

Bayesian Auctions with Efficient Queries*

Jing Chen[†] Bo Li[†] Yingkai Li[†] Pinyan Lu[‡]

[†]Department of Computer Science, Stony Brook University
Stony Brook, NY 11794, USA

{jingchen, boli2, yingkli}@cs.stonybrook.edu

[‡]Institute for Theoretical Computer Science, Shanghai University of Finance and Economics
Shanghai 200433, China

lu.pinyan@mail.shufe.edu.cn

Abstract

Designing dominant-strategy incentive compatible (DSIC) mechanisms to generate (approximately) optimal revenue is a fundamental problem in Bayesian mechanism design. However, most existing studies assume that the seller knows the entire distribution. Unfortunately, this assumption may not hold in reality: for example, when the distributions have exponentially large supports or do not have succinct representations. In this work we consider, for the first time, the *query complexity* of Bayesian mechanisms. The seller only has limited oracle accesses to the players' distributions, via *quantile queries* and *value queries*. We prove *logarithmic* lower-bounds for any constant approximation DSIC mechanisms in a large class of auction settings. For single-item auctions and multi-item auctions with unit-demand or additive valuation functions, we prove *tight* upper-bounds via efficient query schemes. Thus, in those settings the seller needs to access much less than the entire distribution to achieve approximately optimal revenue.

Keywords: mechanism design, the complexity of Bayesian mechanisms, query complexity, quantile queries, value queries

*This work has been partially supported by NSF CAREER Award (No. 1553385), National NSF of China (No. 61741209) and the Fundamental Research Funds for the Central Universities. Part of this work was done when the first three authors were visiting Shanghai University of Finance and Economics.

1 Introduction

An important problem in Bayesian mechanism design is to design auctions that (approximately) maximize the seller’s expected revenue. More precisely, in a Bayesian multi-item auction a seller has m heterogeneous items to sell to n players. Each player i has a private value for each item j , v_{ij} ; and each v_{ij} is independently drawn from some prior distribution \mathcal{D}_{ij} . When the prior distribution $\mathcal{D} \triangleq \times_{ij} \mathcal{D}_{ij}$ is of *common knowledge* to both the seller and the players, optimal Bayesian incentive-compatible (BIC) mechanisms have been discovered for various auction settings [31, 16, 7, 8], where all players reporting their true values forms a Bayesian Nash equilibrium. When there is no common prior but the seller knows \mathcal{D} , many (approximately) optimal dominant-strategy incentive-compatible (DSIC) Bayesian mechanisms have been designed [31, 32, 12, 29, 35, 9], where it is each player’s *dominant strategy* to report his true values.

However, the *complexity* for the seller to carry out such mechanisms is largely unconsidered in the literature. Most existing Bayesian mechanisms require that the seller has full access to the prior distribution \mathcal{D} and is able to carry out all required optimizations based on \mathcal{D} , so as to compute the allocation and the prices. Unfortunately the seller may not be so knowledgeable or powerful in real-world scenarios. If the supports of the distributions are exponentially large (in m and n), or if the distributions are continuous and do not have succinct representations, it is hard for the seller to write out “each single bit” of the distributions or precisely carry out arbitrary optimization tasks based on them. Even with a single player and a single item, when the value distribution is irregular, computing the optimal price in time that is much smaller than the size of the support is not an easy task. Thus, a natural and important question to ask is *how much the seller should know about the distributions in order to obtain approximately optimal revenue*.

In this work we consider, for the first time, the *query complexity* of Bayesian mechanisms. In particular, the seller can only access the distributions by making oracle queries. Two natural types of queries are allowed, *quantile queries* and *value queries*. That is, the seller queries the oracle with specific quantiles (respectively, values), and the oracle returns the corresponding values (respectively, quantiles) in the underlying distributions. These two types of queries happen a lot in market study. Indeed, the seller may wish to know what is the price he should set so that half of the consumers would purchase his product; or if he sets the price to be 200 dollars, how many consumers would buy it. Another important scenario where such queries naturally come up is in databases. Indeed, although the seller may not know the distribution, some powerful institutes, say the Office for National Statistics, may have such information figured out and stored in its database. As in most database applications, it may be neither necessary nor feasible for the seller to download the whole distribution to his local machines. Rather, he would like to access the distribution via queries to the database. Other types of queries are of course possible, and will be considered in future works.

In this work we focus on *non-adaptive* queries. That is, the seller makes all oracle queries simultaneously, before the auction starts. This is also natural in both database and market study scenarios, and in Section 3, we will show that the performance of adaptive queries cannot be improved up to a logarithmic factor.

1.1 Main Results

We would like to understand both lower- and upper-bounds for the query complexity of approximately optimal Bayesian auctions. In this work, we mainly consider three widely studied settings: single-item auctions and multi-item auctions with unit-demand or additive valuation functions. Our main results are summarized in Table 1.

Note that we allow arbitrary unbounded distributions that satisfy *small-tail assumptions*, with

formal definitions provided in Section 5.1. Similar assumptions are widely adopted in sampling mechanisms [33, 18], to deal with irregular distributions with unbounded supports. Since distributions with bounded supports automatically satisfy the small-tail assumptions, the lower-bounds listed for the former apply to the latter as well.

	Query	Distributions	
	Complexity	Bounded in $[1, H]$	Unbounded & Small Tail
Auctions	Single-Item	$\Theta(n\epsilon^{-1} \log H)$	
	Unit-Demand	$\forall c > 1: \Omega(\frac{mn \log H}{\log c})$	$\forall c > 24: O(\frac{mn \log H}{\log(c/24)})$
	Additive	$\forall c > 1: \Omega(\frac{mn \log H}{\log c})$	$\forall c > 8: O(\frac{mn \log H}{\log(c/8)})$
	Single-Item	Regular Distributions: $\Omega(n\epsilon^{-1}), O(n\epsilon^{-1} \log \frac{n}{\epsilon})$	

Table 1: Our main results. Here $h(\cdot) < 1$ is the tail function in the small-tail assumptions. For single-item auctions, the revenue is a $(1 + \epsilon)$ -approximation to the optimal BIC revenue, with ϵ sufficiently small. For multi-item auctions with unit-demand or additive valuation functions, the revenue is a c -approximation for some constant c .

Also note that our lower- and upper-bounds on query complexity are *tight* for bounded distributions. As will become clear in Section 3, our lower-bounds allow the seller to make both value and quantile queries, and apply to any multi-player multi-item auctions where each player’s valuation function is *succinct sub-additive*: formal definitions in Section 3. The lower-bounds also allow randomized queries and randomized mechanisms.

For the upper-bounds, all our query schemes are deterministic and only make one type of queries: value queries for bounded distributions and quantile queries in the other cases; see Sections 4 and 5. We show that our schemes, despite of being very efficient, only loses a small fraction of revenue compared with when the seller has full access to the distributions.

1.2 Discussions

Sample Complexity. A closely related area to our work is sampling mechanisms [15, 3, 28, 19, 30, 18, 21, 22, 6]. It assumes that the seller does not know \mathcal{D} but observes independent samples from \mathcal{D} before the auction begins. The *sample complexity* measures how many samples the seller needs so as to obtain a good approximation to the optimal Bayesian revenue. The best-known sample complexity results are summarized in Table 2.

Auctions	Single-Item (regular) [18]	Single-Item (bounded) [18]	Unit-Demand (bounded) [30]	Additive (bounded)[6]
Sample Complexity	$\tilde{O}(n\epsilon^{-4})$	$\tilde{O}(nH\epsilon^{-3})$	$\tilde{O}(nm^2H^2\epsilon^{-2})$	$\tilde{O}(nm^2H^2\epsilon^{-2})$
Approximations	$1 + \epsilon$	$1 + \epsilon$	27	32

Table 2: Sample Complexity. When a distribution is bounded, it is bounded within $[1, H]$. For multi-item auctions with unit-demand or additive valuation functions, the revenue has an extra ϵ additive loss.

Oracle queries can be seen as *targeted samples*, where the seller actively asks the information he needs rather than passively learns about it from random samples. As such, it is intuitive that queries are more efficient than samples, but it is a priori unclear how efficient queries can be. Our results answer this question quantitatively and show that query complexity can be exponentially smaller than sample complexity: the former is *logarithmic* in the “size” of the distributions, while the latter is polynomial.

Finally, the design of query mechanisms facilitates the design of sampling mechanisms. If the seller observes enough samples from \mathcal{D} , then he can mimic quantile queries and apply query mechanisms: see Section 7 for more details.

Parametric Auctions. Parametric mechanisms [2, 1] assume the seller only knows some specific parameters about the distributions, such as the mean, the median (or a single quantile), and the variance. Note that using quantile or value queries, one can get the exact value of the median and the approximate value of the mean, and then apply parametric mechanisms. However, existing parametric mechanisms only consider single-parameter auctions, where the distributions are regular or have monotone hazard rate. Since our mechanisms make non-adaptive oracle queries, our results imply parametric mechanisms in multi-parameter settings with general distributions, where the “parameters” are the oracle’s answers to our query schemes. Our lower-bounds imply that knowing only the median is not enough to achieve the same approximation ratios as we do. Finally, it remains unknown whether constant approximations can be achieved for multi-parameter auctions or general distributions, knowing only the mean and the variance.

Distributions within Bounded Distance. Recently, [6] studies auctions where the true prior distribution is unknown to the seller, but he is given a distribution that is close to the true prior, as measured by the Kolmogorov distance. On the one hand, the learnt distributions from our query schemes can be far from the true prior in terms of the Kolmogorov distance, thus their mechanisms do not apply. On the other hand, although a distribution close to the true prior may be learnt via sufficiently many oracle queries, our lower-bounds imply that the query complexity of this approach will not be better than ours.

Using Experts as Oracles. If the players’ value distributions are known by some experts, then the seller can use the experts as oracles. Indeed, we are able to design proper scoring rules [5, 11] for the seller to elicit truthful answers from the experts for his queries. If the experts are actually players in the auctions, then they have their own stakes about the final allocation and prices, and it would be interesting to see how the seller can still use them as oracles and get truthful answers for his queries, while keeping them truthful about their own values. See [14] for more discussions on this front.

1.3 Other Related Works

The complexity of auctions is an important topic in the literature, and several complexity measures have been considered. Following the taxation principle [24, 23], [26] defines the *menu complexity* of truthful auctions. For a single additive buyer, [17] shows the optimal Bayesian auction for revenue can have an infinite menu size or a continuum of menu entries, and [4] shows a constant approximation under finite menu complexity. Recently, [20] considers the taxation, communication, query and menu complexities of truthful combinatorial auctions, and shows important connections among them. The queries considered there are totally different from ours: we are concerned with the

complexity of accessing the players’ value distributions in Bayesian settings, while [20] is concerned with the complexity of accessing the players’ valuation functions in non-Bayesian settings.

1.4 Future Directions

Many interesting questions about the query complexity of Bayesian auctions are worth exploring. First, as mentioned, there is a logarithmic gap between adaptive queries and non-adaptive queries for multi-item auctions. It is intriguing to design approximately optimal Bayesian mechanisms with matching query complexity using adaptive queries. Another interesting direction is when the answers of the oracle contain noise. In this case, the distributions learnt by the seller may be within a small distance from the “true distributions” defined by oracle answers without noise. This is related to [6] and it would be interesting to design mechanisms to handle such noise.

2 Preliminaries

2.1 Bayesian Auctions

In a multi-item auction there are m items, denoted by $M = \{1, \dots, m\}$, and n players, denoted by $N = \{1, \dots, n\}$. Each player $i \in N$ has a non-negative value for each item $j \in M$, v_{ij} , which is independently drawn from distribution \mathcal{D}_{ij} . Player i ’s *true valuation* is $(v_{ij})_{j \in [m]}$. To simplify the notations, we may write v_i for $(v_{ij})_{j \in [m]}$ and v for $(v_i)_{i \in [n]}$. Letting $\mathcal{D}_i = \times_{j \in M} \mathcal{D}_{ij}$ and $\mathcal{D} = \times_{i \in N} \mathcal{D}_i$, we use $\mathcal{I} = (N, M, \mathcal{D})$ to denote the corresponding Bayesian auction instance and $OPT(\mathcal{I})$ the optimal BIC revenue of \mathcal{I} . When \mathcal{I} is clear from the context, we write OPT for short.

We will consider several classes of widely studied auctions. A *single-item* auction has $m = 1$. When $m > 1$, a bidder i being *unit-demand* means his value for a subset S of items is $\max_{j \in S} v_{ij}$, and a bidder i being *additive* means his value for S is $\sum_{j \in S} v_{ij}$. When all bidders are unit-demand (respectively, additive), we call such an auction a *unit-demand auction* (respectively, an *additive auction*) for short.

2.2 Query Complexity

In this work, we only allow the seller to access the prior distributions via two types of oracle queries: *value queries* and *quantile queries*. Given a distribution D over reals, in a value query, the seller sends a value $v \in \mathbb{R}$ and the oracle returns the corresponding quantile $q(v) \triangleq \Pr_{x \sim D}[x \geq v]$. In a quantile query, the seller sends a quantile $q \in [0, 1]$ and the oracle returns the corresponding value $v(q)$ such that $\Pr_{x \sim D}[x \geq v(q)] = q$. With *non-adaptive* queries, the seller first sends all his queries to the oracle, gets the answers back, and then runs the auction. The *query complexity* is the number of queries made by the seller.

Note that the answer to a value query is unique. The quantile queries are a bit tricky, as for discrete distributions there may be multiple values corresponding to the same quantile q , or there may be none. When there are multiple values, to resolve the ambiguity, let the output of the oracle be the largest one: that is, $v(q) = \arg \max_z \{\Pr_{x \sim D}[x \geq z] = q\}$. When there is no value corresponding to q , the oracle returns the largest value whose corresponding quantile is larger than q : that is, $v(q) = \arg \max_z \{\Pr_{x \sim D}[x \geq z] > q\}$. So for any quantile query q , $v(q) = \arg \max_z \{\Pr_{x \sim D}[x \geq z] \geq q\}$ in general. For any discrete distribution D and quantile query $q > 0$, $v(q)$ is always in the support of D . When $q = 0$, $v(q)$ may be $+\infty$.

3 Lower Bounds

In this section, we prove lower bounds for the query complexity of Bayesian mechanisms, and we focus on DSIC mechanisms. As a building block for our general lower bound, we first have the following for *single-item single-player* auctions. Note that although the main focus of this paper is about non-adaptive queries, here we prove a stronger result by showing that even with adaptive queries, our bounds for *single-item single-player* auctions are still tight. Formally, we have the following lemma.

Lemma 1. *For any constant $c > 1$, there exists a constant C such that, for any large enough H , any Bayesian mechanism \mathcal{M} making less than $C \log_c H$ adaptive value and quantile queries to the oracle, there exists a single-player single-item Bayesian auction instance $\mathcal{I} = (N, M, \mathcal{D})$ where the values are bounded in $[1, H]$, such that $\text{Rev}(\mathcal{M}(\mathcal{I})) < \frac{\text{OPT}(\mathcal{I})}{c}$.*

Proof. Here we consider the equal revenue curve, i.e., $F(v) = 1 - \frac{1}{v}$, $v \in [1, H]$ and $F(H) = 1$. For this distribution, posting any price from $[1, H]$ has exactly the same revenue 1. For any constant H , let $k \triangleq \lfloor \frac{1}{4} \log_{(4c)}^{4c+2} H \rfloor$. We divide the value interval $[1, H]$ into $k + 1$ sub-intervals as follows: from right to left, $u_k = H$, and $u_s = \frac{u_{s+1}}{(4c)^{4c+2}}$ for each $s \in \{k-1, \dots, 0\}$. For each value interval (u_s, u_{s+1}) , there exists a corresponding quantile interval $(q_s, q_{s+1}) \triangleq (\frac{1}{u_s}, \frac{1}{u_{s+1}})$.

For each value $s \in \{k-1, \dots, 0\}$, consider pair of interval (u_s, u_{s+1}) and (q_s, q_{s+1}) , we construct $\lceil 4c \rceil$ distributions for it. More precisely, the distribution \mathcal{D}_z^s for each $z \in \lceil 4c \rceil$ is defined as follows. For $v_z < u_s$ or $u_{s+1} < v_z < H$, the density of v_z is $\frac{1}{v_z^2}$. The probability of $v_z = H$ is $\frac{1}{H}$. The probability of $(4c)^z u_s$ is $q_{s+1} - q_s$.

By construction, given parameter s , all $\lceil 4c \rceil$ distributions coincide outside the value and quantile range (u_s, u_{s+1}) and (q_s, q_{s+1}) . Therefore, only a query inside the range can distinguish those distributions from each other. Moreover, for any two parameter $s, s' \in \{k-1, \dots, 0\}$, the value and quantile queries will get the exactly same response for queries outside their own value and quantile intervals. Thus, for any adaptive value and quantile queries, a single query can distinguish at most one set of $\lceil 4c \rceil$ distributions. Letting $c' \triangleq 1 - \frac{1}{2c}$ and $C \triangleq \frac{1-c'}{8(4c+2) \log_c(4c)} = \frac{1}{16c(4c+2) \log_c(4c)}$, we have $C \log_c H < k(1 - c')$. Accordingly, for any Bayesian mechanism \mathcal{M} that makes less than $C \log_c H$ adaptive value and quantile queries, there exists a value s such that, with probability at least c' , \mathcal{M} cannot distinguish \mathcal{D}_z^s 's from each other.

We now analyze the optimal BIC revenue for those instances. For any $\mathcal{I}_z = (N, M, \mathcal{D}_z)$, Myerson's mechanism is optimal: it sets a (randomized) threshold for the unique player, if the player bids at least the threshold then he gets the item and pays the threshold payment, otherwise the item is unsold. It is not hard to verify that $\text{OPT}(\mathcal{I}_z) = (4c)^z u_s q_{s+1}$ for each \mathcal{I}_z .

Next, we analyze the revenue of \mathcal{M} . Since \mathcal{M} is DSIC, the allocation rule must be monotone in the player's bid, and he will pay the threshold payment set by \mathcal{M} , denoted by P . Here P may also be randomized. Note that for all instances, setting $P < 4cu_s$ is strictly worse than setting $P = 4cu_s$, and setting $P > (4c)^{\lceil 4c \rceil} u_s$ is strictly worse than setting $P = (4c)^{\lceil 4c \rceil} u_s < u_{s+1}$. Also, for any instance \mathcal{I}_z and any $z' \in \{1, \dots, \lceil 4c \rceil - 1\}$, setting $P \in ((4c)^{z'} u_s, (4c)^{z'+1} u_s)$ is strictly worse than setting $P = (4c)^{z'+1} u_s$. Thus, when mechanism \mathcal{M} cannot distinguish the \mathcal{I}_z 's, it must use the same P for all \mathcal{I}_z 's, and the best it can do is to set $P = (4c)^z u_s$ with some probability ρ_z for each $z \in \lceil 4c \rceil$. Because $\sum_{z \in \lceil 4c \rceil} \rho_z = 1$, there exists z^* such that $\rho_{z^*} \leq \frac{1}{4c}$. Thus we have

$$\begin{aligned} \text{Rev}(\mathcal{M}(\mathcal{I}_{z^*})) &\leq \frac{1}{4c} \cdot (4c)^{z^*} \cdot u_s \cdot q_{s+1} + (1 - \frac{1}{4c})(4c)^{z^*-1} \cdot u_s \cdot q_{s+1} \\ &< \frac{1}{2c} \cdot (4c)^{z^*} \cdot u_s \cdot q_{s+1} = \frac{1}{2c} \text{OPT}(\mathcal{I}_{z^*}), \end{aligned}$$

where the first inequality is because for any threshold other than $(4c)^{z^*} u_s$, the resulting expected revenue is no larger than that with the threshold being $(4c)^{z^*-1} u_s$. That is, when \mathcal{M} cannot distinguish the \mathcal{I}_z 's, it cannot be a $2c$ -approximation for \mathcal{I}_{z^*} .

As the revenue of \mathcal{M} under \mathcal{I}_{z^*} is at most $OPT(\mathcal{I}_{z^*})$ when it is able to distinguish \mathcal{I}_{z^*} from all the other instances, we have

$$Rev(\mathcal{M}(\mathcal{I}_{z^*})) \leq (1 - \frac{1}{2c}) \frac{1}{2c} OPT(\mathcal{I}_{z^*}) + \frac{1}{2c} OPT(\mathcal{I}_{z^*}) < \frac{1}{c} OPT(\mathcal{I}_{z^*}).$$

Thus \mathcal{M} is not a c -approximation for \mathcal{I}_{z^*} , and Lemma 1 holds. \square

We extend this lemma to arbitrary multi-player multi-item Bayesian auctions with *succinct sub-additive* valuations, as follows. To make our exposition clearer, we first introduce some notations. A very broad class of Bayesian auctions, (*monotone*) *sub-additive* auctions, is such that each player i has a valuation function $v_i : 2^{[m]} \rightarrow \mathbb{R}$, which satisfies $v_i(S) + v_i(T) \geq v_i(S \cup T) \geq v_i(S) \geq 0$ for any subsets of items S and T . As such a valuation function in general needs 2^m values to describe, following the conventions in Bayesian auction design [34, 13, 10], we will consider *succinct sub-additive* auctions, where only the item-values, that is, the v_{ij} 's, are independently drawn from the underlying distribution $\mathcal{D} = \times_{i \in [n], j \in [m]} \mathcal{D}_{ij}$. Given $(v_{ij})_{j \in [m]}$, it is publicly known how to compute player i 's value for any subset of items. That is, the valuation function v_i now takes a vector of item-values $(v_{ij})_{j \in [m]}$ and a subset $S \subseteq [m]$ as inputs, such that for any vector $(v_{ij})_{j \in [m]}$, the resulting function $v_i((v_{ij})_{j \in [m]}, \cdot)$ is sub-additive and $v_i((v_{ij})_{j \in [m]}, \{j\}) = v_{ij}$ for each item j . Note that such auctions include single-item, unit-demand and additive auctions as special cases.

Theorem 1. *For any constant $c > 1$, there exists a constant C such that, for any $n \geq 1, m \geq 1$, for any large enough H , any monotone sub-additive valuation function profile $v = (v_i)_{i \in [n]}$, and any Bayesian mechanism \mathcal{M} making less than $Cnm \log_m \ln H$ adaptive value and quantile queries to the oracle, there exists a multi-item Bayesian auction instance $\mathcal{I} = (N, M, \mathcal{D})$, where $|N| = n, |M| = m$ and the values are bounded in $[1, H]$, such that $Rev(\mathcal{M}(\mathcal{I})) < \frac{OPT(\mathcal{I})}{c}$.*

Proof. Similar to the proof of Lemma 1, for any H , let $k \triangleq \lfloor \frac{1}{4} \log_{(4cx)}^{4c+2} H \rfloor$, where x is a parameter to be determined later. Let $C' \triangleq \frac{1}{24c(4c+2) \log_c(4cx)}$ and H be large enough so that $k \geq 1$. We divide the value interval $[1, H]$ into $k + 1$ sub-intervals as follows: from right to left, $u_k = H$, and $u_s = \frac{u_{s+1}}{(4cx)^{4c+2}}$ for each $s \in \{k-1, \dots, 0\}$. For each value interval (u_s, u_{s+1}) , there exists a corresponding quantile interval $(q_s, q_{s+1}) \triangleq (\frac{1}{u_s}, \frac{1}{u_{s+1}})$.

It is easy to see that $C'nm \log_c H < \frac{nmk}{3c}$. Thus, for any Bayesian mechanism \mathcal{M} making less than $C'nm \log_c H$ adaptive value and quantile queries, with probability at least $1 - \frac{1}{3c}$, there exists a player-item pair (i^*, j^*) , a value interval (u_s, u_{s+1}) with corresponding quantile interval (q_s, q_{s+1}) such that \mathcal{M} does not query these two intervals for $\mathcal{D}_{i^*j^*}$ and does not distinguish $\mathcal{D}_{i^*j^*;s}^z$'s from each other.

Therefore, for each $i \in [n], j \in [m], s \in [k]$, we construct $\lceil 4c \rceil$ Bayesian instances $\{\mathcal{I}^z = (N, M, \mathcal{D}^z)\}_{z \in \lceil [4c] \rceil}$, where each \mathcal{D}_{ij}^z is equal revenue distribution if $i \neq i^*$ or $j \neq j^*$. For any $z \in \lceil [4c] \rceil$, we construct the distribution $\mathcal{D}_{ij;s}^z$ such that for $v_z < u_s$ or $u_{s+1} < v_z < H$, the density of v_z is $\frac{1}{v_z}$. The probability of $v_z = H$ is $\frac{1}{H}$. The probability of $(4cx)^z u_s$ is $q_{s+1} - q_s$. All distributions other than \mathcal{D}_{ij} are equal revenue distribution. For any queries outside the interval (u_s, u_{s+1}) and (q_s, q_{s+1}) , it cannot distinguish those instances. Thus, with less than $Cnm \log_c H$ adaptive value and quantile queries, with probability at least $1 - \frac{1}{3c}$, the mechanism cannot distinguish those instances from each other.

We now analyze the optimal BIC revenue for those instances. For any \mathcal{I}_z , Myerson's mechanism is optimal: it sets a (randomized) threshold for the unique player, if the player bids at least the threshold then he gets the item and pays the threshold payment, otherwise the item is unsold. Letting $\delta \triangleq \frac{1}{H}$, it is not hard to verify that $OPT(\mathcal{I}_z) = (4cx)^z u_s q_s$ for each \mathcal{I}_z .

Given any succinct sub-additive valuation function profile $v = (v_i)_{i \in [n]}$ where each v_i takes a vector of item-values $(v_{ij})_{j \in [m]}$ as part of its input, we would like to compare the optimal revenue for the sub-additive instances defined by the \mathcal{I}^z 's with the corresponding expected revenue of \mathcal{M} . By construction, the \mathcal{D}^z 's differ only at the $\mathcal{D}_{i^*j^*}^z$'s, within the value interval (u_s, u_{s+1}) and the quantile interval (q_s, q_{s+1}) . Accordingly, with probability at least $1 - \frac{1}{3c}$, mechanism \mathcal{M} cannot distinguish the \mathcal{I}^z 's from each other. Eventually, we will analyze the revenue of \mathcal{M} conditional on this event happening.

For now, to compare the optimal revenue and that of \mathcal{M} , let us first introduce some notations. For any item-value profile $\hat{v} = (\hat{v}_{ij})_{i \in [n], j \in [m]}$, when the players bid \hat{v} , we denote by $x_i(\hat{v})$ the (randomized) allocation of \mathcal{M} to a player i . It is defined by the probabilities $\sigma_{iS}(\hat{v})$ for all the subsets $S \subseteq [m]$: each $\sigma_{iS}(\hat{v})$ is the probability that player i receives S under bid \hat{v} . Accordingly, the expected value of player i for allocation $x_i(\hat{v})$ is $v_i((\hat{v}_{ij})_{j \in [m]}, x_i(\hat{v})) = \sum_S v_i((\hat{v}_{ij})_{j \in [m]}, S) \cdot \sigma_{iS}(\hat{v})$. Moreover, for each item j , let $x_{ij}(\hat{v})$ be the probability that player i receives item j according to $x_i(\hat{v})$: that is, $x_{ij}(\hat{v}) = \sum_{S: j \in S} \sigma_{iS}(\hat{v})$.

We upper-bound the revenue of \mathcal{M} in three steps. To begin with, we reduce the multi-player sub-additive instances to *single-player* sub-additive instances, and construct a DSIC Bayesian mechanism \mathcal{M}^* that only sells the items to player i^* . Given any instance \mathcal{I}^z , mechanism \mathcal{M}^* runs on the single-player sub-additive instance $\mathcal{I}_{i^*}^z = (\{i^*\}, M, \mathcal{D}_{i^*}^z)$. It first simulates the item values of players in $N \setminus \{i^*\}$, which are all 1's, and then runs \mathcal{M} . Mechanism \mathcal{M}^* answers the oracle queries of \mathcal{M} truthfully. The allocation and the payment for player i^* under \mathcal{M}^* is the same as those under \mathcal{M} . For any player $i \neq i^*$, mechanism \mathcal{M}^* assigns nothing to him and charges him 0, because i is an imaginary player to \mathcal{M}^* . It is easy to see that mechanism \mathcal{M}^* is DSIC. Moreover,

$$Rev(\mathcal{M}^*(\mathcal{I}_{i^*}^z)) \geq Rev(\mathcal{M}(\mathcal{I}^z)) - \mathbb{E}_{\hat{v} \sim \mathcal{D}^z} \sum_{i \neq i^*} v_i((\hat{v}_{ij})_{j \in [m]}, x_i(\hat{v})), \quad (1)$$

because the revenue generated by \mathcal{M} from players in $N \setminus \{i^*\}$ is at most their total value for the allocation.

Next, we reduce the single-player sub-additive instances to single-player *additive* instances, and construct a DSIC Bayesian mechanism \mathcal{M}^+ that runs on the single-player additive instances $\mathcal{I}_{i^*}^{+z} = (\{i^*\}, M, \mathcal{D}_{i^*}^z)$, with $z \in \llbracket 4c \rrbracket$. Note that each $\mathcal{I}_{i^*}^{+z}$ has the same item-value distributions as $\mathcal{I}_{i^*}^z$, but player i^* 's value for any subset of items is additive.

For each single-player sub-additive instance defined by $\mathcal{I}_{i^*}^z$ and the valuation function profile v , by the taxation principle [23], mechanism \mathcal{M}^* is equivalent to providing a menu of options to player i^* and then letting i^* choose a menu entry maximizing his expected utility according to his true valuation. Given any instance $\mathcal{I}_{i^*}^{+z}$, mechanism \mathcal{M}^+ provides the same menu as mechanism \mathcal{M}^* under $\mathcal{I}_{i^*}^z$ and v , except that the payment in each entry is discounted by a multiplicative $1 - \hat{\epsilon}$. Here $\hat{\epsilon}$ is a sufficiently small constant in $(0, 1)$ to be determined later in the analysis. The truthfulness of \mathcal{M}^+ is immediate, because it lets i^* choose a menu entry maximizing his expected utility under his true additive values. Let

$$\bar{\delta} \triangleq \mathbb{E}_{\hat{v}_{i^*} \sim \mathcal{D}_{i^*}^z} \max_{S \subseteq [m]} \left(\sum_{j \in S} \hat{v}_{i^*j} - v_{i^*}((\hat{v}_{i^*j})_{j \in [m]}, S) \right),$$

the expected maximum difference between the additive values and the succinct sub-additive values. Following Lemma 3.4 in [34], which compares the revenue in the sub-additive instance with that in

the corresponding additive instance, we have

$$\text{Rev}(\mathcal{M}^+(\mathcal{I}_{i^*}^{+z})) \geq (1 - \hat{\epsilon})(\text{Rev}(\mathcal{M}^*(\mathcal{I}_{i^*}^z)) - \bar{\delta}/\hat{\epsilon}). \quad (2)$$

Finally, we reduce the single-player additive instances to single-player *single-item* instances, and construct a DSIC Bayesian mechanism \mathcal{M}' that only sells item j^* to player i^* . Mechanism \mathcal{M}' runs on the single-player single-item instances $\mathcal{I}_{i^*j^*}^z = (\{i^*\}, \{j^*\}, \mathcal{D}_{i^*j^*}^z)$, with $z \in [[4c]]$. Given any $\mathcal{I}_{i^*j^*}^z$, it first lets player i^* report $\hat{v}_{i^*j^*}$. Then it simulates the \hat{v}_{i^*j} 's from $\mathcal{D}_{i^*j}^z$ for $j \neq j^*$, which are all 1's, and runs \mathcal{M}^+ on the augmented additive instance $\mathcal{I}_{i^*}^{+z}$ to obtain allocation $x_{i^*}^+(\hat{v}_{i^*})$ and payment $p_{i^*}^+(\hat{v}_{i^*})$. For each item j , let $x_{i^*j}^+(\hat{v}_{i^*})$ be the probability that player i^* receives item j in the allocation. Mechanism \mathcal{M}' sets its outcome to be the following:

- $x'_{i^*j^*}(\hat{v}_{i^*j^*}) = x_{i^*j^*}^+(\hat{v}_{i^*})$; and
- $p'_{i^*}(\hat{v}_{i^*j^*}) = p_{i^*}^+(\hat{v}_{i^*}) - \sum_{j \in [m] \setminus \{j^*\}} \hat{v}_{i^*j} x_{i^*j}^+(\hat{v}_{i^*})$.

Note that $p'_{i^*}(\hat{v}_{i^*j^*})$ may be negative. By Lemma 21 of [25], mechanism \mathcal{M}' is DSIC and

$$\text{Rev}(\mathcal{M}'(\mathcal{I}_{i^*j^*}^z)) \geq \text{Rev}(\mathcal{M}^+(\mathcal{I}_{i^*}^{+z})) - \sum_{j \neq j^*} \mathbb{E}_{\hat{v}_{i^*j} \sim \mathcal{D}_{i^*j}^z} \hat{v}_{i^*j}. \quad (3)$$

Now we combine the above three reduction steps together and consider the event when mechanism \mathcal{M} cannot distinguish the \mathcal{I}^z 's from each other. When this happens, mechanism \mathcal{M} produces the same outcome for all the instances. Accordingly, although mechanism \mathcal{M}^* is given the distributions $\mathcal{D}_{i^*}^z$, by simulating \mathcal{M} , it still produces the same outcome for all the $\mathcal{I}_{i^*}^z$'s, thus the same menu for all of them. So mechanism \mathcal{M}^+ also produces the same menu for all the $\mathcal{I}_{i^*}^{+z}$'s: that is, the menu produced by \mathcal{M}^* with the payments discounted by $1 - \hat{\epsilon}$. As a result, although mechanism \mathcal{M}' is given the $\mathcal{D}_{i^*j^*}^z$'s, it still cannot “distinguish” the $\mathcal{I}_{i^*j^*}^z$'s from each other and produces the same outcome for all of them. Using the similar argument in the proof of Lemma 1, in this case there exists $z^* \in [[4c]]$ such that

$$\text{Rev}(\mathcal{M}'(\mathcal{I}_{i^*j^*}^{z^*})) < \frac{1}{2c} \text{OPT}(\mathcal{I}_{i^*j^*}^{z^*}).$$

Combining this inequality with Equations 1, 2 and 3, we have

$$\begin{aligned} \text{Rev}(\mathcal{M}(\mathcal{I}^{z^*})) &\leq \text{Rev}(\mathcal{M}^*(\mathcal{I}_{i^*}^{z^*})) + \mathbb{E}_{\hat{v} \sim \mathcal{D}^{z^*}} \sum_{i \neq i^*} v_i((\hat{v}_{ij})_{j \in [m]}, x_i(\hat{v})) \\ &\leq \frac{\text{Rev}(\mathcal{M}^+(\mathcal{I}_{i^*}^{+z^*}))}{1 - \hat{\epsilon}} + \bar{\delta}/\hat{\epsilon} + \mathbb{E}_{\hat{v} \sim \mathcal{D}^{z^*}} \sum_{i \neq i^*} v_i((\hat{v}_{ij})_{j \in [m]}, x_i(\hat{v})) \\ &\leq \frac{1}{1 - \hat{\epsilon}} \left(\text{Rev}(\mathcal{M}'(\mathcal{I}_{i^*j^*}^{z^*})) + \sum_{j \neq j^*} \mathbb{E}_{\hat{v}_{i^*j} \sim \mathcal{D}_{i^*j}^{z^*}} \hat{v}_{i^*j} \right) + \bar{\delta}/\hat{\epsilon} + \mathbb{E}_{\hat{v} \sim \mathcal{D}^{z^*}} \sum_{i \neq i^*} v_i((\hat{v}_{ij})_{j \in [m]}, x_i(\hat{v})) \\ &< \frac{1}{1 - \hat{\epsilon}} \left(\frac{1}{2c} \text{OPT}(\mathcal{I}_{i^*j^*}^{z^*}) + \sum_{j \neq j^*} \mathbb{E}_{\hat{v}_{i^*j} \sim \mathcal{D}_{i^*j}^{z^*}} \hat{v}_{i^*j} \right) + \bar{\delta}/\hat{\epsilon} + \mathbb{E}_{\hat{v} \sim \mathcal{D}^{z^*}} \sum_{i \neq i^*} v_i((\hat{v}_{ij})_{j \in [m]}, x_i(\hat{v})). \quad (4) \end{aligned}$$

Note that $\text{OPT}(\mathcal{I}_{i^*j^*}^{z^*}) \leq \text{OPT}(\mathcal{I}^{z^*})$, since selling a single item to a single player is a feasible outcome. Moreover, since $\mathcal{D}_{ij}^{z^*}$ is constantly 1 when $i \neq i^*$ or $j \neq j^*$, and since the valuation function

profile v is succinct sub-additive, we have

$$\begin{aligned} \sum_{j \neq j^*} \mathbb{E}_{\hat{v}_{i^*j} \sim \mathcal{D}_{i^*j}^{z^*}} \hat{v}_{i^*j} &= (m-1) \ln H, \\ \bar{\delta}/\hat{\epsilon} &= \frac{1}{\hat{\epsilon}} \mathbb{E}_{\hat{v}_{i^*} \sim \mathcal{D}_{i^*}^{z^*}} \max_{S \subseteq [m]} \left(\sum_{j \in S} \hat{v}_{i^*j} - v_{i^*}((\hat{v}_{i^*j})_{j \in [m]}, S) \right) \leq \frac{m \ln H}{\hat{\epsilon}}, \\ \mathbb{E}_{\hat{v} \sim \mathcal{D}^{z^*}} \sum_{i \neq i^*} v_i((\hat{v}_{ij})_{j \in [m]}, x_i(\hat{v})) &\leq m \ln H. \end{aligned}$$

Here the second equation is because $\sum_{j \in S} \hat{v}_{i^*j} - v_{i^*}((\hat{v}_{i^*j})_{j \in [m]}, S) \leq \sum_{j \in [m]} \hat{v}_{i^*j}$ for any \hat{v}_{i^*} and S . The third equation is because $\sum_{i \neq i^*} v_i((\hat{v}_{ij})_{j \in [m]}, x_i(\hat{v})) \leq \sum_j \sum_{i \neq i^*} x_{ij}(\hat{v}) \hat{v}_{ij}$ for any \hat{v} : indeed, each item can be sold to at most one player, generating expected value $\ln H$.

Combining the equations above with Equation 4, we have

$$\text{Rev}(\mathcal{M}(\mathcal{I}^{z^*})) < \frac{1}{1-\hat{\epsilon}} \left(\frac{1}{2c} \text{OPT}(\mathcal{I}^{z^*}) + (m-1) \ln H \right) + \frac{m \ln H}{\hat{\epsilon}} + m \ln H.$$

Setting $\hat{\epsilon} = \frac{1}{4}$, we have

$$\text{Rev}(\mathcal{M}(\mathcal{I}^{z^*})) < \frac{2}{3c} \text{OPT}(\mathcal{I}^{z^*}) + \frac{19m \ln H}{3}. \quad (5)$$

Now we combine Equation 5 with the probability that \mathcal{M} cannot distinguish the \mathcal{I}^z 's. As $\text{OPT}(\mathcal{I}^{z^*}) \geq 4cx$, when $x > \frac{57m \cdot c \cdot \ln H}{8}$, we have $\text{OPT}(\mathcal{I}^{z^*}) > \frac{57}{2} m \cdot c^2 \cdot \ln H$ and

$$\text{Rev}(\mathcal{M}(\mathcal{I}^{z^*})) \leq \left(1 - \frac{1}{3c}\right) \left(\frac{2}{3c} \text{OPT}(\mathcal{I}^{z^*}) + \frac{19m \ln H}{3}\right) + \frac{1}{3c} \text{OPT}(\mathcal{I}^{z^*}) < \frac{1}{c} \text{OPT}(\mathcal{I}^{z^*}).$$

Letting $C = \frac{1}{24c(4c+2)}$, which is a constant with respect to c , we have $Cnm \log_{m \ln H} H \leq C'nm \log_c H$. Thus finishes the proof of Theorem 1. \square

When there are multiple items to sell, for non-adaptive queries, we can make some changes to the constructed distributions to improve the query complexity bound. Formally, consider the $\lceil 4c \rceil$ Bayesian instances $\{\mathcal{I}^z = (N, M, \mathcal{D}^z)\}_{z \in \lceil [4c] \rceil}$ that are indistinguishable using only $Cnm \log_c H$ non-adaptive queries. Here for $i \neq i^*$ or $j \neq j^*$, \mathcal{D}_{ij}^z is the distribution that is constantly 1. For distribution $\mathcal{D}_{i^*j^*}^z$, we construct them by shifting them in the interval with no value or quantile queries, i.e. (u_s, u_{s+1}) and (q_s, q_{s+1}) , based on the equal revenue distribution with cumulative probability function $\max\{0, 1 - \frac{\sqrt{H}}{v}\}$. Note that the optimal revenue for those distributions is at least \sqrt{H} . Applying the same analysis as in Theorem 1, we have the following theorem, with proof omitted.

Theorem 2. *For any constant $c > 1$, there exists a constant C such that, for any $n \geq 1, m \geq 1$, any large enough H , any succinct sub-additive valuation function profile $v = (v_i)_{i \in [n]}$, and any DSIC Bayesian mechanism \mathcal{M} making less than $Cnm \log_c H$ non-adaptive value and quantile queries to the oracle, there exists a multi-item Bayesian auction instance $\mathcal{I} = (N, M, \mathcal{D})$ with valuation profile v , where $|N| = n, |M| = m$ and the item values are bounded in $[1, H]$, such that $\text{Rev}(\mathcal{M}(\mathcal{I})) < \frac{\text{OPT}(\mathcal{I})}{c}$.*

Succinct sub-additive valuations is a very broad class and contains single-item, unit-demand, and additive auctions as special cases. Thus Theorem 1 and 2 automatically apply to those cases. We also note that it is shown in [36] that the optimal BIC revenue exceeds the optimal DSIC revenue by a constant factor even for two i.i.d. additive bidders and two identical items. So even with infinite samples, there exist constants $c > 1$ such that no c -approximation to OPT is possible. However, Theorem 2 is stronger: for every constant $c > 1$, one needs at least the given number of queries to get a c -approximation.

4 The Query Complexity for Bounded Distributions

In this section, we consider settings where all distributions are bounded within $[1, H]$, and we construct efficient query mechanisms whose query complexity matches our lower-bounds. We show that it is sufficient to use only value queries, and we define in Section 4.1 a universal query scheme \mathcal{A}_V , which will be used as a black-box in our mechanisms. The seller uses algorithm \mathcal{A}_V to learn a distribution $\mathcal{D}' = \times_{i \in N, j \in M} \mathcal{D}'_{ij}$ that approximates the prior distribution \mathcal{D} and is stochastically dominated by \mathcal{D} . The seller then runs existing DSIC Bayesian mechanisms using \mathcal{D}' , while the players' values are drawn from \mathcal{D} . In this sense, all our mechanisms are simple, but they are not given a true Bayesian instance as input.

The multi-player single-item setting is already non-trivial, but still easy, since we have a good understanding of the optimal mechanism, which is Myerson's auction [31]. In particular, in the analysis it suffices to apply the *revenue monotonicity* theorem of [18]. The situation for unit-demand auctions and additive auctions is much more subtle. The optimal auction could be very complicated and may involve lotteries and bundling, and revenue monotonicity may not hold [27]. Even (disregarding complexity issues and) assuming we can design an optimal Bayesian mechanism for \mathcal{D}' , it is unclear how much revenue it guarantees when the players' values come from the true distribution \mathcal{D} . To overcome this difficulty, we rely on recent developments on *simple* Bayesian mechanisms with approximately optimal revenue.

The mechanism for unit-demand auctions is sequential post-price [29] and the analysis is relatively easy. For additive auctions, the Bayesian mechanism either runs Myerson's auction separately for each item or runs the VCG mechanism with a per-player entry fee [35, 9]. However, an easy and direct analysis would lose a factor of m in the query complexity. To achieve a tight upper-bound, we need to really open the box and analyze the mechanism differently in several crucial places, exploring its behavior under oracle queries.

To sum up, given our query scheme, our mechanisms are black-box reductions to simple Bayesian mechanisms, thus are simple, natural, and easy to implement in practice, while the analysis is non-black-box, non-trivial and reveals interesting connections between Bayesian mechanisms and query schemes.

4.1 The Value-Query Algorithm

The query algorithm \mathcal{A}_V is defined in Algorithm 1. Here $D \in \Delta(\mathbb{R})$ is the distribution to be queried. The algorithm takes two parameters, the value bound H and the precision factor $\delta > 0$, makes $O(\log_{1+\delta} H)$ value queries to the oracle, and then returns a discrete distribution D' . It is easy to verify that D' is stochastically dominated by D . Moreover, it is worth mentioning that the mean of D' approximates that of D . Indeed, $mean(D) = \int_1^H v dF(v) \leq \sum_{l=0}^{k-1} v_{l+1} Pr[v_l \leq v < v_{l+1}] + v_k D(v_k) \leq (1 + \delta) \sum_{l=0}^k v_l D'(v_l) = (1 + \delta) mean(D')$. Therefore, by directly applying the parametric mechanism in [1] with parameter $mean(D')$ (for single-parameter auctions where the distributions are regular or MHR), we will get at least a $(1 + \delta)$ fraction of their revenue.

4.2 Single-Item Auctions and Unit-Demand Auctions

Denoting by \mathcal{M}_{MRS} Myerson's mechanism for single-item auctions, Mechanism 2 defines our *efficient value Myerson* mechanism \mathcal{M}_{EVM} .

The query complexity of \mathcal{M}_{EVM} is $O(n \log_{1+\epsilon} H)$, since each distribution \mathcal{D}_i needs $O(\log_{1+\epsilon} H)$ value queries in \mathcal{A}_V . When ϵ is sufficiently small, $O(n \log_{1+\epsilon} H) \approx O(n\epsilon^{-1} \log H)$. Also, \mathcal{M}_{EVM} is DSIC since \mathcal{M}_{MRS} is so.

Algorithm 1 The Value-Query Algorithm \mathcal{A}_V

Input: The value bound H and the precision factor δ .

- 1: Let $k = \lceil \log_{1+\delta} H \rceil$ and define the *value vector* as $v = (v_0, v_1, \dots, v_{k-1}, v_k) = (1, (1 + \delta), (1 + \delta)^2, \dots, (1 + \delta)^{k-1}, H)$.
- 2: Query the oracle for D with v , and receive a non-increasing quantile vector $q = (q(v_0), \dots, q(v_k)) = (q_l)_{l \in \{0, \dots, k\}}$. Note $q_0 = 1$.
- 3: Construct a discrete distribution D' as follows: $D'(v_l) = q_l - q_{l+1}$ for any $l \in \{0, \dots, k\}$, where $q_{k+1} \triangleq 0$.

Output: Distribution D' .

Mechanism 2 Efficient Value Myerson Mechanism \mathcal{M}_{EVM}

- 1: Given the value bound H and a constant $\epsilon > 0$, run the value-query algorithm \mathcal{A}_V with H and $\delta = \epsilon$ for each player i 's distribution \mathcal{D}_i . Denote by \mathcal{D}'_i the returned distribution. Let $\mathcal{D}' = \times_{i \in N} \mathcal{D}'_i$.
 - 2: Run \mathcal{M}_{MRS} with \mathcal{D}' and the players' reported values, $b = (b_i)_{i \in N}$, to get allocation $x = (x_i)_{i \in N}$ and price profile $p = (p_i)_{i \in N}$ as the outcome.
-

In this section and throughout the paper, we often analyze “mismatching” cases where a Bayesian mechanism \mathcal{M} uses distribution \mathcal{D}' while the actual Bayesian instance is $\mathcal{I} = (N, M, \mathcal{D})$ (i.e., the players' true values are drawn from \mathcal{D}). We use $Rev(\mathcal{M}(\mathcal{I}; \mathcal{D}'))$ to denote the expected revenue in this case. By construction, $Rev(\mathcal{M}_{EVM}(\mathcal{I})) = Rev(\mathcal{M}_{MRS}(\mathcal{I}; \mathcal{D}'))$.

Because the distribution \mathcal{D}' constructed in \mathcal{M}_{EVM} is stochastically dominated by \mathcal{D} , letting $\mathcal{I}' = (N, M, \mathcal{D}')$ be the Bayesian instance under \mathcal{D}' , by revenue monotonicity [18] we have $Rev(\mathcal{M}_{MRS}(\mathcal{I}; \mathcal{D}')) \geq Rev(\mathcal{M}_{MRS}(\mathcal{I}'))$. By Lemma 5 of [18], $Rev(\mathcal{M}_{MRS}(\mathcal{I}')) \geq \frac{OPT(\mathcal{I})}{1+\epsilon}$. Thus we have the following simple result.

Theorem 3. $\forall \epsilon > 0$, for any single-item instance $\mathcal{I} = (N, M, \mathcal{D})$ with values in $[1, H]$, mechanism \mathcal{M}_{EVM} is DSIC, has query complexity $O(n \log_{1+\epsilon} H)$, and $Rev(\mathcal{M}_{EVM}(\mathcal{I})) \geq \frac{OPT(\mathcal{I})}{1+\epsilon}$.

The construction for unit-demand auctions is similar, except the seller uses as a blackbox the DSIC mechanism of [29], denoted by \mathcal{M}_{UD} : see Mechanism 3.

Mechanism 3 Mechanism \mathcal{M}_{EVUD} for Unit-Demand Auctions

- 1: Given H and $\epsilon > 0$, run the value-query algorithm \mathcal{A}_V with H and $\delta = \epsilon$ for each player i 's distribution \mathcal{D}_{ij} for each item j . Denote by \mathcal{D}'_{ij} the returned distribution. Let $\mathcal{D}'_i = \times_{j \in M} \mathcal{D}'_{ij}$ and $\mathcal{D}' = \times_{i \in N} \mathcal{D}'_i$.
 - 2: Run \mathcal{M}_{UD} with \mathcal{D}' and the players' reported values, $b = (b_{ij})_{i \in N, j \in M}$, to get allocation $x = (x_{ij})_{i \in N, j \in M}$ and price profile $p = (p_i)_{i \in N}$ as the outcome.
-

The main difficulty for unit-demand auctions is that we no longer have revenue monotonicity as in single-item auctions. Our analysis then comes in a non-blackbox way and relies on the *COPIES setting* [12, 29], which provides an upper-bound for the optimal BIC revenue. By properly upper-bounding the optimal revenue in the COPIES setting under \mathcal{D}' , we are able to upper-bound the optimal revenue in unit-demand auctions using the expected revenue of \mathcal{M}_{EVUD} . More precisely, we have the following theorem.

Theorem 4. $\forall \epsilon > 0$, for any unit-demand instance $\mathcal{I} = (N, M, \mathcal{D})$ with values in $[1, H]$, mechanism \mathcal{M}_{EVUD} is DSIC, has query complexity $O(mn \log_{1+\epsilon} H)$, and $Rev(\mathcal{M}_{EVUD}(\mathcal{I})) \geq \frac{OPT(\mathcal{I})}{24(1+\epsilon)}$.

Before analyzing mechanism \mathcal{M}_{EVUD} , let us first recall the *sequential post-price* mechanism \mathcal{M}_{UD} . This mechanism processes the players one by one according to an arbitrary order, computes a price for each player i based on remaining items, remaining players and the prior distribution, and lets i choose his utility-maximizing item (or choose none). The revenue of this mechanism is analyzed by reducing the unit-demand instance to the *COPIES* setting, which we introduce below.

For a unit-demand auction instance $\mathcal{I} = (N, M, \mathcal{D})$, the corresponding COPIES instance is denoted by $\mathcal{I}^{CP} = (N^{CP}, M^{CP}, \mathcal{D})$, where each player $i \in N$ has m copies and each item $j \in M$ has n copies, and player i 's copy j is only interested in item j 's copy i , with value v_{ij} drawn independently from \mathcal{D}_{ij} . Thus $N^{CP} = M^{CP} = N \times M$, and \mathcal{I}^{CP} is a single-parameter instance. Denote by N_i the set of player i 's copies and by M_j the set of item j 's copies. Note that both $\{N_i\}_{i \in N}$ and $\{M_j\}_{j \in M}$ are partitions of N^{CP} (and M^{CP}). Two natural constraints are imposed on feasible allocations under the COPIES setting, so as to connect it with the original unit-demand setting: (1) for each player i , at most one of his copies gets an item; and (2) for each item j , at most one of its copies gets allocated. Accordingly, letting q_s be the probability that a feasible mechanism allocates an item to a player copy $s \in N^{CP}$, we have $\sum_{s \in N_i} q_s \leq 1$ for each $i \in N$ and $\sum_{s \in M_j} q_s \leq 1$ for each $j \in M$.

The corresponding mechanism \mathcal{M}_{UD}^{CP} for the COPIES setting works in the same way as \mathcal{M}_{UD} , except that it considers an arbitrary order of the players in N^{CP} , thus different copies of the same player may not be processed together. When evaluating the performance of mechanism \mathcal{M}_{UD}^{CP} , the order of the players is chosen by an *online adaptive adversary*, who tries to minimize the expected revenue of the mechanism. Because this adversary is the worst-case for mechanism \mathcal{M}_{UD}^{CP} ,

$$Rev(\mathcal{M}_{UD}(\mathcal{I}; \mathcal{D}')) \geq Rev(\mathcal{M}_{UD}^{CP}(\mathcal{I}^{CP}; \mathcal{D}'))$$

for any distribution \mathcal{D}' , where the latter is the expected revenue of \mathcal{M}_{UD}^{CP} under the online adaptive adversary. Indeed, mechanism \mathcal{M}_{UD} can be considered as \mathcal{M}_{UD}^{CP} under a specific order where all copies of each player come together, thus the revenue is at least that when the order of N^{CP} is adaptively chosen by the adversary. Now we are ready to prove Theorem 4.

Proof of Theorem 4. It is easy to see that the query complexity of \mathcal{M}_{EVUD} is $O(mn \log_{1+\epsilon} H)$, since each distribution \mathcal{D}_{ij} needs $O(\log_{1+\epsilon} H)$ value queries. Also, it is immediate that \mathcal{M}_{EVUD} is DSIC.

Below we prove the revenue bound. By construction,

$$Rev(\mathcal{M}_{EVUD}(\mathcal{I})) = Rev(\mathcal{M}_{UD}(\mathcal{I}; \mathcal{D}')). \quad (6)$$

Let $\mathcal{I}' = (N, M, \mathcal{D}')$ and $\mathcal{I}'^{CP} = (N^{CP}, M^{CP}, \mathcal{D}')$. We state the following key lemma, which is proved after the proof of Theorem 4.

Lemma 2. $Rev(\mathcal{M}_{UD}(\mathcal{I}; \mathcal{D}')) \geq Rev(\mathcal{M}_{UD}^{CP}(\mathcal{I}^{CP}; \mathcal{D}')) = Rev(\mathcal{M}_{UD}^{CP}(\mathcal{I}'^{CP}))$.

By Theorem 1 of [29], the sequential post-price mechanism is at least a 6-approximation to the optimal BIC revenue in the COPIES setting. Thus

$$Rev(\mathcal{M}_{UD}^{CP}(\mathcal{I}'^{CP})) \geq \frac{1}{6} OPT(\mathcal{I}'^{CP}). \quad (7)$$

Next, because the COPIES setting is a single-parameter setting, and because of the way we discretize the value space in algorithm \mathcal{A}_V , by Lemma 5 of [18] we have

$$OPT(\mathcal{I}'^{CP}) \geq \frac{1}{1+\epsilon} OPT(\mathcal{I}^{CP}). \quad (8)$$

Finally, by Theorem 6 of [9], the optimal BIC revenue in the COPIES setting is a 4-approximation to the optimal BIC revenue in the original unit-demand setting. Thus

$$OPT(\mathcal{I}^{CP}) \geq \frac{1}{4}OPT(\mathcal{I}). \quad (9)$$

Combining Equations 6, 7, 8, 9 and Lemma 2, Theorem 4 holds. \square

Proof of Lemma 2. The inequality is already explained. Now we prove the equality. For any value profile $v \sim \mathcal{D}$, let v' be v rounded down to the support of \mathcal{D}' . That is, for each v_{ij} , v'_{ij} is the largest value in the support of \mathcal{D}'_{ij} that is less than or equal to v_{ij} . Recall that the support of \mathcal{D}'_{ij} is the set $\{v_0, \dots, v_k\}$ as defined in the query algorithm \mathcal{A}_V . By the definition of \mathcal{D}'_{ij} , for any $0 \leq l \leq k-1$,

$$\Pr_{v_{ij} \sim \mathcal{D}'_{ij}} [v'_{ij} = v_l] = \Pr_{v_{ij} \sim \mathcal{D}_{ij}} [v_{ij} \geq v_l] - \Pr_{v_{ij} \sim \mathcal{D}_{ij}} [v_{ij} \geq v_{l+1}] = q(v_l) - q(v_{l+1}) = q_l - q_{l+1} = \mathcal{D}'_{ij}(v_l),$$

and

$$\Pr_{v_{ij} \sim \mathcal{D}'_{ij}} [v'_{ij} = v_k] = \Pr_{v_{ij} \sim \mathcal{D}_{ij}} [v_{ij} \geq v_k] = q(v_k) = q_k = \mathcal{D}'_{ij}(v_k).$$

That is, if v is distributed according to \mathcal{D} then v' is distributed according to \mathcal{D}' .

For any value profile v and the corresponding v' , arbitrarily fix an order σ of the players in N^{CP} , which is a bijection from $\{1, \dots, mn\}$ to $\{1, \dots, mn\}$. Without loss of generality, each player $\sigma(s)$ gets the corresponding item $\sigma(s)$ whenever his true value is greater than or equal to the posted price for him. Below we show that mechanism \mathcal{M}_{UD}^{CP} produces the same outcome no matter the players' true values are v or v' . That is, for any $s \in \{1, \dots, mn\}$, (1) \mathcal{M}_{UD}^{CP} produces the same price $p_{\sigma(s)}$ under v and v' for player $\sigma(s)$, and (2) $v_{\sigma(s)} \geq p_{\sigma(s)}$ if and only if $v'_{\sigma(s)} \geq p_{\sigma(s)}$.

To prove these two properties, note that by the construction of mechanism \mathcal{M}_{UD}^{CP} , the price $p_{\sigma(s)}$ posted to $\sigma(s)$ depends only on the distribution \mathcal{D}' and the set $A_{\sigma(s)}$ of items sold to the players arriving before $\sigma(s)$. Here $p_{\sigma(s)}$ may be randomized if $\mathcal{D}'_{\sigma(s)}$ is irregular, but it always takes value in the support of $\mathcal{D}'_{\sigma(s)}$ (except that, if selling the corresponding item $\sigma(s)$ to player $\sigma(s)$ is not feasible anymore, then $p_{\sigma(s)} = +\infty$).

We prove the two desired properties by induction. When $s = 1$, property (1) trivially holds, because $A_{\sigma(1)} = \emptyset$ under both value profiles. Furthermore, because a realization of $p_{\sigma(1)}$ is always in the support of $\mathcal{D}'_{\sigma(1)}$, and because $v'_{\sigma(1)}$ is $v_{\sigma(1)}$ rounded down to the support of $\mathcal{D}'_{\sigma(1)}$, property (2) holds when $s = 1$.

Now assume (1) and (2) hold for any $s \leq t$ with $t < mn$. We show they also hold for $s = t + 1$. Indeed, the inductive hypothesis implies that for any $s \leq t$, $A_{\sigma(s)}$ is the same under the two value profiles. In particular, $A_{\sigma(t+1)}$ is the same, which means the price $p_{\sigma(t+1)}$ is the same. Thus property (1) holds. Property (2) also holds because a realization of $p_{\sigma(t+1)}$ is always in the support of $\mathcal{D}'_{\sigma(t+1)}$. In sum, for any order σ , mechanism \mathcal{M}_{UD}^{CP} produces the same outcome under the two value profiles v and v' , thus the same revenue.

Accordingly, under the online adaptive adversary for $(\mathcal{I}^{CP}; \mathcal{D}')$, the revenue $Rev(\mathcal{M}_{UD}^{CP}(\mathcal{I}^{CP}; \mathcal{D}'))$ is the same as the revenue when the players' true values are obtained by rounding $v \sim \mathcal{D}$ to v' . Because the resulting v' is distributed according to \mathcal{D}' , $Rev(\mathcal{M}_{UD}^{CP}(\mathcal{I}^{CP}; \mathcal{D}'))$ is at least the expected revenue of \mathcal{M}_{UD}^{CP} under the online adaptive adversary for \mathcal{I}^{CP} . Indeed, a randomized adversary for \mathcal{I}^{CP} can simulate the adversary for $(\mathcal{I}^{CP}; \mathcal{D}')$: in each step, given v'_s with $s \in N^{CP}$ being the player in this step, the former first samples v_s from \mathcal{D}_s conditional on v_s rounded down to v'_s , and then uses the latter to decide which player arrives next. Thus,

$$Rev(\mathcal{M}_{UD}^{CP}(\mathcal{I}^{CP}; \mathcal{D}')) \geq Rev(\mathcal{M}_{UD}^{CP}(\mathcal{I}^{CP})).$$

Similarly,

$$\text{Rev}(\mathcal{M}_{UD}^{CP}(\mathcal{I}^{CP}; \mathcal{D}')) \leq \text{Rev}(\mathcal{M}_{UD}^{CP}(\mathcal{I}'^{CP})).$$

Therefore $\text{Rev}(\mathcal{M}_{UD}^{CP}(\mathcal{I}^{CP}; \mathcal{D}')) = \text{Rev}(\mathcal{M}_{UD}^{CP}(\mathcal{I}'^{CP}))$ and Lemma 2 holds. \square

Letting $c = 24(1 + \epsilon)$, we have the query complexity in Table 1.

4.3 Additive Auctions

For additive auctions, the DSIC Bayesian mechanism in [35, 9] chooses between two mechanisms, whichever generates higher expected revenue under the true prior \mathcal{D} . The first is the ‘‘individual Myerson’’ mechanism, denoted by \mathcal{M}_{IM} , which sells each item separately using Myerson’s mechanism. The second is the VCG mechanism with optimal per-player entry fees, denoted by \mathcal{M}_{BVCG} .

In our mechanism \mathcal{M}_{EVA} , the seller queries about \mathcal{D} using algorithm \mathcal{A}_V with properly chosen parameters. Given the resulting distribution \mathcal{D}' , the seller either runs \mathcal{M}_{IM} or runs \mathcal{M}_{BVCG} as a blackbox, resulting in query mechanisms \mathcal{M}_{EVIM} and \mathcal{M}_{EVBVCG} . We only define the latter in Mechanism 4, and the former simply replaces \mathcal{M}_{BVCG} with \mathcal{M}_{IM} . Note that $\text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I})) = \text{Rev}(\mathcal{M}_{IM}(\mathcal{I}; \mathcal{D}'))$ and $\text{Rev}(\mathcal{M}_{EVBVCG}(\mathcal{I})) = \text{Rev}(\mathcal{M}_{BVCG}(\mathcal{I}; \mathcal{D}'))$. However, the seller cannot compute these two revenue and choose the better one, because he does not know \mathcal{D} . Thus he randomly chooses between the two, according to probabilities defined in our analysis, to optimize the approximation ratio. We have the following theorem.

Theorem 5 is harder to show. Indeed, one cannot use revenue monotonicity or the COPIES setting to easily upper-bound the optimal BIC revenue. Our analysis is based on the duality framework of [9] for Bayesian auctions, properly adapted for the query setting. Finally, letting $c = 8(1 + \epsilon)$, we have the query complexity in Table 1.

Mechanism 4 Mechanism \mathcal{M}_{EVBVCG} to Approximate \mathcal{M}_{BVCG} via Value Queries

- 1: Given H and $\epsilon > 0$, run the value-query algorithm \mathcal{A}_V with H and $\delta = \sqrt{\epsilon + 1} - 1$ for each player i ’s distribution \mathcal{D}_{ij} for each item j . Denote by \mathcal{D}'_{ij} the returned distribution. Let $\mathcal{D}'_i = \times_{j \in M} \mathcal{D}'_{ij}$ and $\mathcal{D}' = \times_{i \in N} \mathcal{D}'_i$.
 - 2: Run \mathcal{M}_{BVCG} with \mathcal{D}' and the players’ reported values, $b = (b_{ij})_{i \in N, j \in M}$, to get allocation $x = (x_{ij})_{i \in N, j \in M}$ and price profile $p = (p_i)_{i \in N}$ as the outcome.
-

Theorem 5. $\forall \epsilon > 0$, for any additive instance $\mathcal{I} = (N, M, \mathcal{D})$ with values in $[1, H]$, mechanism \mathcal{M}_{EVA} is DSIC, has query complexity $O(mn \log_{1+\epsilon} H)$, and $\text{Rev}(\mathcal{M}_{EVA}(\mathcal{I})) \geq \frac{\text{OPT}(\mathcal{I})}{8(1+\epsilon)}$.

Proof. First, it is easy to see that the query complexity of mechanism \mathcal{M}_{EVA} is $O(mn \log_{1+\delta} H)$, since there are in total mn distributions and each one of them needs $O(\log_{1+\delta} H)$ value queries in the algorithm \mathcal{A}_V . Since $\delta = \sqrt{\epsilon + 1} - 1$, $O(mn \log_{1+\delta} H) = O(mn \log_{1+\epsilon} H)$. Second, since mechanisms \mathcal{M}_{BVCG} and \mathcal{M}_{IM} are both DSIC, \mathcal{M}_{EVA} is DSIC.

Recall that mechanism \mathcal{M}_{EVA} randomly chooses between running \mathcal{M}_{EVIM} and running \mathcal{M}_{EVBVCG} . Therefore, to upper-bound the optimal revenue $\text{OPT}(\mathcal{I})$ using $\text{Rev}(\mathcal{M}_{EVA}(\mathcal{I}))$, we only need to upper-bound $\text{OPT}(\mathcal{I})$ using $\text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I}))$ and $\text{Rev}(\mathcal{M}_{EVBVCG}(\mathcal{I}))$.

As in [9], we only need to consider the prior distribution \mathcal{D} with finite support. Let V_{ij} be the support of \mathcal{D}_{ij} for each player i and item j , $V_i = \times_{j \in M} V_{ij}$ and $V = \times_{i \in N} V_i$. In the optimal BIC mechanism, when player i bids v_i , let $\pi_{ij}(v_i)$ be the probability for him to get item j and $p_i(v_i)$ be his expected payment, taken over the randomness of the other players’ values and the randomness of the mechanism. Let $\pi = (\pi_{ij}(v_i))_{i \in N, j \in M, v_i \in V_i}$ and $p = (p_i(v_i))_{i \in N, v_i \in V_i}$. The pair (π, p) is called the *reduced form* (of the optimal BIC mechanism) [7].

Denote by $\tilde{\varphi}_{ij}(v_{ij})$ Myerson's (ironed) virtual value when player i 's value on item j is v_{ij} . For any value sub-profile v_{-i} of the players other than i , let $\beta_{ij}(v_{-i}) = \max_{i' \neq i} v_{i'j}$: that is, the highest bid on item j excluding player i . Moreover, let $r_{ij}(v_{-i}) = \max_{x \geq \beta_{ij}(v_{-i})} \{x \cdot \Pr_{v_{ij} \sim \mathcal{D}_{ij}}[v_{ij} \geq x]\}$, $r_i(v_{-i}) = \sum_j r_{ij}(v_{-i})$, $r_i = \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}}[r_i(v_{-i})]$, and finally $r = \sum_i r_i$. Note that r is the expected revenue by running the 1-look-ahead mechanism of [32] for each item separately, and $r \leq \text{Rev}(\mathcal{M}_{IM}(\mathcal{I}))$.

Next, we use a different method from [9] to partition each player i 's value space V_i into $m + 1$ subsets. More precisely, given $\delta > 0$ and v_{-i} , let $R_0^{(v_{-i})} = \{v_i \in V_i \mid v_{ij} < (1 + \delta)\beta_{ij}(v_{-i}), \forall j\}$. For any $v_i \notin R_0^{(v_{-i})}$, let $j = \arg \max\{v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})\}$ with ties broken lexicographically, and add v_i to the set $R_j^{(v_{-i})}$: note that $v_{ij} - (1 + \delta)\beta_{ij}(v_{-i}) \geq 0$ in this case. Similar to Theorem 3 of [9], the optimal BIC revenue can be upper-bounded by the sum of the following terms, where $\mathcal{D}_i(v_i)$ and $\mathcal{D}_{-i}(v_{-i})$ are respectively the probabilities of v_i and v_{-i} under \mathcal{D} , and \mathbf{I} is the indicator function:

$$\text{OPT}(\mathcal{I}) \leq \text{Single} + \text{Under} + \text{Over} + \text{Tail} + \text{Core}, \quad (10)$$

where

$$\text{Single} = \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \tilde{\varphi}_{ij}(v_{ij}) \cdot \Pr_{v_{-i} \sim \mathcal{D}_{-i}}[v_i \in R_j^{(v_{-i})}],$$

$$\text{Under} = \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \sum_{v_{-i} \in V_{-i}} v_{ij} \cdot \mathcal{D}_{-i}(v_{-i}) \mathbf{I}_{v_{ij} < (1+\delta)\beta_{ij}(v_{-i})},$$

$$\text{Over} = \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \sum_{v_{-i} \in V_{-i}} (1 + \delta)\beta_{ij}(v_{-i}) \mathcal{D}_{-i}(v_{-i}) \mathbf{I}_{v_{ij} \geq (1+\delta)\beta_{ij}(v_{-i})},$$

$$\begin{aligned} \text{Tail} = & \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j \sum_{v_{ij} > (1+\delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})} \mathcal{D}_{ij}(v_{ij}) \cdot (v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})) \\ & \cdot \Pr_{v_{i,-j} \sim \mathcal{D}_{i,-j}}[\exists k \neq j, v_{ik} - (1 + \delta)\beta_{ik}(v_{-i}) \geq v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})], \end{aligned}$$

and

$$\begin{aligned} \text{Core} = & \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j \sum_{(1+\delta)\beta_{ij}(v_{-i}) \leq v_{ij} \leq (1+\delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})} \mathcal{D}_{ij}(v_{ij}) \\ & \cdot (v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})). \end{aligned}$$

In the following, we bound these terms in Inequality 10 separately. Note that when \mathcal{M}_{EVIM} uses the value-query algorithm \mathcal{A}_V to learn a distribution, the parameters are also set to be H and $\delta = \sqrt{\epsilon + 1} - 1$. Thus, applying Theorem 3 to each item, we have

$$\text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) \leq (1 + \delta)\text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I})).$$

For the terms Single, Under, Over and Tail, we are able to upper-bound them using $\text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I}))$. Following Lemma 13 of [9], although the term Single has changed from its original form, we still have

$$\begin{aligned} \text{Single} &= \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \tilde{\varphi}_{ij}(v_{ij}) \cdot \Pr_{v_{-i} \sim \mathcal{D}_{-i}}[v_i \in R_j^{(v_{-i})}] \\ &\leq \text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) \leq (1 + \delta)\text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I})). \end{aligned} \quad (11)$$

Next, using Lemmas 14 and 15 of [9], we upper-bound the term Under as follows:

$$\begin{aligned}
\text{Under} &= \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \cdot v_{ij} \cdot \mathbf{I}_{v_{ij} < (1+\delta)\beta_{ij}(v_{-i})} \\
&= \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \cdot v_{ij} \\
&\quad \cdot (\mathbf{I}_{v_{ij} < \beta_{ij}(v_{-i})} + \mathbf{I}_{\beta_{ij}(v_{-i}) \leq v_{ij} < (1+\delta)\beta_{ij}(v_{-i})}) \\
&\leq \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \\
&\quad \cdot (v_{ij} \cdot \mathbf{I}_{v_{ij} < \beta_{ij}(v_{-i})} + (1 + \delta)\beta_{ij}(v_{-i}) \cdot \mathbf{I}_{v_{ij} \geq \beta_{ij}(v_{-i})}) \\
&\leq \text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) + (1 + \delta)\text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) \leq 2(1 + \delta)^2 \text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I})).
\end{aligned}$$

The second inequality above is by Lemmas 14 and 15 of [9], which respectively upper-bound the term Over and the term Under in the original setting. Indeed, we split our term Under into the sum of the original terms Under and Over. Using the above equation, the approximation ratio to $OPT(\mathcal{I})$ will be $9(1 + \epsilon)$ eventually. To get the desired $8(1 + \epsilon)$ -approximation, we prove a variant of Lemma 15 of [9], which directly upper-bounds our term Under as

$$\text{Under} \leq (1 + \delta)\text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) \leq (1 + \delta)^2 \text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I})). \quad (12)$$

The actual proof of this alternative lemma is tedious and does not provide new insights to our result, thus has been omitted.

Next, we upper-bound the term Over:

$$\begin{aligned}
\text{Over} &= \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \sum_{v_{-i} \in V_{-i}} (1 + \delta)\beta_{ij}(v_{-i})\mathcal{D}_{-i}(v_{-i})\mathbf{I}_{v_{ij} \geq (1+\delta)\beta_{ij}(v_{-i})} \\
&\leq (1 + \delta) \sum_i \sum_{v_i \in V_i} \sum_j \mathcal{D}_i(v_i) \cdot \pi_{ij}(v_i) \cdot \sum_{v_{-i} \in V_{-i}} \beta_{ij}(v_{-i})\mathcal{D}_{-i}(v_{-i})\mathbf{I}_{v_{ij} \geq \beta_{ij}(v_{-i})} \\
&\leq (1 + \delta)\text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) \leq (1 + \delta)^2 \text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I})). \quad (13)
\end{aligned}$$

The second inequality above is by Lemma 14 of [9].

Next, we upper-bound the term Tail, which is similar to the analysis of [9], but with the threshold

price $\beta_{ij}(v_{-i})$ scaled up by a factor of $(1 + \delta)$.

$$\begin{aligned}
\text{Tail} &= \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j \sum_{v_{ij} > (1+\delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})} \mathcal{D}_{ij}(v_{ij}) \cdot (v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})) \\
&\quad \cdot \Pr_{v_{i,-j} \sim \mathcal{D}_{i,-j}} [\exists k \neq j, v_{ik} - (1 + \delta)\beta_{ik}(v_{-i}) \geq v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})] \\
&\leq \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j \sum_{v_{ij} > (1+\delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})} \mathcal{D}_{ij}(v_{ij}) \cdot (v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})) \\
&\quad \cdot \Pr_{v_{i,-j} \sim \mathcal{D}_{i,-j}} [\exists k \neq j, v_{ik} - \beta_{ik}(v_{-i}) \geq v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})] \\
&\leq \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j \sum_{v_{ij} > (1+\delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})} \mathcal{D}_{ij}(v_{ij}) \cdot (v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})) \\
&\quad \cdot \sum_{k=1}^m \Pr_{v_{ik} \sim \mathcal{D}_{ik}} [v_{ik} \geq v_{ij} - (1 + \delta)\beta_{ij}(v_{-i}) + \beta_{ik}(v_{-i})] \\
&\leq \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j \sum_{v_{ij} > (1+\delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})} \mathcal{D}_{ij}(v_{ij}) \\
&\quad \cdot \sum_{k=1}^m (v_{ij} - (1 + \delta)\beta_{ij}(v_{-i}) + \beta_{ik}(v_{-i})) \Pr_{v_{ik} \sim \mathcal{D}_{ik}} [v_{ik} \geq v_{ij} - (1 + \delta)\beta_{ij}(v_{-i}) + \beta_{ik}(v_{-i})] \\
&\leq \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j \sum_{v_{ij} > (1+\delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})} \mathcal{D}_{ij}(v_{ij}) \sum_{k=1}^m r_{ik}(v_{-i}) \\
&= \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j r_i(v_{-i}) \sum_{v_{ij} > (1+\delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})} \mathcal{D}_{ij}(v_{ij}) \\
&\leq \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j ((1 + \delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})) \\
&\quad \cdot \Pr_{v_{ij} \sim \mathcal{D}_{ij}} [v_{ij} > (1 + \delta)\beta_{ij}(v_{-i}) + r_i(v_{-i})] \\
&\leq \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j r_{ij}(v_{-i}) = \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) r_i(v_{-i}) = \sum_i r_i \\
&= r \leq \text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) \leq (1 + \delta) \text{Rev}(\mathcal{M}_{EVIM}(\mathcal{I})). \tag{14}
\end{aligned}$$

The second inequality above is by union bound. The fourth and sixth inequalities use twice the definition of $r_{ij}(v_{-i})$, which sets the optimal price to maximize the expected revenue generated by selling item j to i . The second equality is by the definition of $r_i(v_{-i})$.

Finally, we upper-bound the term Core. The Core part is the most complicated, and we use \mathcal{M}_{EVBVCG} and \mathcal{M}_{EVIM} together to upper-bound it. To do so, below we rewrite Core into a different form. Similar to [9], arbitrarily fixing v_{-i} and letting $v_{ij} \sim \mathcal{D}_{ij}$, define the following two new random variables, which again scale the threshold price $\beta_{ij}(v_{-i})$ up by a factor of $(1 + \delta)$:

$$b_{ij}(v_{-i}) = (v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})) \mathbf{I}_{v_{ij} \geq (1+\delta)\beta_{ij}(v_{-i})},$$

and

$$c_{ij}(v_{-i}) = b_{ij}(v_{-i}) \mathbf{I}_{b_{ij}(v_{-i}) \leq r_i(v_{-i})}.$$

Therefore, we have

$$\text{Core} = \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \sum_j \mathbb{E}_{v_{ij} \sim \mathcal{D}_{ij}} [c_{ij}(v_{-i})].$$

Letting $e_i(v_{-i}) = \sum_j \mathbb{E}_{v_{ij} \sim \mathcal{D}_{ij}} [c_{ij}(v_{-i})] - 2r_i(v_{-i})$, following the proof of Lemma 12 in [9], we still have

$$\Pr\left[\sum_j b_{ij}(v_{-i}) \geq e_i(v_{-i})\right] \geq \frac{1}{2}.$$

In the following, we use the revenue of mechanisms \mathcal{M}_{EVBVCG} and \mathcal{M}_{EVIM} to bound the Core. To do so, first note that by the construction of mechanism \mathcal{M}_{EVBVCG} ,

$$\text{Rev}(\mathcal{M}_{EVBVCG}(\mathcal{I})) = \text{Rev}(\mathcal{M}_{BVCG}(\mathcal{I}; \mathcal{D}')).$$

Let V'_{ij} be the support of \mathcal{D}'_{ij} , $V'_i = \times_{j \in M} V'_{ij}$, $V' = \times_{i \in N} V'_i$. As before, given $v_i \sim \mathcal{D}_i$, denote by $v'_i \in V'_i$ the value vector obtained by rounding v_i down to the support of \mathcal{D}'_i . That is, each v'_{ij} is the largest value in V'_{ij} that is less than or equal to v_{ij} . Then,

$$\begin{aligned} \text{Rev}(\mathcal{M}_{BVCG}(\mathcal{I}; \mathcal{D}')) &\geq \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \text{Rev}(\mathcal{M}_{BVCG}(v'_i, v_{-i}; \mathcal{D}')) \\ &= \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v'_i \sim \mathcal{D}'_i} \text{Rev}(\mathcal{M}_{BVCG}(v'_i, v_{-i}; \mathcal{D}')). \end{aligned}$$

The inequality is because each player i can potentially buy item j only when j is in his winning set (i.e., he is the highest bidder for j), and i 's winning set under v'_i is a subset of his winning set under v_i . Moreover, the entry fee of i is the same under both (v_i, v_{-i}) and (v'_i, v_{-i}) , as it only depends on \mathcal{D}'_i and v_{-i} . Thus the revenue inside the expectation does not increase when v_i is replaced by v'_i . The equality is again because drawing v_i from \mathcal{D}_i and then rounding down to v'_i is equivalent to drawing v'_i from \mathcal{D}'_i directly.

Next, we lower-bound $\sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v'_i \sim \mathcal{D}'_i} \text{Rev}(\mathcal{M}_{BVCG}(v'_i, v_{-i}; \mathcal{D}'))$. As before, arbitrarily fixing v_{-i} and letting $v'_{ij} \sim \mathcal{D}'_{ij}$, define

$$b'_{ij}(v_{-i}) = (v'_{ij} - \beta_{ij}(v_{-i})) \mathbf{I}_{v'_{ij} \geq \beta_{ij}(v_{-i})}.$$

Note that $b'_{ij}(v_{-i})$ is a random variable that represents player i 's utility in the second price mechanism on item j with value $v'_{ij} \sim \mathcal{D}'_{ij}$, when the other players' bids are $v_{-i,j}$. Also note that \mathcal{M}_{BVCG} uses the optimal entry fee for each i with respect to v_{-i} and \mathcal{D}' , which generates expected revenue from i (over \mathcal{D}'_i) greater than or equal to that by using the following entry fee,

$$e'_i(v_{-i}) = \frac{e_i(v_{-i})}{1 + \delta}.$$

Now we show player i accepts the entry fee $e'_i(v_{-i})$ with probability at least $\frac{1}{2}$. Indeed, for any v_i and the corresponding v'_i ,

$$\begin{aligned} \sum_j b'_{ij}(v_{-i}) &= \sum_j (v'_{ij} - \beta_{ij}(v_{-i})) \mathbf{I}_{v'_{ij} \geq \beta_{ij}(v_{-i})} \geq \sum_j \left(\frac{v_{ij}}{1 + \delta} - \beta_{ij}(v_{-i})\right) \mathbf{I}_{\frac{v_{ij}}{1 + \delta} \geq \beta_{ij}(v_{-i})} \\ &= \frac{1}{1 + \delta} \sum_j (v_{ij} - (1 + \delta)\beta_{ij}(v_{-i})) \mathbf{I}_{v_{ij} \geq (1 + \delta)\beta_{ij}(v_{-i})} = \frac{1}{1 + \delta} \sum_j b_{ij}(v_{-i}). \end{aligned}$$

The inequality is because $v'_{ij} \geq \frac{v_{ij}}{1 + \delta}$, and because $\frac{v_{ij}}{1 + \delta} \geq \beta_{ij}(v_{-i})$ implies $v'_{ij} \geq \beta_{ij}(v_{-i})$. Therefore

$$\begin{aligned} \Pr_{v'_i \sim \mathcal{D}'_i} \left[\sum_j b'_{ij}(v_{-i}) \geq e'_i(v_{-i}) \right] &\geq \Pr_{v_i \sim \mathcal{D}_i} \left[\frac{1}{1 + \delta} \sum_j b_{ij}(v_{-i}) \geq \frac{e_i(v_{-i})}{1 + \delta} \right] \\ &= \Pr_{v_i \sim \mathcal{D}_i} \left[\sum_j b_{ij}(v_{-i}) \geq e_i(v_{-i}) \right] \geq \frac{1}{2}, \end{aligned}$$

as desired. Thus we have

$$\begin{aligned}
& Rev(\mathcal{M}_{EVBVCG}(\mathcal{I})) \\
& \geq \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v'_i \sim \mathcal{D}'_i} Rev(\mathcal{M}_{BVCG}(v'_i, v_{-i}; \mathcal{D}')) \\
& \geq \frac{1}{2} \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \cdot \frac{e_i(v_{-i})}{1 + \delta} \\
& = \frac{1}{2(1 + \delta)} \sum_i \sum_{v_{-i} \in V_{-i}} \mathcal{D}_{-i}(v_{-i}) \left(\sum_j \mathbb{E}_{v_{ij} \sim \mathcal{D}_{ij}} [c_{ij}(v_{-i})] - 2r_i(v_{-i}) \right) \\
& = \frac{1}{2(1 + \delta)} \text{Core} - \frac{r}{1 + \delta}.
\end{aligned}$$

That is,

$$\begin{aligned}
\text{Core} & \leq 2(1 + \delta) Rev(\mathcal{M}_{EVBVCG}(\mathcal{I})) + 2r \\
& \leq 2(1 + \delta) [Rev(\mathcal{M}_{EVBVCG}(\mathcal{I})) + Rev(\mathcal{M}_{EVIM}(\mathcal{I}))].
\end{aligned} \tag{15}$$

Combining Inequalities 10, 11, 12, 13, 14 and 15,

$$\begin{aligned}
OPT(\mathcal{I}) & \leq (1 + \delta)^2 (2Rev(\mathcal{M}_{EVBVCG}(\mathcal{I})) + 6Rev(\mathcal{M}_{EVIM}(\mathcal{I}))) \\
& = (1 + \epsilon) (2Rev(\mathcal{M}_{EVBVCG}(\mathcal{I})) + 6Rev(\mathcal{M}_{EVIM}(\mathcal{I}))).
\end{aligned}$$

Accordingly, by running mechanism \mathcal{M}_{EVBVCG} with probability $\frac{1}{4}$ and mechanism \mathcal{M}_{EVIM} with probability $\frac{3}{4}$, the expected revenue of mechanism \mathcal{M}_{EVA} is

$$Rev(\mathcal{M}_{EVA}(\mathcal{I})) \geq \frac{1}{8(1 + \epsilon)} OPT(\mathcal{I}).$$

This finishes the proof of Theorem 5. □

5 The Query Complexity for Unbounded Distributions

Next, we construct efficient query mechanisms for arbitrary distributions whose supports can be unbounded. For a mechanism to approximate the optimal Bayesian revenue using *finite* non-adaptive queries to such distributions, it is intuitive that some kind of *small-tail assumption* for the distributions is needed. Indeed, given any mechanism with query complexity C , there always exists a distribution that has a sufficiently small probability mass around a sufficiently large value, such that the mechanism cannot find it using C queries. If this probability mass is where all the revenue comes (e.g., all the remaining probability mass is around value 0), then the mechanism cannot be a good approximation to OPT . Following the literature [33, 18], the small-tail assumptions are such that the expected revenue generated from the “tail” of the distributions is negligible compared to the optimal revenue; see Section 5.1. Distributions with bounded supports automatically satisfy these assumptions, so are regular distributions in single-item auctions.

Even with small-tail assumptions, it is hard to generate good revenue from unbounded distributions with finite *value* queries. Instead, we show it is sufficient to use only *quantile* queries. As before, the seller uses our quantile-query algorithm \mathcal{A}_Q (defined in Section 5.2) to learn a distribution \mathcal{D}' that approximates \mathcal{D} , and then reduces to simple mechanisms under \mathcal{D}' . However, even

for single-item auctions, it is not so simple to show why the combination of these two parts work. Indeed, under value queries it is easy to “under-price” the item so that the probability of sale is the same as in the optimal mechanism for \mathcal{D} . Under quantile queries, under-pricing may lose a large amount of revenue because, for given quantiles, there is no guarantee on where the corresponding values are. Instead, the main idea in using quantile queries is to “over-price” the item. This is risky in many auction design scenarios, because it may significantly reduce the probability of sale, and thus lose a lot of revenue. We prove a key technical lemma in Lemma 3 for single-item auctions, where we show that by discretizing the quantile space properly, we can over-price the item while almost preserving the probability of sale as in the optimal mechanism under \mathcal{D} . In Lemma 4, we prove another technical lemma showing that proper over-pricing can also be done in additive auctions.

Note that we can get the median of a distribution simply by querying the quantile 1/2. Then, for single-parameter auctions with regular distributions, using the parametric mechanism in [1] we get the same revenue as theirs. However, our query mechanisms can handle multi-parameter auctions and irregular distributions.

5.1 Small-Tail Assumptions

A Bayesian auction instance \mathcal{I} satisfies the *Small-Tail Assumption 1* if there exists a function¹ $h : (0, 1) \rightarrow (0, 1)$ such that, for any constant $\delta_1 \in (0, 1)$ and any BIC mechanism \mathcal{M} , letting $\epsilon_1 = h(\delta_1)$, we have

$$\mathbb{E}_{v \sim \mathcal{D}} \mathbf{I}_{\exists i, j, q_{ij}(v_{ij}) \leq \epsilon_1} \text{Rev}(\mathcal{M}(v; \mathcal{I})) \leq \delta_1 \text{OPT}(\mathcal{I}). \quad (16)$$

Here $q_{ij}(v_{ij})$ is the quantile of v_{ij} under distribution \mathcal{D}_{ij} , $\text{Rev}(\mathcal{M}(v; \mathcal{I}))$ is the revenue of \mathcal{M} under the Bayesian instance \mathcal{I} when the true valuation profile is v , and \mathbf{I} is the indicator function. For discrete distributions, Equation 16 is imposed on the ϵ_1 probability mass over the highest values.

Equation 16 immediately implies the following weaker *Small-Tail Assumption 2*: there exists a function $h : (0, 1) \rightarrow (0, 1)$ such that, for any constant $\delta_1 \in (0, 1)$, letting $\epsilon_1 = h(\delta_1)$, we have

$$\mathbb{E}_{v \sim \mathcal{D}} \mathbf{I}_{\exists i, j, q_{ij}(v_{ij}) \leq \epsilon_1} \text{Rev}_{\text{OPT}}(v; \mathcal{I}) \leq \delta_1 \text{OPT}(\mathcal{I}). \quad (17)$$

Here $\text{Rev}_{\text{OPT}}(v; \mathcal{I})$ is the revenue generated by the optimal BIC mechanism for \mathcal{I} when the true valuation profile is v . Similar assumptions are widely adopted in sampling mechanisms to deal with irregular distributions with unbounded supports.

5.2 The Quantile-Query Algorithm

We define our quantile-query algorithm \mathcal{A}_Q in Algorithm 5. As before, $D \in \Delta(\mathbb{R})$ is the distribution to be queried. The algorithm takes two parameters, the tail length ϵ_1 and the precision factor δ , makes $O(\log_{1+\delta} \frac{1}{\epsilon_1})$ quantile queries to the oracle, and then returns a discrete distribution D' .

5.3 Single-Item Auctions

Mechanism 6 defines our *efficient quantile Myerson* mechanism \mathcal{M}_{EQM} .

Theorem 6. $\forall \epsilon > 0$, any single-item instance $\mathcal{I} = (N, M, \mathcal{D})$ satisfying *Small-Tail Assumption 2*, \mathcal{M}_{EQM} is DSIC, has query complexity $O(-n \log_{1+\frac{\epsilon}{3}} h(\frac{2\epsilon}{3(1+\epsilon)}))$, and $\text{Rev}(\mathcal{M}_{EQM}(\mathcal{I})) \geq \frac{\text{OPT}(\mathcal{I})}{1+\epsilon}$.

¹If computation complexity is a concern, then one can further require that the function is efficiently computable.

Algorithm 5 The Quantile-Query Algorithm \mathcal{A}_Q

Input: the tail length ϵ_1 and the precision factor δ .

- 1: Let $k = \lceil \log_{1+\delta} \frac{1}{\epsilon_1} \rceil$ and define the *quantile vector* as $q = (q_0, q_1, \dots, q_{k-1}, q_k) = (1, \epsilon_1(1 + \delta)^{k-1}, \dots, \epsilon_1(1 + \delta), \epsilon_1)$.
- 2: Query the oracle for D with q , and receive a non-decreasing value vector $(v_l)_{l \in \{0, \dots, k\}}$.
- 3: Construct a distribution D' as follows: $D'(v_l) = q_l - q_{l+1}$ for each $l \in \{0, \dots, k\}$, where $q_{k+1} \triangleq 0$.

Output: Distribution D' .

Mechanism 6 Efficient Quantile Myerson Mechanism \mathcal{M}_{EQM}

- 1: Given $\epsilon > 0$, run algorithm \mathcal{A}_Q with $\delta = \frac{\epsilon}{3}$ and $\epsilon_1 = h(\frac{2\epsilon}{3(1+\epsilon)})$ (i.e., $\delta_1 = \frac{2\epsilon}{3(1+\epsilon)}$ for Small Tail Assumption 2), for each player i 's distribution \mathcal{D}_i . Denote by \mathcal{D}'_i the returned distribution. Let $\mathcal{D}' = \times_{i \in N} \mathcal{D}'_i$.
 - 2: Run Myerson's mechanism \mathcal{M}_{MRS} with \mathcal{D}' and the players' reported values, $b = (b_i)_{i \in N}$, to get allocation $x = (x_i)_{i \in N}$ and price profile $p = (p_i)_{i \in N}$ as the outcome.
-

Before proving Theorem 6, we first claim the following key lemma via an imaginary Bayesian mechanism that “over-prices”. Recall $\mathcal{I}' = (N, M, \mathcal{D}')$ is the instance under \mathcal{D}' .

Lemma 3. $Rev(\mathcal{M}_{MRS}(\mathcal{I}')) \geq \frac{1}{1+\epsilon} OPT(\mathcal{I})$.

Proof. For each player i , denote the support of \mathcal{D}'_i by $V'_i = (v'_{i;l})_{l \in \{0, \dots, k\}}$. We first define a way to couple the values $v'_i \sim \mathcal{D}'_i$ with the values $v_i \sim \mathcal{D}_i$.

The randomized round-down scheme. For any value $v_i \geq v'_{i;0}$, let v_i^- be v_i rounded down to the support of \mathcal{D}'_i , such that v_i^- is distributed according to \mathcal{D}'_i whenever v_i is distributed according to \mathcal{D}_i . Recall that under value queries, v_i^- is simply the largest value in V'_i that is less than or equal to v_i , no matter whether \mathcal{D}_i is continuous or discrete. Under quantile queries, when \mathcal{D}_i is continuous, the same deterministic round-down scheme still works. However, the situation is more subtle when \mathcal{D}_i is discrete, and we need a randomized round-down scheme to ensure the relationship between v_i and v_i^- . More precisely, by the definition of quantile queries, V'_i is a subset of \mathcal{D}_i 's support. If v_i is not in V'_i , then it is still deterministically rounded down as before. If v_i is in V'_i , say $v_i = v'_{i;l}$, then by the definition of quantile queries and the construction of \mathcal{D}'_i , we have $\Pr_{x \sim \mathcal{D}_i}[x \geq v_i] \geq q_l = \Pr_{x \sim \mathcal{D}'_i}[x \geq v_i]$. In this case, v_i is rounded down to $v'_{i;l-1}$ (i.e., $v_i^- = v'_{i;l-1}$) with probability

$$\frac{\Pr_{x \sim \mathcal{D}_i}[x \geq v_i] - \Pr_{x \sim \mathcal{D}'_i}[x \geq v_i]}{\mathcal{D}_i(v_i)},$$

and to $v'_{i;l}$ (i.e., $v_i^- = v'_{i;l}$) with probability

$$1 - \frac{\Pr_{x \sim \mathcal{D}_i}[x \geq v_i] - \Pr_{x \sim \mathcal{D}'_i}[x \geq v_i]}{\mathcal{D}_i(v_i)}.$$

Following this scheme, it is not hard to verify that $\Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq v'_{i;l}] = q_l$ for any $l \in \{0, \dots, k\}$, thus v_i^- is distributed according to \mathcal{D}'_i , as desired.

No matter what v_i^- is, let v_i^+ be the smallest value in V'_i that is strictly larger than v_i^- (if no such value exists, then $v_i^+ = +\infty$). That is, $v_i^+ \geq v_i$ and v_i^+ is v_i “rounded up”, which was not needed under value queries and is new for quantile queries.

The randomized resampling scheme. For any value $v'_i \sim \mathcal{D}'_i$, let v_i be *resampled* from \mathcal{D}_i conditional on “ v_i rounded down to v'_i ”, so that v_i is distributed according to \mathcal{D}_i whenever v'_i is distributed according to \mathcal{D}'_i . Again, under value queries, the resampling is simply conditional on $v_i \in [v'_{i,l}, v'_{i,l+1})$ when $v'_i = v'_{i,l}$, no matter whether \mathcal{D}_i is continuous or discrete. Under quantile queries, this resampling scheme still works when \mathcal{D}_i is continuous. When \mathcal{D}_i is discrete, we need to “undo” the randomized round-down scheme defined above. More precisely, letting $v'_i = v'_{i,l}$, v_i is set to be $v'_{i,l+1}$ with probability

$$p_1 = \frac{\Pr_{x \sim \mathcal{D}_i}[x \geq v'_{i,l+1}] - q_{l+1}}{\mathcal{D}'_i(v'_{i,l})};$$

is resampled from \mathcal{D}_i conditional on $v_i \in (v'_{i,l}, v'_{i,l+1})$ with probability

$$p_2 = \frac{\Pr_{x \sim \mathcal{D}_i}[v'_{i,l} < x < v'_{i,l+1}]}{\mathcal{D}'_i(v'_{i,l})};$$

and is set to be $v'_{i,l}$ with probability

$$p_3 = \frac{\Pr_{x \sim \mathcal{D}_i}[x \leq v'_{i,l}] - \Pr_{x \sim \mathcal{D}'_i}[x < v'_{i,l}]}{\mathcal{D}'_i(v'_{i,l})} = \frac{\mathcal{D}_i(v'_{i,l}) - \Pr_{x \sim \mathcal{D}_i}[x \geq v'_{i,l}] + q_l}{\mathcal{D}'_i(v'_{i,l})}.$$

Following this resampling scheme, it is not hard to verify that v_i is distributed according to \mathcal{D}_i whenever v'_i is distributed according to \mathcal{D}'_i .

Given the round-down and the resampling schemes above, we consider the Bayesian mechanism \mathcal{M}^* defined in Mechanism 7 for \mathcal{I}' , and compare its revenue with that of \mathcal{M}_{MRS} . We first claim that \mathcal{M}^* is a DSIC mechanism. Because \mathcal{M}_{MRS} is DSIC, each x_i is monotone in v_i . Although v_i is a random variable given v'_i , it is easy to see that for any two different values $v'_i \in V'_i$ and $\hat{v}'_i \in V'_i$, the corresponding resampled values v_i and \hat{v}_i are such that $v'_i < \hat{v}'_i$ implies $v_i \leq \hat{v}_i$. Thus x_i is monotone in v'_i as well. Moreover, let θ_i be player i 's threshold payment in \mathcal{M}_{MRS} given v_{-i} and \mathcal{D} . If $v'_i > \theta_i$ then $v_i > \theta_i$, thus player i gets the item at price $p_i = \theta_i$. If $v'_i < \theta_i$ then player i does not get the item and $p_i = 0$, no matter whether $v_i < \theta_i$ or not. Accordingly, θ_i is also player i 's threshold payment in \mathcal{M}^* under v_{-i} and \mathcal{D}' . Since v_{-i} does not depend on v'_i , \mathcal{M}^* is DSIC as desired.

Mechanism 7 A Bayesian mechanism \mathcal{M}^* for instance \mathcal{I}'

- 1: Each player i reports his value v'_i , and the mechanism discards the report that is not in V'_i .
 - 2: For each player i , generate value v_i according to v'_i using our resampling scheme.
 - 3: Run \mathcal{M}_{MRS} with the value profile v and the prior distribution \mathcal{D} , to get the price p_i and the allocation $x_i \in \{0, 1\}$ for each player i .
 - 4: If $x_i = 1$ and $p_i \leq v'_i$, sell the item to i and charge him p_i ; otherwise, set $x_i = 0$ and $p_i = 0$.
-

To analyze the revenue of \mathcal{M}^* , note that by construction, when v'_i is distributed according to \mathcal{D}'_i , the resampled v_i in \mathcal{M}^* is distributed according to \mathcal{D}_i . Moreover, each v'_i is distributed as if we first sample v_i from \mathcal{D}_i and then setting $v'_i = v_i^-$.

Thus, mechanism \mathcal{M}^* on instance \mathcal{I}' essentially generates the same expected revenue as \mathcal{M}_{MRS} on instance \mathcal{I} , except for the case when $v'_i < p_i \leq v_i$ for the winner i . Fortunately, we are able to upper-bound the probability of this event and thus upper-bound the expected revenue loss. More precisely, for each player i , we write p_i as $p_i(v_{-i}; \mathcal{D})$ to emphasize that it is the threshold payment

for i given v_{-i} and \mathcal{D} , and does not depend on v_i or v'_i . We have

$$\begin{aligned} \text{Rev}(\mathcal{M}^*(\mathcal{I}')) &= \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} p_i(v_{-i}; \mathcal{D}) \mathbf{I}_{v_i^- \geq p_i(v_{-i}; \mathcal{D})} \\ &= \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} p_i(v_{-i}; \mathcal{D}) \cdot \Pr_{v_i \sim \mathcal{D}_i} [v_i^- \geq p_i(v_{-i}; \mathcal{D})]. \end{aligned} \quad (18)$$

Here the first equality holds because of the relationship between \mathcal{D}' and \mathcal{D} as established by our rounding and resampling schemes, and because each player i in \mathcal{M}^* pays the same threshold price as in mechanism \mathcal{M}_{MRS} whenever v'_i is at least the threshold, and pays 0 otherwise. By the construction of the distribution \mathcal{D}' , we have the following claim, which is proved in Appendix A.

Claim 1. $\Pr_{v_i \sim \mathcal{D}_i} [v_i \geq p_i(v_{-i}; \mathcal{D}) | q_i(v_i) > \epsilon_1] \leq (1 + \delta) \Pr_{v_i \sim \mathcal{D}_i} [v_i^- \geq p_i(v_{-i}; \mathcal{D})]$.

Combining Equation 18, Claim 1 and Small Tail Assumption 2, we are able to lower-bound the revenue of \mathcal{M}^* as follows, which is also proved in Appendix A.

Claim 2. $\text{Rev}(\mathcal{M}^*(\mathcal{I}')) \geq \frac{1}{1+\epsilon} \text{OPT}(\mathcal{I})$.

By the optimality of \mathcal{M}_{MRS} , $\text{Rev}(\mathcal{M}_{MRS}(\mathcal{I}')) \geq \text{Rev}(\mathcal{M}^*(\mathcal{I}'))$, and Lemma 3 holds. \square

Proof of Theorem 6. First, mechanism \mathcal{M}_{EQM} is DSIC because \mathcal{M}_{MRS} is DSIC. Second, the query complexity of \mathcal{M}_{EQM} is $O(-n \log_{1+\frac{\epsilon}{3}} h(\frac{2\epsilon}{3(1+\epsilon)}))$, because there are $k+1 = \lceil \log_{1+\frac{\epsilon}{3}} \frac{1}{h(\frac{2\epsilon}{3(1+\epsilon)})} \rceil + 1$ quantile queries for each player and there are n players in total. By definition,

$\text{Rev}(\mathcal{M}_{EQM}(\mathcal{I})) = \text{Rev}(\mathcal{M}_{MRS}(\mathcal{I}; \mathcal{D}'))$. By construction, \mathcal{D}' is stochastically dominated by \mathcal{D} . Thus by revenue monotonicity $\text{Rev}(\mathcal{M}_{MRS}(\mathcal{I}; \mathcal{D}')) \geq \text{Rev}(\mathcal{M}_{MRS}(\mathcal{I}'))$. Combining these two equations with Lemma 3, Theorem 6 holds. \square

Mechanism \mathcal{M}_{EQM} and Theorem 6 immediately extend to single-parameter downward-closed settings. Finally, when the distributions are regular, we are able to prove an even better query complexity and a matching lower-bound; see Section 6.

5.4 Unit-Demand Auctions

The unit-demand mechanism \mathcal{M}_{EQUD} is similar (see Mechanism 8), and we have the following.

Theorem 7. $\forall \epsilon > 0$, any unit-demand instance $\mathcal{I} = (N, M, \mathcal{D})$ satisfying Small-Tail Assumption 2, \mathcal{M}_{EQUD} is DSIC, has query complexity $O(-mn \log_{1+\frac{\epsilon}{3}} h(\frac{2\epsilon}{3(1+\epsilon)}))$, and $\text{Rev}(\mathcal{M}_{EQUD}(\mathcal{I})) \geq \frac{\text{OPT}(\mathcal{I})}{24(1+\epsilon)}$.

Mechanism 8 Mechanism \mathcal{M}_{EQUD} for Unit-Demand Auctions

- 1: Given $\epsilon > 0$, run algorithm \mathcal{A}_Q with $\delta = \frac{\epsilon}{3}$ and $\epsilon_1 = h(\frac{2\epsilon}{3(1+\epsilon)})$ (i.e., $\delta_1 = \frac{2\epsilon}{3(1+\epsilon)}$ for Small Tail Assumption 2), for each player i 's distribution \mathcal{D}_{ij} on each item j . Denote by \mathcal{D}'_{ij} the returned distribution. Let $\mathcal{D}'_i = \times_{j \in M} \mathcal{D}'_{ij}$ and $\mathcal{D}' = \times_{i \in N} \mathcal{D}'_i$.
 - 2: Run mechanism \mathcal{M}_{UD} with \mathcal{D}' and the players' reported values, $b = (b_{ij})_{i \in N, j \in M}$, to get allocation $x = (x_{ij})_{i \in N, j \in M}$ and price profile $p = (p_i)_{i \in N}$ as the outcome.
-

The proof of Theorem 7 is similar to that of Theorem 4, but Lemma 3 above is used instead of Lemma 5 of [18], and the round-down scheme is replaced by the randomized round-down scheme designed in the proof of Lemma 3. The details have been omitted.

5.5 Additive Auctions

For additive auctions, we cannot use Small-Tail Assumption 2, because it does not imply that the revenue loss on the tail by running \mathcal{M}_{BVCG} is much less than the revenue of the optimal mechanism. To approximate \mathcal{M}_{BVCG} , not only we need Small-Tail Assumption 1, but we also approximate \mathcal{D} by running the quantile-query algorithm \mathcal{A}_Q with different parameters. The resulting mechanism \mathcal{M}_{EQBVCG} is defined in Mechanism 9, and the mechanism \mathcal{M}_{EQIM} simply replaces \mathcal{M}_{BVCG} with \mathcal{M}_{IM} . Again, in the final mechanism \mathcal{M}_{EQA} the seller randomly chooses between the two query mechanisms, according to probabilities defined in the analysis. We have the following theorem.

Mechanism 9 Mechanism \mathcal{M}_{EQBVCG} for Additive Auctions

- 1: Given $\epsilon > 0$, run algorithm \mathcal{A}_Q with $\delta = (1 + \frac{\epsilon}{5})^{1/m} - 1$ and $\epsilon_1 = h(\frac{\epsilon}{10(1+\epsilon)})$ (i.e., $\delta_1 = \frac{\epsilon}{10(1+\epsilon)}$ for Small Tail Assumption 1), for each player i 's distribution \mathcal{D}_{ij} on each item j . Denote by \mathcal{D}'_{ij} the returned distribution. Let $\mathcal{D}'_i = \times_{j \in M} \mathcal{D}'_{ij}$ and $\mathcal{D}' = \times_{i \in N} \mathcal{D}'_i$.
 - 2: Run \mathcal{M}_{BVCG} with \mathcal{D}' and the players' reported values, $b = (b_{ij})_{i \in N, j \in M}$, to get allocation $x = (x_{ij})_{i \in N, j \in M}$ and price profile $p = (p_i)_{i \in N}$ as the outcome.
-

Theorem 8. $\forall \epsilon > 0$, any additive instance $\mathcal{I} = (N, M, \mathcal{D})$ satisfying Small-Tail Assumption 1, \mathcal{M}_{EQA} is DSIC, has query complexity $O(-m^2 n \log_{1+\frac{\epsilon}{5}} h(\frac{\epsilon}{10(1+\epsilon)}))$, and $Rev(\mathcal{M}_{EQA}(\mathcal{I})) \geq \frac{OPT(\mathcal{I})}{8(1+\epsilon)}$.

Before proving Theorem 8, we first analyze mechanism \mathcal{M}_{EQBVCG} , and we have the following.

Lemma 4. $\forall \epsilon > 0$, for any additive Bayesian instance $\mathcal{I} = (N, M, \mathcal{D})$ satisfying Small-Tail Assumption 1, \mathcal{M}_{EQBVCG} is DSIC, has query complexity $O(-m^2 n \log_{1+\frac{\epsilon}{5}} h(\frac{\epsilon}{10(1+\epsilon)}))$, and

$$Rev(\mathcal{M}_{EQBVCG}(\mathcal{I})) \geq \frac{1}{1 + \frac{\epsilon}{5}} \left(Rev(\mathcal{M}_{BVCG}(\mathcal{I})) - \frac{\epsilon}{10(1+\epsilon)} OPT(\mathcal{I}) \right).$$

Proof. First, mechanism \mathcal{M}_{EQBVCG} is DSIC because \mathcal{M}_{BVCG} is DSIC. The query complexity is also immediate.

We now focus on the revenue of this mechanism. We explicitly write $\mathcal{M}_{BVCG}(\mathcal{I}; \mathcal{D}')$ to emphasize the fact that the seller runs mechanism \mathcal{M}_{BVCG} on the true valuation profile $v \sim \mathcal{D}$, but uses \mathcal{D}' to compute the entry fees e_i . Given a player i and a valuation profile v , $p_i(v_i, \mathcal{D}_i, v_{-i})$ is the price for i under \mathcal{D}_i : that is,

$$p_i(v_i, \mathcal{D}_i, v_{-i}) = \mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} (e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}}),$$

where we omit v_{-i} in $\beta_{ij}(v_{-i})$ when v_{-i} is clear from the context. ² The price $p_i(v_i, \mathcal{D}'_i, v_{-i})$ is similarly defined. By the definition of the mechanism, we have

$$Rev(\mathcal{M}_{EQBVCG}(\mathcal{I})) = Rev(\mathcal{M}_{BVCG}(\mathcal{I}; \mathcal{D}')) = \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} p_i(v_i, \mathcal{D}'_i, v_{-i}). \quad (19)$$

Next, let V'_{ij} be the support of \mathcal{D}'_{ij} , $V'_i = \times_{j \in M} V'_{ij}$, round v_i down to the closest valuation v'_i in V'_i and compare the two valuation profiles (v'_i, v_{-i}) and (v_i, v_{-i}) . By definition, $v'_{ij} \geq \beta_{ij}$ implies $v_{ij} \geq \beta_{ij}$. Moreover, the entry fee of i is the same under both valuation profiles, as it

²If there are ties in the players' values, then we distinguish between β_{ij}^+ and β_{ij}^- , depending on the identity of the player with the highest bid for j in $N \setminus \{i\}$.

only depends on \mathcal{D}'_i and v_{-i} . Similarly, the reserve price β_{ij} is the same for any item j . Thus we have $e(\mathcal{D}'_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \geq e(\mathcal{D}'_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v'_{ij} \geq \beta_{ij}}$ and $\mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \geq \mathbf{I}_{\sum_{j: v'_{ij} \geq \beta_{ij}} (v'_{ij} - \beta_{ij}) \geq e(\mathcal{D}'_i, v_{-i})}$. Therefore

$$\mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} p_i(v_i, \mathcal{D}'_i, v_{-i}) \geq \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} p_i(v'_i, \mathcal{D}'_i, v_{-i}) = \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v'_i \sim \mathcal{D}'_i} p_i(v'_i, \mathcal{D}'_i, v_{-i}), \quad (20)$$

where the equality is again because drawing v_i from \mathcal{D}_i and then rounding down to v'_i is equivalent to drawing v'_i from \mathcal{D}'_i directly.

In mechanism \mathcal{M}_{BVCG} , given v_{-i} and \mathcal{D}'_i , $e(\mathcal{D}'_i, v_{-i})$ is the optimal entry fee for maximizing the expected revenue generated from i , where the expectation is taken over \mathcal{D}'_i . Accordingly,

$$\mathbb{E}_{v'_i \sim \mathcal{D}'_i} p_i(v'_i, \mathcal{D}'_i, v_{-i}) \geq \mathbb{E}_{v'_i \sim \mathcal{D}'_i} p_i(v'_i, \mathcal{D}_i, v_{-i}). \quad (21)$$

Combining Equations 19, 20 and 21, we have

$$Rev(\mathcal{M}_{EQBVCG}(\mathcal{I})) \geq \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v'_i \sim \mathcal{D}'_i} p_i(v'_i, \mathcal{D}_i, v_{-i}). \quad (22)$$

Thus we will use $\sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v'_i \sim \mathcal{D}'_i} p_i(v'_i, \mathcal{D}_i, v_{-i})$ to upper-bound $Rev(\mathcal{M}_{BVCG}(\mathcal{I}))$.

To do so, first, for any player i , item j and value v_{ij} , if $v_{ij} < v'_{ij;k}$ where $v'_{ij;k}$ is the largest value in V'_{ij} , then denote by \bar{v}_{ij} the smallest value in V'_{ij} that is strictly larger than v_{ij} ; otherwise, let $\bar{v}_{ij} = v_{ij}$. Moreover, denote by \underline{v}_{ij} the largest value in V'_{ij} that is weakly smaller than v_{ij} . The valuation \bar{v}_i and \underline{v}_i are defined correspondingly given v_i . Then We have

$$\begin{aligned} Rev(\mathcal{M}_{BVCG}(\mathcal{I})) &= \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} p_i(v_i, \mathcal{D}_i, v_{-i}) \\ &= \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right) \\ &= \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \mathbf{I}_{\forall j, q_{ij}(v_{ij}) > \epsilon_1} \mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right) \\ &+ \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \mathbf{I}_{\exists j, q_{ij}(v_{ij}) \leq \epsilon_1} \mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right). \end{aligned} \quad (23)$$

Below we upper-bound the last two lines in Equation 23 separately. For the first part, we have

$$\begin{aligned} &\sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \mathbf{I}_{\forall j, q_{ij}(v_{ij}) > \epsilon_1} \mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right) \\ &\leq \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \mathbf{I}_{\forall j, q_{ij}(v_{ij}) > \epsilon_1} \mathbf{I}_{\sum_{j: \bar{v}_{ij} \geq \beta_{ij}} (\bar{v}_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{\bar{v}_{ij} \geq \beta_{ij}} \right) \\ &= \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \sum_{u_i \in V'_i: \sum_{j: u_{ij} \geq \beta_{ij}} (u_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \Pr_{v_i \sim \mathcal{D}_i} [\bar{v}_i = u_i] \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{u_{ij} \geq \beta_{ij}} \right). \end{aligned} \quad (24)$$

The inequality above is because $v_{ij} \leq \bar{v}_{ij}$ for each player i and item j , which implies

$$\mathbf{I}_{\sum_j: v_{ij} \geq \beta_{ij} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \leq \mathbf{I}_{\sum_j: \bar{v}_{ij} \geq \beta_{ij} (\bar{v}_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \text{ and } \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \leq \sum_j \beta_{ij} \mathbf{I}_{\bar{v}_{ij} \geq \beta_{ij}}.$$

Next, by the definition of the quantile vector q , for any $u_{ij} \in V'_{ij}$ we have

$$\Pr_{v_{ij} \sim \mathcal{D}_{ij}} [\bar{v}_{ij} = u_{ij}] \leq (1 + \delta) \Pr_{v_{ij} \sim \mathcal{D}_{ij}} [\underline{v}_{ij} = u_{ij}].$$

Indeed, when $u_{ij} = v'_{ij;0}$, $\Pr[v_{ij} < u_{ij}] = 0 < (1 + \delta)(1 - \epsilon_1(1 + \delta)^{k-1}) = (1 + \delta)(q_0 - q_1) = (1 + \delta) \Pr[v_{ij} \in [v'_{ij;0}, v'_{ij;1}]]$. When $u_{ij} = v'_{ij;l}$ with $0 < l < k$, $\Pr(v_{ij} \in [v'_{ij;l-1}, v'_{ij;l}]) = q_{l-1} - q_l \leq (1 + \delta)q_l - q_l = \delta q_l = (1 + \delta)\delta q_{l+1} = (1 + \delta)((1 + \delta)q_{l+1} - q_{l+1}) = (1 + \delta)(q_l - q_{l+1}) = (1 + \delta) \Pr[v_{ij} \in [v'_{ij;l}, v'_{ij;l+1}]]$. And when $u_{ij} = v'_{ij;k}$, $\Pr[v_{ij} \in [v'_{ij;k-1}, v'_{ij;k}]] = q_{k-1} - q_k = \delta \epsilon_1 < \epsilon_1 = \Pr[v_{ij} \geq v'_{ij;k}]$. Since all distributions are independent, for any $u_i \in V'_i$ we have

$$\Pr_{v_i \sim \mathcal{D}_i} [\bar{v}_i = u_i] \leq (1 + \delta)^m \Pr_{v_i \sim \mathcal{D}_i} [\underline{v}_i = u_i]. \quad (25)$$

Combining Equations 24 and 25, we have

$$\begin{aligned} & \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \mathbf{I}_{\forall j, q_{ij}(v_{ij}) > \epsilon_1} \mathbf{I}_{\sum_j: v_{ij} \geq \beta_{ij} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right) \\ & \leq \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{u_i \in V'_i: \sum_{j: u_{ij} - \beta_{ij} \geq e(\mathcal{D}_i, v_{-i})}} (1 + \delta)^m \cdot \Pr_{v_i \sim \mathcal{D}_i} [\underline{v}_i = u_i] \cdot \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{u_{ij} \geq \beta_{ij}} \right) \\ & = (1 + \delta)^m \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v'_i \sim \mathcal{D}'_i} \mathbf{I}_{\sum_j: v'_{ij} \geq \beta_{ij} (v'_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v'_{ij} \geq \beta_{ij}} \right) \\ & = (1 + \delta)^m \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v'_i \sim \mathcal{D}'_i} p_i(v'_i, \mathcal{D}_i, v_{-i}) \leq (1 + \delta)^m \text{Rev}(\mathcal{M}_{ECBVC}(\hat{\mathcal{I}})). \end{aligned} \quad (26)$$

The first equality above holds because drawing v_i from \mathcal{D}_i and rounding down to the support of \mathcal{D}'_i is equivalent to drawing v'_i from \mathcal{D}'_i . The second equality is by the definition of $p_i(v'_i, \mathcal{D}_i, v_{-i})$, and the last inequality holds by Equation 22.

By Equations 23 and 26, we have

$$\begin{aligned} & \text{Rev}(\mathcal{M}_{BVC}(\mathcal{I})) \\ & \leq (1 + \delta)^m \text{Rev}(\mathcal{M}_{EQBVC}(\mathcal{I})) \\ & + \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \mathbf{I}_{\exists j, q_{ij}(v_{ij}) \leq \epsilon_1} \mathbf{I}_{\sum_j: v_{ij} \geq \beta_{ij} (v_{ij} - \beta_{ij}) \geq e(\mathcal{D}_i, v_{-i})} \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right). \end{aligned} \quad (27)$$

For the last line of Equation 27, we have

$$\begin{aligned}
& \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} \mathbf{I}_{\exists j, q_{ij}(v_{ij}) \leq \epsilon_1} \mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq \epsilon}(\mathcal{D}_i, v_{-i}) \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right) \\
&= \mathbb{E}_{v \sim \mathcal{D}} \sum_i \mathbf{I}_{\exists j, q_{ij}(v_{ij}) \leq \epsilon_1} \mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq \epsilon}(\mathcal{D}_i, v_{-i}) \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right) \\
&\leq \mathbb{E}_{v \sim \mathcal{D}} \mathbf{I}_{\exists i, j, q_{ij}(v_{ij}) \leq \epsilon_1} \sum_i \mathbf{I}_{\sum_{j: v_{ij} \geq \beta_{ij}} (v_{ij} - \beta_{ij}) \geq \epsilon}(\mathcal{D}_i, v_{-i}) \left(e(\mathcal{D}_i, v_{-i}) + \sum_j \beta_{ij} \mathbf{I}_{v_{ij} \geq \beta_{ij}} \right) \\
&= \mathbb{E}_{v \sim \mathcal{D}} \mathbf{I}_{\exists i, j, q_{ij}(v_{ij}) \leq \epsilon_1} \text{Rev}(\mathcal{M}_{BVCG}(v; \mathcal{I})) \leq \frac{\epsilon}{10(1+\epsilon)} \text{OPT}(\mathcal{I}). \tag{28}
\end{aligned}$$

The first inequality above is because, for each player i and valuation profile v , $\mathbf{I}_{\exists j, q_{ij}(v_{ij}) \leq \epsilon_1} \leq \mathbf{I}_{\exists i, j, q_{ij}(v_{ij}) \leq \epsilon_1}$. The second inequality is by the Small-Tail Assumption 1.

Combining Equations 27 and 28, we have

$$\text{Rev}(\mathcal{M}_{BVCG}(\mathcal{I})) \leq (1 + \delta)^m \text{Rev}(\mathcal{M}_{EQBVCG}(\mathcal{I})) + \frac{\epsilon}{10(1+\epsilon)} \text{OPT}(\mathcal{I}).$$

By the construction of Mechanism 9, $(1 + \delta)^m = 1 + \frac{\epsilon}{5}$. Therefore Lemma 4 holds. \square

Proof of Theorem 8. First, as both \mathcal{M}_{EQBVCG} and \mathcal{M}_{EQIM} are DSIC, \mathcal{M}_{EQA} is DSIC. Second, note that \mathcal{M}_{EQA} runs both mechanisms with $\delta = (1 + \frac{\epsilon}{5})^{1/m} - 1$ and $\epsilon_1 = h(\frac{\epsilon}{10(1+\epsilon)})$. To ease the analysis, when running mechanism \mathcal{M}_{EQIM} , let $\delta = \frac{\epsilon}{15}$ and $\epsilon_1 = h(\frac{2\epsilon}{3(5+\epsilon)})$: that is, set $\epsilon' = \frac{\epsilon}{5}$ and run mechanism \mathcal{M}_{EQM} with parameter ϵ' for each item. By Theorem 6, with $O(-mn \log_{1+\frac{\epsilon}{15}} h(\frac{2\epsilon}{3(5+\epsilon)}))$ queries,

$$\text{Rev}(\mathcal{M}_{EQIM}(\mathcal{I})) \geq \frac{1}{1 + \frac{\epsilon}{5}} \text{Rev}(\mathcal{M}_{IM}(\mathcal{I})).$$

By Lemma 4, with $O(-m^2 n \log_{1+\frac{\epsilon}{5}} h(\frac{\epsilon}{10(1+\epsilon)}))$ queries,

$$\text{Rev}(\mathcal{M}_{EQBVCG}(\mathcal{I})) \geq \frac{1}{1 + \frac{\epsilon}{5}} \left(\text{Rev}(\mathcal{M}_{BVCG}(\mathcal{I})) - \frac{\epsilon}{10(1+\epsilon)} \text{OPT}(\mathcal{I}) \right).$$

Note that the total query complexity is still $O(-m^2 n \log_{1+\frac{\epsilon}{5}} h(\frac{\epsilon}{10(1+\epsilon)}))$.

Let mechanism \mathcal{M}_{EQA} run \mathcal{M}_{EQBVCG} with probability $\frac{1}{4}$ and \mathcal{M}_{EQIM} with probability $\frac{3}{4}$. We have

$$\begin{aligned}
\text{Rev}(\mathcal{M}_{EQA}(\mathcal{I})) &= \frac{1}{4} \text{Rev}(\mathcal{M}_{EQBVCG}(\mathcal{I})) + \frac{3}{4} \text{Rev}(\mathcal{M}_{EQIM}(\mathcal{I})) \\
&\geq \frac{1}{4(1 + \frac{\epsilon}{5})} \left(\text{Rev}(\mathcal{M}_{BVCG}(\mathcal{I})) - \frac{\epsilon}{10(1+\epsilon)} \text{OPT}(\mathcal{I}) \right) + \frac{3}{4(1 + \frac{\epsilon}{5})} \text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) \\
&\geq \frac{1}{1 + \frac{\epsilon}{5}} \left(\frac{1}{4} \text{Rev}(\mathcal{M}_{BVCG}(\mathcal{I})) + \frac{3}{4} \text{Rev}(\mathcal{M}_{IM}(\mathcal{I})) - \frac{\epsilon}{10(1+\epsilon)} \text{OPT}(\mathcal{I}) \right) \\
&\geq \frac{1}{1 + \frac{\epsilon}{5}} \left(\frac{1}{8} \text{OPT}(\mathcal{I}) - \frac{\epsilon}{10(1+\epsilon)} \text{OPT}(\mathcal{I}) \right) = \frac{1}{8(1+\epsilon)} \text{OPT}(\mathcal{I}).
\end{aligned}$$

The last inequality above holds because $2\mathcal{M}_{BVCG}(\mathcal{I}) + 6\mathcal{M}_{IM}(\mathcal{I}) \geq \text{OPT}(\mathcal{I})$ [9]. Thus Theorem 8 holds. \square

The main advantage of using quantile queries is to handle unbounded distributions. In addition, we can use the resulting query mechanisms to construct *sampling* mechanisms; see Section 7. As shown in Theorem 8, the query complexity of mechanism \mathcal{M}_{EQA} has an extra factor of m compared with that of \mathcal{M}_{EVA} (and the lower bound). It would be interesting to see whether our lower-bounds can be improved in this scenario.

5.6 Using Quantile Queries for Bounded Distributions

As a corollary, Theorems 6, 7 and 8 also provide another way to approximate the optimal BIC revenue using only quantile queries when the distributions are bounded. More precisely, we have the following.

Corollary 1. *For any $\epsilon > 0$, $H > 1$, and prior distribution \mathcal{D} with each \mathcal{D}_{ij} bounded within $[1, H]$, there exist DSIC mechanisms that use $O(mn \log_{1+\epsilon} \frac{nmH(1+\epsilon)}{\epsilon})$ quantile queries for single-item auctions and unit-demand auctions, and use $O(m^2n \log_{1+\epsilon} \frac{nmH(1+\epsilon)}{\epsilon})$ quantile queries for additive auctions, whose approximation ratios to OPT are respectively $1 + \epsilon$, $24(1 + \epsilon)$ and $8(1 + \epsilon)$.*

Proof. We only need to show that the Small-Tail Assumptions 1 and 2 are naturally satisfied when the distributions have bounded supports. For example, consider additive auctions where all values are in $[1, H]$, as considered in [28, 15]. Then mH and 1 are straightforward upper- and lower-bounds for $OPT(\mathcal{I})$, respectively. Moreover, by individual rationality, mH is an upper-bound for the revenue generated under any valuation profiles. Given δ_1 , let $\epsilon_1 = h(\delta_1) = \frac{\delta_1}{m^2nH}$ and denote by E the event that there exist at least one player i and one item j with $q_{ij}(v_{ij}) \leq \epsilon_1$. By the union bound, $\Pr[E] \leq mn\epsilon_1 = mn \cdot \frac{\delta_1}{m^2nH} = \frac{\delta_1}{mH}$. Therefore

$$\mathbb{E}_{v \sim \mathcal{D}} \mathbf{I}_{\exists i, j, q_{ij}(v_{ij}) \leq \epsilon_1} \text{Rev}(\mathcal{M}(v; \mathcal{I})) \leq mH \cdot \Pr[E] \leq \delta_1 \leq \delta_1 OPT(\mathcal{I}).$$

Combining this observation with Theorems 6, 7 and 8, we have Corollary 1 when the values are all bounded in $[1, H]$. \square

6 Single-Item Auctions with Regular Distributions

In this section, we show that when we only consider regular distributions for single-item auctions, the query complexity can be much lower. In fact, we no longer need the small-tail assumptions even when the supports are unbounded. Here our lower- and upper-bounds are *tight* upto a logarithmic factor, and require different techniques from previous sections.

For the lower-bound, recall that in Section 3 we allow the distributions to be irregular. To construct the desired distributions, we can first find the un-queried quantile interval and then move the probability mass from its end points to internal points. Because the distributions can be irregular, we have complete control on where to put the probability mass. However, if the distributions have to be regular then this cannot be done. Instead, we start from two different single-peak revenue curves and construct regular distributions from them. We still want to move probability mass from the end points of the un-queried quantile interval to internal points, but such moves must be continuous in order to preserve regularity.

For the upper-bound, we show that regular distributions satisfy the small-tail property with a properly defined tail function. Thus our techniques for distributions with small-tails directly apply here.

6.1 Lower Bound

With regular distributions, by [19] it is sufficient to use a single sample to achieve 2-approximation in revenue for *single-player* single-item auctions. Because every distribution is a uniform distribution in the quantile space, a sample for such auctions can be obtained by first choosing a quantile q uniformly at random from $[0, 1]$ and then making a quantile query. Thus, a single query is also sufficient for 2-approximation in this case. As such, unlike Theorem 2 where we have proved lower bounds for the query complexity for arbitrary constant approximations, for regular distributions we consider lower bounds for $(1 + \epsilon)$ -approximations, where ϵ is sufficiently small. More precisely, we have the following.

Theorem 9. *For any constant $\epsilon \in (0, \frac{1}{64})$, there exists a constant C such that, for any $n \geq 1$, any DSIC Bayesian mechanism \mathcal{M} making less than $Cn\epsilon^{-1}$ non-adaptive value and quantile queries to the oracle, there exists a multi-player single-item Bayesian auction instance $\mathcal{I} = (N, M, \mathcal{D})$ where $|N| = n$ and \mathcal{D} is regular, such that $\text{Rev}(\mathcal{M}(\mathcal{I})) < \frac{\text{OPT}(\mathcal{I})}{1+\epsilon}$.*

We only prove Theorem 9 for the single-player case, as in the following lemma. The lower bound for general multi-player single-item auctions can be proved using the same technique as in Theorem 2, thus the full proof has been omitted.

Lemma 5. *For any constant $\epsilon \in (0, \frac{1}{64})$, there exists a constant C such that, for any DSIC Bayesian mechanism \mathcal{M} making less than C/ϵ non-adaptive value and quantile queries to the oracle, there exists a single-player single-item Bayesian auction instance $\mathcal{I} = (N, M, \mathcal{D})$ where \mathcal{D} is regular, such that $\text{Rev}(\mathcal{M}(\mathcal{I})) < \frac{\text{OPT}(\mathcal{I})}{1+\epsilon}$.*

Proof. Since the distributions are unbounded, we can always construct the distributions such that for any finite number of value queries, the responses for the value queries have almost none contribution to the optimal revenue. Thus we only need to focus on the lower bound for quantile queries.

Letting $k \triangleq \lceil \frac{1}{\delta\epsilon} \rceil$ and $C \triangleq \frac{1-2\delta\epsilon}{2\delta}$. Here δ is a constant to be determined later and δ, ϵ satisfies that $k \geq 2$. In our construction, we divide the quantile interval $[0, 1]$ into $k + 1$ sub-intervals each, with the right-end points defined as follows: from left to right, $q_0 = 0$, $q_{t+1} = q_t + \delta\epsilon$ for each $t \in \{0, \dots, k-1\}$.

Accordingly, for any Bayesian mechanism \mathcal{M} that makes less than $\frac{C}{\epsilon}$ non-adaptive quantile queries, there exists a quantile interval (q_t, q_{t+1}) such that, $q_{t+1} \leq 1 - 2\delta\epsilon$ and with probability at least $\frac{1}{2}$, no quantile in (q_t, q_{t+1}) is queried. Indeed, if this is not the case, then with probability at least $\frac{1}{2}$, all the quantile intervals except $(1 - 2\delta\epsilon, 1 - \delta\epsilon)$ and $(1 - \delta\epsilon, 1)$ are queried. Since there are at least $k - 2$ quantile intervals, the expected total number of queries made by \mathcal{M} is at least $\frac{k}{2} - 1 \geq \frac{1-2\delta\epsilon}{2\delta\epsilon} = \frac{C}{\epsilon}$, a contradiction.

We now construct two different single-player single-item Bayesian instances

$$\{\mathcal{I}_z = (N, M, \mathcal{D}_z)\}_{z \in \{1, 2\}},$$

where the distributions outside the quantile range (q_t, q_{t+1}) are all the same. Thus with probability at least $\frac{1}{2}$, mechanism \mathcal{M} cannot distinguish the \mathcal{I}_z 's from each other. We then show that when this happens, mechanism \mathcal{M} cannot be a $(1 + 3\epsilon)$ -approximation for all instances \mathcal{I}_z .

Let R be a parameter that is large enough such that no value query will get any useful response. Then the first distribution \mathcal{D}_1 with value bounded within $[0, \frac{R}{q_t}]$ is defined as follows, where $F_1(\cdot)$ is the cumulative probability function of \mathcal{D}_1 .

$$F_1(v) = \begin{cases} 1 - \frac{R}{(1-q_{t+1})v+R}, & 0 \leq v < \frac{R}{q_t}, \\ 1, & v = \frac{R}{q_t}. \end{cases}$$

That is there is a probability mass $\frac{q_t}{1-\delta\epsilon}$ at value $\frac{R}{q_t}$ and within interval $[0, \frac{R}{q_t})$ it is a continuous distribution. Then for any quantile in range $(0, \frac{q_t}{1-\delta\epsilon}]$, the oracle will response $\frac{R}{q_t}$. For quantile q in range $(\frac{q_t}{1-\delta\epsilon}, 1]$, the oracle will response $v(q) = \frac{R}{1-q_{t+1}}(\frac{1}{q} - 1)$. Therefore the revenue function with related to the quantile q is

$$R_1(q) = \begin{cases} \frac{R}{1-q_{t+1}}(1-q), & \frac{q_t}{1-\delta\epsilon} < q \leq 1, \\ \frac{R}{1-\delta\epsilon}, & q = \frac{q_t}{1-\delta\epsilon}. \end{cases}$$

The revenue curve $R_1(q)$ is illustrated figure 1.

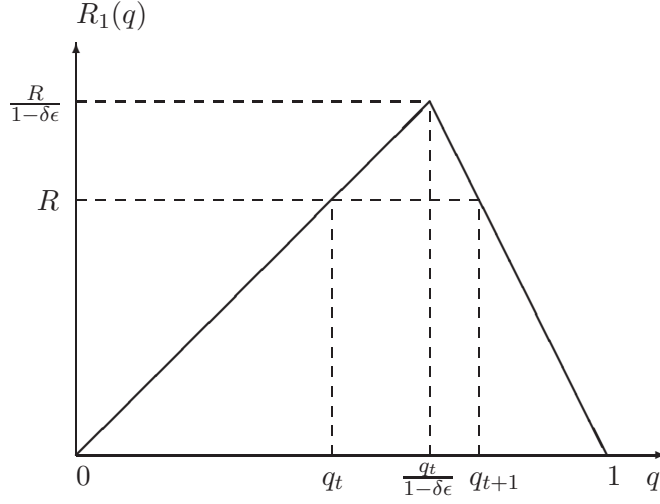


Figure 1: The revenue curve of \mathcal{D}_1 .

The second distribution \mathcal{D}_2 with value bounded within $[0, \frac{R}{q_t}]$ is defined as follows, where $F_2(\cdot)$ is the cumulative probability function of \mathcal{D}_2 . Let $v^* = \frac{R(2-\delta\epsilon)}{2(1-\delta\epsilon)-(2-\delta\epsilon)(1-q_{t+1})}$. Since $q_{t+1} \leq 1 - 2\delta\epsilon$, $v^* > 0$ is well defined and it is easy to check $v^* < \frac{R}{q_t}$.

$$F_2(v) = \begin{cases} 1 - \frac{R}{(1-q_{t+1})v+R}, & 0 \leq v < v^*, \\ 1 - \frac{R(1-\delta\epsilon)}{(1+q_t-\delta\epsilon)v-R}, & v^* \leq v < \frac{R}{q_t}, \\ 1, & v = \frac{R}{q_t}. \end{cases}$$

That is, there is a probability mass q_t at value $\frac{R}{q_t}$ and a two-step continuous distribution within $[q_t, q^*]$ and $[q^*, q_{t+1}]$. Thus for any quantile in range $(0, q_t]$, the oracle will response $\frac{R}{q_t}$. It can be calculated that the quantile of value v^* is $q^* = 1 - \frac{2-\delta\epsilon}{2(1-\delta\epsilon)} \cdot (1 - q_{t+1})$. Then for quantile q in range $(q_t, q^*]$, the oracle will response $v(q) = \frac{R}{q}(1 - \frac{q_t}{1+q_t-\delta\epsilon}) + \frac{R}{1+q_t-\delta\epsilon}$. For quantile q in range $(q^*, 1]$, the oracle will response $v(q) = \frac{R}{1-q_{t+1}}(\frac{1}{q} - 1)$. Therefore the revenue function with related to the quantile q is

$$R_2(q) = \begin{cases} \frac{R}{1-q_{t+1}}(1-q), & q^* < q \leq 1, \\ \frac{R}{1+q_t-\delta\epsilon}(1+q-\delta\epsilon), & q_t \leq q < q^*, \\ R, & q = q_t. \end{cases}$$

The revenue curve $R_2(q)$ is illustrated figure 2.

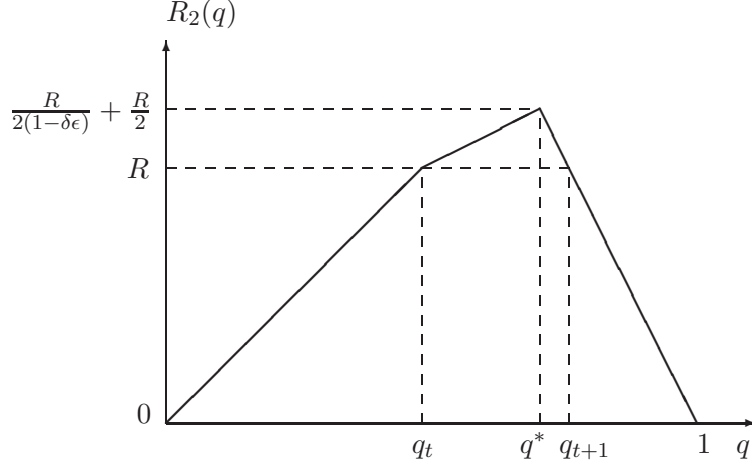


Figure 2: The revenue curve of \mathcal{D}_2 .

Indeed when the quantile query is from $[0, q_t] \cup [q_{t+1}, 1]$, the oracle's answers for all distributions are the same. Accordingly, with probability at least $\frac{1}{2}$, mechanism \mathcal{M} cannot distinguish \mathcal{D}_z 's from each other, which means it cannot distinguish \mathcal{I}_z 's from each other, as desired.

Since \mathcal{M} is truthful, the allocation rule for the player must be monotone and he will pay the threshold payment set by \mathcal{M} , denoted by P . Let $P^* = \frac{(4-\delta\epsilon)R}{4(1-\delta\epsilon)-(4-\delta\epsilon)(1-q_{t+1})}$. Here P may be randomized. Recall that $OPT(\mathcal{I}_1) = \frac{R}{1-\delta\epsilon}$. If with probability $\frac{1}{2}$ setting the price $P \leq P^*$, then for instance \mathcal{I}_1 , we have

$$\begin{aligned} Rev(\mathcal{M}(\mathcal{I}_1)) &\leq \frac{1}{2}OPT(\mathcal{I}_1) + \frac{1}{2}\left(\frac{3R}{4(1-\delta\epsilon)} + \frac{R}{4}\right) \\ &= \frac{7R}{8(1-\delta\epsilon)} + \frac{R}{8} = \frac{R}{1-\delta\epsilon} \left(1 - \frac{1}{8}\delta\epsilon\right) < \frac{OPT(\mathcal{I}_1)}{1+4\epsilon} \end{aligned}$$

when $\delta \geq 32$. On the other hand, recall that $OPT(\mathcal{I}_2) = \frac{R}{2(1-\delta\epsilon)} + \frac{R}{2} = \frac{(2-\delta\epsilon)R}{2(1-\delta\epsilon)}$. If with probability $\frac{1}{2}$, the price $P > P^*$, for instance \mathcal{I}_2 , we have

$$\begin{aligned} Rev(\mathcal{M}(\mathcal{I}_2)) &< \frac{1}{2}OPT(\mathcal{I}_2) + \frac{(4-\delta\epsilon)R}{2(4-2\delta\epsilon)} = \frac{(2-\delta\epsilon)R}{4(1-\delta\epsilon)} + \frac{(4-\delta\epsilon)R}{2(4-2\delta\epsilon)} \\ &= \frac{(2-\delta\epsilon)R}{2(1-\delta\epsilon)} \left(\frac{1}{2} + \frac{(4-\delta\epsilon)(1-\delta\epsilon)}{2(2-\delta\epsilon)^2}\right) = OPT(\mathcal{I}_2) \left(1 - \frac{\delta\epsilon}{2(2-\delta\epsilon)^2}\right) < \frac{OPT(\mathcal{I}_2)}{1+4\epsilon} \end{aligned}$$

when $\delta \geq 32$. Thus for any mechanism \mathcal{M} with $O(\frac{1}{\epsilon})$ quantile queries, there exists $z^* \in \{0, 1\}$ such that when $\epsilon < \frac{1}{64}$ and $\delta = 32$,

$$Rev(\mathcal{M}(\mathcal{I}_{z^*})) \leq \frac{OPT(\mathcal{I}_{z^*})}{2} + \frac{OPT(\mathcal{I}_{z^*})}{2(1+4\epsilon)} < \frac{OPT(\mathcal{I}_{z^*})}{1+\epsilon}.$$

Therefore Lemma 5 holds. \square

6.2 Upper Bound

Our mechanism \mathcal{M}_{EMR} (i.e., ‘‘Efficient quantile Myerson mechanism for Regular distributions’’) first constructs the distribution \mathcal{D}' that approximates \mathcal{D} using the quantile-query algorithm \mathcal{A}_Q with

parameters $\delta = \frac{\epsilon}{4}$ and $\epsilon_1 = \frac{\epsilon^2}{256n}$; and then runs Myerson's mechanism \mathcal{M}_{MRS} on \mathcal{D}' . Formally, we have the following theorem.

Theorem 10. $\forall \epsilon \in (0, 1)$, and for any single-item instance $\mathcal{I} = (N, M, \mathcal{D})$ where \mathcal{D} is regular, mechanism \mathcal{M}_{EMR} is DSIC, has query complexity $O(n \log_{1+\epsilon} \frac{n}{\epsilon})$, and $Rev(\mathcal{M}_{EMR}(\mathcal{I})) \geq \frac{OPT(\mathcal{I})}{1+\epsilon}$.

Mechanism 10 Efficient quantile Myerson mechanism for regular distributions, \mathcal{M}_{EMR}

- 1: Given $\epsilon > 0$, run algorithm \mathcal{A}_Q with $\delta = \frac{\epsilon}{4}$ and $\epsilon_1 = \frac{\epsilon^2}{256n}$ for each player i 's distribution \mathcal{D}_i , with the returned distribution denoted by \mathcal{D}'_i . Let $\mathcal{D}' = \times_{i \in N} \mathcal{D}'_i$.
 - 2: Run \mathcal{M}_{MRS} with \mathcal{D}' and the players' reported values, $b = (b_i)_{i \in N}$, to get allocation $x = (x_i)_{i \in N}$ and price profile $p = (p_i)_{i \in N}$ as the outcome.
-

Proof. Consider the quantile value $q^* = \frac{\epsilon^2}{256n}$ and $v_i^* = F_i^{-1}(1 - q^*)$. Let $\hat{v}_i = \max\{v_i^*, \frac{16OPT(\mathcal{I})}{\epsilon}\}$, and $\bar{\mathcal{D}}_1, \dots, \bar{\mathcal{D}}_n$ be imaginary distributions obtained by truncating $\mathcal{D}_1, \dots, \mathcal{D}_n$ at \hat{v}_i (i.e., a sample \bar{v}_i from $\bar{\mathcal{D}}_i$ is obtained by first sampling v_i from \mathcal{D}_i and then rounding down to $\bar{v}_i = \min\{v_i, \hat{v}_i\}$). Finally, denote by $\bar{\mathcal{I}} = (N, M, \bar{\mathcal{D}})$ the imaginary Bayesian instance where players' values are drawn from $\bar{\mathcal{D}}$.

Note that \mathcal{D}' is also a discretization distribution for $\bar{\mathcal{D}}$, following the proof and notations of Theorem 6, letting v_i^- be the value first sampled from $\bar{\mathcal{D}}_i$ then rounding down to the support of \mathcal{D}' , we have \mathcal{M}_{EMR} is truthful and using the technique of Mechanism 7, we have

$$\begin{aligned}
Rev(\mathcal{M}_{EMR}(\mathcal{I})) &= Rev(\mathcal{M}_{MRS}(v, \mathcal{D}')) \geq Rev(\mathcal{M}_{MRS}(v', \mathcal{D}')) \\
&\geq \sum_i \mathbb{E}_{\bar{v}_i \sim \bar{\mathcal{D}}_i} p_i(\bar{v}_i; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [v_i^- \geq p_i(\bar{v}_i; \bar{\mathcal{D}})] \\
&= \sum_i \mathbb{E}_{\bar{v}_i \sim \bar{\mathcal{D}}_i} p_i(\bar{v}_i; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [v_i^- \geq p_i(\bar{v}_i; \bar{\mathcal{D}})] \cdot (\mathbf{I}_{v_i^* \leq \frac{16OPT(\mathcal{I})}{\epsilon}} + \mathbf{I}_{v_i^* > \frac{16OPT(\mathcal{I})}{\epsilon}}). \tag{29}
\end{aligned}$$

We bound the indicators separately.

$$\begin{aligned}
&\sum_i \mathbb{E}_{\bar{v}_i \sim \bar{\mathcal{D}}_i} p_i(\bar{v}_i; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [v_i^- \geq p_i(\bar{v}_i; \bar{\mathcal{D}})] \cdot \mathbf{I}_{v_i^* \leq \frac{16OPT(\mathcal{I})}{\epsilon}} \\
&= \sum_i \mathbb{E}_{\bar{v}_i \sim \bar{\mathcal{D}}_i} p_i(\bar{v}_i; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [v_i^- \geq p_i(\bar{v}_i; \bar{\mathcal{D}})] \\
&\quad \cdot \mathbf{I}_{v_i^* \leq \frac{16OPT(\mathcal{I})}{\epsilon}} \cdot (\mathbf{I}_{p_i(\bar{v}_i; \bar{\mathcal{D}}) < v_i^*} + \mathbf{I}_{p_i(\bar{v}_i; \bar{\mathcal{D}}) \geq v_i^*}) \\
&\geq \sum_i \mathbb{E}_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [p_i(\bar{v}_i; \bar{\mathcal{D}}) \cdot \frac{1}{1 + \frac{\epsilon}{4}} \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_i; \bar{\mathcal{D}})] \cdot \mathbf{I}_{v_i^* \leq \frac{16OPT(\mathcal{I})}{\epsilon}} \cdot \mathbf{I}_{p_i(\bar{v}_i; \bar{\mathcal{D}}) < v_i^*} \\
&\quad + (p_i(\bar{v}_i; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_i; \bar{\mathcal{D}})] - \frac{16OPT(\mathcal{I})}{\epsilon} \cdot \frac{\epsilon^2}{256n}) \cdot \mathbf{I}_{v_i^* \leq \frac{16OPT(\mathcal{I})}{\epsilon}} \cdot \mathbf{I}_{p_i(\bar{v}_i; \bar{\mathcal{D}}) \geq v_i^*}] \\
&\geq \frac{1}{1 + \frac{\epsilon}{4}} \sum_i \mathbb{E}_{\bar{v}_i \sim \bar{\mathcal{D}}_i} p_i(\bar{v}_i; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_i; \bar{\mathcal{D}})] \cdot \mathbf{I}_{v_i^* \leq \frac{OPT(\mathcal{I})}{16\epsilon}} - \frac{\epsilon}{16} \cdot OPT(\mathcal{I}). \tag{30}
\end{aligned}$$

The first inequality here holds because for price $p_i(\bar{v}_i; \bar{\mathcal{D}}) < v_i^*$, we have

$$\Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [v_i^- \geq p_i(\bar{v}_i; \bar{\mathcal{D}})] \geq \frac{1}{1 + \frac{\epsilon}{4}} \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_i; \bar{\mathcal{D}})]$$

due to the structure of the quantile queries for \mathcal{D}' . For price $p_i(\bar{v}_{-i}; \bar{\mathcal{D}}) \geq v_i^*$, when $v_i^* \leq \frac{16OPT(\mathcal{I})}{\epsilon}$, by the regularity of \mathcal{D}_i , the optimal reserve corresponds to the quantile interval $(\frac{\epsilon^2}{256n}, 1]$. Thus we have

$$\begin{aligned} & p_i(\bar{v}_{-i}; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [v_i^- \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}})] \geq 0 \\ & \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}})] - v_i^* \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq v_i^*] \\ & \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}})] - \frac{16OPT(\mathcal{I})}{\epsilon} \cdot \frac{\epsilon^2}{256n} \end{aligned}$$

since the expected revenue is non-decreasing for quantile range $[0, \frac{\epsilon^2}{256n}]$. Thus Equation 30 holds. Then for the second indicator for Equation 29, we have

$$\begin{aligned} & \sum_i \mathbb{E}_{\bar{v}_{-i} \sim \bar{\mathcal{D}}_{-i}} p_i(\bar{v}_{-i}; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [v_i^- \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}})] \cdot \mathbf{I}_{v_i^* > \frac{16OPT(\mathcal{I})}{\epsilon}} \\ & \geq \frac{1}{1 + \frac{\epsilon}{4}} \sum_i \mathbb{E}_{\bar{v}_{-i} \sim \bar{\mathcal{D}}_{-i}} p_i(\bar{v}_{-i}; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}})] \cdot \mathbf{I}_{v_i^* > \frac{16OPT(\mathcal{I})}{\epsilon}} \end{aligned} \quad (31)$$

also by the construction the quantile queries for \mathcal{D}' . Combining Equation 29, 30 and 31, we have

$$\begin{aligned} & Rev(\mathcal{M}_{EMR}(\mathcal{I})) \\ & \geq \frac{1}{1 + \frac{\epsilon}{4}} \sum_i \mathbb{E}_{\bar{v}_{-i} \sim \bar{\mathcal{D}}_{-i}} p_i(\bar{v}_{-i}; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}})] \cdot \mathbf{I}_{v_i^* \leq \frac{16OPT(\mathcal{I})}{\epsilon}} - \frac{\epsilon}{16} \cdot OPT(\mathcal{I}) \\ & \quad + \frac{1}{1 + \frac{\epsilon}{4}} \sum_i \mathbb{E}_{\bar{v}_{-i} \sim \bar{\mathcal{D}}_{-i}} p_i(\bar{v}_{-i}; \bar{\mathcal{D}}) \cdot \Pr_{\bar{v}_i \sim \bar{\mathcal{D}}_i} [\bar{v}_i \geq p_i(\bar{v}_{-i}; \bar{\mathcal{D}})] \cdot \mathbf{I}_{v_i^* > \frac{16OPT(\mathcal{I})}{\epsilon}} \\ & = \frac{1}{1 + \frac{\epsilon}{4}} Rev(\mathcal{M}_{MRS}(\bar{v}, \bar{\mathcal{D}})) - \frac{\epsilon}{16} \cdot OPT(\mathcal{I}) \end{aligned}$$

By Lemma 2 of [18], for $0 \leq \epsilon \leq 1$, $Rev(\mathcal{M}_{MRS}(\bar{v}, \bar{\mathcal{D}})) = OPT(\bar{\mathcal{I}}) \geq (1 - \frac{\epsilon}{4})OPT(\mathcal{I})$. Thus we have

$$\begin{aligned} & Rev(\mathcal{M}_{EMR}(\mathcal{I})) \\ & \geq \frac{1}{1 + \frac{\epsilon}{4}} (1 - \frac{\epsilon}{4})OPT(\mathcal{I}) - \frac{\epsilon}{16} \cdot OPT(\mathcal{I}) \geq \frac{1}{1 + \epsilon} OPT(\mathcal{I}). \end{aligned}$$

Thus Theorem 10 holds. \square

Following [15, 28, 18], the sample complexity for single-item auction with regular distributions is bounded between $\Omega(\max\{n\epsilon^{-1}, \epsilon^{-3}\})$ and $\tilde{O}(n\epsilon^{-4})$. However, each sample is a valuation profile of the players, and thus contains n values. When ϵ is small, the query complexity in this setting is $O(n\epsilon^{-1} \log \frac{n}{\epsilon})$. Thus the query complexity is still much lower than the sample complexity.

7 Applications: Sampling Mechanisms

Using our techniques for query complexity, we can easily construct sampling mechanisms for multi-parameter auctions. Currently, the sample complexity for unit-demand auctions and additive distributions has been upper-bounded in [3, 30, 21, 6] for bounded auctions. In this section, we provide

another way to explicitly construct sampling mechanisms for both unit-demand and additive auctions, for arbitrary distributions with small-tails (and for bounded distributions).

The idea is to use samples to approximate *quantile* queries. Mechanism 11 defines our sampling mechanism \mathcal{M}_{SM} . Recall that mechanisms \mathcal{M}_{MRS} , \mathcal{M}_{UD} and \mathcal{M}_A are known (approximately) optimal DSIC mechanisms for single-item, unit-demand and additive auctions respectively. Note that in mechanism \mathcal{M}_{SM} , we use a different method to discretize the quantile space for additive auctions, so as to further reduce its sample complexity. In particular, we have the following theorem.

Mechanism 11 Sampling Mechanism \mathcal{M}_{SM}

- 1: For single-item auctions and unit-demand auctions, given $\epsilon > 0$, set $\delta = \frac{\epsilon}{6}$, $\epsilon_1 = h(\frac{2\epsilon}{3(1+\epsilon)})$ and $k = \lceil \log_{1+\delta} \frac{1}{\epsilon_1} \rceil$; define the *quantile vector* as $q = (q_0, q_1, \dots, q_{k-1}, q_k) = (1, \epsilon_1(1 + \delta)^{k-1}, \dots, \epsilon_1(1 + \delta), \epsilon_1)$.
For additive auctions, given $\epsilon > 0$, set $\epsilon_1 = h(\frac{\epsilon}{10(1+\epsilon)})$ and $k = \lfloor \frac{1}{\epsilon_1} \rfloor$; define the *quantile vector* as $q = (q_0, q_1, \dots, q_{k-1}, q_k) = (1, k\epsilon_1, \dots, 2\epsilon_1, \epsilon_1)$.
 - 2: For each player i and item j , given t samples $V_{ij}^t = \{v_{ij}^1, \dots, v_{ij}^t\}$, without loss of generality assume $v_{ij}^1 \geq v_{ij}^2 \geq \dots \geq v_{ij}^t$. For each quantile q_l , set $v_{ij}^{tq_l}$ to be the value corresponding to the quantile query q_l . (If tq_l is not an integer then the mechanism takes $\lceil tq_l \rceil$.)
 - 3: Construct a discrete distribution \mathcal{D}'_{ij} as follows: $\mathcal{D}'_{ij}(v_{ij}^{tq_l}) = q_l - q_{l+1}$ for each $l \in \{0, \dots, k-1\}$, and $\mathcal{D}'_{ij}(v_{ij}^{tq_k}) = \epsilon_1$. Finally, let $\mathcal{D}'_i = \times_{j \in M} \mathcal{D}'_{ij}$ for each player i and let $\mathcal{D}' = \times_{i \in N} \mathcal{D}'_i$.
 - 4: Run $\mathcal{M}_{MRS}/\mathcal{M}_{UD}/\mathcal{M}_A$ with distribution \mathcal{D}' and the players' reported values.
-

Theorem 11. $\forall \epsilon > 0$ and $\gamma \in (0, 1)$, for any Bayesian instance $\mathcal{I} = (N, M, \mathcal{D})$,

- for single-item auctions satisfying the Small-Tail Assumption 2, with $\tilde{O}(h^{-2}(\frac{2\epsilon}{3(1+\epsilon)}) \cdot (\frac{\epsilon}{1+\epsilon})^{-2})$ samples, mechanism \mathcal{M}_{SM} achieves revenue at least $\frac{1}{1+\epsilon} OPT(\mathcal{I})$ with probability at least $1 - \gamma$;
- for unit-demand auctions satisfying the Small-Tail Assumption 2, with $\tilde{O}(h^{-2}(\frac{2\epsilon}{3(1+\epsilon)}) \cdot (\frac{\epsilon}{1+\epsilon})^{-2})$ samples, mechanism \mathcal{M}_{SM} achieves revenue at least $\frac{1}{24(1+\epsilon)} OPT$ with probability at least $1 - \gamma$;
- for additive auctions satisfying the Small-Tail Assumption 1, with $\tilde{O}(h^{-2}(\frac{\epsilon}{10(1+\epsilon)}) (\frac{1}{2} - \frac{1}{1+(1+\frac{\epsilon}{6})^{1/m}}))^{-2})$ samples, mechanism \mathcal{M}_{SM} achieves revenue at least $\frac{1}{8(1+\epsilon)} OPT$ with probability at least $1 - \gamma$.

Proof. After constructing the distributions, we simply run the existing DSIC mechanisms as a Blackbox, and if the constructed distribution satisfies the property that for any quantile q_l ,

$$q_{ij}(v_{ij}^{tq_{l+1}}) \geq \frac{1}{1 + \frac{\epsilon}{3}} \left(q_{ij}(v_{ij}^{tq_l}) \right). \quad (32)$$

all our query complexity results for single-item and unit-demand auctions directly apply here.

Since here for sampling mechanism, we slice the quantile interval uniformly, in the ideal case, the selected sampled values correspond to the desired quantiles and $\mathcal{D}_{ij}(v_{ij}^{tq_l}) = \mathcal{D}_{ij}(v_{ij}^{tq_{l+1}})$. However, since these samples are random, we may not obtain the ideal case. In fact, given parameter $d = \frac{12+3\epsilon}{\epsilon}$, if for any quantile q_l ,

$$q_l - \frac{q_l}{d} \leq q_{ij}(v_{ij}^{tq_l}) \leq q_l + \frac{q_l}{d}, \quad (33)$$

then

$$\frac{q_{ij}(v_{ij}^{tq_{l+1}})}{q_{ij}(v_{ij}^{tq_l})} \geq \frac{q_{l+1}(1 - \frac{1}{d})}{q_l(1 + \frac{1}{d})} \geq \frac{\frac{1}{1+\frac{\epsilon}{6}}(1 - \frac{1}{d})}{1 + \frac{1}{d}} = \frac{1}{1 + \frac{\epsilon}{3}},$$

for any $\epsilon > 0$, that is, Equation 32 holds. In the following, we show how many samples are enough to obtain Inequality 33.

First, we bound the probability that $v_{ij}^{t \cdot q_l}$ locates in the quantile interval $[q_l - \frac{q_l}{d}, q_l + \frac{q_l}{d}]$. Let $E_{ij,l}^{left}$ be the event that $v_{ij}^{t \cdot q_l}$ locates in the quantile interval $[0, q_l - \frac{q_l}{d}]$, and $E_{ij,l}^{right}$ be the event that $v_{ij}^{t \cdot q_l}$ locates in the quantile interval $[q_l + \frac{q_l}{d}, 1]$. Then

$$\Pr[E_{ij,l}^{left}] = \sum_{s=0}^{t \cdot q_l} \binom{t}{s} \left(q_l - \frac{q_l}{d}\right)^s \left(1 - q_l + \frac{q_l}{d}\right)^{t-s},$$

and

$$\Pr[E_{ij,l}^{right}] = \sum_{s=0}^{t \cdot q_l} \binom{t}{s} \left(1 - q_l - \frac{q_l}{d}\right)^s \left(q_l + \frac{q_l}{d}\right)^{t-s}.$$

By Chernoff's inequality and $\forall i, j, l$, letting $\Pr[E_{ij,l}^{left}]$ and $\Pr[E_{ij,l}^{right}]$ be no more than $\frac{\gamma}{2mn(k+1)}$, $t = \tilde{O}((\frac{\epsilon}{d})^{-2}) = \tilde{O}((\frac{\epsilon \cdot \epsilon}{1+\epsilon})^{-2})$. That is with $\tilde{O}(h^{-2}(\frac{2\epsilon}{3(1+\epsilon)}) \cdot (\frac{\epsilon}{1+\epsilon})^{-2})$ samples, the probability that $v_{ij}^{t \cdot q_l}$ does not locate in the quantile interval $[q_l - \frac{q_l}{d}, q_l + \frac{q_l}{d}]$ is less than $\frac{\gamma}{mn(k+1)}$. By union bound, there exists one $v_{ij}^{t \cdot q_l}$ for all $i \in [n], j \in [m], l \in [k+1]$ does not locate in the quantile interval $[q_l - \frac{q_l}{d}, q_l + \frac{q_l}{d}]$ is less than γ . Then with probability $1 - \gamma$, Inequality 33 holds.

For additive auctions, if the constructed distribution satisfies the property that for any quantile q_l ,

$$q_{ij}(v_{ij}^{t \cdot q_l}) - q_{ij}(v_{ij}^{t \cdot q_{l+1}}) \geq \frac{1}{(1 + \frac{\epsilon}{5})^{1/m}} \left(q_{ij}(v_{ij}^{t \cdot q_{l+1}}) - q_{ij}(v_{ij}^{t \cdot q_{l+2}})\right), \quad (34)$$

all our query complexity results for additive auctions directly apply here. In fact, if for any quantile q_l ,

$$q_l - \epsilon_1 \left(\frac{1}{2} - \frac{1}{1 + (1 + \frac{\epsilon}{5})^{1/m}}\right) \leq q_{ij}(v_{ij}^{t \cdot q_l}) \leq q_l + \epsilon_1 \left(\frac{1}{2} - \frac{1}{1 + (1 + \frac{\epsilon}{5})^{1/m}}\right),$$

then,

$$\frac{q_{ij}(v_{ij}^{t \cdot q_l}) - q_{ij}(v_{ij}^{t \cdot q_{l+1}})}{q_{ij}(v_{ij}^{t \cdot q_{l+1}}) - q_{ij}(v_{ij}^{t \cdot q_{l+2}})} \geq \frac{\epsilon_1 - \epsilon_1 \left(1 - \frac{2}{1 + (1 + \frac{\epsilon}{5})^{1/m}}\right)}{\epsilon_1 + \epsilon_1 \left(1 - \frac{2}{1 + (1 + \frac{\epsilon}{5})^{1/m}}\right)} = \frac{1}{(1 + \frac{\epsilon}{5})^{1/m}},$$

Using the same technique of applying the Chernoff's inequality, with $\tilde{O}(h^{-2}(\frac{\epsilon}{10(1+\epsilon)})(\frac{1}{2} - \frac{1}{1 + (1 + \frac{\epsilon}{5})^{1/m}}))^{-2}$ samples, Equation 34 holds with probability $1 - \gamma$. Thus Theorem 11 holds. \square

Remark. Following the convention in the literature, a logarithmic factor depending on γ has been absorbed in $\tilde{O}(\cdot)$. If the values are bounded in $[1, H]$, by defining the tail function h according to H , the resulting sample complexity is $\tilde{O}(m^4 n^2 H^2 (1 + \epsilon)^4 \epsilon^{-4})$ for unit-demand auctions and $\tilde{O}(m^4 n^2 H^2 (\frac{1+\epsilon}{\epsilon})^2 (\frac{1}{2} - \frac{1}{1 + (1 + \frac{\epsilon}{5})^{1/m}}))^{-2})$ for additive auctions.

We note that our resulting sampling mechanisms are not better than the best-known sample complexity results as shown in Table 2. However, our mechanisms are able to deal with unbounded distributions satisfy small-tail assumptions.

A Missing Proofs for Section 5

We now prove the claims used above.

Claim 1 (restated). $\Pr_{v_i \sim \mathcal{D}_i}[v_i \geq p_i(v_{-i}; \mathcal{D}) | q_i(v_i) > \epsilon_1] \leq (1 + \delta) \Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq p_i(v_{-i}; \mathcal{D})]$.

Proof. By definition, $q_i(v_i) > \epsilon_1$ implies $v_i \leq v'_{i;k}$, where $v'_{i;k}$ is the largest value in V'_i , the support of distribution \mathcal{D}'_i . Note that $v_i^- \leq v'_{i;k}$ for any v_i . If $p_i(v_{-i}; \mathcal{D}) > v'_{i;k}$, then both probabilities are 0 and the inequality holds.

Below we consider the case $p_i(v_{-i}; \mathcal{D}) \leq v'_{i;k}$. Let $v'_{i;-1} = -1$ and $l \in \{0, 1, \dots, k\}$ be such that $v'_{i;l} \geq p_i(v_{-i}; \mathcal{D})$ and $v'_{i;l-1} < p_i(v_{-i}; \mathcal{D})$. We have

$$\begin{aligned}
& \Pr_{v_i \sim \mathcal{D}_i}[v_i \geq p_i(v_{-i}; \mathcal{D}) | q_i(v_i) > \epsilon_1] \\
\leq & \Pr_{v_i \sim \mathcal{D}_i}[v_i^+ \geq p_i(v_{-i}; \mathcal{D}) | q_i(v_i) > \epsilon_1] \\
= & \Pr_{v_i \sim \mathcal{D}_i}[v_i^+ \geq v'_{i;l} | q_i(v_i) > \epsilon_1] \\
= & \Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq v'_{i;l-1} | q_i(v_i) > \epsilon_1] \\
= & \Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq v'_{i;\max\{0, l-1\}} | q_i(v_i) > \epsilon_1] \\
= & \frac{\Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq v'_{i;\max\{0, l-1\}} \text{ and } q_i(v_i) > \epsilon_1]}{\Pr_{v_i \sim \mathcal{D}_i}[q_i(v_i) > \epsilon_1]} \\
= & \frac{\Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq v'_{i;\max\{0, l-1\}}] - \Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq v'_{i;\max\{0, l-1\}} \text{ and } q_i(v_i) \leq \epsilon_1]}{\Pr_{v_i \sim \mathcal{D}_i}[q_i(v_i) > \epsilon_1]} \\
= & \frac{\Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq v'_{i;\max\{0, l-1\}}] - \Pr_{v_i \sim \mathcal{D}_i}[q_i(v_i) \leq \epsilon_1]}{\Pr_{v_i \sim \mathcal{D}_i}[q_i(v_i) > \epsilon_1]} \\
\leq & \Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq v'_{i;\max\{0, l-1\}}] \\
= & \Pr_{v'_i \sim \mathcal{D}'_i}[v'_i \geq v'_{i;\max\{0, l-1\}}] = q_{\max\{0, l-1\}} \leq (1 + \delta)q_l \\
= & (1 + \delta) \Pr_{v'_i \sim \mathcal{D}'_i}[v'_i \geq v'_{i;l}] = (1 + \delta) \Pr_{v'_i \sim \mathcal{D}'_i}[v'_i \geq p_i(v_{-i}; \mathcal{D})] \\
= & (1 + \delta) \Pr_{v_i \sim \mathcal{D}_i}[v_i^- \geq p_i(v_{-i}; \mathcal{D})],
\end{aligned}$$

as desired. Indeed, the first inequality is because $v_i^+ > v_i$, and the first equality is because $v_i^+ \in V'_i \cup \{+\infty\}$ and thus $v_i^+ \geq p_i(v_{-i}; \mathcal{D})$ if and only if $v_i^+ \geq v'_{i;l}$. Similarly, the second equality is because (v_i^-, v_i^+) and $(v'_{i;l-1}, v'_{i;l})$ are two pairs of consecutive values in $V'_i \cup \{-1, +\infty\}$, thus $v_i^+ \geq v'_{i;l}$ if and only if $v_i^- \geq v'_{i;l-1}$. The third equality is because $v_i^- \geq v'_{i;0}$ always. The sixth equality is because $q_i(v_i) \leq \epsilon_1$ implies $v_i \geq v'_{i;k} \geq v'_{i;l}$, thus $v_i^- \geq v'_{i;\max\{0, l-1\}}$. The seventh equality is by the definition of the round-down scheme. The following two equalities and the inequality are by the construction of \mathcal{D}'_i and the definition of the quantile vector q . Indeed, $(1 + \delta)q_0 = 1 + \delta > 1 = q_0$, $(1 + \delta)q_1 = \epsilon_1(1 + \delta)^k \geq \epsilon_1(1 + \delta)^{\log_{1+\delta} \frac{1}{\epsilon_1}} = \epsilon_1 \cdot \frac{1}{\epsilon_1} = 1 = q_0$, and $(1 + \delta)q_l = q_{l-1}$ for any $l \geq 2$. The second-last equality is because $v'_i \in V'_i$, thus $v'_i \geq v'_{i;l}$ if and only if $v'_i \geq p_i(v_{-i}; \mathcal{D})$. Finally, the last equality is again by the definition of the round-down scheme. \square

Claim 2 (restated). $Rev(\mathcal{M}^*(\mathcal{I}^l)) \geq \frac{1}{1+\epsilon} OPT(\mathcal{I})$.

Proof. Combining Equation 18 and Claim 1, we have

$$Rev(\mathcal{M}^*(\mathcal{I}')) \geq \frac{1}{1+\delta} \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} p_i(v_{-i}; \mathcal{D}) \cdot \Pr_{v_i \sim \mathcal{D}_i} [v_i \geq p_i(v_{-i}; \mathcal{D}) | q_i(v_i) > \epsilon_1].$$

Accordingly,

$$\begin{aligned} Rev(\mathcal{M}^*(\mathcal{I}')) &\geq \frac{1}{1+\delta} \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} p_i(v_{-i}; \mathcal{D}) \cdot \Pr_{v_i \sim \mathcal{D}_i} [v_i \geq p_i(v_{-i}; \mathcal{D}) | q_i(v_i) > \epsilon_1] \\ &\geq \frac{1}{1+\delta} \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} p_i(v_{-i}; \mathcal{D}) \cdot \Pr_{v_i \sim \mathcal{D}_i} [q_i(v_i) > \epsilon_1 \text{ and } v_i \geq p_i(v_{-i}; \mathcal{D})] \\ &= \frac{1}{1+\delta} \sum_i \mathbb{E}_{v_{-i} \sim \mathcal{D}_{-i}} \mathbb{E}_{v_i \sim \mathcal{D}_i} p_i(v_{-i}; \mathcal{D}) \cdot \mathbf{I}_{q_i(v_i) > \epsilon_1} \cdot \mathbf{I}_{v_i \geq p_i(v_{-i}; \mathcal{D})} \\ &= \frac{1}{1+\delta} \mathbb{E}_{v \sim \mathcal{D}} \sum_i p_i(v_{-i}; \mathcal{D}) \cdot \mathbf{I}_{q_i(v_i) > \epsilon_1} \cdot \mathbf{I}_{v_i \geq p_i(v_{-i}; \mathcal{D})} \\ &\geq \frac{1}{1+\delta} \mathbb{E}_{v \sim \mathcal{D}} \mathbf{I}_{\forall i, q_i(v_i) > \epsilon_1} \cdot \sum_i p_i(v_{-i}; \mathcal{D}) \mathbf{I}_{v_i \geq p_i(v_{-i}; \mathcal{D})} \\ &= \frac{1}{1+\delta} \mathbb{E}_{v \sim \mathcal{D}} \mathbf{I}_{\forall i, q_i(v_i) > \epsilon_1} \cdot Rev_{OPT}(v; \mathcal{I}) \\ &\geq \frac{1-\delta_1}{1+\delta} OPT(\mathcal{I}). \end{aligned} \tag{35}$$

Here the second last equality holds by the definition of $p_i(v_{-i}; \mathcal{D})$ and $Rev_{OPT}(v; \mathcal{I})$, and last inequality holds by the Small-Tail Assumption 2. Since $\delta = \frac{\epsilon}{3}$ and $\delta_1 = \frac{2\epsilon}{3(1+\epsilon)}$, we have

$$\frac{1-\delta_1}{1+\delta} = \frac{1}{1+\epsilon},$$

thus Claim 2 holds. □

References

- [1] P. Azar, C. Daskalakis, S. Micali, and S. M. Weinberg. Optimal and efficient parametric auctions. In *24th Symposium on Discrete Algorithms (SODA'13)*, pages 596–604, 2013.
- [2] P. Azar and S. Micali. Parametric digital auctions. In *4rd Innovations in Theoretical Computer Science Conference (ITCS'13)*, pages 231–232, 2013.
- [3] P. D. Azar, R. Kleinberg, and S. M. Weinberg. Prophet inequalities with limited information. In *25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'14)*, pages 1358–1377, 2014.
- [4] M. Babaioff, N. Immorlica, B. Lucier, and S. M. Weinberg. A simple and approximately optimal mechanism for an additive buyer. In *55th Annual IEEE Symposium on Foundations of Computer Science (FOCS'14)*, pages 21–30, 2014.
- [5] G. W. Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Weview*, 78(1):1–3, 1950.
- [6] Y. Cai and C. Daskalakis. Learning multi-item auctions with (or without) samples. In *58th Symposium on Foundations of Computer Science (FOCS'17)*, pages 516–527, 2017.
- [7] Y. Cai, C. Daskalakis, and S. M. Weinberg. An algorithmic characterization of multi-dimensional mechanisms. In *44th Annual ACM Symposium on Theory of Computing (STOC'12)*, pages 459–478, 2012.
- [8] Y. Cai, C. Daskalakis, and S. M. Weinberg. Optimal multi-dimensional mechanism design: Reducing revenue to welfare maximization. In *53rd Symposium on Foundations of Computer Science (FOCS'12)*, pages 130–139, 2012.
- [9] Y. Cai, N. R. Devanur, and S. M. Weinberg. A duality based unified approach to Bayesian mechanism design. In *48th Annual ACM Symposium on Theory of Computing (STOC'16)*, pages 926–939, 2016.
- [10] Y. Cai and M. Zhao. Simple mechanisms for subadditive buyers via duality. In *49th Symposium on Theory of Computing (STOC'17)*, pages 170–183, 2017.
- [11] J. Cervera and J. Munoz. Proper scoring rules for fractiles. *Bayesian Statistics*, 5:513–519, 1996.
- [12] S. Chawla, J. D. Hartline, D. L. Malec, and B. Sivan. Multi-parameter mechanism design and sequential posted pricing. In *43th ACM Symposium on Theory of Computing (STOC'10)*, pages 311–320, 2010.
- [13] S. Chawla and J. B. Miller. Mechanism design for subadditive agents via an ex ante relaxation. In *17th Conference on Economics and Computation (EC'16)*, pages 579–596. ACM, 2016.
- [14] J. Chen, B. Li, and Y. Li. From Bayesian to crowdsourced Bayesian auctions. *arXiv:1702.01416*, 2016.
- [15] R. Cole and T. Roughgarden. The sample complexity of revenue maximization. In *46th Annual ACM Symposium on Theory of Computing (STOC'14)*, pages 243–252, 2014.

- [16] J. Cremer and R. P. McLean. Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–1257, 1988.
- [17] C. Daskalakis, A. Deckelbaum, and C. Tzamos. Mechanism design via optimal transport. In *14th Conference on Electronic Commerce (EC'13)*, pages 269–286, 2013.
- [18] N. R. Devanur, Z. Huang, and C.-A. Psomas. The sample complexity of auctions with side information. In *48th Annual ACM Symposium on Theory of Computing (STOC'16)*, pages 426–439, 2016.
- [19] P. Dhangwatnotai, T. Roughgarden, and Q. Yan. Revenue maximization with a single sample. *Games and Economic Behavior*, 91:318–333, 2015.
- [20] S. Dobzinski. Computational efficiency requires simple taxation. In *57th Symposium on Foundations of Computer Science (FOCS'16)*, pages 209–218, 2016.
- [21] K. Goldner and A. R. Karlin. A prior-independent revenue-maximizing auction for multiple additive bidders. In *12th International Conference on Web and Internet Economics (WINE'16)*, pages 160–173, 2016.
- [22] Y. A. Gonczarowski and N. Nisan. Efficient empirical revenue maximization in single-parameter auction environments. In *49th Symposium on Theory of Computing (STOC'17)*, pages 856–868, 2017.
- [23] R. Guesnerie. On taxation and incentives: further remarks on the limits to redistribution. *University of Bonn*, 89, 1981.
- [24] P. J. Hammond. Straightforward individual incentive compatibility in large economies. *The Review of Economic Studies*, 46(2):263–282, 1979.
- [25] S. Hart and N. Nisan. Approximate revenue maximization with multiple items. In *13th ACM Conference on Electronic Commerce (EC'12)*, pages 656–656, 2012.
- [26] S. Hart and N. Nisan. The menu-size complexity of auctions. In *14th ACM Conference on Electronic Commerce (EC'13)*, pages 565–566, 2013.
- [27] S. Hart and P. J. Reny. Maximal revenue with multiple goods: Nonmonotonicity and other observations. *Theoretical Economics*, 10:893–922, 2015.
- [28] Z. Huang, Y. Mansour, and T. Roughgarden. Making the most of your samples. In *16th ACM Conference on Economics and Computation (EC'15)*, pages 45–60, 2015.
- [29] R. Kleinberg and S. M. Weinberg. Matroid prophet inequalities. In *44th Annual ACM Symposium on Theory of Computing (STOC'12)*, pages 123–136, 2012.
- [30] J. Morgenstern and T. Roughgarden. Learning simple auctions. In *29th Conference on Learning Theory (COLT'16)*, pages 1298–1318, 2016.
- [31] R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.
- [32] A. Ronen. On approximating optimal auctions. In *3rd ACM Conference on Electronic Commerce (EC'01)*, pages 11–17, 2001.

- [33] T. Roughgarden and O. Schrijvers. Ironing in the dark. In *17th ACM Conference on Economics and Computation (EC'16)*, pages 1–18, 2016.
- [34] A. Rubinstein and S. M. Weinberg. Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. In *16th ACM Conference on Economics and Computation (EC'15)*, pages 377–394, 2015.
- [35] A. C.-C. Yao. An n-to-1 bidder reduction for multi-item auctions and its applications. In *26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'15)*, pages 92–109, 2015.
- [36] A. C.-C. Yao. Dominant-strategy versus bayesian multi-item auctions: Maximum revenue determination and comparison. In *18th Conference on Economics and Computation (EC'17)*, pages 3–20, 2017.