = 1 - \Pr[X = 0] - \Pr[X = 1] \geq 1 - 2 \exp(-3/4) > 1/20$. Hence, SPA extracts a revenue of at least $\text{OPT}/40$.

3.2 A tighter bound for single-item auctions

We now optimize the parameters in the argument outlined in Section 3.1 and obtain main result for single-item auctions:

Theorem 3.1. *In a single item auction with n bidders whose values are drawn independently from (non-identical) regular distributions F_1, \dots, F_n , there exists i such that the second price auction with the same n bidders and an additional bidder $n+1$ whose value is drawn independently from F_i achieves at least 0.108 fraction of the ex ante optimal revenue with the original n bidders.*

Recall from Section 2 that $\tilde{q}_1, \dots, \tilde{q}_n$ are quantiles that solve the ex ante revenue maximization problem with the original n bidders. We use $\text{OPT} = \sum_i \text{Rev}_i(\tilde{q}_i)$ to denote the ex ante optimal revenue, which upper bounds the optimal revenue (without duplicates).

Lemma 3.2 is the technical heart of the proof. Using regularity, i.e. the concavity of the revenue curves, it shows that either there is a bidder i who bids a high value with a constant probability, or the *sum* of each bidder's probability of bidding a high value must be large.

Lemma 3.2. *Let $\alpha, \beta \in [0, 1]$ be constants. Consider a single item auction with n bidders whose values are drawn independently from regular distributions F_1, \dots, F_n . Let OPT be the ex ante optimal revenue. Then exactly one of the following statements is true.*

1. *There exists $i \in [n]$ such that $v_i(\beta) \geq \alpha \cdot \text{OPT}$.*
2. *For all i , $v_i(\beta) < \alpha \cdot \text{OPT}$ but $\sum_{i \in [n]} q_i(\alpha \cdot \text{OPT}) \geq \frac{1-\alpha}{\alpha} \cdot (1 - \beta)$.*

Remark 3.3. *The proof will show that if the first condition does not hold in the lemma then the second condition holds which implies that at least one of the two statements in the lemma is true. However, the two statements are mutually exclusive so this also implies that exactly one of the statements is true.*

Proof of Lemma 3.2. Let $\tilde{q}_1, \dots, \tilde{q}_n$ be an optimal solution to the ex ante relaxation. In particular, $\sum_{i=1}^n \text{Rev}_i(\tilde{q}_i) = \text{OPT}$. Let $S = \{i : v_i(\tilde{q}_i) \geq \alpha \cdot \text{OPT}\}$. Since $\sum_{i \notin S} \tilde{q}_i v_i(\tilde{q}_i) \leq \alpha \cdot \text{OPT}$ (because $\sum_{i \notin S} \tilde{q}_i \leq 1$), we have $\sum_{i \in S} \tilde{q}_i v_i(\tilde{q}_i) \geq (1 - \alpha) \cdot \text{OPT}$.

Suppose that the first statement of the lemma does not hold, i.e. $v_i(\beta) < \alpha \cdot \text{OPT}$ for all i . Recall that $q_i(\cdot)$ is a non-increasing function. So for any $i \in S$, $\tilde{q}_i \leq q_i(\alpha \cdot \text{OPT}) \leq \beta$ by definition. By the concavity of the revenue curve (see Claim A.1), we have $\text{Rev}_i(q_i(\alpha \cdot \text{OPT})) \geq (1 - \beta) \text{Rev}_i(\tilde{q}_i)$ for all $i \in S$. Therefore

$$\sum_{i \in S} \text{Rev}_i(q_i(\alpha \cdot \text{OPT})) \geq (1 - \beta) \sum_{i \in S} \text{Rev}_i(\tilde{q}_i) \geq (1 - \beta)(1 - \alpha) \text{OPT}.$$

But $\text{Rev}_i(q_i(\alpha \cdot \text{OPT}))$ is just $(\alpha \cdot \text{OPT})q_i(\alpha \cdot \text{OPT})$. Rearranging, we have $\sum_{i \in S} q_i(\alpha \cdot \text{OPT}) \geq \frac{1-\alpha}{\alpha} \cdot (1 - \beta)$, as claimed. \square

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. Let $\alpha, \beta \in [0, 1]$ be constants such that $\frac{1-\alpha}{\alpha} \cdot (1 - \beta) \geq 1$. We will determine their values later. The proof amounts to analyzing the two cases given in Lemma 3.2.

³ This follows from standard calculus. Indeed, note that $f'(x) = (1 - x) \exp(-x)$. Hence, $f(x)$ is non-decreasing on $(-\infty, 1]$ and non-increasing on $[1, \infty)$. So $x = 1$ is the maximizer of f and $f(1) = 1/e$.

Case 1: There exists $i \in [n]$ such that $v_i(\beta) \geq \alpha \cdot \text{OPT}$. In this case, posting a price of $v_i(\beta)$ *only* to bidder i obtains a revenue of $\beta v_i(\beta) \geq (\alpha\beta) \cdot \text{OPT}$. Hence, we can apply the Bulow-Klemperer Theorem (Theorem 2.2) to assert that duplicating bidder i suffices to get an $(\alpha\beta)$ -fraction of the ex ante optimal revenue. Finally, adding the remaining bidders back in can only increase the revenue so this obtains at least an $(\alpha\beta)$ -fraction of the ex ante optimal revenue.

Case 2: For all i , $v_i(\beta) < \alpha \cdot \text{OPT}$ but $\sum_{i \in [n]} q_i(\alpha \cdot \text{OPT}) \geq \frac{1-\alpha}{\alpha} \cdot (1-\beta)$. To ease notation, let $q'_i := q_i(\alpha \cdot \text{OPT})$. Also, observe that the condition $v_i(\beta) < \alpha \cdot \text{OPT}$ implies that $q'_i \leq \beta$.

In this case, let us duplicate any bidder in $\arg \max_{i \in [n]} q'_i$, say i^* , and write $S' = [n] \cup \{i^*\}$. Now, we will compute the probability that there are at least two bidders who bid at least $\alpha \cdot \text{OPT}$. The probability that no bidder bids above $\alpha \cdot \text{OPT}$ is at most

$$\prod_{i \in [n]} (1 - q'_i) \leq \exp \left(- \sum_{i \in [n]} q'_i \right) \leq \exp \left(- \frac{1-\alpha}{\alpha} \cdot (1-\beta) \right). \quad (9)$$

Since we duplicated a bidder in $\arg \max_{i \in [n]} q'_i$, the probability that exactly one bidder bids at least $\alpha \cdot \text{OPT}$ is

$$\begin{aligned} \sum_{i \in S'} q'_i \prod_{j \in S' \setminus \{i\}} (1 - q'_j) &\leq \sum_{i \in S'} q'_i \exp \left(- \sum_{j \in S' \setminus \{i\}} q'_j \right) \\ &\leq \sum_{i \in S'} q'_i \exp \left(- \sum_{j \in [n]} q'_j \right) \quad (\text{since } q'_{i^*} \in \arg \max_{j \in S} q'_j) \\ &= \left(\sum_{i \in S'} q'_i \right) \cdot \exp \left(- \sum_{i \in [n]} q'_i \right) \\ &\leq \left(\beta + \frac{1-\alpha}{\alpha} \cdot (1-\beta) \right) \cdot \exp \left(- \frac{1-\alpha}{\alpha} \cdot (1-\beta) \right) \quad (\text{since } q'_{i^*} \leq \beta), \end{aligned} \quad (10)$$

where the last inequality uses the assumption that $\frac{1-\alpha}{\alpha} \cdot (1-\beta) \geq 1$ and the function $x \exp(-x)$ is decreasing for $x \geq 1$. Define

$$\eta = 1 - \left(1 + \beta + \frac{1-\alpha}{\alpha} \cdot (1-\beta) \right) \cdot \exp \left(- \frac{1-\alpha}{\alpha} \cdot (1-\beta) \right),$$

which is the probability that at least two bidders bid at least $\alpha \cdot \text{OPT}$. We have shown that one can always duplicate one bidder to obtain in the Vickrey auction an $\alpha \min\{\beta, \eta\}$ -fraction of the optimal revenue. Choosing $\alpha = 0.27$ and $\beta = 0.4$ and verifying $\frac{1-\alpha}{\alpha} \cdot (1-\beta)$ is indeed greater than 1, we complete the proof of Theorem 3.1. \square

3.3 Choosing the distribution to duplicate

Bulow and Klemperer's result and the extension by Hartline and Roughgarden are also seen as prototypical results in *prior-independent* mechanism design. These results guarantee performances of "detail-free" auctions, such as the Vickrey auction, as long as the underlying distribution satisfies some mild property. As we pointed out in the Introduction, one may not duplicate a single bidder with complete ignorance on the value distributions, since arbitrary or uniformly random duplication easily fails. We show in this section that the information requirement for choosing the duplication is resilient and minimal. We give a few variants of the forms of information sufficient for our purpose.

Theorem 3.4. Consider a single item auction with n bidders whose values are drawn independently from (non-identical) regular distributions F_1, \dots, F_n . Let OPT be the ex ante optimal revenue for this setting.

1. Suppose for some $\beta \in [0, 1]$, the auctioneer has access to $\text{Rev}_i(\beta)$ (or equivalently, $v_i(\beta)$) for all i . Then, by duplicating any bidder i that maximizes $\text{Rev}_i(\beta)$ and running the Vickrey auction on the $n + 1$ bidders, the auctioneer can extract a revenue of at least $c_1(\beta) \cdot \text{OPT}$, where $c_1(\beta)$ is a constant depending only on β . In particular, $c_1(0.355) \geq 0.099$.
2. More generally, suppose that the auctioneer has access to the following oracle for each bidder i for some $\beta \in [0, 1/2], \epsilon \in [0, 1]$: the oracle returns $\text{Rev}_i(\beta'_i)$ with the promise that $|\beta'_i - \beta| \leq \epsilon\beta$. Then, by duplicating any bidder i that maximizes $\text{Rev}_i(\beta'_i)$ and running the Vickrey auction on the $n + 1$ bidders, the auctioneer can extract a revenue of at least $c_2(\beta, \epsilon) \cdot \text{OPT}$, where $c_2(\beta, \epsilon)$ is a constant depending only on β and ϵ . Moreover, $c_2(\beta, \epsilon) \geq (1 - \epsilon)c_1(\beta)$ where $c_1(\beta)$ is the constant as from part 1.
3. Suppose that the auctioneer can draw a sample s_i , independently, from each distribution F_i . Then duplicating any bidder in $\arg \max_i s_i$ and running VCG on the resulting $n + 1$ bidders yields 0.044 fraction of the ex ante optimal revenue.

As we emphasized in the Introduction, revenue-wise, running a Vickrey auction with a duplicate may not be the best way to use the forms of prior information in Theorem 3.4, but the Vickrey auction has the advantage of being simple and anonymous, and may be particularly useful when individual reserve prices cannot be placed. We see Theorem 3.1 as the main result of this work, and Theorem 3.4 as a natural consequence of our techniques.

Proof of Theorem 3.4 part 2. We prove part 2 directly, and part 1 is implied by the proof.

Let $\alpha \in [0, 1]$ be any constant that satisfies $(1 - \alpha)(1 - \beta)/\alpha \geq 1$. As in the proof of Theorem 3.1, we analyze the two cases given by the conclusion of Lemma 3.2.

Case 1: There exists $i \in [n]$ such that $v_i(\beta) \geq \alpha \cdot \text{OPT}$. Let $i^* \in \arg \max_{i \in [n]} \text{Rev}_i(\beta)$. Since we duplicate a bidder in $\arg \max_{i \in [n]} \text{Rev}_i(\beta'_i)$, say j , we can apply the Bulow-Klemperer Theorem (Theorem 2.2) to assert that second price auction with only bidder j already obtains a revenue of at least $\text{Rev}_j(\beta_j)$. However, note that $\text{Rev}_j(\beta'_j) \geq \text{Rev}_{i^*}(\beta'_{i^*}) \geq (1 - \epsilon) \text{Rev}_{i^*}(\beta) \geq (1 - \epsilon)\alpha\beta \text{OPT}$, where the second inequality uses Claim A.2.

Case 2: For all i , $v_i(\beta) < \alpha \cdot \text{OPT}$ but $\sum_{i \in [n]} q_i(\alpha \cdot \text{OPT}) \geq \frac{1 - \alpha}{\alpha} \cdot (1 - \beta)$. To ease notation, let $q'_i := q_i(\alpha \cdot \text{OPT})$. In this case, we will show that the probability that at least two bidders bid at least $\alpha \cdot \text{OPT}$ is quite high even if we do not duplicate any bidder. Indeed, the probability that no bidder bids at least $\alpha \cdot \text{OPT}$ is

$$\prod_{i \in [n]} (1 - q'_i) \leq \exp \left(- \sum_{i \in [n]} q'_i \right) \leq \exp \left(- \frac{1 - \alpha}{\alpha} \cdot (1 - \beta) \right)$$

and the probability that exactly one bidder bids at least $\alpha \cdot \text{OPT}$ is

$$\begin{aligned}
\sum_{i \in [n]} q'_i \prod_{j \in [n] \setminus \{i\}} (1 - q'_j) &\leq \sum_{i \in [n]} q'_i \exp \left(- \sum_{j \in [n] \setminus \{i\}} q'_j \right) \\
&\leq \sum_{i \in [n]} q'_i \exp \left(\beta - \sum_{j \in [n]} q'_j \right) \quad (\text{since } q'_i \leq \beta \text{ for all } i) \\
&= \left(\sum_{i \in [n]} q'_i \right) \cdot \exp \left(\beta - \sum_{i \in [n]} q'_i \right) \\
&\leq \left(\frac{1 - \alpha}{\alpha} \cdot (1 - \beta) \right) \exp(\beta) \cdot \exp \left(- \frac{1 - \alpha}{\alpha} \cdot (1 - \beta) \right).
\end{aligned} \tag{11}$$

Here, we assume that $\left(\frac{1 - \alpha}{\alpha} \cdot (1 - \beta)\right) \geq 1$ and that $x \cdot \exp(x)$ is non-increasing on $[1, \infty)$. Hence, the probability that at least two bidders bid above $\alpha \cdot \text{OPT}$ is at least

$$\eta(\alpha, \beta) = 1 - \left(1 + \frac{1 - \alpha}{\alpha} \cdot (1 - \beta) \right) \exp(\beta) \cdot \exp \left(- \frac{1 - \alpha}{\alpha} \cdot (1 - \beta) \right).$$

Hence, the revenue *without* any extra bidders is at least $\alpha \eta(\alpha, \beta) \text{OPT}$, so if we add any duplicate, we are still guaranteed at least this quantity.

Combining the two cases implies that we obtain a revenue of at least $\alpha \cdot \min\{(1 - \epsilon)\beta, \eta(\alpha, \beta)\} \text{OPT}$. In particular, one can take

$$c(\beta, \epsilon) = \max_{\alpha \in [0, 1]: (1 - \alpha)(1 - \beta)/\alpha \geq 1} \alpha \cdot \min\{(1 - \epsilon)\beta, \eta(\alpha, \beta)\},$$

completing the proof. \square

Proof of Theorem 3.4, part 3. Let $\alpha, \beta, \gamma \in (0, 1)$ be constants. For technical reasons, we also need that $(1 - \alpha)(1 - \beta)/\alpha \geq 1$. As in the proof of Theorem 3.1, we analyze the two cases given by the conclusion of Lemma 3.2.

Case 1: There exists $i \in [n]$ such that $v_i(\beta) \geq \alpha \cdot \text{OPT}$. Let $H = \{i \in [n] : v_i(\beta) \geq \alpha \cdot \text{OPT}\}$ and $L = [n] \setminus H$. We break things down into two smaller cases.

Case 1a: $\Pr[v_i \geq \alpha \cdot \text{OPT} \text{ for some } i \in L] < \gamma$. In this case, with probability at least $\beta(1 - \gamma)$, if we draw a single sample from each bidder then the value of the highest sample will be in H . Conditioned on this event, we duplicate a bidder in H and obtain a revenue of at least $(\alpha\beta) \cdot \text{OPT}$. So in this case, we get at least $(\alpha\beta^2(1 - \gamma)) \cdot \text{OPT}$.

Case 1b: $\Pr[v_i \geq \alpha \cdot \text{OPT} \text{ for some } i \in L] \geq \gamma$. In this case, even if we do not duplicate any bidders, the probability that at least two bidders bid at least $\alpha \cdot \text{OPT}$ is at least $\beta\gamma$. Hence, the second price auction already extracts at least $(\alpha\beta\gamma) \cdot \text{OPT}$.

Case 2: For all i , $v_i(\beta) < \alpha \cdot \text{OPT}$ but $\sum_{i \in [n]} q_i(\alpha \cdot \text{OPT}) \geq \frac{1 - \alpha}{\alpha} \cdot (1 - \beta)$. In this case, the argument is identical to that of part 2 of Theorem 3.4. Let $\eta(\alpha, \beta)$ be as in the proof of part 2.

Putting all the cases together, we see that this mechanism achieves a revenue of at least $\alpha \cdot \min\{\beta^2(1 - \gamma), \beta\gamma, \eta(\alpha, \beta)\} \cdot \text{OPT}$. Choosing $\alpha = 0.26, \beta = 0.51, \gamma = 0.34$ yields a revenue of at least 0.0446OPT . \square

4 The VCG Auction with k Duplicates in k -Items Auctions

In this section we describe our results for k -items auctions with unit demand bidders. The generalization of the Vickrey auction in this setting is the VCG auction, as we explained in Section 2. Again we will show that, in order for the VCG auction to secure a constant fraction of the optimal auction’s revenue, one needs to duplicate far fewer bidders than required by Theorem 2.5. However, one must think carefully about conditions that should be imposed on the duplication environment. As we showed in the Introduction, the example with one bidder having value 1 and all other bidders having value 0 shows that even if one duplicates all bidders, the VCG auction still has revenue 0 whereas the optimal revenue is 1.

The VCG auction therefore needs more power in this setting than simply duplicating bidders at most once. We study two settings that grant the VCG auction different powers. In the first setting, which we deem more natural and term as the “free” setting, the VCG auction may duplicate the same bidder more than once. In the example above, it is easily seen that the VCG auction needs to duplicate the same bidder at least k times to be a constant approximation. Our main theorem in this section shows that duplicating k bidders is sufficient.

Theorem 4.1. *In a k -item auction with n unit demand bidders whose values are drawn independently from (non-identical) regular distributions F_1, \dots, F_n , there exists i such that the VCG auction with the same n bidders and additional bidders $n+1, \dots, n+k$ whose values are drawn independently from F_i achieves at least 0.009 fraction of the ex ante optimal revenue with the original n bidders.*

We relegate the proof of Theorem 4.1 to Subsection 4.1. Although the proof follows a similar structure to the proof of Theorem 3.1, there are some subtleties that one needs to take into account. First, it is not difficult to see that if there are a small number of bidders that are high bidders *or* the probability that many bidders bid above a certain threshold is large then one can extract a large fraction of the revenue. The tricky part is to deal with the middle ground: when there is no bidder with high revenue *and* it is not often the case that many bidders bid above a certain threshold. In this case, we show that the expected number of high bids is large from which the desired result will follow from an application of a theorem due to Hoeffding [12].

The second setting was introduced by Hartline and Roughgarden [11], which we term as “constrained”. Here the VCG auction can only duplicate each bidder at most once, but each original bidder and her duplicate can win at most one item. With this constraint, in the example above, with the first bidder duplicated, the VCG auction recovers the optimal revenue. The following example shows that $\Omega(k)$ duplicates are necessary, even with the additional constraint, for the VCG auction to recover a constant fraction of the optimal revenue: $k/3$ bidders have value 1, and the rest of the bidders have value 0. The optimal revenue here is $k/3$, and the VCG auction’s revenue is equal to the number of duplicates it recruits. We show that this is the worst case: k duplicates always suffices for the VCG auction to extract a constant fraction of the optimal revenue.

Theorem 4.2. *Let OPT be the ex ante optimal revenue in a k -items auction with n unit-demand bidders with values drawn independently from regular (non-identical) distributions. There exist k bidders such that the VCG auction with a duplicate of each of these k bidders, where at most one of an original bidder and her duplicate can win an item, extracts revenue that is at least 0.1 OPT.*

We relegate the proof of Theorem 4.2 to Subsection 4.2. Let us also remark that in this setting, we can also obtain an analagous result to Theorem 3.4. We relegate the details to Appendix B.

The proofs of Theorem 4.1 and Theorem 4.2 both make use of the following analog of Lemma 3.2 for k -item auctions.

Lemma 4.3. Let $\beta, \gamma, \delta \in (0, 1)$ be constants that satisfy $\frac{(1-\gamma)(1-\beta)-\delta}{\gamma} \geq \frac{3}{2}$. In a k -items auction with n unit demand bidders whose values are drawn independently from regular distributions F_1, \dots, F_n at least one of the following three statements must be true:

1. There exists $q_1, \dots, q_n \in [\beta, 1)$ and a set $H \subseteq [n]$ with $|H| \leq k$ such that $v_i(\beta) \geq \frac{\gamma}{k} \text{OPT}$ for all $i \in H$ and $\sum_{i \in H} \text{Rev}_i(q_i) \geq \delta \cdot \text{OPT}$.
2. There exists a set $H \subseteq [n]$ with $|H| = k$ such that $v_i(\beta) \geq \frac{\gamma}{k} \text{OPT}$.
3. With probability $1/2$, at least $k+1$ bidders bid at least $\frac{\gamma}{k} \text{OPT}$.

Proof. Assume that the first two statements do not hold. Let $\{\tilde{q}_i\}_{i=1}^n$ be an optimal solution to the ex ante relaxation and $S = \{i \in [n] : v_i(\tilde{q}_i) \geq \frac{\gamma}{k} \text{OPT}\}$. Then $\sum_{i \notin S} \text{Rev}_i(\tilde{q}_i) \leq \gamma \text{OPT}$ because $\sum_{i \notin S} \tilde{q}_i \leq k$. Hence, $\sum_{i \in S} \text{Rev}_i(\tilde{q}_i) \geq (1-\gamma) \cdot \text{OPT}$.

For $i \in S$, define q'_i as follows. If $\tilde{q}_i > \beta$ then set $q'_i = \tilde{q}_i$ but note that $v_i(\beta) \geq v_i(q'_i) \geq \frac{\gamma}{k} \text{OPT}$ since $v_i(\cdot)$ is decreasing in $[0, 1]$. Otherwise, set $q'_i = \min\{\beta, 1 - F_i(\frac{\gamma}{k} \text{OPT})\}$. Note that if $q'_i < \beta$ then $v_i(q'_i) = \frac{\gamma}{k} \text{OPT}$.

If $q'_i \leq \beta$ then Claim A.1 implies that $\text{Rev}_i(q'_i) \geq (1-\beta) \text{Rev}_i(\tilde{q}_i)$ and if $q'_i > \beta$ then $q'_i = \tilde{q}_i$ so $\text{Rev}_i(q'_i) = \text{Rev}_i(\tilde{q}_i)$. Hence $\sum_{i \in S} \text{Rev}_i(q'_i) \geq (1-\gamma)(1-\beta) \cdot \text{OPT}$. Let $S' = \{i \in S : v_i(\beta) \geq \frac{\gamma}{k} \text{OPT}\}$. If $i \notin S'$ we have $v_i(\beta) < \frac{\gamma}{k} \text{OPT}$, which means that $q'_i \leq \beta$ (if $q'_i > \beta$ then $q'_i = \tilde{q}_i$). If $q'_i = \beta$, then $v_i(q'_i) < \frac{\gamma}{k} \text{OPT}$ because $i \notin S'$. Otherwise, $v_i(q'_i) = v_i(1 - F_i(\frac{\gamma}{k} \text{OPT})) = \frac{\gamma}{k} \text{OPT}$. Hence, for $i \notin S'$, we have $v_i(q'_i) \leq \frac{\gamma}{k} \text{OPT}$. Since the second item in the claim does not hold, we have $|S'| < k$. Since the first item in the claim does not hold, we have

$$\sum_{i \in S \setminus S'} \text{Rev}_i(q'_i) = \sum_{i \in S} \text{Rev}_i(q'_i) - \sum_{i \in S'} \text{Rev}_i(q'_i) \geq ((1-\gamma)(1-\beta) - \delta) \text{OPT}.$$

As we observed above, $v_i(q'_i) \leq \frac{\gamma}{k} \text{OPT}$ for all $i \in S \setminus S'$. Plugging this into the above inequality with the fact that $\text{Rev}_i(q'_i) = q'_i v_i(q'_i)$ gives

$$\sum_{i \in S \setminus S'} q'_i \geq k \cdot \frac{(1-\gamma)(1-\beta) - \delta}{\gamma} \geq \frac{3k}{2},$$

where the last inequality is an assumption made in the claim.

Let X be the number of bidders in $S \setminus S'$ that bid at least $\frac{\gamma}{k} \text{OPT}$. Then $\mathbf{E}[X] \geq \sum_{i \in S \setminus S'} q'_i \geq \frac{3k}{2} \geq k+1$ since $k \geq 2$. By Theorem A.3 and Fact A.4, we have that the probability that at least $k+1$ bidders in $S \setminus S'$ bid at least $\frac{\gamma}{k} \text{OPT}$ is at least $1/2$. \square

4.1 Proof of Theorem 4.1

Proof of Theorem 4.1. Let $\beta, \gamma, \delta \in (0, 1)$ be constants to be chosen later. We will consider the three cases in the conclusion of Lemma 4.3.

Case 1: There exists $q_1, \dots, q_n \in [\beta, 1)$ and a set $H \subseteq [n]$ with $|H| \leq k$ such that $\sum_{i \in H} \text{Rev}_i(q_i) \geq \delta \cdot \text{OPT}$. By shuffling the indices, we may assume without loss of generality that $\text{Rev}_1(q_1) \geq \dots \geq \text{Rev}_n(q_n)$. Let $r \leq k$ be such that $\text{Rev}_i(q_i) \geq \frac{\delta \text{OPT}}{2k}$ for all $i \leq r$ and $\text{Rev}_i(q_i) < \frac{\delta \text{OPT}}{2k}$ for $i \in \{r+1, \dots, k\}$. Then $\sum_{i=1}^r \text{Rev}_i(q_i) \geq \frac{\delta \text{OPT}}{2}$ because $\sum_{i=r+1}^k \text{Rev}_i(q_i) < k \frac{\delta \text{OPT}}{2k} = \frac{\delta \text{OPT}}{2}$.

Define $R_i = \frac{\text{Rev}_i(q_i)}{\text{OPT}}$. We will consider two subcases based on the value of R_1 .

Case 1a: $R_1 \geq \frac{\delta}{4}$. First, note that $v_1(1 - 1/k) \geq \frac{\delta}{4k}$ OPT. To see this, if $q_1 > 1 - 1/k$ then $v_1(1 - 1/k) \geq v_1(q_1) \geq \text{Rev}_1(q_1) \geq \frac{\delta}{4}$ OPT. On the other hand, if $q_1 \leq 1 - 1/k$ then Claim A.1 implies that $\text{Rev}_1(1 - 1/k) \geq \text{Rev}_1(q_1)/k \geq \frac{\delta}{4k}$ OPT and hence $v_1(1 - 1/k) \geq \frac{\delta}{4k}$ OPT.

Now suppose we obtain k duplicates of bidder 1. The probability that any one of them bids above $\frac{\delta}{4k}$ OPT is at least $1 - 1/k$ so the probability that all $k + 1$ of them do is at least $(1 - 1/k)^{k+1} \geq 1/8$ since $k \geq 2$. Hence, VCG extracts a revenue of at least $\frac{\delta \text{OPT}}{32}$.

Case 1b: $R_1 \leq \frac{\delta}{4}$. We begin with the following claim which gives a lower bound on the probability that a bidder bids above a threshold of $\frac{\delta \text{OPT}}{8k}$.

Claim 4.4. For all $i \in [r]$, we have $v_i\left(1 - \frac{\delta}{8kR_i}\right) \geq \frac{\delta \text{OPT}}{8k}$.

Proof. Suppose first that $q_i > 1 - \frac{\delta}{8kR_i}$. Then since $v_i(\cdot)$ is non-increasing on $[0, 1]$, we have

$$v_i\left(1 - \frac{\delta}{8kR_i}\right) \geq v_i(q_i) \geq \text{Rev}_i(q_i) \geq \frac{\delta \text{OPT}}{2k}$$

where the last inequality is by definition of r .

On the other hand, suppose that $q_i \leq 1 - \frac{\delta}{8kR_i}$. Then we can apply Claim A.1 to get that

$$\text{Rev}_i\left(1 - \frac{\delta}{8kR_i}\right) \geq \frac{\delta}{8kR_i} \text{Rev}_i(q_i) = \frac{\delta \text{OPT}}{8k}.$$

Consequently, $v_i\left(1 - \frac{\delta}{8kR_i}\right) \geq \frac{\delta \text{OPT}}{8k}$. □

Consider obtaining k duplicates of bidder 1. Let X be the number of bidders amongst the original r bidders and the k duplicates that bid at least $\frac{\delta \text{OPT}}{8k}$. Then

$$\mathbf{E}[X] \geq k \left(1 - \frac{\delta}{8kR_1}\right) + \sum_{i=1}^r \left(1 - \frac{\delta}{8kR_i}\right) \geq k - \frac{\delta}{8R_1} + \frac{3r}{4},$$

where the last inequality is because $R_i \geq \frac{\delta}{2k}$ for $1 \leq i \leq r$ so $1 - \frac{\delta}{8kR_i} \geq \frac{3}{4}$. It now remains to lower bound r .

Claim 4.5. $r \geq \frac{\delta}{2R_1}$.

Proof. Indeed, we have

$$r = \sum_{i=1}^r R_i \frac{1}{R_i} \geq \frac{1}{R_1} \sum_{i=1}^r R_i \geq \frac{\delta}{2R_1}.$$

The first inequality is because $R_i \leq R_1$ for all i . The second inequality is because $\sum_{i=1}^r \text{Rev}_i(q_i) \geq \frac{\delta \text{OPT}}{2}$ is equivalent to $\sum_{i=1}^r R_i \geq \frac{\delta}{2}$. □

Hence, we have

$$\mathbf{E}[X] \geq k - \frac{\delta}{8R_1} + \frac{3r}{4} \geq k - \frac{\delta}{8R_1} + \frac{3\delta}{8R_1} = k + \frac{\delta}{4R_1}.$$

Since $R_1 \leq \delta/4$, we have that $\mathbf{E}[X] \geq k + 1$.

Now observe that X is a sum of independent Bernoulli random variables, albeit with non-identical means. Combining Theorem A.3 and Fact A.4, we conclude that at least $k + 1$ bidders bid at least $\frac{\delta \text{OPT}}{8k}$ with probability at least $1/2$. Hence, in this case VCG obtains a revenue of at least $\frac{\delta \text{OPT}}{16}$.

Case 2: There exists a set $H \subseteq [n]$ with $|H| = k$ such that $v_i(\beta) \geq \frac{\gamma}{k} \text{OPT}$. Note that for $i \in H$, we have $\text{Rev}_i(\beta) \geq \frac{\beta\gamma}{k} \text{OPT}$. We claim that this implies $v_i(3/4) \geq \frac{\beta\gamma \text{OPT}}{3k}$. Indeed, if $\beta \geq 3/4$ then $v_i(3/4) \geq v_i(\beta) = \text{Rev}_i(\beta)/\beta \geq \text{Rev}_i(\beta) \geq \frac{\beta\gamma}{k} \text{OPT}$. On the other hand, if $\beta \leq 3/4$ then $\text{Rev}_i(3/4) \geq \frac{\beta\gamma}{4k} \text{OPT}$ by Claim A.1. Hence, $v_i(3/4) = \frac{\text{Rev}_i(3/4)}{3/4} \geq \frac{\beta\gamma \text{OPT}}{3k}$.

Now consider obtaining k duplicates of any bidder in H . Then the average number of bidders from H and the duplicates that bid at least $\frac{\beta\gamma \text{OPT}}{3k}$ is at least $\frac{3}{4} \cdot 2k = \frac{3k}{2} \geq k + 1$. By Theorem A.3 and Fact A.4 again, this implies that VCG obtains a revenue of at least $\frac{\beta\gamma}{6} \text{OPT}$.

Case 3: With probability at least $1/2$ there exists at least $k + 1$ bidders who bid at least $\frac{\gamma}{k} \text{OPT}$. The assumption implies that VCG without duplicates already obtains a revenue of at least $\frac{\gamma}{2} \text{OPT}$; duplicating bidders will never hurt the revenue.

Let $\mathcal{A} = \left\{ (\beta, \gamma, \delta) : \frac{(1-\gamma)(1-\beta)-\delta}{\gamma} \geq \frac{3}{2} \right\}$. Putting all the cases together, we get the revenue of VCG with k duplicates is at least

$$\text{OPT} \cdot \max_{(\beta, \gamma, \delta) \in \mathcal{A}} \min \left\{ \frac{\delta}{32}, \frac{\beta\gamma}{6}, \frac{\gamma}{2} \right\}.$$

By setting $\beta = 0.377, \delta = 0.3, \gamma = 0.15$, we see that this is at least $0.009 \cdot \text{OPT}$. \square

Remark 4.6. *One can also modify the above proof so that the choice of the bidder is independent of k , albeit with a decay in the approximation factor.*

Indeed, notice that in case 1 of the above proof, one can duplicate any bidder in

$$\arg \max_{i \in [n]} \max_{q_i \in [\beta, 1]} \text{Rev}_i(q_i)$$

and that this choice is independent of k . By reordering, suppose that bidder 1 is in the argmax of the previous expression.

Note that it may not be the case that $v_1(\beta) \geq \frac{\gamma}{k} \text{OPT}$ so we may not be able to duplicate bidder 1 in case 2 of the above proof. However, if $q_1^ \in \arg \max_{q_1 \in [\beta, 1]} \text{Rev}_1(q_1)$ then $\text{Rev}_1(q_1^*) \geq \max_{i \in [n]} \text{Rev}_i(\beta) \geq \frac{\beta\gamma}{k} \text{OPT}$. Hence,*

$$v_1(\beta) \geq v_1(q_1^*) \geq q_1^* v_1(q_1^*) = \text{Rev}_1(q_1^*) \geq \frac{\beta\gamma}{k} \text{OPT}.$$

So one can repeat the argument in case 2 of the above proof with a set $H \subseteq [n]$ and $|H| = k$ such that $1 \in H$ and $v_i(\beta) \geq \frac{\beta\gamma}{k} \text{OPT}$ for all $i \in H$. Since we can duplicate any bidder in H , we can duplicate bidder 1 as well which would give $\frac{\beta^2\gamma}{6}$ -approximation.

Case 3 of the above proof remains unchanged.

Hence, the approximation ratio is at least $\min \left\{ \frac{\delta}{32}, \frac{\beta^2\gamma}{6}, \frac{\gamma}{2} \right\}$ (provided that β, γ, δ satisfy the conditions of Theorem 4.1). Choosing $\beta = 0.4, \delta = 0.2, \gamma = 0.19$ gives an approximation ratio of > 0.005 .

4.2 Proof of Theorem 4.2

Proof of Theorem 4.2. We will consider the three cases in the conclusion of Lemma 4.3.

Case 1: There exists $q_1, \dots, q_n \in (0, 1)$ and a set $H \subseteq [n]$ with $|H| \leq k$ such that $\sum_{i \in H} \text{Rev}_i(q_i) \geq \delta \cdot \text{OPT}$. Suppose we duplicate every bidder in H and run VCG with just these bidders and their duplicates, while ensuring that a bidder and her duplicate do not both win. In this setting, the VCG auction simply runs a single-item second-price auction for each pair. Observe that if the auctioneer had posted a price of $\text{Rev}_i(q_i)/q_i = v_i(q)$ for bidder i then the revenue extracted from bidder i would have been at least $\text{Rev}_i(q_i)$. Applying Theorem 2.2, it follows that the single-item auction with bidder i and her duplicate yields a revenue of at least $\text{Rev}_i(q_i)$. Hence, the revenue of VCG is at least $\sum_{i \in H} \text{Rev}_i(q_i) \geq \delta \cdot \text{OPT}$.

Case 2: There exists a set $H \subseteq [n]$ with $|H| = k$ such that $v_i(\beta) \geq \frac{\gamma}{k} \text{OPT}$. This case is very similar to the previous case. Again, we duplicate every bidder in H and note that VCG with only the bidders in H and their duplicates is exactly a single-item second-price auction for each pair. Since $\text{Rev}_i(\beta) \geq \frac{\beta\gamma}{k} \text{OPT}$ for $i \in H$, we can apply Theorem 2.2 to assert that we extract at least $\frac{\beta\gamma}{k} \text{OPT}$ revenue from bidder i or her duplicate. Hence, the total revenue is at least $\beta\gamma \text{OPT}$.

Case 3: With probability at least $1/2$ there exists at least $k+1$ bidders who bid at least $\frac{\gamma}{k} \text{OPT}$. In this case, the VCG mechanism without duplicates already gets a $\gamma/2$ fraction of the ex ante optimal revenue with the original n bidders. Now observe that VCG with the original n bidders and their duplicates (with the constraint that a bidder and her duplicate cannot simultaneously win) earns at least as much revenue of VCG for the original n bidders. Hence, VCG in the duplicate setting also achieves at least $\gamma/2$ fraction of the ex ante optimal revenue.

Choosing $\beta = 0.5, \delta = 0.1, \gamma = 0.2$ and checking that $\frac{(1-\gamma)(1-\beta)-\delta}{\gamma} = \frac{3}{2}$ gives the desired bound of $0.1 \cdot \text{OPT}$. \square

5 Tighter Analysis for the Second Price Auction with n Single Duplicated Bidders

In this section, we strengthen Theorem 2.3 [11] in another direction.

Theorem 5.1. *Consider an auction with 2 bidders with v_1, v_2 drawn independently from regular distributions F_1 and F_2 . Let OPT be the ex ante optimal revenue. The expected revenue of the Vickrey auction with 4 bidders, with v_1, v_3 drawn from F_1 and v_2, v_4 from F_2 , all independently, is at least $\frac{3}{4} \text{OPT}$.*

The proof of Theorem 5.1 appears in Subsection 5.2. This ratio matches a lower bound given by Hartline and Roughgarden [11] (see Example 5.5), and therefore is tight for $n = 2$. We first show that the Vickrey auction with n duplicates generates at least as much revenue as the *lookahead auction* without the duplicates. We will use the following auction as a middle step, whose revenue obviously lower bounds the Vickrey auction with duplicates.

Definition 5.2 (The Second Price Auction with Late Duplicate (SPALD)). *Solicit bids from bidders 1 to n , and let i^* be the highest bidder among them. Then run the second price auction for bidders 1, 2, \dots , n and $n + i^*$.*

Lemma 5.3. *Let LA denote the revenue of the lookahead auction with n bidders, where each bidder i 's value is independently drawn from a regular distribution F_i . The Vickrey auction with n duplicates extracts a revenue at least as much as LA.*

The proof of Lemma 5.3 is in Subsection 5.1. It draws ingredients from both Ronen [15]’s analysis of the lookahead auction and Dhangwatnotai et al. [6]’s use of the revenue curve.⁴ Conditioning on \mathbf{v}_{-i} , the profile of values except bidder i ’s, it expresses SPALD’s revenue from bidder i and her duplicate in terms of an area under a curve closely related to bidder i ’s revenue curve. The comparison with the lookahead auction’s revenue then follows from the concavity of the curve.

Remark 5.4. *Theorem 2.4 and Lemma 5.3 can be used to give a short and transparent proof of Theorem 2.3. Indeed, Theorem 2.4 asserts that the revenue of the lookahead auction is a 2-approximation to the optimal revenue and Lemma 5.3 implies that the Vickrey auction with n duplicates extracts at least as much revenue as the lookahead auction. Hence, the Vickrey auction with n duplicates is a 2-approximation to the optimal revenue with the original n bidders.*

To prove Theorem 5.1, we again switch our benchmark to the stronger ex ante optimal revenue. It was observed in Alaei et al. [1] that, given an optimal solution $\tilde{q}_1, \dots, \tilde{q}_n$ to the ex ante revenue maximization problem, the solution remains the same if one were to replace the revenue curves by triangles that each peaks at \tilde{q}_i with revenue $\text{Rev}_i(\tilde{q}_i)$. On the other hand, for the Vickrey auctions, the triangle revenue curves are also the worst-case distributions (Lemma C.2). This allows us to focus on triangle revenue curves and then analytically minimize the ratio between the SPALD’s revenue and the ex ante optimal revenue.

We recall the example from Hartline and Roughgarden [11] showing that this ratio is tight for $n = 2$:

Example 5.5. *Let F_1 be the point mass on $v_1 = 1$, and let F_2 be the “slanted” equal revenue distribution $F_2(v) = 1 - \frac{1}{v+1}$. If we offer a price of $H \gg 1$ to bidder 2 and, if not taken, offer a price of 1 to bidder 1 then we obtain a revenue of $H/(H+1) + 1 - (1/(H+1)) \approx 2$. On the other hand, in the second price auction with duplicates of each distribution, if we denote by w the second highest value, then the revenue is*

$$\mathbf{E}[w] = \int_0^\infty \Pr[w \geq v] \, dv = 1 + \int_1^\infty \left(\frac{1}{v+1}\right)^2 \, dv = \frac{3}{2}.$$

Remark 5.6. *Note that in Example 5.5, the two revenue curves are two triangle revenue curves with $R_1 = R_2 = 1$, $\tilde{q}_1 = 1$ and $\tilde{q}_2 = 0$.⁵*

Remark 5.7. *Example 5.5 also provides a lower bound on the approximation of the revenue of the Vickrey auction with a single duplicate which is greater than with both duplicates. Indeed, if one duplicates only the first bidder, then the resulting Vickrey auction has revenue 1. On the other hand, duplicating the second bidder extracts revenue strictly less than 1/2. To see this, note that if both of the duplicates bid strictly less than 1 then the revenue extracted in this case is strictly less than 1 (whereas if we duplicate both bidders then the revenue would still be 1 in this case). With more detailed calculations, we show in Appendix C that the revenue is $\ln 4$ and hence, the approximation ratio is ≈ 1.44 . This is the largest gap we know for the Vickrey auction with a single duplicate.*

We believe that the Vickrey auction with n duplicates in general give a 4/3-approximation to the optimal revenue, although the following example shows that for $n > 2$ it does *not* 4/3-approximate the ex ante optimal revenue.

⁴In fact, the case $n = 1$ is the Bulow-Klemperer Theorem for a single bidder for which Dhangwatnotai et al. [6] gave an elegant proof. The proof of Lemma 5.3 can be seen as generalizing the proof of Dhangwatnotai et al. [6] to more than a single bidder.

⁵The point mass can be seen to have a triangle revenue curve peaking at $q = 1$ by being approximated by a uniform distribution on $[1, 1 + \epsilon]$ with arbitrarily small ϵ .

Example 5.8. Consider three distributions, the first having a triangle revenue curve peaking at $(0, 1)$ (i.e., having cdf $F_1(v) = 1 - \frac{1}{v+1}$), and the other two having triangle revenue curve peaking at $(\frac{1}{2}, \frac{1}{2})$. The ex ante optimal revenue is 2, but we show in Appendix C that the Vickrey auction with 3 duplicates has a revenue strictly less than 1.5.

5.1 Proof of Lemma 5.3.

Proof of Lemma 5.3. We say bidder $n + i$ is bidder i 's duplicate. We will show that SPALD has a revenue no less than LA. Let us condition on the realization of $v_1, \dots, v_{i^*-1}, v_{i^*+1}, \dots, v_n$ (which henceforth we denote as $\mathbf{v}_{[n]-i^*}$), and denote by v_{sh} the highest among them. We show that, conditioning on this, the lookahead auction (without duplicates) obtains no more revenue from bidder i^* than SPALD does from bidder i^* and her duplicate, bidder $n + i^*$.

We first analyze the posted price faced by i^* in the lookahead auction. Recall that the lookahead auction posts the revenue-optimal price for i^* conditioning on $v_{i^*} \geq v_{\text{sh}}$. Conditioning on $v_{i^*} \geq v_{\text{sh}}$, the revenue by posting any price $p \geq v_{\text{sh}}$ to bidder i^* is $p(1 - F_{i^*}(p))/(1 - F_{i^*}(v_{\text{sh}}))$, which is simply the unconditioned revenue at that price $p(1 - F_{i^*}(p))$ amplified by a constant factor $1/(1 - F_{i^*}(v_{\text{sh}}))$. Therefore the shape of the revenue curve conditioned on $v_{i^*} \geq v_{\text{sh}}$ is simply bidder i^* 's revenue curve to the left of $q(v_{\text{sh}})$ and amplified by a constant factor. Let r_i^* denote the monopoly reserve for bidder i . Since the original revenue curve is concave, it is nondecreasing to the left of $q_i^*(r_i^*)$, the quantile of the monopoly reserve price, and non-increasing to the right. Therefore, if v_{sh} is higher than $r_{i^*}^*$, the lookahead auction will simply post v_{sh} as the take-it-or-leave-it price for bidder i^* ; otherwise, the lookahead auction uses $r_{i^*}^*$ as the price.⁶

If $v_{\text{sh}} \geq r_{i^*}^*$, the lookahead auction extracts v_{sh} from bidder i^* when $v_{i^*} \geq v_{\text{sh}}$; but whenever $v_{i^*} \geq v_{\text{sh}}$, SPALD extracts a revenue at least v_{sh} from bidders i^* and $n + i^*$.

The case where $v_{\text{sh}} < r_{i^*}^*$ needs more calculation. Let R_{i^*} be the monopoly revenue extractable from bidder i^* (i.e., $r_{i^*}^*(1 - F_{i^*}(r_{i^*}^*))$). The lookahead auction, by posting the monopoly reserve price $r_{i^*}^*$, extracts revenue R_{i^*} from bidder i^* . We show that SPALD gets at least as much from bidder i^* and her duplicate.

Recall that the duplicate bidder's value v_{n+i^*} is drawn from the same distribution F_{i^*} as bidder i^* . With probability $F_{i^*}(v_{\text{sh}})$, v_{n+i^*} is no more than v_{sh} , and SPALD extracts a revenue of $\text{Rev}_{i^*}(q_{i^*}(v_{\text{sh}}))$ from bidder i^* . When $v_{n+i^*} \geq v_{\text{sh}}$, bidder i^* faces in SPALD a take-it-or-leave-it price of v_{n+i^*} , and the expected revenue from bidder i^* is precisely the height of Rev_{i^*} at the quantile of v_{n+i^*} . Since v_{n+i^*} is drawn from the same distribution, its quantile in F_{i^*} , namely, $q_{i^*}(v_{n+i^*}) = 1 - F_{i^*}(v_{n+i^*})$, is uniformly distributed on $[0, 1]$. Therefore the expected revenue when $v_{n+i^*} \geq v_{\text{sh}}$ is simply the area under bidder i^* 's revenue curve between 0 and $q_{i^*}(v_{\text{sh}})$. (See Figure 1 for an illustration.⁷) The expected revenue from the duplicate bidder $n + i^*$ is the same area: when v_{i^*} is at least v_{sh} , bidder $n + i^*$ faces a take-it-or-leave-it price of v_{i^*} , whose quantile in F_{n+i^*} is uniform between 0 and $q_{n+i^*}(v_{\text{sh}})$, and therefore the expected revenue from bidder $n + i^*$ is exactly the area under the revenue curve Rev_{i^*} between quantiles 0 and $q_{i^*}(v_{\text{sh}})$.

Summing up the parts, the revenue of SPALD from bidders i^* and $n + i^*$ when $v_{\text{sh}} \leq r_{i^*}^*$ is twice the area under the revenue curve Rev_{i^*} between 0 and $q_{i^*}(v_{\text{sh}})$, plus the area of a rectangle with height $\text{Rev}_{i^*}(q_{i^*}(v_{\text{sh}}))$ and width $1 - q_{i^*}(v_{\text{sh}})$. If one replaces the part of the revenue curve Rev_{i^*} between $q_{i^*}(r_{i^*}^*)$ and 1 by a straight line, the resulting curve is still concave, with the highest value

⁶This analysis essentially shows that the lookahead auction in our setting is the second price auction with monopoly reserve prices.

⁷This conversion from the expected revenue to an area under the revenue curve is an elegant technique from Dhangwatnotai et al. [6].

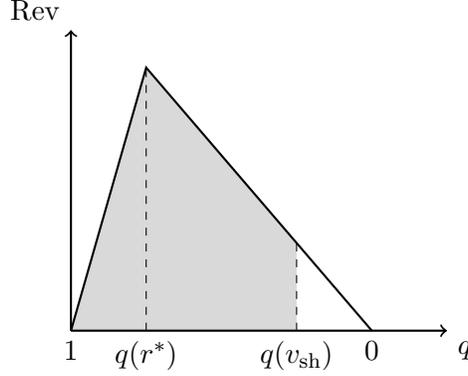


Figure 1: Illustration of a concave revenue curve. In this figure $v_{\text{sh}} < r^*$, and the shaded area is the expected revenue from this bidder when both she and her duplicate bidder bid above v_{sh} .

at R_{i^*} . Twice the area under this new curve is precisely SPALD's revenue from i^* and $n + i^*$. By the concavity of the curve, this is still at least R_{i^*} . This completes the proof. \square

5.2 Proof of Theorem 5.1

Proof of Theorem 5.1. Given any two regular distributions with revenue curves Rev_1 and Rev_2 , let \tilde{q}_1, \tilde{q}_2 be the solution to the ex ante revenue maximization problem, and let R_1 be $\text{Rev}_1(\tilde{q}_1)$ and R_2 be $\text{Rev}_2(\tilde{q}_2)$. Then the ex ante optimal revenue is $R_1 + R_2$.

If we replace the two distributions so that bidder i 's revenue curve is a triangle with the three vertices at $(0, 0)$, (\tilde{q}_i, R_i) and $(1, 0)$, then by Lemma C.2, the expected revenue of the second price auction with two duplicates is smaller whereas, by Lemma C.3, the ex ante optimal revenue remains the same. We therefore only need to show that the second price auction with duplicates extracts at least 0.75 fraction of the ex ante optimal revenue in this setting with triangular revenue curves.

By Lemma 5.3, it suffices to analyze the revenue of the lookahead auction (without duplicates) in this setting. Without loss of generality, suppose $\frac{R_1}{\tilde{q}_1} \geq \frac{R_2}{\tilde{q}_2}$. By the same analysis as in the proof of Lemma 5.3, bidder 1 always faces a take-it-or-leave-it price of R_1/\tilde{q}_1 , by which the auction extracts a revenue of R_1 from her. As for bidder 2, when $v_1 < R_2/\tilde{q}_2$, she is faced with a price at R_2/\tilde{q}_2 ; otherwise she faces a price of $v_1 > \tilde{q}_2$, which she can afford with probability 0. Therefore the expected revenue extracted from bidder 2 is R_2 times $\Pr[v_1 \leq R_2/\tilde{q}_2]$. This probability can be easily calculated from the geometry of the revenue curves: let this probability be s , then

$$\frac{sR_1}{1 - \tilde{q}_1} = \frac{(1 - s)R_2}{\tilde{q}_2} \Rightarrow s = \frac{\alpha(1 - \tilde{q}_1)}{\tilde{q}_2 + (1 - \tilde{q}_1)\alpha},$$

where α denotes the ratio $\frac{R_2}{R_1}$. The ratio between the revenue of the lookahead auction and that of the ex ante optimal revenue is therefore

$$\frac{R_1 + R_2s}{R_1 + R_2} = \frac{1 + \frac{\alpha^2(1 - \tilde{q}_1)}{\tilde{q}_2 + (1 - \tilde{q}_1)\alpha}}{1 + \alpha}.$$

We are interested in the minimum value of this ratio. It is evident that, everything else fixed, the ratio decreases with \tilde{q}_2 . Therefore, given the constraint $\tilde{q}_1 + \tilde{q}_2 \leq 1$, we may take $\tilde{q}_2 = 1 - \tilde{q}_1$. This

simplifies the ratio to

$$\frac{1 + \frac{\alpha^2 \tilde{q}_2}{\tilde{q}_2 + \alpha \tilde{q}_2}}{1 + \alpha} = \frac{1 + \alpha + \alpha^2}{(1 + \alpha)^2}.$$

This ratio is minimized at $\alpha = 1$, and evaluates to $\frac{3}{4}$. □

6 Conclusion and Open Questions

In this work we closed the gap in the number of duplicates needed for the VCG auction to yield a constant fraction of the optimal revenue. For both the single-item and k -item auctions, we showed tight bounds: instead of duplicating all bidders as required by Hartline and Roughgarden [11], only one (or k , respectively) duplicate is needed. These match the numbers in the i.i.d. settings in Bulow and Klemperer’s original result, and show that the only loss, when one removes symmetry from the i.i.d. setting, is a constant fraction of revenue, without need to grow the number of duplicates (although one needs to be more careful in choosing a duplicate).

We leave several interesting questions for future study. We have focused on obtaining a constant approximation, and our technique is unlikely to yield the tightest approximation achievable by a single duplicate. In fact, in the single-item auction, the worst revenue loss with a *single* duplicate we know of is only 0.31 fraction of the optimal revenue (see Remark 5.7 in Section 5). We believe the approximation ratio is closer to 2 than to 10.

We extended our results to k -items auctions in two settings. In the “free” setting, a bidder can be duplicated multiple times and all duplicates participate as all other bidders. This setting is most natural for a k -items auction, but cannot be extended to settings with more complex feasibility constraints. The “constrained” setting allows each bidder to be duplicated at most once, but allows at most one of a bidder and her duplicate to win. This can be generalized to domains such as matroid settings. Hartline and Roughgarden [11] showed their 2-approximation for n -duplicates in any matroid setting. It would be interesting to know whether k duplicates always suffice for an auction with a rank k matroid feasibility constraint.

We made progress on the ten-year-old open problem of the approximation ratio of the Vickrey auction with n duplicates. We showed that $\frac{4}{3}$ is the correct bound for $n = 2$ but as we remarked in Section 5, new techniques would be needed to tighten the analysis for $n > 2$.

References

- [1] Alaei, S., Hartline, J. D., Niazadeh, R., Pountourakis, E., and Yuan, Y. (2015). Optimal auctions vs. anonymous pricing. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 1446–1463.
- [2] Allouah, A. and Besbes, O. (2018). Prior-independent optimal auctions. In *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018*, page 503.
- [3] Bulow, J. and Klemperer, P. (1996). Auctions versus negotiations. *The American Economic Review*, 86(1):180–194.
- [4] Bulow, J. and Roberts, J. (1989). The simple economics of optimal auctions. *Journal of Political Economy*, 97(5):1060–90.
- [5] Chawla, S., Hartline, J. D., Malec, D. L., and Sivan, B. (2013). Prior-independent mechanisms for scheduling. In *Symposium on Theory of Computing Conference, STOC’13, Palo Alto, CA, USA, June 1-4, 2013*, pages 51–60.
- [6] Dhangwatnotai, P., Roughgarden, T., and Yan, Q. (2015). Revenue maximization with a single sample. *Games and Economic Behavior*, 91:318–333.
- [7] Eden, A., Feldman, M., Friedler, O., Talgam-Cohen, I., and Weinberg, S. M. (2017). The competition complexity of auctions: A bulow-klemperer result for multi-dimensional bidders. In *Proceedings of the 2017 ACM Conference on Economics and Computation, EC ’17, Cambridge, MA, USA, June 26-30, 2017*, page 343.
- [8] Feldman, M., Friedler, O., and Rubinstein, A. (2018). 99% revenue via enhanced competition. In *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018*, pages 443–460.
- [9] Fu, H., Immorlica, N., Lucier, B., and Strack, P. (2015). Randomization beats second price as a prior-independent auction. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC ’15, Portland, OR, USA, June 15-19, 2015*, page 323.
- [10] Hartline, J. (2017). *Mechanism Design and Approximation*.
- [11] Hartline, J. D. and Roughgarden, T. (2009). Simple versus optimal mechanisms. In *Proceedings 10th ACM Conference on Electronic Commerce (EC-2009), Stanford, California, USA, July 6–10, 2009*, pages 225–234.
- [12] Hoeffding, W. (1956). On the distribution of the number of successes in independent trials. *Ann. Math. Statist.*, 27(3):713–721.
- [13] Kaas, R. and Buhrman, J. M. (1980). Mean, median and mode in binomial distributions. *Statistica Neerlandica*, 34(1):13–18.
- [14] Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):pp. 58–73.
- [15] Ronen, A. (2001). On approximating optimal auctions. In *ACM Conference on Electronic Commerce*, pages 11–17.

- [16] Roughgarden, T., Talgam-Cohen, I., and Yan, Q. (2012). Supply-limiting mechanisms. In *ACM Conference on Electronic Commerce, EC '12, Valencia, Spain, June 4-8, 2012*, pages 844–861.
- [17] Sivan, B. and Syrgkanis, V. (2013). Vickrey auctions for irregular distributions. In *Web and Internet Economics - 9th International Conference, WINE 2013, Cambridge, MA, USA, December 11-14, 2013, Proceedings*, pages 422–435.
- [18] Vickrey, W. (1961). Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance*, pages 8–37.

A Technical results

Claim A.1. *Suppose that $0 \leq q \leq q' \leq \beta \leq 1$. Then for any regular distribution, we have $\text{Rev}(q') \geq (1 - \beta) \text{Rev}(q)$.*

Proof. By the concavity of the revenue curves, we have

$$\text{Rev}(q') \geq \frac{1 - q'}{1 - q} \text{Rev}(q) + \frac{q' - q}{1 - q} \text{Rev}(1) \geq (1 - \beta) \text{Rev}(q),$$

since $\text{Rev}(1) \geq 0$. □

Claim A.2. *Let $q \in [0, 1/2]$ and $q' \in [0, 1]$ be such that $|q - q'| \leq \epsilon q$ for some $\epsilon \in [0, 1]$. Then $(1 - \epsilon) \text{Rev}(q) \leq \text{Rev}(q') \leq \frac{1}{1 - \epsilon} \text{Rev}(q)$.*

Proof. First suppose that $q' > q$. By concavity of the revenue curve, we have

$$\begin{aligned} \text{Rev}(q') &\geq \frac{1 - q'}{1 - q} \text{Rev}(q) + \frac{q - q'}{1 - q} \text{Rev}(1) \\ &\geq \frac{1 - q'}{1 - q} \text{Rev}(q) \\ &\geq \frac{1 - q + q - q'}{1 - q} \text{Rev}(q) \\ &= \left(1 - \frac{q' - q}{1 - q}\right) \text{Rev}(q). \end{aligned}$$

Since $q \leq 1/2$, we have $q \leq 1 - q$. So

$$\begin{aligned} \text{Rev}(q') &\geq \left(1 - \frac{q' - q}{1 - q}\right) \text{Rev}(q) \\ &\geq \left(1 - \frac{\epsilon q}{1 - q}\right) \text{Rev}(q) \\ &\geq \left(1 - \frac{\epsilon(1 - q)}{1 - q}\right) \text{Rev}(q) \\ &= (1 - \epsilon) \text{Rev}(q). \end{aligned}$$

This proves the lower bound on $\text{Rev}(q')$. To get the upper bound, we use concavity again and note that

$$\begin{aligned} \text{Rev}(q) &\geq \frac{q}{q'} \text{Rev}(q') + \frac{q' - q}{q'} \text{Rev}(0) \\ &\geq \frac{q}{q'} \text{Rev}(q') \\ &\geq \frac{q' - \epsilon q}{q'} \text{Rev}(q') \\ &= (1 - \epsilon q/q') \text{Rev}(q') \\ &\geq (1 - \epsilon) \text{Rev}(q'), \end{aligned}$$

where the last inequality is because $q < q'$. So $\text{Rev}(q') \leq \frac{1}{1 - \epsilon} \text{Rev}(q)$.

The case of $q' < q$ follows from calculations which are similar to those above. Indeed

$$\begin{aligned} \text{Rev}(q') &\geq \frac{q'}{q} \text{Rev}(q) + \frac{q - q'}{q} \text{Rev}(0) \\ &\geq \frac{q'}{q} \text{Rev}(q) \\ &\geq \frac{(1 - \epsilon)q}{q} \text{Rev}(q) \\ &= (1 - \epsilon) \text{Rev}(q) \end{aligned}$$

and

$$\text{Rev}(q) \geq \left(1 - \frac{q - q'}{1 - q'} \text{Rev}(q')\right).$$

Since $q' < q \leq 1/2$, we have $q \leq 1 - q \leq 1 - q'$. So $q - q' \leq \epsilon q \leq \epsilon(1 - q')$ and this shows that $\text{Rev}(q) \geq (1 - \epsilon) \text{Rev}(q')$ as desired. \square

Let $\text{Ber}(p)$ denote the Bernoulli distribution with mean p and $\text{Binom}(n, p)$ be the binomial distribution.

Theorem A.3 (Theorem 4 in [12]). *Let X_1, \dots, X_n be random variables where $X_i \sim \text{Ber}(p_i)$. Let $S = \sum_{i=1}^n X_i$ and set $p = n^{-1} \mathbf{E}[S]$. Then for all $c \leq np - 1$, $\Pr[S \leq c] \leq \Pr[\text{Binom}(n, p) \leq c]$.*

Fact A.4 (Corollary 1 in [13]). *The median of $\text{Binom}(n, p)$ is at least $\lfloor np \rfloor$.*

Let X_1, \dots, X_n be random variables where $X_i \sim \text{Ber}(p_i)$, $S = \sum_{i=1}^n X_i$, and $p = n^{-1} \mathbf{E}[S]$. Since $np - 1 \leq \lfloor np \rfloor$, Theorem A.3 and Fact A.4 imply that

$$\Pr[S \leq np - 1] \leq \Pr[S < \lfloor np \rfloor] \leq 1/2.$$

Hence, $\Pr[S \geq np] \geq 1/2$.

B Additional Results for k -item Auctions

Theorem 4.2 only guarantees the existence of at most k bidders which can be duplicated so that VCG in the duplication environment approximates the ex-ante optimal revenue. The next result shows that under mild assumptions and (noisy) knowledge about the revenue curve, the auctioneer can determine which bidders ought to be duplicated.

Theorem B.1. *Consider a k -item auction with n bidders whose values are drawn independently from (non-identical) regular distributions F_1, \dots, F_n . Suppose that the auctioneer has access to the following oracle for each bidder i for some $\beta \in (0, 1/2]$ and $\epsilon \in (0, 1)$: the oracle returns $\text{Rev}_i(\beta'_i)$ with the promise that $|\beta - \beta_i| \leq \epsilon\beta$. Let S be the k highest bidders in terms of $\text{Rev}_i(\beta'_i)$. Then duplicating every bidder in S and running the VCG mechanism (restricting that a bidder and her duplicate cannot both win) on the $n + k$ bidders, the auctioneer can extract a revenue of at least $c(\beta, \epsilon) \cdot \text{OPT}$ where $c(\beta, \epsilon)$ is a constant depending only on β and ϵ . Moreover, $c(0.5, \epsilon) \geq (1 - \epsilon)^3 \cdot 0.1$.*

Proof. Let γ, δ satisfy $\frac{1}{\gamma}((1 - \beta)(1 - \gamma) - \delta) \geq \frac{3}{2}$. By shuffling the indices, let us assume that $\text{Rev}_1(\beta) \geq \dots \geq \text{Rev}_n(\beta)$.

Again, we consider the following three cases given by Lemma 4.3. However, we will have to be a little careful with how we set up the cases (note that the cases are not necessarily a partition).

Case 1: $v_k(\beta) < (1 - \epsilon)^2 \frac{\gamma}{k}$ but there exists $q_1, \dots, q_n \in [\beta, 1)$ and a set $H \subseteq [n]$ with $|H| < k$ such that $v_i(\beta) \geq \frac{\gamma}{k}$ for all $i \in H$ and $\sum_{i \in H} \text{Rev}_i(q_i) \geq \delta \cdot \text{OPT}$. The assumptions in the case imply that we can assume $H = \{i \in S : v_i(\beta) \geq \frac{\gamma \text{OPT}}{k}\}$. Let S be the top k highest bidders in terms of $\text{Rev}_i(\beta'_i)$. We claim that H is a subset of S .

Claim B.2. *It holds that $H \subseteq S$.*

Proof. Suppose $i \in H$. Then $v_i(\beta) \geq \frac{\gamma}{k} \text{OPT}$ so $\text{Rev}_i(\beta) \geq \frac{\gamma^\beta}{k} \text{OPT}$. By Claim A.2, this implies that $\text{Rev}_i(\beta'_i) \geq (1 - \epsilon) \frac{\gamma^\beta}{k} \text{OPT}$.

On the other hand, suppose that $i \geq k$. Then $v_i(\beta) < (1 - \epsilon)^2 \frac{\gamma}{k} \text{OPT}$ so $\text{Rev}_i(\beta) < (1 - \epsilon)^2 \frac{\gamma^\beta}{k} \text{OPT}$. Hence, applying Claim A.2 again, we have that $\text{Rev}_i(\beta'_i) < (1 - \epsilon) \frac{\gamma^\beta}{k} \text{OPT}$.

The two previous paragraphs imply that if $i \in H$ then $\text{Rev}_i(\beta'_i) \geq \text{Rev}_j(\beta'_j)$ for all $j \geq k$. Since S contains the top k bidders in terms of $\text{Rev}_i(\beta'_i)$ this implies that S contains H completing the proof. \square

Since we duplicate every bidder in H , Theorem 2.2 implies that a lower bound on the revenue of VCG on the resulting environment is $\sum_{i \in H} \text{Rev}_i(\beta) \geq \delta \cdot \text{OPT}$.

Case 2: There exists a set $H \subseteq [n]$ with $|H| \geq k$ such that $v_i(\beta) \geq (1 - \epsilon)^2 \frac{\gamma}{k} \text{OPT}$. In this case, we have $\text{Rev}_i(\beta) \geq (1 - \epsilon)^2 \frac{\beta \gamma}{k} \text{OPT}$ for $1 \leq i \leq k$. Let H be the k highest bidders in terms of $\text{Rev}_i(\beta'_i)$. Then

$$\begin{aligned} \sum_{i \in H} \text{Rev}_i(\beta'_i) &\geq \sum_{i=1}^k \text{Rev}_i(\beta'_i) \\ &\geq \sum_{i=1}^k (1 - \epsilon) \text{Rev}_i(\beta) \quad (\text{by Claim A.2}) \\ &\geq \sum_{i=1}^k (1 - \epsilon)^3 \frac{\beta \gamma}{k} \text{OPT} \\ &\geq (1 - \epsilon)^3 \beta \gamma \text{OPT}. \end{aligned}$$

Hence, duplicating the bidders in H obtains and running VCG extracts a revenue of at least $(1 - \epsilon)^3 \beta \gamma \text{OPT}$.

Case 3: With probability at least $1/2$ there exists at least $k+1$ bidders who bid at least $\frac{\gamma}{k} \text{OPT}$. This case is straightforward since without duplicating, VCG already extracts a revenue of $\frac{\gamma}{2} \text{OPT}$.

Combining the three cases we see that we can obtain a revenue of at least $\min\{(1 - \epsilon)\beta\delta, (1 - \epsilon)\beta\gamma, \gamma(1/2)\} \text{OPT}$. In particular, one can take

$$c(\beta, \epsilon) = \max_{\substack{\gamma, \delta \in [0, 1]: \\ \frac{1}{\gamma}((1-\beta)(1-\gamma)-\delta) \geq 3/2}} \min\{\delta, (1 - \epsilon)^3 \gamma \beta, \gamma/2\}.$$

Finally, observe that if one chooses $\beta = 0.5, \delta = 0.1, \gamma = 0.2$ then one obtains that $c(0.5, \epsilon) \geq (1 - \epsilon)^3 \cdot 0.1$. \square

C Missing proofs from Section 5

Lemma C.1. *Let R and \tilde{R} be the revenue curves of two value distributions F and \tilde{F} . F stochastically dominates \tilde{F} if and only if R pointwise dominates \tilde{R} in the sense that for any $q \in [0, 1]$, $R(q) \geq \tilde{R}(q)$.*

Proof. Suppose F stochastically dominates \tilde{F} , for any quantile q , $F^{-1}(1 - q) \geq \tilde{F}^{-1}(1 - q)$, and so $R(q) = qF^{-1}(1 - q) \geq q\tilde{F}^{-1}(1 - q) = \tilde{R}(q)$. Conversely, if R pointwise dominates \tilde{R} , for any value v , let the quantile of v in F be q , and that in \tilde{F} be \tilde{q} . Then $\Pr_{w \sim F}[w \geq v] = \frac{R(q)}{v} \geq \frac{\tilde{R}(q)}{v} = \Pr_{w \sim \tilde{F}}[w \geq v]$, which shows that F stochastically dominates \tilde{F} . \square

Lemma C.2. *In a second price auction where all bidders' values are drawn independently, if we replace bidder i by another bidder i' whose revenue curve is dominated by that of bidder i 's, the new auction extracts no more expected revenue than the original.*

Proof. The revenue of the second price auction is the expected value of the second highest bid. By Lemma C.1, the operation in the lemma statement replaces one value distribution by another stochastically dominated by the former. It is immediate that this decreases the expected value of any order statistics, including the revenue of the second price auction. \square

Lemma C.3. *In a single item auction setting with values independently drawn from regular distributions whose revenue curves are $\text{Rev}_1, \dots, \text{Rev}_n$, suppose the optimal solution to the ex ante revenue maximization problem are the quantiles $\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_n$, these remain an optimal solution (to the problem of ex ante revenue maximization) when each bidder i 's value distribution is replaced by one whose revenue curve is a triangle whose three vertices are $(0, 0), (1, 0)$ and $(\tilde{q}_i, \text{Rev}_i(\tilde{q}_i))$.*

This lemma was proved and used in Alaei et al. [1]. For completeness we give a proof.

Proof. Note that $\tilde{q}_1, \dots, \tilde{q}_n$ constitute a feasible solution to the ex ante revenue in the setting with triangle revenue curves. Therefore the ex ante optimal revenue in the new setting is no less than before. However, it cannot be more either. This is because each revenue curve in the new setting is pointwise dominated by the original one, by the concavity of the latter, so any feasible solution gives weakly less revenue in the new setting. This shows the optimality of $\tilde{q}_1, \dots, \tilde{q}_n$. \square

Calculation of Remark 5.7. If we duplicate the deterministic bidder then the revenue is 1.

On the other, suppose that we duplicate the bidder whose cdf is $F_2(v) = 1 - \frac{1}{v+1}$. Let us analyze the revenue of the Vickrey auction in this case. Let v_1 be the value of the deterministic bidder (so $v_1 = 1$ with probability 1) and $v_2, v_3 \sim F_2$.

If $v_2, v_3 \leq 1$ (which occurs with probability 1/4) then the Vickrey auction extracts $\max\{v_2, v_3\}$. So the average revenue, conditioned on $v_2, v_3 \leq 1$ is $\int_0^1 1 - \left(2 - \frac{2}{t+1}\right)^2 dt = 4 \ln(4) - 5$. Hence, this case contributes $\ln(4) - 5/4$ to the revenue.

If $v_2 \leq 1 \leq v_3$ or $v_3 \leq 1 \leq v_2$ (each occurring with probability 1/4) then the revenue is 1. These two cases contribute 1/2 to the revenue.

Finally, if $v_2, v_3 \geq 1$ (which occurs with probability 1/4) then the revenue is $1 + \int_1^\infty \frac{4}{(t+1)^2} dt = 1 + 2$. So this case contributes 3/4 to the revenue.

Hence, the average revenue is $\ln(4)$.

The revenue of the optimal auction with the original bidders is ≈ 2 so the approximation ratio is $\frac{2}{\ln 4} \approx 1.44$. \square

Calculation of Example 5.8. Note that the highest value in the supports of the latter two distributions is 1. Let u be the highest value among the bidders with values drawn from the latter two distributions. Conditioning on any u , the revenue of the Vickrey auction with duplicates is at most $u + \int_u^\infty \frac{1}{(v+1)^2} dv = u + \frac{1}{u+1}$. For $u \geq 0$ this is increasing in u , and is 1.5 when u is 1. However, with constant probability u is strictly smaller than, say, 0.9, and therefore in expectation the Vickrey auction with duplicates extracts a revenue strictly smaller than 1.5. \square