SIAM J. COMPUT.
Vol. 47, No. 3, pp. 651–674

MAKING THE MOST OF YOUR SAMPLES*

ZHIYI HUANG[†], YISHAY MANSOUR[‡], AND TIM ROUGHGARDEN[§]

Abstract. We study the problem of setting a price for a potential buyer with a valuation drawn from an unknown distribution D. The seller has "data" about D in the form of $m \geq 1$ independent and identically distributed samples, and the algorithmic challenge is to use these samples to obtain expected revenue as close as possible to what could be achieved with advance knowledge of D. Our first set of results quantifies the number of samples m that are necessary and sufficient to obtain a $(1-\epsilon)$-approximation. For example, for an unknown distribution that satisfies the monotone hazard rate (MHR) condition, we prove that $\tilde{\Theta}(\epsilon^{-3/2})$ samples are necessary and sufficient. Remarkably, this uses fewer samples than is necessary to accurately estimate the expected revenue obtained for such a distribution by even a single reserve price. We also prove essentially tight sample complexity bounds for regular distributions, bounded-support distributions, and a wide class of irregular distributions. Our lower bound approach, which applies to all randomized pricing strategies, borrows tools from differential privacy and information theory, and we believe it could find further applications in auction theory. Our second set of results considers the single-sample case. While no deterministic pricing strategy is better than $\frac{1}{2}$-approximate for regular distributions, for MHR distributions we show how to do better: there is a simple deterministic pricing strategy that guarantees expected revenue at least 0.589 times the maximum possible. We also prove that no deterministic pricing strategy achieves an approximation guarantee better than $\frac{e}{4} \approx .68$.

Key words. pricing, auctions, sample complexity, information theory

AMS subject classifications. 68Q25, 91A99

DOI. 10.1137/16M1065719

1. Introduction.

We study the basic pricing problem of making an "optimal" take-it-or-leave-it price to a potential buyer with an unknown willingness to pay (also known as *valuation*). Offering a price of p to a buyer with valuation v yields revenue p if $v \geq p$, and 0 otherwise. The traditional approach in theoretical computer science to such problems is to assume as little as possible about the buyer's valuation—for example, only lower and upper bounds on its value—and to compare the performance of different prices using worst-case analysis. The traditional approach in economics is to assume that the buyer's valuation is drawn from a distribution D that is known to the seller, and to use average-case analysis. In the latter case, the optimal solution is clear—it is the *monopoly price* $\max_{p \geq 0} p \cdot (1 - F(p))$, where F is the cumulative distribution function (c.d.f.) of D.

Cole and Roughgarden [9] recently proposed adapting the formalism of learning theory [25] to interpolate between the traditional worst- and average-case approaches,

*Received by the editors March 14, 2016; accepted for publication (in revised form) January 22, 2018; published electronically May 8, 2018. A preliminary version [19] of this work appeared in the Sixteenth ACM Conference on Economics and Computation, ACM, New York, 2015, pp. 45–60.
http://www.siam.org/journals/sicomp/47-3/M106571.html
Funding: The second author was supported in part by The Israeli Centers of Research Excellence (I-CORE) program, (Center No. 4/11), by a grant from the Israel Science Foundation (ISF); by a grant from United States-Israel Binational Science Foundation (BSF), and by a grant from the Israeli Ministry of Science (MoS). The third author was supported in part by NSF grants CCF-1016885 and CCF-1215965, and an ONR PECASE Award.

[†]The University of Hong Kong, Pokfulam, Hong Kong (zhiyi@cs.hku.hk). This work was done while the author was at Stanford University.
[‡]Tel Aviv University and Microsoft Research, Hertzelia, Israel (mansour@tau.ac.il).
[§]Department of Computer Science, Stanford University, 474 Gates Building, 353 Serra Mall, Stanford, CA 94305 (tim@cs.stanford.edu).

in the context of single-item auction design. The idea is to parameterize a seller's knowledge about an unknown distribution D through a number m of independent and identically distributed (i.i.d.) samples from D. When $m = 0$ this is equivalent to the worst-case approach, and as $m \to \infty$ it becomes equivalent to the average-case approach. The benchmark is the maximum expected revenue obtainable when the distribution D is known a priori. The algorithmic challenge is to use the m samples from D to get expected revenue as close to this benchmark as possible, no matter what the underlying distribution D is.[1]

This "hybrid" model offers several benefits. First, it is a relatively faithful model of many realistic computer science applications, where data from the past is assumed to be a reasonable proxy for future inputs and guides the choice of an algorithm. For example, in Yahoo!'s keyword auctions, the choice of reserve prices is guided by past bid data in a natural way [21]. Second, the model is a potential "sweet spot" between worst-case and average-case analysis, inheriting much of the robustness of the worst-case model (since we demand guarantees for every underlying D) while allowing very good approximation guarantees with respect to a strong benchmark. Third, by analyzing the trade-offs between the number of samples m available from D and the best-possible worst-case approximation guarantee, the analysis framework implicitly quantifies the value of data (i.e., of additional samples). It becomes possible, for example, to make statements like "4 times as much data improves our revenue guarantee from 80% to 90%." Finally, proving positive results in this model involves rigorously justifying natural methods of incorporating past data into an algorithm, and this task is interesting in its own right.

1.1. Our results. Formally, we study a single seller of some good, and a single buyer with a private valuation v for the good drawn from an unknown distribution D. The seller has access to $m \geq 1$ i.i.d. samples $v_1, \ldots, v_m$ from D. The goal is to identify, among all *m-pricing strategies*—functions from a sample $v_1, \ldots, v_m$ to a price p—the strategy that has the highest expected revenue. The expectation here is over $m + 1$ i.i.d. draws from D—the samples $v_1, \ldots, v_m$ and the unknown valuation v of the buyer—and the randomness of the pricing strategy. The *approximation guarantee* of a pricing strategy $p(\cdot)$ for a set $\mathcal{D}$ of distributions is its worst-case (over $\mathcal{D}$) approximation of the (optimal) expected revenue obtained by the monopoly price:

$$\inf_{D \in \mathcal{D}} \frac{\mathbf{E}_{v_1,\ldots,v_m \sim D}[p(v_1, \ldots, v_m) \cdot (1 - F(p(v_1, \ldots, v_m)))]}{\max_p p(1 - F(p))} ,$$

where F is the c.d.f. of D.

We first describe our results that quantify the inherent trade-off between the number of samples m and the best-possible approximation guarantee of an m-pricing strategy; see also Table 1. Some restriction on the class $\mathcal{D}$ of allowable distributions is necessary for the existence of pricing strategies with any nontrivial approximation of the optimal expected revenue.[2] We give essentially tight bounds on the number of samples that are necessary and sufficient to achieve a target approximation of $1 - \epsilon$ for

[1] There are, of course, other ways one can parameterize partial knowledge about valuations. See, e.g., [2, 8] for alternative approaches.

[2] To appreciate this issue, consider all distributions that take on a value M^2 with probability $\frac{1}{M}$ and 0 with probability $1 - \frac{1}{M}$. The optimal price for such a distribution earns expected revenue at least M. It is not difficult to prove that, for every m, there is no way to use m samples to achieve near-optimal revenue for every such distribution—for sufficiently large M, all m samples are 0 with high probability and the algorithm has to resort to an uneducated guess for M.

Table 1

Sample complexity of a $(1-\epsilon)$-approximation. For bounded-support distributions, the support is a subset of $[1, H]$. For general distributions, the benchmark is the optimal revenue of prices with sale probability at least δ.

	Upper bound		Lower bound	
MHR	$O(\epsilon^{-3/2} \log \epsilon^{-1})$	(Thm. 3.2)	$\Omega(\epsilon^{-3/2})$	(Thm. 4.12)
Regular	$O(\epsilon^{-3} \log \epsilon^{-1})$	(see [11])	$\Omega(\epsilon^{-3})$	(Thm. 4.8)
General	$O(\delta^{-1}\epsilon^{-2} \log(\delta^{-1}\epsilon^{-1}))$	(Thm. 3.6)	$\Omega(\delta^{-1}\epsilon^{-2})$	(Thm. 4.6)
Bounded support	$O(H\epsilon^{-2} \log(H\epsilon^{-1}))$	(Thm. 3.7 and [4])	$\Omega(H\epsilon^{-2})$	(Thm. 4.7)

all of the choices of the class $\mathcal{D}$ that are common in auction theory. For example, when $\mathcal{D}$ is the set of distributions that satisfy the monotone hazard rate (MHR) condition,[3] we prove that $m = \Omega(\epsilon^{-3/2})$ samples are necessary and that $m = O(\epsilon^{-3/2} \log \epsilon^{-1})$ samples are sufficient to achieve an approximation guarantee of $1 - \epsilon$.[4] This bound holds more generally for the class of "α-strongly regular distributions" introduced in [9] (for fixed $\alpha > 0$). When $\mathcal{D}$ is the (larger) set of regular distributions,[5] we prove that the sample complexity is $\tilde{\Theta}(\epsilon^{-3})$. When $\mathcal{D}$ is the set of arbitrary distributions with support contained in $[1, H]$, the sample complexity is $\tilde{\Theta}(H\epsilon^{-2})$. We also give essentially optimal sample complexity bounds for distributions that are parameterized by the probability of a sale at the monopoly price; see section 1.2 for more discussion. On the upper bound side, our primary contribution is the bound for MHR and strongly regular distributions.[6] All of our lower bounds, which are information theoretic and apply to arbitrary randomized pricing strategies, are new.

Our second set of results considers the regime where the seller has only one sample ($m = 1$) and wants to use it in the optimal deterministic way.[7] Dhangwatnotai, Roughgarden, and Yan [11] observed that an elegant result from auction theory, the Bulow–Klemperer theorem on auctions versus negotiations [6], implies that the 1-sample pricing strategy $p(v) = v$ has an approximation guarantee of $\frac{1}{2}$ when $\mathcal{D}$ is the set of regular distributions.[8] It is not hard to prove that there is no better deterministic pricing strategy for this set of distributions. We show how to do better, however, when $\mathcal{D}$ is the smaller set of MHR distributions: a simple 1-pricing strategy of the form $p(v) = cv$ for some $c < 1$ has an approximation guarantee of 0.589. We also prove that no deterministic 1-pricing strategy is better than an $\frac{e}{4}$-approximation for MHR distributions, and that no continuously differentiable such strategy is better than a 0.677-approximation.

[3] D satisfies the monotone hazard condition if $\frac{f(v)}{1-F(v)}$ is nondecreasing; see section 2 for details.

[4] We suppress only universal constant factors, which do not depend on the specific distribution $D \in \mathcal{D}$. Such uniform sample complexity bounds are desirable because the valuation distribution is unknown. Law-of-large-numbers–type arguments do not generally give uniform bounds.

[5] D is regular if the "virtual valuation function" $v - \frac{1-F(v)}{f(v)}$ is nondecreasing; see section 2 for details.

[6] The upper bound for regular distributions was proved in [11] and the upper bound for bounded valuations can be deduced from [4].

[7] We offer the problems of determining the best randomized pricing strategy for $m = 1$ and the best way to use a small $m \geq 2$ number of samples as challenging and exciting directions for future work. See [14] for recent progress on these questions.

[8] Dhangwatnotai, Roughgarden, and Yan [11] observed this in the context of the design and analysis of prior-independent auctions. Plugging our better bounds for single-sample pricing strategies with MHR distributions into the framework of [11] immediately yields analogously better prior-independent mechanisms.

1.2. A few technical highlights. This section singles out a few of our results and techniques that seem especially useful or motivating for follow-up work. First, recall that we prove that $O(\epsilon^{-3/2} \log \epsilon^{-1})$ samples from an unknown MHR distribution—or more generally, an unknown α-strongly regular distribution [9]—are sufficient to achieve expected revenue at least $1 - \epsilon$ times that of the monopoly price. Remarkably, this is fewer than the $\approx \epsilon^{-2}$ samples necessary to accurately estimate the expected revenue obtained by even a single fixed price for such a distribution![9] In this sense, we prove that near-optimal revenue maximization is strictly easier than accurately learning even very simple statistics of the underlying distribution. The most important idea in our upper bound is that, because of the structure of the revenue-maximization problem, the estimation errors of different competing prices are usefully correlated. For example, if the estimated expected revenue of the true monopoly price is significantly less than its actual expected revenue (because of a higher-than-expected number of low samples), then this probably also holds for prices that are relatively close to the monopoly price. Moreover, these are precisely the incorrect prices that an algorithm is most likely to choose by mistake. The second ingredient is the fact that MHR distributions have strongly concave "revenue curves," and this limits how many distinct prices can achieve expected revenue close to that of the monopoly price.

Second, recall that we prove essentially matching lower bounds for all of our sample complexity upper bounds. For example, there is no $(1 - \epsilon)$-approximate pricing strategy (deterministic or randomized) for MHR distributions when $m = o(\epsilon^{-3/2})$ or for regular distributions when $m = o(\epsilon^{-3})$. For both of these lower bounds, we reduce the existence of a $(1 - \epsilon)$-optimal pricing strategy to that of a classifier that distinguishes between two similar distributions. We borrow methodology from the differential privacy literature to construct two distributions with small Kullback–Leibler (KL) divergence and disjoint sets of near-optimal prices, and use Pinsker's inequality to derive the final sample complexity lower bounds. This lower bound approach is novel in the context of auction theory and we expect it to find further applications.

Third, we offer a simple and novel approach for reasoning about irregular distributions. We noted above the problematic irregular distributions that place a very low probability on a very high value. Regularity can also fail for more reasonable distributions, such as mixtures of common distributions. In section 3.3, we consider a benchmark R^*_δ defined as the maximum expected revenue achievable for the underlying distribution using a price that sells with probability at least δ, and prove essentially tight sample complexity bounds for approximating this benchmark. As a special case, if for every distribution in $\mathcal{D}$ the monopoly price sells with probability at least δ—as is the case for sufficiently small δ and typical reasonable distributions, even irregular ones—then approximating R^*_δ is equivalent to approximating the optimal revenue. Even if not all distributions of $\mathcal{D}$ satisfy this property, this benchmark enables parameterized sample complexity bounds that do not require blanket distributional restrictions such as regularity. We believe that this parameterized approach will find more applications.[10]

[9] It is well known (e.g., [1, Lemma 5.1]) that, given a coin that either has bias $\frac{1}{2} - \epsilon$ or bias $\frac{1}{2} + \epsilon$, $\Omega(\epsilon^{-2} \log \frac{1}{\delta})$ coin flips are necessary to distinguish between the two cases with probability at least $1 - \delta$. This lower bound is information theoretic and applies to arbitrarily sophisticated learning methods. This sample complexity lower bound translates straightforwardly to the problem of estimating, by any means, the expected revenue of a fixed price for an unknown MHR distribution up to a factor of $(1 \pm \epsilon)$.

[10] See, e.g., [18, Appendix D], [16, Chapter 4], and [24] for alternative approaches to parameterizing irregularity.

TABLE 2
Optimal approximation ratio with a single sample.

	Positive result	Negative result
Regular	≥ 0.5 (see [6, 11])	≤ 0.5 (Thm. 6.1)
MHR	≥ 0.589 (Thm. 5.2)	≤ 0.68 (Thm. 6.3)

1.3. Further related work. We already mentioned the related work of Cole and Roughgarden [9]; the present work follows the same formalism. Specializing the results in [9] to the sample complexity questions that we study here yields much weaker results than the ones we prove—only a lower bound of $\epsilon^{-1/2}$ and an upper bound of ϵ^{-c} for a large constant c. Two of the upper bounds in Table 2 follow from previous work. The upper bound of $O(\epsilon^{-3}\log\epsilon^{-1})$ for regular distributions was proved in [11]. (They also proved a bound of $O(\epsilon^{-2}\log\epsilon^{-1})$ for MHR distributions, which is subsumed by our nearly tight bound of $O(\epsilon^{-3/2}\log\epsilon^{-1})$.) The upper bound of $O(H\epsilon^{-2}\log H\epsilon^{-1})$ for bounded valuations can be deduced from Balcan et al. [4].[11] We emphasize that, in addition to our new upper bound results in the large-sample regime, there is no previous work on sample complexity lower bounds for our pricing problem nor on the best-possible approximation given a single sample.

There are many less related previous works that also use the idea of independent samples in the context of auction design. For example, some previous works study the asymptotic (in the number of samples) convergence of an auction's revenue to the optimal revenue, without providing any uniform sample complexity bounds. See Neeman [20], Segal [23], Baliga and Vohra [5], and Goldberg et al. [15] for several examples. Some recent and very different uses of samples in auction design include Fu et al. [13], who use samples to extend the Crémer–McLean theorem [10] to partially known valuation distributions, and Chawla, Hartline, and Nekipelov [7], who design auctions that have both near-optimal revenue and enable accurate inference about the valuation distribution from samples.

2. Preliminaries. Suppose the buyer's value is drawn from a publicly known distribution D whose support is a continuous interval. Let F be the c.d.f. of D. If F is differentiable, let f be the probability density function (p.d.f.) of D. Let $q(v) = 1 - F(v)$ be the quantile of value v, i.e., the sale probability of reserve price v.[12] Let $v(q)$ be the value with quantile q.

The first set of distributions we study are those satisfying standard small-tail assumptions such as regularity, MHR, and α-strong regularity [9]. We explain these assumptions in more detail next. For these distributions, we assume F is differentiable and f exists.

Let $R(q) = qv(q)$ be the revenue as a function over the quantiles space (meaning the expected revenue obtained from a price $v(q)$ with quantile q on a random draw from D). We have

$$R'(q) = v(q) + q\frac{dv}{dq} = v - \frac{q(v)}{f(v)} \ .$$

[11] The paper by Balcan et al. [4] studies a seemingly different problem—the design of digital good auctions with n buyers in a prior-free setting (with bounded valuations). But if one instantiates their model with bidders with i.i.d. valuations from a distribution D, then their performance analysis of their random sampling optimal offer (RSO) mechanism essentially gives a performance guarantee for the empirical monopoly price for D with $n/2$ samples, relative to the expected revenue of the monopoly price with a single bidder.

[12] This terminology is for consistency with the recent literature on related topics. "Reverse quantile" might be a more accurate term.

The *virtual valuation function* is defined to be $\phi(v) = v - \frac{1-F(v)}{f(v)} = R'(q)$. A distribution D is *regular* if for all values v in its support,

$$\frac{d\phi}{dv} \geq 0. \tag{1}$$

Note that $v(q)$ is decreasing in q. A distribution is regular if $R'(q) = \phi(v)$ is decreasing in q and, thus, $R(q)$ is concave. In this case, $R(q)$ is maximized whenever $R'(q) = \phi(v(q)) = 0$. Let q^* and $v^* = v(q^*)$ be the revenue-optimal quantile and reserve prices, respectively.

A distribution D satisfied the MHR condition if for all values v in its support,

$$\frac{d\phi}{dv} \geq 1. \tag{2}$$

We'll make use of the following property of MHR distributions.

LEMMA 2.1 (see [17]). *For every MHR distribution,* $q^* \geq \frac{1}{e}$.

Cole and Roughgarden [9] defined *α-strong regular* distributions (where $\alpha \in [0, 1]$) to interpolate between (1) and (2):

$$\frac{d\phi}{dv} \geq \alpha. \tag{3}$$

Many properties of MHR distributions carry over to α-strongly regular distributions with different constants. For example, we have the following.

LEMMA 2.2 (see [9]). *For every α-strongly regular distribution with $\alpha > 0$,* $q^* \geq \alpha^{1/(1-\alpha)}$.

To reason about general (irregular) distributions, we require an alternative benchmark (recall footnote 2). We propose

$$R^*_\delta = \max_{q \geq \delta} qv(q),$$

the optimal revenue if we only consider reserve prices with sale probability at least δ. Here, we expect the sample complexity to depend on both ϵ and δ.

3. Asymptotic upper bounds.

We now present our positive results in the asymptotic regime.

DEFINITION 3.1. *Given m samples $v_1 \geq v_2 \geq \cdots \geq v_m$, the* empirical reserve *is*

$$\arg\max_{i \geq 1} i \cdot v_i.$$

If we only consider $i \geq cm$ for some parameter c, it is called the *c-guarded empirical reserve.*

3.1. MHR upper bound.

We next prove the following.

THEOREM 3.2. *The empirical reserve with $m = \Theta(\epsilon^{-3/2} \log \epsilon^{-1})$ samples is $(1 - \epsilon)$-approximate for all MHR distributions.*

We also give a matching lower bound (up to the log factor) in section 4.

For simplicity of presentation, we prove Theorem 3.2 for the $\frac{1}{e}$-guarded empirical reserve. (Recall that $q^* \geq \frac{1}{e}$ for MHR distributions.) The unguarded version is similar but requires some extra care on the small quantiles.

To show Theorem 3.2, we use two properties of MHR distributions. First, the optimal quantile of an MHR distribution is at least e^{-1} (Lemma 2.1). Second, the revenue decreases quadratically in how much the reserve price deviates from the optimal one in quantile space, which we formulate as the following lemma.

LEMMA 3.3. *For every* $0 \le q' \le 1$, *we have* $R(q^*) - R(q') \ge \frac{1}{4}(q^* - q')^2 R(q^*)$.

Proof. First, consider the case when $q' > q^*$. By the optimality of q^*, for any q s.t. $q^* \le q \le q'$, we have $qv(q) \le q^* v(q^*)$ and, hence,

$$v(q) \le \frac{q^*}{q} v(q^*). \tag{4}$$

Further, since the MHR assumption implies that $\frac{d\phi(v)}{dv} \ge 1$ for every $q^* \le q \le q'$, we have

$$\phi(v(q)) \le \phi(v(q^*)) + v(q) - v(q^*) = v(q) - v(q^*).$$

Combining this with the inequality (4), we get that

$$\phi(v(q)) \le \frac{q^* - q}{q} v(q^*).$$

Therefore,

$$R(q^*) - R(q') = \int_{q^*}^{q'} -R'(q)dq = \int_{q^*}^{q'} -\phi(v(q))dq \ge \int_{q^*}^{q'} \frac{q - q^*}{q} v(q^*)dq.$$

Note that $\frac{q-q^*}{q} \ge 0$ for every $q^* \le q \le q'$. Moreover, for any $q \ge \frac{q'+q^*}{2}$, we have $\frac{q-q^*}{q} \ge \frac{q'-q^*}{q'+q^*}$. Hence,

$$R(q^*) - R(q') \ge \int_{\frac{q'+q^*}{2}}^{q'} \frac{q' - q^*}{q' + q^*} v(q^*)dq = \frac{(q' - q^*)^2}{2(q' + q^*)} v(q^*) = \frac{(q' - q^*)^2}{2q^*(q' + q^*)} R(q^*).$$

The lemma then follows from the fact that $0 \le q', q^* \le 1$.

Next, we consider the case when $q' < q^*$. The high-level proof idea of this case is similar to the previous case, but requires some subtle changes in the inequalities. For completeness, we include the proof below.

By concavity of the revenue curve, for any $q' \le q \le q^*$, we have

$$qv(q) \ge \frac{q - q'}{q^* - q'} q^* v(q^*) + \frac{q^* - q}{q^* - q'} q' v(q').$$

Dividing both sides by q, we have

$$v(q) \ge \frac{q^* v(q^*) - q' v(q')}{q^* - q'} + \frac{q^* q'}{q(q^* - q')} \big(v(q') - v(q^*)\big). \tag{5}$$

Further, by the MHR assumption,

$$\phi(v(q)) \ge \phi(v(q^*)) + v(q) - v(q^*) \ = \ v(q) - v(q^*).$$

Note that the direction of this inequality is the opposite of its counterpart in the first case (because $v(q) > v(q^*)$ rather than $v(q) < v(q^*)$). Combining this with inequality (5), we get

$$\begin{aligned}\phi(v(q)) &\geq \frac{q^*v(q^*) - q'v(q')}{q^* - q'} + \frac{q^*q'}{q(q^* - q')}\big(v(q') - v(q^*)\big) - v(q^*)\\ &= \frac{q'(q^* - q)}{q(q^* - q')}\big(v(q') - v(q^*)\big)\\ &\geq \frac{q'(q^* - q)}{q^*(q^* - q')}\big(v(q') - v(q^*)\big),\end{aligned}$$

where the last inequality is due to $q \leq q^*$. Hence, we have

$$\begin{aligned}R(q^*) - R(q') &= \int_{q'}^{q^*} R'(q)dq \qquad (6)\\ &= \int_{q'}^{q^*} \phi(v(q))dq\\ &\geq \int_{q'}^{q^*} \frac{q'(q^* - q)}{q^*(q^* - q')}\big(v(q') - v(q^*)\big)dq\\ &= \frac{q'}{2q^*}(q^* - q')\big(v(q') - v(q^*)\big). \qquad (7)\end{aligned}$$

On the other hand, we have

$$R(q^*) - R(q') = q^*v(q^*) - q'v(q'). \qquad (8)$$

Taking the convex combination of $\frac{2q^*}{3q^*-q'}$ times the expression in (7) plus $\frac{q^*-q'}{3q^*-q'}$ times the right-hand side of (8) yields

$$R(q^*) - R(q') \geq \frac{(q^* - q')^2}{3q^* - q'}v(q^*) = \frac{1}{q^*(3q^* - q')}(q^* - q')^2R(q^*) \geq \frac{1}{3}(q^* - q')^2R(q^*),$$

where the last inequality holds because $0 \leq q^*, q' \leq 1$. ∎

Next we show how to use the lemma (and additional ideas) to prove Theorem 3.2.

Proof of Theorem 3.2. We first show that for any two samples v_1 and v_2 with quantiles q_1, q_2 such that either $q_1 < q_2 < q^*$ or $q^* < q_1 < q_2$, if the revenue of one of v_1, v_2 is at least $1 - \frac{\epsilon}{2}$ times smaller than that of the other, then with probability at least $1 - o(\frac{1}{m^2})$, the algorithm correctly determines which of v_1, v_2 is the price with the higher revenue. Further, with high probability, there is at least one sample that is $\frac{\epsilon}{2}$-close to q^* in quantile space both among samples with quantile at least q^* and among those with quantile at most q^*. By concavity of the revenue curve, such samples are $(1 - \frac{\epsilon}{2})$-optimal. So the theorem follows from union bound.

Let us focus on the case when $q_1 < q_2 < q^*$ and $v_1q_1 < (1-\epsilon)v_2q_2$; the other case is almost identical. Suppose $R(q_1) = (1 - \Delta)R(q_2)$ and $q_1 = q_2 - \delta$. By concavity of the revenue curve and Lemma 3.3, we have

$$R(q_2) - R(q_1) \geq R(q^*) - R(q^* - q_2 + q_1) \geq \frac{1}{4}(q_2 - q_1)^2R(q^*) \geq \frac{1}{4}(q_2 - q_1)^2R(q_2).$$

So we have $\Delta = \Omega(\delta^2)$.

Let $\tilde{q}_i m$ be the number of samples with value at least v_i, $i = 1, 2$. The goal is to show that $\tilde{q}_1 v_1 < \tilde{q}_2 v_2$ with probability at least $1 - o(\frac{1}{m^2})$. The straightforward argument does not work because we would need $\tilde{\Omega}(\epsilon^{-2})$ samples to estimate q_i up to a $1 - \epsilon$ factor.

Before diving into the technical proof, let us explain informally how to get away with fewer samples. A bad scenario for the straightforward argument is when, say, $\tilde{q}_1 > (1 + \Delta)q_1$ and $\tilde{q}_2 < (1 - \Delta)q_2$. We observe that such a bad scenario is very unlikely due to correlation between $\tilde{q}_1$ and $\tilde{q}_2$: the samples used to estimate q_1 and q_2 are the same; those that cause the algorithm to overestimate q_1 also contribute to the estimation of q_2; for the bad scenario to happen, it must be that the number of samples *between* q_1 and q_2 is much smaller than its expectation (as we will formulate as (9)), and this probability is tiny.

Now we proceed with the formal proof. Since we consider the $\frac{1}{e}$-guarded empirical reserve, $q_1, q_2 \geq \frac{1}{e}$. By the Chernoff bound, with $\Theta(\epsilon^{-3/2} \log \epsilon^{-1})$ samples, we have $\tilde{q}_i \geq (1 - \epsilon^{3/4})q_i = \Omega(1)$ and $\tilde{q}_i \leq (1 + \epsilon^{3/4})q_i$ with high probability.

If $\tilde{q}_1 v_1 \geq \tilde{q}_2 v_2$, then

$$\frac{\tilde{q}_1}{\tilde{q}_2} \geq \frac{v_2}{v_1} = \frac{R(q_2)}{R(q_1)} \frac{q_1}{q_2} = (1 - \Delta)^{-1} \frac{q_1}{q_2} = \frac{q_1}{q_2} + \Omega(\Delta) \ .$$

So

$$\begin{aligned} \tilde{q}_2 - \tilde{q}_1 &= \left(1 - \frac{\tilde{q}_1}{\tilde{q}_2}\right) \tilde{q}_2 \\ &\leq \left(1 - \frac{\tilde{q}_1}{\tilde{q}_2}\right) \left(1 + \epsilon^{3/4}\right) q_2 \\ &\leq \left(1 - \frac{q_1}{q_2} - \Omega(\Delta)\right) \left(1 + \epsilon^{3/4}\right) q_2 \\ &= q_2 - q_1 + \delta \epsilon^{3/4} - \Omega(\Delta). \end{aligned}$$

Since $\delta = O(\sqrt{\Delta})$ and $\Delta \geq \frac{\epsilon}{2}$, we have $\delta \epsilon^{3/4} = o(\Delta)$. So

$$\tilde{q}_2 - \tilde{q}_1 \leq q_2 - q_1 - \Omega(\Delta) \ . \tag{9}$$

That is, the number of samples that fall between q_1 and q_2 is smaller than its expectation by at least $\Omega(\Delta m)$. By the Chernoff bound, the probability of this event is at most $\exp(-\frac{\Delta^2 m}{\delta})$. Recall that $\Delta = \Omega(\delta^2)$, $\Delta \geq \frac{\epsilon}{2}$, and $m = \Theta(\epsilon^{-3/2} \log \epsilon^{-1})$. So this probability is at most $\exp(-\Omega(\log \epsilon^{-1})) = o(\frac{1}{m^2})$ with an appropriate choice of parameters. ∎

3.2. α-strongly regular upper bound. Our proof of Theorem 3.2 can be extended to α-strongly regular distributions with $\alpha > 0$. We present the formal statement and sketch the necessary changes below.

THEOREM 3.4. *For every $\alpha > 0$, the empirical reserve with $m = \Theta(\epsilon^{-3/2} \log \epsilon^{-1})$ samples is $(1 - \epsilon)$-approximate for all α-strongly regular distributions (with the hidden constant depending on α).*

The proof of Theorem 3.2 relies on two properties: the monopoly price having at least constant sale probability (Lemma 2.1), and strict concavity of the revenue curve at the monopoly quantile (Lemma 3.3). The proof of Theorem 3.4 is identical, modulo using weaker versions of the lemmas. Specifically, we will replace Lemma 2.1

by Lemma 2.2, and Lemma 3.3 by the following lemma, whose proof is almost identical to that of Lemma 3.3.

LEMMA 3.5. *For any $0 \leq q' \leq 1$, we have $R(q^*) - R(q') \geq \frac{\alpha}{3}(q^* - q')^2 R(q^*)$.*

Proof of Theorem 3.4. We use the same setup and notation as in the MHR case. Using the weaker lemmas, we have $q_i, \tilde{q}_i = \Omega(\alpha^{1/(1-\alpha)})$ and $\Delta \geq \Omega(\alpha\delta^2)$.

If $\tilde{q}_1 v_1 \geq \tilde{q}_2 v_2$, then

$$\frac{\tilde{q}_1}{\tilde{q}_2} \geq \frac{v_2}{v_1} = \frac{R(q_2)}{R(q_1)}\frac{q_1}{q_2} = (1-\Delta)^{-1}\frac{q_1}{q_2} = \frac{q_1}{q_2} + \Omega\left(\alpha^{1/(1-\alpha)}\Delta\right) .$$

So

$$\begin{aligned}
\tilde{q}_2 - \tilde{q}_1 &= \left(1 - \frac{\tilde{q}_1}{\tilde{q}_2}\right)\tilde{q}_2 \\
&\leq \left(1 - \frac{\tilde{q}_1}{\tilde{q}_2}\right)\left(1 + \epsilon^{3/4}\right) q_2 \\
&\leq \left(1 - \frac{q_1}{q_2} - \Omega(\alpha^{1/(1-\alpha)}\Delta)\right)\left(1 + \epsilon^{3/4}\right) q_2 \\
&= q_2 - q_1 + \delta\epsilon^{3/4} - \Omega\left(\alpha^{1/(1-\alpha)}\Delta\right).
\end{aligned}$$

Since $\delta = O(\sqrt{\Delta/\alpha})$ and $\Delta \geq \frac{\epsilon}{2}$, we have $\delta\epsilon^{3/4} = o(\Delta)$. So

$$\tilde{q}_2 - \tilde{q}_1 \leq q_2 - q_1 - \Omega\left(\alpha^{1/(1-\alpha)}\Delta\right) . \tag{10}$$

That is, the number of samples that fall between q_1 and q_2 differs from its expectation by at least $\Omega(\alpha^{1/(1-\alpha)}\Delta m)$. By the Chernoff bound, the probability of this event is at most $\exp(-\frac{(\alpha^{1/(1-\alpha)}\Delta)^2 m}{\delta})$. Recall that $\Delta = \Omega(\alpha\delta^2)$, $\Delta \geq \frac{\epsilon}{2}$, and set $m = \Theta(\alpha^{-2/(1-\alpha)-1/2}\epsilon^{-3/2}\log\epsilon^{-1})$. Then, this probability is at most $\exp(-\Omega(\log\epsilon^{-1})) = o(\frac{1}{m^2})$ with a suitable choice of parameters. ∎

3.3. General upper bounds. Next, we present a sample complexity upper bound for general distributions using R^*_δ as a benchmark. Recall that R^*_δ is the optimal revenue by prices with sale probability at least δ.

THEOREM 3.6. *The $\frac{\delta}{2}$-guarded empirical reserve with $m = \Theta(\delta^{-1}\epsilon^{-2}\log(\delta^{-1}\epsilon^{-1}))$ gives revenue at least $(1-\epsilon)R^*_\delta$ for all distributions.*

Sketch. The proof is standard so we present only a sketch here. Let $q^*_\delta = \arg\max_{q\geq\delta} qv(q)$ be the optimal reserve price with sale probability at least δ. With high probability, there exists at least one sampled price with quantile between $(1-\frac{\epsilon}{3})q^*$ and q^*: this price has revenue at least $(1-\frac{\epsilon}{3})R^*_\delta$. Further, since $q^* \geq \delta$, this price has rank at least $\frac{\delta}{2}$ (among sampled prices) with high probability and thus is considered by the empirical reserve algorithm. Finally, with high probability, any sampled price with rank at least $\frac{\delta}{2}$ has sale probability at least $\frac{\delta}{4}$; for prices with sale probability at least $\frac{\delta}{4}$, the algorithm estimates their sale probability up to a $1-\frac{\epsilon}{3}$ factor with high probability with $m = \Theta(\delta^{-1}\epsilon^{-2}\log(\delta^{-1}\epsilon^{-1}))$ samples. The theorem then follows. ∎

We remark that one can also derive a bound for a single sample (i.e., $m = 1$) which guarantees expected revenue at least $(\delta/2)R^*_\delta$.

Note that for distributions with support $[1, H]$, the optimal sale probability is at least $1/H$. So we have the following theorem as a direct corollary of Theorem 3.6. This bound can also be deduced from [4]; we include it for completeness.

THEOREM 3.7. *The empirical reserve with* $m = \Theta(H\epsilon^{-2}\log(H\epsilon^{-1}))$ *samples is* $(1-\epsilon)$*-approximate for all distributions with support* $[1, H]$.

4. Asymptotic lower bounds. This section gives asymptotically tight (up to a log factor) sample complexity lower bounds. These lower bounds are information theoretic and apply to all possible pricing strategies, including randomized strategies. We first present a general framework for proving sample complexity lower bounds, and then instantiate it for each of the classes of distributions listed in Table 1.

4.1. Lower bound framework: Reducing pricing to classification. The high-level plan is to reduce the pricing problem to a classification problem. We will construct two distributions D_1 and D_2 and show that given any pricing algorithm that is $(1-\epsilon)$-approximate for both D_1 and D_2, we can construct a classification algorithm that can distinguish D_1 and D_2 with constant probability, say, $\frac{1}{3}$, using the same number of samples as the pricing algorithm. Further, we will construct D_1 and D_2 to be similar enough and use tools from information theory to show a lower bound on the number of samples needed to distinguish the two distributions.

Information theory preliminaries. Consider two distributions P_1 and P_2 over a sample space Ω. Let p_1 and p_2 be the density functions. The *statistical distance* between P_1 and P_2 is

$$\delta(P_1, P_2) = \frac{1}{2}\int_{\Omega} \big|p_1(\omega) - p_2(\omega)\big| d\omega.$$

In information theory, it is known (e.g., [3]) that no classification algorithm $A : \Omega \to \{1, 2\}$ can distinguish P_1 and P_2 correctly with probability strictly better than $\frac{\delta(P_1,P_2)+1}{2}$, i.e., there exists $i \in \{1, 2\}$, $\mathbf{Pr}_{\omega\sim P_i}[A(\omega) = i] \leq \frac{\delta(P_1,P_2)+1}{2}$. This lower bound applies to arbitrary randomized classification algorithms.

Suppose we want to show a sample complexity lower bound of m. Then we will let $P_i = D_i^m$ and then upper bound $\delta(P_1, P_2)$. However, the statistical distance is hard to bound directly when we have multiple samples: $\delta(D_1^m, D_2^m)$ cannot be written as a function of m and $\delta(D_1, D_2)$. In particular, the statistical distance does not grow linearly with the number of samples.

In order to derive an upper bound on the statistical distance with multiple samples, it is many times convenient to use the *KL divergence*, which is defined as follows:

$$D_{KL}(P_1 \| P_2) = \mathbf{E}_{\omega\sim P_1}\left[\ln \frac{p_1(\omega)}{p_2(\omega)}\right] .$$

In information theory, the KL divergence can be viewed as the redundancy in the encoding in the case that the true distribution is P_1 and we use the optimal encoding for distribution P_2. One nice property of the KL divergence is that it is additive over samples: if $P_1 = D_1^m$ and $P_2 = D_2^m$ are the distributions over m samples of D_1 and D_2, then the KL divergence of P_1 and P_2 is m times $D_{KL}(D_1 \| D_2)$.

We can relate the KL divergence to the statistical distance through Pinsker's inequality [22], which states that

$$\delta(P_1, P_2) \leq \sqrt{\tfrac{1}{2} D_{KL}(P_1 \| P_2)}.$$

By symmetry, we also have $\delta(P_1, P_2) \leq \sqrt{\frac{1}{2} D_{KL}(P_2 \| P_1)}$, so

$$\delta(P_1, P_2) \leq \frac{1}{2}\sqrt{D_{KL}(P_1 \| P_2) + D_{KL}(P_2 \| P_1)}.$$

This implies that we can upper bound the statistical distance of m samples from D_1 and D_2 by $\frac{1}{2}\sqrt{m \cdot (D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1))}$. To get a statistical distance of at least, say, $\frac{1}{3}$, we need $m = \frac{4}{9} \frac{1}{D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1)}$ samples.

Reducing pricing to classification. Next, we present the reduction from pricing to classification. Given a value distribution D and $\alpha < 1$, its *α-optimal price set* is defined to be the set of reserve prices that induce at least an α fraction of the optimal revenue.

LEMMA 4.1. *If value distributions D_1 and D_2 have disjoint $(1-3\epsilon)$-approximate price sets, and there is a pricing algorithm that is $(1-\epsilon)$-approximate for both D_1 and D_2, then there is a classification algorithm that distinguishes P_1 and P_2 correctly with probability at least $\frac{2}{3}$, using the same number of samples as the pricing algorithm.*

We omit the straightforward proof. Note that to distinguish P_1 and P_2 correctly with probability at least $\frac{2}{3}$, the statistical distance between P_1 and P_2 is least $\frac{1}{3}$. So we have the following theorem.

THEOREM 4.2. *If value distributions D_1 and D_2 have disjoint $(1-3\epsilon)$-approximate price sets, and there is a pricing algorithm that is $(1-\epsilon)$-approximate for both D_1 and D_2, then the algorithm uses at least $\frac{4}{9} \frac{1}{D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1)}$ samples.*

A tool for constructing distributions with small KL divergence. Given Theorem 4.2, our goal is to construct a pair of distributions with small relative entropy subject to having disjoint approximately optimal price sets. Here we introduce a lemma from the differential privacy literature that is useful for constructing pairs of distributions with small KL divergence.

LEMMA 4.3 (see [12, Lemma III.2]). *If distributions D_1 and D_2 with densities f_1 and f_2 satisfy that $(1+\epsilon)^{-1} \le \frac{f_1(\omega)}{f_2(\omega)} \le (1+\epsilon)$ for every $\omega \in \Omega$, then*

$$D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1) \le \epsilon^2 \ .$$

For completeness, we include the proof.

Proof. By the definition of KL divergence, we have

$$\begin{aligned} &D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1) \\ &\quad = \int_\Omega \left[p_1(\omega) \ln \frac{p_1(\omega)}{p_2(\omega)} + p_2(\omega) \ln \frac{p_2(\omega)}{p_1(\omega)} \right] d\omega \\ &\quad = \int_\Omega \left[p_1(\omega) \left(\ln \frac{p_1(\omega)}{p_2(\omega)} + \ln \frac{p_2(\omega)}{p_1(\omega)} \right) + \big(p_2(\omega) - p_1(\omega)\big) \ln \frac{p_2(\omega)}{p_1(\omega)} \right] d\omega \\ &\quad \le \int_\Omega \Big[0 + |p_2(\omega) - p_1(\omega)| \ln(1+\epsilon) \Big] d\omega. \end{aligned}$$

The last inequality follows by $(1+\epsilon)^{-1} \le \frac{f_1(\omega)}{f_2(\omega)} \le (1+\epsilon)$. Further, this condition also implies that

$$|p_2(\omega) - p_1(\omega)| \le \big((1+\epsilon) - 1\big) \min\{p_1(\omega), p_2(\omega)\} = \epsilon \min\{p_1(\omega), p_2(\omega)\}.$$

Thus, we have

$$\begin{aligned} D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1) &\le \epsilon \ln(1+\epsilon) \int_\Omega \min\{p_1(\omega), p_2(\omega)\} d\omega \\ &\le \epsilon \ln(1+\epsilon) \\ &\le \epsilon^2. \end{aligned}$$

∎

The following two useful variants have similar proofs.

LEMMA 4.4. *If distributions D_1 and D_2 satisfy the condition in Lemma* 4.3, *and further there is a subset of outcomes Ω' such that $p_1(\omega) = p_2(\omega)$ for every $\omega \in \Omega'$, then*

$$D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1) \le \epsilon^2 \big(1 - p_1(\Omega')\big) \ .$$

LEMMA 4.5. *If distributions D_1 and D_2 satisfy that $(1+\epsilon)^{-1} \le \frac{f_1(\omega)}{f_2(\omega)} \le (1+\epsilon)$ for every $\omega \in \Omega$ and $(1+\epsilon')^{-1} \le \frac{f_1(\omega)}{f_2(\omega)} \le (1+\epsilon')$ for any $\omega \in \Omega' \subseteq \Omega$, then*

$$D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1) \le \epsilon^2 p_1(\Omega \setminus \Omega') + (\epsilon')^2 p_1(\Omega') \ .$$

4.2. Applications. Inspired by the above lemmas, we will aim to construct D_1 and D_2 such that the densities of all values are close in the two distributions.

General lower bound. As a warm-up case, we demonstrate how to use the above framework to derive a tight (up to a log factor) sample complexity lower bound for general distributions using R^*_δ as benchmark. Recall that R^*_δ is the optimal revenue by prices with sale probability at least δ.

THEOREM 4.6. *Every pricing algorithm that guarantees at least $(1-\epsilon)R^*_\delta$ revenue for all distributions uses $\Omega(\delta^{-1}\epsilon^{-2})$ samples.*

Proof. Let D_1 and D_2 be two distributions with support $\{H = \delta^{-1}, 2, 1\}$: D_1 takes values H with probability $\frac{1+4\epsilon}{H}$, 2 with probability $\frac{1-4\epsilon}{H}$, and 1 with probability $1 - \frac{2}{H}$; D_2 takes values H with probability $\frac{1-4\epsilon}{H}$, 2 with probability $\frac{1+4\epsilon}{H}$, and 1 with probability $1 - \frac{2}{H}$. Clearly, D_1 and D_2 have disjoint $(1-3\epsilon)$-approximate price sets. Further,

$$D_{KL}(D_1 \| D_2) = D_{KL}(D_2 \| D_1) = \tfrac{1+4\epsilon}{H} \ln \tfrac{1+4\epsilon}{1-4\epsilon} + \tfrac{1-4\epsilon}{H} \ln \tfrac{1-4\epsilon}{1+4\epsilon} = \tfrac{8\epsilon}{H} \ln \tfrac{1+4\epsilon}{1-4\epsilon} = O\left(\tfrac{\epsilon^2}{H}\right).$$

So the claim follows from Theorem 4.2. ∎

The way we prove Theorem 4.6 also implies a tight (up to a log factor) sample complexity lower bound for distributions with support in $[1, H]$.

THEOREM 4.7. *Every pricing algorithm that is $(1-\epsilon)$-approximate for all distributions with support in $[1, H]$ uses $\Omega(H\epsilon^{-2})$ samples.*

Regular lower bound. We now show that $(1-\epsilon)$-approximate pricing for regular distributions requires $\Omega(\epsilon^{-3})$ samples.

THEOREM 4.8. *Any pricing algorithm that is $(1-\epsilon)$-approximate for all regular distributions uses at least $\frac{(1-6\epsilon)^2}{486\epsilon^3} = \Omega(\frac{1}{\epsilon^3})$ samples.*

This result implies, for example, that we need at least 12 samples to guarantee 95 percent of the optimal revenue, and at least 1935 samples to guarantee 99 percent.

We next describe the two distributions that we use and explain the lower bound for regular distributions. Let D_1 be the distribution with c.d.f. $F_1(v) = 1 - \frac{1}{v+1}$ and p.d.f. $f_1(v) = \frac{1}{(v+1)^2}$. Let $\epsilon_0 = 3\epsilon$ and let D_2 be the distribution with c.d.f.

$$F_2(v) = \begin{cases} 1 - \frac{1}{v+1} & \text{if } 0 \le v \le \frac{1-2\epsilon_0}{2\epsilon_0}, \\ 1 - \frac{(1-2\epsilon_0)^2}{v-(1-2\epsilon_0)} & \text{if } v > \frac{1-2\epsilon_0}{2\epsilon_0}, \end{cases}$$

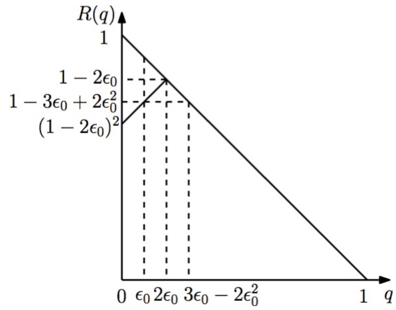

FIG. 1. *D_1 is the distribution with revenue curve R_1 that goes from $(0,1)$ to $(1,0)$. D_2 is identical to D_1 for quantiles from $2\epsilon_0$ to 1; for quantiles from 0 to $2\epsilon_0$, D_2's revenue curve goes from $(0,(1-2\epsilon_0)^2)$ to $(2\epsilon_0, 1-2\epsilon_0)$.*

and p.d.f.

$$f_2(v) = \begin{cases} \frac{1}{(v+1)^2} & \text{if } 0 \le v \le \frac{1-2\epsilon_0}{2\epsilon_0}, \\ \frac{(1-2\epsilon_0)^2}{(v-(1-2\epsilon_0))^2} & \text{if } v > \frac{1-2\epsilon_0}{2\epsilon_0}. \end{cases}$$

We have

(11)
$$\frac{f_1}{f_2} = \begin{cases} 1 & \text{if } 0 \le v \le \dfrac{1-2\epsilon_0}{2\epsilon_0}, \\ \dfrac{1}{(1-2\epsilon_0)^2}\dfrac{(v-(1-2\epsilon_0))^2}{(v+1)^2} \in [(1-2\epsilon_0)^2, (1-2\epsilon_0)^{-2}] & \text{if } v > \dfrac{1-2\epsilon_0}{2\epsilon_0}. \end{cases}$$

The revenue curves of D_1 and D_2 are summarized in Figure 1.

LEMMA 4.9.

$$D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1) \le \frac{8\epsilon_0^3}{(1-2\epsilon_0)^2}.$$

Proof. By (11), we have $(1-2\epsilon_0) \le \frac{f_1}{f_2} \le (1-2\epsilon_0)^{-1}$. Further, note that a $1-2\epsilon$ fraction (w.r.t. quantile) of D_1 and D_2 are identical. The lemma follows from Lemma 4.4. ∎

Let R_1 and R_2 be the revenue curves of D_1 and D_2. Let R_1^* and R_2^* be the corresponding optimal revenues. The following lemmas follow directly from the definition of D_1 and D_2.

LEMMA 4.10. *$R_1(v) \ge (1-\epsilon_0)R_1^*$ if and only if $v \ge \frac{1}{\epsilon_0} - 1$.*

LEMMA 4.11. *$R_2(v) \ge (1-\epsilon_0)R_1^*$ if and only if $\frac{1}{\epsilon_0} - 3 + 2\epsilon_0 \ge v \ge \frac{1}{2\epsilon_0 - \epsilon_0^2} - 1$.*

Recall that $\epsilon_0 = 3\epsilon$. The $(1-3\epsilon)$-optimal price sets of D_1 and D_2 are disjoint. Theorem 4.8 now follows from Theorem 4.2 and Lemma 4.9.

MHR lower bound. We now turn to MHR distributions and show that $(1-\epsilon)$-approximate pricing for MHR distributions requires $\Omega(\epsilon^{-3/2})$ samples.

THEOREM 4.12. *Any pricing algorithm that is $(1-\epsilon)$-approximate for all MHR distributions uses at least $\Omega(\epsilon^{-3/2})$ samples.*

We first describe the two distributions that we use. Let D_1 be the uniform distribution over $[1,2]$. Let $\epsilon_0 = c\epsilon$, where c is a sufficiently large constant to be

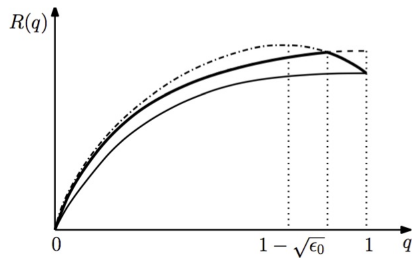

FIG. 2. *R_1 (the lower solid curve) is a quadratic curve that peaks at $q = 1$, $R_1 = 1$, and passes through $q = 0$, $R_1 = 0$. To construct R_2, first draw the revenue curves of the uniform distributions over $[1+\frac{2\epsilon_0}{1-\sqrt{\epsilon_0}+2\epsilon_0}, 2]$ (the dashed curve) and $[1, 1+\frac{1}{1-2\sqrt{\epsilon_0}}]$ (the dash-dotted curve). R_2 (the bold solid curve) is the lower envelope of the two curves.*

determined later. Define D_2 by scaling up the density (of D_1) in $v \in [1+\sqrt{\epsilon_0}, 2]$ by a factor of $1+\frac{2\epsilon_0}{1-\sqrt{\epsilon_0}}$ and scaling down the density in $v \in [1, 1+\sqrt{\epsilon_0}]$ by a factor $1-2\sqrt{\epsilon_0}$, i.e.,

$$f_2(v) = \begin{cases} 1-2\sqrt{\epsilon_0} & \text{if } 1 \le v \le 1+\sqrt{\epsilon_0}, \\ 1+\frac{2\epsilon_0}{1-\sqrt{\epsilon_0}} & \text{if } 1+\sqrt{\epsilon_0} < v \le 2. \end{cases}$$

We summarize the revenue curves of D_1 and D_2 in Figure 2.

LEMMA 4.13.

$$D_{KL}(D_1 \| D_2) + D_{KL}(D_2 \| D_1) = O\left(\epsilon_0^{3/2}\right).$$

Proof. By our choice of D_1 and D_2, D_1 and D_2 differ by $1+\frac{2\epsilon_0}{1-\sqrt{\epsilon_0}}$ for $v \in [1+\sqrt{\epsilon_0}, 2]$, and by $1-2\sqrt{\epsilon_0}$ for $v \in [1, 1+\sqrt{\epsilon_0}]$. The lemma follows from Lemma 4.5. ◻

Let q_i^* be the revenue optimal quantile of R_i, and let $R_i^* = R_i(q_i^*)$ be the optimal revenue.

LEMMA 4.14. *$q_1^* = 1$ and $R_1^* = 1$.*

Proof. $R_1(q) = q(2-q)$ is a quadratic curve that peaks at $q = 1$ with $R_1(1) = 1$. ◻

LEMMA 4.15. *$q_2^* = 1-\sqrt{\epsilon_0}+2\epsilon_0$ and $R_2^* = 1+\epsilon_0+2\epsilon^{3/2}$.*

Proof. For $1 \le v \le 1+\sqrt{\epsilon_0}$ and $1-\sqrt{\epsilon_0}+2\epsilon_0 \le q(v) \le 1$, R_2 is identical to the revenue curve of the uniform distribution over $[1, 1+\frac{1}{1-2\sqrt{\epsilon_0}}]$, i.e., $q(\frac{2-2\sqrt{\epsilon_0}}{1-2\sqrt{\epsilon_0}} - \frac{1}{1-2\sqrt{\epsilon_0}}q)$, which is maximized at $q = 1-\sqrt{\epsilon_0}$.

For $1+\sqrt{\epsilon_0} \le v \le 2$ and $0 \le q(v) \le 1-\sqrt{\epsilon_0}+2\epsilon_0$, R_2 is identical to the revenue curve of the uniform distribution over $[1+\frac{2\epsilon_0}{1-\sqrt{\epsilon_0}+2\epsilon_0}, 2]$, i.e., $q(2-\frac{1-\sqrt{\epsilon_0}}{1-\sqrt{\epsilon_0}+2\epsilon_0}q)$, which is maximized at $q = 1$ in interval $[0, 1]$.

R_2 is maximized at the intersection of the two parts, where $q = 1-\sqrt{\epsilon_0}+2\epsilon_0$, and $R_2(1-\sqrt{\epsilon_0}+2\epsilon_0) = 1+\epsilon_0+2\epsilon^{3/2}$. ◻

The following lemmas follow from our construction of D_1 and D_2.

LEMMA 4.16. *If $R_1(q) \ge (1-3\epsilon)R_1^*$, then $q \ge 1-\sqrt{3\epsilon}$ and $v(q) \le \frac{1-3\epsilon}{1-\sqrt{3\epsilon}}R_1^*$.*

LEMMA 4.17. *If $R_2(q) \ge (1-3\epsilon)R_2^*$, then $q \le 1-\sqrt{\epsilon_0}+2\epsilon_0+\sqrt{3\epsilon}$ and $v(q) \ge \frac{1-3\epsilon}{1-\sqrt{\epsilon_0}+2\epsilon_0+\sqrt{3\epsilon}}R_2^*$.*

Therefore, when $\epsilon_0 = c\epsilon$ for sufficiently large constant c, we have $1 - \sqrt{3\epsilon} > 1 - \sqrt{\epsilon_0} + 2\epsilon_0 + \sqrt{3\epsilon}$. Further note that $R_2^* > R_1^*$, so the $(1 - 3\epsilon)$-optimal price sets of D_1 and D_2 are disjoint. Theorem 4.12 then follows from Theorem 4.2 and Lemma 4.13.

5. The single-sample regime: Beating identity pricing for MHR distributions. This section considers deterministic 1-sample pricing strategies. Recall from the introduction that "identity pricing," meaning $p(v_i) = v_1$, has an approximation guarantee of $\frac{1}{2}$ for the class of regular distributions. There is no better 1-sample deterministic pricing strategy for the class of regular distributions (Theorem 6.1). Also, identity pricing is no better than $\frac{1}{2}$-approximate even for the special case of MHR distributions.

THEOREM 5.1. *The identity pricing algorithm is no better than a $\frac{1}{2}$-approximation for MHR distributions.*

Proof. Consider the uniform distribution over $[1 - \epsilon, 1]$. The optimal revenue is $1-\epsilon$, with reserve price $1-\epsilon$. The identity pricing algorithm gets revenue $\frac{1}{\epsilon^2}\int_{1-\epsilon}^{1} v(1-v)dv = \frac{1}{2} - \frac{1}{3}\epsilon$. The approximation ratio approaches $\frac{1}{2}$ as ϵ goes to zero. ∎

We remark that the lower bound still holds if the support must start from 0 because we can add a little mass on $[0, 1-\epsilon]$ without changing the nature of the lower bound. The same trick applies to the lower bounds in section 6.

Our next goal is to show that scaling down the sampled value, i.e., $p(v) = cv$ for some constant $c < 1$, achieves an approximation ratio better than $\frac{1}{2}$ for MHR distributions.

THEOREM 5.2. *$p(v) = 0.85v$ is 0.589-approximate for MHR distributions.*

The intuition is as follows. We divide the quantile space into two subsets: those that are larger than the quantile of the optimal reserve, i.e., q^*, and those that are smaller.

- First, consider those that are larger. We recall the argument that identity pricing is $\frac{1}{2}$-approximate: the expected revenue of identity pricing is the area under the revenue curve; by concavity of the revenue curve, this is at least half the height and, thus, half the optimal revenue. We show that the revenue curve of an MHR distribution is at least as concave as that of an exponential distribution. This implies that identity pricing is strictly better than $\frac{1}{2}$-approximate for quantiles larger than q^*. Furthermore, scaling down the price by a factor of c decreases the revenue by at most a factor of c. For $c < 1$ close enough to 1, the expected revenue of this part is still strictly better than one-half of the optimal.
- Next, consider quantiles that are smaller than q^*. A worst-case scenario for identity pricing is an approximate point mass, where the sale probability of identity pricing is only $\frac{1}{2}$ on average. By scaling down the sampled value by a little, we double the selling probability w.r.t. a point mass without changing the price by much. So the expected revenue of this part is also strictly better than one-half of the optimal.

We now proceed with the formal argument.

Let $\hat{R}(q)$ be the expected revenue of reserve price $p(v(q)) = cv(q)$ w.r.t. revenue curve R. Define $\hat{R}^{\exp}(q)$ similarly w.r.t. the exponential distribution $D^{\exp}$. For ease of presentation, we will without loss of generality scale the values so that $v(q^*) = v^{\exp}(q^*)$ and, thus, $R(q^*) = R^{\exp}(q^*)$.

Technical lemmas about MHR distributions. Given the revenue at quantile q^*, $R^{\exp}$ minimizes the revenue at any quantile larger than q^* among all MHR distributions.

LEMMA 5.3. *For any* $1 \geq q \geq q^*$, $R(q) \geq R^{\exp}(q)$.

Proof. Suppose it is not true. Let q_0 be a quantile such that $R(q_0) < R^{\exp}(q_0)$. Since $R(q^*) = R^{\exp}(q^*)$ and R and $R^{\exp}$ are continuous, there exists $q_1 \in [q^*, q_0]$ such that $R(q_1) = R^{\exp}(q_1)$ and $R'(q_1) \leq (R^{\exp})'(q_1)$. Similarly, since $R(1) \geq 0 = R^{\exp}(1)$, there exists $q_2 \in [q_0, 1]$ such that $R(q_2) < R^{\exp}(q_2)$ and $R'(q_2) > (R^{\exp})'(q_2)$.

Recall that $R'(q_1) = \phi(v(q_1))$ and $R'(q_2) = \phi(v(q_2))$. By the MHR assumption,

$$R'(q_1) - R'(q_2) = \phi(v(q_1)) - \phi(v(q_2)) \geq v(q_1) - v(q_2) = \frac{R(q_1)}{q_1} - \frac{R(q_2)}{q_2}. \tag{12}$$

By the definition of $R^{\exp}$, the above relation holds with equality for $R^{\exp}$:

$$(R^{\exp})'(q_1) - (R^{\exp})'(q_2) = \frac{R^{\exp}(q_1)}{q_1} - \frac{R^{\exp}(q_2)}{q_2}. \tag{13}$$

Subtracting (13) from (12) and using that $R(q_1) = R^{\exp}(q_1)$, we have

$$R'(q_1) - R'(q_2) - (R^{\exp})'(q_1) + (R^{\exp})'(q_2) \geq \frac{R^{\exp}(q_2)}{q_2} - \frac{R(q_2)}{q_2} > 0 \ ,$$

contradicting $R'(q_1) \leq (R^{\exp})'(q_1)$ and $R'(q_2) > (R^{\exp})'(q_2)$. ◻

As a corollary of Lemma 5.3, $R^{\exp}$ minimizes the value at each quantile larger than q^* (Lemma 5.4), and it maximizes the quantile at each value smaller than $v(q^*)$ (Lemma 5.5).

LEMMA 5.4. *For any* $1 \geq q \geq q^*$, $v(q) \geq v^{\exp}(q)$.

LEMMA 5.5. *For any* $0 \leq v \leq v(q^*)$, $q(v) \geq q^{\exp}(v)$.

Expected revenue of large quantiles. We show that for quantiles between q^* and 1, $R^{\exp}$ is indeed the worst-case scenario.

LEMMA 5.6. *For any* $1 \geq q \geq q^*$, $\hat{R}(q) \geq \hat{R}^{\exp}(q)$.

Proof. We abuse notation and let $R(v) = R(q(v))$. By Lemma 5.4, $v(q) \geq v^{\exp}(q)$. So we have $cv(q) \geq cv^{\exp}(q)$. Since $R(v)$ is nondecreasing when $v > v(q^*)$ and $v(q^*) > cv(q) \geq cv^{\exp}(q)$, we have $\hat{R}(q) = R(cv(q)) \geq R(cv^{\exp}(q))$.

Further, by Lemma 5.5, $R^{\exp}$ minimizes the sale probability at any price $v \leq v(q^*)$ and, thus, minimizes the revenue at price v. So we have $R(cv^{\exp}(q)) \geq R^{\exp}(cv^{\exp}(q)) = \hat{R}^{\exp}(q)$ and the lemma follows. ◻

As a direct corollary of Lemma 5.6, we can lower bound the expected revenue of quantiles between q^* and 1 by the expected revenue of the worst-case distribution $R^{\exp}$.

LEMMA 5.7. $\int_{q^*}^1 \hat{R}(q)dq \geq \int_{q^*}^1 \hat{R}^{\exp}(q)dq$.

Expected revenue of small quantiles. Next, we consider quantiles between 0 and q^*. Let q_0 be such that $cv(q_0) = v(q^*)$, i.e., the reserve price is smaller than $v(q^*)$

if and only if the sample has quantile larger than q_0. We will first lower bound the expected revenue of quantiles smaller than q_0, and then handle the other quantiles.

LEMMA 5.8. *For any* $0 \le q' \le q_0$, $\int_0^{q'} \hat{R}(q)dq \ge \frac{1+\sqrt{1-c}}{2} q' R(q')$.

Proof. Let $c_0 = \frac{c}{2}$. For any $i \ge 0$, let $c_{i+1} = 1 - \frac{c}{4c_i}$. We will inductively show that $\int_0^{q'} \hat{R}(q)dq \ge c_i q' R(q')$, and then prove that c_i converges to $\frac{1+\sqrt{1-c}}{2}$.

Base case. By concavity of the revenue curve, $\int_0^{q'} R(q)dq \ge \frac{1}{2} R(q')$. Further, lower reserve prices have larger sale probability. So $\hat{R}(q) \ge cR(q)$ and the base case follows.

Inductive step. Let q_1 be such that $cv(q_1) = v(q')$, i.e., the reserve price is smaller than $v(q')$ if and only if the sample has quantile larger than q_1. We have

$$R(q_1) = v(q_1)q_1 = \frac{v(q')q_1}{c} = \frac{q_1}{cq'} q' v(q') = \frac{q_1}{cq'} R(q').$$

For the interval from 0 to q_1, by the inductive hypothesis we have

$$\int_0^{q_1} \hat{R}(q)dq \ge c_i q_1 R(q_1) = \frac{c_i q_1^2}{cq'} R(q').$$

Next, consider the interval from q_1 to q'. For any $q_1 \le q \le q'$, by that $q' \le q_0$ and our choice of q_0, we have $v(q^*) < cv(q) < cv(q_1) = v(q')$. So $\hat{R}(q) \ge R(q')$. Thus,

$$\int_{q_1}^{q'} \hat{R}(q)dq \ge (q' - q_1)R(q').$$

Putting everything together, we have

$$\int_0^{q'} \hat{R}(q)dq \ge \left(\frac{c_i}{c} \left(\frac{q_1}{q'} \right)^2 + \left(1 - \frac{q_1}{q'} \right) \right) q' R(q').$$

Minimizing the right-hand side over $0 \le q_1 \le q'$, we have

$$\int_0^{q'} \hat{R}(q)dq \ge \left(1 - \frac{c}{4c_i} \right) q' R(q') = c_{i+1} q' R(q').$$

Convergence of c_i. There is only one stable stationary point, $\frac{1+\sqrt{1-c}}{2}$, for the recursion $c_{i+1} = 1 - \frac{c}{4c_i}$. The other, nonstable, stationary point is $\frac{1-\sqrt{1-c}}{2}$. Note that for any $0 < c < 1$, $c_0 = \frac{c}{2} > \frac{1-\sqrt{1-c}}{2}$. So c_i converges to $\frac{1+\sqrt{1-c}}{2}$ as i goes to infinity. ∎

LEMMA 5.9. *If* $c = 0.85$, *then* $\int_0^{q^*} \hat{R}(q)dq \ge 0.656 q^* R(q^*)$.

Proof. Recall that $cv(q_0) = v(q^*)$. So

$$R(q_0) = q_0 v(q_0) = \frac{q_0}{cq^*} q^* v(q^*) = \frac{q_0}{cq^*} R(q^*).$$

Plugging $c = 0.85$ and $\frac{1+\sqrt{1-c}}{2} \geq 0.693$ into Lemma 5.8, we have

$$\int_0^{q_0} \hat{R}(q)dq \geq 0.693 q_0 R(q_0) = 0.693 \frac{1}{c}\left(\frac{q_0}{q^*}\right)^2 q^* R(q^*) \geq 0.815 \left(\frac{q_0}{q^*}\right)^2 q^* R(q^*). \tag{14}$$

On the other hand, for every $q_0 \leq q \leq q^*$, by concavity of the revenue curve, we have

$$qv(q) \geq \frac{q - q_0}{q^* - q_0} q^* v(q^*) + \frac{q^* - q}{q^* - q_0} q_0 v(q_0).$$

Thus,

$$v(q) \geq \frac{q^* v(q^*) - q_0 v(q_0)}{q^* - q_0} + \frac{1}{q}\frac{q_0 q^*}{q^* - q_0}\big(v(q_0) - v(q^*)\big).$$

Further, by our choice of q_0, the quantile of $cv(q)$ is at least q^*. So we have

$$\hat{R}(q) \geq cv(q)q^* \geq \left(\frac{q^* v(q^*) - q_0 v(q_0)}{q^* - q_0} + \frac{1}{q}\frac{q_0 q^*}{q^* - q_0}\big(v(q_0) - v(q^*)\big)\right) cq^*.$$

Let $x = \frac{q_0}{q^*}$. Plugging in $cv(q_0) = v(q^*)$, $R(q_0) = \frac{q_0}{cq^*} R(q^*)$, and $c = 0.85$, we have

$$\left(\frac{0.85 - x}{q^* - q_0} + \frac{1}{q}\frac{(1 - 0.85)x}{1 - x}\right) q^* R(q^*).$$

Integrating over q from q_0 to q^*, we have

$$\int_{q_0}^{q^*} \hat{R}(q)dq \geq \left((0.85 - x) + \ln\left(\frac{1}{x}\right)\frac{0.15x}{1 - x}\right) q^* R(q^*). \tag{15}$$

Summing up (14) and (15) gives

$$\int_0^{q^*} \hat{R}(q)dq \geq \left((0.85 - x) + \ln\left(\frac{1}{x}\right)\frac{0.15x}{1 - x} + 0.815x^2\right) q^* R(q^*).$$

We would like to minimize $f(x)$ for $x \in [0, 1]$, where

$$f(x) = (0.85 - x) + \ln\left(\frac{1}{x}\right)\frac{0.15x}{1 - x} + 0.815x^2.$$

Taking the derivative we have

$$f'(x) = -1 - \frac{0.15}{1 - x} - \frac{0.15\ln(x)}{(1 - x)^2} + 1.63x.$$

This function has two roots in $x \in [0, 1]$, at $x \in [0.546, 0.547]$ ($f'(0.546) < -3 * 10^{-5}$

and $f'(0.547) > 10^{-3}$) and at $x \in [4.7 \cdot 10^{-4}, 4.8 \cdot 10^{-4}]$ ($f'(4.7 \cdot 10{-}4) > 1.3 \cdot 10^{-3}$ and $f'(4.8 \cdot 10^{-4}) < 1.9 \cdot 10^{-3}$). Testing the second derivative

$$f''(x) = -\frac{0.15}{(1-x)^2} - \frac{0.3\ln(x)}{(1-x)^3} - \frac{0.15}{x(1-x)^2} + 1.63$$

we have that $x \approx 0.546$ is a minimum point ($f''(0.546) \approx 1.5$) and $x \approx 4.7 * 10^{-4}$ is a local maximum ($f''(4.7 * 10^{-4}) \approx -315$). Therefore, the only remaining point we need to test is $x = 0$ (the end of the interval $[0,1]$) and we have $\lim_{x\to 0} f(x) = 0.85$. ∎

Proof of Theorem 5.2. Plugging in the exponential distribution in Lemma 5.7, we obtain

$$\int_{q^*}^1 \hat{R}(q)dq \geq \frac{c}{(c+1)^2}\frac{1-(q^*)^{c+1}+(q^*)^{c+1}\ln(q^*)^{c+1}}{-(q^*)\ln(q^*)}R(q^*).$$

Combining this with Lemma 5.9 yields

$$\int_0^1 \hat{R}(q)dq \geq \left(0.656q^* + \frac{c}{(c+1)^2}\frac{1-(q^*)^{c+1}+(q^*)^{c+1}\ln(q^*)^{c+1}}{-(q^*)\ln(q^*)}\right)R(q^*).$$

We would like to lower bound $f(x)$, where

$$f(x) = 0.656x + \frac{c}{(c+1)^2}\left(-x^{-1}\ln^{-1}(x) + x^c\ln^{-1}(x) - (c+1)x^c\right).$$

The derivative $f'(x)$ is

$$0.656 + \frac{c}{(c+1)^2}\left(x^{-2}\ln^{-1}(x) + x^{-2}\ln^{-2}(x) + cx^{c-1}\ln^{-1}(x) - x^{c-1}\ln^{-2}(x) - c(c+1)x^c\right).$$

For $c = 0.85$, we have a root at some $x \in [0.544, 0.545]$ as $f'(0.544) < -1.7 \cdot 10^{-4}$ and $f'(0.545) > 9 \cdot 10^{-4}$. By testing the second derivative, this is a local minimum and the only root in the interval. The minimal value is at least 0.589 and the theorem follows. ∎

6. Single-sample negative results. First, we note that identity pricing is an optimal deterministic strategy for regular distributions.

THEOREM 6.1. *No deterministic* 1*-pricing strategy is better than a* $\frac{1}{2}$*-approximation for regular distributions.*

Proof. Distributions with triangle revenue curves (with vertices $(0,0)$, $(1,0)$, and $(q^*, R(q^*))$) are commonly considered to be the worst-case regular distributions because they have the least concave revenue curves. In particular, we consider two such distributions: a point mass at v, whose revenue curve is a triangle with $(q^*, R(q^*)) = (1, v)$, and the distribution $F(v) = 1 - \frac{1}{v+1}$, whose revenue curve is a triangle with $(q^*, R(q^*)) = (0, 1)$.

To achieve a nontrivial approximation ratio when the distribution is a point mass at v, $p(v) \leq v$ must hold. Then, consider the second distribution. The revenue at price v is $\frac{v}{v+1}$, which is strictly increasing in v. So every deterministic pricing algorithm that satisfies $p(v) \leq v$ gets revenue less than or equal to that of the identity pricing algorithm $p(v) = v$, which is $\frac{1}{2}$-approximate for the second distribution. ∎

Next, we turn to MHR distributions. We first present a negative result that holds for every continuously differentiable pricing.

THEOREM 6.2. *No continuously differentiable* 1*-pricing strategy is better than* 0.677*-approximate for MHR distributions.*

Proof. Let us first assume that the pricing function is linear, i.e., $b(v) = cv$. As in the proof of Theorem 6.1, the algorithm has a finite approximation ratio only if $c \le 1$.

We consider exponential distributions and truncated exponential distributions, where all values higher than some threshold v^* in an exponential distribution are replaced with a uniform distribution over $[v^*, v^* + \alpha]$, where v^* and α are parameters to be determined later. We show that no scaling parameter c can achieve better than a 0.67-approximation in both the truncated and untruncated cases.

Exponential distribution. Consider the approximation ratio w.r.t. the exponential distribution $D^{\exp}$. Given a sample with quantile q, its value is $v^{\exp}(q) = -\ln q$. So $b(v(q)) = -c \ln q = -\ln q^c$ and the sale probability is q^c. The expected revenue of $b(v)$ is

$$\int_0^1 -q^c \ln q^c dq = \frac{c}{(c+1)^2}.$$

Recall that $R^{\exp}(q) = -q \ln q$, which is maximized at $q = \frac{1}{e}$ with optimal revenue $\frac{1}{e}$. So the approximation ratio w.r.t. $D^{\exp}$ is $\frac{ec}{(c+1)^2}$. Note that this immediately gives a lower bound of $\frac{e}{4} \approx 0.68$. If $c \le 0.878$, then $\frac{ec}{(c+1)^2} < 0.677$. For now on, we assume that $c \ge 0.878$.

Truncated exponential distribution. Let $q^* = 0.43$ and consider an exponential distribution such that $v(q^*) = 1$. Truncate the exponential distribution at q^* with a uniform distribution over $[1, 1+\alpha]$ with $\alpha = 0.74$. Hence, for $q \in [q^*, 1]$, $v(q) = \frac{\ln q}{\ln q^*}$; for $q \in [0, q^*]$, $v(q) = 1 + \alpha(1 - \frac{q}{q^*})$. It is easy to check that the revenue is maximized at $q = q^*$ with maximal revenue q^*.

Next, we upper bound the expected revenue of $b(v)$. Consider first the contribution from quantiles $q \in [\frac{1}{e}, 1]$. The analysis is similar to that of $D^{\exp}$, except that the values are scaled up by $-\frac{1}{\ln q^*}$. So this part contributes

$$\begin{aligned} -\frac{1}{\ln q^*} \int_{q^*}^1 -q^c \ln q^c dq &= -\frac{1}{\ln q^*} \frac{c}{(c+1)^2} \left(1 - (q^*)^{c+1} + (q^*)^{c+1} \ln(q^*)^{c+1} \right) \\ &\le 2.7556 \frac{c}{(c+1)^2} \left(1 - (q^*)^{c+1} + (q^*)^{c+1} \ln(q^*)^{c+1} \right) q^* \ . \end{aligned}$$

Now consider the quantiles that are smaller than q^* and have values at most $\frac{1}{c}$, i.e., the corresponding reserve price is smaller than 1. These are quantiles $q \in [\frac{c\alpha + c - 1}{c\alpha} q^*, q^*]$. We upper bound the expected revenue by the optimal revenue q^* when the sampled quantile is in this interval. So this part contributes at most

$$\left(q^* - \frac{c\alpha + c - 1}{c\alpha} q^* \right) q^* = \left(\frac{1-c}{c\alpha} \right) (q^*)^2 \le 0.5811 \left(\frac{1}{c} - 1 \right) q^* \ .$$

Finally, consider quantiles $q \in [0, \frac{c\alpha + c - 1}{c\alpha} q^*]$. The corresponding value is $v(q) = 1 + \alpha(1 - \frac{q}{q^*})$ and the reserve price is $b(v) = c(1 + \alpha(1 - \frac{q}{q^*}))$. The sale probability at this price is $\frac{1 + \alpha - b(v)}{\alpha} q^*$. So this part contributes

$$\int_0^{\frac{c\alpha + c - 1}{c\alpha} q^*} b(v) \frac{1 + \alpha - b(v)}{\alpha} q^* dq.$$

Note that $db(v) = -\frac{c\alpha}{q^*}dq$, $b = 1$ when $q = \frac{c\alpha+c-1}{c\alpha}q^*$, and $b = c(1+\alpha)$ when $q = 0$. So the above equals

$$\begin{aligned}
&\frac{(q^*)^2}{c\alpha^2}\int_0^{c(1+\alpha)} b(v)(1+\alpha-b(v))db(v)\\
&\quad= \frac{(q^*)^2}{c\alpha^2}\left(\frac{1+\alpha}{2}\left(\left(c(1+\alpha)\right)^2-1\right)-\frac{1}{3}\left(\left(c(1+\alpha)\right)^3-1\right)\right)\\
&\quad= \frac{(q^*)^2}{c\alpha^2}\left(\left(\frac{c^2}{2}-\frac{c^3}{3}\right)(1+\alpha)^3-\frac{1+\alpha}{2}+\frac{1}{3}\right)\\
&\quad\le \left(-1.3789c^2+2.0683c-0.4215\frac{1}{c}\right)q^*.
\end{aligned}$$

Putting everything together and dividing the expected revenue by the optimal revenue q^*, the approximation ratio is at most

$$\begin{aligned}
2.7556\frac{c}{(c+1)^2}\left(1-(q^*)^{c+1}+(q^*)^{c+1}\ln(q^*)^{c+1}\right)+0.5811\left(\frac{1}{c}-1\right)&\\
+\left(-1.3768c^2+2.2683c-0.4215\frac{1}{c}\right)&.
\end{aligned}$$

We numerically maximize the above function over $c \in [0.878, 1]$. It is decreasing in the interval $[0.878, 1]$ and takes value about 0.6762 at $c = 0.878$. This completes the proof for linear pricing functions.

Continuously differentiable pricing functions. Next, we explain how to reduce the case of a continuously differentiable pricing function to the linear pricing function case. Since $b(v)$ is continuously differentiable, for any $\delta > 0$, there exists $\epsilon > 0$ such that for any $v \in [0, \epsilon]$, $|b'(v) - b'(0)| < \delta$, i.e., in this neighborhood of 0, $b(v)$ behaves like a linear function with slope approximately $b'(0)$ (up to error δ). So we can handle them like the linear case.

Formally, if $b'(0) < \frac{1}{2}$, then $b(v) \le (\frac{1}{2}+\delta)v$ for $v \le \epsilon$. So its approximation ratio w.r.t. a point mass at v is at most $\frac{1}{2}+\delta < 0.677$ for sufficiently small δ.

Next, assume $b'(0) > \frac{1}{2}$. Let us scale down the values in the exponential distribution and the truncated exponential distribution from the linear case such that all values are less than ϵ.[13] For any sampled value v, the expected revenue is $b(v)q(b(v))$, where $(f'(0)-\delta)v < b(v) < (f'(0)+\delta)v$. So

$$b(v)q(b(v)) \le (f'(0)+\delta)vq((f'(0)-\delta)v) = \frac{f'(0)+\delta}{f'(0)-\delta}(f'(0)-\delta)vq((f'(0)-\delta)v)\ ,$$

which is at most $\frac{f'(0)+\delta}{f'(0)-\delta}$ times larger than the revenue of linear pricing function $(f'(0)-\delta)v$. Letting δ go to zero completes the proof. ∎

Next, we present a slightly weaker bound that applies to all deterministic 1-pricing algorithms (continuously differentiable or not).

THEOREM 6.3. *No deterministic* 1*-pricing strategy is better than* $\frac{e}{4} \approx 0.68$*-approximate for MHR distributions.*

[13]Note that the support of the exponential distribution spans all nonnegative real numbers. So instead of scaling to make all values smaller than ϵ, we will make sure $1-10^{-5}$ fraction of the values are smaller than ϵ; the remaining values can change the approximation ratio by at most 10^{-5}.

Proof. By previous arguments, it suffices to consider only (deterministic) pricing functions with $p(v) \leq v$ for all v. We consider a distribution over value distributions. Draw λ from an exponential distribution with parameter γ, i.e., the density of λ is $\gamma e^{-\gamma\lambda}$, where γ is a parameter to be determined later. Let the value distribution be an exponential distribution with parameter λ, i.e., the density of v is $\lambda e^{-\lambda v}$.

We first compute the best response pricing algorithm $p(v)$ subject to $p(v) \leq v$ for this case. The expected revenue of $p(v)$ is

$$\begin{aligned}
R &= \int_0^\infty \left[\int_0^\infty \lambda e^{-\lambda v} p(v) e^{-\lambda p(v)} dv\right] \gamma e^{-\gamma\lambda} d\lambda \\
&= \gamma \int_0^\infty \left[\int_0^\infty \lambda e^{-\lambda(v+p(v)+\gamma)} p(v) d\lambda\right] dv \\
&= \gamma \int_0^\infty \left[\int_0^\infty \lambda e^{-\lambda(v+p(v)+\gamma)} p(v) d\lambda\right] dv \\
&= \gamma \int_0^\infty \left[\frac{p(v)}{v+p(v)+\gamma} \int_0^\infty (v+p(v)+\gamma)\lambda e^{-\lambda(v+p(v)+\gamma)} d\lambda\right] dv.
\end{aligned}$$

Note that $\frac{p(v)}{(v+p(v)+\gamma)^2}$ is maximized at $p(v) = v$ for $p(v) \leq v$. So the best response is $p(v) = v$. Given any λ, the optimal revenue is $\frac{1}{e\lambda}$, and the expected revenue of $p(v) = v$ is $\frac{1}{4\lambda}$. So the approximation ratio is at most $\frac{e}{4} \approx 0.68$, as desired. $\square$

REFERENCES

[1] M. Anthony and P. L. Bartlett, *Neural Network Learning: Theoretical Foundations*, Cambridge University Press, Cambridge, 1999.

[2] P. Azar, C. Daskalakis, S. Micali, and S. M. Weinberg, *Optimal and efficient parametric auctions*, in Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), SIAM, Philadelphia, 2013, pp. 596–604.

[3] T. Baigneres, P. Junod, and S. Vaudenay, *How far can we go beyond linear cryptanalysis?*, in Advances in Cryptology (Asiacrypt), Springer, Berlin, 2004, pp. 432–450.

[4] M.-F. Balcan, A. Blum, J. D. Hartline, and Y. Mansour, *Reducing mechanism design to algorithm design via machine learning*, J. Comput. System Sci., 74 (2008), pp. 1245–1270.

[5] S. Baliga and R. Vohra, *Market research and market design*, Adv. Theor. Econ., 3 (2003), 1059.

[6] J. Bulow and P. Klemperer, *Auctions versus negotiations*, Adv. Theor. Econ., 86 (1996), pp. 180–194.

[7] S. Chawla, J. D. Hartline, and D. Nekipelov, *Mechanism design for data science*, in Proceedings of the 15th ACM Conference on Economics and Computation, ACM, New York, 2014, pp. 711–712.

[8] A. Chiesa, S. Micali, and Z. A. Zhu, *Mechanism design with approximate valuations*, in Proceedings of the Innovations in Theoretical Computer Science Conference, ACM, New York, 2012, pp. 34–38.

[9] R. Cole and T. Roughgarden, *The sample complexity of revenue maximization*, in Proceedings of the 46th Annual ACM Symposium on Theory of Computing, ACM, New York, 2014, pp. 243–252.

[10] J. Crémer and R. P. McLean, *Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent*, Econometrica, 53 (1985), pp. 345–361.

[11] P. Dhangwatnotai, T. Roughgarden, and Q. Yan, *Revenue maximization with a single sample*, Games Econom. Behav., 91 (2015), pp. 318–333.

[12] C. Dwork, G. N. Rothblum, and S. Vadhan, *Boosting and differential privacy*, in 51st Annual IEEE Symposium on Foundations of Computer Science, FOCS, IEEE Computer Society, Los Alamitos, CA, 2010, pp. 51–60.

[13] H. Fu, N. Haghpanah, J. D. Hartline, and R. Kleinberg, *Optimal auctions for correlated buyers with sampling*, in Proceedings of the 15th ACM Conference on Economics and Computation, ACM, New York, 2014, pp. 23–36.

[14] H. Fu, N. Immorlica, B. Lucier, and P. Strack, *Randomization beats second price as a prior-independent auction*, in Proceedings of the 16th ACM Conference on Economics and Computation, ACM, New York, 2015, p. 323.
[15] A. V. Goldberg, J. D. Hartline, A. R. Karlin, M. Saks, and A. Wright, *Competitive auctions*, Games Econom. Behav., 55 (2006), pp. 242–269.
[16] J. D. Hartline, *Mechanism Design and Approximation*, manuscript.
[17] J. D. Hartline, V. S. Mirrokni, and M. Sundararajan, *Optimal marketing strategies over social networks*, in Proceedings of the 17th International Conference on World Wide Web, Curran, Red Hook, NY, 2008, pp. 189–198.
[18] J. D. Hartline and T. Roughgarden, *Optimal Platform Design*, preprint, arXiv:1412.8518, 2014.
[19] Z. Huang, Y. Mansour, and T. Roughgarden, *Making the most of your samples*, in Proceedings of the Sixteenth ACM Conference on Economics and Computation, ACM, New York, 2015, pp. 45–60.
[20] Z. Neeman, *The effectiveness of English auctions*, Games Econom. Behav., 43 (2003), pp. 214–238.
[21] M. Ostrovsky and M. Schwarz, *Reserve prices in internet advertising auctions: A field experiment*, in Proceedings of the 12th ACM Conference on Economics and Computation, ACM, New York, 2011, pp. 59–60.
[22] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day, San Francisco, 1964.
[23] I. Segal, *Optimal pricing mechanisms with unknown demand*, Amer. Econ. Rev., 93 (2003), pp. 509–529.
[24] B. Sivan and V. Syrgkanis, *Vickrey auctions for irregular distributions*, in Proceedings of the 9th International Workshop on Internet and Network Economics, Springer, Berlin, 2013, pp. 422–435.
[25] L. G. Valiant, *A theory of the learnable*, Commun. ACM, 27 (1984), pp. 1134–1142.