arXiv:1907.04220v3 [cs.GT] 28 Apr 2022

Robust Revenue Maximization Under Minimal Statistical Information*

Yiannis Giannakopoulos[†] Diogo Poças[‡] Alexandros Tsigonias-Dimitriadis[§]

April 28, 2022

Abstract

We study the problem of multi-dimensional revenue maximization when selling m items to a buyer that has additive valuations for them, drawn from a (possibly correlated) prior distribution. Unlike traditional Bayesian auction design, we assume that the seller has a very restricted knowledge of this prior: they only know the mean μ_j and an upper bound σ_j on the standard deviation of each item's marginal distribution. Our goal is to design mechanisms that achieve good revenue against an ideal optimal auction that has *full* knowledge of the distribution in advance. Informally, our main contribution is a tight quantification of the interplay between the dispersity of the priors and the aforementioned robust approximation ratio. Furthermore, this can be achieved by very simple selling mechanisms.

More precisely, we show that selling the items via separate price lotteries achieves an $O(\log r)$ approximation ratio where $r = \max_j(\sigma_j/\mu_j)$ is the maximum coefficient of variation across the items. To prove the result, we leverage a price lottery for the single-item case. If forced to restrict ourselves to deterministic mechanisms, this guarantee degrades to $O(r^2)$. Assuming independence of the item valuations, these ratios can be further improved by pricing the full bundle. For the case of identical means and variances, in particular, we get a guarantee of $O(\log(r/m))$ which converges to optimality as the number of items grows large. We demonstrate the optimality of the above mechanisms by providing matching lower bounds. Our tight analysis for the single-item deterministic case resolves an open gap from the work of Azar and Micali [ITCS'13].

As a by-product, we also show how one can directly use our upper bounds to improve and extend previous results related to the *parametric auctions* of Azar et al. [SODA'13].

*Supported by the Alexander von Humboldt Foundation with funds from the German Federal Ministry of Education and Research (BMBF). The second author further acknowledges the support of FCT via LASIGE Research Unit, ref. UIDB/00408/2020 and ref. UIDP/00408/2020. The last author further acknowledges the support of the German Research Foundation (DFG) within the Research Training Group AdONE (GRK 2201). Part of this work was done while the first two authors were members of the Operations Research group at the Technical University of Munich.

An extended abstract of this paper appeared in WINE'20 [35].

[†]Friedrich-Alexander-Universität Erlangen-Nürnberg, Department of Data Science. Email: yiannis.giannakopoulos@fau.de

[‡]LASIGE, Faculdade de Ciências, Universidade de Lisboa. Email: dmpocas@fc.ul.pt

[§]Technical University of Munich. Email: alexandros.tsigonias@tum.de. Associated researcher with the Research Training Group "Advanced Optimization in a Networked Economy" (GRK 2201), funded by the German Research Foundation (DFG).

1 Introduction

Optimal auction design is one of the most well-studied and fundamental problems in (algorithmic) mechanism design. In the traditional Myersonian [57] setting, an auctioneer has a single item for sale and there are n interested bidders. Each bidder has a (private) valuation for the item which, intuitively, represents the amount of money they are willing to spend to buy it. The standard Bayesian approach is to assume that the seller has only an incomplete knowledge of these valuations, in the form of a prior joint distribution F. A selling mechanism receives bids from the buyers and then decides to whom the item should be allocated (which, in general, can be a randomized rule) and for what price. The goal is to design a truthful[1] selling mechanism that maximizes the auctioneer's revenue, in expectation over F.

Myerson [57] provided a complete and very elegant solution for this problem when bidder valuations are independent, that is, F is a product distribution. In particular, when the distributions are identical and further satisfy a regularity assumption, the optimal mechanism takes the very satisfying form of a second-price (Vickrey) auction with a reserve price. Unfortunately, in general these characterizations collapse when we move to multi-dimensional environments where there are $m > 1$ items for sale. Multi-item optimal auction design is one of the most challenging and currently active research areas of mechanism design. Given that the exact description of the revenue maximizing auctions in such settings is a notoriously hard task, there is an impressive stream of recent papers, predominantly from the algorithmic game theory community, that try to provide good approximation guarantees to the optimal revenue.

The critical common underlying assumption throughout the aforementioned optimal auction design settings is that the seller has *full knowledge* of the prior joint distribution F of the bidders' valuations. In many applications though, this might arguably be an unrealistic assumption to make: usually an auctioneer can derive some distributional properties about the bidder population, but to completely determine the actual distribution would require enormous resources. Thus, inspired by the parametric auctions of Azar and Micali [2] for the single-dimensional case, we would like to be able to design robust auctions that (1) make only use of *minimal statistical information* about the valuation distribution, namely its mean and variance; and (2) still provide good revenue guarantees even in the worst case against an adversarial selection of the actual distribution F; in particular, *no further assumptions* (e.g., independence of item valuations or regularity) should in general be made about F. This is our main goal in this paper.

1.1 Related Work

As mentioned in the introduction, there has been an impressive stream of recent work on optimal [16, 29, 34, 39, 51] and approximately-optimal [5, 15, 20, 41, 48, 61, 67] multi-dimensional auction design, which tries to extend the traditional, single-dimensional auction setting studied in the seminal paper of Myerson [57]. A prominent characteristic that can often be seen in these papers is the "simplicity vs optimality" approach: knowing the computational hardness [23, 24, 28] and structural complexity [29, 40] of describing exact optimality, emphasis is placed on designing both simple and practical mechanisms that can still provide good revenue guarantees. Of course, this idea can be traced back to the work of Hartline and Roughgarden [43] and Bulow and Klemperer [13] for the single-dimensional setting. For a more thorough overview we refer to the recent review article of Roughgarden and Talgam-Cohen [60] and the textbook of Hartline [42].

Related to this, and placed under the general theme of what has come to be known as "Wilson's doctrine" [66] (see also [52, Section 5.2]), there has also been significant effort towards

[1] In such mechanisms it is to the buyers' best interest to honestly report their actual private valuation when bidding. For formal definitions, see Section 2.1. We notice here that this is essentially without loss to the revenue maximization objective, due to the Revelation Principle (see, e.g., [58]).

the direction of *robust* revenue maximization: designing auctions that make as few assumptions as possible on the seller's prior knowledge about the bidders' valuations for the items. Examples include models where the auctioneer can perform quantile queries [22] or knows some estimate of the actual prior [9, 14, 49]. Another line of work studies robustness with respect to the correlation of valuations across bidders or items [8, 18, 38]. Other approaches regarding the parameterization of partial distributional knowledge were considered by Dütting et al. [32] and Bandi and Bertsimas [6]. See also the recent survey by Carroll [19].

Most relevant to our work in the present paper is the model of *parametric auctions*, introduced by Azar and Micali [2]. More specifically, they study single-dimensional (digital goods and single-item) auction settings with independent item valuations, under the assumption that the seller has only access to the mean μ_i and the variance σ_i^2 of each buyer's i prior distribution. Using Chebyshev-like tail bounds, they show that for the special single-bidder, single-item case, deterministically pricing at a multiple of the standard deviation below the mean, i.e. offering a take-it-or-leave-it price of $\mu - k \cdot \sigma$, guarantees an approximation ratio of $\tilde{\rho}(r)$, where $\tilde{\rho}$ is an increasing function taking values in $[1, \infty)$ and $r = \sigma/\mu$. In Appendix C, we actually quantify this bound and show that it grows quadratically. Under an extra assumption of Monotone Hazard Rate (MHR), they show how the even simpler selling mechanism that just prices at μ achieves an approximation ratio of e.

It is interesting to notice here that Azar and Micali [2] provide an *exact* solution, for deterministic mechanisms, to the robust optimization problem of maximizing the expected revenue. Then, they use this maximin revenue-optimal mechanism and compare it to the optimal social welfare (which is trivially also an upper bound on the optimal revenue), to finally derive their upper bound guarantee on the approximation ratio of revenue. As such, their results are not tailored to be tight for the *ratio* benchmark. As a matter of fact, in [4] the authors also provide an explicit lower bound that can be written as $1 + r^2$. This is an important motivating factor for our work, since one of our main goals in this paper is to close these gaps and provide tight approximation ratio bounds.

Azar et al. [3] use a clever reduction (see also the work of Chawla et al. [21]) to show how these results can be paired with the work of Dhangwatnotai et al. [30] regarding the VCG mechanism with reserves, in order to design parametric auctions for very general single-dimensional settings. In particular, they show how in matroid-constrained environments with the extra assumption of regularity on the prior distributions (or MHR for more general downward-closed environments), using the aforementioned parametric prices as lazy reserves guarantees a $2\tilde{\rho}(r)$-approximation to the optimal (Myersonian) revenue and a $\tilde{\rho}(r)$-approximation to the optimal social welfare. Here $r = \max_i \sigma_i/\mu_i$.

Another work which is close to ours is that of Carrasco et al. [17]. The authors essentially extend the model of Azar and Micali [2] to randomized mechanisms, solving the maximin robust optimization problem with respect to revenue. Again, in principle their results cannot be immediately translated to tight bounds for the approximation ratio; however, unlike the deterministic case for which in the present paper we have to design a new mechanism in order to achieve ratio optimality, we will show that the maximin optimal lottery of Carrasco et al. [17] is actually also optimal for the ratio benchmark.

Sample access vs knowledge of moments Another stream of research studies models where the auctioneer has sample access to the distribution [1, 26, 27, 30, 33, 37, 44, 45, 54, 64]. It is not hard to imagine scenarios where such access to individual past data might be infeasible or impractical, e.g. due to data protections and privacy restrictions. Furthermore, there might also exist computational limitations in representing a distribution, or storing and reasoning with a large number of samples. In such settings, it is more natural to assume access to only some statistical aggregates of the underlying data, such as the mean and the standard deviation.

From a theoretical perspective, the sample access model is incomparable with the moment-based model of the present paper, as they rely on different distributional assumptions. In particular, independence, regularity and/or upper bounds on the support are standard assumptions in the aforementioned sample complexity papers. As a matter of fact, these are necessary to derive non-trivial results (see e.g. the counterexample of Cole and Roughgarden [26, Footnote 3]). Furthermore, if independence is dropped, Dughmi et al. [31] demonstrate that an exponential number of samples is required in order to achieve a constant-factor approximation to the optimal revenue. In our setting, on the other hand, we require none of the above. However, we do assume (as a design principle) exact knowledge of the mean and an upper bound on the standard deviation. This information cannot be retrieved exactly via any finite amount of samples, although intervals of confidence can be used to estimate it; we leave as future work the study of the revenue maximization problem when having only approximate knowledge of the distribution moments.

Maximin robustness for approximation ratio vs revenue The literature so far has focused on solving the maximin robust optimization problem with respect to revenue, but we chose to formulate and use the robust approximation ratio instead. Apart from being a natural choice for a computer scientist, since ratios of this sort are often used in algorithm design, it can also complement the existing objective and offer further insights.

Both quantities have strengths from a theoretical and practical perspective, and their comparison can be subject to a broader discussion (see also Section 7). In any case, we believe that the robust approximation ratio will come in handy in some scenarios. For instance, in large markets, where the seller's task of approximately quantifying the revenue that they expect to obtain might become daunting, the "scale-free" approximation ratio can be helpful. It is a more interpretable objective because the seller can observe their loss due to the limited statistical information as just a percentage of the full knowledge benchmark. Moreover, they can easily follow how changes in the mean or the standard deviation drive the optimal they can achieve with a robust mechanism away from the revenue of an ideal optimal auction.

Similarly, the robust approximation ratio is probably the suitable benchmark for understanding the effect of the knowledge of higher moments. An ambitious, meaningful open question is to show how the seller should design robustly optimal mechanisms when they know up to N moments of the distribution. By assuming in our current model knowledge of also, e.g., the third moment, the seller's revenue will increase since she learns more about the underlying distribution. The ratio benchmark can show us *at which rate* the maximin revenue is improving every time we add the knowledge of a higher moment and when it becomes near-optimal.

1.2 Results and Techniques

The main focus of our paper is a multi-dimensional auction setting where a single bidder has additive valuations for m items, drawn from a joint probability distribution F. We make no further assumptions on F; in particular, we do not require F to be a product distribution nor do we enforce any kind of regularity. The seller knows only the mean μ_j and (an upper bound on) the standard deviation σ_j of each item's j marginal distribution. Based on this limited statistical information, they are asked to fix a truthful (possibly randomized) mechanism to sell the items. Then, an adversary chooses the actual distribution F (respecting, of course, the statistical (μ_j, σ_j)-information) and the seller realizes the expected revenue of the auction, in the standard Bayesian way, in expectation with respect to F. The main quantity of interest, which we call the *robust approximation ratio* is the ratio of the optimal revenue (which has full knowledge of F in advance) to this revenue.

Our worst-case, min-max approach is similar in spirit to the previous work of Azar et al. [3], Azar and Micali [4] and Carrasco et al. [17]. However, the critical difference in the present paper is that our main goal is to optimize the *ratio* against the optimal revenue and not just the expected revenue of the selling mechanism on its own. It turns out that, similarly to the aforementioned previous work, our bounds can be stated with respect to the ratio $r_j = \sigma_j/\mu_j$ of each item's marginal distribution. This is an important statistical quantity called the *coefficient of variation (CV)*; it is essentially a "unit-independent" measure of the dispersion of the distribution (see, e.g., [55] or [46, Sec. 2.21]).

In Section 2 we formally introduce our model and necessary notation. In the following two sections we focus on the single-item case, since this will be the building block for all our results. In particular, in Section 3 we show that the robust approximation ratio of deterministic mechanisms is *exactly* $\rho_D(r) \approx 1 + 4r^2$ (see Definition 1), closing a gap open from the work of Azar and Micali [4]. Similarly to previous work, in order to achieve this we solve exactly the corresponding min-max problem (see Lemma 2); however, the method and the solution itself have to be different, since we are dealing with the ratio, which is a more "sensitive" quantity than the revenue on its own. By "sensitive" we mean that its value changes in a less smooth and more unpredictable way for small perturbations of the distribution and the mechanism.

Next, in Section 4 we deal with general randomized auctions and we show that a lottery proposed by Carrasco et al. [17], which we term *log-lottery*, although designed for a different objective achieves an approximation ratio of $\rho(r) \approx 1 + \ln(1 + r^2)$ (see Definition 1) in our setting, which is asymptotically optimal. We start with a quantitative analysis of the log-lottery mechanism (Theorem 2). In particular, we show an upper bound to the robust approximation ratio that grows logarithmically in r. This bound already establishes a strong separation between the power of deterministic and randomized mechanisms. The question then becomes if a different randomized selling mechanism can achieve a sublogarithmic or even constant upper bound. We answer this in the negative by showing that the logarithmic upper bound is asymptotically tight. The construction of the lower bound instance (Theorem 3) is arguably the most technically challenging part of our paper, and is based on a novel utilization of Yao's minimax principle (see also Appendix D) that might be of independent interest for deriving robust approximation lower bounds in other Bayesian mechanism design settings as well. Informally, the adversary offers a distribution over two-point mass distributions, finely-tuned such that the resulting mixture becomes a truncated "equal-revenue style" distribution (see Fig. 2c). The main difference to other settings in the literature where Yao's principle is applied is that the adversary has to randomize over probability distributions, which form an infinite-dimensional space. We can imagine this as a space of "distributions over distributions". This introduces new technical challenges since the adversary's model of randomization needs to be properly defined, and more importantly, Yao's principle does not hold anymore. Thus, our goal is twofold: we need to carefully describe how the adversary constructs a space of distributions over distributions and then show that we can extend Yao's principle to such spaces.

It is important to restate that we work under the assumption that we know an *upper bound* on the standard-deviation σ and not its exact value. Although this makes our upper bounds more powerful, it is not a source of "artificial" additional power for the adversary when designing our lower bounds. We formalize this in Lemma 5. Furthermore, this helps us to formally demonstrate (see Proposition 2) that our aforementioned, Yao-based, lower bound construction lies at the "border of simplicity" of any non-trivial lower bound.

In Section 5 we demonstrate how the $O(\log r)$-approximate mechanism of the single-item case can be utilized to provide optimal approximation ratios for the multi-dimensional case of m items as well. More specifically, we show that selling each item j separately using the log-lottery guarantees an approximation ratio of $\rho(r_{\max})$ where $r_{\max} = \max_j r_j$ is the maximum CV across the items. If the seller has extra information that item valuations are independent (that is, F is a product distribution), then switching to a lottery that offers all items in a single full bundle can give an improved approximation ratio of $\rho(\bar{r})$, where $\bar{r} = \sqrt{\sum_j \sigma_j^2}/\sum_j \mu_j$ is the CV of the average valuation. We complement these upper bounds by tight lower bounds in Theorem 5; these constructions have at their core the single-item lower bound, but they take care of delicately assigning valuations to the remaining items so that they respect independence and the common prior statistical information. We want to highlight that the lower bound of Theorem 5 is strong enough to hold for *any* number of items and *any* choice of coefficients of variation $r_1, r_2, \dots, r_m$. An interesting corollary of our upper bounds (Corollary 1) is that for the special case of independent valuations with the same mean and variance, the approximation ratio is at most $\rho\left(\frac{\sigma}{\mu\sqrt{m}}\right)$, converging to optimality as the number of items grows large.

In Section 6.1 we diverge from our main model to discuss some additional "peripheral" results that can be deduced as direct corollaries of previous work combined with our upper bounds, in a "black-box" way. First, we study the single-dimensional, multi-bidder setting of *parametric auctions* introduced by Azar and Micali [4]. More specifically, we show how the positive results derived in Azar et al. [3, Theorem 4.3] can be further improved: running VCG with lazy reserve prices drawn from the log-lottery guarantees a $2\rho(r)$ approximation to the optimal Myersonian revenue (Corollary 2).

Secondly, in Section 6.2 we discuss how a relaxation of our model that only assumes knowledge of the mean (that is, without any information about the variance σ^2) can still produce good robust approximation ratios under an extra regularity assumption. More precisely, in Proposition 3 we give an upper bound on the approximation ratio of the mechanism that just offers the mean μ as a take-it-or-leave-it price, under the extra assumption that the item's valuation distribution is λ-regular (see Fig. 3a); this is a general notion of regularity that interpolates smoothly between regularity à la Myerson ($\lambda = 1$) and the Monotone Hazard Rate (MHR) condition ($\lambda = 0$); see, e.g., [36, 62]. This result extends the e-approximation for MHR distributions of Azar and Micali [2, Theorem 3]. Finally, we provide a more detailed characterization of the relationship between the knowledge of λ-regularity and knowledge of σ, with respect to the resulting robust approximation ratio upper bound (see Fig. 3b).

Size of the coefficient of variation It is worth discussing briefly the implications of the size of the CV, our main quantity of interest, for our results. We can observe that our upper bounds do *not* increase with the number of items m; as a matter of fact, for the case of independently distributed items with the same mean and variance, the upper bound even decreases with respect to m. Although the CV of a distribution could be arbitrarily large in general, one could argue that, for many practical scenarios, it is unlikely to encounter data with very large dispersion. From a theoretical perspective, note that the CV is actually bounded for important special classes of distributions, like MHR (which include, e.g., the truncated normal, uniform, exponential and gamma [7]) and, more generally, λ-regular for a fixed $\lambda < 1/2$ (see (14)). Furthermore, for general distributions, if one assumes that the CV of the item marginals are bounded by a universal constant, then our bounds yield a constant robust approximation ratio to the optimal pricing, even for correlated distributions (and regardless of the number of items).

2 Preliminaries

2.1 Model and Notation

A real nonnegative random variable will be called (μ, σ)-*distributed* if its expectation is μ and its standard deviation is *at most* σ. We let $\mathbb{F}_{\mu,\sigma}$ denote the class of (μ, σ) distributions. We shall also briefly (see Lemma 5) discuss the restriction to distributions with standard deviation of *exactly* σ; this subclass will be denoted by $\mathbb{F}^{=}_{\mu,\sigma}$.

For the most part of this paper we study auctions with m items and a single additive bidder, whose valuations $(v_1, \dots, v_m)$ for the items are drawn from a joint distribution F over $\mathbb{R}^m_{\geq 0}$. We denote the marginal distribution of v_j by F_j, and assume that it has finite mean and variance. In general, we make no further assumptions for F; in particular, we do not assume independence of the random variables $v_1, \dots, v_m$ nor do we enforce any regularity or continuity assumption.

For vectors $\vec{\mu} = (\mu_1, \dots, \mu_m) \in \mathbb{R}^m_{>0}$, $\vec{\sigma} = (\sigma_1, \dots, \sigma_m) \in \mathbb{R}^m_{\geq 0}$ we denote by $\mathbb{F}_{\vec{\mu},\vec{\sigma}}$ the class of all m-dimensional distributions whose j-th marginal is (μ_j, σ_j)-distributed, for all $j = 1, \dots, m$.

A (direct revelation, possibly randomized) selling *mechanism* for a single bidder and m items is defined by a pair (x, π) where $x : \mathbb{R}^m_{\geq 0} \to [0,1]^m$ is the *allocation rule* and $\pi : \mathbb{R}^m_{\geq 0} \to \mathbb{R}_{\geq 0}$ is the *payment rule*. If the buyer submits as bid a valuation vector of $\vec{v}$, then they receive each item i with probability $x_i(\vec{v})$, and are charged (a total of) $\pi(\vec{v})$. We restrict our study to *truthful* mechanisms, which are characterized by the conditions

$$x(\vec{v}) \cdot \vec{v} - \pi(\vec{v}) \geq x(\vec{w}) \cdot \vec{v} - \pi(\vec{w}) \qquad \text{for all } \vec{v}, \vec{w}; \tag{1}$$

$$x(\vec{v}) \cdot \vec{v} - \pi(\vec{v}) \geq 0 \qquad \text{for all } \vec{v}. \tag{2}$$

Informally, the first condition states that the bidder can not be "better off" by misreporting their true valuation; the second condition, known as *individual rationality*, ensures that the bidder cannot harm themselves by truthfully participating in the mechanism.

Let $\mathbb{A}_m$ denote the space of all truthful selling mechanisms. Then, given an m-dimensional distribution F, we denote by

- $\mathrm{REV}(A; F) = \mathbb{E}_{\vec{v}\sim F}[\pi(\vec{v})]$, the expected *revenue* of A (the expectation is taken w.r.t. F);
- $\mathrm{WEL}(A; F) = \mathbb{E}_{\vec{v}\sim F}[x(\vec{v}) \cdot \vec{v}]$, the expected *welfare of* A;
- $\mathrm{OPT}(F) = \sup_{A\in\mathbb{A}_m} \mathrm{REV}(A; F)$, the optimum revenue;
- $\mathrm{VAL}(F) = \sup_{A\in\mathbb{A}_m} \mathrm{WEL}(A; F)$, the optimum welfare. By definition, this is also the welfare of a VCG auction; moreover, for a single additive bidder with a joint distribution in $\mathbb{F}_{\vec{\mu},\vec{\sigma}}$, this is just the sum of the marginal expectations, $\mathrm{VAL}(F) = \sum_{j=1}^m \mu_j$.

Note that, due to (2), we immediately have the so-called *welfare bounds* for the above quantities: for any mechanism and distribution,

$$\mathrm{REV}(A; F) \leq \mathrm{WEL}(A; F) \qquad \text{and} \qquad \mathrm{OPT}(F) \leq \mathrm{VAL}(F).$$

Our goal in this paper is to quantify the following benchmark

$$\mathrm{APX}(\vec{\mu}, \vec{\sigma}) = \inf_{A\in\mathbb{A}_m} \sup_{F\in\mathbb{F}_{\vec{\mu},\vec{\sigma}}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(A; F)}, \tag{3}$$

which we call the *robust approximation ratio.* The semantics are the following: a seller chooses the best (revenue-maximizing) selling mechanism A, given only knowledge of the means $\vec{\mu}$ and standard deviations $\vec{\sigma}$ and then an adversary ("nature") responds by choosing a worst-case "valid" distribution that respects the statistical information $\vec{\mu}$ and $\vec{\sigma}$. At some parts of our paper, we restrict our attention to *deterministic* mechanisms A; that is, mechanisms whose allocation rule satisfies $x(\vec{v}) \in \{0,1\}^m$, for all $\vec{v}$. Under this additional constraint, the quantity in (3) will be denoted by $\mathrm{DAPX}(\vec{\mu}, \vec{\sigma})$.

For the special case of a single item ($m = 1$), we know from the seminal work of Myerson [57] that an auction $A \in \mathbb{A}_1$ is truthful if and only if its allocation rule is monotone nondecreasing and the payment rule is given by $\pi(v) = v \cdot x(v) - \int_0^v x(z)\,dz$. In particular, this implies that every deterministic mechanism $A \in \mathbb{A}_1$ is completely determined by a single take-it-or-leave-it price $p \geq 0$; thus, we will feel free to sometimes abuse notation and write $\mathrm{REV}(p; F)$ instead of $\mathrm{REV}(A; F)$ if A is the deterministic auction that sells at price p.

Most importantly for our work, every randomized auction for a single item can be seen as a nonnegative random variable over prices (see Carrasco et al. [17, Footnote 10]). In particular, since the allocation rule is monotone and takes values in $[0, 1]$, it can be interpreted as the cumulative distribution of a certain randomization over prices, which assigns the item with the same probability as the original mechanism.[2] In this way, for a randomized single-item auction we can abuse notation and write $p \sim A$ to denote that a price p is sampled according to A. In this way, $\mathrm{REV}(A; F) = \mathbb{E}_{p\sim A}[\mathrm{REV}(p; F)]$.

Finally, from Myerson [57] we also know that for single-item settings the optimum revenue can always be achieved by a deterministic mechanism, that is,

$$\mathrm{OPT}(F) = \sup_{p\geq 0} \mathrm{REV}(p; F) = \sup_{p\geq 0} p \cdot (1 - F(p-)) \tag{4}$$

where we use $F(\cdot)$ for the cumulative function (cdf) of distribution F and $F(p-) = \Pr[X < p] = \lim_{x\to p^-} F(x)$, where $X \sim F$. We shall call $\mathrm{OPT}(\cdot)$ the *Myerson operator* and for now we simply observe that this is a functional mapping distributions to real nonnegative numbers.

2.2 Determinism vs Randomization

We would like to give some basic intuition on how randomization helps to hedge uncertainty. To this end, we present a simple example where a randomized strategy beats every price.

Example 1. Assume that we are facing a very restricted adversary who can choose between two distributions. Distribution A has just a point mass at 1. Distribution B is a two-point mass distribution, which returns either 0 or 2 with probability 1/2 each.

If the seller is restricted to deterministic pricing rules, it is not hard to see that their best strategy is to post a price equal to 1 (and for the adversary to choose distribution B), for a worst-case expected revenue of $\frac{1}{2}$. If the seller posts anything above 1, then the adversary will always respond with distribution A, resulting in zero revenue. Consider now the following randomization over prices: The seller posts a price of 1 with probability 2/3, and a price of 2 with probability 1/3. If the adversary chooses Distribution A, then the expected revenue will be $1 \cdot \frac{2}{3} = \frac{2}{3}$. Similarly if Distribution B is chosen, then the expected revenue becomes $1 \cdot \frac{2}{3} \cdot \frac{1}{2} + 2 \cdot \frac{1}{3} \cdot \frac{1}{2} = \frac{2}{3}$.

Regardless of the adversarial response, a randomization over two prices strictly outperforms the best deterministic pricing. In subsequent sections we formalize this intuition, by showing a significant separation between the power of deterministic and randomized mechanisms. A separation between determinism and randomization in single-dimensional settings has been demonstrated by Fu et al. [33] under a sample access model, and by Bergemann and Schlag [10] when the seller only knows the support of the buyer's valuation distribution. Beyond the context of auction design, the observation than ambiguity averse individuals can randomize over their choices to hedge uncertainty can be traced back decades ago in the work of Raiffa [59].

2.3 Auxiliary Functions and Distributions

To state our bounds, it will be convenient to define the following auxiliary functions. We will use function ρ_D in Section 3 and ρ in Sections 4 to 6.

Definition 1 (Functions ρ_D, ρ)**.** For any $r \geq 0$, let $\rho_D(r) = \rho$, resp. $\rho(r) = \rho$, be the unique positive solution of equation

$$\frac{(\rho-1)^3}{(2\rho-1)^2} = r^2, \qquad \text{resp.} \qquad \frac{1}{\rho^2}\left(2e^{\rho-1} - 1\right) = r^2 + 1.$$

[2] There are only two subtle technical issues that need to be taken into account; x need not be right-continuous, and $\lim_{v\to\infty} x(v)$ need not equal 1; we can assume these without loss of generality. Otherwise, one could take the right-continuous closure of x, and either assign the remainder probability to high prices, or apply a suitable scaling, which would only increase expected revenue.

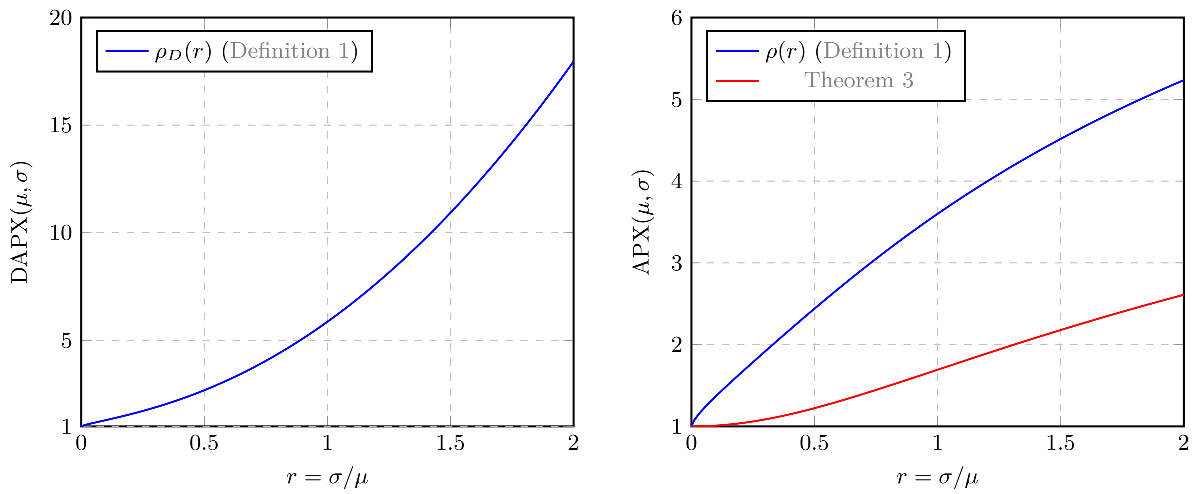


Figure 1: The robust approximation ratio for deterministic (left) and randomized (right, blue) selling mechanisms for a single (μ,σ)-distributed item, for small values of the coefficient of variation $r=\sigma/\mu$. The former is tight and given in Theorem 1. The latter is the upper bound given by Theorem 2; it is asymptotically matching the lower bound (red) of Theorem 3.

Plots of these functions, for small values of r, can be seen in Fig. 1. Their asymptotic behaviour is given in the following lemma, whose proof is deferred to Appendix A (Lemmas 6 and 7).

Lemma 1. *For the functions ρ_D, ρ defined in Definition 1, we have the bounds and asymptotics,*

$$1+4r^2 \le \rho_D(r) \le 2+4r^2 \qquad \text{for all } r \ge 0; \qquad\qquad \rho(r)=1+(1+o(1))\ln(1+r^2).$$

The asymptotics for ρ hold for large enough values of r and that is why there is a gap in Fig. 1 between the upper bound of ρ and the lower bound of $1+\ln(1+r^2)$. The bounds asymptotically match; to be precise, they are within an $(1+o(1))$-constant factor.

3 Single Item: Deterministic Pricing

In this section we begin our study of robust revenue maximization by looking at the simplest case: one item and deterministic pricing rules. Note that Azar and Micali [4] already established a lower bound of $1+r^2$ for this setting, together with an upper bound which can be shown to be $1+\left(\frac{27}{4}+o(1)\right)r^2$ (they actually characterized the upper bound via the solution of a cubic equation; we provide the exact asymptotics of that solution in Appendix C). Our result (Theorem 1) is a refined analysis that captures the exact robustness ratio (and in particular the "correct" constant in the quadratic term).

Our first observation (Lemma 2) will be that the worst-case adversarial response (for a specific selling price) can be characterized in terms of a two-point mass distribution, which allows the problem to be solved exactly. These types of distributions have appeared already in the results of Azar and Micali [2] and Carrasco et al. [17], and we will start by introducing some notation to reason about them.

A two-point mass distribution F takes some value x with probability α and some value y with probability $1-\alpha$, where without loss $x<y$. When the distribution is constrained to have mean μ and variance exactly equal to σ^2, only one free parameter remains, i.e. F can be characterized by the position x of its first point mass. The other two parameters can be obtained as

$$y(x)=\mu+\frac{\sigma^2}{\mu-x} \quad \text{and} \quad \alpha(x)=\frac{\sigma^2}{\sigma^2+(\mu-x)^2},$$

by solving the first and second moment conditions $\mu = \alpha x + (1-\alpha)y$ and $\mu^2 + \sigma^2 = \alpha x^2 + (1-\alpha)y^2$. For the remainder, we let F_x, $x \in [0, \mu)$, denote this distribution. Note that the limiting case $x \to \mu$ corresponds to $\alpha(x) \to 1$ and $y(x) \to \infty$, meaning that F_x weakly converges to μ.

By first solving the innermost optimization problem in (3), i.e. by characterizing the worst-case adversarial response against a specific *deterministic* pricing, we can derive the robustness ratio for deterministic mechanisms:

Lemma 2. *For any choice of mean μ and variance σ^2, and any deterministic pricing scheme, the worst-case robust approximation ratio is achieved over a limiting two-point mass distribution. Formally, for any μ, σ, and any price p,*

1. *if $p \geq \mu$, then the worst-case response corresponds to playing F_x with $x \to \mu^-$, and*

$$\sup_{F \in \mathbb{F}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(p; F)} = \infty;$$

2. *if $0 < p < \mu$, then the worst-case response corresponds to playing F_x with $x \to p^-$, and*

$$\sup_{F \in \mathbb{F}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(p; F)} = \max\left\{1 + \frac{\sigma^2}{(\mu-p)^2}, \frac{\mu}{p} + \frac{\sigma^2}{p(\mu-p)}\right\}.$$

Proof. If $p \geq \mu$, then the worst-case robust approximation ratio can become arbitrarily large by taking $x \to \mu^-$, that is, x arbitrarily close to μ, so that $\alpha(x) \to 1$. Indeed, we have that $\mathrm{REV}(p; F_x) \leq p(1-\alpha(x)) \to 0$, whereas $\mathrm{OPT}(F_x) \geq x \to \mu$, so that the supremum of the ratio is unbounded.

Next, let us suppose that $0 < p < \mu$. First, we compute the limit of the approximation ratio for distribution F_x, as $x \to p^-$. Observe that $\mathrm{OPT}(F_x) = \max\{x, (1-\alpha(x))y(x)\}$; and since $x < p$, we sell the item with probability $1 - \alpha(x)$, to obtain $\mathrm{REV}(p; F_x) = p(1-\alpha(x))$. Therefore,

$$\begin{aligned}
\lim_{x \to p^-} \frac{\mathrm{OPT}(F_x)}{\mathrm{REV}(p, F_x)} &= \lim_{x \to p^-} \frac{\max\{x, (1-\alpha(x))y(x)\}}{p(1-\alpha(x))} \\
&= \max\left\{\frac{1}{1-\alpha(p)}, \frac{y(p)}{p}\right\} \\
&= \max\left\{1 + \frac{\sigma^2}{(\mu-p)^2}, \frac{\mu}{p} + \frac{\sigma^2}{p(\mu-p)}\right\}.
\end{aligned}$$

Thus, it only remains to show that for any random variable X drawn from a (μ, σ) distribution F, we have that

$$\frac{\mathrm{OPT}(F)}{\mathrm{REV}(p; F)} \leq \max\left\{\frac{1}{1-\alpha(p)}, \frac{y(p)}{p}\right\}.$$

We first derive a lower bound on the probability of selling the item at price p via a one-sided version of Chebyshev's inequality, also called Cantelli's inequality[3] (see, e.g., [12, p. 46]),

$$\Pr[X \geq p] = \Pr[X - \mu \geq -(\mu - p)] \geq 1 - \frac{\sigma^2}{\sigma^2 + (\mu-p)^2} = 1 - \alpha(p). \tag{5}$$

Let p^* denote the optimal take-it-or-leave-it price for distribution F, so that $\mathrm{OPT}(F) = p^* \Pr[x \geq p^*]$. Again, we consider two cases: if $p^* \leq p$, then we have

$$\frac{\mathrm{OPT}(F)}{\mathrm{REV}(p, F)} = \frac{p^* \Pr[X \geq p^*]}{p \Pr[X \geq p]} \leq \frac{1}{1-\alpha(p)} \leq \max\left\{\frac{1}{1-\alpha(p)}, \frac{y(p)}{p}\right\}$$

[3] Although the original statement of Cantelli's inequality is for a random variable with variance *equal to* σ^2, by monotonicity the same holds if σ^2 is instead an upper bound on the variance.

where in the first inequality we used (5) and the bounds $p^* \leq p$, $\Pr[X \geq p^*] \leq 1$.

Next, consider the case $p^* > p$. By looking at the conditional random variable $(X|X \geq p)$, we observe that

$$\frac{p^* \Pr[X \geq p^*]}{\Pr[X \geq p]} = p^* \Pr\left[X \geq p^* | X \geq p\right] = \mathrm{REV}(p^*; F|X \geq p) \leq \mathbb{E}\left[X|X \geq p\right]; \tag{6}$$

the inequality holds because the social welfare is always an upper bound to the revenue.

In order to bound the conditional expectation, we use a result in Mallows and Richter [50, Eq. (1.2)]. It states that if X is a real-valued random variable with mean μ and variance σ^2 and E is a non-zero probability event, then

$$\mathbb{E}[X \mid E] - \mu \leq \sigma\sqrt{\frac{1 - \Pr[E]}{\Pr[E]}}.$$

In our case, we use $E = (X \geq p)$, together with the lower bound in (5), to get

$$\mathbb{E}\left[X|X \geq p\right] \leq \mu + \sigma\sqrt{\frac{1}{\Pr[X \geq p]} - 1} \leq \mu + \sigma\sqrt{\frac{1}{1-\alpha(p)} - 1} = \mu + \frac{\sigma^2}{\mu - p} = y(p).$$

Finally, combining the above with (6) yields

$$\frac{\mathrm{OPT}(F)}{\mathrm{REV}(p,F)} = \frac{p^* \Pr[X \geq p^*]}{p \Pr[X \geq p]} \leq \frac{\mathbb{E}\left[X|X \geq p\right]}{p} \leq \frac{y(p)}{p} \leq \max\left\{\frac{1}{1-\alpha(p)}, \frac{y(p)}{p}\right\},$$

which concludes the proof. □

Theorem 1. *The* deterministic *robust approximation ratio of selling a single (μ, σ)-distributed item is* exactly *equal to*

$$\mathrm{DAPX}(\mu, \sigma) = \rho_D(r) \approx 1 + 4 \cdot r^2,$$

where $r = \sigma/\mu$ and function $\rho_D(\cdot)$ is given in Definition 1. In particular, this is achieved by offering a take-it-or-leave-it price of $p = \frac{\rho_D(r)}{2\rho_D(r)-1} \cdot \mu$.

Proof. For fixed μ and σ, Lemma 2 gives the worst-case approximation ratio for any choice of p. Thus, from the seller's perspective, it is clear that one should offer a price below the mean, and furthermore the outermost optimization problem reduces to finding

$$\rho = \inf_{0<p<\mu} \max\left\{1 + \frac{\sigma^2}{(\mu-p)^2}, \frac{\mu}{p} + \frac{\sigma^2}{p(\mu-p)}\right\}.$$

In Lemma 6 (see Appendix A), we prove that this quantity is minimized when both branches coincide; that it corresponds to the unique positive solution of the equation

$$\frac{(\rho-1)^3}{(2\rho-1)^2} = \left(\frac{\sigma}{\mu}\right)^2;$$

the desired asymptotics; and also the characterization of the selling price p in terms of ρ. □

4 Single Item: Lotteries

In this section, we continue to focus on a single-item setting, but now we study the robust approximation ratio that can be achieved by a randomized mechanism, i.e. by randomizing over posted prices. We first define a specific randomized selling mechanism, which essentially corresponds to the lottery proposed by Carrasco et al. [17, Prop. 4]:

Definition 2 (Log-Lottery)**.** Fix any $\mu > 0$ and $\sigma \geq 0$. A *log-lottery* is a randomized mechanism that sells at a price $P^{\log}_{\mu,\sigma}$, which is distributed over the nonnegative interval support $[\pi_1, \pi_2]$ according to the cdf

$$F^{\log}_{\mu,\sigma}(x) = \frac{\pi_2 \ln \frac{x}{\pi_1} - (x - \pi_1)}{\pi_2 \ln \frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)},$$

where parameters π_1, π_2 are the (unique) solutions of the system

$$\begin{cases} \pi_1 \left(1 + \ln \dfrac{\pi_2}{\pi_1}\right) = \mu & \text{(7a)} \\ \pi_1(2\pi_2 - \pi_1) = \mu^2 + \sigma^2. & \text{(7b)} \end{cases}$$

We will sometimes slightly abuse notation and use $P^{\log}_{\mu,\sigma}$ to refer both to the log-lottery mechanism and the corresponding random variable of the prices.

Carrasco et al. [17] have given the explicit solution to the robust *absolute* revenue problem,

$$\sup_{A \in \mathbb{A}_1} \inf_{F \in \mathbb{F}_{\mu,\sigma}} \mathrm{REV}(A; F). \tag{8}$$

We state below a proposition that can be directly derived from their work and which would be very useful for our setting (the detailed derivation can be found in Appendix B).

Proposition 1. *For $\mu > 0$, $\sigma \geq 0$, the value of the maximin problem (8) is given by*

$$\sup_{A \in \mathbb{A}_1} \inf_{F \in \mathbb{F}_{\mu,\sigma}} \mathrm{REV}(A; F) = \pi_1,$$

where π_1 is derived by the unique solution of the system (7a)-(7b). Moreover, this value is achieved by the log-lottery $P^{\log}_{\mu,\sigma}$ described in Definition 2.

An intuitive interpretation of the result is the following: Since the seller is playing a game against an adversary, and since the seller is randomizing over prices in the equilibrium, they should expect the same revenue regardless of which value of the randomizing interval is sampled. In particular, we can evaluate the maximin expected revenue taking the lowest possible price π_1, in which case the item is always sold, and, thus, the resulting revenue is π_1. The above characterization can be directly used to derive a logarithmic upper bound on the robust approximation ratio:

Theorem 2. *The robust approximation ratio of selling a single (μ, σ)-distributed item is at most*

$$\mathrm{APX}(\mu, \sigma) \leq \rho(r) \approx 1 + \ln(1 + r^2),$$

where $r = \sigma/\mu$ and function ρ is given in Definition 1. In particular, this is achieved by the log-lottery described in Definition 2.

Proof. By Proposition 1, if A is the log-lottery from Definition 2, then for any (μ, σ) distribution F we have that $\mathrm{REV}(A; F) \geq \pi_1$. Thus, using the trivial upper bound of $\mathrm{OPT}(F) \leq \mu$ for the optimal revenue, we can derive an upper bound of $\frac{\mu}{\pi_1}$ on the approximation ratio. For convenience, let us denote this by $\rho \equiv \mu/\pi_1$.

Manipulating (7a) we get

$$\pi_1 \left(1 + \ln \frac{\pi_2}{\pi_1}\right) = \mu \quad \Longleftrightarrow \quad \ln \frac{\pi_2}{\pi_1} = \frac{\mu}{\pi_1} - 1 \quad \Longleftrightarrow \quad \frac{\pi_2}{\pi_1} = e^{\rho - 1}$$

and so from (7b) we can derive

$$\pi_1(2\pi_2 - \pi_1) = \mu^2 + \sigma^2 \quad \Longleftrightarrow \quad \frac{\pi_1^2}{\mu^2}\left(2\frac{\pi_2}{\pi_1} - 1\right) = \frac{\sigma^2}{\mu^2} + 1 \quad \Longleftrightarrow \quad \frac{1}{\rho^2}\left(2e^{\rho-1} - 1\right) = r^2 + 1,$$

which is exactly the equation in Definition 1. The asymptotic behaviour follows from Lemma 1. ∎

By looking at the proof of the previous theorem, it is not difficult to see that our upper bound is also an upper bound with respect to welfare (which for a single (μ, σ) distribution is simply given by μ). If we were interested in comparing the revenue of our auction to the maximum welfare, then it immediately follows from Proposition 1 that the bound is exact and tight. However, our main goal in the present paper is to provide tight bounds with respect to the optimal *revenue*, and achieving this requires some extra work. The rest of our section is devoted to proving and discussing the following lower bound, which asymptotically matches that of Theorem 2. Note that even though there is a gap between the lower and upper bound for small values of r (see also Fig. 1), the two bounds are asymptotically within a $(1 + o(1))$-factor of each other.

Theorem 3. *For a single (μ, σ)-distributed item, the robust approximation ratio is at least*

$$\mathrm{APX}(\mu, \sigma) \geq 1 + \ln(1 + r^2),$$

where $r = \sigma/\mu$.

Before we go into the actual construction of our lower bound instances, we need some technical preliminaries and to recall Yao's principle (see, e.g., [11, Sec. 8.3] or [56, Sec. 2.2.2]). As we already mentioned (see Section 2.1), a randomized mechanism $A \in \mathbb{A}_1$ can be interpreted as a randomization over prices $p \sim A$. From (3), we are interested in the value of a game in which the mechanism designer plays first, randomizing over posted prices, and the adversary plays second, choosing a worst-case distribution. Intuitively, Yao's principle states that this is at least the value of another game in which the adversary plays first, randomizing over their choices, and the mechanism designer plays second, choosing a deterministic response, i.e. a single posted price.

However, to define this second game formally, we would have to first explain what it means for the adversary to randomize over probability distributions, which form an infinite-dimensional space. In order to avoid technical or measure-theoretical issues, we focus on a specific model of randomization, which in the literature gives rise to the concept of *mixture* or *contagious* distribution (see, e.g., Mood et al. [53, Ch. III.4]).

Definition 3. Let $\mathfrak{F}$ be a class of cumulative distribution functions over the nonnegative reals, and consider any measure space over a ground set T. By an *$\mathfrak{F}$-mixture with parameter space T*, we mean a pair (Θ, F), where Θ is a probability measure in T, and F is a measurable function of type $F : \mathbb{R}_{\geq 0} \times T \to \mathbb{R}$, whose sections are in $\mathfrak{F}$; i.e. for any parameter $\theta \in T$, the function

$$F_\theta : \mathbb{R}_{\geq 0} \to \mathbb{R}, \quad F_\theta(x) = F(x; \theta),$$

is a cumulative distribution in $\mathfrak{F}$.

Given an $\mathfrak{F}$-mixture (Θ, F), we denote its *posterior* distribution by $\mathbb{E}_{\theta\sim\Theta}[F_\theta]$; this is specified by the cdf

$$\mathop{\mathbb{E}}_{\theta\sim\Theta}[F_\theta](z) = \int F(z;\theta) d\Theta(\theta) = \mathop{\mathbb{E}}_{\theta\sim\Theta}[F_\theta(z)].$$

When $\mathfrak{F} = \mathbb{F}_{\mu,\sigma}$ is the class of (μ, σ) distributions, we shall let $\Delta_{\mu,\sigma}$ denote the class of (μ, σ) mixtures, that is, the class of mixtures over $\mathbb{F}_{\mu,\sigma}$ (with arbitrary, unspecified parameter space). We can interpret (Θ, F) as a convex combination of distributions, so that the cdf of $\mathbb{E}_{\theta\sim\Theta}[F_\theta]$ is the convex combination of the corresponding cdfs; alternatively, $\mathbb{E}_{\theta\sim\Theta}[F]$ can be seen as the cdf of a random variable that first samples a distribution F_θ according to $\theta \sim \Theta$, and then samples a value z according to F_θ.

Now that we have carefully described the adversarial model, we can formally state a version of Yao's principle (Lemma 3 below) that will help us prove lower bounds. Since this applies on "non-standard" continuous spaces, for completeness we need to formally derive it "from scratch"; we give such a self-contained proof in Appendix D.

Lemma 3. *For any μ, σ, we have the following lower bound on the robust approximation ratio,*

$$\inf_{A \in \mathbb{A}_1} \sup_{F \in \mathbb{F}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(A;F)} \geq \sup_{(\Theta,F) \in \Delta_{\mu,\sigma}} \inf_{p \geq 0} \frac{\mathbb{E}_{\theta\sim\Theta}[\mathrm{OPT}(F_\theta)]}{\mathbb{E}_{\theta\sim\Theta}[\mathrm{REV}(p;F_\theta)]}.$$

Note that, by using (4), we can rewrite the denominator of the previous quantity as follows:

$$\begin{aligned}
\sup_{p\geq 0} \mathop{\mathbb{E}}_{\theta\sim\Theta} [\mathrm{REV}(p;F_\theta)] &= \sup_{p\geq 0} \mathop{\mathbb{E}}_{\theta\sim\Theta} [p(1-F_\theta(p-))] \\
&= \sup_{p\geq 0} p\left(1 - \mathop{\mathbb{E}}_{\theta\sim\Theta}[F_\theta(p-)]\right) \\
&= \sup_{p\geq 0} p\left(1 - \mathop{\mathbb{E}}_{\theta\sim\Theta}[F_\theta](p-)\right) \\
&= \sup_{p\geq 0} \mathrm{REV}\left(p; \mathop{\mathbb{E}}_{\theta\sim\Theta}[F_\theta]\right) \\
&= \mathrm{OPT}\left(\mathop{\mathbb{E}}_{\theta\sim\Theta}[F_\theta]\right).
\end{aligned}$$

The second equality comes from linearity of expectation and the third one follows from the definition of a mixture distribution. Putting all these together, we arrive at the following key technical result:

Lemma 4. *For any μ, σ, the robust approximation ratio is lower bounded by*

$$\mathrm{APX}(\mu,\sigma) \geq \sup_{(\Theta,F)\in\Delta_{\mu,\sigma}} \frac{\mathbb{E}_{\theta\sim\Theta}[\mathrm{OPT}(F_\theta)]}{\mathrm{OPT}(\mathbb{E}_{\theta\sim\Theta}[F_\theta])}. \tag{9}$$

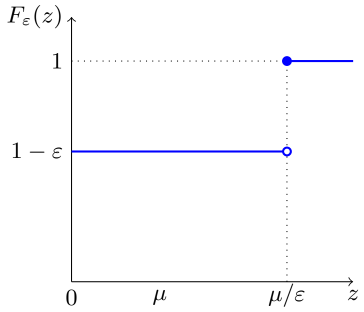
From a practical perspective, the above result has a positive consequence. It allows us to obtain lower bounds by constructing a single (μ,σ) mixture, (Θ, F), and calculating the expected optimal revenue before and after the realization of $\theta \sim \Theta$. Our goal is to make this ratio as high as possible and, ideally, match the competitive ratio of the log-lottery pricing. From this, we can gain some insight into how to construct a "good" mixture. By looking at the right-hand side of the inequality in Lemma 3, we would intuitively expect that different posted prices p yield similar revenues of $\mathbb{E}_{\theta\sim\Theta}[\mathrm{REV}(p;F_\theta)] = \mathrm{REV}(p;\mathbb{E}_{\theta\sim\Theta}[F_\theta])$. Thus, we would aim for a mixture (Θ,F) for which the posterior distribution has this property for at least some subset of its support.

From a theoretical perspective, the quantity in (9) is interesting by itself. One can check that the Myerson operator is convex, that is, the revenue achieved by a convex combination of distributions can only be smaller than the convex combinations of the corresponding revenues. Thus, by Jensen's inequality, the ratio in (9) is always at least 1. On the other hand, for a linear functional $\mathcal{L}$, we have that $\mathbb{E}_{\theta\sim\Theta}[\mathcal{L}(F_\theta)] = \mathcal{L}(\mathbb{E}_{\theta\sim\Theta}[F_\theta])$. Thus, (9) somehow attempts to quantify the extent to which OPT is nonlinear, or in other words, it can be understood as a *measure of convexity of the Myerson operator*. In any case, we can use this result to construct lower bound instances and prove the main result of this section:

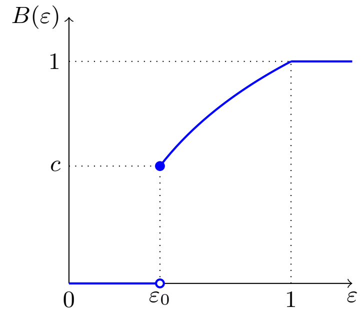
Proof of Theorem 3. We shall construct a (μ,σ) mixture over two-point mass distributions. Each two-point mass distribution F_ε is given by a unique choice of parameter $\varepsilon \in (0,1]$; F_ε returns 0 with probability $1-\varepsilon$ and μ/ε with probability ε. Note that F_ε has mean μ and variance $\mu^2(1/\varepsilon - 1)$. The upper bound of σ^2 on the variance implies that we can only take values of $\varepsilon \geq \varepsilon_0 \equiv \frac{1}{1+r^2}$, where r is the coefficient of variation (our quantity of interest).

Our next step is to describe the convex mixture of these distributions. Define a random variable with support $[\varepsilon_0, 1]$ and distributed according to B as follows:

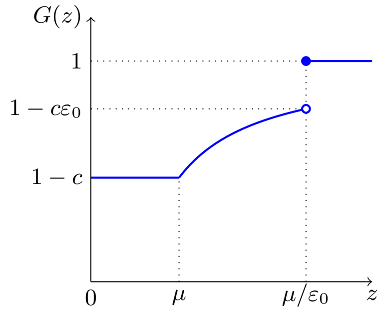
- B has a point mass at ε_0 of size c;



(a) Two-point mass distribution with one mass at 0 and another at μ/ε.



(b) Mixing in parameter space; each ε corresponds to a two-point mass distribution as in Fig. 2a. Note that this distribution has a mass at ε_0.



(c) Posterior distribution from the mixture obtained via Figs. 2a and 2b.

Figure 2: The cdfs of the various distributions used in the lower bound construction of Theorem 3.

- B is continuous over $(\varepsilon_0, 1]$, with density $\beta(\varepsilon) = c/\varepsilon$.

The value of c is given by $c = \frac{1}{1+\ln(1+r^2)}$ and is chosen as a normalizing constant; indeed,

$$1 = \mathop{\mathbb{E}}_{\varepsilon \sim B}[1] = c + c \ln \frac{1}{\varepsilon_0} = c\left(1 + \ln\left(1 + r^2\right)\right).$$

Our (μ, σ) mixture distribution thus corresponds to sampling F_ε where $\varepsilon \sim B$. Next, we describe the posterior distribution $G = \mathbb{E}_{\varepsilon \sim B}[F_\varepsilon]$. Its cumulative function can be seen in Fig. 2c.

- Mass at 0: as each F_ε has a point mass at 0, so does G. The value of this mass is given by
$$\mathop{\mathbb{E}}_{\varepsilon \sim B}[\text{mass of } F_\varepsilon \text{ at } 0] = \int_{\varepsilon_0}^{1} (1-\varepsilon)\beta(\varepsilon) d\varepsilon + (1-\varepsilon_0)c = c \ln \frac{1}{\varepsilon_0} = 1 - c;$$
- Mass at μ/ε_0: as B has a point mass at ε_0 and F_{ε_0} has a point mass at μ/ε_0, this implies that G has a point mass at μ/ε_0 of size $c\varepsilon_0$;
- cdf in $[\mu, \mu/\varepsilon_0)$: for each $z \in [\mu, \mu/\varepsilon_0)$, $F_\varepsilon(z)$ is $(1-\varepsilon)$ for $\varepsilon < \mu/z$ and 1 for $\varepsilon \geq \mu/z$; thus the cdf of G can be computed as
$$G(z) = \int_{\varepsilon_0}^{\mu/z} (1-\varepsilon)\beta(\varepsilon) d\varepsilon + \int_{\mu/z}^{1} \beta(\varepsilon) d\varepsilon + (1-\varepsilon_0)c = 1 - \frac{c\mu}{z}.$$

We can interpret $G(z)$ as a *truncated equal-revenue* distribution over the interval $[\mu, \mu/\varepsilon_0)$, *with additional point masses* at 0 and μ/ε_0. In particular, every posted price in $[\mu, \mu/\varepsilon]$ yields the same (optimal) revenue, and $\text{OPT}(G) = c\mu = \frac{\mu}{1+\ln(1+r^2)}$. On the other hand, note that for every $\varepsilon > 0$ we have $\text{OPT}(F_\varepsilon) = \mu$, so $\mathbb{E}_{\varepsilon \sim B}[\text{OPT}(F_\varepsilon)] = \mu$. Plugging these into (9) yields a lower bound of $1/c = 1 + \ln(1 + r^2)$ as desired. □

From the previous proof, some further discussion and remarks are in order. Note that our mixture uses distributions F_ε, which for $\varepsilon > \varepsilon_0$ have a variance *strictly smaller* than σ^2. Since we have defined our adversarial model to play (μ, σ) distributions, such instances are allowed. However, one may wish to ensure that the adversary only picks distributions in $\mathbb{F}^{=}_{\mu,\sigma}$ (i.e. with *exact* equality on the variance); this might be relevant, for example, if the seller had extra information about the exact value of σ; or, from a theoretical perspective, such a restriction of the adversary would only make our lower bound more "clear" and powerful. We shall now argue that indeed our assumption on having just a bound on the standard deviation, is not only a technical convenience (and, arguably, more realistic), but also is without loss of generality for our bounds. Intuitively, for any mechanism A and any (μ, σ) distribution F, one can "perturb" F into a distribution in $\mathbb{F}^{=}_{\mu,\sigma}$ having nearly the same approximation ratio. Below we formalize this intuition for single-item settings, although it is not hard to see how to generalize it to higher dimensions.

Lemma 5. *For single-item settings, the restriction of the robust approximation problem from (μ, σ) distributions to distributions in $\mathbb{F}^{=}_{\mu,\sigma}$ does not change its value. Formally, for any $\mu > 0$, $\sigma \geq 0$, and any mechanism A, we have*

$$\sup_{F \in \mathbb{F}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(A; F)} = \sup_{F \in \mathbb{F}^{=}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(A; F)};$$

and hence

$$\inf_{A \in \mathbb{A}_1} \sup_{F \in \mathbb{F}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(A; F)} = \inf_{A \in \mathbb{A}_1} \sup_{F \in \mathbb{F}^{=}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(A; F)}.$$

Proof. Let μ and σ be given, and let A be any mechanism and F_0 any (μ, σ) distribution. Suppose that the variance of F_0 is $\tilde{\sigma}^2 < \sigma^2$. For each $\delta \in (0, 1]$, let us define the perturbed distribution F_δ as the following convex combination of distributions:

- with probability $1 - \delta$, sample a value according to F_0;
- with probability δ, sample a value according to the rare event distribution that is 0 with probability $1 - \varepsilon$ and μ/ε with probability ε;
- the value of ε is chosen so that F_δ has variance *exactly equal to* σ^2; in other words, it is obtained by solving the system

$$(1 - \delta)(\mu^2 + \tilde{\sigma}^2) + \delta\mu^2/\varepsilon = \mu^2 + \sigma^2 \quad \Longrightarrow \quad \varepsilon = \frac{\delta\mu^2}{\delta\mu^2 + \sigma^2 - (1 - \delta)\tilde{\sigma}^2}.$$

Note that, for each δ, F_δ has the desired mean of μ as it is the convex combination of two distributions of mean μ. Moreover, as $\delta \to 0$, also $\varepsilon \to 0$, so that F_δ weakly converges to F_0. Finally, we have the trivial bounds

$$\mathrm{REV}(A; F_\delta) \leq (1 - \delta)\mathrm{REV}(A; F_0) + \delta\mu; \qquad \mathrm{OPT}(F_\delta) \geq (1 - \delta)\mathrm{OPT}(F_0),$$

which can be combined to yield

$$\frac{\mathrm{OPT}(F_\delta)}{\mathrm{REV}(A; F_\delta)} \geq \frac{(1 - \delta)\mathrm{OPT}(F_0)}{(1 - \delta)\mathrm{REV}(A; F_0) + \delta\mu}.$$

By letting δ go to 0, we have

$$\sup_{F \in \mathbb{F}^{=}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathrm{REV}(A; F)} \geq \lim_{\delta \to 0} \frac{(1 - \delta)\mathrm{OPT}(F_0)}{(1 - \delta)\mathrm{REV}(A; F_0) + \delta\mu} = \frac{\mathrm{OPT}(F_0)}{\mathrm{REV}(A; F_0)}.$$

Taking suprema over F_0 on the right-hand side yields the first statement of our lemma; and taking infima over A on both sides yields the last statement. □

It should also be mentioned that, in principle, we could accommodate the proof of Theorem 3 to handle distributions with exact equality with respect to σ, with minor technical modifications. More precisely, one would define $F_{\varepsilon,\delta}$ as a perturbation of F_ε as in the proof of Lemma 5. This would yield an approximation ratio that depends on δ, which would then be taken in the limit $\delta \to 0$.

Another observation is that the "bad instances" that we used for Theorem 3 were two-point mass distributions, with one of the points being 0. Note that these differ from the instances we used in the deterministic lower bounds (Lemma 2, Theorem 1), which were two-point mass distributions with exact variance of σ^2. These latter instances were actually shown in [17] to be worst-case distributions for their objective function, and they were also used in Azar and Micali [4] to prove maximin optimality in their model. Thus, it would be natural to wonder whether such instances could have been actually enough to prove a matching lower bound in the randomized setting. Below we answer this question in the negative; in other words, we prove a *constant* upper bound when the adversary is forced to pick one of these distributions.

Proposition 2. *For every choice of μ, σ, there is a randomized mechanism A that achieves (at least) a $\frac{1}{4}$-fraction of the optimal revenue on* any *distribution F that is a two-point mass with mean μ and variance σ^2. In particular, A is the mechanism that offers price $\frac{1}{2}\mu$ with probability $\frac{1}{2}$ and $\mu + \frac{\sigma^2}{\mu}$ with probability $\frac{1}{2}$.*

Proof. Let us analyse the performance of A on a two-point mass distribution F_x, say with a point mass at x of size $\alpha(x)$ and another at $y(x)$ of size $1-\alpha(x)$, with $x < \mu < y(x)$. If $\frac{1}{2}\mu \leq x$ then the mechanism chooses with probability $1/2$ a price that always sells, guaranteeing revenue of $\frac{\mu}{4}$, which is also a $1/4$-fraction of $\mathrm{OPT}(F)$. Next, suppose that $x \leq \frac{1}{2}\mu$. This implies

$$1-\alpha(x) = \frac{(\mu-x)^2}{\sigma^2+(\mu-x)^2}, \quad y(x) = \mu + \frac{\sigma^2}{\mu - x} \leq \mu + 2\frac{\sigma^2}{\mu},$$

since $y(x)$ is a nondecreasing function. Moreover, we have that

$$(1-\alpha(x))y(x) = \frac{\sigma^2(\mu-x)+(\mu-x)^2\mu}{\sigma^2+(\mu-x)^2} \geq \frac{\mu}{2}\frac{\sigma^2+2(\mu-x)^2}{\sigma^2+(\mu-x)^2} \geq \frac{\mu}{2} \geq x,$$

so that $\mathrm{OPT}(F)$ is achieved by pricing at $y(x)$. Our mechanism A chooses with probability $1/2$ a price of $\mu + \frac{\sigma^2}{\mu}$, which sells with probability $1-\alpha(x)$. Thus the approximation ratio is at least

$$\frac{\frac{1}{2}(1-\alpha(x))\left(\mu+\frac{\sigma^2}{\mu}\right)}{(1-\alpha(x))y(x)} \geq \frac{1}{2}\frac{\mu+\frac{\sigma^2}{\mu}}{\mu+2\frac{\sigma^2}{\mu}} = \frac{1}{4}\frac{\sigma^2+\mu^2}{\sigma^2+\frac{1}{2}\mu^2} > \frac{1}{4};$$

so that the mechanism achieves a $1/4$-fraction of $\mathrm{OPT}(F)$ in this case as well. ☐

The proposition above implies that the lower bound from Theorem 3 would break down, if the adversary is restricted to the family of two-point mass distributions with exact variance of σ^2.

5 Multiple Items

In this section we finally consider the more general setting of a single additive buyer with valuations for m items. As it turns out, the main tools developed in Section 4 can be leveraged very naturally to produce similar upper and lower bounds. We begin by proving upper bounds for both correlated and independent item valuations.

Theorem 4. *The robust approximation ratio of selling m (possibly correlated) $(\vec{\mu}, \vec{\sigma})$-distributed items is at most*

$$\mathrm{APX}(\vec{\mu}, \vec{\sigma}) \leq \rho(r_{\max}), \qquad \text{where} \quad r_{\max} = \max_{j=1,\dots,m} r_j, \quad r_j = \frac{\sigma_j}{\mu_j}$$

and function ρ is given in Definition 1. This is achieved by selling each item j separately *using the log-lottery $P^{\log}_{\mu_j,\sigma_j}$ from Definition 2.*

Furthermore, if the items are independently *distributed, the above bound improves to*

$$\mathrm{APX}(\vec{\mu},\vec{\sigma}) \leq \rho(\bar{r}), \qquad \textit{where} \quad \bar{r} = \frac{\bar{\sigma}}{\bar{\mu}}, \;\; \bar{\mu} = \sum_{j=1}^{m} \mu_j, \;\; \bar{\sigma} = \sqrt{\sum_{j=1}^{m} \sigma_j^2},$$

achieved by selling the items in a single full-bundle *using the log-lottery* $P^{\log}_{\bar{\mu},\bar{\sigma}}$ *from Definition 2.*

Proof. Let X_j, $j = 1, \dots, m$, be (μ_j, σ_j)-distributed random variables corresponding to the marginals of the joint m-dimensional valuation distribution F. Their sum $Y = \sum_{i=1}^m X_i$ has an expected value of $\mathbb{E}[Y] = \sum_{j=1}^m \mu_j = \bar{\mu} = \mathrm{VAL}(F)$. Furthermore, if $X_1, \dots, X_j$ are independent, its variance is $\mathrm{Var}[Y] = \sum_{j=1}^m \mathrm{Var}[X_j] \leq \sum_{j=1}^m \sigma_j^2 = \bar{\sigma}^2$. Denote the distribution of Y by F_Y. Also, recall that the optimal revenue of F cannot exceed the expected welfare, thus we have the trivial upper bound of

$$\mathrm{OPT}(F) \leq \mathrm{VAL}(F) = \sum_{j=1}^{m} \mu_j,$$

no matter if the distributions are independent or not.

For our general upper bound first, observe that selling item j using a lottery A_j, where $A_j = P^{\log}_{\mu_j,\sigma_j}$ is the log-lottery of Definition 2, guarantees (Theorem 2) a revenue of at least

$$\mathrm{REV}(A_j; F_j) \geq \frac{\mu_j}{\rho(r_j)}. \tag{10}$$

Thus, if A is the mechanism that sells independently each item j using A_j, we can get the following approximation ratio upper bound for our total revenue

$$\frac{OPT(F)}{\mathrm{REV}(A;F)} = \frac{OPT(F)}{\sum_{j=1}^m \mathrm{REV}(A_j;F_j)} \leq \frac{\sum_{j=1}^m \mu_j}{\sum_{j=1}^m \frac{\mu_j}{\rho(r_j)}} \leq \rho(r_{\max}),$$

where the last inequality holds due to the monotonicity of $\rho(\cdot)$: $\rho(r_j) \leq \rho(r_{\max})$ for all j.

For the case of independent valuations, observe that a feasible selling mechanism for our items is to bundle them all together and treat them as a single item, i.e. price their sum of valuations Y. Since Y is $(\bar{\mu},\bar{\sigma})$-distributed, offering a log-lottery $A = P^{\log}_{\bar{\mu},\bar{\sigma}}$ for Y results in an approximation ratio guarantee of

$$\mathrm{APX}(\vec{\mu},\vec{\sigma}) \leq \frac{OPT(F)}{\mathrm{REV}(A;F_Y)} \leq \frac{\mathbb{E}[Y]}{\frac{1}{\rho(\bar{r})}\,\mathbb{E}[Y]} = \rho(\bar{r}),$$

for $\bar{r} = \bar{\sigma}/\bar{\mu}$.

Finally, to verify that $\rho(\bar{r}) \leq \rho(r_{\max})$, due to the monotonicity of $\rho(\cdot)$ it is enough to see that

$$\bar{r} = \frac{\bar{\sigma}}{\bar{\mu}} = \frac{\left(\sum_{j=1}^m \sigma_j^2\right)^{1/2}}{\bar{\mu}} \leq \frac{\sum_{j=1}^m \sigma_j}{\bar{\mu}} = \frac{\sum_{j=1}^m \mu_j r_j}{\sum_{j=1}^m \mu_j}$$

is a weighted average of $r_1, r_2, \dots, r_m$, and thus at most $r_{\max}$. ◻

Corollary 1. *The robust approximation ratio of selling m independently (μ,σ)-distributed items is at most*

$$\mathrm{APX}(\vec{\mu},\vec{\sigma}) \leq \rho\left(\frac{r}{\sqrt{m}}\right),$$

where $r = \sigma/\mu$, achieved by selling the items in a single full-bundle *using the mechanism given in Theorem 2.*

Proof. In the proof of Theorem 4, if $X_1, \dots, X_m$ are independent random variables with mean μ and standard deviation at most σ, then for their sum Y we have $\bar{\mu} = m \cdot \mu$ and $\bar{\sigma} \le \sqrt{m\sigma^2} = \sqrt{m}\sigma$. $\square$

Remark. For deterministic mechanisms, it is not difficult to see that the robust approximation ratio of selling m (possibly correlated) $(\vec{\mu}, \vec{\sigma})$-distributed items is at most $\mathrm{DAPX}(\vec{\mu}, \vec{\sigma}) \le \tilde{\rho}(r_{\max})$ (where $\tilde{\rho}$ is given in Appendix C); just replace ρ by $\tilde{\rho}$ in the proof of Theorem 4. In particular, the validity of (10) is implied by (20).

We make a few observations at this point. Notice that when moving from a single item to many items, our approximation guarantees do not degrade; in particular, the robust approximation ratio is at most that of the "worst" item (i.e. the item with the highest coefficient of variation). In fact, for m independently (μ, σ)-distributed items the approximation ratio even converges to optimality (Corollary 1); this can be seen as a reinterpretation of the known result that full-bundling is asymptotically optimal for an additive bidder and many i.i.d. items (see Hart and Nisan [41, A.5.]), but in our framework of minimal statistical information.

Although the mechanisms presented in Theorem 4 are extremely simple (lotteries over separate pricing or bundle pricing), we can actually show asymptotically matching lower bounds for *any* choice of the coefficients of variation:

Theorem 5. *Fix any positive integer m and positive real numbers $r_1, \dots, r_m$, and let $r = \max_j r_j$. Then, for any $\varepsilon > 0$, there exist $\vec{\mu} = (\mu_1, \dots, \mu_m) \in \mathbb{R}^m_{>0}$, $\vec{\sigma} = (\sigma_1, \dots, \sigma_m) \in \mathbb{R}^m_{\ge 0}$ with $r_j = \sigma_j / \mu_j$, such that*

$$\mathrm{APX}(\vec{\mu}, \vec{\sigma}) \ge 1 - \varepsilon + \ln(1 + r^2).$$

Furthermore, this lower bound is achieved by independent *(μ_j, σ_j)-distributions.*

Proof. Let $m, r_1, \dots, r_m, \varepsilon$ be as in the statement of the theorem, and without loss assume $\max_j r_j = r_1$. Let $\delta > 0$ be chosen such that $\delta \ln(1 + r^2)(1 + \ln(1 + r^2))^2 < \varepsilon$. We shall choose the values for the mean and variance as

$$\mu_1 = 1, \quad \sigma_1 = r_1,$$
$$\mu_j = \frac{\delta}{m-1}, \quad \sigma_j = r_j \frac{\delta}{m-1} \quad \text{for } j \ge 2.$$

The idea is that we create a "bad" instance in which items $2, \dots, m$ are rare event distributions with very little welfare and so their contribution to the revenue will be negligible. To that end, we must first introduce some notation. For every item $j \ge 2$, denote

$$p_j = \frac{1}{1 + r_j^2}, \quad \alpha_j = (1 + r_j^2)\frac{\delta}{m-1},$$

and for every $S \subseteq \{2, \dots, n\}$, i.e. for every subset of the "low" items,

$$p_S = \prod_{j \in S} p_j \cdot \prod_{j \notin S \cup \{1\}} (1 - p_j).$$

Also, define the event

$$E_S = \left[\bigwedge_{j \in S} (v_j = \alpha_j)\right] \wedge \left[\bigwedge_{j \notin S \cup \{1\}} (v_j = 0)\right].$$

Next, let A be any m-dimensional truthful mechanism, i.e. a mechanism for selling m items to a single bidder. For each $S \subseteq \{2, \dots, n\}$, let A_S be the 1-dimensional mechanism induced by event E_S; intuitively, this mechanism allocates according to A with the values v_j set as in

E_S, but discounting the payment by the welfare from items in S. Formally, if A is defined by allocation and payment rules, $A = (\vec{x}, \pi)$, then $A_S = (x_S, \pi_S)$ can be defined as

$$x_S(v_1) = x_1(v_1, \vec{v}_{-1}), \qquad \pi_S(v_1) = \pi(v_1, \vec{v}_{-1}) - \vec{v}_{-1} \cdot \vec{x}_{-1}(0, \vec{v}_{-1}),$$

where $\vec{v}_{-1} = (v_2, \dots, v_m)$ and, for $j \geq 2$, we have $v_j = \alpha_j$ if $j \in S$; and $v_j = 0$ if $j \notin S$. One can directly check that A_S defines a truthful mechanism.

Now define $\bar{A} = \sum_S p_S A_S$ to be the convex combination of mechanisms A_S. This can be interpreted as the one-dimensional mechanism that samples a subset $S \subseteq \{2, \dots, n\}$ with probability p_S and then runs mechanism A_S. Finally, we apply Theorem 3 that ensures the existence of a "bad" single-item distribution for mechanism A_S, i.e. a distribution F_1 with mean μ_1 and standard deviation σ_1 such that

$$\mathrm{REV}(\bar{A}; F_1) \leq \frac{\mathrm{OPT}(F_1)}{1 + \ln(1 + r^2)}. \tag{11}$$

Each of the remaining distributions, F_j for $j = 2, \dots, m$, is a rare event distribution that assigns a mass of p_j on value α_j, and a mass of $1 - p_j$ on value 0. It is not hard to see that F_j has the desired mean of μ_j and variance of σ_j^2. To conclude the proof, let $F = F_1 \times \cdots \times F_m$ be the product distribution corresponding to item-independent valuations; it only remains to show that

$$\frac{\mathrm{OPT}(F)}{\mathrm{REV}(A; F)} \geq 1 - \varepsilon + \ln(1 + r_1^2).$$

We first recall a standard revenue-decomposition inequality (see the proof of Hart and Nisan [41, Lemma 8]). For any $S \subseteq \{2, \dots, n\}$, we know that

$$\mathrm{REV}(A; F_1 \times \cdots \times F_m | E_S) \leq \mathrm{REV}(A_S; F_1) + \mathrm{VAL}(F_2 \times \cdots \times F_m | E_S).$$

By the construction of our two-point mass distributions F_j, $j \geq 2$, we know that E_S form a partition of all possible valuation profiles, each event occurring with probability p_S; in this way, we can sum over the conditional expected revenues,

$$\begin{aligned}
\mathrm{REV}(A; F) &= \sum_S p_S \mathrm{REV}(A; F_1 \times \cdots \times F_m | E_S) \\
&\leq \sum_S p_S \left(\mathrm{REV}(A_S; F_1) + \mathrm{VAL}(F_2 \times \cdots \times F_m | E_S)\right) \\
&= \mathrm{REV}\left(\sum_S p_S A_S; F_1\right) + \sum_S p_S \mathrm{VAL}(F_2 \times \cdots \times F_m | E_S) \\
&\leq \frac{\mathrm{OPT}(F_1)}{1 + \ln(1 + r^2)} + \mathrm{VAL}(F_2 \times \cdots \times F_m).
\end{aligned} \tag{12}$$

Next, we consider two cases. If $\mathrm{REV}(A; F) \leq \frac{1}{(1+\ln(1+r^2))^2}$, then recall that by the mechanism presented in Theorem 2 one can extract revenue of at least $\frac{1}{1+\ln(1+r^2)}$ from F_1, hence

$$\frac{\mathrm{OPT}(F)}{\mathrm{REV}(A; F)} \geq \frac{1/(1 + \ln(1 + r^2))}{1/(1 + \ln(1 + r^2))^2} = 1 + \ln(1 + r^2).$$

Hence we can assume that $\mathrm{REV}(A; F) \geq \frac{1}{(1+\ln(1+r^2))^2}$. Note that by selling the items separately, and in particular using a price of α_j for items $j = 2, \dots, m$ we can lower bound the optimal revenue by

$$\mathrm{OPT}(F_1, F_2, \dots, F_m) \geq \mathrm{OPT}(F_1) + \sum_{j=2}^{m} \mathrm{OPT}(F_j) = \mathrm{OPT}(F_1) + \mathrm{VAL}(F_2 \times \cdots \times F_m). \tag{13}$$

Using this bound, together with the derivation in (12) and the fact that $\mathrm{VAL}(F_2 \times \cdots \times F_m) = \delta$, yields

$$\begin{aligned}
\frac{\mathrm{OPT}(F)}{\mathrm{REV}(A;F)} &\geq \frac{\mathrm{OPT}(F_1) + \delta}{\mathrm{REV}(A;F)} \\
&\geq \frac{(1+\ln(1+r^2))(\mathrm{REV}(A;F) - \delta) + \delta}{\mathrm{REV}(A;F)} \\
&= 1 + \ln(1+r^2) - \delta \frac{\ln(1+r^2)}{\mathrm{REV}(A;F)} \\
&\geq 1 + \ln(1+r^2) - \delta \ln(1+r^2)(1+\ln(1+r^2))^2 \\
&\geq 1 + \ln(1+r^2) - \varepsilon,
\end{aligned}$$

as we wanted to prove. □

One observation at this point is that our result for multiple items is in line with the main result of Carroll [18], but for the robust approximation ratio objective and in our framework of minimal statistical information. Carroll also considers a multi-dimensional setting with m items and a single additive buyer. In contrast to ours, the seller has full knowledge of the marginal distributions (but again does not know the joint distribution) and wants to optimize the maximin expected revenue. A crucial common point with our model is that the seller knows nothing about the correlation between the items. Similar to our main result, he proves that selling the items separately is maximin optimal. In other words, with no information regarding correlations, the seller chooses to *never* bundle items. A possible interpretation of this result is the following: We know that for some correlation structures, bundling works fine, while for others, it can be very bad. Thus, the seller, who wants to be robust against an unknown, possibly correlated joint distribution, might hesitate to sell as a single unit items with no information about their correlation. At the same time, the seller can calculate the expected revenue from selling each item separately in Carroll's model. Combining these two facts intuitively makes selling separately a natural candidate for maximin optimality of the expected revenue. Our result supports this interpretation for the ratio objective and partial distributional knowledge of the marginals. Even when facing uncertainty for the revenue from a single item, the seller still chooses not to bundle items when the correlation structure is entirely unknown.

6 Further Results

6.1 Parametric Auctions with Lazy Reserves

In this section, we present (Corollary 2) an additional immediate consequence of our results to the setting of Azar et al. [3]. Since this is not the main focus of our work, we refer to the above papers, as well as Hartline [42, Ch. 4] for formal definitions. The key components are that we consider a single-dimensional, matroid-constrained environment with n bidders, meaning that the set of feasible allocations forms a matroid over $\{1, \dots, n\}$. A class of mechanisms of particular interest are called Lazy-VCG with reserve prices $(P_1, \dots, P_n)$, where $P_1, \dots, P_n$ are nonnegative random variables. This auction works by first selecting a welfare-maximizing set W of candidate winners (i.e. running a VCG auction) and then offering to an agent $i \in W$ a take-it-or-leave-it price sampled according to P_i. An important result in this setting is the following black-box reduction from many bidders to one bidder with good performance guarantees (see also Chawla et al. [21, Thm. A.3]):

Theorem 6 (Azar et al. [3])**.** *Assume a single-dimensional, matroid-constrained environment with n bidders having valuations drawn independently from regular distributions $F_1, F_2, \dots, F_n$.*

If $P_1, \dots, P_n$ are nonnegative random variables such that for all players i

$$\mathbb{E}_{p\sim P_i}[\mathrm{REV}(p; F_i)] \geq c_1 \cdot \mathrm{OPT}(F_i) \qquad \textit{and} \qquad \mathbb{E}_{p\sim P_i}[\mathrm{WEL}(p; F_i)] \geq c_2 \cdot \mathrm{VAL}(F_i)$$

for constants $c_1, c_2 \in [0, 1]$, then Lazy-VCG with random reserves $(P_1, \dots, P_n)$ guarantees (in expectation) a $\frac{1}{2}c_1$-fraction of the optimal revenue and a c_2-fraction of the optimal welfare.

As an immediate consequence, since our log-lotteries from Section 4 satisfy the conditions of Theorem 6 with a suitable choice of c_1, c_2, we get the following:

Corollary 2. *Assume a single-dimensional, matroid-constrained environment with n bidders having independent regular valuations with mean μ_i and standard deviation σ_i. Then Lazy-VCG with a reserve for player i drawn from the log-lottery $P^{\log}_{\mu_i,\sigma_i}$ (see Definition 2) guarantees a $2\rho(r)$-approximation to the optimal revenue and a $\rho(r)$-approximation to the optimal welfare, where $r = \max_i \frac{\sigma_i}{\mu_i}$ and function $\rho(\cdot)$ is defined in Definition 1.*

Proof. Take $c_1 = c_2 = \frac{1}{\rho(r)} \leq \frac{1}{\rho(\sigma_i/\mu_i)}$ for all i. Note that the welfare bounds come "for free" since for any mechanism $A \in \mathbb{A}_1$ we have $\mathrm{WEL}(A; F_i) \geq \mathrm{REV}(A; F_i)$ and the upper bound in Theorem 2 was derived with respect to $\mathrm{VAL}(F_i) = \mu_i$. □

6.2 Regularity vs Dispersion

Note that regularity plays an important role in the previous Corollary 2, as it enables the black-box reduction of Azar et al. [3] to achieve meaningful upper bounds on the robust approximation ratio for a class of multi-bidder auctions. Given this observation, an obvious question would be whether additional knowledge of regularity can help us design better auctions, even for the single-item, single-bidder setting of Sections 3 and 4. In this section, we consider the notion of λ-regularity, which has already been studied in the context of revenue maximization, e.g. by Schweizer and Szech [62] and Cole and Rao [25][4], prove some basic results (Corollary 3) and discuss some interesting implications.

Consider a continuous distribution F supported over an interval D_F of nonnegative reals, and a real parameter $\lambda \in [0, 1]$. Let f denote the density function of F. We will say that F is *λ-regular* if its *λ-generalized virtual valuation* function

$$\phi_\lambda(x) \equiv \lambda \cdot x - \frac{1 - F(x)}{f(x)}$$

is monotonically nondecreasing in D_F.

It is not difficult to see that, for any $0 \leq \lambda \leq \lambda' \leq 1$, any λ-regular distribution is also λ'-regular. For the special case of $\lambda = 1$, the above definition recovers exactly the standard notion of regularity à la Myerson [57]. On the other extreme of the range, for $\lambda = 0$ we get the definition of *Monotone Hazard Rate (MHR)* distributions. Intuitively, MHR distributions have exponentially decreasing tails. Although they represent the strictest class within the λ-regularity hierarchy, they are still general enough to give rise to a wide family of natural distributions, such as the uniform, exponential, (truncated) normal and gamma.

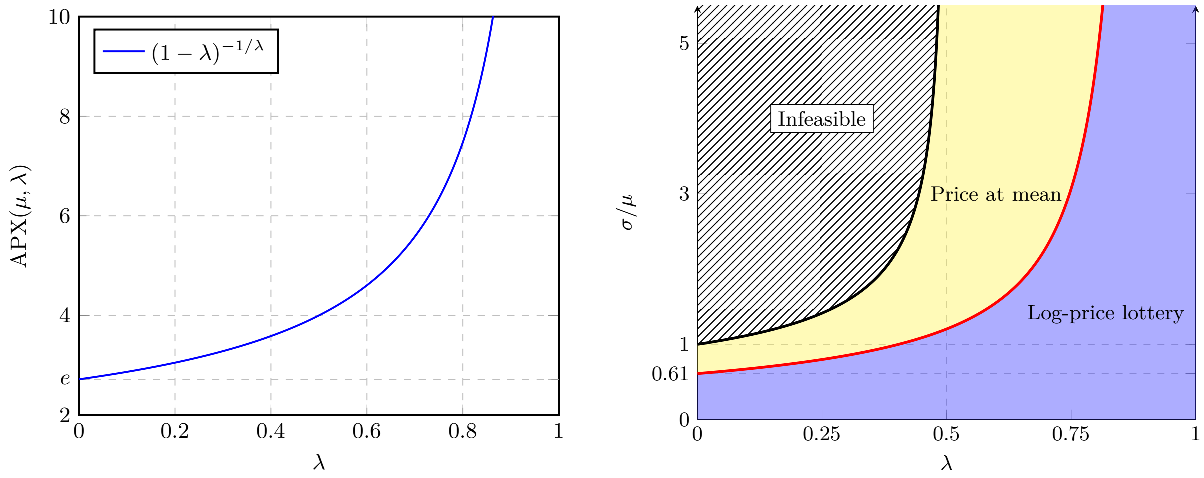
We will also need the following auxiliary results for λ-regular distributions, which follow from Propositions 2 and 4, and their corresponding proofs, of [62].

Proposition 3 (Schweizer and Szech [62])**.**

1. *Let F be λ-regular for some $\lambda \in [0, 1)$. Then F has a finite mean, say μ, and we have the inequality*

$$P(X > \mu) \geq (1 - \lambda)^{\frac{1}{\lambda}} \quad \textit{for } \lambda \neq 0, \qquad P(X > \mu) \geq \frac{1}{e} \quad \textit{for } \lambda = 0.$$

[4]To be precise, Cole and Rao [25] use the notion of α-strong regularity, originally introduced by Cole and Roughgarden [26]; this corresponds exactly to the notion of λ-regularity used in [62] and this paper, for $\alpha = 1 - \lambda$.



(a) Approximation ratio for λ-regular distributions.

(b) Knowledge of λ-regularity vs variance σ^2.

Figure 3: (a) The robust approximation ratio upper bound when pricing at the mean μ of a λ-regular distribution. (b) Description of our proposed single-item, single-bidder mechanism under knowledge of (μ, σ, λ). Note that $\lambda < 1/2$ already implies an upper bound on the coefficient of variation σ/μ (black curve). Moreover, if σ/μ is sufficiently small (in particular, smaller than the function of λ given by the red curve), then offering a lottery over prices (blue area) guarantees a better approximation ratio than simply pricing at the mean (yellow area).

2. *Let F be λ-regular for some $\lambda \in [0, 1/2)$. Then F has a finite variance, say σ^2, and we have the inequality*
$$\sigma^2 \leq \frac{\mu^2}{1-2\lambda}.$$

Now we can state our main result in this section:

Corollary 3. *Consider a single-item, single-bidder setting in which the seller has knowledge of the mean μ and an upper bound on the regularity $\lambda \in (0, 1]$ of distribution F. Then we can achieve a robust approximation ratio of $(1-\lambda)^{-1/\lambda}$ by offering the mean as a selling price.*

Proof. Using an upper bound of μ on the revenue of an optimal auction, and the lower bound on the selling probability given by Proposition 3, the result immediately follows as
$$\frac{\mathrm{OPT}(F)}{\mathrm{REV}(\mu; F)} \leq \frac{\mu}{\mu(1-\lambda)^{1/\lambda}} = (1-\lambda)^{-1/\lambda}.$$
□

Note that Corollary 3 gives an upper bound that degrades from e at $\lambda = 0$ (MHR), to ∞ at $\lambda = 1$ (regular); see Fig. 3a for a plot of this quantity. Next, we compare this ratio against the logarithmic ratio from Theorem 2. In other words, consider a model in which the bidder has information about three quantities of the distribution F: its mean μ, an upper bound of σ^2 on its variance, and an upper bound of λ on its "regularity". Combining our results so far, we can postulate a selling strategy, summarized in Fig. 3b. The first observation is that some triples (μ, σ, λ) are infeasible in the following sense: if the seller knows an upper bound on λ, and furthermore $\lambda < 1/2$, then this immediately implies an upper bound on the coefficient of variation by Proposition 3; in particular, the seller would know that
$$\sigma/\mu \leq \sqrt{1/(1-2\lambda)} \tag{14}$$
Thus, we can assume without loss that triple (μ, σ, λ) obeys this inequality.

Next, we compare the robust approximation ratios of our two candidate strategies, to determine when the log-lottery of Definition 2 outperforms the pricing-at-the-mean from Corollary 3. This amounts to solving the inequality

$$\rho\left(\frac{\sigma}{\mu}\right) \leq \frac{1}{(1-\lambda)^{1/\lambda}}.$$

Since ρ is strictly increasing, this is equivalent to

$$\frac{\sigma}{\mu} \leq \rho^{-1}\left(\frac{1}{(1-\lambda)^{1/\lambda}}\right) = \sqrt{\frac{1}{(1-\lambda)^{2/\lambda}}\left(2e^{(1-\lambda^{1/\lambda}-1}-1\right)-1},$$

where for the last equality we simply rewrote the equation in Definition 1 in terms of r. The conclusion is that the upper bound for the log-lottery is better than the upper bound for pricing-at-the-mean iff σ/μ is below a certain cutoff point (which depends on λ). Note that Fig. 3b does not show the actual approximation ratio, but rather it partitions the $(\lambda, \sigma/\mu)$-space into regions where (the approximation guarantee of) each mechanism is better.

Some additional observations about Fig. 3b are in order. First, in the limit $\lambda \to 1$, our best guarantee comes from offering the log-lottery mechanism (i.e. knowledge of 1-regularity does not improve the currently best known approximation guarantees for a single-item and a single-bidder); secondly, there is a value of σ/μ, approximately equal to 0.61, below which offering the log-lottery mechanism achieves a better guarantee than that provided by pricing-at-the-mean, *regardless of* the regularity parameter $\lambda \in [0,1]$. Intuitively, one could say that *knowing that the standard deviation of F is at most* 61% *its mean gives better revenue guarantees than knowing that F is MHR*, at least for single-item, single-bidder settings.

7 Discussion and Future Directions

In this paper, we studied the robust approximation ratio of revenue maximization under minimal statistical information of the bidders' prior distribution on the item valuations. The fundamental quantities of interest turn out to be the coefficients of variation (CV), $r_j = \sigma_j/\mu_j$, of the marginal distributions. For the single-item, single-bidder case, we completely characterized this ratio for deterministic mechanisms (quadratic in r) and gave asymptotically tight bounds for randomized mechanisms (logarithmic in r). This yields natural upper bounds for the multi-item, single-additive-bidder setting. The tight lower bound is particularly powerful as it works for any choice of the r_j. Moreover, the results hold for a possibly correlated prior distribution F over the items, with only knowledge of the mean and an upper bound on the standard deviation of each marginal. The optimal mechanism turns out to be very simple: sell the items separately using the optimal randomized mechanism for the single-item case. It is also worth mentioning that although the upper bounds for the single item generalize straightforwardly to multiple items via the welfare bounds (which are trivial upper bounds to the optimal revenue), proving that these are the "correct" bounds requires careful technical work. At the heart of our analysis lies a new version of Yao's principle, which applies to the "non-standard" continuous spaces that arise in the single-item setting and might be of independent interest. As an interesting consequence, we have observed how our results can be immediately applied to the single-dimensional, multi-bidder setting proposed by Azar et al. [3], and also made a short digression into a setting in which additional information on the regularity is assumed.

We believe that the general topic of "robust revenue with minimal statistical information" gives rise to many interesting questions and variants; below we propose directions for possible future work.

Approximation ratio vs absolute revenue As we already mentioned in the paper, besides the robust approximation *ratio* in (3), another quantity of independent interest is that given in (8):

$$\sup_{A \in \mathbb{A}_1} \inf_{F \in \mathbb{F}} \mathrm{REV}(A; F).$$

This can be seen as a "vanilla" notion of robust revenue maximization, and it was considered in Azar and Micali [4] (where they proved maximin optimality for deterministic mechanisms); it was also of central interest in the work by Carrasco et al. [17] and in other works in the economics and management science literature (e.g., [18, 47, 63]).

It is perhaps subjective to ask which, if any, of the two quantities is "better", as both have their merits. From a theoretical perspective, the *absolute revenue* in (8) is a simpler quantity (e.g. it behaves linearly with respect to convex combinations of mechanisms and distributions) and thus probably easier to extend to other settings; furthermore, it might be more appealing to an economist. On the other hand, the *approximation ratio* in (3) is "dimensionless" or "scale-free", and arguably rather natural for a computer scientist.

Consider the following thought experiment, that highlights this comparison from a more practical perspective. You are the head of a selling platform and your marketing team offers you two possible selling mechanisms:

- Mechanism A (in expectation) guarantees 10$ on each item for sale, but only 25% of the optimal revenue.
- Mechanism B (in expectation) guarantees 50% of the optimal revenue, but only 5$ on each item for sale.

One possible answer could be that "it (almost) doesn't matter": for single-item randomized mechanisms, we proved that the maximin optimal lottery of Carrasco et al. [17] yields asymptotically the best possible guarantee for the robust approximation ratio. However, it is not at all clear if, in general settings, the maximin optimal auction always achieves a guarantee "similar to" (say, a constant away of) the robust ratio-optimal auction. Providing an answer to the debate above is of course beyond the scope of the present paper. Nevertheless, we briefly presented it here as a potentially stimulating topic for future work and discussion, both from a theoretical and an empirical/behavioural point of view.

Tighter bounds Although the single-item, single-bidder case is almost completely solved, our tight analysis still has an "ε-gap" between the upper (Theorem 2) and lower (Theorem 3) bounds; it would be worth trying to close this gap, either by providing stronger lower bound instances, or by improving the upper bound analysis; similar comments apply to the multi-item, single-bidder setting, where it would be interesting to quantify the finer dependencies of the approximation ratio on various choices of $(\vec{\mu}, \vec{\sigma})$.

Multiple bidders We would also like to point out a qualitative change between the many-items and many-bidders settings, when moving to them from the basic single-bidder, single-item scenario: for a single bidder and many items, the approximation guarantee does not degrade; it is essentially bounded by the approximation guarantee of the "worst" item (see Theorem 4). For a single item and many bidders, however, even with the assumption of independent, regular distributions, we gain an extra factor of 2 (see Corollary 2), coming from the general black box reduction in Azar et al. [3]. It would be interesting to see if this factor can be dropped (or alternatively, provide stronger lower bounds). We believe that a promising way to attack this question would be to study existing or novel bounds on the coefficient of variation of the *maximum order statistic* of random variables, which may be of independent interest to statisticians. Of course, the most ambitious extension would be to consider multi-dimensional, multi-bidder settings (a generalization of both our work and that of Azar et al. [3], Azar and Micali [4]).

Finally, below we propose alternative, or more general, models of limited statistical information that might be interesting for future work:

Intervals of confidence In some situations, it may be impractical to assume *exact* knowledge of the mean μ of a distribution; instead, one could assume an estimate on the mean (via historical data), say $\mu \in [\underline{\mu}, \overline{\mu}]$ or $\mu \in (1 \pm \varepsilon)\mu_0$. It would be interesting to incorporate this type of information to the results presented in the present paper, and study how they affect the approximation guarantees.

Broader classes of value functions An interesting next case would be to study the setting of, say, a single unit-demand bidder and many items, or perhaps more generally, other valuation models such as constrained additivity or submodularity.

Regularity vs dispersion In Section 6.2, we considered a very simple mechanism (price at the mean) which achieves good performance guarantees for λ-regular distributions. Note that λ was not used for designing the mechanism, but only for its analysis. Could we perhaps use the knowledge of λ to design (possibly randomized) mechanisms with better performance? And could we combine knowledge of λ and σ to design better mechanisms in a non-trivial way?

Higher-order moments Carrasco et al. [17] already looked at a single-item, single-bidder case for the "vanilla" revenue maximization problem (8) under knowledge of the first N moments of the valuation distribution; they characterized the solution in terms of an N-dimensional optimization problem, and briefly described it for the case of $N = 3$. The most intriguing question in this line of work would be to understand the dependence of the approximation guarantee on the number of moments N and, specifically, whether it converges to optimality and at what rate. In other words, what would be *the "moment complexity" of robust revenue maximization*?

A Technical Lemmas

Lemma 6. *For any $\mu > 0$ and $\sigma \geq 0$, let $r = \sigma/\mu$ and*

$$\rho = \inf_{0<p<\mu} \max \left\{ 1 + \frac{\sigma^2}{(\mu - p)^2}, \frac{\mu}{p} + \frac{\sigma^2}{p(\mu - p)} \right\}. \tag{15}$$

Then

- *the infimum in (15) occurs at a point p^* that is the unique positive solution of equation*

$$\frac{(\mu - p)^3}{2p - \mu} = \sigma^2.$$

- *the value of the infimum, ρ, is the unique solution over $[1, \infty)$ of the equation*

$$\frac{(\rho - 1)^3}{(2\rho - 1)^2} = r^2,$$

and further satisfies

$$1 + 4r^2 \leq \rho \leq 2 + 4r^2 \qquad \text{and} \qquad p^* = \frac{\rho}{2\rho - 1} \cdot \mu.$$

Proof. We begin by analysing when the first branch of the maximum in (15) is higher than the second branch. Some algebraic manipulation yields

$$1+\frac{\sigma^2}{(\mu-p)^2}\geq\frac{\mu}{p}+\frac{\sigma^2}{p(\mu-p)}\quad\iff\quad(\mu-p)^3\leq\sigma^2(2p-\mu). \tag{16}$$

When $p\leq\mu/2$, the right expression is nonpositive and hence the second branch of the maximum is highest. Next, observe that $\frac{(\mu-p)^3}{2p-\mu}$ is decreasing over $p\in(\mu/2,\mu)$, with a positive pole at $p=\mu/2$, and vanishing at $p=\mu$. Hence, for any choice of μ,σ, there is a unique point p^* at which (16) holds with equality. It follows that for $p\geq p^*$ the maximum is achieved on the first branch and for $p\leq p^*$ the maximum is achieved on the second branch.

Next, observe that $1+\frac{\sigma^2}{(\mu-p)^2}$ is increasing on $p\in(p^*,\mu)$. To see that the second branch of the maximum in (15) is decreasing on $p\in(0,p^*)$, we take its derivative

$$\frac{d}{dp}\left(\frac{\mu}{p}+\frac{\sigma^2}{p(\mu-p)}\right)=\frac{-\mu(\mu-p)^2+\sigma^2(2p-\mu)}{p^2(\mu-p)^2}.$$

When $p\leq p^*$ we have (by definition of p^*) that $\sigma^2(2p-\mu)\leq(\mu-p)^3$ and hence the above quantity is at most $-1/p$, which is negative. We conclude that the minimum occurs when both branches intersect, i.e. at $p=p^*$; using the fact that $(\mu-p^*)^3=\sigma^2(2p^*-\mu)$, we can further express the value of the minimum as

$$\rho=1+\frac{\sigma^2}{(\mu-p^*)^2}=1+\frac{\mu-p^*}{2p^*-\mu}=\frac{p^*}{2p^*-\mu}.$$

We can now use this to express p^* in terms of ρ,

$$p^*=\frac{\rho}{2\rho-1}\cdot\mu;\qquad\mu-p^*=\frac{\mu(\rho-1)}{2\rho-1};\qquad 2p^*-\mu=\frac{\mu}{2\rho-1}.$$

Putting these together, we get

$$\sigma^2=\frac{(\mu-p^*)^3}{2p^*-\mu}=\mu^2\frac{(\rho-1)^3}{(2\rho-1)^2}\quad\iff\quad\frac{(\rho-1)^3}{(2\rho-1)^2}=\left(\frac{\sigma}{\mu}\right)^2\equiv r^2.$$

One can directly check that the expression $\frac{(\rho-1)^3}{(2\rho-1)^2}$ is increasing and goes from 0 at $\rho=1$ to ∞ at $\rho\to\infty$, so that for any nonnegative r there is a unique solution $\rho\in[1,\infty)$ to the above equation. Moreover, we can write

$$r^2=\frac{(\rho-1)^3}{(2\rho-1)^2}=\frac{1}{4}\rho-\frac{1}{4}-\frac{(\rho-\frac{3}{4})(\rho-1)}{4(\rho-\frac{1}{2})^2}\quad\iff\quad\rho=1+4r^2+\frac{(\rho-\frac{3}{4})(\rho-1)}{(\rho-\frac{1}{2})^2};$$

since the fraction appearing on the right-hand side takes values between 0 and 1 (for $\rho\in[1,\infty)$), this gives us the desired global bounds. ☐

Lemma 7. *For $r>0$, let $\rho(r)$ denote the (unique) positive solution of the equation*

$$\frac{1}{\rho^2}\left(2e^{\rho-1}-1\right)=r^2+1.$$

Then, for any $\varepsilon>0$, it holds that

$$\rho(r)\leq1+(1+\varepsilon)\ln(1+r^2)$$

for large enough values of r.

Proof. Fix an $\varepsilon > 0$. For convenience, define the functions $f, g : (0, \infty) \longrightarrow \mathbb{R}$ with

$$f(x) = \frac{1}{x^2}\left(2e^{x-1} - 1\right) \qquad \text{and} \qquad g(x) = 1 + (1+\varepsilon)\ln(1+x^2).$$

By considering their derivatives, it is straightforward to see that both f and g are increasing functions. So, to prove our lemma, it is enough to show that

$$f(g(r)) \geq r^2 + 1$$

for large enough values of r.

Indeed, taking r large enough we can guarantee that

$$g(r) = 1 + (1+\varepsilon)\ln(1+r^2) \leq (1+r^2)^{\varepsilon/2},$$

since $\ln(1+x) = o(x^{\varepsilon/2})$. Thus we have

$$f(g(r)) = \frac{2e^{g(r)-1} - 1}{[g(r)]^2} \geq \frac{2e^{(1+\varepsilon)\ln(1+r^2)} - 1}{\left[(1+r^2)^{\varepsilon/2}\right]^2} = \frac{2(1+r^2)^{1+\varepsilon} - 1}{(1+r^2)^{\varepsilon}} = 2(1+r^2) - \frac{1}{(1+r^2)^{\varepsilon}}$$

which is greater than $1 + r^2$ for large enough r, since $\frac{1}{x^\varepsilon} = o(x)$. □

B Proof of Proposition 1

In this proof we refer to multiple points in the paper from Carrasco et al. [17]. The optimal mechanism for (8) is given by the allocation rule (see their Proposition 4)

$$x(v) = \begin{cases} 0, & \text{for} \quad v \leq \pi_1, \\ \lambda_1 \ln \frac{v}{\pi_1} + 2\lambda_2(v - \pi_1), & \text{for} \ \ \pi_1 \leq v \leq \pi_2, \\ 1, & \text{for} \ \ \pi_2 \leq v, \end{cases} \tag{17}$$

and the value of the maximin problem (8) is (see end of page 274[5])

$$\sup_{A \in \mathbb{A}_1} \inf_{F \in \mathbb{F}_{\mu,\sigma}} \mathrm{REV}(A; F) = \lambda_0 + \lambda_1 \mu + \lambda_2(\mu^2 + \sigma^2), \tag{18}$$

where the values of $\lambda_0, \lambda_1, \lambda_2$ are given by (see (B.4-B.6))

$$\lambda_0 = -\frac{\pi_1(2\pi_2 - \pi_1)}{2\left(\pi_2 \ln \frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)\right)}; \lambda_1 = \frac{\pi_2}{\pi_2 \ln \frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)}; \lambda_2 = -\frac{1}{2\left(\pi_2 \ln \frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)\right)}. \tag{19}$$

Note that, as we explained at the end of Section 2.1, the allocation rule $x(v)$ from (17) can be interpreted as the cdf of a randomization over prices which forms an equivalent mechanism. Moreover, by replacing the values of $\lambda_0, \lambda_1, \lambda_2$ as in (19) we get

$$x(v) = \frac{\pi_2 \ln \frac{v}{\pi_1} - (v - \pi_1)}{\pi_2 \ln \frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)},$$

which is exactly the log-lottery of Definition 2.

Finally, by replacing the values of $\lambda_0, \lambda_1, \lambda_2$ from (19), and the values of μ and σ from (7a),(7b), into (18), the value of the maximin problem can be greatly simplified to

$$\lambda_0 + \lambda_1 \mu + \lambda_2(\mu^2 + \sigma^2)$$

[5]Carrasco et al. [17] define their solutions in terms of the moments $k_1 \equiv \mu$ and $k_2 \equiv \mu^2 + \sigma^2$.

$$= -\frac{\pi_1(2\pi_2 - \pi_1)}{2\left(\pi_2 \ln\frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)\right)} + \frac{\pi_2\pi_1\left(1 + \ln\frac{\pi_2}{\pi_1}\right)}{\pi_2 \ln\frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)} - \frac{\pi_1(2\pi_2 - \pi_1)}{2\left(\pi_2 \ln\frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)\right)}$$
$$= \frac{\pi_1\left(\pi_2 + \pi_2 \ln\frac{\pi_2}{\pi_1} - 2\pi_2 + \pi_1\right)}{\pi_2 \ln\frac{\pi_2}{\pi_1} - (\pi_2 - \pi_1)} = \pi_1,$$

as we wanted to prove.

C Asymptotics of the Mechanism by Azar and Micali [2]

In this section we look at the upper bound proposed in Azar and Micali [2, Thm. 1]. They propose a deterministic mechanism with selling price $p = \mu - k(r)\sigma$, where $k(r)$ is the unique positive solution of the cubic equation $\frac{1}{r} = \frac{1}{2}(3k + k^3)$. They derive an approximation guarantee which in our setting can be expressed as

$$\mathrm{APX}(\mu, \sigma) \leq \frac{\mu}{\mathrm{REV}(p; F)} \leq \frac{1}{1 - \frac{3}{2}rk(r)} \equiv \tilde{\rho}(r). \tag{20}$$

We have the following global bounds and asymptotics:

Lemma 8. *For any $\mu > 0$ and $\sigma \geq 0$, let $r = \sigma/\mu$ and let k denote the unique real solution of $\frac{1}{r} = \frac{1}{2}(3k + k^3)$. Furthermore, let $\tilde{\rho} = \frac{1}{1 - \frac{3}{2}rk}$ and $p = \mu - k\sigma$. Then $\tilde{\rho}$ is the unique solution over $[1, \infty)$ of the equation*

$$\frac{27}{4}r^2 = \frac{(\tilde{\rho} - 1)^3}{\tilde{\rho}^2},$$

and further satisfies

$$1 + \frac{27}{4}r^2 \leq \tilde{\rho} \leq 3 + \frac{27}{4}r^2 \qquad \textit{and} \qquad p = \frac{\tilde{\rho} + 2}{3\tilde{\rho}} \cdot \mu.$$

Proof. We begin by rewriting k in terms of $\tilde{\rho}$,

$$\tilde{\rho} = \frac{1}{1 - \frac{3}{2}rk} \quad \iff \quad k = \frac{2}{3r}\frac{\tilde{\rho} - 1}{\tilde{\rho}};$$

plugging this in the cubic equation for k, and doing some manipulation, gives

$$\frac{1}{r} = \frac{1}{2}\left(\frac{2}{r}\frac{\tilde{\rho} - 1}{\tilde{\rho}} + \frac{8}{27r^3}\frac{(\tilde{\rho} - 1)^3}{\tilde{\rho}^3}\right) \quad \iff \quad \frac{27}{4}r^2 = \frac{(\tilde{\rho} - 1)^3}{\tilde{\rho}^2}.$$

One can directly check that the expression $\frac{(\tilde{\rho}-1)^3}{\tilde{\rho}^2}$ is increasing and goes from 0 at $\tilde{\rho} = 1$ to ∞ at $\tilde{\rho} \to \infty$, so that for any nonnegative r there is a unique solution $\tilde{\rho} \in [1, \infty)$ to the above equation. Moreover, we can write

$$p = \mu - k\sigma = \mu - \frac{2}{3}\frac{\sigma}{r}\frac{\tilde{\rho} - 1}{\tilde{\rho}} = \frac{\tilde{\rho} + 2}{3\tilde{\rho}} \cdot \mu$$

and

$$\frac{27}{4}r^2 = \frac{(\tilde{\rho} - 1)^3}{\tilde{\rho}^2} = \tilde{\rho} - 1 - \frac{(2\tilde{\rho} - 1)(\tilde{\rho} - 1)}{\tilde{\rho}^2} \quad \iff \quad \tilde{\rho} = 1 + \frac{27}{4}r^2 + \frac{(2\tilde{\rho} - 1)(\tilde{\rho} - 1)}{\tilde{\rho}^2}.$$

Since the fraction appearing on the right-hand side takes values between 0 and 2 (for $\tilde{\rho} \in [1, \infty)$), this gives us the desired global bounds. $\square$

D Yao's Principle for Arbitrary Measures

Lemma 9. *Let $(X, \Sigma_X, \mathcal{F})$ and $(Y, \Sigma_Y, \mathcal{G})$ be arbitrary probability spaces,*[6] *i.e.*

- *Σ_X and Σ_Y are σ-algebras over X and Y respectively;*
- *$\mathcal{F}$ and $\mathcal{G}$ are probability measures over (X, Σ_X) and (Y, Σ_Y) respectively.*

Let also $h : X \times Y \to \mathbb{R}_{\geq 0}$, $g : Y \to \mathbb{R}_{>0}$ be measurable functions. Then[7]

$$\sup_{y \in Y} \frac{g(y)}{\mathbb{E}_{x\sim\mathcal{F}}[h(x,y)]} \geq \inf_{x \in X} \frac{\mathbb{E}_{y\sim\mathcal{G}}[g(y)]}{\mathbb{E}_{y\sim\mathcal{G}}[h(x,y)]}.$$

Proof. Let α be an arbitrary nonnegative real number, and suppose that

$$\inf_{x \in X} \frac{\mathbb{E}_{y\sim\mathcal{G}}[g(y)]}{\mathbb{E}_{y\sim\mathcal{G}}[h(x,y)]} \geq \alpha, \tag{21}$$

that is, $\mathbb{E}_{y\sim\mathcal{G}}[g(y)] \geq \alpha \sup_{x\in X} \mathbb{E}_{y\in\mathcal{G}}[h(x,y)]$. This implies that, for every $x \in X$, we have $\mathbb{E}_{y\sim\mathcal{G}}[g(y)] \geq \alpha \mathbb{E}_{y\in\mathcal{G}}[h(x,y)]$. Hence, by sampling x according to $\mathcal{F}$, we also have

$$\mathbb{E}_{y\sim\mathcal{G}}[g(y)] \geq \alpha \mathbb{E}_{x\sim\mathcal{F}}[\mathbb{E}_{y\sim\mathcal{G}}[h(x,y)]] = \alpha \mathbb{E}_{y\sim\mathcal{G}}[\mathbb{E}_{x\sim\mathcal{F}}[h(x,y)]];$$

the equality holds due to Tonelli's theorem (see, e.g., Tao [65, Theorem 1.7.15]), since h is measurable and nonnegative, and $\mathcal{F}$, $\mathcal{G}$ are finite measures. By the previous inequality between expectations, we must conclude that it holds for some realization of $\mathcal{G}$, that is, there must exist $y \in \mathrm{supp}(\mathcal{G})$ such that $g(y) \geq \alpha \mathbb{E}_{x\sim\mathcal{F}}[h(x,y)]$. This implies that

$$\frac{g(y)}{\mathbb{E}_{x\sim\mathcal{F}}[h(x,y)]} \geq \alpha, \quad \text{and hence} \quad \sup_{y\in Y} \frac{g(y)}{\mathbb{E}_{x\sim\mathcal{F}}[h(x,y)]} \geq \alpha.$$

As α was any real number that satisfies (21), the desired inequality follows. ☐

Lemma (Lemma 3)**.** *For $\mu > 0$, $\sigma \geq 0$, let $\Delta_{\mu,\sigma}$ denote the class of (μ,σ) mixtures, that is, the class of mixtures over $\mathbb{F}_{\mu,\sigma}$. Then*

$$\inf_{A\in\mathbb{A}_1} \sup_{F\in\mathbb{F}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathbb{E}_{p\sim A}[\mathrm{REV}(p;F)]} \geq \sup_{(\Theta,F)\in\Delta_{\mu,\sigma}} \inf_{p\geq 0} \frac{\mathbb{E}_{\theta\sim\Theta}[\mathrm{OPT}(F_\theta)]}{\mathbb{E}_{\theta\sim\Theta}[\mathrm{REV}(p;F_\theta)]}.$$

Proof. Start by fixing an arbitrary truthful mechanism $A \in \mathbb{A}_1$ and an arbitrary (μ,σ) mixture (Θ, F) over parameter space T. Since A can be interpreted as a randomization over prices, $(\mathbb{R}_{\geq 0}, \mathcal{L}, A)$ is a well-posed probability space.

Next, define the functions

$$h : \mathbb{R}_{\geq 0} \times T \to \mathbb{R}, \quad g : T \to \mathbb{R};$$

$$h(p,\theta) = \mathrm{REV}(p;F_\theta); \quad g(\theta) = \mathrm{OPT}(F_\theta).$$

Clearly, h is nonnegative and g is positive since F_θ is (μ,σ)-distributed. We just need to argue that both are measurable. Note that

$$h(p,\theta) = \mathrm{REV}(p;F_\theta) = p(1 - F_\theta(p-)) = \inf_{y<p} p(1 - F(y;\theta)).$$

[6] For formal definitions of the measure-theoretic notions used in this lemma see, e.g., Tao [65].

[7] Throughout this lemma, we handle ratios of the form $\frac{g}{h}$ where $g > 0$ and $h \geq 0$. For convenience, if $h = 0$ we interpret the ratio as being equal to ∞. This means that, for any nonnegative real number α, we have the following relation, even when $h = 0$:

$$\frac{g}{h} \geq \alpha \quad \iff \quad g \geq \alpha \cdot h.$$

As F is measurable and taking extrema preserves measurability, so is h. In a similar way, g is measurable as it can be expressed as the supremum

$$g(\theta) = \mathrm{OPT}(F_\theta) = \sup_{p \geq 0} \mathrm{REV}(p; F_\theta).$$

Hence, we can directly apply Lemma 9 and conclude that

$$\sup_{F \in \mathbb{F}_{\mu,\sigma}} \frac{\mathrm{OPT}(F)}{\mathbb{E}_{p \sim A}[\mathrm{REV}(p; F)]} \geq \sup_{\theta \in T} \frac{\mathrm{OPT}(F_\theta)}{\mathbb{E}_{p \sim A}[\mathrm{REV}(p; F_\theta)]} \geq \inf_{p \geq 0} \frac{\mathbb{E}_{\theta \sim \Theta}[\mathrm{OPT}(F_\theta)]}{\mathbb{E}_{\theta \sim \Theta}[\mathrm{REV}(p; F_\theta)]}.$$

As A and (Θ, F) were arbitrary, we can take the supremum on the right-hand side over (μ, σ) mixtures, and the infimum on the left-hand side over truthful mechanisms; the result follows. ☐

References

[1] A. Allouah, A. Bahamou, and O. Besbes. Pricing with samples. *Operations Research*, 2022. doi:10.1287/opre.2021.2200. Forthcoming.

[2] P. Azar and S. Micali. Optimal parametric auctions. Technical Report MIT-CSAIL-TR-2012-015, MIT CSAIL, Oct. 2012. URL http://hdl.handle.net/1721.1/70556.

[3] 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)*, pages 596–604, 2013. doi:10.1137/1.9781611973105.43.

[4] P. D. Azar and S. Micali. Parametric digital auctions. In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science (ITCS)*, pages 231–232, 2013. doi:10.1145/2422436.2422464.

[5] M. Babaioff, N. Immorlica, B. Lucier, and S. M. Weinberg. A simple and approximately optimal mechanism for an additive buyer. In *Proceedings of the 55th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 21–30, 2014. doi:10.1109/FOCS.2014.11.

[6] C. Bandi and D. Bertsimas. Optimal design for multi-item auctions: A robust optimization approach. *Mathematics of Operations Research*, 39(4):1012–1038, 2014. doi:10.1287/moor.2014.0645.

[7] R. E. Barlow and F. Proschan. Failure distributions. In *Mathematical Theory of Reliability*, chapter 2. Society for Industrial and Applied Mathematics, 1996. doi:10.1137/1.9781611971194.

[8] X. Bei, N. Gravin, P. Lu, and Z. G. Tang. Correlation-robust analysis of single item auction. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 193–208, 2019. doi:10.1137/1.9781611975482.13.

[9] D. Bergemann and K. Schlag. Robust monopoly pricing. *Journal of Economic Theory*, 146 (6):2527 – 2543, 2011. doi:10.1016/j.jet.2011.10.018.

[10] D. Bergemann and K. H. Schlag. Pricing without priors. *Journal of the European Economic Association*, 6(2-3):560–569, 2008. doi:10.1162/jeea.2008.6.2-3.560.

[11] A. Borodin and R. El-Yaniv. *Online Computation and Competitive Analysis*. Cambridge University Press, 1998.

[12] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[13] J. Bulow and P. Klemperer. Auctions versus negotiations. *The American Economic Review*, 86(1):180–194, 1996. URL http://www.jstor.org/stable/2118262.

[14] Y. Cai and C. Daskalakis. Learning multi-item auctions with (or without) samples. In *Proceedings of the 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 516–527, Oct 2017. doi:10.1109/FOCS.2017.54.

[15] Y. Cai and M. Zhao. Simple mechanisms for subadditive buyers via duality. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 170–183, 2017. doi:10.1145/3055399.3055465.

[16] Y. Cai, N. R. Devanur, and S. M. Weinberg. A duality based unified approach to Bayesian mechanism design. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 926–939. ACM Press, 2016. doi:10.1145/2897518.2897645.

[17] V. Carrasco, V. F. Luz, N. Kos, M. Messner, P. Monteiro, and H. Moreira. Optimal selling mechanisms under moment conditions. *J. Econ. Theory*, 177:245–279, 2018. doi:10.1016/j.jet.2018.05.005.

[18] G. Carroll. Robustness and separation in multidimensional screening. *Econometrica*, 85 (2):453–488, 2017. doi:10.3982/ECTA14165.

[19] G. Carroll. Robustness in mechanism design and contracting. *Annual Review of Economics*, 11(1):139–166, 2019. doi:10.1146/annurev-economics-080218-025616.

[20] S. Chawla, J. D. Hartline, D. L. Malec, and B. Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the 42nd ACM symposium on Theory of Computing (STOC)*, pages 311–320, 2010. doi:10.1145/1806689.1806733.

[21] S. Chawla, H. Fu, and A. R. Karlin. Approximate revenue maximization in interdependent value settings. *CoRR*, abs/1408.4424, 2014. URL https://arxiv.org/abs/1408.4424.

[22] J. Chen, B. Li, Y. Li, and P. Lu. Bayesian auctions with efficient queries. *CoRR*, abs/1804.07451, 2018. URL https://arxiv.org/abs/1804.07451.

[23] X. Chen, I. Diakonikolas, A. Orfanou, D. Paparas, X. Sun, and M. Yannakakis. On the complexity of optimal lottery pricing and randomized mechanisms. In *Proceedings of 56th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1464–1479, Oct. 2015. doi:10.1109/FOCS.2015.93.

[24] X. Chen, G. Matikas, D. Paparas, and M. Yannakakis. On the complexity of simple and optimal deterministic mechanisms for an additive buyer. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2036–2049, 2018. doi:10.1137/1.9781611975031.133.

[25] R. Cole and S. Rao. Applications of α-strongly regular distributions to bayesian auctions. *ACM Trans. Econ. Comput.*, 5(4):18:1–18:29, Dec. 2017. doi:10.1145/3157083.

[26] R. Cole and T. Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC)*, pages 243–252, 2014. doi:10.1145/2591796.2591867.

[27] C. Daskalakis and M. Zampetakis. More revenue from two samples via factor revealing sdps. In *Proceedings of the 21st ACM Conference on Economics and Computation*, page 257–272, 2020. doi:10.1145/3391403.3399543.

[28] C. Daskalakis, A. Deckelbaum, and C. Tzamos. The complexity of optimal mechanism design. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1302–1318, 2013. doi:10.1137/1.9781611973402.96.

[29] C. Daskalakis, A. Deckelbaum, and C. Tzamos. Strong duality for a multiple-good monopolist. *Econometrica*, 85(3):735–767, 2017. doi:10.3982/ECTA12618.

[30] P. Dhangwatnotai, T. Roughgarden, and Q. Yan. Revenue maximization with a single sample. *Games and Economic Behavior*, 91(C):318–333, 2014.

[31] S. Dughmi, L. Han, and N. Nisan. Sampling and representation complexity of revenue maximization. In *Proceedings of the 10th International Conference on Web and Internet Economics (WINE)*, pages 277–291, 2014. doi:10.1007/978-3-319-13129-0_22.

[32] P. Dütting, T. Roughgarden, and I. Talgam-Cohen. Simple versus optimal contracts. In *Proceedings of the 20th ACM Conference on Economics and Computation (EC)*, pages 369–387, 2019. doi:10.1145/3328526.3329591.

[33] 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 (EC)*, pages 323–323, 2015. doi:10.1145/2764468.2764489.

[34] Y. Giannakopoulos and E. Koutsoupias. Duality and optimality of auctions for uniform distributions. *SIAM Journal on Computing*, 47(1):121–165, 2018. doi:10.1137/16M1072218.

[35] Y. Giannakopoulos, D. Poças, and A. Tsigonias-Dimitriadis. Robust revenue maximization under minimal statistical information. In *Proceedings of the 16th Conference on Web and Internet Economics (WINE)*, pages 177–190, 2020. doi:10.1007/978-3-030-64946-3_13.

[36] Y. Giannakopoulos, D. Poças, and K. Zhu. Optimal pricing for MHR and λ-regular distributions. *ACM Trans. Econ. Comput.*, 9(1):2:1–2:28, Dec. 2020. doi:10.1145/3434423.

[37] Y. A. Gonczarowski and S. M. Weinberg. The sample complexity of up-to-ε multi-dimensional revenue maximization. *Journal of the ACM*, 68(3):1–28, 2021. doi:10.1145/3439722.

[38] N. Gravin and P. Lu. Separation in correlation-robust monopolist problem with budget. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2069–2080, 2018. doi:10.1137/1.9781611975031.135.

[39] N. Haghpanah and J. Hartline. Reverse mechanism design. In *Proceedings of the 16th ACM Conference on Economics and Computation (EC)*, pages 757–758, 2015. doi:10.1145/2764468.2764498.

[40] S. Hart and N. Nisan. The menu-size complexity of auctions. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC)*, pages 565–566, 2013. doi:10.1145/2482540.2482544.

[41] S. Hart and N. Nisan. Approximate revenue maximization with multiple items. *J. Econ. Theory*, 172:313–347, 2017. doi:10.1016/j.jet.2017.09.001.

[42] J. D. Hartline. Mechanism design and approximation. Manuscript, 2013. URL `http://jasonhartline.com/MDnA/`.

[43] J. D. Hartline and T. Roughgarden. Simple versus optimal mechanisms. In *Proceedings of the 10th ACM Conference on Electronic Commerce (EC)*, pages 225–234, 2009. doi:10.1145/1566374.1566407.

[44] Y. Hu, Z. Huang, Y. Shen, and X. Wang. Targeting makes sample efficiency in auction design. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, page 610–629, 2021. doi:10.1145/3465456.3467631.

[45] Z. Huang, Y. Mansour, and T. Roughgarden. Making the most of your samples. *SIAM Journal on Computing*, 47(3):651–674, 2018. doi:10.1137/16M1065719.

[46] M. G. Kendall. *The Advanced Theory of Statistics*, volume I. Charles Griffin, 4th edition, 1948.

[47] Ç. Koçyiğit, G. Iyengar, D. Kuhn, and W. Wiesemann. Distributionally robust mechanism design. *Management Science*, 66(1):159–189, 2020. doi:10.1287/mnsc.2018.3219.

[48] X. Li and A. C.-C. Yao. On revenue maximization for selling multiple independently distributed items. *Proceedings of the National Academy of Sciences*, 110(28):11232–11237, 2013. doi:10.1073/pnas.1309533110.

[49] Y. Li, P. Lu, and H. Ye. Revenue maximization with imprecise distribution. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1582–1590, 2019. URL http://dl.acm.org/citation.cfm?id=3306127.3331877.

[50] C. L. Mallows and D. Richter. Inequalities of chebyshev type involving conditional expectations. *The Annals of Mathematical Statistics*, 40(6):1922–1932, 1969. URL http://www.jstor.org/stable/2239511.

[51] A. M. Manelli and D. R. Vincent. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic Theory*, 137(1):153–185, 2007. doi:10.1016/j.jet.2006.12.007.

[52] P. Milgrom. *Putting Auction Theory to Work*. Cambridge University Press, 2004. doi:10.1017/CBO9780511813825.009.

[53] A. M. Mood, F. A. Graybill, and D. C. Boes. *Introduction to the Theory of Statistics*. McGraw-Hill, 3rd edition, 1974.

[54] J. Morgenstern and T. Roughgarden. Learning simple auctions. In *Proceedings of the 29th Annual Conference on Learning Theory (COLT)*, pages 1298–1318, 2016. URL http://proceedings.mlr.press/v49/morgenstern16.html.

[55] S. Moriguti. Extremal properties of extreme value distributions. *The Annals of Mathematical Statistics*, 22(4):523–536, 1951. URL http://www.jstor.org/stable/2236921.

[56] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[57] R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. doi:10.1287/moor.6.1.58.

[58] N. Nisan. Introduction to mechanism design (for computer scientists). In N. Nisan, T. Roughgarden, É. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*, chapter 9. Cambridge University Press, 2007.

[59] H. Raiffa. Risk, ambiguity, and the savage axioms: Comment. *The Quarterly Journal of Economics*, 75(4):690–694, 1961. doi:10.2307/1884326.

[60] T. Roughgarden and I. Talgam-Cohen. Approximately optimal mechanism design. *Annual Review of Economics*, 11(1), 2019. doi:10.1146/annurev-economics-080218-025607.

[61] A. Rubinstein and S. M. Weinberg. Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. *ACM Trans. Econ. Comput.*, 6(3-4):19:1–19:25, Oct. 2018. doi:10.1145/3105448.

[62] N. Schweizer and N. Szech. Performance bounds for optimal sales mechanisms beyond the monotone hazard rate condition. *Journal of Mathematical Economics*, 82:202–213, 2019.

[63] A. Suzdaltsev. An optimal distributionally robust auction. *CoRR*, abs/2006.05192, 2020. URL https://arxiv.org/abs/2006.05192.

[64] V. Syrgkanis. A sample complexity measure with applications to learning optimal auctions. In *Proccedings of the 31st Conference on Neural Information Processing Systems (NIPS)*, pages 5352–5359, 2017. URL http://papers.nips.cc/paper/7119-a-sample-complexity-measure-with-applications-to-learn

[65] T. Tao. *An Introduction to Measure Theory*, volume 126 of *Graduate Studies in Mathematics*. American Mathematical Society, 2011.

[66] R. Wilson. Game-theoretic analyses of trading processes. In *Advances in Economic Theory: Fifth World Congress*, Econometric Society Monographs, pages 33–70. Cambridge University Press, 1987. doi:10.1017/CCOL0521340446.002.

[67] A. C.-C. Yao. An n-to-1 bidder reduction for multi-item auctions and its applications. In *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 92–109, 2015. doi:10.1137/1.9781611973730.8.