Extreme-Value Theorems for Optimal Multidimensional Pricing

Yang Cai
EECS, MIT

Constantinos Daskalakis
EECS, MIT

June 6, 2011

Abstract

We provide a Polynomial Time Approximation Scheme for the *multi-dimensional unit-demand pricing problem*, when the buyer's values are independent (but not necessarily identically distributed.) For all $\epsilon > 0$, we obtain a $(1+\epsilon)$-factor approximation to the optimal revenue in time polynomial, when the values are sampled from Monotone Hazard Rate (MHR) distributions, quasi-polynomial, when sampled from regular distributions, and polynomial in $n^{\text{poly}(\log r)}$, when sampled from general distributions supported on a set $[u_{min}, ru_{min}]$. We also provide an additive PTAS for all bounded distributions.

Our algorithms are based on novel extreme value theorems for MHR and regular distributions, and apply probabilistic techniques to understand the statistical properties of revenue distributions, as well as to reduce the size of the search space of the algorithm. As a byproduct of our techniques, we establish structural properties of optimal solutions. We show that, for all $\epsilon > 0$, $g(1/\epsilon)$ distinct prices suffice to obtain a $(1+\epsilon)$-factor approximation to the optimal revenue for MHR distributions, where $g(1/\epsilon)$ is a quasi-linear function of $1/\epsilon$ that does not depend on the number of items. Similarly, for all $\epsilon > 0$ and $n > 0$, $g(1/\epsilon \cdot \log n)$ distinct prices suffice for regular distributions, where n is the number of items and $g(\cdot)$ is a polynomial function. Finally, in the i.i.d. MHR case, we show that, as long as the number of items is a sufficiently large function of $1/\epsilon$, a single price suffices to achieve a $(1+\epsilon)$-factor approximation.

1 Introduction

Here is a natural pricing problem: A seller has n items to sell to a buyer who is interested in buying a single item. The seller wants to maximize her profit from the sale, and wants to leverage stochastic knowledge she has about the buyer to achieve this goal. In particular, we assume that the seller has access to a distribution $\mathcal{F}$ from which the values $(v_1, \ldots, v_n)$ of the buyer for the items are drawn. Given this information, she needs to compute prices $p_1, \ldots, p_n$ for the items to maximize her revenue, assuming that the buyer is *quasi-linear*—i.e. will buy the item i maximizing $v_i - p_i$, as long as this difference is positive. Hence, the seller's expected payoff from a price vector $P = (p_1, \ldots, p_n)$ is

$$\mathcal{R}_P = \sum_{i=1}^{n} p_i \cdot \Pr\left[(i = \arg\max\{v_j - p_j\}) \ \wedge \ (v_i - p_i \geq 0)\right], \tag{1}$$

where we assume that the arg max breaks ties in a consistent way, if there are multiple maximizers. A more sophisticated seller could try to improve her payoff by pricing lotteries over items, i.e. price randomized allocations (see [3],) albeit this may be less natural than item pricing.

While the problem looks simple, it exhibits a rich behavior depending on the nature of $\mathcal{F}$. For example, if $\mathcal{F}$ assigns the same value to all the items with probability 1, i.e. when the buyer always values all items equally, the problem degenerates to—what Economists call—a *single-dimensional* setting. In this setting, it is obvious that lotteries do not improve the revenue and that an optimal price vector should assign the same price to all items. This observation is a special case of a more general, celebrated result of Myerson [14] on optimal mechanism design (i.e. the multi-buyer version of the above problem, and generalizations thereof.) Myerson's result provides a closed-form solution to this generalized problem in a single sweep that covers many settings, but only works under the same limiting assumption that every buyer is single-dimensional, i.e. receives the same value from all the items (in general, the same value from all outcomes where she is provided service.)

Following Myerson's work, a large body of research in both Economics and Engineering has been devoted to extending this result to the *multi-dimensional setting*, i.e. when the buyers' values come from general distributions. And while there has been sporadic progress (see survey [13] and its references,) it appears that we are far from an optimal multi-dimensional mechanism, generalizing Myerson's result. In particular, there is no optimal solution known to even the single-buyer problem presented above. Even the ostensibly easier version of that problem, where the values of the buyer for the n items are independent and supported on a set of cardinality 2 appears difficult. [1]

Motivated by the importance of the problem to Economics, and intrigued by its simplicity and apparent hardness, we devote this paper to the multi-dimensional pricing problem. Our main contribution is to develop the *first near-optimal algorithms for this problem,* when the buyer's values are independent (but not necessarily identically distributed) random variables.

Previous work on this problem by Chawla et al. [5, 6] provides factor 2 approximation to the revenue achieved by the optimal price vector. The elegant observation enabling this result is to consider the following mental experiment: suppose that the unit-demand buyer is split into n "copies" $t_1, \ldots, t_n$. Copy t_i is only interested in item i and her value for that item is drawn from the distribution $\mathcal{F}_i$ (where $\mathcal{F}_i$ is the marginal of $\mathcal{F}$ on item i), independently from the values of the other copies. On the other hand, the seller has the same feasibility constraints as before: only one item can be sold in this auction. It is intuitively obvious and can be formally established that the seller in the latter scenario is in better shape: there is more competition in the market and this can be exploited to extract more revenue. So the revenue of the seller in the original scenario can be upper bounded by the revenue in the hypothetical scenario. Moreover, the latter is a single-parameter setting; hence,

[1] Incidentally, the problem is trickier than it originally seems, and various intuitive properties that one would expect from an optimal solution fail to hold. See Appendix A for an example.

we understand exactly how its optimal revenue behaves by Myerson's result. So we can go back to our original setting and design a mechanism whose revenue comes close to Myerson's revenue in the hypothetical scenario. Using this approach [6] obtains a 2-approximation to the optimal revenue. Moreover, if the distributions $\{\mathcal{F}_i\}_i$ are *regular* (this is a commonly studied class of distributions in Economics,) the corresponding price vector can be computed efficiently.

Nevertheless, there is an inherent loss in the approach outlined above, as the revenue obtained by the sought-after mechanism will eventually be compared to a revenue that is not the optimal achievable revenue in the real setting, but the optimal revenue in a hypothetical setting; and as far as we know this could be up to a factor of 2 larger than the real one. So it could be that this approach is inherently limited to constant factor approximations. We are interested instead in efficient pricing mechanisms that achieve a $(1-\epsilon)$-fraction of the optimal revenue, for arbitrarily small ϵ. We show

Theorem 1 (PTAS for MHR Distributions)**.** *For all $\epsilon > 0$, there is a Polynomial Time Approximation Scheme* [2] *for computing a price vector whose revenue is a $(1+\epsilon)$-factor approximation to the optimal revenue, when the values of the buyer are independent and drawn from Monotone Hazard Rate distributions. (This is a commonly studied class of distributions in Economics—see Section 2.) For all $\epsilon > 0$, the algorithm runs in time $n^{\text{poly}(1/\epsilon)}$.*

Theorem 2 (Quasi-PTAS for Regular Distributions)**.** *For all $\epsilon > 0$, there is a Quasi-Polynomial Time Approximation Scheme* [3] *for computing a price vector whose revenue is a $(1+\epsilon)$-factor approximation to the optimal revenue, when the values of the buyer are independent and drawn from regular distributions. (These contain MHR and are also commonly studied in Economics—see Section 2.) For all $\epsilon > 0$, the algorithm runs in time $n^{\text{poly}(\log n, 1/\epsilon)}$.*

Theorem 3 (General Algorithm)**.** *For all $\epsilon > 0$, there is an algorithm for computing a price vector whose revenue is a $(1+\epsilon)$-factor approximation to the optimal revenue, whose running time is $n^{\text{poly}(\frac{1}{\epsilon}, \log r)}$ when the values of the buyer are independent and distributed in an interval $[u_{min}, r u_{min}]$.* [4]

Theorem 4 (Additive PTAS–General Distributions)**.** *For all $\epsilon > 0$, there is a PTAS for computing a price vector whose revenue is within an additive ϵ of the optimal revenue, when the values of the buyer are independent and distributed in $[0,1]$.*

Structural Theorems. Our approach is different than that of [5, 6] in that we study directly the optimal revenue (as a random variable,) rather than only relating its expectation to a benchmark that may be off by a constant factor. Clearly, the optimal revenue is a function of the values (which are random) and the optimal price vector (which is unknown). Hence it may be hard to pin down its distribution exactly. Nevertheless, we manage to understand its statistical properties sufficiently to deduce the following interesting structural theorems.

Theorem 5 (**Structural 1:** A Constant Number of Distinct Prices Suffice for MHR Distributions)**.** *There exists a (quasi-linear) function $g(\cdot)$ such that, for all $\epsilon > 0$ and all $n > 0$, $g(1/\epsilon)$ distinct prices suffice for a $(1+\epsilon)$-approximation to the optimal revenue when the buyer's values for the n items are independent and MHR. These distinct prices can be computed efficiently from the value distributions.*

Theorem 6 (**Structural 2:** A Polylog Number of Distinct Prices Suffice for Regular Distributions)**.** *There exists a (polynomial) function $g(\cdot)$ such that, for all $\epsilon > 0$ and $n > 0$, $g(1/\epsilon \cdot \log n)$ distinct prices suffice for a $(1+\epsilon)$-approximation to the optimal revenue, when the buyer's values for the n items are independent and regular. These prices can be computed efficiently from the value distributions.*

[2]Recall that a Polynomial Time Approximation Scheme (PTAS) is a family of algorithms $\{\mathcal{A}_\epsilon\}_\epsilon$, indexed by a parameter $\epsilon > 0$, such that for every fixed $\epsilon > 0$, $\mathcal{A}_\epsilon$ runs in time polynomial in the size of its input.

[3]Recall that a Quasi Polynomial Time Approximation Scheme (Quasi-PTAS) is a family of algorithms $\{\mathcal{A}_\epsilon\}_\epsilon$, indexed by a parameter $\epsilon > 0$, such that for every fixed $\epsilon > 0$, $\mathcal{A}_\epsilon$ runs in time quasi-polynomial in the size of its input.

[4]We point out that a straightforward application of the discretization proposed by Nisan (see [5]) or Hartline and Koltun [11] would only give a $\left(\frac{1}{\epsilon}\log r\right)^{O(n)}$-time algorithm.

Theorem 5 shows that, when the values are MHR independent, then only the desired approximation ϵ dictates the number of distinct prices that are necessary to achieve a $(1+\epsilon)$-approximation to the optimal revenue, and the number of items n as well as the range of the distributions *are irrelevant* (!) Theorem 6 generalizes this to a mild dependence on n for regular distributions. Establishing these theorems is quite challenging, as it relies on a deep understanding of the properties of the tails of MHR and regular distributions. For this purpose, we develop novel *extreme value theorems* for these classes of distributions (Theorems 12 and 14 in Sections 3 and 4 respectively.) Our theorems bound the size of the tail of the *maximum* of n independent (but not necessarily identically distributed) random variables, which are MHR or regular respectively, and are instrumental in establishing the following truncation property: truncating all the values into a common interval of the form $[\alpha, \text{poly}(1/\epsilon)\alpha]$ in the MHR case, and $[\alpha, \text{poly}(n, 1/\epsilon)\alpha]$ in the regular case, for some α that depends on the value distributions, only loses a fraction of ϵ of the optimal revenue. This is quite remarkable, especially in the case that the value distributions are non-identical. Why should most of the contribution to the optimal revenue come from a restricted set as above, when each of the underlying value distributions may concentrate on different supports? We expect that our extreme value theorems will be useful in future work, and indeed they have already been used [8]. As a final remark, we would like to point out that extreme value theorems *have* been obtained in Statistics for large classes of distributions [9], and indeed those theorems have been applied earlier in optimal mechanism design [2]. Nevertheless, known extreme value theorems are typically asymptotic, only hold for maxima of i.i.d. random variables, and are not known to hold for all MHR or regular distributions.

Covers of Revenue Distributions. Our structural theorems enable us to significantly reduce the search space for an (approximately) optimal price vector. Nevertheless, our value distributions are not necessarily identically distributed, so the search space remains exponentially large even for the MHR case, where a constant (function of ϵ only) number of distinct prices suffice by Theorem 5. Even if there are only 2 possible prices, how can we efficiently decide what price to give to each item if the items are not i.i.d? The natural approach would be to cluster the distributions into a small number of buckets, containing distributions with similar statistical properties, and proceed to treat all items in a bucket as essentially identical. However, the problem at hand is not sufficiently smooth for us to perform such bucketing and several intuitive bucketing approaches fail. We can obtain a delicate discretization of the support of the distributions into a small set (Lemma 48), but cannot discretize the probabilities used by these distributions into coarse-enough accuracy, arriving at an impasse with discretization ideas.

Our next conceptual idea is to shift the focus of attention from the space of *input value distributions*, which is inherently exponential, to the space of all possible *revenue distributions*, which are *scalar* random variables. (As we mentioned earlier, the revenue from a given price vector can be viewed as a random variable that depends on the values.) There are still exponentially many possible revenue distributions (one for each price vector,) but we find a way to construct a sparse δ-cover of this space under the total variation distance between distributions. The cover is implicit, i.e. it has no succinct closed-form description. We argue instead that it can be produced by a dynamic program, which considers prefixes of the items and constructs sub-covers for the revenue distributions induced by these prefixes, pruning down the size of the cover before growing it again to include the next item. Once a cover of the revenue distributions is obtained in this way, we argue that there is only a δ-fraction of revenue lost by replacing the optimal revenue distribution with its proxy in the cover. The high-level structure of the argument is provided in Section 6, and the details are in Section 7. Finally, the proofs of our algorithmic results (Theorems 1, 2, 3 and 4) are given in Section J.

Extensions and Related Work A natural conjecture is that, when the distributions are not widely different, a single price should suffice for extracting a $(1-\epsilon)$-fraction of the optimal revenue;

that is, as long as there is a sufficient number of items for sale. We show such a result in the case that the buyer's values are i.i.d. according to a MHR distribution. See Appendix K.

Theorem 7 (**Structural 3 (i.i.d.):** A Single Price Suffices for MHR Distributions)**.** *There is a function $g(\cdot)$ such that, for any $\epsilon > 0$, if the number of items $n > g(1/\epsilon)$ then a single price suffices for a $(1+\epsilon)$-factor approximation to the optimal revenue, if the buyer's values are i.i.d. and MHR.*

Another interesting byproduct of our techniques is that any constant-factor approximation to the optimal pricing can be converted into a PTAS or a quasi-PTAS respectively in the case of MHR or regular value distributions. This result (whose proof is given in Appendix J) is a direct product of our extreme value theorems, which can be boot-strapped with a constant factor approximation to OPT. Having such approximation would obviate the need to use our generic algorithm, outlined in the proofs of Theorems 5 and 6.

Theorem 8 (Constant Factor to Near-Optimal Approximation)**.** *If we have a constant-factor approximation to the optimal revenue of an instance of the pricing problem where the values are either MHR or regular, we can use this to speed-up our algorithms of Theorems 1 and 2.*

Future and Related Work. In conclusion, this paper provides the first near-optimal efficient algorithms for interesting instances of the multi-dimensional mechanism design problem, for a unit-demand bidder whose values are independent (but not necessarily identically distributed.) Our results provide algorithmic, structural and probabilistic insights into the properties of the optimal deterministic mechanism for the case of MHR, regular, and more general distributions. It would be interesting to extend our results (algorithmic and/or structural) to more general distributions, to mechanisms that price lotteries over items [17, 3], to bundle-pricing [12] and to budgets [1, 16]. We can certainly obtain such extensions, albeit when sizes of lotteries, bundles, etc. are a constant. We believe that our extreme value theorems, and our probabilistic view of the problem in terms of revenue distributions will be helpful in obtaining more general results. We also leave the complexity of the exact problem as an open question, and conjecture that it is NP-hard, referring the reader to [4] for hardness results in the case of correlated distributions.

Finally, it is important to solve the multi-bidder problem, extending Myerson's celebrated mechanism to the multi-dimensional setting, and the results of [1, 6] beyond constant factor approximations. In recent work, Daskalakis and Weinberg [8] have made progress in this front obtaining efficient mechanisms for multi-bidder multi-item auctions. These results are neither subsumed, nor subsume the results in the present paper. Indeed, we are more general here in that we allow the buyer to have values for the items that are not necessarily i.i.d., an assumption needed in [8] if the number of items is large. On the other hand, we are less general in that (a) we solve the single-bidder problem and (b) are near-optimal with respect to all deterministic (i.e. item-pricing), but not necessarily randomized (lottery-pricing) mechanisms. Strikingly, the techniques of the present paper are essentially orthogonal to those of [8]. The approach of [8] uses randomness to symmetrize the solution space, coupling this symmetrization with Linear Programming formulations of the problem. Our paper takes instead a probabilistic approach, developing extreme value theorems to characterize the optimal solution, and designing covers of revenue distributions to obtain efficient algorithmic solutions. It is tempting to conjecture that our approach here, combined with that of [8] would lead to more general results. Indeed, our extreme value theorems found use in [8], but we expect that significant technical work is required to go forward.

2 Preliminaries

For a random variable X we denote by $F_X(x)$ the cumulative distribution function of X, and by $f_X(x)$ its probability density function. We also let $u_{min}^X = \sup\{x|F_X(x) = 0\}$ and $u_{max}^X = \inf\{x|F_X(x) = 1\}$.

u^X_{max} may be $+\infty$, but we assume that $u^X_{min} \geq 0$, since the distributions we consider in this paper represent value distributions of items. Moreover, we often drop the subscript or superscript of X, if X is clear from context. A natural question is how distributions are provided as input to an algorithm (explicitly or with oracle access). We discuss this technical issue in Appendix C. We also define precisely what it means for an algorithm to be "efficient" in each case. We continue with the precise definition of *Monotone Hazard Rate (MHR)* and *Regular* distributions, which are both commonly studied classes of distributions in Economics.

Definition 9 (Monotone Hazard Rate Distribution)**.** *We say that a one-dimensional differentiable distribution F has* Monotone Hazard Rate*, shortly* MHR*, if $\frac{f(x)}{1-F(x)}$ is non-decreasing in $[u_{min}, u_{max}]$.*

Definition 10 (Regular Distribution)**.** *A one-dimensional differentiable distribution F is called* regular *if $x - \frac{1-F(x)}{f(x)}$ is non-decreasing in $[u_{min}, u_{max}]$.*

It is worth noticing that all MHR distributions are also regular distributions, but there are regular distributions that are not MHR. The family of MHR distributions includes such familiar distributions as the Normal, Exponential, and Uniform distributions. The family of regular distributions contains a broader range of distributions, such as fat-tail distributions $f_X(x) \sim x^{-(1+\alpha)}$ for $\alpha \geq 1$ (which are not MHR). In Appendix D and E we establish important properties of MHR and regular distributions. These properties are instrumental in establishing our extreme value theorems (Theorems 12 and 14 in the following sections).

We conclude this section by defining two computational problems. For the value distributions that we consider, we can show that they are well-defined (i.e. have finite optimal solutions.)

PRICE: **Input:** A collection of mutually independent random variables $\{v_i\}_{i=1}^n$, and some $\epsilon > 0$.
Output: A vector of prices $(p_1, \ldots, p_n)$ such that the expected revenue $\mathcal{R}_P$ under this price vector, defined as in Eq. (1), is within a $(1+\epsilon)$-factor of the optimal revenue achieved by any price vector.

RESTRICTEDPRICE: **Input:** A collection of mutually independent random variables $\{v_i\}_{i=1}^n$, a discrete set $\mathcal{P} \subset \mathbb{R}_+$, and some $\epsilon > 0$.
Output: A vector of prices $(p_1, \ldots, p_n) \in \mathcal{P}^n$ such that the expected revenue $\mathcal{R}_P$ under this price vector is within a $(1+\epsilon)$-factor of the optimal revenue achieved by any vector in $\mathcal{P}^n$.

3 Extreme Values of MHR Distributions

We reduce the problem of finding a near-optimal price vector for MHR distributions to finding a near-optimal price vector for value distributions supported on a common, balanced interval, where the imbalance of the interval is only a function of the desired approximation $\epsilon > 0$. More precisely,

Theorem 11 (From MHR to Balanced Distributions)**.** *Let $\mathcal{V} = \{v_i\}_{i\in[n]}$ be a collection of mutually independent (but not necessarily identically distributed) MHR random variables. Then there exists some $\beta = \beta(\mathcal{V}) > 0$ such that for all $\epsilon \in (0, 1/4)$, there is a reduction from* PRICE$(\mathcal{V}, c\epsilon \log(\frac{1}{\epsilon}))$ *to* PRICE$(\tilde{\mathcal{V}}, \epsilon)$*, where $\tilde{\mathcal{V}} := \{\tilde{v}_i\}_i$ is a collection of mutually independent random variables supported on the set $[\frac{\epsilon}{2}\beta, 2\log\frac{1}{\epsilon}\beta]$, and c is some absolute constant.*

Moreover, β is efficiently computable from the distributions of the X_i's (whether we are given the distributions explicitly, or we have oracle access to them,) and for every ϵ the running time of the reduction is polynomial in the size of the input and $\frac{1}{\epsilon}$. In particular, if we have oracle access to the distributions of the v_i's, then the forward reduction produces oracles for the distributions of the $\tilde{v}_i$'s, which run in time polynomial in n, $1/\epsilon$, the input to the oracle and the desired oracle precision.

We discuss the essential elements of this reduction below. Most crucially, the reduction is enabled by the following characterization of the extreme values of a collection of independent, but not necessarily identically distributed, MHR distributions.

Theorem 12 (Extreme Values of MHR distributions)**.** *Let $X_1, \ldots, X_n$ be a collection of independent (but not necessarily identically distributed) random variables whose distributions are MHR. Then there exists some anchoring point β such that $\Pr[\max_i\{X_i\} \geq \beta/2] \geq 1 - \frac{1}{\sqrt{e}}$ and*

$$\int_{2\beta\log 1/\epsilon}^{+\infty} t \cdot f_{\max_i\{X_i\}}(t)dt \leq 36\beta\epsilon\log 1/\epsilon, \text{ for all } \epsilon \in (0, 1/4). \tag{2}$$

Moreover, β is efficiently computable from the distributions of the X_i's (whether we are given the distributions explicitly, or we have oracle access to them.)

Theorem 12 (whose proof is in Appendix F.1) shows that, for all ϵ, at least a $(1-O(\epsilon\log\frac{1}{\epsilon}))$-fraction of $\mathbb{E}[\max_i X_i]$ is contributed to by values that are no larger than $\mathbb{E}[\max_i X_i]\cdot\log\frac{1}{\epsilon}$. Our result is quite surprising, especially for the case of non-identically distributed MHR random variables. Why should most of the contribution to $\mathbb{E}[\max_i X_i]$ come from values that are close (*within a function of ϵ only*) to the expectation, when the underlying random variables X_i may concentrate on widely different supports? To obtain the theorem one needs to understand how the tails of the distributions of a collection of independent but not necessarily identically distributed MHR random variables contribute to the expectation of their maximum. Our proof technique is rather intricate, defining a tournament between the tails of the distributions. Each round of the tournament ranks the distributions according to the size of their tails, and eliminates the lightest half. The threshold β is then obtained by some side-information that the algorithm records in every round.

Given our understanding of the extreme values of MHR distributions, our reduction of Theorem 11 from MHR to Balanced distributions proceeds in the following steps:

- We start with the computation of the threshold β specified by Theorem 12. This computation can be done efficiently, as stated in the statement of the theorem. Given that $\Pr[\max_i\{X_i\} \geq \beta/2]$ is bounded away from 0, β provides a lower bound to the optimal revenue. See Section F.2.1 for the precise lower bound we obtain. Such lower bound is useful as it implies that, if our transformation loses revenue that is a small fraction of β, this corresponds to a small fraction of optimal revenue lost.
- Next, using (2) we show that, for all $\epsilon > 0$, if we restrict the prices to lie in the balanced interval $[\epsilon\cdot\beta, 2\log(\frac{1}{\epsilon})\cdot\beta]$, we only lose a $O(\epsilon\log 1/\epsilon)$ fraction of the optimal revenue; this step is detailed in Section F.2.2.
- Finally, we show that we can efficiently transform the given MHR random variables $\{v_i\}_{i\in[n]}$ into a new collection of random variables $\{\tilde{v}_i\}_{i\in[n]}$ that take values in $[\frac{\epsilon}{2}\cdot\beta, 2\log(\frac{1}{\epsilon})\cdot\beta]$ and satisfy the following: a near-optimal price vector for the setting where the buyer's values are distributed as $\{\tilde{v}_i\}_{i\in[n]}$ can be efficiently transformed into a near-optimal price vector for the original setting, i.e. where the buyer's values are distributed as $\{v_i\}_{i\in[n]}$. This step is detailed in Section F.2.3.

4 Extreme Values of Regular Distributions

Our goal is to reduce the problem of finding a near-optimal pricing for a collection of independent (but not necessarily identical) *regular* value distributions to the problem of finding a near-optimal pricing for a collection of independent distributions, which are supported on a common finite interval $[u_{min}, u_{max}]$, where $u_{max}/u_{min} \leq 16n^8/\epsilon^4$, where n is the number of distributions and ϵ is the desired approximtion. It is important to notice that our bound on the ratio u_{max}/u_{min} does not depend on the distributions at hand, just their number and the required approximation. We also emphasize that the input regular distributions may be supported on $[0, +\infty)$, so it is a priori not clear if we can truncate these distributions to any finite set (even of exponential imbalance) without losing revenue.

Theorem 13 (Reduction from Regular to $Poly(n)$-Balanced Distributions)**.** *Let $\mathcal{V} = \{v_i\}_{i\in[n]}$ be a collection of mutually independent (but not necessarily identically distributed)* regular *random variables. Then there exists some $\alpha = \alpha(\mathcal{V}) > 0$ such that, for any $\epsilon \in (0,1)$, there is a reduction from* $\textsc{Price}(\mathcal{V}, \epsilon)$ *to* $\textsc{Price}(\tilde{\mathcal{V}}, \epsilon - \Theta(\epsilon/n))$*, where $\tilde{\mathcal{V}} = \{\tilde{v}_i\}_{i\in[n]}$ is a collection of mutually independent random variables that are supported on $[\frac{\epsilon\alpha}{4n^4}, \frac{4n^4\alpha}{\epsilon^3}]$.*

Moreover, we can compute α in time polynomial in n and the size of the input (whether we have the distributions of the v_i's explicitly, or have oracle access to them.) For all ϵ, the reduction runs in time polynomial in n, $1/\epsilon$ and the size of the input. In particular, if we have oracle access to the distributions of the v_i's, then the forward reduction produces oracles for the distributions of the $\tilde{v}_i$'s, which run in time polynomial in n, $1/\epsilon$, the input to the oracle and the desired oracle precision.

Our reduction is based on the following extreme value theorem for regular distributions, proved in Appendix G.1. See Appendix G.2 for a discussion of what this theorem means.

Theorem 14 (Homogenization of the Extreme Values of Regular Distributions)**.** *Let $\{X_i\}_{i\in[n]}$ be a collection of mutually independent (but not necessarily identically distributed) regular random variables, where $n \geq 2$. Then there exists some $\alpha = \alpha(\{X_i\}_i)$ such that:*

1. *α has the following "anchoring" properties:*
   - *for all $\ell \geq 1$, $\Pr[X_i \geq \ell\alpha] \leq 2/(\ell n^3)$, for all $i \in [n]$;*
   - *$\alpha/n^3 \leq c \cdot \max_z(z \cdot \Pr[\max_i\{X_i\} \geq z])$, where c is an absolute constant.*
2. *for all $\epsilon \in (0,1)$, the tails beyond $\frac{2n^2\alpha}{\epsilon^2}$ can be "homogenized", i.e.*
   - *for any integer $m \leq n$, thresholds $t_1, \ldots, t_m \geq t \geq \frac{2n^2\alpha}{\epsilon^2}$, and index set $\{a_1, \ldots, a_m\} \subseteq [n]$:*

$$\sum_{i=1}^{m} t_i \Pr[X_{a_i} \geq t_i] \leq \left(t - \frac{2\alpha}{\epsilon}\right) \cdot \Pr\left[\max_i\{X_{a_i}\} \geq t\right] + \frac{7\epsilon}{n} \cdot \left(\frac{2\alpha}{\epsilon} \cdot \Pr\left[\max_i\{X_{a_i}\} \geq \frac{2\alpha}{\epsilon}\right]\right).$$

Finally, α is efficiently computable from the distributions of the X_i's (whether we are given the distributions explicitly, or have oracle access to them.)

Given our homogenization theorem, our reduction of Theorem 13 is obtained as follows.

- First, we compute the threshold α specified in Theorem 14. This can be done efficiently as stated in Theorem 14. Now given the second anchoring property of α, we obtain an $\Omega(\alpha/n^3)$ lower bound to the optimal revenue. Such a lower bound is useful as it implies that we can ignore prices below some $O(\epsilon\alpha/n^3)$.
- Next, using our homogenization Theorem 14, we show that if we restrict a price vector to lie in $[\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$, we only lose a $O(\frac{\epsilon}{n})$ fraction of the optimal revenue. This step is detailed in Appendix G.3.1.
- Finally, we show that we can efficiently transform the input regular random variables $\{v_i\}_{i\in[n]}$ into a new collection of random variables $\{\tilde{v}_i\}_{i\in[n]}$ that are supported on $[\frac{\epsilon\alpha}{4n^4}, \frac{4n^4\alpha}{\epsilon^3}]$ and satisfy the following: a near-optimal price vector for when the buyer's values are distributed as $\{\tilde{v}_i\}_{i\in[n]}$ can be efficiently transformed into a near-optimal price vector for when the buyer's values are distributed as $\{v_i\}_{i\in[n]}$. This step is detailed in Appendix G.3.2, while Appendix G.3.3 concludes the proof of Theorem 13.

5 From Continuous to Discrete Distributions

We argued that the expected revenue can be sensitive even to small perturbations of the prices and the probability distributions. So it is a priori not clear whether there is a coarse discretization of the input and the search space that does not cost a lot of revenue. We show that, if done delicately, there is in fact such coarse discretization. Our discretization result is summarized in Theorem 15. Notice that the obtained discretization does not eliminate the exponentiality of the search or the input space.

Theorem 15 (Price/Value Distribution Discretization)**.** *Let $\mathcal{V} = \{v_i\}_{i\in[n]}$ be a collection of mutually independent random variables supported on a finite set $[u_{min}, u_{max}] \subset \mathbb{R}_+$, and let $r = \frac{u_{max}}{u_{min}} \geq 1$. For any $\epsilon \in \left(0, \frac{1}{(4\lceil \log r\rceil)^{1/6}}\right)$, there is a reduction from* $\textsc{Price}(\mathcal{V}, \epsilon)$ *to* $\textsc{RestrictedPrice}(\hat{\mathcal{V}}, \mathcal{P}, \Theta(\epsilon^8))$*, where*

- *$\hat{\mathcal{V}} = \{\hat{v}_i\}_{i\in[n]}$ is a collection of mutually independent random variables that are supported on a common set of cardinality $O\left(\frac{\log r}{\epsilon^{16}}\right)$;*
- *$|\mathcal{P}| = O\left(\frac{\log r}{\epsilon^{16}}\right)$.*

Moreover, assuming that the set $[u_{min}, u_{max}]$ is specified in the input, [5] *we can compute the (common) support of the distributions of the variables $\{\hat{v}_i\}_i$ as well as the set of prices $\mathcal{P}$ in time polynomial in $\log u_{min}$, $\log u_{max}$ and $1/\epsilon$. We can also compute the distributions of the variables $\{\hat{v}_i\}_{i\in[n]}$ in time polynomial in the size of the input and $1/\epsilon$, if we have the distributions of the variables $\{v_i\}_{i\in[n]}$ explicitly. If we have oracle access to the distributions of the variables $\{v_i\}_{i\in[n]}$, we can construct an oracle for the distributions of the variables $\{\hat{v}_i\}_{i\in[n]}$, running in time polynomial in $\log u_{min}$, $\log u_{\max}$, $1/\epsilon$, the input to the oracle and the desired precision.*

That prices can be discretized follows immediately from a discretization lemma attributed to Nisan [5] (see also a related discretization in [11],) and our result is summarized in Lemma 44 of Appendix H.1. The discretization of the value distributions is inspired by Nisan's lemma, but requires an intricate twist in order to reduce the size of the support to be linear in $\log r$ rather than linear in $r^2 \log r$ which is what a straightforward modification of the lemma gives. (Indeed, quite some effort is needed to get the former bound.) The achieved discretization in the value distributions is summarized in Lemma 48 of Appendix H.2.

6 Probabilistic Covers of Revenue Distributions

Let $\mathcal{V} := \{v_i\}_i$ be an instance of $\textsc{Price}$, where the v_i's are mutually independent random variables distributed on a finite set $[u_{min}, u_{max}]$ according to distributions $\{F_i\}_i$, and let $\mathcal{R}_{OPT}$ be the optimal expected revenue for $\mathcal{V}$. Our goal is to compute a price vector with expected revenue $(1-\epsilon)\mathcal{R}_{OPT}$. Theorem 15 of Section 5 provides an efficient reduction of this problem to the $(1-\delta)$ approximation of a discretized problem, where both the values and the prices come from discrete sets whose cardinality is $O(\log r/\delta^2)$, where $r = \frac{u_{max}}{u_{min}}$ and $\delta = O(\epsilon^8)$. For convenience, we denote by $\{\hat{F}_i\}_i$ the resulting discretized distributions, by $\{\hat{v}_i\}_i$ a collection of mutually independent random variables distributed according to the $\hat{F}_i$'s, by $\{v^{(1)}, v^{(2)}, \ldots, v^{(k_1)}\}$ the (common) support of all the $\hat{F}_i$'s, and by $\{p^{(1)}, p^{(2)}, \ldots, p^{(k_2)}\}$ the set of available price levels, where both k_1 and k_2 are $O(\log r/\delta^2)$. It is worth noting that the set of prices satisfies $\min\{p^{(i)}\} \geq u_{min}/(1+\delta)$ and $\max\{p^{(i)}\} \leq u_{max}$, and that these prices are points of a geometric sequence of ratio $1/(1-\delta^2)$. (See Lemmas 41 and 44 in the Appendix.)

Having discretized the support sets of values and prices, a natural idea that one would like to use to go forward is to further discretize the distributions $\{\hat{F}_i\}_i$ by rounding the probabilities they assign to every point in their support to integer multiples of some fraction $\sigma = \sigma(\epsilon, r) > 0$, i.e. a fraction that does not depend on n. If such discretization were feasible, the problem would be greatly

[5] The requirement that the set $[u_{min}, u_{max}]$ is specified as part of the input is only relevant if we have oracle access to the distributions of the v_i's, as if we have them explicitly we can easily find $[u_{min}, u_{max}]$.

simplified. For example, if additionally r were an absolute constant or a function of ϵ only (as it happens for MHR distributions by virtue of Theorem 31), there would only be a constant number of possible value distributions (as both the cardinality of the support of the distributions and the number of available probability levels would be a function of ϵ only.) In such case, we could try to develop an algorithm tailored to a constant number of available value distributions. This is still not easy to do (as we don't even know how to solve the i.i.d. case of our problem), but is definitely easier to dream of. Nevertheless, the approach breaks down as preserving the revenue while doing a coarse rounding of the probabilities appears difficult, and the best discretization we can obtain is given in Lemma 49 of Appendix H.4, where the accuracy is inverse polynomial in n.

Given the apparent impasse towards eliminating the exponentiality from the input space of our problem, our solution evolves in a radically different direction. To explain our approach, let us view our problem in the graphical representation of Figure 1 of Appendix B. Circuit C takes as input a price vector $p_1, \ldots, p_n$ and outputs the distribution $F_{\hat{R}_P}$ of the revenue of the seller under this price vector. Indeed, the revenue of the seller is a random variable $\hat{R}_P$ whose value depends on the variables $\{\hat{v}_i\}_{i \in [n]}$. So in order to compute the distribution of the revenue the circuit also uses the distributions $\{\hat{F}_i\}_{i \in [n]}$, which are hard-wired into the circuit. Let us denote the expectation of $\hat{R}_P$ as $\hat{\mathcal{R}}_P$.

Given our restriction of the prices to the finite set $\{p^{(1)}, p^{(2)}, \ldots, p^{(k_2)}\}$, there are k_2^n possible inputs to the circuit, and a corresponding k_2^n number of possible revenue distributions that the circuit can produce. Our main conceptual idea is this: *instead of worrying about the set of inputs to circuit C, we focus on the revenue distribution directly, constructing a probabilistic cover (under an appropriate metric) of all the possible revenue distributions that can be output by the circuit.* The two crucial properties of our cover are the following: (a) it has cardinality $O(n^{\mathrm{poly}(\frac{1}{\epsilon}, \log r)})$, and (b) for any possible revenue distribution that the circuit may output, there exists a revenue distribution in our cover with approximately the same expectation.

Details of the Cover. At a high level, the way we construct our cover is via dynamic programming, whose steps are interleaved with coupling arguments pruning the size of the DP table before proceeding to the next step. Intuitively, our dynamic program sweeps the items from 1 through n, maintaining a cover of the revenue distributions produced by all possible pricings on a prefix of the items. More precisely, for each prefix of the items, our DP table keeps track of all possible feasible collections of $k_1 \times k_2$ probability values, where $\Pr_{i_1, i_2}$, $i_1 \in [k_1], i_2 \in [k_2]$, denotes the probability that the item with the largest value-minus-price gap (i.e. the item of the prefix that would have been sold in a sale that only sales the prefix of items) has value $v^{(i_1)}$ for the buyer and is assigned price $p^{(i_2)}$ by the seller. I.e. we memoize all possible (winning-value, winning-price) distributions that can arise from each prefix of items. The reasons we decide to memoize these distributions are the following:

- First, if we have these distributions, we can compute the expected revenue that the seller would obtain, if we restricted our sale to the prefix of items.
- Second, when our dynamic program considers assigning a particular price to the next item, then having the (winning-value, winning-price) distribution on the prefix suffices to obtain the new (winning-value,winning-price) distribution that also includes the next item. I.e., if we know these distributions, we do not need to keep track of anything else in the history to keep going. Observe that it is crucial here to maintain the joint distribution of both the winning-value and the winning-price, rather than just the distribution of the winning-price.
- In the end of the program, we can look at all feasible (winning-value,winning-price) distributions for the full set of items to find the one achieving the best revenue; we can then follow back-pointers stored in our DP table to uncover a price vector consistent with the optimal distribution.

All this is both reasonable, and fun, but thus far we have achieved nothing in terms of reducing the number of distributions $F_{\hat{R}_p}$ in our cover. Indeed, there could be exponentially many (winning-value,winning-price) distributions consistent with each prefix, so that the total number of distributions that we have to memoize in the course of the algorithm is exponentially large. To obtain a polynomially small cover we show that we can be coarse in our bookkeeping of the (winning-value, winning-price) distributions, without sacrificing much revenue. Indeed, it is exactly here where viewing our problem in the "upside-down" way illustrated in Figure 1 (i.e. targeting a cover of the output of circuit C rather than figuring out a sparse cover of the input) is important: we show that, as far as the expected revenue is concerned, we can discretize probabilities into multiples of $\frac{1}{(nr)^3}$ after each round of the DP without losing much revenue, and while keeping the size of the DP table from exploding. That the loss due to pruning the search space is not significant follows from a joint application of the coupling lemma and the optimal coupling theorem (see, e.g., [10]), after each step of the Dynamic Program.

7 The Algorithm for the Discrete Problem

In this section, we formalize our ideas from the previous section, providing our main algorithmic result. We assume that the pricing problem at hand is discrete: the value distributions are supported on a discrete set $\mathcal{S} = \{v^{(1)}, v^{(2)}, \ldots, v^{(k_1)}\}$, and the sought after price vector also lies in a discrete set $\left\{p^{(1)}, \ldots, p^{(k_2)}\right\}^n$, where both $\mathcal{S}$ and $\mathcal{P} := \left\{p^{(1)}, \ldots, p^{(k_2)}\right\}$ are given explicitly as part of the input, while our access to the value distributions may still be either explicit or via an oracle. We denote by OPT the optimal expected revenue for this problem, when the prices are restricted to set $\mathcal{P}$.

The Algorithm. As a first step, we invoke Lemma 49 of Appendix H.4, obtaining a polynomial-time reduction of our problem into a new one, where additionally the probabilities that the value distributions assign to each point in $\mathcal{S}$ is an integer multiple of $1/(rn)^3$, where $r = \max\left\{\frac{p^{(j)}}{p^{(i)}}\right\}$. The loss in revenue from this reduction is at most an additive $\frac{4k_1}{rn^2}\min\{p^{(i)}\}$. Moreover, the construction of Lemma 49 is explicit, so from now on we can assume that we know the value distributions explicitly. Let us denote by $\{\hat{F}_i\}_i$ the rounded distributions and set $m := rn$ throughout this section.

The second phase of our algorithm is the Dynamic Program outlined in Section 6. We provide some further details on this next. Our program computes a Boolean function $g(i, \widehat{\Pr})$, whose arguments lie in the following range: $i \in [n]$ and $\widehat{\Pr} = (\widehat{\Pr}_{1,1}, \widehat{\Pr}_{1,2}, \ldots, \widehat{\Pr}_{k_1,k_2})$, where each $\widehat{\Pr}_{i_1,i_2} \in [0,1]$ is an integer multiple of $\frac{1}{m^3}$. The function g is stored in a table that has one cell for every setting of i and $\widehat{\Pr}$, and the cell contains a 0 or a 1 depending on the value of g at the corresponding input. In the terminology of the previous section, argument i indexes the last item in a prefix of the items and $\widehat{\Pr}$ is a (winning-value, winning-price) distribution in multiples of $\frac{1}{m^3}$. If $\widehat{\Pr}$ can arise from some pricing of the items $1 \ldots i$ (up to discretization of probabilities into multiples of $\frac{1}{m^3}$), we intend to store $g(i, \widehat{\Pr}) = 1$; otherwise we store $g(i, \widehat{\Pr}) = 0$.

Due to lack of space we postpone the straightforward details of the Dynamic Program to Appendix I.1. Very briefly, the table is filled in a bottom-up fashion from $i = 1$ through n. At the end of the i-th iteration, we have computed all feasible "discretized" (winning-value,winning-price) distributions for the prefix $1 \ldots i$, where "discretized" means that all probabilities have been rounded into multiples of $1/m^3$. For the next iteration, we try all possible prices $p^{(j)}$ for item $i+1$ and compute how each of the feasible discretized (winning-value,winning-price) distributions for the prefix $1 \ldots i$ evolves into a discretized distribution for the prefix $1 \ldots i+1$, setting the corresponding cell of layer $g(i+1, \cdot)$ of the DP table to 1. Notice, in particular, that *we lose accuracy in every step of the Dynamic Program*, as each step involves computing how a discretized distribution for items $1 \ldots i$ evolves into a distribution for items $1 \ldots i+1$ and then rounding the latter back again into multiples of $1/m^3$. We show in the analysis of our algorithm that the error accumulating from these roundings can be controlled via coupling arguments.

After computing g's table, we look at all cells such that $g(n, \widehat{\Pr}) = 1$ and evaluate the expected revenue resulting from the distribution $\widehat{\Pr}$, i.e.

$$\mathcal{R}_{\widehat{\Pr}} = \sum_{i_1 \in [k_1], i_2 \in [k_2]} p^{(i_2)} \cdot \widehat{\Pr}_{i_1,i_2} \cdot \mathbb{1}_{v^{(i_1)} \geq p^{(i_2)}}.$$

Having located the cell whose $\mathcal{R}_{\widehat{\Pr}}$ is the largest, we follow back-pointers to obtain a price vector consistent with $\widehat{\Pr}$. At some steps of the back-tracing, there may be multiple choices; we pick an arbitrary one to proceed.

Running Time and Correctness. Next we bound the algorithm's running time and revenue.

Lemma 16. *Given an instance of* RestrictedPrice*, where the value distributions are supported on a discrete set $\mathcal{S}$ of cardinality k_1 and the prices are restricted to a discrete set $\mathcal{P}$ of cardinality k_2, the algorithm described in this section produces a price vector with expected revenue at least*

$$OPT - \left(\frac{2k_1k_2}{(nr)^2} + \frac{16}{n} \right) \cdot \min\{\mathcal{P}\},$$

where OPT is the optimal expected revenue, $\min\{\mathcal{P}\}$ *is the lowest element of $\mathcal{P}$, and r is the ratio of the largest to the smallest element of $\mathcal{P}$.*

Lemma 17. *The running time of the algorithm is polynomial in the size of the input and $(nr)^{O(k_1k_2)}$.*

Due to space limitations, we postpone the proofs of these lemmas to Appendix I. Intuitively, if we did not perform any rounding of distributions, our algorithm would have been *exact*, outputting an optimal price vector in $\{p^{(1)}, \ldots, p^{(k_2)}\}^n$. What we show is that the roundings performed at the steps of the dynamic program are fine enough that do not become detrimental to the revenue. To show this, we use the probabilistic concepts of total variation distance and coupling of random variables, invoking the coupling lemma and the optimal coupling theorem after each step of the algorithm. (See Lemma 50 in Appendix I.2.) This way, we show that the rounded (winning-value,winning-price) distributions maintained by the algorithm for each price vector are close in total variation distance to the corresponding exact distributions arising from these price vectors, culminating in Lemma 16.

Using Lemmas 16 and 17 and our work in previous sections, we obtain our main algorithmic results in this paper (Theorems 1, 2, 3, and 4). See Appendix J for the proof of these theorems.

References

[1] S. Bhattacharya, G. Goel, S. Gollapudi and K. Munagala. Budget constrained auctions with heterogeneous items. *Proceedings of the ACM Symposium on Theory of Computing, STOC* 2010.

[2] L. Blumrosen and T. Holenstein. Posted prices vs. negotiations: an asymptotic analysis. *Proceedings of the ACM Conference on Electronic Commerce, EC* 2008.

[3] P. Briest, S. Chawla, R. Kleinberg and S. M. Weinberg. Pricing Randomized Allocations. *Proceedings of SODA* 2010.

[4] P. Briest. Uniform Budgets and the Envy-Free Pricing Problem. *Proceedings of ICALP* 2008.

[5] S. Chawla, J. D. Hartline and R. D. Kleinberg. Algorithmic Pricing via Virtual Valuations. *Proceedings of the ACM Conference on Electronic Commerce,* EC 2007.

[6] S. Chawla, J. D. Hartline, D. Malec and B. Sivan. Multi-Parameter Mechanism Design and Sequential Posted Pricing. *Proceedings of the ACM Symposium on Theory of Computing, STOC* 2010.

[7] S. Chawla, D. Malec and B. Sivan. The Power of Randomness in Bayesian Optimal Mechanism Design. *Proceedings of the ACM Conference on Electronic Commerce, EC*, 2010.

[8] C. Daskalakis and S. M. Weinberg. On Optimal Multi-Dimensional Mechanism Design. *Manuscript,* 2011.

[9] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction.* Springer Series in Operations Research, 2006.

[10] R. Durrett. *Random Graph Dynamics.* Cambridge University Press, 2006.

[11] J. D. Hartline and V. Koltun. Near-Optimal Pricing in Near-Linear Time. *Proceedings of WADS,* 2005.

[12] A. M. Manelli and D. R. Vincent. Bundling as an Optimal Selling Mechanism for a Multiple-Good Monopolist. *Journal of Economic Theory,* 127(1):1–35, 2006.

[13] 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.

[14] R. B. Myerson. Optimal Auction Design. *Mathematics of Operations Research,* 1981.

[15] N. Nisan, T. Roughgarden, E. Tardos and V. V. Vazirani (eds.). *Algorithmic Game Theory.* Cambridge University Press, 2007.

[16] M. Pai and R. Vohra. Optimal auctions with financially constrained bidders. *Working Paper,* 2008.

[17] J. Thanassoulis. Haggling over substitutes. *Journal of Economic Theory,* 117(2):217245, 2004.

[18] R. B. Wilson. *Nonlinear Pricing.* Oxford University Press, 1997.

Appendix

A An interesting example

A natural property than one may expect to hold is that, when the value distributions are discrete, there always exists an optimal solution that uses prices from the support of the value distributions. It turns out that this is not true. Here is an example:

Suppose that the seller has two items to sell, and the buyer's values for the items are v_1, which is uniform on $\{1, 5\}$, and v_2, which is uniform on $\{3, 3.5\}$. Moreover, assume that, if there is a tie between the value-minus-price gap for the two items, the buyer tie-breaks in favor of item 1. We claim that in this case the price vector $P = (4.5, 3)$ achieves higher revenue than any price vector that uses prices from the set $\{1, 3, 3.5, 5\}$ (where the values are drawn from.) Let us do the calculation. All our calculations are written in the form

$$\mathcal{R}_P = p_1 \times \Pr[\textit{item } 1 \textit{ is the winner}] + p_2 \times \Pr[\textit{item } 2 \textit{ is the winner}].$$

1. When $P = (4.5, 3)$
   $\mathcal{R}_P = 4.5 \times (1/2 \times 1) + 3 \times (1/2 \times 1) = 30/8$
2. When $P \in \{1, 3, 3.5, 5\}^2$:
   - If $P = (5, 3.5)$ then
     $\mathcal{R}_P = 5 \times (1/2 \times 1) + 3.5 \times (1/2 \times 1/2) = 27/8 < 30/8$
   - If $P = (5, 3)$ then
     $\mathcal{R}_P = 5 \times (1/2 \times 1/2) + 3 \times (1 \times 1/2 + 1/2 \times 1/2) = 28/8 < 30/8$
   - For any other price vector, the maximum revenue is bounded by $3.5 = 28/8 < 30/8$.

B Figures

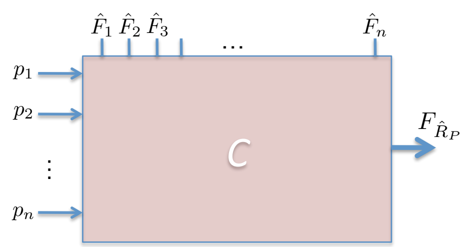

Figure 1: The Revenue Distribution.

C Access to Value Distributions

In this paper, we consider two ways that a distribution may be input to an algorithm.

- **Explicitly:** In this case the distribution has to be discrete, and we are given its support as a list of numbers, and the probability that the distribution places on every point in the support. If a distribution is provided explicitly to an algorithm, the algorithm is said to be efficient, if it runs in time polynomial the description complexity of the numbers required to specify the distribution.

- **As an Oracle:** In this case, we are given an oracle that answers queries about the value of the cumulative distribution function on a queried point. In particular, a query to the oracle consists of a point x and a precision ϵ, and the oracle outputs a value of bit complexity polynomial in the description of x and ϵ, which is within ϵ from the value of the cumulative distribution function at point x. Moreover, we assume that we are given an *anchoring point* x^* such that the value of the cumulative distribution at that point is between two a priori known absolute constants c_1 and c_2, such that $0 < c_1 < c_2 < 1$. Having such a point is necessary, as otherwise it would be impossible to find any interesting point in the support of the distribution (i.e. any point where the cumulative is different than 0 or 1).

  If a distribution is provided to an algorithm as an oracle, the algorithm is said to be efficient, if it runs in time polynomial in its other inputs and the bit complexity of x^*, ignoring the time spent by the oracle to answer queries (since this is not under the algorithm's control).

  If we have a closed form formula for our input distribution, e.g. if our distribution is $\mathcal{N}(\mu, \sigma^2)$, we think of it as given to us as an oracle, answering queries of the form (x, ϵ) as specified above. In most common cases, such an oracle can be implemented so that it also runs efficiently in the description of the query.

D Properties of MHR Distributions

Definition 18. *For a random variable X, we define $\alpha_1 = u_{min}$, and for every real number $p \in (1, +\infty)$, we define $\alpha_p = \inf\left\{x | F(x) \geq 1 - \frac{1}{p}\right\}$.*

The following lemma establishes an interesting property of MHR distributions. Intuitively, the lemma provides a lower bound on the speed of the decay of the tail of a MHR distribution. We prove the lemma by showing that the function $\log\big(1 - F(x)\big)$ is concave if F is MHR, and exploiting this concavity (see Appendix D.1).

Lemma 19. *If the distribution of a random variable X satisfies MHR, $m \geq 1$ and $d \geq 1$, $d \cdot \alpha_m \geq \alpha_{m^d}$.*

Next we study the expectation of a random variable that satisfies MHR. We show that the contribution to the expectation from values $\geq m$, is $O(m \cdot \Pr[X \geq m])$. We start with a definition.

Definition 20. *For a random variable X, let $Con[X \geq x] = \mathbb{E}[X | X \geq x] \cdot \Pr\{X \geq x\}$ be the contribution to expectation of X from values which are no smaller than x, i.e.*

$$Con[X \geq x] = \int_x^{+\infty} x \cdot f(x) dx.$$

It is an obvious fact that for any random variable X and any two points $x_1 \leq x_2$, $Con[X \geq x_1] \geq Con[X \geq x_2]$. Using the bound on the tail of a MHR distribution obtained in Lemma 19, we bound the contribution to the expectation of X by the values at the tail of the distribution. The proof is given in Appendix D.

Lemma 21. *Let X be a random variable whose distribution satisfies MHR. For all $m \geq 2$, $Con[X \geq \alpha_m] \leq 6\alpha_m / m$.*

D.1 Prooofs

Proof of Lemma 19: It is not hard to see that $f(x) > 0$, for all $x \in (u_{min}, u_{max})$. For a contradiction, assume this is not true, that is, for some $x' \in (u_{min}, u_{max})$, $f(x') = 0$. We know $1 - F(x') > 0$.

Thus $\frac{f(x')}{1-F(x')} = 0$. Since the distribution satisfies MHR and $1 - F(x)$ is positive for all $x \in (u_{min}, x')$, $f(x) = 0$ in this interval. Hence, it must also be that $F(x) = 0$ in $[u_{min}, x')$. Since $x' > u_{min}$, it follows that $u_{min} \neq \sup\{x | F(x) = 0\}$, a contradiction.

Since $f(x) > 0$ in (u_{min}, u_{max}), $F(x)$ is monotone in (u_{min}, u_{max}). So we can define the inverse $F^{-1}(x)$ in (u_{min}, u_{max}). It is not hard to see that for any $p \in [1, +\infty)$, $F(\alpha_p) = 1 - 1/p$ and $\alpha_p = F^{-1}(1 - 1/p)$.

Now let $G(x) = \log(1 - F(x))$. We will show that $G(x)$ is a concave function.

Let us consider the derivative of $G(x)$. By the definition of MHR, $G'(x) = \frac{-f(x)}{1-F(x)}$ is monotonically non-increasing. Therefore, $G(x)$ is concave. It follows that, for every m, by the concavity of $G(x)$, the following inequality holds:

$$G\left(\frac{d-1}{d} \cdot \alpha_1 + \frac{1}{d} \cdot \alpha_{m^d}\right) \geq \frac{d-1}{d} G(\alpha_1) + \frac{1}{d} G(\alpha_{m^d}).$$

Let us rewrite the RHS as follows

$$\begin{aligned}
&\frac{d-1}{d} G(\alpha_1) + \frac{1}{d} G(\alpha_{m^d}) \\
=&\frac{d-1}{d} \log 1 + \frac{1}{d} \log(1 - F(\alpha_{m^d})) \\
=&\frac{1}{d} \log\left(\frac{1}{m^d}\right) \\
=&\log\left(\frac{1}{m}\right)
\end{aligned}$$

Hence, we have the following:

$$\begin{aligned}
&G\left(\frac{d-1}{d} \cdot \alpha_1 + \frac{1}{d} \cdot \alpha_{m^d}\right) \geq \log\left(\frac{1}{m}\right) \\
\Longrightarrow&\log\left(1 - F\left(\frac{d-1}{d} \cdot \alpha_1 + \frac{1}{d} \cdot \alpha_{m^d}\right)\right) \geq \log\left(\frac{1}{m}\right) \\
\Longrightarrow&1 - F\left(\frac{d-1}{d} \cdot \alpha_1 + \frac{1}{d} \cdot \alpha_{m^d}\right) \geq \frac{1}{m} \\
\Longrightarrow&1 - F\left(\frac{d-1}{d} \cdot \alpha_1 + \frac{1}{d} \cdot \alpha_{m^d}\right) \geq 1 - F(\alpha_m) \\
\Longrightarrow&F(\alpha_m) \geq F\left(\frac{d-1}{d} \cdot \alpha_1 + \frac{1}{d} \cdot \alpha_{m^d}\right) \\
\Longrightarrow&\alpha_m \geq \frac{d-1}{d} \cdot \alpha_1 + \frac{1}{d} \cdot \alpha_{m^d} \quad (F \textit{ is monotone increasing}) \\
\Longrightarrow&\alpha_m \geq \frac{1}{d} \cdot \alpha_{m^d} \quad (u_{min} \geq 0) \\
\Longrightarrow&d \cdot \alpha_m \geq \alpha_{m^d}.
\end{aligned}$$

□

Proof of Lemma 21: Let $S = Con[X \geq \alpha_m]$, and consider the sequence $\{\beta_i := \alpha_{m^{(2^i)}}\}$, defined for all non-negative integers i. It can easily be seen that $\lim_{i \to +\infty} \alpha_{m^{(2^i)}} = u_{max}$; hence, $\lim_{i \to +\infty} \beta_i = u_{max}$ and by continuity $\lim_{i \to +\infty} F(\beta_i) = F(u_{max}) = 1$.

Also,

$$\int_{\beta_i}^{\beta_{i+1}} x \cdot f(x)dx \le \beta_{i+1}(1 - F(\beta_i)) = \beta_{i+1}/m^{(2^i)}.$$

Moreover, Lemma 19 implies that $\beta_i \le 2\beta_{i-1}$; thus, $\beta_i \le 2^i\beta_0 \le 2^i\alpha_m$. Hence, we have the following:

$$\begin{aligned}
S = \int_{\alpha_m}^{u_{max}} x \cdot f(x)dx &\le \sum_{i=0}^{+\infty} \frac{\beta_{i+1}}{m^{(2^i)}} \le \sum_{i=0}^{+\infty} \frac{2^{i+1}\alpha_m}{m^{(2^i)}} \\
&\le \frac{2\alpha_m}{m} + \sum_{i=1}^{+\infty} \frac{2^{(i+1)}\alpha_m}{m^{(2i)}} = \frac{2\alpha_m}{m} + \frac{4\alpha_m}{m^2}\sum_{i=0}^{+\infty}\left(\frac{2}{m^2}\right)^i \\
&= \frac{2\alpha_m}{m} + \frac{4\alpha_m}{m^2} \cdot \frac{1}{1 - 2/m^2} \\
&\le \frac{2\alpha_m}{m} + \frac{4\alpha_m}{m} \\
&\le \frac{6\alpha_m}{m}.
\end{aligned}$$

□

E Properties of Regular Distributions

If F is a differentiable continuous regular distribution, it is not hard to see the following: if $f(x) = 0$ for some x, then $f(x') = 0$ for all $x' \ge x$ (as otherwise the definition of regularity would be violated.) Hence, if X is a random variable distributed according to F, it must be that $f(x) > 0$ for $x \in [u_{min}^X, u_{max}^X]$. So we can define F^{-1} on $[u_{min}^X, u_{max}^X]$, and it will be differentiable, since F is differentiable and f is non-zero. Now we can make the following definition, capturing the revenue of a seller who prices an item with value distribution F, so that the item is bought with probability exactly q.

Definition 22 (Revenue Curve)**.** *For a differentiable continuous regular distribution F, define $R_F : [0,1] \to \mathbb{R}$ as follows*

$$R_F(q) = q \cdot F^{-1}(1 - q).$$

The following is well-known. We include its short proof for completeness.

Lemma 23. *If F is regular, $R_F(q)$ is a concave function on $(0,1]$.*

Proof. The derivative of $R_F(q)$ is

$$R'_F(q) = F^{-1}(1 - q) - \frac{q}{f\big(F^{-1}(1 - q)\big)}.$$

Notice that $F^{-1}(1 - q)$ is monotonically non-increasing in q. This observation and the regularity of F imply that $R'_F(q)$ is monotonically non-increasing in q. (To see this try the change of variable $x(q) = F^{-1}(1 - q)$.) This implies that $R_F(q)$ is concave. □

Lemma 24. *For any regular distribution F, if $0 < \tilde{q} \le q \le p < 1$, then*

$$R_F(\tilde{q}) \le \frac{1}{1 - p} R_F(q).$$

Proof. Since $q \in [\tilde{q}, 1)$, there exists a $\lambda \in (0, 1]$, such that

$$\lambda \cdot \tilde{q} + (1 - \lambda) \cdot 1 = q.$$

Hence: $\lambda = \frac{1-q}{1-\tilde{q}} \geq \frac{1-p}{1} = 1 - p$. Now, from Lemma 23, we have that $R_F(x)$ is concave. Thus

$$R_F(q) = R_F\big(\lambda \cdot \tilde{q} + (1 - \lambda) \cdot 1\big) \geq \lambda \cdot R_F(\tilde{q}) + (1 - \lambda) \cdot R_F(1).$$

Since $R_F(1) \geq 0$, $R_F(q) \geq \lambda \cdot R_F(\tilde{q}) \geq (1 - p) R_F(\tilde{q})$. Thus, $R_F(\tilde{q}) \leq \frac{1}{1-p} R_F(q)$. □

Corollary 25. *For any regular distribution F, if $\tilde{q} \leq q \leq \frac{1}{n^3}$, then*

$$R_F(\tilde{q}) \leq \frac{n^3}{n^3 - 1} R_F(q).$$

F Details of Section 3: MHR to Balanced Distributions

F.1 Proof of Theorem 12 (the Extreme Value Theorem for MHR Distributions)

We start with some useful notation. For all $i = 1, \ldots, n$, we denote by F_i the distribution of variable X_i. We also let $\alpha_m^{(i)} := \inf\left\{x | F_i(x) \geq 1 - \frac{1}{m}\right\}$, for all $m \geq 1$. Moreover, we assume that n is a power of 2. If not, we can always include at most n additional random variables that are detreministically 0, making the total number of variables a power of 2.

We proceed with the proof of Theorem 12. The threshold β is computed by an algorithm. At a high level, the algorithm proceeds in $O(\log n)$ rounds, indexed by $t \in \{0, \ldots, \log n\}$, eliminating half of the variables at each round. The way the elimination works is as follows. In round t, we compute for each of the variables that have survived so far the threshold $\alpha_{n/2^t}$ beyond which the size of the tail of their distribution becomes smaller than $1/(n/2^t)$. We then sort these thresholds and eliminate the bottom half of the variables, recording the threshold of the last variable that survived this round. The maximum of these records among the $\log n$ rounds of the algorithm is our β. The pseudocode of the algorithm is given below. Given that we may only be given oracle access to the distributions $\{F_i\}_{i \in [n]}$, we allow some slack $\eta \leq \frac{1}{2}$ in the computation of our thresholds so that the computation is efficient. If we know the distributions explicitly, the description of the algorithm simplifies to the case $\eta = 0$.

Algorithm 1 Algorithm for finding β

```
Define the permutation of the variables π_0(i) = i, ∀ i ∈ [n], and the set of remaining variables
  Q_0 = [n].
for t := 0 to log n − 1 do
    For all j ∈ [n/2^t], compute some x_{n/2^t}^{(π_t(j))} ∈ [1−η, 1+η] · α_{n/2^t}^{(π_t(j))}, for a small constant η ∈ [0, 1/2)
    Sort these n/2^t numbers in decreasing order π_{t+1} such that
      x_{n/2^t}^{(π_{t+1}(1))} ≥ x_{n/2^t}^{(π_{t+1}(2))} ≥ ... ≥ x_{n/2^t}^{(π_{t+1}(n/2^t))}
    Q_{t+1} := { π_{t+1}(i) | i ≤ n/2^{t+1} }
    β_t := x_{n/2^t}^{(π_{t+1}(n/2^{t+1}))}
end for
Compute x_2^{(π_{log n}(1))} ∈ [1 − η, 1 + η] · α_2^{(π_{log n}(1))}
Set β_{log n} := x_2^{(π_{log n}(1))}
Output β := max_t β_t
```

Crucial in the proof of the theorem is the following lemma.

Lemma 26. *For all $i \in [n]$ and $\epsilon \in (0,1)$, let $S_i = Con[X_i \geq 2\log(\frac{1}{\epsilon}) \cdot \beta]$, where $Con[\cdot]$ is defined as in Definition 20. Then*

$$\sum_{i=1}^{n} S_i \leq 36\log(1/\epsilon)\epsilon \cdot \beta, \quad \textit{for all } \epsilon \in (0, 1/4).$$

Proof. Let $d = \log(\frac{1}{\epsilon})$ and notice that $d \geq 2$. It is not hard to see that we can divide $[n]$ into $(\log n)+1$ different groups $\{G_t\}_{t\in\{0,\ldots,\log n\}}$ based on the sets Q_t maintained by the algorithm, as follows. For $t \in \{0, \ldots, \log n\}$, set

$$G_t = \begin{cases} Q_t \setminus Q_{t+1} & t < \log n \\ Q_{\log n} & t = \log n \end{cases}$$

Now, it is not hard to see that, for all $t < \log n$ and all $i \in G_t$, $S_i \leq Con[X_i \geq 2d \cdot \beta_t]$, since $\beta_t \leq \beta$. Also for any $i \in G_t$, there must exist some $k \in (n/2^{t+1}, n/2^t]$, such that $i = \pi_{t+1}(k)$. Then by the definition of the algorithm, we know that

$$(1-\eta)\alpha_{n/2^t}^{(i)} \leq x_{n/2^t}^{(i)} \leq x_{n/2^t}^{(\pi_{t+1}(n/2^{t+1}))} = \beta_t.$$

Recall that η is chosen to satisfy $2 \geq 1/(1-\eta)$. Then $d \cdot \alpha_{n/2^t}^{(i)} \leq 2d \cdot \beta_t$. But Lemma 19 gives $d \cdot \alpha_{n/2^t}^{(i)} \geq \alpha_{(n/2^t)^d}^{(i)}$. Hence,

$$2d \cdot \beta_t \geq d \cdot \alpha_{n/2^t}^{(i)} \geq \alpha_{(n/2^t)^d}^{(i)},$$

which implies that

$$Con[v_i \geq 2d \cdot \beta_t] \leq Con[v_i \geq \alpha_{(n/2^t)^d}^{(i)}].$$

Using Lemma 21, we know that

$$Con[v_i \geq \alpha_{(n/2^t)^d}^{(i)}] \leq 6\alpha_{(n/2^t)^d}^{(i)}(2^t/n)^d \leq 12d\beta_t(2^t/n)^d.$$

Now, since $|G_t| = n/2^{t+1}$,

$$\sum_{i\in G_t} S_i \leq 12d\beta_t(2^t/n)^d \times n/2^{t+1} = 6d \cdot \beta_t(2^t/n)^{d-1} = \frac{6d \cdot \beta_t}{n^{d-1}}(2^{d-1})^t.$$

Thus,

$$\begin{aligned}
\sum_{i\in[n]\setminus G_{\log n}} S_i \leq & \sum_{t=0}^{(\log n)-1} \frac{6d \cdot \beta_t}{n^{d-1}}(2^{d-1})^t \\
\leq & \frac{6d \cdot \beta}{n^{d-1}} \cdot \frac{(2^{d-1})^{\log_n} - 1}{2^{d-1} - 1} \\
= & \frac{6d \cdot \beta}{n^{d-1}} \cdot \frac{n^{d-1} - 1}{2^{d-1} - 1} \\
\leq & \frac{12d \cdot \beta}{2^d - 2} \\
\leq & \frac{24d \cdot \beta}{2^d} \\
= & 24\log(1/\epsilon)\epsilon \cdot \beta
\end{aligned}$$

Let i be the unique element in $G_{\log n}$. Then $\beta_{\log n} = x_2^{(i)}$. Using Lemma 19 and the definition of $x_2^{(i)}$, we obtain

$$2d \cdot \beta \geq 2d \cdot \beta_{\log n} \geq 2d \cdot x_2^{(i)} \geq 2(1-\eta)d \cdot \alpha_2^{(i)} \geq d \cdot \alpha_2^{(i)} \geq \alpha_{2^d}^{(i)} = \alpha_{1/\epsilon}^{(i)}.$$

Using the above and Lemma 21 we get

$$S_i \leq Con[v_i \geq \alpha_{1/\epsilon}^{(i)}] \leq 6\epsilon \cdot \alpha_{1/\epsilon}^{(i)} \leq 12\epsilon d \cdot \beta.$$

Putting everything together,

$$\sum_{i=1}^{n} S_i \leq 36 \log(1/\epsilon)\epsilon \cdot \beta.$$

□

Using Lemma 26, we obtain

$$\int_{2\beta \log 1/\epsilon}^{+\infty} t \cdot f_{\max_i\{X_i\}}(t)dt \leq \sum_{i=1}^{n} S_i \leq 36 \log(1/\epsilon)\epsilon \cdot \beta.$$

It remains to show that

$$\Pr[\max_i\{X_i\} \geq \beta/2] \geq 1 - \frac{1}{e^{1/2}}. \tag{3}$$

We show that, for all t, $\Pr\left[\max_i\{X_i\} \geq \frac{\beta_t}{1+\eta}\right] \geq 1 - \frac{1}{e^{1/2}}$, where η is the parameter used in Algorithm 1. This is sufficient to imply (3), as $\eta \leq 1/2$. Observe that for all $i \in [n/2^{t+1}]$,

$$(1+\eta) \cdot \alpha_{n/2^t}^{(\pi_{t+1}(i))} \geq x_{n/2^t}^{(\pi_{t+1}(i))} \geq \beta_t,$$

where π_{t+1} is the permutation constructed in the t-th round of the algorithm. This implies

$$\alpha_{n/2^t}^{(\pi_{t+1}(i))} \geq \frac{\beta_t}{1+\eta}.$$

Hence, for all $i \in [n/2^{t+1}]$, $\Pr[X_{\pi_{t+1}(i)} \leq \frac{\beta_t}{1+\eta}] \leq 1 - 2^t/n$. Thus,

$$\begin{aligned}\Pr\left[\max_i\{X_i\} \geq \frac{\beta_t}{1+\eta}\right] &\geq \Pr\left[\exists i \in [n/2^{t+1}],\ X_{\pi_{t+1}(i)} \geq \frac{\beta_t}{1+\eta}\right] \\ &\geq 1 - (1 - 2^t/n)^{n/2^{t+1}} \\ &\geq 1 - \frac{1}{e^{1/2}}.\end{aligned}$$

Eq. (3) now follows.

F.2 Proof of Theorem 11 (the Reduction from MHR to Balanced Distributions)

Recall that we represent by $\{v_i\}_{i\in[n]}$ the values of the buyer for the items. We will denote their distributions by $\{F_i\}_{i\in[n]}$ throughout this section.

F.2.1 Relating OPT to β

We demonstrate that the anchoring point β of Theorem 12 provides a lower bound to the optimal revenue. In particular, we show that the optimal revenue satisfies $OPT = \Omega(\beta)$. This lemma justifies the relevance of β.

Lemma 27. *If β is the anchoring point of Theorem 12, then $OPT \geq \left(1 - \frac{1}{\sqrt{e}}\right)\frac{\beta}{2}$.*

Proof of Lemma 27: Suppose we priced all items at $\frac{\beta}{2}$. The revenue we would get from such price vector would be at least

$$\frac{\beta}{2}\Pr\left[\max\{v_i\} \geq \frac{\beta}{2}\right] \geq \frac{\beta}{2}\left(1 - \frac{1}{\sqrt{e}}\right),$$

where we used Theorem 12. Hence, $OPT \geq \left(1 - \frac{1}{\sqrt{e}}\right)\frac{\beta}{2}$. $\square$

For simplicity, we set $c_1 := \frac{1}{2}\left(1 - \frac{1}{\sqrt{e}}\right)$ for the next sections, keeping in mind that c_1 is an absolute constant.

F.2.2 Restricting the Prices

This section culminates in Lemma 30 (given below), which states that we can constrain our prices to the set $[\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]$ without hurting the revenue by more than a fraction of $\frac{\epsilon + c_2(\epsilon)}{c_1}$, where $c_2(\epsilon) := 36\log(\frac{1}{\epsilon})\epsilon$ and c_1 is the constant defined in Section F.2.1. We prove this in two steps. First, exploiting our extreme value theorem for MHR distributions (Theorem 12), we show that for a given price vector, if we lower the prices that are above $2\log(\frac{1}{\epsilon}) \cdot \beta$ to $2\log(\frac{1}{\epsilon}) \cdot \beta$, the loss in revenue is bounded by $c_2(\epsilon) \cdot \beta$, namely

Lemma 28. *Given any price vector P we define P' as follows: we set $p'_i = p_i$, if $p_i \leq 2\log(\frac{1}{\epsilon}) \cdot \beta$, and $p'_i = 2\log(\frac{1}{\epsilon}) \cdot \beta$ otherwise, where $\epsilon \in (0, 1/4)$. Then the expected revenues $\mathcal{R}_P$ and $\mathcal{R}_{P'}$ achieved by price vectors P and P' respectively satisfy: $\mathcal{R}_{P'} \geq \mathcal{R}_P - c_2(\epsilon) \cdot \beta$.*

We complement Lemma 28 with the following lemma, establishing that, for any given price vector P, if we increase all prices below $\epsilon \cdot \beta$ to $\epsilon \cdot \beta$, the loss in revenue is at most $\epsilon \cdot \beta$.

Lemma 29. *Let P be any price vector and define P' as follows: set $p'_i = p_i$, if $p_i \geq \alpha$, and $p'_i = \alpha$ otherwise. The expected revenues $\mathcal{R}_P$ and $\mathcal{R}_{P'}$ from these price vectors satisfy $\mathcal{R}_{P'} \geq \mathcal{R}_P - \alpha$.*

Combining these lemmas, we obtain Lemma 30. Observe that we can make the loss in revenue arbitrarily small be taking ϵ sufficiently small.

Lemma 30. *For all $\epsilon \in (0, 1/4)$, there exists a price vector $P^* \in [\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]^n$, such that the revenue from this price vector satisfies $\mathcal{R}_{P^*} \geq \left(1 - \frac{\epsilon + c_2(\epsilon)}{c_1}\right) OPT$, where OPT is the optimal revenue under any price vector.*

All proofs of this section are in Appendix F.3.1.

F.2.3 Truncating the Value Distributions

Exploiting Lemma 30, i.e. that we can constrain the prices to $[\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]$ without hurting the revenue, we show Theorem 11, i.e. that we can also constrain the support of the value distributions into a balanced range. In particular, we show that we can "truncate" the value distributions to the range $[\frac{\epsilon}{2} \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]$, where for our purposes "truncating" means this: for every distribution F_i, we shift all probability mass from $(2\log(\frac{1}{\epsilon})\cdot\beta, +\infty)$ to the point $2\log(\frac{1}{\epsilon})\cdot\beta$, and all probability mass from $(-\infty, \epsilon\cdot\beta)$ to $\frac{\epsilon}{2}\cdot\beta$. Clearly, this modification can be computed in polynomial time if F_i is known explicitly; if we have oracle access to F_i, we can produce in polynomial time another oracle whose output behaves according to the modified distribution. We show that our modification does not hurt the revenue. That is, we establish a polynomial-time reduction from the problem of computing a near-optimal price vector when the buyer's value distributions are arbitrary MHR distributions to the case where the buyer's value distributions are supported on a balanced interval $[u_{min}, c\cdot u_{min}]$, where $c = c(\epsilon) = 4\frac{1}{\epsilon}\log(\frac{1}{\epsilon})$ is a constant that only depends on the desired approximation ϵ. The proof of Theorem 31 is given in Appendix F.3.2.

Theorem 31 (Reduction from MHR to Balanced Distributions)**.** *Given $\epsilon \in (0, 1/4)$ and a collection of mutually independent random variables $\{v_i\}_i$ that are MHR, let us define a new collection of random variables $\{\tilde{v}_i\}_i$ via the following coupling: for all $i \in [n]$, set $\tilde{v}_i = \frac{\epsilon}{2}\cdot\beta$ if $v_i < \epsilon\cdot\beta$, set $\tilde{v}_i = 2\log(\frac{1}{\epsilon})\cdot\beta$ if $v_i \geq 2\log(\frac{1}{\epsilon})\cdot\beta$, and set $\tilde{v}_i = v_i$ otherwise, where $\beta = \beta(\{v_i\}_i)$ is the anchoring point of Theorem 12 computed from the distributions of the variables $\{v_i\}_i$. Let also $\widetilde{OPT}$ be the optimal revenue of the seller when the buyer's values are distributed as $\{\tilde{v}_i\}_{i\in[n]}$ and OPT the optimal revenue when the buyer's values are distributed as $\{v_i\}_{i\in[n]}$. Then given a price vector that achieves revenue $(1-\delta)\cdot\widetilde{OPT}$ when the buyer's values are distributed as $\{\tilde{v}_i\}_{i\in[n]}$, we can efficientlly compute a price vector with revenue*

$$\left(1 - \delta - \frac{2\epsilon + 3c_2(\epsilon)}{c_1}\right) OPT$$

when the buyer's values are distributed as $\{v_i\}_{i\in[n]}$.

Theorem 11 follows from Theorem 31.

F.3 Proofs Omitted from Section F.2

F.3.1 Restricting the Price Range for MHR Distributions: the Proofs

Proof of Lemma 28: We first show that, given a price vector, if we make all prices that are above α equal to $+\infty$, then the loss in revenue can bounded by the sum, overall items whose price was turned into $+\infty$, of the contribution to this item's expected value by points above α. Formally,

Lemma 32. *Let $\alpha > 0$ and $S(\alpha) = Con[\max_i v_i \geq \alpha]$, where $Con[\cdot]$ is defined as in Definition 20. Moreover, for a given price vector P, define P' as follows: set $p'_i = p_i$, if $p_i < \alpha$, and $p'_i = +\infty$, otherwise. Then the expected revenues $\mathcal{R}_P$ and $\mathcal{R}_{P'}$ from P and P' respectively satisfy*

$$\mathcal{R}_{P'} \geq \mathcal{R}_P - S(\alpha).$$

Proof. Let $\pi_i = \Pr[\ \forall j : v_i - p_i \geq v_j - p_j \wedge v_i - p_i \geq 0]$, $\pi'_i = \Pr[\forall j :\ v_i - p'_i \geq v_j - p'_j \wedge v_i - p'_i \geq 0]$ denote respectively the probability that item i is bought under price vectors P and P'.

We partition the items into two sets, expensive and cheap, based on P: $S_{exp} = \{\ i \mid p_i > \alpha\ \}$ and $S_{chp} = \{\ i \mid p_i \leq \alpha\ \}$. If $i \in S_{exp}$, $p'_i = +\infty$, while if $i \in S_{chp}$, $p'_i = p_i$.

For any $i \in S_{chp}$, $\pi_i \leq \pi'_i$. Indeed, for any valuation vector $(v_1, v_2, \ldots, v_n)$: if $v_i - p_i \geq v_j - p_j$, for all j, and $v_i - p_i \geq 0$, it should also be true that $v_i - p'_i \geq v_j - p'_j$, for all j, and $v_i - p'_i \geq 0$ (since we only increased the prices of items different from i). Therefore, the following inequality holds:

$$\sum_{i\in S_{chp}} p_i \cdot \pi_i \leq \sum_{i\in S_{chp}} p'_i \cdot \pi'_i. \quad (1)$$

Notice that, for any $i \in S_{exp}$, whenever the event $C_i = (\forall j :\ v_i - p_i \geq v_j - p_j) \wedge (v_i - p_i \geq 0)$ happens, v_i has to be larger than p_i, which itself is larger than α. Hence, whenever the event $\cup_{i\in S_{exp}} C_i$

happens, it must be that $\max_i\{v_i\} \geq \alpha$. Given that $\max_i\{v_i\}$ is the best revenue we could hope to extract (pointwise for all vectors $(v_1, \ldots, v_n)$), it follows that

$$\sum_{i \in S_{exp}} p_i \cdot \pi_i \leq Con\left[\max_i\{v_i\} \geq \alpha\right] = S(\alpha).$$

Combining this with (1), we obtain

$$\mathcal{R}_{P'} \geq \mathcal{R}_P - S(\alpha).$$

□

We showed that setting the price of expensive items to $+\infty$ is not detrimental to the revenue. We show that the same is true of a less aggressive strategy.

Lemma 33. *Let P be a price vector, define $S_\infty = \{i : p_i = +\infty\}$, and let $p_{max} \geq \max_{i \in [n] \setminus S_\infty} p_i$. Let P' be the following price vector: set $p'_i = p_{max}$, for all $i \in S_\infty$, and $p'_i = p_i$ otherwise. Then the expected revenues $\mathcal{R}_P$ and $\mathcal{R}_{P'}$ from P and P' respectively satisfy*

$$\mathcal{R}_{P'} \geq \mathcal{R}_P.$$

Proof. Take any valuation vector $(v_1, v_2, \ldots, v_n)$ and suppose that i is the winner under price vector P, i.e. $v_i - p_i \geq v_j - p_j$, for all j, and $v_i - p_i \geq 0$. Under P', there are two possibilities: (1) i is still the winner; then the contribution to the revenue is the same under P and P'. (2) i is not the winner; in this case, the winner should be among those items whose price was lowered from P to P'. So the new winner must some $j \in S_\infty$. But notice that $p'_j = p_{max} \geq p_i$. Hence, the contribution of item j to revenue under price vector P' is not smaller than the contribution of item i to the revenue under price vector P. Hence, $\mathcal{R}_{P'} \geq \mathcal{R}_P$. □

Combining Theorem 12 with Lemmas 32 and 33, it is easy to argue that if we truncate a price vector P at value $2\log(\frac{1}{\epsilon}) \cdot \beta$ to obtain a new price vector P' the change in revenue can be bounded as follows:

$$\mathcal{R}_{P'} \geq \mathcal{R}_P - c_2(\epsilon) \cdot \beta.$$

□

Proof of Lemma 29: Let $\pi_i = \Pr[(\forall j : v_i - p_i \geq v_j - p_j) \wedge (v_i - p_i \geq 0)]$ and $\pi'_i = \Pr[(\forall j : v_i - p'_i \geq v_j - p'_j) \wedge (v_i - p'_i \geq 0)]$, and define $S_{low} = \{i : p_i \leq \alpha\}$. For all $i \in [n] \setminus S_{low}$, we claim that $\pi'_i \geq \pi_i$. Indeed, since we only increased the price of items other than i, for any valuation vector $v = (v_1, v_2, \ldots, v_n)$, if $v_i - p_i \geq v_j - p_j$, for all j, and $v_i - p_i \geq 0$, it should also be true that $v_i - p'_i \geq v_j - p'_j$, for all j, and $v_i - p'_i \geq 0$.

Moreover, it is easy to see that

$$\sum_{i \in S_{low}} p_i \cdot \pi_i \leq \alpha.$$

Hence, we have the following inequality:

$$\mathcal{R}_{p'} \geq \sum_{i:[n] \setminus S_{low}} p'_i \cdot \pi'_i \geq \sum_{i:[n] \setminus S_{low}} p_i \cdot \pi_i \geq \mathcal{R}_p - \alpha.$$

□

Proof of Lemma 30: Lemma 28 implies that, if we start from any price vector P, we can modify it into another price vector P' that does not use any price above $2\log(\frac{1}{\epsilon}) \cdot \beta$, and satisfies $\mathcal{R}_{P'} \geq \mathcal{R}_P - c_2(\epsilon) \cdot \beta$.

Then Lemma 29 implies that we can change P' into another vector $P'' \in [\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]^n$, such that $\mathcal{R}_{P''} \geq \mathcal{R}_{P'} - \epsilon \cdot \beta$.

By Lemma 27, we know that $OPT \geq c_1 \cdot \beta$. Hence, if we start with the optimal price vector P and apply the above transformations, we will obtain a price vector $P^* \in [\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]^n$ such that

$$\mathcal{R}_{P^*} \geq OPT - \big(\epsilon + c_2(\epsilon)\big) \cdot \beta \geq \left(1 - \frac{\epsilon + c_2(\epsilon)}{c_1}\right) OPT.$$

□

F.3.2 Bounding the Support of the Distributions: the Proofs

To establish Theorem 31 we show that we can transform $\{v_i\}_{i\in[n]}$ into $\{\tilde{v}_i\}_{i\in[n]}$ such that, for all i, $\tilde{v}_i$ only takes values in $[\frac{\epsilon}{2} \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]$, and for any price vector $P \in [\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]^n$, $|\tilde{\mathcal{R}}_P - \mathcal{R}_P| \leq c_2(\epsilon) \cdot \beta$, where $\mathcal{R}_P$ and $\tilde{\mathcal{R}}_P$ are respectively the revenues of the seller when the buyer's values are distributed as $\{v_i\}_{i\in[n]}$ and $\{\tilde{v}_i\}_{i\in[n]}$. We first show that one side of our truncation works.

Lemma 34. *Given $\epsilon \in (0, 1/4)$ and a collection of random variables $\{v_i\}_i$ that are MHR, let us define a new collection of random variables $\{\hat{v}_i\}_i$ via the following coupling: for all $i \in [n]$, if $v_i \leq 2\log(\frac{1}{\epsilon}) \cdot \beta$, set $\hat{v}_i = v_i$, otherwise set $\hat{v}_i = 2\log(\frac{1}{\epsilon}) \cdot \beta$, where $\beta = \beta(\{v_i\}_i)$ is the anchoring point of Theorem 12 computed from the distributions of the variables $\{v_i\}_i$. Then, for any price vector $P \in [\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]^n$, $|\mathcal{R}_P - \hat{\mathcal{R}}_P| \leq c_2(\epsilon) \cdot \beta$, where $\mathcal{R}_P$ and $\hat{\mathcal{R}}_P$ are respectively the revenues of seller when the buyer's values are distributed as $\{v_i\}_{i\in[n]}$ and as $\{\hat{v}_i\}_{i\in[n]}$.*

Proof. For convenience let $d = \log(\frac{1}{\epsilon})$. Recall that $\{v_i\}_i$ and $\{\hat{v}_i\}_i$ are defined via a coupling. We distinguish two events: (1) $v := (v_1, v_2, \ldots, v_n) \equiv (\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n) =: \hat{v}$; (2) $v_i \neq \hat{v}_i$, for some i.

Under Event (1), the item sold is the same under v and $\hat{v}$. Hence, the revenue is the same under Event (1). So we only need to worry about Event (2). We show that the total contribution to the revenue from this event is very small. Indeed, the probability of this event is $\Pr[\exists i, v_i > 2d \cdot \beta]$, and the maximum price is $2d \cdot \beta$. So using our extreme value theorem (Theorem 12), we can bound the contribution to the revenue from this event as follows.

$$2d \cdot \beta \cdot \Pr[\exists i, v_i > 2d \cdot \beta] = 2d \cdot \beta \cdot \Pr[\max_i v_i \geq 2d \cdot \beta] \leq Con[\max_i v_i \geq 2d \cdot \beta] \leq c_2(\epsilon) \cdot \beta.$$

We obtain $|\mathcal{R}_P - \hat{\mathcal{R}}_P| \leq c_2(\epsilon) \cdot \beta$. □

Next we show that the other side of the truncation works.

Lemma 35. *Given $\epsilon, \beta > 0$ and a collection of random variables $\{\hat{v}_i\}_i$, let us define a new collection of random variables $\{\tilde{v}_i\}_i$ via the following coupling: for all $i \in [n]$, if $\hat{v}_i \geq \epsilon \cdot \beta$, set $\tilde{v}_i = \hat{v}_i$, otherwise set $\tilde{v}_i = \frac{\epsilon}{2} \cdot \beta$. Then, for any price vector $P \in [\epsilon \cdot \beta, +\infty)^n$, $\tilde{\mathcal{R}}_P = \hat{\mathcal{R}}_P$, where $\hat{\mathcal{R}}_P$ and $\tilde{\mathcal{R}}_P$ are respectively the revenues of the seller when the buyer's values are distributed as $\{\hat{v}_i\}_{i\in[n]}$ and as $\{\tilde{v}_i\}_{i\in[n]}$.*

Proof. Recall that $\{\hat{v}_i\}_i$ and $\{\tilde{v}_i\}_i$ are defined via a coupling. For any value of $\hat{v} = (\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n)$, we distinguish the following cases: (1) $\tilde{v} = (\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_n)$ is exactly the same as $\hat{v}$; (2) there exists some i such that $\hat{v}_i \neq \tilde{v}_i$.

In the first case, it is clear that the revenue of the seller is the same under values $\hat{v}$ and $\tilde{v}$. In the second case, we distinguish two further subcases:

- if there is some item i^* sold under values $\hat{v}$, we show that i^* is also sold under values $\tilde{v}$. Indeed, since i^* is the winner under $\hat{v}$, we have that, for all i,

$$\hat{v}_{i^*} - p_{i^*} \geq \hat{v}_i - p_i;$$

and

$$\hat{v}_{i^*} - p_{i^*} \geq 0.$$

Given that $p_{i^*} \geq \epsilon \cdot \beta$, it must be that $\hat{v}_{i^*} \geq \epsilon \cdot \beta$. Therefore, $\tilde{v}_{i^*} = \hat{v}_{i^*}$. Hence, for any i such that $\tilde{v}_i = \hat{v}_i$, $\tilde{v}_{i^*} - p_{i^*} \geq \tilde{v}_i - p_i$; and, for any i such that $\tilde{v}_i \neq \hat{v}_i$, $\tilde{v}_i = \frac{\epsilon}{2} \cdot \beta < p_i$. Thus, $\tilde{v}_{i^*} - p_{i^*} \geq 0 > \tilde{v}_i - p_i$. So i^* is also the winner under $\tilde{v}$.

- if there is no winner under $\hat{v}$, that is $\hat{v}_i - p_i \leq 0$, for all i, then: for any i such that $\tilde{v}_i = \hat{v}_i$, $\tilde{v}_i - p_i \leq 0$, and for any i such that $\tilde{v}_i \neq \hat{v}_i$, $\tilde{v}_i = \frac{\epsilon}{2} \cdot \beta < p_i$, as we argued above. So, $\tilde{v}_i - p_i \leq 0$, for all i.

Hence, we have argued the revenues under $\hat{v}$ and $\tilde{v}$ are the same in the second case too. Therefore,

$$\tilde{\mathcal{R}}_P = \hat{\mathcal{R}}_P.$$

□

Putting these lemmas together we obtain our reduction.

Proof of Theorem 31: Let P be a near-optimal price vector when the values of the buyer are distributed as $\{\tilde{v}_i\}_{i\in[n]}$, i.e. one that satisfies

$$\tilde{\mathcal{R}}_P \geq (1-\delta) \cdot \widetilde{OPT},$$

where $\tilde{\mathcal{R}}_P$ denotes the expected revenue of the seller under price vector P when the buyer's values are $\{\tilde{v}_i\}_{i\in[n]}$. Given that each $\tilde{v}_i$ lies in $[\frac{\epsilon}{2} \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]$, it follows from Lemma 41 that we can (efficiently) transform P into another vector $P' \in [\frac{\epsilon}{2} \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]^n$, such that $\tilde{\mathcal{R}}_P \leq \tilde{\mathcal{R}}_{P'}$.

We can then apply the following efficient transformation to P', to get P'': For any i, if $p'_i < \epsilon \cdot \beta$, set $p''_i = \epsilon \cdot \beta$, and set $p''_i = p'_i$ otherwise. By Lemma 29, we know that,

$$\tilde{\mathcal{R}}_{P''} \geq \tilde{\mathcal{R}}_{P'} - \epsilon \cdot \beta.$$

Now, since P'' is a price vector in $[\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]^n$, by Lemmas 34 and 35, we get

$$\mathcal{R}_{P''} \geq \tilde{\mathcal{R}}_{P''} - c_2(\epsilon) \cdot \beta,$$

where $\mathcal{R}_{P''}$ is the expected revenue of the seller under price vector P'' when the values of the buyers are $\{v_i\}_i$.

On the other hand, suppose that P^* is the optimal price vector in $[\epsilon \cdot \beta, 2\log(\frac{1}{\epsilon}) \cdot \beta]^n$ for values $\{v_i\}_{i\in[n]}$. By Lemma 30, we know that $\mathcal{R}_{P^*} \geq \left(1 - \frac{\epsilon + c_2(\epsilon)}{c_1}\right) OPT$. Now Lemmas 34 and 35 give

$$\tilde{\mathcal{R}}_{P^*} \geq \mathcal{R}_{P^*} - c_2(\epsilon) \cdot \beta \geq \left(1 - \frac{\epsilon + c_2(\epsilon)}{c_1}\right) OPT - c_2(\epsilon) \cdot \beta \geq \left(1 - \frac{\epsilon + 2c_2(\epsilon)}{c_1}\right) OPT,$$

where we used that $OPT \geq c_1 \cdot \beta$, by Lemma 27.

Since $\widetilde{OPT} \geq \tilde{\mathcal{R}}_{P^*}$,

$$\tilde{\mathcal{R}}_{P'} \geq \tilde{\mathcal{R}}_P \geq (1-\delta)\widetilde{OPT} \geq (1-\delta)\left(1 - \frac{\epsilon + 2c_2(\epsilon)}{c_1}\right) OPT \geq \left(1 - \delta - \frac{\epsilon + 2c_2(\epsilon)}{c_1}\right) OPT.$$

Recall that $\mathcal{R}_{P''} \geq \tilde{\mathcal{R}}_{P''} - c_2(\epsilon) \cdot \beta \geq \tilde{\mathcal{R}}_{P'} - \epsilon \cdot \beta - c_2(\epsilon) \cdot \beta$. Therefore,

$$\mathcal{R}_{P''} \geq \left(1 - \delta - \frac{\epsilon + 2c_2(\epsilon)}{c_1}\right) OPT - \epsilon \cdot \beta - c_2(\epsilon) \cdot \beta \geq \left(1 - \delta - \frac{2\epsilon + 3c_2(\epsilon)}{c_1}\right) OPT.$$

So given a near-optimal price vector P for $\{\tilde{v}_i\}_{i\in[n]}$, we can construct a near-optimal price vector P'' for $\{v_i\}_{i\in[n]}$ in polynomial time. □

G Details of Section 4

G.1 Proof of Theorem 14 (the Extreme Value Theorem for Regular Distributions)

We define α explicitly from the distributions $\{F_i\}_i$ of the variables $\{X_i\}_i$. We first need a definition.

Definition 36. *A point x is a (c_1, c_2)-anchoring point of a distribution F, if $F(x) \in [c_1, c_2]$.*

Now fix two arbitrary constants $0 < c_1 < c_2 \leq \frac{7}{8}$, and let, for all i, α_i be a (c_1, c_2)-anchoring point of the distribution F_i. Then define

$$\alpha = \frac{n^3}{c_1} \cdot \max_i \left[\alpha_i \cdot \big(1 - F_i(\alpha_i)\big)\right].$$

Clearly, a collection $\alpha_1, \ldots, \alpha_n$ of (c_1, c_2)-anchoring points can be computed efficiently from the F_i's (whether we have these distributions explicitly or have oracle access to them.) Hence, an α as above can be computed efficiently. We proceed to establish anchoring properties satisfied by α.

Proposition 37. *$\alpha \geq \max_i \alpha^{(i)}_{n^3}$, where $\alpha^{(i)}_p = \inf\left\{x | F_i(x) \geq 1 - \frac{1}{p}\right\}$ as in Definition 18.*

Proof. Because $1/n^3 \leq 1 - c_2 \leq 1 - F(\alpha_i) \leq 1 - c_1$, it follows from Lemma 24 that

$$\frac{1}{c_1} \cdot \alpha_i \cdot \big(1 - F_i(\alpha_i)\big) \geq \alpha^{(i)}_{n^3}/n^3.$$

Hence: $\alpha \geq \frac{n^3}{c_1} \cdot \left[\alpha_i \cdot \big(1 - F_i(\alpha_i)\big)\right] \geq \alpha^{(i)}_{n^3}$. This is true for all i, hence the theorem. $\square$

Proof of Theorem 14: We first show that $\Pr[X_i \geq \ell\alpha] \leq 2/(\ell n^3)$, for any $\ell \geq 1$. By Proposition 37 and Corollary 25, we have that $(\ell\alpha)\Pr[X_i \geq \ell\alpha] \leq \frac{n^3}{n^3-1}\alpha \Pr[X_i \geq \alpha]$. Thus

$$\Pr[X_i \geq \ell\alpha] \leq \frac{n^3}{n^3 - 1} \cdot \frac{1}{\ell} \cdot \Pr[X_i \geq \alpha] \leq 2/(\ell n^3), \tag{4}$$

which establishes the first anchoring property satisfied by α.

Moreover, we have that

$$\alpha/n^3 = \frac{1}{c_1} \cdot \max_i \left[a_i \cdot \big(1 - F_i(a_i)\big)\right] \leq \frac{1}{c_1} \max_z (z \cdot \Pr[\max_i\{X_i\} \geq z]),$$

which establishes the second anchoring property of α.

Finally, we demonstrate the homogenization property of α. We want to show that, for any integer $m \leq n$, thresholds $t_1, \ldots, t_m \geq t \geq \frac{2n^2\alpha}{\epsilon^2}$, index set $S = \{a_1, \ldots, a_m\} \subseteq [n]$, and $\epsilon \in (0, 1)$:

$$\sum_{i=1}^{m} t_i \Pr[X_{a_i} \geq t_i] \leq \left(t - \frac{2\alpha}{\epsilon}\right) \cdot \Pr\left[\max_i\{X_{a_i}\} \geq t\right] + \frac{7\epsilon \cdot (2\alpha/\epsilon \cdot \Pr[\max_i\{X_{a_i}\} \geq 2\alpha/\epsilon])}{n}.$$

For notational simplicity, we define $f_i(z_i) = z_i \cdot \Pr[X_{a_i} \geq z_i]$ and $f^{(S)}_{max}(z) = z \cdot \Pr[\max_i\{X_{a_i}\} \geq z]$. Notice that for any $t_i \geq t \geq 2\alpha/\epsilon$, a double application of Proposition 37, Lemma 24 and Equation (4) gives

$$f_i(t_i) \leq \frac{(n^3/\epsilon)}{(n^3/\epsilon) - 1} f_i(t) \leq \frac{2(n^3/\epsilon)}{(n^3/\epsilon) - 1} f_i\left(\frac{2\alpha}{\epsilon}\right). \tag{5}$$

Thus,

$$\begin{aligned}
LHS &\le \sum_{i=1}^{m} f_i(t) + \frac{1}{(n^3/\epsilon)-1}\sum_{i=1}^{m} f_i(t) \\
&\le \sum_{i=1}^{m} f_i(t) + \frac{2}{(n^3/\epsilon)-1}\sum_{i=1}^{m} f_i\left(\frac{2\alpha}{\epsilon}\right) \\
&\le \sum_{i=1}^{m} f_i(t) + \frac{2n}{(n^3/\epsilon)-1} f_{max}^{(S)}\left(\frac{2\alpha}{\epsilon}\right) \\
&\le \sum_{i=1}^{m} f_i(t) + \frac{2\epsilon}{n} f_{max}^{(S)}\left(\frac{2\alpha}{\epsilon}\right).
\end{aligned}$$

On the other hand, for any $t \ge 2\alpha/\epsilon$: $\Pr[X_{a_i} \ge t] \le \Pr[X_{a_i} \ge 2\alpha/\epsilon] \le \epsilon/n^3$ (using (4)). Thus:

$$\sum_i \Pr[X_{a_i} \ge t] \ge \Pr[\max_i\{X_{a_i}\} \ge t] \ge (1-\epsilon/n^2)\sum_i \Pr[X_{a_i} \ge t], \tag{6}$$

where the last inequality follows from the fact that, for all i, the probability that $X_{a_i} \ge t$, while $X_{a_j} < t$ for all $j \in S \setminus \{i\}$ is at least $\Pr[X_{a_i} \ge t](1-\epsilon/n^3)^{m-1} \ge \Pr[X_{a_i} \ge t](1-\epsilon/n^2)$. Therefore, continuing our upper-bounding from above:

$$\begin{aligned}
LHS \le \quad & \sum_{i=1}^{m} f_i(t) + \frac{2\epsilon}{n} f_{max}^{(S)}\left(\frac{2\alpha}{\epsilon}\right) \\
&\le (t-2\alpha/\epsilon)\Pr[\max_i\{X_{a_i}\} \ge t] + (2\alpha/\epsilon)\Pr[\max_i\{X_{a_i}\} \ge t] + (\epsilon/n^2)\sum_{i=1}^{m} f_i(t) + \frac{2\epsilon}{n} f_{max}^{(S)}\left(\frac{2\alpha}{\epsilon}\right) \\
&\le (t-2\alpha/\epsilon)\Pr[\max_i\{X_{a_i}\} \ge t] + (2\alpha/\epsilon t)\sum_{i=1}^{m} f_i(t) + (2\epsilon/n^2)\sum_{i=1}^{m} f_i\left(\frac{2\alpha}{\epsilon}\right) + \frac{2\epsilon}{n} f_{max}^{(S)}\left(\frac{2\alpha}{\epsilon}\right) \\
&\le (t-2\alpha/\epsilon)\Pr[\max_i\{X_{a_i}\} \ge t] + (\epsilon/n^2)\sum_{i=1}^{m} f_i(t) + \frac{4\epsilon}{n} f_{max}^{(S)}\left(\frac{2\alpha}{\epsilon}\right) \\
&\le (t-2\alpha/\epsilon)\Pr[\max_i\{X_{a_i}\} \ge t] + \frac{6\epsilon}{n} f_{max}^{(S)}\left(\frac{2\alpha}{\epsilon}\right),
\end{aligned}$$

where we got the third inequality by invoking (5) and (6), the fourth inequality by invoking (6) with $t = 2\alpha/\epsilon$, and the fifth inequality by invoking (5) and then (6) with $t = 2\alpha/\epsilon$. This concludes the proof of Theorem 14. □

G.2 Discussion of Theorem 14

In this section we play around with Theorem 14 to gain some intuition about its meaning:

- Suppose that we set all the t_i's equal to $t \ge 2n^2\alpha/\epsilon^2$. In this case, the homogenization property of Theorem 14 essentially states that the union bound is tight for t large enough. Indeed:

$$\begin{aligned}
\Pr\left[\max_i\{X_{a_i}\} \ge t\right] &\le \left(\sum_{i=1}^{m} \Pr[X_{a_i} \ge t]\right) \\
&\le \left(\frac{t-\frac{2\alpha}{\epsilon}}{t}\right)\cdot \Pr\left[\max_i\{X_{a_i}\} \ge t\right] + \frac{7\epsilon}{tn}\cdot\left(\frac{2\alpha}{\epsilon}\cdot \Pr\left[\max_i\{X_{a_i}\} \ge \frac{2\alpha}{\epsilon}\right]\right) \\
&\le \Pr\left[\max_i\{X_{a_i}\} \ge t\right] + \frac{7\epsilon}{tn}\cdot\left(\frac{2\alpha}{\epsilon}\cdot \Pr\left[\max_i\{X_{a_i}\} \ge \frac{2\alpha}{\epsilon}\right]\right).
\end{aligned}$$

This is not surprising, since for all i, the event $X_{a_i} \geq t$ only happens with tiny probability, by the anchoring property of α.

- Now let's try to set all the t_i's to the same value $t' > t \geq 2n^2\alpha/\epsilon^2$. The homogenization property can be used to obtain that the probability of the event $\max_i\{X_{a_i}\} \geq t'$ scales linearly in t'.

$$\begin{aligned}
\Pr\left[\max_i\{X_{a_i}\} \geq t'\right] &\leq \sum_{i=1}^{m} \Pr[X_{a_i} \geq t'] \\
&\leq \left(\frac{t - \frac{2\alpha}{\epsilon}}{t'}\right) \cdot \Pr\left[\max_i\{X_{a_i}\} \geq t\right] + \frac{7\epsilon}{t'n} \cdot \left(\frac{2\alpha}{\epsilon} \cdot \Pr\left[\max_i\{X_{a_i}\} \geq \frac{2\alpha}{\epsilon}\right]\right) \\
&\leq \frac{1}{t'} \cdot \left[t \cdot \Pr\left[\max_i\{X_{a_i}\} \geq t\right] + \frac{7\epsilon}{n} \cdot \left(\frac{2\alpha}{\epsilon} \cdot \Pr\left[\max_i\{X_{a_i}\} \geq \frac{2\alpha}{\epsilon}\right]\right)\right].
\end{aligned}$$

This follows easily from Markov's inequality, if the expression in the brackets is within a constant factor of $\mathbb{E}[\max_i\{X_{a_i}\}]$. The result is surprising as it is totally possible for that expression to be much smaller than $\mathbb{E}[\max_i\{X_{a_i}\}]$.

In the same spirit as the second point above, the theorem has many interesting implications by setting our t_i's to different values. We make heavy use of the theorem in the following sections.

G.3 Proof of Theorem 13 (the Reduction from Regular to $poly(n)$-Balanced Distributions)

G.3.1 Restricting the Prices for the Input Regular Distributions

Lemma 38. *Let $\mathcal{V} = \{v_i\}_{i\in[n]}$ be a collection of independent regular value distributions, $\epsilon \in (0,1)$, and c the absolute constant in the statement of Theorem 14. For any price vector P, we can construct a new price vector $\hat{P} \in [\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$, such that $\mathcal{R}_{\hat{P}} \geq \mathcal{R}_P - \frac{(c+10)\epsilon\mathcal{R}_{OPT}}{n}$, where $\mathcal{R}_P$ and $\mathcal{R}_{\hat{P}}$ are respectively the expected revenues under price vectors P and $\hat{P}$, and $\mathcal{R}_{OPT}$ is the optimal expected revenue for $\mathcal{V}$.*

Proof. **First step:** We first construct a price vector $P' \in [0, 2n^2\alpha/\epsilon^2]^n$ based on P, such that the revenue under P' is at most an additive $O(\frac{\epsilon\cdot\mathcal{R}_{OPT}}{n})$ smaller than the revenue under P.

We define P' as follows. Let $S = \{i \,\big|\, p_i > 2n^2\alpha/\epsilon^2\}$. For any $i \in S$ set $p'_i = \frac{2(n^2/\epsilon-1)\alpha}{\epsilon}$, while if $i \notin S$ set $p'_i = p_i$. Now assume $|S| = m$. For notational convenience we assume that $S = \{a_i \,|\, i \in [m]\}$, and set $X_{a_i} = v_{a_i}$. Moreover, let $t = \frac{2n^2\alpha}{\epsilon^2}$ and $t_i = p_{a_i}$.

Clearly, the contribution to $\mathcal{R}_P$ from items in S is upper bounded by $\sum_{i=1}^{m} t_i \Pr[X_{a_i} \geq t_i]$. We proceed to analyze the contribution to revenue $\mathcal{R}'_P$ from items in S. Notice that, when $\max_{i\in S}\{v_i\} = \max_i\{X_{a_i}\} \geq t$, the largest value-minus-price gap for items in S is at least $2\alpha/\epsilon$ (given our subtle choice of prices for items in S above). Hence, for the item of S achieving this gap not to be the winner, it must be that some item in $[n] \setminus S$ has a larger value-minus-price gap. For this to happen, the value for this item has to be higher than $2\alpha/\epsilon$. However, the probability that there exists an item in $[n] \setminus S$ with value greater than $2\alpha/\epsilon$ is smaller than $n \cdot \epsilon/n^3 = \epsilon/n^2$ (by Theorem 14). Thus, when $\max_i\{X_{a_i}\} \geq t$, then with probability at least $1 - \epsilon/n^2$, the item in S achieving the largest value-minus-price gap is the item bought by the buyer. So when the price vector is P', the revenue from the items in S is lower bounded by $(t - 2\alpha/\epsilon)\Pr[\max_i\{X_{a_i}\} \geq t](1 - \epsilon/n^2)$ (where we used independence and the fact that $p'_i = t - 2\alpha/\epsilon$ for all $i \in S$.)

Clealry, $(t - 2\alpha/\epsilon)\Pr[\max\{X_{a_i}\} \geq t] \leq t\Pr[\max_i\{X_{a_i}\} \geq t] \leq \mathcal{R}_{OPT}$. To see this, notice that the first inequality is obvious and the second follows from the observation that we could set the prices of all items in S to t and of all other items to $+\infty$ to achieve revenue $t\Pr[\max_i\{X_{a_i}\} \geq t]$. So $\mathcal{R}_{OPT}$

should be larger than this revenue. Similarly, we see that $2\alpha/\epsilon \cdot \Pr[\max_i X_{a_i} \geq 2\alpha/\epsilon] \leq \mathcal{R}_{OPT}$. Using these observations and Theorem 14 we get

$$
\begin{aligned}
&(t - 2\alpha/\epsilon)\Pr[\max_i\{X_{a_i}\} \geq t](1-\epsilon/n^2) + \frac{9\epsilon \cdot \mathcal{R}_{OPT}}{n} \\
\geq &(t - 2\alpha/\epsilon)\Pr[\max_i\{X_{a_i}\} \geq t] + \frac{7\epsilon \cdot (2\alpha/\epsilon \cdot \Pr[\max_i\{X_{a_i}\} \geq 2\alpha/\epsilon])}{n} \\
\geq &\sum_{i=1}^{m} t_i \Pr[X_{a_i} \geq t_i].
\end{aligned}
$$

The above imply that the contribution to $\mathcal{R}_{P'}$ from the items in S is at most an additive $\frac{9\epsilon \cdot \mathcal{R}_{OPT}}{n}$ smaller than the contribution to $\mathcal{R}_P$ from the items in S.

We proceed to compare the contributions from the items in $[n] \setminus S$ to $\mathcal{R}_P$ and $\mathcal{R}_{P'}$. We start with $\mathcal{R}_P$. The contribution from the items in $[n] \setminus S$ is no greater than the total revenue when we ignore the existence of the items in S (e.g. by setting the prices of these items to $+\infty$), since this only boosts the winning probabilities of each item in $[n] \setminus S$.

Under price vector P', $\forall i \in S$, $\Pr[v_i \geq p'_i] \leq \frac{\epsilon}{n^3}$ (Theorem 14). So with probability at least $1 - \frac{\epsilon}{n^2}$, no item in S has a positive value-minus-price gap and the item that has the largest positive gap among the items in $[n] - S$ is the item that is bought by the buyer. Hence, by independence the contribution to $\mathcal{R}_{P'}$ from the items in $[n] - S$ is at least a $1 - \frac{\epsilon}{n^2}$ fraction of the revenue when the items of S are ignored.

By the above discussion, the contribution to $\mathcal{R}_{P'}$ from the items in $[n] - S$ is at most an additive $\frac{\epsilon \mathcal{R}_{OPT}}{n^2}$ smaller than the contribution to $\mathcal{R}_P$ from the items in $[n] - S$.

Putting everything together, we get that $\mathcal{R}_{P'} \geq \mathcal{R}_P - \frac{10\epsilon \mathcal{R}_{OPT}}{n}$.

Second step: To truncate the lower prices, we invoke Lemma 29. This implies that we can set all the prices below $\epsilon\alpha/n^4$ to $\epsilon\alpha/n^4$, only hurting our revenue by an additive $\epsilon\alpha/n^4 \leq \frac{c\epsilon}{n} \cdot \max_z(z \cdot \Pr[\max_i\{X_i\} \geq z]) \leq c\epsilon\mathcal{R}_{OPT}/n$ (where we used Theorem 14 for the first inequality).

Hence, we can define $\hat{P}$ as follows: if $p'_i \leq \epsilon\alpha/n^4$, set $\hat{p}_i = \epsilon\alpha/n^4$, otherwise set $\hat{p}_i = p'_i$. It follows from the above that $\mathcal{R}_{\hat{P}} \geq \mathcal{R}_P - \frac{(c+10)\epsilon\mathcal{R}_{OPT}}{n}$.

□

Thus, we have reduced the problem of finding a near-optimal price vector in $[0, +\infty]^n$ to the problem of finding a near-optimal price vector in the set $[\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$.

G.3.2 Truncating the Support of the Input Regular Distributions

We show that we can truncate the support of the distributions if the price vectors are restricted. Namely

Lemma 39. *Given a collection of independent regular random variables $\mathcal{V} = \{v_i\}_{i \in [n]}$ and any $\epsilon \in (0,1)$, let us define a new collection of random variables $\tilde{\mathcal{V}} = \{\tilde{v}_i\}_{i \in [n]}$ via the following coupling: for all $i \in [n]$, set $\tilde{v}_i = \frac{\epsilon\alpha}{4n^4}$ if $v_i < \frac{\epsilon\alpha}{2n^4}$, set $\tilde{v}_i = 4n^4\alpha/\epsilon^3$, if $v_i \geq 4n^4\alpha/\epsilon^3$, and $\tilde{v}_i = v_i$ otherwise. Also, let c be the absolute constant defined in Theorem 14. For any price vector $P \in [\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$, $|\mathcal{R}_P(\mathcal{V}) - \mathcal{R}_P(\tilde{\mathcal{V}})| \leq \frac{c\epsilon\mathcal{R}_{OPT}(\mathcal{V})}{n}$, where $\mathcal{R}_P(\mathcal{V})$ and $\mathcal{R}_P(\tilde{\mathcal{V}})$ are respectively the revenues of the seller under price vector P when the values of the buyer are $\mathcal{V}$ and $\tilde{\mathcal{V}}$.*

Proof. First, let us define another collection of mutually independent random variables $\hat{\mathcal{V}} = \{\hat{v}_i\}_{i \in [n]}$ via the following coupling: for all $i \in [n]$ set $\hat{v}_i = 4n^4\alpha/\epsilon^3$ if $v_i \geq 4n^4\alpha/\epsilon^3$, and set $\hat{v}_i = v_i$ otherwise.

By Theorem 14, we know that for every i, $\Pr[v_i \geq 4n^4\alpha/\epsilon^3] \leq \frac{\epsilon^3}{2n^7}$. Hence, the probability of the event that there exists an i such that $v_i \geq 4n^4\alpha/\epsilon^3$ is no greater than $n \times \epsilon^3/2n^7 = \epsilon^3/2n^6$. Thus the difference between the contributions of this event to the revenues $\mathcal{R}_P(\mathcal{V})$ and $\mathcal{R}_P(\hat{\mathcal{V}})$ is no greater than $2n^2\alpha/\epsilon^2 \cdot (\epsilon^3/2n^6) = \frac{\epsilon\alpha}{n^4} \leq \frac{c\epsilon}{n} \cdot \max_z(z \cdot \Pr[\max_i\{v_i\} \geq z]) \leq \frac{c\epsilon\mathcal{R}_{OPT}}{n}$, given that the largest price is at most $2n^2\alpha/\epsilon^2$.

Now let us consider the event: $v_i \leq 4n^4\alpha/\epsilon^3$, for all i. In this case $\hat{v}_i = v_i$ for all i. So the contribution of this event to the revenues $\mathcal{R}_P(\mathcal{V})$ and $\mathcal{R}_P(\hat{\mathcal{V}})$ is the same.

Thus, $|\mathcal{R}_P(\mathcal{V}) - \mathcal{R}_P(\hat{\mathcal{V}})| \leq \frac{c\epsilon\mathcal{R}_{OPT}}{n}$.

Now it follows from Lemma 35 that the seller's revenue under any price vector in $[\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$ is the same when the buyer's value distributions are $\hat{\mathcal{V}}$ and $\tilde{\mathcal{V}}$. □

The above lemma shows that we can reduce the problem of finding a near-optimal price vector in $[\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$ for the original value distributions $\mathcal{V}$ to the problem of finding a near-optimal price vector in the set $[\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$ for a collection of value distributions $\tilde{\mathcal{V}}$ supported on the set $[\frac{\epsilon\alpha}{4n^4}, 4n^4\alpha/\epsilon^3]$. Next, we establish that the latter problem can be reduced to finding any (i.e. not necessarily restricted) near-optimal price vector for the distributions $\tilde{\mathcal{V}}$.

Lemma 40. *Given a collection of independent regular random variables $\mathcal{V} = \{v_i\}_{i\in[n]}$ and any $\epsilon \in (0,1)$, let us define a new collection of random variables $\tilde{\mathcal{V}} = \{\tilde{v}_i\}_{i\in[n]}$ via the following coupling: for all $i \in [n]$, set $\tilde{v}_i = \frac{\epsilon\alpha}{4n^4}$ if $v_i < \frac{\epsilon\alpha}{2n^4}$, set $\tilde{v}_i = 4n^4\alpha/\epsilon^3$ if $v_i \geq 4n^4\alpha/\epsilon^3$, and set $\tilde{v}_i = v_i$ otherwise. Let also c be the absolute constant defined in Theorem 14. For any price vector P, we can efficiently construct a new price vector $\hat{P} \in [\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$, such that $\mathcal{R}_{\hat{P}}(\tilde{\mathcal{V}}) \geq \mathcal{R}_P(\tilde{\mathcal{V}}) - \frac{(c+10)\epsilon\cdot\mathcal{R}_{OPT}(\tilde{\mathcal{V}})}{n}$.*

The proof is essentially the same as the proof of Lemma 38 and we skip it. Combining Lemmas 38, 39 and 40 we obtain Theorem 13. The proof is given in the next section.

G.3.3 Finishing the Reduction

Proof of Theorem 13: We start with computing α. This can be done efficiently as specified in the statement of Theorem 14. Now let us define $\tilde{\mathcal{V}}$ via the following coupling: for all $i \in [n]$, set $\tilde{v}_i = \frac{\epsilon\alpha}{4n^4}$ if $v_i < \frac{\epsilon\alpha}{2n^4}$, set $\tilde{v}_i = 4n^4\alpha/\epsilon^3$ if $v_i \geq 4n^4\alpha/\epsilon^3$, and set $\tilde{v}_i = v_i$ otherwise. It is not hard to see that the distributions of the $\tilde{v}_i$'s can be computed in time polynomial in n, $1/\epsilon$ and the description complexity of the distributions of the v_i's, if these are given to us explicitly. If we have oracle access to these distributions, we can construct oracles for the distributions of the $\tilde{v}_i$'s that run in time polynomial in n, $1/\epsilon$ and the desired oracle accuracy.

Now let P be a price vector such that $\mathcal{R}_P(\tilde{\mathcal{V}}) \geq (1 - \epsilon + \frac{(4c+21)\epsilon}{n}) \cdot \mathcal{R}_{OPT}(\tilde{\mathcal{V}})$. It follows from Lemma 40 that we can efficiently construct a price vector $P' \in [\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$, such that

$$\mathcal{R}_{P'}(\tilde{\mathcal{V}}) \geq \left(1 - \epsilon + \frac{(4c+21)\epsilon}{n}\right) \cdot \mathcal{R}_{OPT}(\tilde{\mathcal{V}}) - \frac{(c+10)\epsilon}{n}\mathcal{R}_{OPT}(\tilde{\mathcal{V}}) \geq \left(1 - \epsilon + \frac{(3c+11)\epsilon}{n}\right) \cdot \mathcal{R}_{OPT}(\tilde{\mathcal{V}}).$$

Lemma 38 implies that there exists a price vector $\hat{P} \in [\epsilon\alpha/n^4, 2n^2\alpha/\epsilon^2]^n$, such that $\mathcal{R}_{\hat{P}}(\mathcal{V}) \geq \left(1 - \frac{(c+10)\epsilon}{n}\right) \cdot \mathcal{R}_{OPT}(\mathcal{V})$. By Lemma 39, we know that

$$\mathcal{R}_{OPT}(\tilde{\mathcal{V}}) \geq \mathcal{R}_{\hat{P}}(\tilde{\mathcal{V}}) \geq \mathcal{R}_{\hat{P}}(\mathcal{V}) - \frac{c\epsilon}{n}\mathcal{R}_{OPT}(\mathcal{V}) \geq \left(1 - \frac{(2c+10)\epsilon}{n}\right) \cdot \mathcal{R}_{OPT}(\mathcal{V}).$$

So $\mathcal{R}_{P'}(\tilde{\mathcal{V}}) \geq (1 - \epsilon + \frac{c\epsilon}{n}) \cdot \mathcal{R}_{OPT}(\mathcal{V})$. We can now apply Lemma 39 again, and get

$$\mathcal{R}_{P'}(\mathcal{V}) \geq \mathcal{R}_{P'}(\tilde{\mathcal{V}}) - \frac{c\epsilon}{n}\mathcal{R}_{OPT}(\mathcal{V}) \geq (1 - \epsilon) \cdot \mathcal{R}_{OPT}(\mathcal{V}).$$

□

H From Continuous to Discrete Distributions: the Details

We give the details of our discretization results, culminating in the proof of Theorem 15. Throughout this section we assume that the values $\{v_i\}_{i\in[n]}$ of the buyer lie in a *known finite* (but not necessarily discrete) interval $[u_{min}, u_{max}]$, where $u_{max} = r \cdot u_{min}$ for some $r \geq 1$. Our goal is to establish that the prices, as well as the values can be restricted to discrete sets.

H.1 Discretization in the Price Domain

For starters, it is an easy observation that, when the value distributions are supported in the set $[u_{min}, u_{max}]$, it is sufficient to consider prices that lie in the same range, without any sacrifice in the revenue. This is the point of the next lemma.

Lemma 41 (Price Restriction)**.** *Suppose that the value distributions of the items are supported in $[u_{min}, u_{max}]$ and let P be any price vector. Suppose we modify the price vector P to a new price vector P' as follows: $p'_i = u_{max}$, if $p_i > u_{max}$; $p'_i = u_{min}$, if $p'_i < u_{min}$; and $p'_i = p_i$ otherwise. The expected revenue $\mathcal{R}_P$ and $\mathcal{R}_{P'}$ achieved by the price vectors P and P' respectively satisfies $\mathcal{R}_{P'} \geq \mathcal{R}_P$.*

Proof of Lemma 41: The proof follows easily by combining Lemmas 42 and 43 below.

Lemma 42. *Let P be a price vector and let $S_{over} = \{i : p_i > u_{max}\}$. We define P' to be the following price vector: $p'_i = u_{max}$, for all $i \in S_{over}$, and $p'_i = p_i$ otherwise. We have $\mathcal{R}_{P'} \geq \mathcal{R}_P$.*

Proof. Observe that, for any valuation vector $(v_1, v_2, \ldots, v_n)$ where no item has value larger than or equal to its price in P, the revenue is 0. If that's the case, the revenue under P' can be no worse than 0. So we only consider valuation vectors where some item i is sold under P. Then it must be that $v_i - p_i \geq 0$ and, for all j, $v_i - p_i \geq v_j - p_j$. Under price vector P', there are two possibilities: (1) i is still the item sold; in this case the revenue is the same under P and P'; and (2) i is not the winner anymore; the new winner must be among those items whose price was decreased going from P to P', i.e. some item in $j \in S_{over}$. Observe that $p'_j = u_{max} \geq p_i$ (since for i to be a winner under P, p_i must be no greater than u_{max}, as otherwise $v_i - p_i < 0$). Hence, under P' an item j that is at least as expensive as i is sold. Given that the above is true point-wise, $\mathcal{R}_{P'} \geq \mathcal{R}_P$. □

Lemma 43. *Let P be a price vector and let $S_{below} = \{i : p_i < u_{min}\}$. We define P' to be the following price vector: $p'_i = u_{min}$, for all $i \in S_{below}$, and $p'_i = p_i$ otherwise. We have $\mathcal{R}_{P'} \geq \mathcal{R}_P$.*

Proof. For any valuation vector $(v_1, v_2, \ldots, v_n)$ where no item has value larger than or equal to its price in P, the revenue is 0. If that's the case, the revenue under P' can be no worse than 0. So we only consider valuation vectors where some item i is sold under P. Suppose first that $i \notin S_{below}$. Then i is still the winner under price vector P', since we only increased the prices of items different than i going from P to P'. If the winner $i \in S_{below}$, there are two possibilities: (1) i is still the winner; then the price paid is higher under P', since $p'_i = u_{min} \geq p_i$. (2) i is not the winner anymore; nevertheless, there still needs to be a winner since $v_i \geq u_{min} = p'_i$, and the price paid is at least $u_{min} \geq p_i$. Given that the above hold point-wise, $\mathcal{R}_{P'} \geq \mathcal{R}_P$. □

□

Combining Lemma 41 with a price discretization lemma attributed to Nisan [5], allows us to discretize the set of prices to a set of cardinality $O(\frac{\log r}{\epsilon^2})$, as follows.

Lemma 44 (Price Discretization)**.** *Suppose that the value distributions of the items are supported in $[u_{min}, u_{max}]$. For any $\epsilon \in (0, 1/2)$, consider the following finite set of prices:*

$$\mathcal{P}_\epsilon = \left\{ p \,\middle|\, p = \frac{1 + \epsilon^2 - \epsilon}{(1 - \epsilon^2)^i} \cdot u_{min},\ i \in \left[\left\lfloor \log_{\frac{1}{(1-\epsilon^2)}} (u_{max}/u_{min}) \right\rfloor\right] \right\}.$$

For any price vector $P \in [u_{min}, u_{max}]^n$, we can construct a price vector P' such that $p'_i \in \mathcal{P}_\epsilon$ and $p'_i \in [1-\epsilon, 1+\epsilon^2-\epsilon] \cdot p_i$, for all i. The expected revenue achieved by the two price vectors satisfies $\mathcal{R}_{P'} \geq (1-2\epsilon)\mathcal{R}_P$.

Proof of Lemma 44: Our proof exploits the following Lemma, due to Nisan.

Lemma 45 (Nisan)**.** *For any $\epsilon \in (0,1)$, let P and P' be price vectors that satisfy $p'_i \in [1-\epsilon, 1+\epsilon^2-\epsilon] \cdot p_i$, for all i. Then the expected revenue achieved by the two price vectors satisfies $\mathcal{R}_{P'} \geq (1-2\epsilon)\mathcal{R}_P$.*

Coming to the proof of Lemma 44, for every p_i define

$$p'_i = \frac{1+\epsilon^2-\epsilon}{(1-\epsilon^2)^{\left\lfloor \log_{1/(1-\epsilon^2)}(p_i/u_{min}) \right\rfloor}} \cdot u_{min}.$$

Observe that

$$\frac{1}{(1-\epsilon^2)^{\lfloor \log_{1/(1-\epsilon^2)}(p_i/u_{min}) \rfloor}} \cdot u_{min} \in [1-\epsilon^2, 1] \cdot p_i.$$

On the other hand, $(1-\epsilon^2)(1+\epsilon^2-\epsilon) = 1-\epsilon+\epsilon^3-\epsilon^4 \geq 1-\epsilon$. Thus, $p'_i \in [1-\epsilon, 1+\epsilon^2-\epsilon] \cdot p_i$, for all i. Now Lemma 45 implies that $\mathcal{R}_{P'} \geq (1-2\epsilon)R_P$. □

H.2 "Horizontal" Discretizations in the Value Domain

To enable discretizations in the value domain, we prove an analogue of Nisan's lemma for the value distributions. A straightforward modification of Nisan's approach would result in a discrete support of the value distributions of size linear in $r^2 \log r$, where $r = u_{max}/u_{min}$. With a more intricate argument, we obtain a reduction to a support of size linear in $\log r$. Our discretization result is summarized in Lemma 48 and is obtained via an application of Lemma 46.

Lemma 46. *Let $\{v_i\}_{i\in[n]}$ and $\{\hat{v}_i\}_{i\in[n]}$ be two collections of mutually independent random variables, where all v_i's are supported on a common set $[u_{min}, u_{max}] \subset \mathbb{R}_+$, and let $r = u_{max}/u_{min}$. Let also $\delta \in \left(0, \frac{1}{(4\lceil \log r \rceil)^{1/(2a-1)}}\right]$, where $a \in (1/2, 1)$, and suppose that we can couple the two collections of random variables so that, for all $i \in [n]$, $\hat{v}_i \in [1+\delta-\delta^2, 1+\delta] \cdot v_i$ with probability 1. Finally, let $\mathcal{R}_{OPT}$ be the optimal expected revenue from any pricing when the buyer's values are $\{v_i\}_{i\in[n]}$. Then, for any price vector $P \in [u_{min}, u_{max}]^n$, such that $\mathcal{R}_P(\{v_i\}_i) \geq \mathcal{R}_{OPT}/2$, it holds that*

$$\mathcal{R}_P(\{\hat{v}_i\}_i) \geq (1-3\delta^{1-a})\mathcal{R}_P(\{v_i\}_i),$$

where $\mathcal{R}_P(\{v_i\}_i)$ is the expected revenue of a seller using the price vector P when the values of the buyer are $\{v_i\}_{i\in[n]}$, while $\mathcal{R}_P(\{\hat{v}_i\}_i)$ is the revenue under P when the buyer's values are $\{\hat{v}_i\}_{i\in[n]}$.

Proof of Lemma 46: For notational convenience, throughout this proof we use $\mathcal{R}_P := \mathcal{R}_P(\{v_i\}_i)$ and $\hat{\mathcal{R}}_P := \mathcal{R}_P(\{\hat{v}_i\}_i)$.

Consider now the joint distribution of $\{v_i\}_{i\in[n]}$ and $\{\hat{v}_i\}_{i\in[n]}$ satisfying $\hat{v}_i \in [1+\delta-\delta^2, 1+\delta] \cdot v_i$, for all i, with probability 1. For every point in the support of the joint distribution, we show that the revenue of the seller under price vector P is approximately equal in "Scenario A", where the values of the buyer are $\{v_i\}_{i\in[n]}$, and in "Scenario B", where the values are $\{\hat{v}_i\}_{i\in[n]}$. In particular, we argue first that the winning prices in the two scenarios are within $\delta \cdot u_{max}$ from each other with probability 1. Indeed, for every point in the support of the joint distribution, we distinguish two cases:

1. The items sold are the same in the two scenarios. In this case, the winning prices are also the same.

2. The items sold are different in the two scenarios. In this case, we show that the winning prices are close. Since $\hat{v}_i$ is greater than v_i for all i, if there is a winner in Scenario A, there is a winner in Scenario B. Let i be the winner in Scenario A, and j be the winner in Scenario B. We have the following two inequalities:

$$v_i - p_i \geq v_j - p_j$$
$$\hat{v}_j - p_j \geq \hat{v}_i - p_i$$

The two inequalities imply that

$$\hat{v}_j - v_j \geq \hat{v}_i - v_i.$$

Since $\hat{v}_j \in [1+\delta-\delta^2, 1+\delta] \cdot v_j$, it follows that $\hat{v}_j - v_j \leq \delta \cdot v_j$. Using the same starting condition for i, we can show that $\hat{v}_i - v_i \geq (\delta - \delta^2) \cdot v_i$.

Hence,

$$\delta \cdot v_j \geq (\delta - \delta^2) \cdot v_i.$$

Also we know that

$$p_j \geq p_i + v_j - v_i.$$

Therefore,

$$p_j \geq p_i + v_j - v_i \geq p_i + (1-\delta) \cdot v_i - v_i = p_i - \delta \cdot v_i.$$

The above establishes that with probability 1 the winning prices in the two scenarios are within an *additive* δu_{max} from each other. We proceed to convert this additive approximation guarantee into a multiplicative approximation guarantee. Observe that whenever $p_i \geq \delta^a v_i$, $p_i - \delta \cdot v_i \geq (1-\delta^{1-a})p_i$. Hence, if we can show that most of the revenue $\mathcal{R}_P$ is contributed by value-price pairs (v_i, p_i) satisfying $p_i \geq \delta^a v_i$, we can convert our additive approximation to a $(1-\delta^{1-a})$ multiplicative approximation. Indeed, we argue next that when a price vector P satisfies $\mathcal{R}_P \geq \mathcal{R}_{OPT}/2$, the contribution to the revenue from the event

$$S = \{\textit{the sold item } k \textit{ satisfies } p_k < \delta^a v_k\}$$

is small. More precisely,

Proposition 47. *If $\mathcal{R}_P \geq \mathcal{R}_{OPT}/2$, then the contribution to $\mathcal{R}_P$ from the event S is no greater than $2\delta^{1-a}\mathcal{R}_P$.*

Proof. The proof is by contradiction. For all $i \in [\lceil \log r \rceil]$, define the event

$$S_i = \{(\textit{the sold item } k \textit{ has price } p_k < \delta^a v_k) \wedge (p_k \in [2^{i-1}u_{min}, 2^i u_{min}))\}.$$

Note that S_i and S_j are disjoint for all $i \neq j$. Let $n_p = \lceil \log r \rceil$ and note that $S = \cup_{i=1}^{n_p} S_i$. [6] Assuming that the contribution to $\mathcal{R}_P$ from the event S is larger than $2\delta^{1-a}\mathcal{R}_P$, there must exist some i such that the contribution to $\mathcal{R}_P$ from S_i is at least $2\delta^{1-a}\mathcal{R}_P/n_p \geq \delta^{1-a}\mathcal{R}_{OPT}/n_p$. For this i, let us modify the price vector P to P' in the following fashion:

$$p'_k = \begin{cases} +\infty & p_k \notin [2^{i-1}u_{min}, 2^i u_{min}) \\ \frac{2^{i-1}u_{min}}{\delta^a} & \textit{otherwise} \end{cases}$$

We claim that for all outcomes $(v_1, v_2, \ldots, v_n) \in S_i$, there always exists an item sold under P'. Indeed, let k be the winner under P. Then $p_k < \delta^a v_k$. By the definition of p'_k, we know that

$$p'_k = \frac{2^{i-1}u_{min}}{\delta^a} \leq p_k/\delta^a < v_k.$$

[6] To be more accurate, replace the set $[2^{i-1}u_{min}, 2^i u_{min})$ by $[2^{i-1}u_{min}, 2^i u_{min}]$ for the definition of the event S_{n_p}.

Thus, an item has to be sold. Moreover, the sold item has price $\frac{2^{i-1}u_{min}}{\delta^a}$, as all the other prices are set to $+\infty$. Hence, we can lower bound $\mathcal{R}_{P'}$ as follows

$$\mathcal{R}_{P'} \geq \Pr[S_i] \cdot \frac{2^{i-1}u_{min}}{\delta^a} \geq \frac{\textit{Contribution of } S_i \textit{ to } \mathcal{R}_P}{2\delta^a} \geq \frac{\delta^{1-a}\mathcal{R}_{OPT}}{2n_p\delta^a}.$$

Given that $\delta \leq (\frac{1}{4n_p})^{1/(2a-1)}$, the above implies $\mathcal{R}_{P'} \geq 2\mathcal{R}_{OPT}$, which is impossible, i.e. we get a contradiction. This concludes the proof of the proposition. $\square$

Given the proposition, at least $(1-2\delta^{1-a})$ fraction of $\mathcal{R}_P$ is contributed by value-price pairs (v_i, p_i) satisfying $p_i \geq \delta^a v_i$. Recalling our earlier discussion, this implies that $\hat{\mathcal{R}}_P \geq (1-2\delta^{1-a})(1-\delta^{1-a})\mathcal{R}_P \geq (1-3\delta^{1-a})\mathcal{R}_P$. $\square$

Lemma 46 allows us to discretize the support of the value distributions into a discrete set of bounded cardinality without harming the revenue, as specified by the following lemma.

Lemma 48 (Horizontal Discretization of Values)**.** *Let $\{v_i\}_{i\in[n]}$ be a collection of mutually independent random variables supported on a set $[u_{min}, u_{max}] \subset \mathbb{R}_+$, and let $r = \frac{u_{max}}{u_{min}}$. For any $\delta \in \left(0, \frac{1}{(4\lceil\log r\rceil)^{4/3}}\right)$, there exists another collection of mutually independent random variables $\{\hat{v}_i\}_{i\in[n]}$, which are supported on a discrete set of cardinality $O\left(\frac{\log r}{\delta^2}\right)$ and satisfy the following properties.*

1. *The optimal revenue when the buyer's values are $\{\hat{v}_i\}_{i\in[n]}$ is at least a $(1-3\delta^{1/8})$-fraction of the optimal revenue when the values are $\{v_i\}_{i\in[n]}$. I.e. $\hat{\mathcal{R}}_{OPT} \geq (1-3\delta^{1/8})\mathcal{R}_{OPT}$, where $\mathcal{R}_{OPT} = \max_P \mathcal{R}_P(\{v_i\}_i)$ and $\hat{\mathcal{R}}_{OPT} = \max_P \mathcal{R}_P(\{\hat{v}_i\}_i)$.*
2. *Moreover, for any constant $\rho \in (0, 1/2)$ and any price vector P such that $\mathcal{R}_P(\{\hat{v}_i\}_i) \geq (1-\rho)\hat{\mathcal{R}}_{OPT}$, we can construct in time polynomial in the description of P, $1/\delta$ and $\log u_{min}$ another price vector $\tilde{P}$ such that $\mathcal{R}_{\tilde{P}}(\{v_i\}_i) \geq (1-7\delta^{1/8}-\rho)\mathcal{R}_{OPT}$.*

Moreover, assuming that the set $[u_{min}, u_{max}]$ is specified in the input,[7] we can compute the support of the distributions of the variables $\{\hat{v}_i\}_i$ in time polynomial in $\log u_{min}$, $\log u_{max}$ and $1/\delta$. We can also compute the distributions of the variables $\{\hat{v}_i\}_{i\in[n]}$ in time polynomial in the size of the input and $1/\epsilon$, if we have the distributions of the variables $\{v_i\}_{i\in[n]}$ explicitly. If we have oracle access to the distributions of the variables $\{v_i\}_{i\in[n]}$, we can construct an oracle for the distributions of the variables $\{\hat{v}_i\}_{i\in[n]}$, running in time polynomial in $\log u_{min}$, $\log u_{\max}$, $1/\delta$, the input to the oracle and the desired precision.

Proof of Lemma 48: We begin with the description of the random variables $\{\hat{v}_i\}_{i\in[n]}$. We will use $\{F_i\}_{i\in[n]}$ and $\{\hat{F}_i\}_{i\in[n]}$ to denote respectively the cumulative distribution functions of the variables $\{v_i\}_{i\in[n]}$ and $\{\hat{v}_i\}_{i\in[n]}$. Taking $\xi = \frac{\delta^2}{1+\delta-\delta^2}$, our new variables $\{\hat{v}_i\}_{i\in[n]}$ will only be supported on the set

$$\left\{a_j = (1+\delta)(1+\xi)^j u_{min} \;\middle|\; j \in \left\{0, \ldots, \lfloor \log_{1+\xi} \frac{u_{max}}{u_{min}} \rfloor\right\}\right\}.$$

Moreover, for all i, the probability mass that $\hat{F}_i$ assigns to every point in its support is:

$$\hat{F}_i(a_j) = F_i(a_j/(1+\delta-\delta^2)) - F_i(a_j/(1+\delta)).$$

Now, for all i, we couple v_i with $\hat{v}_i$ as follows: If $v_i \in [a_j/(1+\delta), a_j/(1+\delta-\delta^2))$, we set $\hat{v}_i = a_j$. Given our definition of the $\hat{F}_i$'s, this defines a valid coupling of the collections $\mathcal{V} = \{v_i\}_i$ and $\hat{\mathcal{V}} = \{\hat{v}_i\}_i$.

[7] The requirement that the set $[u_{min}, u_{max}]$ is specified as part of the input is only relevant if we have oracle access to the distributions of the v_i's, as if we have them explicitly we can easily find $[u_{min}, u_{max}]$.

Moreover, by definition, our coupling satisfies $\hat{v}_i \in [1+\delta-\delta^2, 1+\delta] \cdot v_i$, for all i, with probability 1, and all the $\hat{v}_i$'s are supported on $[(1+\delta)u_{min}, (1+\delta)u_{max}]$.

We are now ready to establish the first part of the lemma. Using Lemma 46 and the property of our coupling it follows immediately that

$$\mathcal{R}_P(\hat{\mathcal{V}}) \geq (1 - 3\delta^{1/8})\mathcal{R}_P(\mathcal{V}),$$

for any price vector $P \in [u_{min}, u_{max}]^n$ s.t. $\mathcal{R}_P(\mathcal{V}) \geq \frac{1}{2}\mathcal{R}_{OPT}$. Lemma 41 implies that the optimal revenue for $\mathcal{V}$ is achieved by some price vector in $[u_{min}, u_{max}]^n$. Hence, we get from the above that $\hat{\mathcal{R}}_{OPT} \geq (1 - 3\delta^{1/8})\mathcal{R}_{OPT}$.

We proceed to show the second part of the lemma. We do this by defining another collection of random variables $\tilde{\mathcal{V}} = \{\tilde{v}_i\}_{i\in[n]}$. These are defined implicitly via the following coupling between $\{\tilde{v}_i\}_{i\in[n]}$ and $\{\hat{v}_i\}_{i\in[n]}$: for all i, we set

$$\tilde{v}_i = \frac{\hat{v}_i}{(1+\delta-\delta^2)(1+\delta)}.$$

It follows that the $\tilde{v}_i$'s are supported on $[u_{min}/(1+\delta-\delta^2), u_{max}/(1+\delta-\delta^2)]$.

Moreover, for any price vector P, let us construct another price vector $\tilde{P}$ as follows:

$$\tilde{p}_i = \frac{p_i}{(1+\delta-\delta^2)(1+\delta)}. \tag{7}$$

Under our coupling between $\{\tilde{v}_i\}_{i\in[n]}$ and $\{\hat{v}_i\}_{i\in[n]}$, it is not hard to see that if we use price vector P when the buyer's values are $\{\hat{v}_i\}_{i\in[n]}$ and price vector $\tilde{P}$ when the buyer's values are $\{\tilde{v}_i\}_{i\in[n]}$, then the index of the item that the buyer buys is the same in the two cases, with probability 1. Hence:

$$\mathcal{R}_{\tilde{P}}(\tilde{\mathcal{V}}) = \frac{\mathcal{R}_P(\hat{\mathcal{V}})}{(1+\delta-\delta^2)(1+\delta)}. \tag{8}$$

This follows from the fact that both $\tilde{P}$ and $\{\tilde{v}_i\}_{i\in[n]}$ are the same linear transformations of P and $\{\hat{v}_i\}_{i\in[n]}$ respectively.

Composing the coupling between v_i and $\hat{v}_i$ and the coupling between $\hat{v}_i$ with $\tilde{v}_i$, we obtain a coupling between v_i and $\tilde{v}_i$. We show that this coupling satisfies $v_i \in [1+\delta-\delta^2, 1+\delta] \cdot \tilde{v}_i$, with probability 1. Since $(1+\delta-\delta^2)v_i \leq \hat{v}_i \leq (1+\delta)v_i$,

$$v_i/(1+\delta) \leq \hat{v}_i/(1+\delta-\delta^2)(1+\delta) = \tilde{v}_i \leq v_i/(1+\delta-\delta^2).$$

$$v_i \in [1+\delta-\delta^2, 1+\delta] \cdot \tilde{v}_i.$$

Given that $v_i \in [1+\delta-\delta^2, 1+\delta] \cdot \tilde{v}_i$ with probability 1, an application of Lemma 46 implies that, for any price vector $\tilde{P} \in [u_{min}/(1+\delta-\delta^2), u_{max}/(1+\delta-\delta^2)]^n$ satisfying $\mathcal{R}_{\tilde{P}}(\tilde{V}) \geq \frac{1}{2}\mathcal{R}_{OPT}(\tilde{V})$:

$$\mathcal{R}_{\tilde{P}}(\mathcal{V}) \geq (1 - 3\delta^{1/8})\mathcal{R}_{\tilde{P}}(\tilde{\mathcal{V}}). \tag{9}$$

Now let P be a price vector satisfying $\mathcal{R}_P(\hat{\mathcal{V}}) \geq (1-\rho)\hat{\mathcal{R}}_{OPT}$. Lemma 41 implies that WLOG we can assume that $P \in [(1+\delta)u_{min}, (1+\delta)u_{max}]^n$ (as if the given price vector is not in this set, we can efficiently convert it into one that is in this set without losing any revenue.) Then the vector $\tilde{P}$ obtained from P via Eq. (7) is in $[u_{min}/(1+\delta-\delta^2), u_{max}/(1+\delta-\delta^2)]^n$, and clearly satisfies

$\mathcal{R}_{\tilde{P}}(\tilde{\mathcal{V}}) \geq (1-\rho)\mathcal{R}_{OPT}(\tilde{\mathcal{V}})$, as $\tilde{P}$ and $\tilde{\mathcal{V}}$ are the same linear transformations of P and $\hat{\mathcal{V}}$ respectively. Hence, Equations (8) and (9) give

$$\begin{aligned}
\mathcal{R}_{\tilde{P}}(\mathcal{V}) &\geq \left((1-3\delta^{1/8})\Big/(1+\delta)(1+\delta-\delta^2)\right)\mathcal{R}_P(\hat{\mathcal{V}}) \\
&\geq (1-3\delta^{1/8})(1-2\delta)\mathcal{R}_P(\hat{\mathcal{V}}) \\
&\geq (1-4\delta^{1/8})\mathcal{R}_P(\hat{\mathcal{V}}) \\
&\geq (1-4\delta^{1/8})(1-\rho)\hat{\mathcal{R}}_{OPT} \\
&\geq (1-4\delta^{1/8})(1-\rho)(1-3\delta^{1/8})\mathcal{R}_{OPT} \qquad \text{(using the first part of the theorem)} \\
&\geq (1-7\delta^{1/8}-\rho)\mathcal{R}_{OPT}.
\end{aligned}$$

□

H.3 Proof of Theorem 15

Proof of Theorem 15: Lemma 48 implies that we can reduce the problem $\textsc{Price}(\mathcal{V}, \epsilon)$ to the problem $\textsc{Price}(\hat{\mathcal{V}}, (\epsilon/8)^8)$, where $\hat{\mathcal{V}} = \{\hat{v}_i\}_i$ is a collection of mutually independent random variables supported on a common discrete set $\mathcal{S} = \{s^{(1)}, \ldots, s^{(k_1)}\} \subset [(1+(\epsilon/8)^8)u_{min}, (1+(\epsilon/8)^8)u_{max}]$ of cardinality $k_1 = O(\frac{\log r}{\epsilon^{16}})$. Now, Lemmas 41 and 44 imply that we can reduce the problem $\textsc{Price}(\hat{\mathcal{V}}, (\epsilon/8)^8)$ to the problem $\textsc{ResrtictedPrice}(\hat{\mathcal{V}}, \mathcal{P}, 0.5(\epsilon/8)^8)$, where $\mathcal{P}$ is a discrete set of prices of cardinality $O(\frac{\log r}{\epsilon^{16}})$. We omit the analysis of the running time of the reduction, as it is straightforward. □

H.4 "Vertical" Discretizations in the Value Domain

Lemma 48 allows us to reduce our problem into one where the support of the value distributions is a bounded set of cardinality $O(\frac{\log r}{\delta^2})$, where $r = u_{max}/u_{min}$ without worsening the revenue by more than a fraction of $\delta^{1/8}$. The following lemma allows us to also round the probabilities assigned by the value distributions to every point in their support. But, this result is weaker as we require the probabilities to scale polynomially in both the number of distributions n and r. Still, the lemma is handy in the design of our algorithms as it helps keep our calculations in finite precision.

Lemma 49 (Vertical Discretization of Values)**.** *Let $\{v'_i\}_{i\in[n]}$ be a set of mutually independent random variables supported on a discrete set $S = \{s_1, \ldots, s_m\}$ and let $[u_{min}, u_{max}] \subset \mathbb{R}$ and $r = \frac{u_{max}}{u_{min}}$. It is assumed that S as well as u_{min}, u_{max} are specified as part of the input. Then we can construct another collection of mutually independent random variables $\{v''_i\}_{i\in[n]}$, which are supported on the same set S and are such that the probability that v''_i equals any value in S is an integer multiple of $\frac{1}{(rn)^3}$ for all i. Moreover, for any price vector $P \in [u_{min}, u_{max}]^n$, if $\mathcal{R}'_P$ is the seller's expected revenue under price vector P when the buyer's values are $\{v'_i\}_{i\in[n]}$, and $\mathcal{R}''_P$ is the revenue under price vector P when the buyer's values are $\{v''_i\}_{i\in[n]}$, then*

$$|\mathcal{R}'_P - \mathcal{R}''_P| \leq \frac{4m}{rn^2} \cdot u_{min}.$$

Moreover, the distributions of the variables v''_i are constructed explicitly, and the running time of the construction is polynomial in n, m, $\log r$, the description complexity of S, and the description complexity of the distributions of the v'_i's, if these are given explicitly in the input.

Proof of Lemma 49: If we know the distributions explicitly, then, for all i, we construct the distribution of v''_i based on the distribution of v'_i as follows. Let $\pi'_{s_j} = \Pr[v'_i = s_j]$, and

$$\delta'_{s_j} = \pi'_{s_j} - \frac{1}{(rn)^3} \cdot \lfloor \pi'_{s_j} \cdot (rn)^3 \rfloor.$$

Denoting $\Pr[v''_i = s_j]$ by π''_{s_j} (for convenience), we set $\pi''_{s_j} := \pi'_{s_j} - \delta'_{s_j}$, for $j \geq 2$, and $\pi''_{s_1} := 1 - \sum_{j\geq 2} \pi''_{s_j}$.

It is not hard to see now that we can couple the two sets of random variables in the following way. For every i: if $v'_i = s_1$, we let $v''_i = s_1$; if $v'_i = s_j$ for $j \geq 2$, we let $v''_i = s_j$ with probability $1 - \frac{\delta'_{s_j}}{\pi'_{s_j}}$, and let $v''_i = s_1$ with probability $\frac{\delta'_{s_j}}{\pi'_{s_j}}$. Notice that this coupling satisfies:

$$\Pr[v'_i \neq v''_i] = \sum_{j=2}^{m} \pi'_{s_j} \cdot \frac{\delta'_{s_j}}{\pi'_{s_j}} = \sum_{j=2}^{m} \delta'_{s_j} \leq \frac{m}{(rn)^3}.$$

Now taking a union bound over all i, the probability that the vector $v' = (v'_1, v'_2, \ldots, v'_n)$ is different from $v'' = (v''_1, v''_2, \ldots, v''_n)$ is at most $\frac{m}{(rn)^2}$. In other words, with probability at least $1 - \frac{m}{(rn)^2}$, $v' = v''$. Clearly, for all draws from the distribution such that $v' = v''$, the revenues are the same. When $v' \neq v''$, the difference between the revenues is at most u_{max}, since $P \in [u_{min}, u_{max}]^n$. And this only happens with probability at most $\frac{m}{(rn)^2}$. Therefore, the difference between the expected revenues under the two distributions should be no greater than $\frac{m}{(rn)^2} \cdot u_{max} \leq \frac{m}{rn^2} \cdot u_{min}$, i.e. $|\mathcal{R}'_P - \mathcal{R}''_P| \leq \frac{m}{rn^2} \cdot u_{min}$.

Clearly, we can compute the distributions of the v''_i's in time polynomial in n, m, $\log r$ and the description complexity of the distributions of the variables v'_i's, if these distributions are given to us explicitly. If we have oracle access to the distributions of the v'_i's we can query our oracle with high enough precision, say $1/(rn)^3$, to obtain a function $g_i : S \to [0, 1]$ that satisfies $\sum_{x\in S} g_i(x) = 1 \pm \frac{m}{(rn)^3}$. Using g_i as a proxy for the distribution of v'_i we can follow the algorithm outlined above to define the distribution of v''_i. It is not hard to argue that the total variation distance between v'_i and v''_i can be bounded by $\frac{4m}{(rn)^3}$. Hence, we can couple v'_i and v''_i so that

$$\Pr[v'_i \neq v''_i] \leq \frac{4m}{(rn)^3}$$

and proceed as above. □

I The Algorithm for Discrete Distributions

I.1 The DP Step: Add an Item and Prune the Table

In Section 7, we described what our intended meaning of the Boolean function $g(i, \widehat{\Pr})$. Here we discuss how to compute g via Dynamic Programming. Our dynamic program works bottom-up (i.e. from smaller to larger i's), filling in g's table so that the following recursive conditions are met.

- If $i > 1$, we set $g(i, \widehat{\Pr}) = 1$ iff there is a price $p^{(j)}$ and a distribution $\widehat{\Pr}'$ so that the following hold:
  1. $g\left(i-1, \widehat{\Pr}'\right) = 1$.
  2. Suppose that under some pricing the (winning-value, winning-price) distribution for the prefix $1 \ldots i-1$ of the items is $\widehat{\Pr}'$, and we assign price $p^{(j)}$ to the i-th item. We can compute the resulting (winning-value, winning-price) distribution $\{\widehat{\Pr}''_{i_1,i_2}\}_{i_1\in[k_1],\ i_2\in[k_2]}$ for the prefix $1 \ldots i$

from just $\widehat{\Pr}'_{i_1,i_2}$ and the distribution $\hat{F}_i$. Indeed:

$$\widehat{\Pr}''_{i_1,i_2} = \widehat{\Pr}'_{i_1,i_2} \cdot \Pr_{v_i \sim \hat{F}_i}[v_i - p^{(j)} < v^{(i_1)} - p^{(i_2)}] + \left(\sum_{\substack{j_1,j_2 \\ s.t.\ v^{(j_1)} - p^{(j_2)} \\ \leq v^{(i_1)} - p^{(i_2)}}} \widehat{\Pr}'_{j_1,j_2} \right) \cdot \Pr_{v_i \sim \hat{F}_i}[v_i = v^{(i_1)}] \cdot \mathbb{1}_{p^{(j)} = p^{(i_2)}}.$$

We require that $\widehat{\Pr}$ is a rounded version of $\widehat{\Pr}''$, where all the probabilities are integer multiples of $\frac{1}{m^3}$. The rounding should be of the following canonical form. Setting $\delta_{i_1,i_2} = \widehat{\Pr}''_{i_1,i_2} - \left\lfloor \frac{\widehat{\Pr}''_{i_1,i_2}}{1/m^3} \right\rfloor \cdot \frac{1}{m^3}$, and $l = \left(\sum_{i_1 \in [k_1],\ i_2 \in [k_2]} \delta_{i_1,i_2} \right) \Big/ \left(\frac{1}{m^3}\right)$, we require that the first l probabilities in $\{\widehat{\Pr}''_{i_1,i_2}\}_{i_1 \in [k_1], [i_2] \in k_2}$ in some fixed lexicographic order are rounded up to the closest multiple of $\frac{1}{m^3}$, and the rest are rounded down to the closest multiple of $\frac{1}{m^3}$.

(If Conditions 1 and 2 are met, we also keep a pointer from cell $g(i, \widehat{\Pr})$ to cell $g(i-1, \widehat{\Pr}')$ of the dynamic program recording the price $p^{(j_2)}$ on that pointer. We use these pointers later to recover price vectors consistent with a certain distribution.)

• To fill in the first slice of the table corresponding to $i = 1$, we use the same recursive definition given above, imagining that there is a slice $i = 0$, whose cells are all 0 except for those corresponding to the distributions $\widehat{\Pr}$ that satisfy: $\widehat{\Pr}_{i_1,i_2} = 0$, for all i_1, i_2, except for the lexicographically smallest $(i_1^*, i_2^*) \in \arg\min_{(k_1,k_2)}\ v^{(k_1)} - p^{(k_2)}$, where $\widehat{\Pr}_{i_1^*,i_2^*} = 1$.

While decribed the function g recursively above, we compute it iteratively from $i = 1$ through n.

I.2 Lower Bounding the Revenue

In this section, we justify the correctness of the algorithm presented in Section 7, providing a proof of Lemma 16. Intuitively, if we did not perform any rounding of distributions, our algorithm would have been *exact*, outputting the optimal price vector in $\{p^{(1)}, \ldots, p^{(k_2)}\}^n$. We show next that the rounding is fine enough that it does not become detrimental to our revenue. To show this, we use the probabilistic concepts of *total variation distance* and *coupling of random variables*. Recall that the total variation distance between two distributions $\mathbb{P}$ and $\mathbb{Q}$ over a finite set $\mathcal{A}$ is defined as follows

$$||\mathbb{P} - \mathbb{Q}||_{TV} = \frac{1}{2} \sum_{\alpha \in \mathcal{A}} |\mathbb{P}(\alpha) - \mathbb{Q}(\alpha)|.$$

Similarly, if X and Y are two random variables ranging over a finite set, their total variation distance, denoted $||X - Y||_{TV}$ is defined as the total variation distance between their distributions.

Proceeding to the correctness of our algorithm, let $P = (p_1, p_2, \cdots, p_n) \in \{p^{(1)}, \ldots, p^{(k_2)}\}^n$ be an arbitrary price vector. We can use this price vector to select n cells of our dynamic programming table, picking one cell per layer. The cells are those that the algorithm would have traversed if it made the decision of assigning price p_i to item i. Let us call the resulting cells $cell_1, cell_2, \ldots, cell_n$.

For all i, we intend to compare the distributions $\left\{\Pr^{(i)}_{i_1,i_2}\right\}_{i_1 \in [k_1],\ i_2 \in [k_2]}$ and $\left\{\widehat{\Pr}^{(i)}_{i_1,i_2}\right\}_{i_1 \in [k_1],\ i_2 \in [k_2]}$, which are respectively the (winning-value,winning-price) distribution:

- arising when the prefix $1 \ldots i$ of items with distributions $\{\hat{F}_j\}_{j=1,\ldots,i}$ is priced according to price vector $(p_1, \ldots, p_i)$;

- stored in $cell_i$ by the algorithm.

The following lemma shows that these distributions have small total variation distance.

Lemma 50. *Let $\{X_i\}_{i\in[n]}$ and $\{\hat{X}_i\}_{i\in[n]}$ be two collections of (k_1k_2)-dimensional random unit vectors defined in terms of $\left\{\Pr^{(i)}\right\}_i$ and $\left\{\widehat{\Pr}^{(i)}\right\}_i$ as follows: for all $i \in [n]$, $i_1 \in [k_1]$, $i_2 \in [k_2]$, and $\ell = (i_1 - 1)\cdot k_2 + i_2 \in [k_1k_2]$, we set $\Pr[X_i = e_\ell] = \Pr^{(i)}_{i_1,i_2}$ and $\Pr[\hat{X}_i = e_\ell] = \widehat{\Pr}^{(i)}_{i_1,i_2}$, where e_ℓ is the unit vector along dimension ℓ.*

Then, for all i,

$$||X_i - \hat{X}_i||_{TV} \le nk_1k_2/m^3.$$

Proof. We prove this by induction. For the base case ($i = 1$), observe that $||X_1 - \hat{X}_1||_{TV} \le k_1k_2/m^3$, because $\widehat{\Pr}^{(1)}$ is just a rounding of $\Pr^{(1)}$ into probabilities that are multiples of $\frac{1}{m^3}$, whereby the probability of every point in the support is not modified by more than an additive $\frac{1}{m^3}$.

For the inductive step, it suffices to argue that for all $i \in [n-1]$,

$$||X_{i+1} - \hat{X}_{i+1}||_{TV} - ||X_i - \hat{X}_i||_{TV} \le k_1k_2/m^3.$$

To show this, we are going to consider two auxiliary random variables Y_{i+1} and Z_{i+1}:

- Y_{i+1} is a (k_1k_2)-dimensional random unit vector defined as follows: for $i_2^* \in [k_2]$ such that $p^{(i_2^*)} = p_{i+1}$, if $\ell = (i_1-1)\cdot k_2 + i_2^*$, then $\Pr[Y_{i+1} = e_\ell] = \hat{F}_{i+1}(v^{(i_1)})$, otherwise $\Pr[Y_{i+1} = e_\ell] = 0$. [8]
- Z_{i+1} is a (k_1k_2)-dimensional random unit vector defined as follows

$$\Pr[Z_{i+1} = e_{(i_1-1)\cdot k_2+i_2}] = \sum_{\substack{j_1,j_2 \\ s.t.\ v^{(j_1)}-p^{(j_2)} \\ \le v^{(i_1)}-p^{(i_2)}}} \Pr[\hat{X}_i + Y_{i+1} = e_{(i_1-1)\cdot k_2+i_2} + e_{(j_1-1)\cdot k_2+j_2}], \tag{10}$$

where for the purposes of the above definition $\hat{X}_i$ and Y_{i+1} are taken to be independent. [9]

We claim that

$$||X_{i+1} - Z_{i+1}||_{TV} \le ||(X_i + Y_{i+1}) - (\hat{X}_i + Y_{i+1})||_{TV}. \tag{11}$$

Indeed, we can define X_{i+1} as follows: for all i_1, i_2,

$$\Pr[X_{i+1} = e_{(i_1-1)\cdot k_2+i_2}] = \sum_{\substack{j_1,j_2 \\ s.t.\ v^{(j_1)}-p^{(j_2)} \\ \le v^{(i_1)}-p^{(i_2)}}} \Pr[X_i + Y_{i+1} = e_{(i_1-1)\cdot k_2+i_2} + e_{(j_1-1)\cdot k_2+j_2}], \tag{12}$$

where for the purposes of the above definition X_i and Y_{i+1} are taken to be independent. Now (10), (12) and the definition of the total variation distance imply (11).

To finish our proof, suppose further that we couple X_i and $\hat{X}_i$ optimally. By the *optimal coupling theorem* our joint distribution satisfies $\Pr[X_i \ne \hat{X}_i] = ||X_i - \hat{X}_i||_{TV}$. Defining Y_{i+1} in the same space as X_i and $\hat{X}_i$ so that Y_{i+1} is independent from both X_i and $\hat{X}_i$, the *coupling lemma* implies:

$$\begin{aligned}
||(X_i + Y_{i+1}) - (\hat{X}_i + Y_{i+1})||_{TV} &\le \Pr[(X_i + Y_{i+1}) \ne (\hat{X}_i + Y_{i+1})] \\
&\le \Pr[X_i \ne \hat{X}_i] \\
&= ||X_i - \hat{X}_i||_{TV} \qquad \text{(since } X_i \text{ and } \hat{X}_i \text{ are optimally coupled)}
\end{aligned}$$

On the other hand, it is easy to see that $||Z_{i+1} - \hat{X}_{i+1}||_{TV} \le k_1k_2/m^3$. Indeed:

[8] In other words, Y_{i+1} is the random unit vector analog of $\hat{F}_{i+1}$.

[9] Z_{i+1} is the random unit vector analog of the (winning-value,winning-price) distribution for the prefix $1\ldots i+1$, if item $i+1$ is assigned price p_{i+1} and the (winning-value,winning-price) distribution for the prefix $1\ldots i$ is $\widehat{\Pr}_i$.

- Z_{i+1} is the random unit vector analog of the (winning-value,winning-price) distribution for the prefix $1 \ldots i+1$ of items, if item $i+1$ is assigned price p_{i+1} and the (winning-value,winning-price) distribution for the prefix $1 \ldots i$ is $\widehat{\Pr}_i$;
- $\hat{X}_{i+1}$ is the random unit vector analog of the same (winning-value,winning-price) distribution as above, except after rounding that distribution according to the rounding rule used in our dynamic program;
- the rounding changes every probability in the support by at most an additive $1/m^3$.

Combining the above and using the triangle inequality, we obtain

$$||X_{i+1} - \hat{X}_{i+1}||_{TV} \leq ||Z_{i+1} - \hat{X}_{i+1}||_{TV} + ||X_{i+1} - Z_{i+1}||_{TV} \leq k_1 k_2/m^3 + ||X_i - \hat{X}_i||_{TV}.$$

□

Proof of Lemma 16: Let P^* be the optimal price vector for the instance of the pricing problem obtained after the application of Lemma 49, and let *cell** be the cell at layer n of the DP corresponding to the price vector P^*. Lemma 50 implies that

$$\sum_{i_1 \in [k_1],\ i_2 \in [k_2]} |\Pr^{(n)}_{i_1,i_2} - \widehat{\Pr}^{(n)}_{i_1,i_2}| \leq n k_1 k_2 / m^3,$$

where $\Pr^{(n)}$ is the true (winning-value,winning-price) distribution corresponding to price vector P^* and $\widehat{\Pr}^{(n)}$ is the distribution stored in cell *cell**. Clearly, the expected revenues $\mathcal{R}_{P^*}$ and $\mathcal{R}_{cell^*}$ of these two distributions are related, as follows

$$|\mathcal{R}_{P^*} - \mathcal{R}_{cell^*}| \leq \sum_{i_1 \in [k_1],\ i_2 \in [k_2]} |\Pr^{(n)}_{i_1,i_2} - \widehat{\Pr}^{(n)}_{i_1,i_2}| \cdot p^{(i_2)} \leq \frac{n k_1 k_2}{m^3} \cdot \max\{\mathcal{P}\} \leq \frac{k_1 k_2}{m^2} \cdot \min\{\mathcal{P}\}.$$

Now let *cell*$'$ be the cell at layer n of the dynamic programming table that has the highest revenue, and let P' be the price vector reconstructed from *cell*$'$ by tracing back-pointers. Using the same notation as above, call $\mathcal{R}_{cell'}$ the revenue from the distribution stored at *cell*$'$ and $\mathcal{R}_{P'}$ the revenue from price vector P'. Then we have the following:

$$\mathcal{R}_{cell'} \geq \mathcal{R}_{cell^*}; \qquad \text{(by the optimality of } cell'\text{)} \tag{13}$$

$$|\mathcal{R}_{P'} - \mathcal{R}_{cell'}| \leq \frac{k_1 k_2}{m^2} \cdot \min\{\mathcal{P}\}. \qquad \text{(using Lemma 50, as we did above)} \tag{14}$$

Putting all the above together, we obtain that

$$\mathcal{R}_{P'} \geq \mathcal{R}_{P^*} - \frac{2 k_1 k_2}{m^2} \cdot \min\{\mathcal{P}\}. \tag{15}$$

Hence, the price vector P' output by the dynamic program achieves revenue $\mathcal{R}_{P'}$ that is close to the optimal in the instance of the pricing problem obtained after the first step of our algorithm, i.e. the instance obtained after invoking the reduction of Lemma 49. We still have to relate the revenue that P' achieves in the original instance (i.e. before the reduction of Lemma 49) to the optimal revenue OPT of that instance. For this, let us define the quantities:

- $\mathcal{R}(P^*)$: the revenue achieved by price vector P^* in the original instance;
- $\mathcal{R}(P')$: the revenue achieved by price vector P' in the original instance.

Using Lemma 49 we can easily see that

$$\mathcal{R}(P^*) \geq OPT - \frac{8m}{rn^2} \cdot \min\{\mathcal{P}\}.$$

Moreover,

$$\mathcal{R}(P') \geq \mathcal{R}_{P'} - \frac{4m}{rn^2} \cdot \min\{\mathcal{P}\}; \text{and}$$

$$\mathcal{R}_{P^*} \geq \mathcal{R}(P^*) - \frac{4m}{rn^2} \cdot \min\{\mathcal{P}\}.$$

Combining these with (15), we get

$$\mathcal{R}(P') \geq \mathcal{R}_{P^*} - \frac{2k_1k_2}{m^2} \cdot \min\{\mathcal{P}\} - \frac{4m}{rn^2} \cdot \min\{\mathcal{P}\} \tag{16}$$

$$\geq \mathcal{R}(P^*) - \frac{2k_1k_2}{m^2} \cdot \min\{\mathcal{P}\} - \frac{8m}{rn^2} \cdot \min\{\mathcal{P}\} \tag{17}$$

$$\geq OPT - \frac{2k_1k_2 + 16mr}{m^2} \cdot \min\{\mathcal{P}\}. \tag{18}$$

□

I.3 Analyzing the Running Time

Proof of Lemma 17: Recall that both the support $\mathcal{S} = \{v^{(1)}, v^{(2)}, \ldots, v^{(k_1)}\}$ of the value distributions and the set $\mathcal{P} := \left\{p^{(1)}, \ldots, p^{(k_2)}\right\}$ of prices are explicitly part of the input to our algorithm.

Given this, the reduction of Lemma 49 (used as the first step of our algorithm) takes time polynomial in the size of the input. After this reduction is carried out, the value distributions $\{\hat{F}_i\}_i$ that are provided as input to the dynamic program are known explicitly and the probabilities they assign to every point in $\mathcal{S}$ are integer multiples of $\frac{1}{m^3}$, where $m = nr$, $r = \max p^{(j)}/p^{(i)}$.

We proceed to bound the run-time of the Dynamic Program. First, it is easy to see that its table has at most $n \times (m^3+1)^{k_1k_2}$ cells, since there are n possible choices for i and m^3+1 possible values for each $\Pr_{i_1,i_2}$. Our DP computation proceeds iteratively from layer $i = 1$ to layer $i = n$ of the table. For every cell of layer i, there are at most k_2 different prices we can assign to the next item $i+1$, and for every such price we need to compute a distribution and round that distribution. Hence, the total work we need to do per cell of layer i is polynomial in the input size, since our computation involves probabilities that are integer multiples of $\frac{1}{m^3}$. Indeed the probability distributions maintained by our dynamic program use probabilities that are integer multiples of $\frac{1}{m^3}$, and recall the distributions $\hat{F}_i$ also use probabilities in multiples of $\frac{1}{m^3}$. Hence, the total time we need to spend to fill up the whole table is $m^{O(k_1k_2)}$. In the last phase of the algorithm, we exhaustively search for the cell of layer n with the highest expected revenue. This costs time polynomial in the size of the input and $m^{O(k_1k_2)}$, since there are $m^{O(k_1k_2)}$ cells at layer n, and the expected revenue computation for each cell can be done in time polynomial in the input size. Once we find the cell with the highest expected revenue, we can obtain a corresponding price vector in time $O(n)$, since we just need to follow back-pointers from layer n to layer 1.

Overall, the running time of the algorithm is polynomial in the size of the input and $m^{O(k_1k_2)}$. □

J Summary of Algorithmic Results

In this section, we prove our main algorithmic results (Theorems 1, 2, 3, 4, and 8). We only mildly try to optimize the constants in our running times. We should be able to improve them with a more careful analysis. We start with the proof of Theorem 3, which we repeat here.

Theorem 51 (Finite Support). *Let $\{F_i\}_{i\in[n]}$ be a collection of distributions that are supported on a common set $[u_{min}, u_{max}] \subset \mathbb{R}_+$ which is specified as part of the input,* [10] *and let $r := u_{max}/u_{min}$. Then, for any constant $\epsilon > 0$, there is an algorithm that runs in time polynomial in the size of the input and $n^{\log^9 r/\epsilon^8}$ and computes a price vector P such that*

$$\mathcal{R}_P \geq (1-\epsilon)OPT,$$

where $\mathcal{R}_P$ is the expected revenue under price vector P when the buyer's values for the items are independent draws from the distributions $\{F_i\}_i$ and OPT is the optimal revenue.

Proof of Theorem 51: First set $\hat{\epsilon} = \min\left\{\epsilon, \frac{1}{(4\lceil\log r\rceil)^{1/6}}\right\}$. Clearly, it suffices to find a price vector with expected revenue $(1-\hat{\epsilon})OPT$. Now, let us invoke the reduction of Theorem 15, reducing this task to solving the discrete restricted-price problem $\textsc{RestrictedPrice}(\{\hat{F}_i\}_i, \mathcal{P}, 0.5(\hat{\epsilon}/8)^8)$, where the value distributions $\{\hat{F}_i\}_i$ are supported on a discrete set $\mathcal{S} = \{s^{(1)}, \ldots, s^{(k_1)}\}$ of cardinality $k_1 = O(\log r/\hat{\epsilon}^{16})$ and the prices are also restricted to a discrete set $\mathcal{P} = \{p^{(1)}, \ldots, p^{(k_2)}\}$ of cardinality $k_2 = O(\log r/\hat{\epsilon}^{16})$. It is important to note that $\mathcal{S} \subset [(1+(\hat{\epsilon}/8)^8)u_{min}, (1+(\hat{\epsilon}/8)^8)u_{max}]$ and $\min_i\{p^{(i)}\} \leq \min_i\{s^{(i)}\}$ (this can be checked by a careful study of the proof of Theorem 15). Hence, if $\widehat{OPT}$ is the optimal revenue of the discrete instance resulting from the reduction, we have $\widehat{OPT} \geq \min_i\{p^{(i)}\}$. It is our goal to achieve revenue at least $(1-0.5(\hat{\epsilon}/8)^8)\widehat{OPT}$ in this instance.

To do that, we invoke the algorithm of Section 7 or a modified version of it, depending on the value of n . In particular, if $n \geq \frac{c}{\hat{\epsilon}^{20}}$, for a large enough constant c, we use the algorithm as is, obtaining a price vector with revenue at least

$$\widehat{OPT} - \left(\frac{2k_1k_2}{(n(2r))^2} + \frac{16}{n}\right)\cdot\min_i\{p^{(i)}\} \geq \widehat{OPT}\left(1 - \left(O\left(\frac{\log^2 r}{\hat{\epsilon}^{32}n^2r^2}\right) + \frac{16}{n}\right)\right) \tag{19}$$

$$\geq \widehat{OPT}\left(1 - O(\hat{\epsilon}^8)\right), \tag{20}$$

as we wanted. The running time of the algorithm in this case is polynomial in the input and $(nr)^{\frac{\log^2 r}{\hat{\epsilon}^{32}}}$, that is polynomial in the input and $n^{\frac{\log^3 r}{\hat{\epsilon}^{32}}}$ (assuming $n \geq 2$). If $n \leq \frac{c}{\hat{\epsilon}^{20}}$, we modify the input, introducing dummy items so that the total number of items for sale is $n' = \frac{c}{\hat{\epsilon}^{20}}$. The dummy items have distributions that place a point mass of 1 at value $\min\{s^{(i)}\}/2$ (so we extend the set of values $\{s^{(1)}, \ldots, s^{(k_1)}\}$ by one point $s^{(0)} := \min\{s^{(i)}\}/2$), and hence they do not contribute anything to the revenue, since the minimum price satisfies $\min\{p^{(i)}\} > \min\{s^{(i)}\}/2$ (this is easy to check from the proof of Lemma 44.) Invoking the algorithm of Section 7 on the augmented instance, we obtain a price vector with revenue

$$\widehat{OPT} - \left(\frac{2(k_1+1)k_2}{(n'(4r))^2} + \frac{16}{n'}\right)\cdot\min_i\{p^{(i)}\} \geq \widehat{OPT}\left(1 - \left(O\left(\frac{\log^2 r}{\hat{\epsilon}^{32}(n'r)^2}\right) + \frac{16}{n'}\right)\right) \tag{21}$$

$$\geq \widehat{OPT}\left(1 - O(\hat{\epsilon}^8)\right), \tag{22}$$

as we wanted. The running time of the algorithm in this case is polynomial in the input and $(n'r)^{\frac{\log^2 r}{\hat{\epsilon}^{32}}}$, that is polynomial in the input and $n^{\frac{\log^3 r \log(1/\hat{\epsilon})}{\hat{\epsilon}^{32}}}$ (assuming $n \geq 2$). Being a bit more careful in the application of our discretization lemmas we can obtain a running time of $n^{\log^9 r/\epsilon^8}$. $\square$

Theorems 1 and 2 now follow immediately from Theorem 3 (Theorem 51 in this section) and our structural theorems for Monotone-Hazard-Rate and Regular distributions (Theorems 11 and 13 of Sections 3 and 4 respectively). Here are their statements with some constant optimizations.

[10] The requirement that the set $[u_{min}, u_{max}]$ is specified as part of the input is only relevant if we have oracle access to the distributions of the v_i's, as if we have them explicitly we can easily find $[u_{min}, u_{max}]$.

Corollary 52 (PTAS for MHR Distributions)**.** *Suppose we are given a collection of distributions* $\{F_i\}_{i\in[n]}$ *that are MHR. Then, for any constant* $\epsilon > 0$*, there exists an algorithm that runs in time* $n^{O(\frac{1}{\epsilon^7})}$ *and computes a price vector* P *such that*

$$\mathcal{R}_P \geq (1-\epsilon)\mathcal{R}_{OPT},$$

where $\mathcal{R}_P$ *is the expected revenue under price vector* P *when the buyer's values for the items are independent draws from the distributions* $\{F\}_i$ *and* $\mathcal{R}_{OPT}$ *is the revenue achieved by the optimal price vector.*

Corollary 53 (Quasi-PTAS for Regular Distributions)**.** *Suppose we are given a collection of distributions* $\{F_i\}_{i\in[n]}$ *that are regular. Then, for any constant* $\epsilon > 0$*, there exists an algorithm that runs in time* $n^{O(\frac{(\log(n))^9}{\epsilon^9})}$ *and computes a price vector* P *such that*

$$\mathcal{R}_P \geq (1-\epsilon)\mathcal{R}_{OPT},$$

where $\mathcal{R}_P$ *is the expected revenue under price vector* P *when the buyer's values for the items are independent draws from the distributions* $\{F\}_i$ *and* $\mathcal{R}_{OPT}$ *is the revenue achieved by the optimal price vector.*

We proceed to give the proof of Theorem 4.

Proof of Theorem 4: Denote by $\mathcal{V} = \{v_i\}_{i\in[n]}$ the buyer's values for the items, and let OPT be the optimal revenue. Also, let $\epsilon' = \epsilon/3$.

By Lemma 41, we only need to consider price vectors in $[0,1]^n$. On the other hand, Lemma 29 tells us that if we restrict the prices to be higher than ϵ', we lose at most an additive ϵ' in revenue. So there exists a price vector $\bar{P} \in [\epsilon', 1]^n$, such that $\mathcal{R}_{\bar{P}}(\mathcal{V}) \geq OPT - \epsilon'$

Now, we define a new collection of random variables $\tilde{\mathcal{V}} = \{\tilde{v}_i\}_{i\in[n]}$ via the following coupling: for all $i \in [n]$, set $\tilde{v}_i = \frac{\epsilon'}{2}$ if $v_i < \epsilon'$, and $\tilde{v}_i = v_i$ otherwise. According to Lemma 35, for any price vector P in $[\epsilon', 1]^n$, $\mathcal{R}_P(\tilde{\mathcal{V}}) = \mathcal{R}_P(\mathcal{V})$.

If we are given a price vector $\tilde{P}$, such that $\mathcal{R}_{\tilde{P}}(\tilde{\mathcal{V}}) \geq (1-\epsilon')\mathcal{R}_{OPT}(\tilde{V})$. By Lemma 41 and 29, we can efficiently convert $\tilde{P}$ to $P' \in [\epsilon', 1]^n$, such that $\mathcal{R}_{P'}(\tilde{\mathcal{V}}) \geq \mathcal{R}_{\tilde{P}}(\tilde{\mathcal{V}}) - \epsilon'$.

Combining the inequalities above, we have

$$\mathcal{R}_{OPT}(\tilde{\mathcal{V}}) \geq \mathcal{R}_{\bar{P}}(\tilde{\mathcal{V}}) = \mathcal{R}_{\bar{P}}(\mathcal{V}) \geq OPT - \epsilon',$$

and

$$\mathcal{R}_{P'}(\mathcal{V}) = \mathcal{R}_{P'}(\tilde{\mathcal{V}}) \geq \mathcal{R}_{\tilde{P}}(\tilde{\mathcal{V}}) - \epsilon' \geq (1-\epsilon')\mathcal{R}_{OPT}(\tilde{\mathcal{V}}) - \epsilon'.$$

Thus,

$$\mathcal{R}_{P'}(\mathcal{V}) \geq (1-\epsilon')OPT - 2\epsilon' \geq OPT - 3\epsilon'.$$

So the algorithm is as follows. We first construct $\tilde{\mathcal{V}} = \{\tilde{v}_i\}_{i\in[n]}$, which can be done in time polynomial in n and $1/\epsilon$. Next, we run the PTAS of Theorem 51 on $[\frac{\epsilon'}{2}, 1]$, which runs in time $n^{O(\frac{1}{\epsilon^9})}$ to get $\tilde{P}$. Finally, we convert $\tilde{P}$ to P' in time polynomial in n and $1/\epsilon$. □

We conclude this section with the proof of Theorem 8.

Proof of Theorem 8: We sketch the argument for MHR distributions. The argument for regular distributions is similar. Suppose we have an a-approximation to the optimal revenue for an instance of the pricing problem in which the values are independent and come from MHR distributions. That is, suppose we have β' such that $a\beta' \geq OPT$. We argue that β' can play essentially the same role as β in Theorem 11. More precisely, we show that we can reduce the problem of finding a near-optimal price vector for the original instance to finding a near-optimal price vector for a collection of distributions that are supported on a common interval centered around β'. Hence, we can leverage our knowledge of β' to proceed with our algorithm, obviating the need to compute β.

From Lemma 27, we know that $OPT \geq (1-\frac{1}{\sqrt{e}})\frac{\beta}{2}$. Hence, $\frac{2a}{1-\frac{1}{\sqrt{e}}}\beta' \geq \beta$. Now set $c = \frac{2a}{1-\frac{1}{\sqrt{e}}}$. From Theorem 12, it follows that for any $\epsilon \in (0, 1/4)$

$$Con\left[\max_i\{v_i\} \geq 2\log\left(\frac{1}{\epsilon}\right)(c\beta')\right] \leq 36\epsilon\log\left(\frac{1}{\epsilon}\right)(c\beta').$$

Given this property and that $\beta' \leq OPT$, we can reproduce the proof of Theorem 11 to argue that the resulting distributions can be truncated to the interval $[\epsilon\beta', 2c\log(\frac{1}{\epsilon})\beta']$, without losing more than a $O(\epsilon)$-fraction of the optimal revenue. $\square$

K A Single Price Suffices for I.I.D. MHR Distributions

We provide a faster algorithm for the case where the buyer's values are independent and identically distributed according to some MHR distribution. The main technical idea that goes into the algorithm is this: if the number of items is a sufficiently large function of ϵ, then using a single price suffices to get an ϵ fraction of the optimal revenue. We proceed to the details of our algorithm. To simplify our notation, let us assume that all the v_i's are independent copies of the random variable v, and denote the distribution of v by F. Moreover, let $\alpha_n = \inf\left\{x | F(x) \geq 1 - \frac{1}{n}\right\}$ (as in Definition 18). We start by showing an analogue of Lemma 26.

Lemma 54. *If $S = Con[v \geq (1+\epsilon)\alpha_n]$, then $S \leq \frac{6(1+\epsilon)\alpha_n}{n^{1+\epsilon}}$.*

Using Lemma 54 and Lemma 32, we deduce that if we constrain our prices to be $\leq (1+\epsilon)\alpha_n$, we lose no more most $\frac{6(1+\epsilon)\alpha_n}{n^\epsilon}$ revenue. Given that the optimal revenue with the restriction that all prices be $\leq (1+\epsilon)\alpha_n$ is at most $(1+\epsilon)\alpha_n$, it follows that the optimal revenue without the restriction is at most $(1+\epsilon)\alpha_n + \frac{6(1+\epsilon)\alpha_n}{n^\epsilon} = (1+\epsilon)(1+\frac{6}{n^\epsilon})\alpha_n$. This is very close to α_n if n is a sufficiently large function of ϵ. If that's the case, it suffices to find a price vector achieving revenue close to α_n.

Lemma 55. *If we use the price vector $P = ((1-\epsilon)\alpha_n, (1-\epsilon)\alpha_n, \ldots, (1-\epsilon)\alpha_n)$, we receive revenue at least $\left(1 - e^{(-n^\epsilon)} - \epsilon\right)\alpha_n$.*

Notice that, when $n \geq (1/\epsilon)^{1/\epsilon}$, $n^\epsilon \geq 1/\epsilon$. In this case, we have shown that $OPT \leq (1+\epsilon)(1+6\epsilon)\alpha_n \leq (1+8\epsilon)\alpha_n$. On the other hand, Lemma 55, says that we can achieve revenue at least $(1 - \frac{1}{e^{1/\epsilon}} - \epsilon)\alpha_n$ using a single price. Since $e^{1/\epsilon} \geq 1/\epsilon$, this revenue is at least $(1-2\epsilon)\alpha_n$. Given that $(1+8\epsilon)(1-10\epsilon) \leq (1-2\epsilon)$, we have $(1-2\epsilon)\alpha_n \geq (1-10\epsilon)OPT$. So if we set the price for every item to be $(1-\epsilon)\alpha_n$, we achieve a revenue that is at least $(1-10\epsilon)OPT$.

Theorem 56. *If the values of the buyer are i.i.d. according to a MHR distribution, there is a PTAS for finding a price vector that achieves a $(1-\epsilon)$-fraction of the optimal revenue. The algorithm runs in time linear in $\log(\frac{\log n}{\epsilon})$ and polynomial in $2^{\frac{\log(1/\epsilon)}{\epsilon^8}}$ and the size of the input. Moreover, if $n \geq (12/\epsilon)^{12/\epsilon}$, there exists an efficiently computable price such that, if all items are priced at this price, the resulting revenue is at least $(1-\epsilon)OPT$.*

L Proofs Omitted from Section K

Proof of Lemma 54: By Lemma 19, we know that $(1+\epsilon)\alpha_n \geq \alpha_{n^{1+\epsilon}}$. Thus, $S \leq Con[v \geq \alpha_{n^{1+\epsilon}}]$. But Lemma 21 gives $Con[v \geq \alpha_{n^{1+\epsilon}}] \leq 6\alpha_{n^{1+\epsilon}}/n^{1+\epsilon}$. Hence,

$$S \leq \frac{6\alpha_{n^{1+\epsilon}}}{n^{1+\epsilon}} \leq \frac{6(1+\epsilon)\alpha_n}{n^{1+\epsilon}}.$$

□

Proof of Lemma 55: Let $p = (1-\epsilon)\alpha_n$. By Lemma 19, we know that $\frac{\alpha_{n^{1-\epsilon}}}{(1-\epsilon)} \geq \alpha_n$. Hence, for all i,

$$\Pr[v_i < p] \leq \Pr[v_i < \alpha_{n^{1-\epsilon}}] \leq 1 - \frac{1}{n^{1-\epsilon}}.$$

It follows that

$$\Pr[\exists i, v_i \geq p] \geq 1 - \left(1 - \frac{1}{n^{1-\epsilon}}\right)^n \geq 1 - e^{(-n^{\epsilon})}.$$

Hence, with probability at least $1 - e^{(-n^{\epsilon})}$, the buyer will purchase an item and will pay p. Hence, the revenue is at least $(1 - e^{(-n^{\epsilon})})(1-\epsilon)\alpha_n \geq (1 - e^{(-n^{\epsilon})} - \epsilon)\alpha_n$. □

Proof of Theorem 56: Let $\epsilon' = \epsilon/12$. Depending on the value of n our algorithm proceeds in one of the following ways:

- If $n \geq (1/\epsilon')^{1/\epsilon'}$, we do binary search starting at an anchoring point of the distribution (see Section C) to find some $p \in [1-\epsilon', 1+\epsilon']\alpha_n$. This takes time $O(\log(\frac{\log n}{\epsilon'})$ and the size of the input, since $\alpha_n \leq \alpha_2 \cdot \log_2 n$. We then set every item's price to $(1-2\epsilon')p$. Since $(1-2\epsilon')p \leq (1-\epsilon')\alpha_n$,

  $$\Pr[\exists\ i,\ v_i \geq (1-2\epsilon')p] \geq \Pr[\exists\ i,\ v_i \geq (1-\epsilon')\alpha_n].$$

  On the other hand, $(1-2\epsilon')p \geq (1-2\epsilon')(1-\epsilon')\alpha_n$. Thus, the revenue we obtain if we price all items at $(1-2\epsilon')p$ is at least $(1-2\epsilon')$ times the revenue under price vector $P = ((1-\epsilon')\alpha_n, (1-\epsilon')\alpha_n, \ldots, (1-\epsilon')\alpha_n)$. Hence, the revenue is at least $(1-12\epsilon')OPT = (1-\epsilon)OPT$.
- If $n < (1/\epsilon')^{1/\epsilon'}$, we simply use the algorithm for non-i.i.d. case (Corollary 52).

□