The Power of Two Choices in Dynamic Pricing: As Good As Full Feedback, Instance-Wise

Yan Dai* Tao Lin† Aleksandrs Slivkins‡

November 2026

Abstract

We study the dynamic pricing problem, where a firm sells a product to T myopic customers and their values are independent samples from an unknown value distribution F. With posted prices, the firm only sees whether each customer buys, and the minimax regret is $\widetilde{\Theta}(T^{2/3})$ [KL03]. A *menu* instead offers several options, each selling the product with some probability at some expected payment. Infinitely many menu options allow the seller to fully reveal customer values (i.e., eliciting full feedback), reducing the regret to $\Theta(\sqrt{T})$. How many options are needed?

Perhaps surprisingly, a second menu option suffices, achieving the same regret rates as infinite menus (up to log factors). This holds broadly: without knowing any property of F, our algorithm with two menu options attains $\widetilde{\mathcal{O}}(\sqrt{T})$ regret in general, $\widetilde{\mathcal{O}}(T^{1/3})$ regret for monotone-hazard-rate distributions, and $\widetilde{\mathcal{O}}(T^{d/(d+1)})$ instance-dependent regret when $d \in (0,1)$ is a suitable notion of "dimension". All are optimal even with infinite menus. Moreover, we argue, via a novel lower-bound construction, that our algorithm is near-optimal *for each problem instance*.

Along the way, we study the query complexity of revenue maximization: An algorithm issues a *query* to each customer, from some class of feasible queries, and learns something about their value. The goal is to learn the property of interest (either the best price or the CDF) up to some precision. Query complexity is the smallest number of queries to guarantee this, in the worst case over some class of value distributions. We find query complexity technically intertwined with dynamic pricing, with posted pricing and full feedback corresponding to, resp., threshold queries and value queries. Given these two extremes, how powerful are the "intermediate" query classes?

Our answer focuses on *interval queries*, which reveal whether customer value lies in a given interval. We show that they suffice to achieve the same query complexity as value queries (up to log factors) for both properties (best price and CDF) and across several classes of value distributions. In particular, we provide a new interval-query algorithm for each property, which handles all distribution classes at once. The best-price-estimation algorithm is the key ingredient in our main dynamic pricing algorithm, and the CDF-estimation algorithm also underpins a simpler dynamic pricing algorithm which attains $\widetilde{\mathcal{O}}(\sqrt{T})$ regret with three menu options.

*Operations Research Center, Massachusetts Institute of Technology. Email: yandai20@mit.edu.
Part of the research was done when Yan Dai was a research intern at Microsoft Research New England.

†School of Data Science, The Chinese University of Hong Kong, Shenzhen. Email: lintao@cuhk.edu.cn.
Part of the research was done when Tao Lin was a postdoctoral researcher at Microsoft Research New England.

‡Microsoft Research NYC. Email: slivkins@microsoft.com.

Contents

1 Introduction **1**

2 Setup: Dynamic Pricing with Menus **6**

3 Upper Bounds: Dynamic Pricing with Size-2 Menus **7**
3.1 Connecting Menus and Interval Queries 7
3.2 Size-3 Menus and Interval-Query Complexity of CDF Estimation 8
3.3 Size-2 Menus and Interval-Query Complexity of the Monopoly Price 8
3.4 Upper Bounds for Dynamic Pricing with Size-2 Menus 10

4 Lower Bounds: From Minimax to Instance-Optimality **11**
4.1 Minimax Lower Bounds 11
4.2 Near Instance-Optimality of Algorithm 2 12
4.3 Constructions of Decoy Instances 13
4.4 Anytime Lower Bounds 16
4.5 Dynamic Pricing Compared to Lipschitz Bandits or Experts 17

5 Extension: Adversarial Dynamic Pricing Using Size-3 Menus **18**
5.1 Adversarial Dynamic Pricing Algorithm using Online Learning 19

6 Further Discussion of Related Work **22**

7 Conclusion and Future Research **23**

Technical Appendices **29**

1 Introduction

Consider a firm introducing a new product to the market, whose values to customers – the highest price they are willing to pay – are randomized and unknown. Hence, at a price p, a customer buys with probability $D(p) := \Pr_{v\sim F}\{v \geq p\}$, where F is a fixed but unknown distribution of values; the function $D(\cdot)$ is called the *demand curve*. The firm learns this function (or equivalently, the value distribution F) via dynamic pricing and customers' buying decisions.

Concretely, consider the now-standard formulation of dynamic pricing [KL03]. A revenue-maximizing firm has one product for sale, with unlimited supply. In each round $t \in [T]$, a new customer arrives with value $v_t \in [0,1]$ drawn from F. The firm posts a price $p_t \in [0,1]$ based on the history, according to some algorithm. The customer buys the product if and only if $v_t \geq p_t$. The firm never sees v_t but observes the buying decision $\mathbb{1}[v_t \geq p_t]$. If F were known, the optimal algorithm would post the *monopoly price* p^*_F in every round: the price maximizing expected revenue $R(p) := p \cdot D(p)$ [Mye81]. The firm aims to minimize its *regret*, defined as $T \cdot R(p^*_F)$ minus its expected cumulative revenue in T rounds. One can interpret regret as how much the firm would pay to know the demand curve in advance, e.g., via market research. This problem and many of its variants are well-studied; in particular, the minimax-optimal regret across all problem instances is of the order $\widetilde{\Theta}(T^{2/3})$ [KL03].[1] It is closely related to multi-armed bandits and particularly to Lipschitz bandits.[2]

However, firms commonly offer a *menu* instead of a single price, e.g., different subscription levels at different prices. Within our single-product framework, a menu is defined as a collection of options, where each option $(q,p) \in [0,1]^2$ indicates that the customer gets the item with probability q and pays p in expectation.[3] Option $(0,0)$, the "outside option" of not doing anything, is always implicitly included. Thus, a single posted price p becomes a size-1 menu $\{(1,p)\}$. A customer with value v chooses the option with the highest utility $qv - p$, or buys nothing if all utilities are negative. Such choices can potentially reveal information about the customer's value, allowing the seller to learn the unknown F and optimize future prices. Our central question is:

How beneficial are menus for dynamic pricing?

Thus, we study *dynamic pricing with menus*, where each customer is offered a menu designed by the seller's algorithm based on history, and the chosen option is observable. The optimal algorithm under the known value distribution F remains the same: always post price p^*_F, defined as above, which holds even when infinitely many options (*infinite menus*) are allowed [Mye81].

Warm-Up: Infinite Menus. We achieve regret $\mathcal{O}(\sqrt{T})$ using infinite menus, with a complementary minimax lower bound. This is in stark contrast to the $\widetilde{\Theta}(T^{2/3})$ regret rate without menus. While infinite menus are impractical, we treat them as a useful baseline and a stepping stone for finite-menu results. To elaborate on infinite menus as a baseline, we show that infinite menus are equivalent to *full feedback*, where each customer's value v_t is directly observed after each round.[4] That is, with either infinite menus or full feedback (or both), the seller achieves the same expected revenue in dynamic pricing; this holds for every realization of values $(v_t)_{t\in[T]}$.

[1] The *minimax-optimal regret* is the optimal regret over all algorithms, in the worst case over a given class of problem instances, for a given time horizon T. A lower bound on this quantity is called a *minimax lower bound*.

[2] In multi-armed bandits [LS20; Sli19], in each round the algorithm chooses some "arm" and receives a reward for this arm (drawn from some fixed but unknown distribution). Lipschitz bandits [KSU19; BMSS11] is a well-studied bandit problem where expected rewards obey a Lipschitz condition, such as $|R(p) - R(p')| \leq |p - p'|$, for all $p, p' \in [0,1]$. Dynamic pricing satisfies a one-sided version: whenever $p > p'$. See Related Work (Section 6) for more background.

[3] One could also interpret q as quality, if varying qualities are possible and customer utility is linear in quality.

[4] Generically, *full feedback* refers to observing the outcome of every possible action one could have taken.

Table 1: Regret Bounds under Different Menu Sizes

Value Distribution	With Infinite Choices (Size-∞ Menus)	With a Second Choice (Size-2 Menus)	Only Posted Price (Size-1 Menus)
General	$\Omega(\sqrt{T})$ **(Thm 4.1a)**	$\widetilde{\mathcal{O}}(\sqrt{T})$ **(Thm 3.4a)**	$\widetilde{\Theta}(T^{2/3})$ [KL03]
MHR	$\Omega(T^{1/3})$ **(Thm 4.1b)**	$\widetilde{\mathcal{O}}(T^{1/3})$ **(Thm 3.4b)**	$\widetilde{\Theta}(\sqrt{T})$ [SW24; LSTW23b]
Constant curvature	$\Omega(T^{1/3})$ **(Thm 4.1c)**	$\widetilde{\mathcal{O}}(T^{1/3})$ **(Thm 3.4c)**	$\widetilde{\mathcal{O}}(T^{3/5})$ [KSU19] $\Omega(T^{3/5})$ **(Lm B.29)**
Interval Dimension $d \in (0,1)$	$\Omega(T^{d/(d+1)})$ **(Thm 4.1d)**	$\widetilde{\mathcal{O}}(T^{d/(d+1)})$ **(Thm 3.5)**	$\widetilde{\mathcal{O}}(T^{(d+1)/(d+2)})$ [KSU19] $\Omega(T^{(d+1)/(d+2)})$ **(Thm 4.6)**

Main Contributions. Our main high-level result is that *size-2 menus* are as powerful as *infinite menus*: they achieve the same regret rates that are better than the rates without menus. We prove this in a very broad sense. First, we achieve $\widetilde{\Theta}(\sqrt{T})$ regret for all problem instances. Second, we improve this to $\widetilde{\Theta}(T^{1/3})$ for all instances with monotone hazard rate (MHR) or constant curvature (a similar property defined in [KL03]). Third, we obtain a continuous interpolation: $\widetilde{\Theta}(T^{d/(d+1)})$ regret for all instances with *interval dimension* $d \in [0,1]$, a concept similar to the zooming dimension in Lipschitz bandits [KSU19; BMSS11]. In each scenario, we obtain an upper bound for *size-2 menus*, and a matching minimax lower bound for *infinite menus*. We summarize them in Table 1, along with the rates attainable without menus. All upper bounds are obtained by a single algorithm (Algorithm 2) and via a unified analysis, which is our first key contribution from a technical perspective.

Interestingly, we obtain minimax regret rates $o(\sqrt{T})$, which is uncommon for bandit problems even under full feedback. This is where the special structure of dynamic pricing is essential, as opposed to, e.g., the more generic structure of Lipschitz bandits. The latter incurs $\Omega(\sqrt{T})$ minimax regret rate under full feedback, even when the interval dimension is $d = 0$. In contrast, when only using posted prices, dynamic pricing yields the same minimax regret rates as Lipschitz bandits (see Section 4.5).

While standard minimax lower bounds are tailored to a particular T, we also provide *anytime* lower bounds that apply to infinitely many T's at once. They rule out any algorithm that beats every fixed-T lower bound *eventually*, after some instance-dependent initial time period. They handle problem instances with any fixed interval dimension $d \in (0,1)$, with the same rates as in Table 1.

Furthermore, we argue our algorithm is near-optimal *for every value distribution* F, rather than (just) in the worst case over some class of problem instances. This is due to a more intricate lower bound we prove. The "shape" of such a lower bound must be non-trivial, as something must prevent the algorithm from always posting the monopoly price p^*_F. So, given F and time horizon T, we construct two value distributions (called *decoys*) with the following properties. First, letting V be our algorithm's regret on F, no other algorithm, even with infinite menus, can have regret much better than V simultaneously on F and on both decoys. Second, the decoys are close to F: Their Hellinger distance to F tends to 0 as $T \to \infty$. Both properties can be made more concrete in various ways, under additional assumptions. In particular, the decoys can preserve the properties of F in terms of interval dimension or MHR. These lower-bound constructions, unified by a common idea specifically tailored for dynamic pricing, jointly constitute our second key technical contribution.

Ancillary Contributions: Interval Query Complexity. Dynamic pricing is closely related to query complexity of revenue maximization. Here, an algorithm faces a sequence of customers, with values as described above. In each round t, the algorithm issues a *query* to the customer, from some

Table 2: Sample, Interval-Query, and Threshold-Query Complexities

Property	Value Distribution	Sample Complexity	Interval-Query Complexity	Threshold-Query Complexity
CDF	General	$\mathcal{O}(\epsilon^{-2})$ [DKW56] $\Omega(\epsilon^{-2})$ [HMR15]	$\widetilde{\mathcal{O}}(\epsilon^{-2})$ **(Thm 3.2)**	$\widetilde{\Theta}(\epsilon^{-3})$ [LSTW23b]
	Regular	$\Omega(\epsilon^{-2})$ **(Lm B.16)**		$\widetilde{\mathcal{O}}(\epsilon^{-2.5})$ [LSTW23a] $\Omega(\epsilon^{-2.5})$ [LSTW23b][a]
	MHR			
	Constant Curvature	$\Omega(\epsilon^{-2})$ **(Lm B.17)**		$\widetilde{\Theta}(\epsilon^{-3})$ **(Lm B.27)**
Monopoly Price	General	$\mathcal{O}(\epsilon^{-2})$ [DKW56] $\Omega(\epsilon^{-2})$ [HMR15]	$\widetilde{\mathcal{O}}(\epsilon^{-2})$ **(Thm 3.2)**	$\widetilde{\Theta}(\epsilon^{-3})$ [LSTW23b]
	Regular	$\Omega(\epsilon^{-2})$ **(Lm B.14)**		$\widetilde{\Theta}(\epsilon^{-2})$ [LSTW23b]
	MHR	$\widetilde{\Theta}(\epsilon^{-1.5})$ [HMR15]	$\widetilde{\mathcal{O}}(\epsilon^{-1.5})$ **(Thm 3.4)**	
	Constant Curvature	$\Omega(\epsilon^{-1.5})$ **(Lm B.18)**		$\widetilde{\mathcal{O}}(\epsilon^{-2.5})$ [KSU19] $\Omega(\epsilon^{-2.5})$ **(Lm B.28)**

[a] Their hard instance [LSTW23b, Theorem 3.7] is also MHR, even though they do not state so explicitly.

class of feasible queries, and learns something about value v_t. The goal is to learn a property of interest: either the CDF of value distribution, or just the monopoly price p^*_F. The *query complexity* is the smallest number of queries guaranteed to learn the property up to a given precision $\epsilon > 0$.

Prior work studied *threshold queries* that reveal $\mathbb{1}[v_t \geq p]$ for a given threshold p [LSTW23a; LSTW23b] and *value queries* that directly reveal v_t [DKW56; HMR15].[5] These two query classes correspond to dynamic pricing with size-1 menus and full feedback, respectively, in terms of the information revealed after each round. Their query complexities are well-understood; see Table 2.

While the threshold and value queries represent the two extremes, our concern is "intermediate" queries: ones that reveal more information than threshold queries but less than value queries. How powerful are they? This is poorly understood, much like dynamic pricing with finite menus.

We focus on *interval queries* that reveal whether v_t lies in a given interval $[a, b)$. This is, in some sense, the least informative query class above threshold queries.[6] Our high-level result is that *interval queries are maximally powerful in terms of query complexity, i.e., as powerful as value queries.* This again holds in a very broad sense: for both properties (CDF and monopoly price) and across several classes of value distributions (all distributions, regular distributions, MHR, and those of constant curvature). We provide a new interval-query algorithm for each property, which handles all distribution classes at once. See Table 2 for specific upper and lower bounds.

Most notably, we show that the minimax-optimal interval-query complexity is $\widetilde{\mathcal{O}}(\epsilon^{-2})$ for both properties, matching the lower bounds from [HMR15]. We extend the $\widetilde{\Omega}(\epsilon^{-2})$ lower bound to several classes of instances: MHR and constant-curvature (for CDF estimation) and regular value distributions (for both CDF and monopoly-price estimation). Moreover, we strengthen the monopoly-price *upper* bound to $\widetilde{\mathcal{O}}(\epsilon^{-1.5})$ for MHR and constant-curvature (and show that it is optimal).

Our results on query complexity are closely connected to those on dynamic pricing. We derive and leverage several formal connections between the two problems (see Figure 1). Some of these connections are "black-box reductions", converting an arbitrary algorithm for one problem into an algorithm for another (and likewise for their performance guarantees). Some others are "ad-hoc

[5] Value-query (a.k.a. *sample*) complexity is widely studied in dynamic pricing [CR14; GHZ19; GHTZ21; JKMS24].
[6] If one can only perform N threshold queries on a given customer, interval queries correspond to $N = 2$.

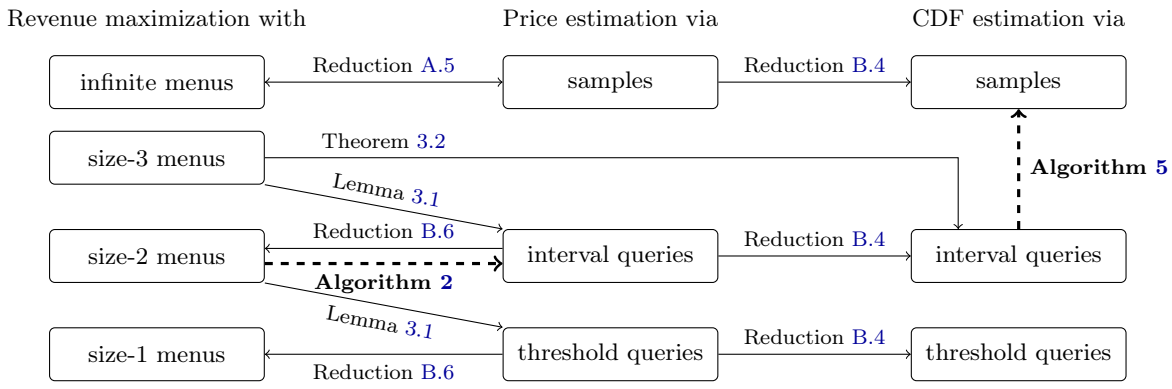


Figure 1: Formal connections between query complexity and dynamic pricing with menus. Arrows stand for black-box reductions, while thick dashed arrows are our ad-hoc algorithms.

reductions", converting a particular approach for one problem into an algorithm for another. In particular, our best-price-estimation algorithm is the key ingredient in our main algorithm for dynamic pricing using size-2 menus, and our CDF-estimation algorithm underpins a simple approach which attains $\widetilde{\mathcal{O}}(\sqrt{T})$ minimax regret with three menu options.

Extension: Adversarial Dynamic Pricing with Menus. We come back to dynamic pricing with menus and consider *adversarial values*: Each value v_t is chosen by an adversary, rather than drawn from a fixed distribution. The choice of v_t can be randomized and adaptive to the history.

Using size-3 menus, we achieve $\widetilde{\mathcal{O}}(\sqrt{T})$ minimax regret against the best fixed price in hindsight, whose minimax optimum is $\widetilde{\Theta}(T^{2/3})$ when only using posted prices [KL03]. This is the best possible even with infinite menus. Leveraging our machinery for i.i.d. values, we plug any suitable algorithm $\mathcal{A}^{\mathrm{OL}}$ for *online learning* (i.e., adversarial bandits with full feedback) into our approach. Using Hedge [FS97], we attain the aforementioned $\widetilde{\mathcal{O}}(\sqrt{T})$ bound; whereas using Fixed-Share [HW98], we obtain $\widetilde{\mathcal{O}}(\sqrt{ST})$ regret against any price sequence that switches the price at most $S \geq 1$ times.

Upper Bound Techniques. Our dynamic pricing algorithms build on those for interval queries. With threshold queries, estimating a CDF up to ϵ takes $\widetilde{\Theta}(\epsilon^{-3})$ queries [LSTW23b]: Within any of the ϵ^{-1} intervals of width ϵ, one can move a mass of order ϵ from its right half to its left half, such that any threshold query outside it tells no difference. Hence one must spend $\Omega(\epsilon^{-2})$ threshold queries on each of these intervals, thus giving $\Omega(\epsilon^{-3})$. An interval query on the left half of it, in contrast, observes a coin whose mean, and hence variance, is $\mathcal{O}(\epsilon)$, so the same change is detectable in $\mathcal{O}(\epsilon^{-1})$ queries. In general, for an interval I with probability mass $D(I)$, its outcome (which is Bernoulli) has variance no more than $D(I)$, and thus empirical Bernstein [MP09] estimates $D(I)$ up to ϵ using $\widetilde{\mathcal{O}}(\frac{D(I)}{\epsilon^2} + \frac{1}{\epsilon})$ queries. We thus build a segment tree whose leaves are the ϵ^{-1} intervals of width ϵ, and apply empirical Bernstein to estimate each tree interval up to $\epsilon/(\log_2 \epsilon^{-1})$. As the intervals of each depth are disjoint, their masses sum to 1; since there are $\log_2 \epsilon^{-1}$ depths, this approach uses $\widetilde{\mathcal{O}}(\epsilon^{-2})$ interval queries (Theorem 3.2). Translating interval queries to menus directly, however, requires menus with 3 options, because we need an extra option to "exploit" the estimated monopoly price $\widehat{p}^*$. It also cannot bypass the worst-case rate of $\widetilde{\mathcal{O}}(\epsilon^{-2})$, due to our near-matching $\Omega(\epsilon^{-2})$ sample complexity lower bounds for regular, MHR, or curvature distributions (see Table 2).

We then directly estimate the monopoly price p^*, which only requires estimating those near-optimal

prices. Our starting point is the one-sided Lipschitzness of the revenue function R: For any $p < p'$,

$$R(p') - R(p) = (p' - p) \Pr_{v \sim F}\{v \geq p'\} - p \Pr_{v \sim F}\{p \leq v < p'\} \leq p' - p. \tag{1}$$

If $p < p'$ are both Δ-optimal, i.e., $R(p), R(p') \geq R(p^*) - \Delta$, then $p \Pr_{v\sim F}\{p \leq v < p'\} \leq p' - p + \Delta$; therefore, *close near-optimal prices also have little mass in between*. This allows a phased elimination approach (Algorithm 2): For a halving $\Delta = \frac{1}{2}, \frac{1}{4}, \ldots$, we start with 4Δ-optimal candidates $p_0 < \cdots < p_N$, estimate all revenue differences up to $\frac{\Delta}{4}$ with a segment tree over candidates, refine our candidate set, and eliminate the sub-optimal ones. Using the one-sided Lipschitzness in Eq. (1), the total queries spent on a Δ is $\widetilde{\mathcal{O}}(\frac{p_N - p_0 + \Delta}{\Delta^2})$ (Lemma 3.3). This not only gives $\widetilde{\mathcal{O}}(\epsilon^{-2})$ in general (via $p_N - p_0 \leq 1$), but also gives $\widetilde{\mathcal{O}}(\epsilon^{-1.5})$ for MHR or curvature distributions whose Δ-optimal region can be no wider than $\sqrt{\Delta}$. Even better, our interval queries always have one Δ-optimal endpoint, thus eliminating the need for an "exploitation price" and being *implementable using size-2 menus*.

In general, for any value distribution F, let its covering number $N_F(\Delta)$ be the number of length-Δ intervals needed to cover the convex hull of all Δ-optimal prices. Algorithm 2 spends $\widetilde{\mathcal{O}}(N_F(4\Delta)/\Delta)$ rounds on Δ (since the 4Δ-optimal candidates satisfy $p_N - p_0 \leq 4\Delta \cdot N_F(4\Delta)$), each with regret $\mathcal{O}(\Delta)$. Therefore, its regret is $\widetilde{\mathcal{O}}(V_T(F))$ with $V_T(F) := \sup_\Delta \min\{T\Delta, N_F(\Delta)\}$. When $N_F(\Delta) = \mathcal{O}(\Delta^{-d})$ for some $d \in [0, 1]$ (when we say F has interval dimension d), this becomes $\widetilde{\mathcal{O}}(T^{d/(d+1)})$ (Theorem 3.5).

Lower Bound Techniques. Our lower bounds hold around *every* F, which rules out the usual approach: Lower bounds for Lipschitz bandits plant a bump at an arbitrary location [KSU19, §5.1], but in pricing, the revenues of all prices are determined by the same value distribution F, so making one price Δ-better than another requires moving masses between them. A key fact is that Eq. (1) is tight in both directions: While two close Δ-optimal prices have little mass in between (enabling our upper bounds), *between two far apart ones*, say the smallest and largest Δ-optimal prices, *the mass is at least* $\Delta N_F(\Delta)$. Hence scaling the density on a half of this interval by $1 \pm \Theta(1/N_F(\Delta))$ makes either end Δ-better than the other, and the resulting decoys G_+, G_- have squared Hellinger distances from F bounded by $\Delta/N_F(\Delta)$. No algorithm, even with infinite menus, tells them apart in $\Theta(N_F(\Delta)/\Delta)$ rounds, hence giving an $\Omega(\min\{T\Delta, N_F(\Delta)\}) = \Omega(V_T(F))$ lower bound on $\{G_+, G_-\}$.

To claim instance-optimality of Algorithm 2, we further require $V_T(G_\pm)$ not to be much larger than $V_T(F)$. But the previous move may lift a plateau of R, creating many more near-optimal prices. We again scale the density by a relative $\Theta(1/N_F(\Delta))$ within the new Δ-optimal region, down on the left of the new $p^*_\pm$ and up on its right, so that the new $R_\pm$ decays at least linearly. This keeps $V_T(G_\pm)$ at most a constant factor larger than $V_T(F)$, giving Theorem 4.2. A more detailed proof sketch, as well as an illustration (Figure 2), can be found in Section 4.3. The refined versions, Theorem 4.4 for interval dimensions and Theorem 4.5 for MHR, follow similar but more delicate constructions.

The decoy instances above depend on T, and we further develop anytime lower bounds using a similar idea. We give a (randomized) value distribution of interval dimension d which almost surely ensures that any algorithm, even with infinite menus, suffers regret $\Omega(T^{d/(d+1)})$ for infinitely many T's. Instead of the single-parameter family under posted prices [KL03], which only gives $\widetilde{\mathcal{O}}(1)$ regret under infinite menus, we construct a fractal-like value distribution F: We plant the two-point construction above at super-exponential scales $\Delta_1 > \Delta_2 > \cdots$ with $\Delta_{k+1} = \Delta_k^q$ for a constant $q > 1$, each on the top of the previous one, with a fresh fair coin deciding which half is better. Then each scale Δ_k forces $\Omega(N(\Delta_k)) = \Omega(T^{d/(d+1)})$ regret at $T \approx N(\Delta_k)/\Delta_k$ with probability $\frac{1}{2}$, and hence we derive an $\Omega(T^{d/(d+1)})$ lower bound that holds infinitely often almost surely (Theorem 4.6).

2 Setup: Dynamic Pricing with Menus

Notations. For a positive integer n, let $[n] := \{1, 2, \ldots, n\}$. For a (possibly infinite) set $\mathcal{X}$, $\triangle\mathcal{X}$ denotes the set of distributions over $\mathcal{X}$. We denote the Dirac measure over p by δ_p. We use standard big-oh notations: $\mathcal{O}, \Omega, \Theta$ hide absolute constants, and $\widetilde{\mathcal{O}}, \widetilde{\Omega}, \widetilde{\Theta}$ further hide logarithmic factors.

Problem Statement. We have a T-round game between a seller and myopic customers. The seller has an infinite supply of a product. In each round $t \in [T]$, a customer with value $v_t \in [0, 1]$ comes, where v_t is independently sampled from a fixed but unknown *value distribution* $F \in \triangle[0, 1]$ (we extend to adversarial values in Section 5). Neither the value v_t nor the value distribution F is known.

The seller offers a menu M_t to this customer. A menu is defined as a collection of options $(q, p) \in [0, 1] \times [0, 1]$, where each option $(q, p) \in M_t$ indicates that the customer choosing it gets the item with probability q and pays p in expectation (i.e., whether or not they get the item). We call $|M_t|$ the *size* of the menu, so that a single posted price p_t is equivalent to a size-1 menu $M_t = \{(1, p_t)\}$.

The payoff of a customer with value v under option (q, p) is quasi-linear as $U(q, p; v) := q \times v - p$. Customer t chooses an option $(q_t, p_t) \in M_t \cup \{(0, 0)\}$ maximizing their utility $U(\,\cdot\,; v_t)$. The "outside option" $(0, 0)$ for not doing anything is always implicitly included in the menu. When several options maximize utility, we assume the customer chooses the one maximizing allocation probability.

The seller observes the chosen option (q_t, p_t) but not the value v_t. Thus, the observable history up to the end of round t is $\mathcal{H}_t = (M_s, q_s, p_s)_{s \in [t]}$. The menu M_t can depend on the history $\mathcal{H}_{t-1}$. We design algorithms for the seller to construct menus, so as to maximize the expected total revenue.

Monopoly Price and Regret. The seller competes with the *monopoly price*, defined as

$$p^*_F := \operatorname*{argmax}_{p \in [0,1]} R_F(p), \qquad \text{where} \quad D_F(p) := \Pr_{v \sim F}\{v \geq p\} \quad \text{and} \quad R_F(p) := p \cdot D_F(p) \tag{2}$$

are the *demand curve* and the *revenue function* of F, respectively. This price always exists since the demand curve is non-increasing and left-continuous in p; if there are several maximizers, p^*_F is any fixed one of them. If F were known, posting p^*_F in all rounds is optimal across all algorithms, even if menus are allowed [Mye81]. We define *regret* w.r.t. the monopoly price:

$$\mathrm{Reg}_F(T) := T \cdot R_F(p^*_F) - \mathbb{E}\left[\textstyle\sum_{t \in [T]} p_t\right],$$

mirroring the standard notion of pseudo-regret from (stochastic) multi-armed bandits. We omit the subscript F when it is clear from the context. We are often interested in *minimax-optimal regret*: the best possible T-round regret achievable by a dynamic pricing algorithm using menu size at most $m \in \mathbb{N} \cup \{\infty\}$, in the worst case over a given class $\mathcal{F}$ of value distributions:

$$\mathrm{Reg}^{(m)}_{\mathcal{F}}(T) := \inf_{\text{algorithms using size-}m\text{ menus}} \sup_{F \in \mathcal{F}} \mathrm{Reg}_F(T).$$

Classes of Interest. We are particularly interested in the following classes of value distributions:

- General distributions ($\mathcal{F}_{\text{all}}$): all distributions over $[0, 1]$, including those without densities.
- Regular distributions [Mye81] ($\mathcal{F}_{\text{regular}}$): value distributions with probability density function (PDF) f such that the *virtual value* function $\phi(v) := v - \frac{D_F(v)}{f(v)}$ is non-decreasing.
- Monotone hazard rate (MHR) [BMP63] ($\mathcal{F}_{\text{MHR}}$): those with a non-decreasing *hazard rate* function $h(v) := \frac{f(v)}{D_F(v)}$. Both regular and MHR are standard in economic theory with $\mathcal{F}_{\text{MHR}} \subseteq \mathcal{F}_{\text{regular}}$.

- Distributions with *curvature* $\kappa \in (0,2)$ [KL03] ($\mathcal{F}_{\text{curv}(\kappa)}$): value distributions which satisfy the property that $R(p^*) - R(p) \geq \kappa(p-p^*)^2$ for all prices $p \in [0,1]$.[7]

Other Notations. For the ease of reading, we write $\text{Reg}^{(m)}_{\text{all}}(T) := \text{Reg}^{(m)}_{\mathcal{F}_{\text{all}}}(T)$, and similarly for the other classes. The *price space* $[0,1]$ is the set of all feasible prices. A *price interval* is an interval in the price space. Slightly abusing the $D_F(\cdot)$ notation, let $D_F(I) := \Pr_{v\sim F}\{v \in I\}$ be the probability of interval I under F. A price p is *ϵ-optimal* if $R(p) \geq R(p^*) - \epsilon$. A distribution $P \in \triangle[0,1]$ over prices is ϵ-optimal if its expected revenue $\mathbb{E}_{p\sim P}[R(p)] \geq R(p^*) - \epsilon$.

Query Complexity. For monopoly-price estimation, given a class $\mathcal{F}$ of value distributions, we define the query complexities $\textsc{Sample}_{\mathcal{F}}(\epsilon)$, $\textsc{Threshold}_{\mathcal{F}}(\epsilon)$, and $\textsc{Interval}_{\mathcal{F}}(\epsilon)$ as, resp., the minimal number of value queries (samples), threshold queries, and interval queries needed to find an ϵ-optimal price distribution. For CDF estimation, the goal is to find a value distribution $\widehat{F}$ that is ϵ-close to F in Lévy distance with probability $\frac{3}{4}$. The corresponding query complexities are denoted, resp., as $\textsc{Sample}^{\text{CDF}}_{\mathcal{F}}(\epsilon)$, $\textsc{Interval}^{\text{CDF}}_{\mathcal{F}}(\epsilon)$, and $\textsc{Threshold}^{\text{CDF}}_{\mathcal{F}}(\epsilon)$. Formal definitions are in Appendix B.1.

3 Upper Bounds: Dynamic Pricing with Size-2 Menus

In this section we present our main algorithmic results, as described in the Introduction. The main theorems are stated in Section 3.4. The rest of the section builds towards these theorems. In Section 3.1, we establish a key property that connects menus and interval queries. In Section 3.2, we design a new interval-query algorithm for CDF estimation, matching the query complexity of value queries. Invoking the reduction from Section 3.1, we obtain a simple algorithm for dynamic pricing with size-3 menus which attains $\widetilde{\mathcal{O}}(\sqrt{T})$ regret. In Section 3.3, we specify our main algorithm (Algorithm 2) for dynamic pricing with size-2 menus, and prove a key technical lemma about its performance. Internally, this algorithm is based on monopoly-price estimation via interval queries (as CDF estimation is wasteful) and can be usefully repurposed for that problem (see Appendix B.5).

3.1 Connecting Menus and Interval Queries

This subsection presents an interesting connection between menus and interval queries. Let P be a *price distribution* supported on m points $0 \leq p_1 \leq \cdots \leq p_m \leq 1$. An interval query on two consecutive support points $[p_i, p_{i+1})$ asks whether the customer's value is in that interval. The key property enabling all our results is that any price distribution P and an interval query on its support points can be "implemented" by a menu. This menu obtains the same expected revenue as a randomized posted price $p \sim P$ and (at least) the same information as the interval query.

Lemma 3.1. *Fix a price distribution P with discrete support $0 \leq p_1 < \cdots < p_m \leq 1$. For $i \in [m]$, let w_i be the probability of p_i under P. Construct a size-m menu $M = \left\{ \left(\sum_{j\leq i} w_j,\ \sum_{j\leq i} w_j p_j \right) \,\middle|\, i \in [m] \right\}$.*

Given M, a customer with value $v \in [p_i, p_{i+1})$ chooses the i-th option, where $p_{m+1} := \infty$, and one with $v < p_1$ chooses the outside option. Hence, the seller observes which of the $(m+1)$ intervals, $[0,p_1), [p_1,p_2), \ldots, [p_m, 1]$, contains v; and the customer gets the item with probability $\Pr_{p\sim P}\{v \geq p\}$ and pays $\mathbb{E}_{p\sim P}[p\,\mathbb{1}[v \geq p]]$ in expectation, the same as posting a random price $p \sim P$.

[7] $\mathcal{F}_{\text{curv}(\kappa)} = \varnothing$ for $\kappa > 2$ and $F \in \mathcal{F}_{\text{curv}(2)}$ is trivial (Lemma B.1); we hence assume $\kappa \in (0,2)$. The original "curvature and smooth" assumption [KL03, Theorem 1.2] assumes R has a unique maximizer $p^* \in (0,1)$, $R \in \mathcal{C}^{(2)}$ near p^*, and $R''(p^*) < 0$. Such value distributions lie in $\mathcal{F}_{\text{curv}(\kappa)}$ for some $\kappa > 0$ [KL03, Lemma 3.11]. Under their assumption, they achieve $\widetilde{\mathcal{O}}(\sqrt{T})$ regret for posted pricing, whereas we prove $\text{Reg}^{(1)}_{\text{curv}(\kappa)}(T) \geq \Omega(T^{3/5})$ (Lemma B.29).

Subroutine 1 EMPIRICALBERNSTEIN(interval I, precision ϵ, failure probability δ)

for $K = 1, 2, \ldots$ **do** ▷ *This loop is adaptive and issues an indefinite number of queries*

  Issue an interval query on I, observe a new feedback $X_K = \mathbb{1}[v \in I]$ with fresh $v \sim F$

  Calculate $\widehat{\mu}_K := \frac{1}{K}\sum_{k=1}^{K} X_k$ and the confidence radius r_K according to Eq. (3)

$$r_K := \sqrt{\frac{2\iota}{K} \cdot \frac{1}{K(K-1)} \sum_{k<k'} (X_k - X_{k'})^2} + \frac{7\iota}{3(K-1)}, \quad \iota = \log \frac{2}{\delta} \tag{3}$$

  Terminate if $K \geq 2$ and $r_K \leq \epsilon$, in which case **return** estimation $\widehat{D}(I) := \widehat{\mu}_K$

Therefore, menus enable the seller to perform interval (instead of threshold) queries on v_t, while still collecting the same expected revenue as posting a random price. Specifically, to implement an interval query $[a, b)$ while exploiting a price r, such as the current estimate of p^*, we apply Lemma 3.1 to the price distribution $P = \gamma\delta_a + \gamma\delta_b + (1-2\gamma)\delta_r$ where $0 < \gamma < 1/2$ is arbitrarily small. The resulting menu has three options, reveals whether the customer's value lies in $[a, b)$, and earns expected revenue at least $R(r) - 2\gamma$. A size-2 menu can also ask an interval query $[a, b)$, but only exploits either a or b. Hence dynamic pricing using size-3 menus reduces to estimating p^* via interval queries, while size-2 menus additionally require every query to have a near-optimal endpoint.

3.2 Size-3 Menus and Interval-Query Complexity of CDF Estimation

Starting with CDF estimation, we provide a new interval-query algorithm, showing that interval queries are as powerful as value queries. This also enables dynamic pricing using size-3 menus.

Theorem 3.2. $\textsc{Interval}_{\text{all}}^{\text{CDF}}(\epsilon) = \mathcal{O}(\epsilon^{-2}\log^4\frac{2}{\epsilon})$ *and* $\textsc{Reg}_{\text{all}}^{(3)}(T) = \mathcal{O}(\sqrt{T}\log^3 T)$.

The key is Subroutine 1 (subroutine is an algorithm that starts at some round t_0, inputs the history $\mathcal{H}_{t_0}$, and controls the subsequent rounds until it exits). For an interval I with $D(I) = \Pr_{v\sim F}\{v \in I\}$, the outcome of an interval query on I has a variance at most $D(I)$. Therefore, by the empirical Bernstein inequality [MP09], Subroutine 1 estimates $D(I)$ up to ϵ using $\widetilde{\mathcal{O}}(D(I)\epsilon^{-2} + \epsilon^{-1})$ queries.

We then estimate the CDF by building a segment tree over $[0, 1)$: Let $N := 2^{\lceil \log_2 \epsilon^{-1} \rceil}$ so that $\frac{1}{N} \leq \epsilon$. In each depth $h = 0, 1, \ldots, \log_2 N$, the tree has the 2^h intervals $[\frac{k}{2^h}, \frac{k+1}{2^h})$. Each non-leaf interval has its two halves as its children, decomposing any $[0, \frac{i}{N})$ as a disjoint union of $\mathcal{O}(\log_2 N)$ tree intervals.

Thus, estimating the probability of every tree interval up to $\frac{\epsilon}{2\log_2(2N)}$ estimates every $\Pr_{v\sim F}\{v < \frac{i}{N}\}$ up to ϵ, giving a value distribution $\widehat{F}$ that is ϵ-close to F in Lévy distance. As the tree intervals of the same depth are disjoint, their probabilities $\sum_{k=0}^{2^h-1} D([\frac{k}{2^h}, \frac{k+1}{2^h}))$ sum to 1. Hence each of the $\log_2 N$ depths takes $\widetilde{\mathcal{O}}(\epsilon^{-2})$ queries in total, proving the first statement of Theorem 3.2. We defer the full pseudocode (Algorithm 5) and how to build segment trees (Subroutine 4) to Appendix C.

The second statement follows from Lemma 3.1: By setting two p's as the two endpoints of an interval query, each with a negligible probability $\gamma = T^{-2}$, and the other one as $\max\{\widehat{p}^* - \epsilon, 0\}$, where $\widehat{p}^*$ is the monopoly price of the current estimated CDF $\widehat{F}$, with a probability $1 - 2\gamma$, we reduce size-3 revenue maximization to CDF estimation with interval queries, thus giving $\textsc{Reg}_{\text{all}}^{(3)}(T) = \widetilde{\mathcal{O}}(\sqrt{T})$.

3.3 Size-2 Menus and Interval-Query Complexity of the Monopoly Price

Focusing on dynamic pricing with size-2 menus, we specify our main algorithm and prove a key technical guarantee which underpins our theorems. Internally, our algorithm is based on monopoly-

price estimation via interval queries, as CDF estimation is wasteful for our purposes (recall Table 2).

To estimate p^*, it suffices to consider the near-optimal regions of the price space. If two prices $p < p'$ are both Δ-optimal, i.e., $R(p), R(p') \geq R(p^*) - \Delta$, the one-sided Lipschitzness of R (Eq. (1)) gives

$$D_F([p, p')) = \Pr_{v \sim F}\{p \leq v < p'\} \leq \frac{p' - p + \Delta}{p}.$$

Hence, two close Δ-optimal prices must have little probability mass in between, so the interval query between them has a small variance. This recovers the key to the improvement in Theorem 3.2.

Utilizing this, our Algorithm 2 runs in layers with halving $\Delta = \frac{1}{2}, \frac{1}{4}, \ldots$. At the beginning of each layer Δ, it maintains a set $\mathcal{S}_{2\Delta}$ of candidate prices, which are all 4Δ-optimal and at least Δ apart from each other. We also ensure one of them lies in $[p^* - \Delta, p^*]$, i.e., $p^* \in \bigcup_{p \in \mathcal{S}_{2\Delta}} [p, p + \Delta]$. Then:

- **Estimate.** Write out the candidate prices $\mathcal{S}_{2\Delta} = \{p_0, p_1, \ldots, p_N\}$ in increasing order. As in Section 3.2, build a segment tree $\mathcal{T}$ over $\mathcal{S}_{2\Delta}$ and estimate the probability mass of every tree interval via empirical Bernstein (precision is specified later). Also estimate the mass above p_N.
- **Refine.** For every candidate $p_i \in \mathcal{S}_{2\Delta}$, add both p_i and $p_i + \frac{\Delta}{2}$ to $\mathcal{S}_\Delta$, and estimate the mass of $[p_i, p_i + \frac{\Delta}{2})$. Together with the tree interval estimates above, using a segment tree decomposition $\mathcal{I}_\mathcal{T}$, we estimate the revenue difference between any two candidate prices up to a precision $\frac{\Delta}{4}$.
- **Eliminate.** Remove those $p \in \mathcal{S}_\Delta$'s whose estimated revenue is Δ-below the best candidate. The remaining prices are 2Δ-optimal, $\frac{\Delta}{2}$ apart from each other, and contain a price in $[p^* - \frac{\Delta}{2}, p^*]$.

This design not only improves the interval-query complexity (by ignoring sub-optimal prices), but also saves a menu option since all interval queries have their left endpoints in $\mathcal{S}_{2\Delta}$. Therefore, we eliminate the need for a dedicated exploitation price: When performing an interval query $[p_i, p_j)$ with $p_i < p_j$ and $p_i \in \mathcal{S}_{2\Delta}$, by Lemma 3.1, the size-2 menu built from the price distribution $(1-\gamma)\delta_{p_i} + \gamma\delta_{p_j}$ with $\gamma = T^{-2}$ reveals which of $[0, p_i), [p_i, p_j), [p_j, 1]$ contains v_t, and collects revenue $(1-\gamma)R(p_i) + \gamma R(p_j) \geq R(p^*) - 4\Delta - \gamma$. Therefore, every round in layer Δ brings regret $\mathcal{O}(\Delta)$.

It remains to count the number of interval queries needed for each layer Δ. In Eq. (5), the first two terms come from the **Estimate** step, where $\mathcal{I}_\mathcal{T}[p_i, p_N)$ consists of $\mathcal{O}(\log N)$ segment tree intervals (Lemma C.3), and the last term comes from the **Refine** step. As the **Eliminate** step only compares two candidates $p \in \{p_i, p_i + \frac{\Delta}{2}\}$ and $p' \in \{p_j, p_j + \frac{\Delta}{2}\}$ with $p_i, p_j \in \mathcal{S}_{2\Delta}$, we only need to estimate all $\widehat{R}(p') - \widehat{R}(p)$ up to $\frac{\Delta}{4}$. This leads to our choices of precisions when calling EmpiricalBernstein in Lines 7, 8, 11, and 12. They bound the number of rounds for each layer Δ, in the following lemma.

Lemma 3.3 (Number of Rounds). *For each $\Delta = 2^{-k}$, Algorithm 2 spends $\widetilde{\mathcal{O}}((p_N - p_0 + \Delta)/\Delta^2)$ rounds where p_0, p_N are both 4Δ-optimal prices, and incurs $\mathcal{O}(\Delta)$ regret in each round. Consequently, for any given T, the regret of Algorithm 2 is upper bounded as*

$$\widetilde{\mathcal{O}}\left(\sup_{\Delta = 2^{-k}, k \geq 0} \min\left\{T\Delta, \frac{\operatorname{diam}\{p \in [0,1] \mid R(p) \geq R(p^*) - \Delta\}}{\Delta}\right\}\right), \tag{4}$$

where $\operatorname{diam}\{A\} := \sup\{A\} - \inf\{A\}$ is the diameter of a set A.

Proof Sketch. Given our choice of empirical Bernstein precisions, estimating each $[p_i, p_j) \in \mathcal{T}$ only takes $\widetilde{\mathcal{O}}(\frac{p_i D([p_i, p_j))}{\Delta^2} + \frac{1}{\Delta})$ queries. Since $p_i D([p_i, p_j)) \leq p_j - p_i + R(p_i) - R(p_j)$ from Eq. (1), the first terms telescope to $p_N - p_0 + R(p_0) - R(p_N) \leq p_N - p_0 + 4\Delta$ over each depth (as $p_0, p_N \in \mathcal{S}_{2\Delta}$ are both 4Δ-optimal); and all those Δ^{-1} terms sum to $\widetilde{\mathcal{O}}(\frac{p_N - p_0 + \Delta}{\Delta^2})$ because all candidates in $\mathcal{S}_{2\Delta}$ are

Algorithm 2 Revenue Maximization via Size-2 Menus

1: Let $\overline{p} \leftarrow$ ScaleSearch in Subroutine 7. From now on, every price p is offered as $\overline{p} \times p$, and F now denotes the distribution of $\min\{\frac{v}{\overline{p}}, 1\}$ with $v \sim F$ ▷ *for an edge case in Footnote 9*
2: Initialize $\mathcal{S}_1$ as the grid $\{0, \frac{1}{2}, 1\}$, such that each $p \in \mathcal{S}_1$ is 1-optimal and p^* is $\frac{1}{2}$-close to $\mathcal{S}_1$
3: **for** $\Delta = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ **do**
4: Write out $\mathcal{S}_{2\Delta} = \{p_0, p_1, \ldots, p_N\}$ in increasing order ▷ *all $p \in \mathcal{S}_{2\Delta}$'s are 4Δ-optimal*
5: Build a segment tree by $(\mathcal{T}, \mathcal{I}_{\mathcal{T}}) \leftarrow$ BuildSegTree$(p_0, p_1, \ldots, p_N)$ ▷ *see Subroutine 4*
6: **for** each tree interval $I = [p_i, p_j) \in \mathcal{T}$ **do**
7: Estimate $\widehat{D}(I) \leftarrow$ EmpiricalBernstein$(I, \frac{\Delta}{80(\log_2 N+2)(p_i+\Delta)}, T^{-3})$ ▷ *see Subroutine 1*
8: Estimate $\widehat{D}(p_N) \leftarrow$ EmpiricalBernstein$([p_N, 1], \frac{\Delta}{8(p_N-p_0+\Delta)}, T^{-3})$
9: Initialize the new set of candidate prices $\mathcal{S}_\Delta := \{p_0, p_0+\frac{\Delta}{2}, p_1, p_1+\frac{\Delta}{2}, \ldots, p_N, p_N+\frac{\Delta}{2}\} \cap [0,1]$
10: **for** $i = 0, 1, \ldots, N-1$ **do**
11: Estimate $\widehat{D}([p_i, p_i+\frac{\Delta}{2})) \leftarrow$ EmpiricalBernstein$([p_i, p_i+\frac{\Delta}{2}), \frac{\Delta}{40(p_i+\Delta)}, T^{-3})$
12: Estimate $\widehat{D}([p_N, p_N+\frac{\Delta}{2})) \leftarrow$ EmpiricalBernstein$([p_N, p_N+\frac{\Delta}{2}), \frac{\Delta}{40(p_N+\Delta)}, T^{-3})$, but terminate this call once $\widehat{\mu}_K - r_K > \frac{2\Delta}{p_N}$ in Subroutine 1 (then remove $p_N+\frac{\Delta}{2}$ from $\mathcal{S}_\Delta$)
13: **for** each candidate price $p \in \mathcal{S}_\Delta$ **do**
14: Estimate its revenue by Eq. (5) with $p_i := \max\{p' \in \mathcal{S}_{2\Delta} \mid p' \le p\}$, where the three terms come from Line 8, Line 7, and Lines 11 and 12 (the decomposition $\mathcal{I}_{\mathcal{T}}$ is defined in Subroutine 4

$$\widehat{R}(p) := p\left(\widehat{D}(p_N) + \sum_{I \in \mathcal{I}_{\mathcal{T}}[p_i, p_N)} \widehat{D}(I) - \widehat{D}([p_i, p))\right) \tag{5}$$

15: Eliminate all $p \in \mathcal{S}_\Delta$'s with $\max_{p' \in \mathcal{S}_\Delta} \widehat{R}(p') - \widehat{R}(p) \ge \Delta$ ▷ *all $p \in \mathcal{S}_\Delta$'s are 2Δ-optimal*

Δ apart. The same telescoping calculation also holds for those $[p_i, p_i+\frac{\Delta}{2})$ with $i < N$. The last interval $[p_N, p_N+\frac{\Delta}{2})$ cannot be absorbed in $(p_N - p_0)$ and thus requires extra care: In Line 12, if we are confident that $p_N+\frac{\Delta}{2}$ earns less than p_N, we directly discard it. If this never happens, then $D([p_N, p_N+\frac{\Delta}{2})) \le \frac{4\Delta}{p_N}$, and the number of queries is then $\widetilde{\mathcal{O}}(\frac{1}{\Delta})$. Otherwise, after $\widetilde{\mathcal{O}}(\frac{1}{\Delta})$ rounds, empirical Bernstein gives a multiplicatively close confidence bound, thus triggering the discard. □

3.4 Upper Bounds for Dynamic Pricing with Size-2 Menus

Building upon Lemma 3.3, in Theorem 3.4, we first derive several minimax regret upper bounds for Algorithm 2. We will further supply an instance-dependent upper bound in Theorem 3.5.

Theorem 3.4 (Upper Bounds)**.** *For dynamic pricing with size-2 menus, Algorithm 2 ensures:*

(a) $\mathrm{Reg}^{(2)}_{\mathrm{all}}(T) = \mathcal{O}(\sqrt{T}\log^4 T)$;
(b) $\mathrm{Reg}^{(2)}_{\mathrm{MHR}}(T) = \mathcal{O}(T^{1/3}\log^4 T)$;
(c) $\mathrm{Reg}^{(2)}_{\mathrm{curv}(\kappa)}(T) = \mathcal{O}((\kappa^{-1/3}T^{1/3} + \log T)\log^4 T)$ *for every curvature* $\kappa \in (0, 2)$.

Theorem 3.4 implies interval-query complexity $\textsc{Interval}_{\mathrm{MHR}}(\epsilon) = \widetilde{\mathcal{O}}(\epsilon^{-1.5})$ and $\textsc{Interval}_{\mathrm{curv}(\kappa)}(\epsilon) = \widetilde{\mathcal{O}}(\kappa^{-1/2}\epsilon^{-1.5})$, matching the corresponding sample complexities (Theorem B.10), whereas we know that $\textsc{Threshold}_{\mathrm{MHR}}(\epsilon) = \widetilde{\Theta}(\epsilon^{-2})$ [LSTW23b] and $\textsc{Threshold}_{\mathrm{curv}(\kappa)}(\epsilon) = \Omega(\epsilon^{-2.5})$ (Lemma B.28).

Looking at Eq. (4), for any F and $\Delta \in (0, 1]$, let $N_F(\Delta) := \lceil \mathrm{diam}\{p \in [0,1] \mid R(p) \ge R(p^*) - \Delta\}/\Delta \rceil$ be the Δ-covering number of the convex hull of all Δ-optimal prices of F. Inspired by the zooming

dimension in Lipschitz bandits [KSU19; BMSS11], we define the interval dimension as follows:

Definition 1. F has interval dimension $d \in [0,1]$ with multiplier C if $N_F(\Delta) \le C\Delta^{-d}$, $\forall \Delta \in (0,1]$.[8]

Unlike Lipschitz bandits, $N_F(\Delta)$ covers the *convex hull* of the Δ-optimal prices rather than the prices themselves. The two concepts coincide when R_F is unimodal (e.g., $F \in \mathcal{F}_{\text{regular}}$). But as we will see shortly in Remark 2, this distinction is necessary rather than an artifact of our analysis. Plugging N_F into Lemma 3.3, we derive the following instance-dependent regret upper bound.

Theorem 3.5 (Instance-Dependent Regret Upper Bound)**.** *For any value distribution F and any $T \ge 2$, the regret of Algorithm 2 is $\mathcal{O}(V_T(F)\log^5 T)$, where*

$$V_T(F) := \sup_{\Delta \in (0,1]} \min\{T\Delta, N_F(\Delta)\}, \tag{6}$$

which implies regret $\mathcal{O}(CT^{d/(d+1)}\log^5 T)$ on any F with interval dimension d and multiplier C.

Proof Sketch of Theorems 3.4 and 3.5. Since we always have $p_N - p_0 \le 1$, the number of rounds spent on layer Δ is $\widetilde{\mathcal{O}}(\Delta^{-2})$ by Lemma 3.3. As the per-round regret is $\mathcal{O}(\Delta)$, this gives $\widetilde{\mathcal{O}}(\sqrt{T})$ regret in the worst case. For MHR or curvature distributions, the near-optimal prices concentrate around p^*: With curvature κ, all Δ-optimal prices are within $\sqrt{\Delta/\kappa}$ of p^* by definition; for MHR distributions, they are also within $4\sqrt{\Delta}$ of p^* when $\Delta \le \frac{R(p^*)}{4}$ (Lemma C.9; a price-space analog of [HMR15, Lemma 3.2]). Hence $p_N - p_0 = \mathcal{O}(\sqrt{\Delta})$, and layer Δ spends $\widetilde{\mathcal{O}}(\Delta^{-1.5})$ rounds, thus giving a refined regret upper bound of $\widetilde{\mathcal{O}}(T^{1/3})$.[9] For any F, the $\mathcal{O}(V_T(F)\log^5 T)$ bound follows from Lemma 3.3, since $\operatorname{diam}\{p \in [0,1] \mid R(p) \ge R(p^*) - \Delta\} \le \Delta \cdot N_F(\Delta)$ by definition. This consequently proves Theorem 3.5, as interval dimension d with multiplier C ensures $V_T(F) \le CT^{d/(d+1)}$: Every $\Delta \le T^{-1/(d+1)}$ has $T\Delta \le T^{d/(d+1)}$, and every other Δ has $N_F(\Delta) \le C\Delta^{-d} < CT^{d/(d+1)}$. □

4 Lower Bounds: From Minimax to Instance-Optimality

We introduce the minimax lower bounds in Section 4.1. Our main lower bounds are in Section 4.2, showing the near-instance-optimality of Algorithm 2: Around every value distribution F, no algorithm, even with infinite menus, can significantly beat Algorithm 2 on a few *decoy instances* close to F. We overview how to construct these decoy instances in Section 4.3. We develop anytime lower bounds in Section 4.4, and compare dynamic pricing with Lipschitz bandits and experts in Section 4.5.

4.1 Minimax Lower Bounds

Theorem 4.1 (Minimax Lower Bounds)**.** *For dynamic pricing with infinite menus and full feedback:*

(a) $\operatorname{Reg}^{(\infty)}_{\text{all}}(T) = \Omega(\sqrt{T})$;
(b) $\operatorname{Reg}^{(\infty)}_{\text{MHR}}(T) = \Omega(T^{1/3})$;

[8] $d = 0$ for a point mass at $\frac{1}{2}$, where R drops linearly; $d = \frac{1}{2}$ for MHR distributions (e.g., Unif$[0,1]$) and curvature distributions, where R drops quadratically around p^*; and $d = 1$ for the equal-revenue distribution $\Pr\{v \ge p\} = \frac{1}{2p}$ on $[\frac{1}{2}, 1]$, where R does not drop at all. Appendix C.6 gives more economically natural examples, such as discrete values, kinked demands, Pareto distributions. In that section, we also give a family of value distributions with every $d \in [0,1)$.

[9] There is an edge case to take care of: Whenever $\Delta \ge R(p^*)$, every price is Δ-optimal, so the early layers always refine the whole $[0,1]$ rather than one with width $\sqrt{\Delta}$, thus bringing an unnecessary term of order $\Delta^{-1} = 1/R(p^*)$ (e.g., $\sqrt{T}$ for Unif$[0, T^{-1/2}] \in \mathcal{F}_{\text{MHR}}$). To avoid this, we first find a $\overline{p} \ge p^*$ by posting the prices $1, \frac{1}{2}, \frac{1}{4}, \ldots$ in turn (which we call ScaleSearch; Subroutine 7), and then only zoom into the smaller interval $[0, \overline{p}]$. This gives a $\overline{p} < 10p^*$ for MHR distributions while only incurring $\mathcal{O}(\log^2 T)$ regret, thus avoiding the unnecessary $1/R(p^*)$. See Appendix C.3.

(c) $\mathrm{Reg}^{(\infty)}_{\mathrm{curv}(\kappa)}(T) = \Omega(\kappa^{-1/3}T^{1/3})$ *for every curvature* $\kappa \in (0,2)$*;*
(d) *the minimax regret over value distributions with interval dimension* $d \in (0,1)$ *is* $\Omega(T^{d/(d+1)})$.

Hence, all rates in Theorem 3.4 are minimax-optimal even with infinite menus. We defer the (mostly standard) proofs to Appendix D. A few remarks before moving on to our main lower bounds:

- Result (d) can also be extended to an *anytime* lower bound, which we will detail in Section 4.4.
- Our hard instances in (a) are regular, hence $\mathrm{Reg}^{(\infty)}_{\mathrm{regular}}(T) = \Omega(\sqrt{T})$, which is as hard as $\mathcal{F}_{\mathrm{all}}$ and strictly harder than $\mathcal{F}_{\mathrm{MHR}}$. This is in contrast to dynamic pricing with posted prices, where $\mathrm{Reg}^{(1)}_{\mathrm{regular}}(T) = \widetilde{\Theta}(\sqrt{T})$ [SW24; LSTW23b], thus strictly easier than $\mathcal{F}_{\mathrm{all}}$ and as easy as $\mathcal{F}_{\mathrm{MHR}}$.
- Our hard instances in (a) are also *quasi-MHR*, a broader-than-MHR class introduced in [FJ25], thus answering their open question on learning reserves for quasi-MHR buyers (see Remark 8).
- As depicted in Table 1, Theorems 3.4 and 4.1 together show that size-2 menus are nearly as good as infinite menus (up to log factors) for all these classes of value distributions. On the other hand, algorithms using only posted prices are polynomially worse for all these classes.
- All these results also imply lower bounds on sample complexities. Hence, for both CDF estimation and monopoly price estimation, interval queries are as powerful as samples (up to log factors), while being strictly stronger than threshold queries (see Table 2; detailed in Appendix B).

4.2 Near Instance-Optimality of Algorithm 2

Are size-2 menus as good as infinite ones in general, or just for these specific classes in Theorem 4.1? We claim that, on *every* possible value distribution F, size-2 menus near-match infinite menus.

To justify such an "instance-optimality" claim, ideally, one would want to lower bound any algorithm's regret on this F alone. But this is impossible, because the algorithm always posting p^*_F (somehow "cheating" by giving up on all other instances) has zero regret on F. A natural rescue, known as the Lai–Robbins-type [LR85], is to only consider algorithms that are good on all instances (e.g., with $o(T)$ regret). But it fails for our full-feedback problem since exploration is free: One can keep playing p^*_F until the empirical CDF leaves the confidence ball of F [DKW56], and then switch to Algorithm 2. This gives $\mathcal{O}(1)$ regret on F while remaining $\widetilde{\mathcal{O}}(\sqrt{T})$ on every other value distribution. Instead, we construct a local-minimax-type lower bound [CL15; ZCDL16] in Theorem 4.2.

Theorem 4.2 (Local Minimax Lower Bound)**.** *Let* F *be such that* $R_F(p^*_F) > 0$ *and* $T \geq 2$*, and let* $c_1 := \frac{R_F(p^*_F)^{12}}{2^{45}}$, $c_2 := \frac{325}{R_F(p^*_F)^3}$, *and* $c_3 := \frac{2^{20}}{R_F(p^*_F)^6}$. *There are value distributions* G_+, G_- *such that:*

(a) *Any algorithm, even with infinite menus, suffers regret at least* $c_1 V_T(F)$ *on* G_+ *or on* G_-*;*
(b) $V_T(G_\pm) \leq c_2 V_T(F)$*; and*
(c) $H^2(F, G_\pm) \leq \frac{c_3}{V_T(F)^2}$ *where* $H^2(F,G) := \frac{1}{2}\int(\sqrt{dF} - \sqrt{dG})^2$ *is the squared Hellinger distance.*

Comparing with Theorem 3.5, we conclude the *near instance-optimality* of Algorithm 2: ***No algorithm, even using infinite menus, beats Algorithm 2 simultaneously on these converging-to-*** F ***decoy instances*** beyond log factors. We summarize it as Corollary 4.3, which is much stronger than minimax lower bounds that only evaluate the algorithm on a few *specific* hard instances.

Corollary 4.3 (Near Instance-Optimality of Algorithm 2)**.** *Under the conditions of Theorem 4.2,*

1. *Algorithm 2 has regret* $\mathcal{O}(V_T(F)\log^5 T)$ *on* F *and* $\mathcal{O}(R_F(p^*_F)^{-3}V_T(F)\log^5 T)$ *on* G_+ *and on* G_-.
2. *Any algorithm, even with infinite menus, suffers* $\Omega(R_F(p^*_F)^{12}V_T(F))$ *regret on* G_+ *or on* G_-.
3. *If* F *does not have interval dimension* 0*, then* $H^2(F, G_\pm) \to 0$ *as* $T \to \infty$ *(see Lemma C.20). Otherwise, Algorithm 2 attains polylogarithmic regret on* F*, hence trivially near instance-optimal.*

Remark 2. Theorem 4.2 implies that $N_F(\Delta)$ must cover the convex hull of all Δ-optimal prices (as in Definition 1), not just these prices themselves (as in zooming dimension of Lipschitz bandits). Indeed, value distribution $F = \frac{1}{2}\delta_{1/2} + \frac{1}{2}\delta_1$ has interval dimension 1 but zooming dimension 0. For all large T, there are two close-to-F value distributions $G_\pm$, both also bimodal with zooming dimension 0, where any algorithm, even with infinite menus, must suffer $\Omega(V_T(F)) = \Omega(\sqrt{T})$ regret.

While general, Corollary 4.3 is a bit away from being perfect: The decoy instances $G_\pm$ do not include F, which does not exclude the possibility that F is actually much easier; and $G_\pm$, while close to F in Hellinger distances, no longer enjoy the good properties of F (e.g., interval dimensions or MHR).

In Theorem 4.4, we resolve them when $N_F(\Delta) = \Theta(\Delta^{-d})$ for all Δ small enough. For T large enough, there now exists a decoy instance G with interval dimension $\le d$ such that $\mathrm{Reg}^{(\infty)}_{\{F,G\}}(T) = \Omega(T^{d/(d+1)})$. Since Algorithm 2 attains $\mathcal{O}(T^{d/(d+1)}\log^5 T)$ regret on both F and G (Theorem 3.5), we can sharpen the near instance-optimality to: ***No algorithm, even using infinite menus, beats Algorithm 2 by*** $\omega(\log^5 T)$ ***factors simultaneously on*** F ***and a local*** G (which converges to F as $T \to \infty$).

Theorem 4.4 (Near Instance-Optimality via Interval Dimensions)**.** *If F has interval dimension $d \in (0,1)$ with multiplier C, there is a constant $c > 0$ only depending on $R_F(p^*_F), C$, and d. For any $T \ge \frac{1}{c}$ satisfying $N_F(\Delta) \ge \frac{\Delta^{-d}}{C}$ with $\Delta = T^{-1/(d+1)}$, there is a value distribution G that satisfies:*

(a) *Any algorithm, even with infinite menus, suffers regret at least $cT^{d/(d+1)}$ on F or on G;*
(b) *G also has interval dimension d with multiplier C/c; and*
(c) *$H^2(F,G) \le \frac{1}{cT}$.*

We also refine Theorem 4.2 to MHR distributions, motivated by the fact that interval (or zooming) dimensions are not the most suitable metric for MHR distributions: In dynamic pricing with posted prices, due to the $\frac{1}{2}$ zooming dimension, MHR distributions seem to bear a $\widetilde{\Theta}(T^{3/5})$ rate [KSU19]; but their minimax regret is actually $\widetilde{\mathcal{O}}(\sqrt{T})$ [SW24]. With menus, while our Theorem 4.4 indeed suggests an $\Omega(T^{1/3})$ lower bound, the decoy instance G may not be MHR. We resolve this in Theorem 4.5.

Theorem 4.5 (Near Instance-Optimality for MHR Distributions)**.** *Let F be MHR. If on an open interval containing p^*_F, its hazard rate function $h(v) = \frac{f(v)}{D_F(v)}$ is differentiable with h' bounded by some $[c, C]$ with $0 < c < C$, then for every $T > 0$, there is an MHR distribution G with $H^2(F,G) \le \frac{1}{64T}$, such that any algorithm, even with infinite menus, suffers regret $\mathrm{Reg}^{(\infty)}_{\{F,G\}}(T) = \Omega(T^{1/3})$.*

4.3 Constructions of Decoy Instances

We overview our constructions of the decoy instances in Theorem 4.2, and how to refine them in Theorems 4.4 and 4.5. We remark that while local minimax type of lower bounds exist in the statistics or learning literature [CL15; ZCDL16], their techniques do not translate to our setting. Instead, our constructions are specifically tailored towards dynamic pricing and are highly non-trivial.

Proof Sketch of Theorem 4.2. Throughout, all non-subscripted p^*, R, and N stand for those of F, and all $\mathcal{O}$, Ω, Θ notations treat $R(p^*)$ and p^* as constants. While we still use the two-point method (standard for minimax bounds), we instead construct decoy instances locally around F.

Let Δ be such that $\min\{T\Delta, N(\Delta)\} \ge \frac{1}{2}V_T(F)$ (Eq. (6)), and assume F has a probability density function (PDF) f for the ease of presentation. Let $\ell \le u$ be the smallest and the largest Δ-optimal prices of F, and call $R_G(p^*_G) - R_G(p)$ the *gap* of a price p under a value distribution G. It suffices to find two value distributions, both close to F, such that every price has a gap $\Omega(\Delta)$ under one of

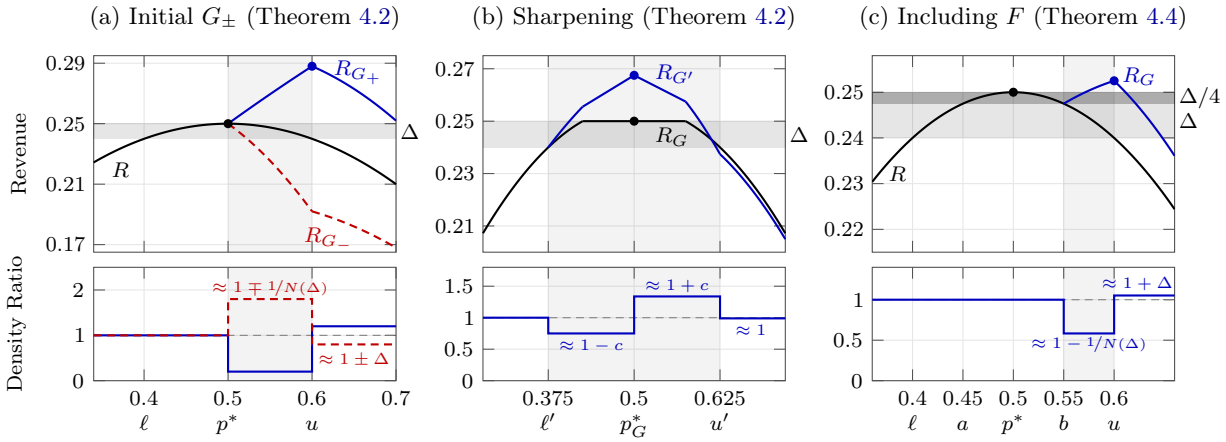


Figure 2: Constructions in Theorems 4.2 and 4.4 (under $F = \text{Unif}[0,1]$ and $\Delta = 0.01$)

them. We do so by moving $\Theta(\Delta)$ probability mass from (a sub-interval of) $[\ell, u]$ to $[u, 1]$, so that price u gains $\Theta(\Delta)$ revenue relative to ℓ (or vice versa, giving two distributions G_+ and G_-). How to ensure $G_\pm$ are close to F? If this sub-interval has total probability mass m in F, then this move rescales f by $1 \mp \Theta(\frac{\Delta}{m})$. Hence $G_\pm$ are indistinguishable for $\Theta(\frac{m}{\Delta^2})$ rounds via the Hellinger distance, giving regret $\Omega(\min\{T\Delta, \frac{m}{\Delta}\})$. Therefore, finding an interval with $m = \Omega(\Delta \cdot N(\Delta))$ suffices. This is made possible by the other direction of Eq. (1), namely far apart prices have large mass in between:

$$\begin{aligned} D_F([\ell, u)) &\overset{(a)}{=} \frac{(u-\ell)D_F(u) - (R(u) - R(\ell))}{\ell} \geq \frac{1}{\ell}\left(\frac{R(p^*)}{2}(u-\ell) - \Delta\right) \\ &\overset{(b)}{\geq} \frac{1}{\ell}\left(\frac{R(p^*)}{2}(N(\Delta)-1)\Delta - \Delta\right) \overset{(c)}{\geq} \frac{R(p^*)}{4}\Delta \cdot N(\Delta) = \Omega(\Delta \cdot N(\Delta)), \end{aligned}$$

whenever $N(\Delta) \geq 2 + \frac{4}{R(p^*)} = \Theta(1)$, where (a) uses Eq. (1), (b) uses the definition that $(u - \ell) \geq (N(\Delta) - 1)\Delta$, and (c) uses $N(\Delta) \geq 2 + \frac{4}{R(p^*)}$ and $\ell \leq 1$. Suppose that $[p^*, u)$ has at least half of this mass (the case of $[\ell, p^*)$ is similar). Construct G_+ and G_-, with PDFs g_+ and g_-, as follows:

$$g_\pm := \begin{cases} f & \text{on } [0, p^*), \\ \left(1 \mp \dfrac{4\Delta/p^*}{D_F([p^*, u))}\right) f & \text{on } [p^*, u), \\ \left(1 \pm \dfrac{4\Delta/p^*}{D_F(u)}\right) f & \text{on } [u, 1]. \end{cases}$$

That is, G_+ moves a total mass $\theta := 4\Delta/p^*$ from $[p^*, u)$ to $[u, 1]$ by rescaling f on both intervals, and G_- does the opposite; see Figure 2 (a) for an illustration. They are indeed close to F:

$$H^2(F, G_\pm) \leq \int_0^1 \frac{(g_\pm(v) - f(v))^2}{f(v)} dv = \frac{\theta^2}{D_F([p^*, u))} + \frac{\theta^2}{D_F(u)} = \mathcal{O}\left(\frac{\Delta}{N(\Delta)} + \Delta^2\right) = \mathcal{O}\left(\frac{1}{V_T(F)^2}\right),$$

where we used $\theta = \Theta(\Delta)$, $D_F([p^*, u)) = \Omega(\Delta \cdot N(\Delta))$, and $D_F(u) = \Theta(1)$. By direct calculation,

$$R_{G_\pm}(p) = R(p) \pm \theta p \cdot \frac{D_F([p^*, p))}{D_F([p^*, u))}, \quad \forall p \in [p^*, u].$$

Hence under G_+, the price u gains $\theta u \geq 4\Delta$, whereas every $p \in [p^*, u]$ with $D_F([p^*, p)) \leq \frac{1}{2} D_F([p^*, u))$ gains at most $\frac{\theta u}{2}$ and thus has a gap at least $\frac{\theta u}{2} - \Delta \geq \Delta$. Under G_-, p^* keeps its revenue, whereas every $p \in [p^*, u]$ with $D_F([p^*, p)) \geq \frac{1}{2} D_F([p^*, u))$ loses at least $\frac{\theta p^*}{2} = 2\Delta$ (hence a gap at least Δ). Finally, no $p > u$ is Δ-optimal under F and hence under G_-; and no $p < p^*$ is Δ-optimal under G_+. Hence, every $p \in [0, 1]$ has a gap at least Δ under G_+ or under G_-, enabling the two-point method.

However, $V_T(G_\pm)$ can be much larger than $V_T(F)$: If R_F decreases linearly at a suitable slope on the upper part of $[p^*, u]$, then G_+ exactly cancels this decrease, so that R_{G_+} is maximized on a plateau of length $\Theta(\Delta \cdot N(\Delta))$ (see Figure 2 (b) for an illustration). We thus perform a *sharpening* step on each $G \in \{G_+, G_-\}$. Let p^*_G be the monopoly price under G, and $\ell' \leq p^*_G \leq u'$ be the smallest and the largest Δ-optimal prices under G. Construct value distribution G' with PDF g' as follows:

$$g' := \begin{cases} g & \text{on } [0, \ell'), \\ (1-c)\, g & \text{on } [\ell', p^*_G), \\ (1+c_1)(1+c)\, g & \text{on } [p^*_G, u'), \\ (1+c_1)(1-c_2)\, g & \text{on } [u', 1], \end{cases}$$

where $c := \min_{\delta \in [\Delta, 1]} \frac{1}{N_F(\delta)}$, and $c_1 := c \cdot \frac{D_G([\ell', p^*_G))}{D_G(p^*_G)}$ and $c_2 := c \cdot \frac{D_G([p^*_G, u'))}{D_G(u')}$ keep the total mass unchanged (both c_1 and c_2 are $\mathcal{O}(\Delta)$). That is, by rescaling g on different intervals, G' moves a c fraction of the mass on $[\ell', p^*_G)$ to $[p^*_G, 1]$, and then scales up $[p^*_G, u')$ by $(1+c)$, taking the extra mass from $[u', 1]$. Then $R_{G'} = R_G$ on $[0, \ell']$, p^*_G remains optimal, no gap decreases, and

$$R_{G'}(p^*_G) - R_{G'}(p) \geq R_G(p^*_G) - R_G(p) + c\, p^*_G\, |D_G(p) - D_G(p^*_G)| = \Omega\left(c |p - p^*_G|\right), \quad \forall p \in [\ell', u'],$$

where the last step applies Eq. (1) to G. Therefore, in both directions of $p^*_{G'} = p^*_G$, our new $R_{G'}$ is decreasing linearly fast throughout the Δ-optimal region $[\ell', u']$. This gives $N_{G'}(\delta) = \mathcal{O}(\frac{1}{c})$ for all $\delta < \Delta$, where $\frac{1}{c} = N_F(\delta')$ for some $\delta' \in [\Delta, 1]$. For $\delta \geq \Delta$, as no gap decreases and $R_G = R_F + \mathcal{O}(\Delta)$, every δ-optimal price under G' is δ'-optimal under F for some $\delta' \geq \delta$, giving $N_{G'}(\delta) = \mathcal{O}(N_F(\delta'))$. Either way, for any $\delta \in (0, 1]$, we have $N_{G'}(\delta) = \mathcal{O}(N_F(\delta'))$ for some $\delta' \geq \delta$. Therefore,

$$\begin{aligned} V_T(G') = \sup_{\delta \in (0,1]} \min\{T\delta, N_{G'}(\delta)\} &\leq \sup_{\delta \in (0,1]} \sup_{\delta' \in [\delta, 1]} \mathcal{O}\left(\min\{T\delta, N_F(\delta')\}\right) \\ &\leq \sup_{\delta' \in (0,1]} \mathcal{O}\left(\min\{T\delta', N_F(\delta')\}\right) = \mathcal{O}\left(V_T(F)\right). \end{aligned}$$

As $g' = (1 \pm \mathcal{O}(c))g$ on $[\ell', u')$ and $g' = (1 \pm \mathcal{O}(\Delta))g$ on $[u', 1]$, similar calculation gives $H^2(F, G') = \mathcal{O}(\frac{\Delta}{N(\Delta)}) = \mathcal{O}(V_T(F)^{-2})$. Hence G'_+ and G'_-, sharpened from G_+ and G_-, are indistinguishable in $\Theta(N(\Delta)/\Delta)$ rounds. Consequently, by the two-point method (Lemma D.1), any algorithm, even with infinite menus, must suffer regret $\Omega(\min\{T\Delta, N(\Delta)\}) = \Omega(V_T(F))$ on one of them. □

Proof Sketch of Theorem 4.4. Compared to the generic lower bound, we now additionally want F itself to replace one of the two decoy instances. Let $\Delta := T^{-1/(d+1)}$, so that $N(\Delta) \geq \frac{\Delta^{-d}}{C}$, and let $a \leq b$ be the smallest and the largest Δ'-optimal prices for $\Delta' := (4C^2)^{-1/(1-d)}\Delta$. Then

$$b - a \leq \Delta' N(\Delta') \leq C\Delta'^{1-d} = \frac{\Delta^{1-d}}{4C} \leq \frac{1}{4}\Delta \cdot N(\Delta).$$

Since $u - \ell = \Omega(\Delta \cdot N(\Delta))$, one of $[\ell, a)$ and $[b, u)$, say $[b, u)$, has length $\Omega(\Delta \cdot N(\Delta))$, and hence probability mass $\Omega(\Delta \cdot N(\Delta))$ by Eq. (1). Similar to Theorem 4.2, let G move a mass $\theta := \frac{\Delta + \Delta'}{u}$

from $[b, u)$ to $[u, 1]$, so $R_G = R$ on $[0, b]$, while $R_G(u) = R(u) + \theta u \geq R(p^*) + \Delta'$ (see Figure 2 (c)). Thus every price in $[0, b]$ has gap at least Δ' under G, every price outside $[a, b]$ has gap more than Δ' under F, and $H^2(F, G) = \mathcal{O}(\frac{\Delta}{N(\Delta)}) = \mathcal{O}(\Delta^{d+1}) = \mathcal{O}(\frac{1}{T})$. Similarly sharpening G gives a G' such that $N_{G'}(\delta) = \mathcal{O}(N_F(\delta'))$ for some $\delta' \geq \delta$, which means G' has interval dimension d. A two-point method (Lemma D.1) on F and G' then yields the desired $\mathrm{Reg}^{(\infty)}_{\{F,G'\}}(T) = \Omega(T^{d/(d+1)})$ lower bound. □

Proof Sketch of Theorem 4.5. Recall that the demand of a value distribution F is $D_F(p) = \Pr_{v\sim F}\{v \geq p\}$. By definition, the hazard rate $h = \frac{f}{D_F} = -(\log D_F)'$, and thus MHR is equivalent to having a concave $\log D_F$. In Theorem 4.2, we rescale the density f by constants on intervals, which creates downward jumps in the hazard rate and breaks MHR. We instead directly work on the log-demand by constructing $\log D_G = \log D_F + \phi$ for some ϕ. Since we want G close to F, we only pick a small interval $[a, b]$ (that contains p^*_F) for ϕ to be non-constant. From the assumption that $h' \geq c$ around p^*_F, in order to make $\log D_F + \phi$ concave, we thus set $\phi(p) = \frac{c}{2}(\min\{p, b\} - a)^2_+$.

It remains to choose a and b such that every price has a large gap under F or G, so the two-point method applies. Let $s = b - p^*_F$. Under F, every price above $p^*_F + \frac{s}{2}$ has a gap $\Omega(s^2)$ as R_F drops quadratically around p^*_F; under G, since $\log R_G = \log p + \log D_G$ is concave, every price below $p^*_F + \frac{s}{2}$ has a gap $\Omega(s^2)$ once $\log R_G(b) - \log R_G(p^*_F + \frac{s}{2}) = \Omega(s^2)$. That is, the increase of ϕ from $p^*_F + \frac{s}{2}$ to b, which is of order $cs(b - a)$, must beat the drop of $\log R_F$ there by $\Omega(s^2)$. Since $h' \leq C$, the drop is only $\mathcal{O}(s^2)$, and thus taking $a := p^*_F - \Theta(s)$ with a large enough constant suffices.

The final step is verifying that G is close to F: The density ratio $\frac{dG}{dF} = (1 - \frac{\phi'}{h})e^{\phi}$ deviates from 1 by $\mathcal{O}(s)$ on $[a, b]$, an interval of length $\Theta(s)$, and by $\mathcal{O}(s^2)$ above b. Thus $H^2(F, G) = \mathcal{O}(s^3)$. Picking $s = \Theta(T^{-1/3})$, the two-point method (Lemma D.1) gives $\Omega(Ts^2) = \Omega(T^{1/3})$ regret on F or on G. □

4.4 Anytime Lower Bounds

The decoy instances in Theorems 4.2, 4.4, and 4.5 are all chosen after seeing T (in this aspect, similar to standard minimax lower bounds). It would be more favorable if a (randomized) instance is hard for infinitely many T's almost surely, which we call *anytime* lower bounds.

In dynamic pricing and the closely related Lipschitz bandits literature, *i)* for curvature and smooth value distributions (see Footnote 7), an $\Omega(\sqrt{T})$ anytime lower bound is available [KL03]; and *ii)* for Lipschitz bandits, $\Omega(T^{(d+1)/(d+2)})$ anytime lower bounds are also derived [KSU19, §5.1]. But they don't directly apply to dynamic pricing with posted prices or infinite menus. Utilizing our construction in Figure 2, we complement the literature by deriving an anytime lower bound for dynamic pricing: We construct a single randomized distribution F, whose interval dimension is a given $d \in (0, 1)$, that forces $\Omega(T^{d/(d+1)})$ with infinite menus and $\Omega(T^{(d+1)/(d+2)})$ with posted prices.

Theorem 4.6 (Anytime Lower Bounds)**.** *For every $d \in (0, 1)$, there is a constant $c > 0$ and a probability distribution μ over value distributions with interval dimension d and multiplier 11, such that for every algorithm, the following holds with probability 1 over the randomized $F \sim \mu$:*

1. *Even with infinite menus, its regret on F is at least $cT^{d/(d+1)}$ for infinitely many T.*
2. *When using posted prices, its regret on F is at least $cT^{(d+1)/(d+2)}$ for infinitely many T.*

Proof Sketch. The construction of Kleinberg and Leighton [KL03] uses a smooth one-parameter family, which can be learned at $\mathcal{O}(\log T)$ regret with infinite menus. We use the multi-scale needle-in-a-haystack idea for Lipschitz bandits [KSU19, §5.1]: Fix an infinite sequence of scales $\Delta_1 > \Delta_2 > \cdots$.

Table 3: Comparison between Dynamic Pricing and Lipschitz Bandits or Experts

Problem	Full Feedback / Infinite Menus	Bandit Feedback / Posted Prices
Dynamic Pricing	$\widetilde{\mathcal{O}}(T^{d/(d+1)})$ using size-2 **(Theorem 3.5)** $\Omega(T^{d/(d+1)})$ **(Theorems 4.2 and 4.4)**	$\widetilde{\mathcal{O}}(T^{(d+1)/(d+2)})$ [KSU19] $\Omega(T^{(d+1)/(d+2)})$ **(Theorem 4.6)**
Lipschitz Bandits	$\widetilde{\mathcal{O}}(\sqrt{T})$ [KSU19] $\Omega(\sqrt{T})$ even if $d = 0$ **(Lemma D.18)**	$\widetilde{\mathcal{O}}(T^{(d+1)/(d+2)})$ [KSU19] $\Omega(T^{(d+1)/(d+2)})$ [Sli14]

Starting from a plateau P_0, each scale k raises a part of the current plateau P_{k-1} by Δ_k. The top of it, whose half-width $w_k = \Delta_{k+1}^{1-d}$, becomes the next plateau P_k. To construct a single instance F that forces infinite menus and posted prices at the same time, we alternate between two types of scales:

1. The first type is a *needle* [KSU19, Theorem 5.7], dividing P_{k-1} into width-Δ_k intervals and raising a uniformly random one of them by Δ_k. For posted pricing, this forces regret $\Omega(T_k\Delta_k)$ in $T_k = \Theta(\frac{w_{k-1}}{\Delta_k^3})$ rounds with probability at least $\frac{1}{2}$. Since $\Delta_k = w_{k-1}^{1/(1-d)}$, we have $T_k\Delta_k = \Theta(T_k^{(d+1)/(d+2)})$. Unlike in Lipschitz bandits, where the needle can be planted anywhere, the result must remain a revenue function, i.e., $D(p) = R(p)/p$ must stay non-increasing, or $D' = (R' - D)/p < 0$. We ensure this by planting all scales on plateaus around $p = \frac{1}{2}$ such that $D(p) \geq 0.41$ and $R'(p) \leq \frac{1}{3}$.
2. The second type is our new construction, which we call a *tilt*. We divide P_{k-1} into two halves, choose one using fresh randomness, and raise it by Δ_k by moving masses within it similar to Figure 2 (a): the revenue rises linearly from both ends of this half, so that the middle part has a plateau P_k that is Δ_k higher. These two halves are indistinguishable with infinite menus for $T_k = \Theta(\frac{w_{k-1}}{\Delta_k^2})$ rounds, thus forcing $T_k\Delta_k = \Omega(T_k^{d/(d+1)})$ regret (where again, $\Delta_k = w_{k-1}^{1/(1-d)}$).

There is a unique challenge due to infinite menus. In posted prices, the finer scales are invisible to the seller unless a query falls into the small plateau P_k that contains them, as the outcomes are otherwise identically distributed. With infinite menus, the seller sees every value (as if a threshold query is placed everywhere), so it could detect the finer scales within the first T_k rounds. Because they sit inside P_k, this would reveal our choice at scale k. We therefore let the scales decay super-exponentially, $\Delta_{k+1} = \Delta_k^q$ for a large constant q, so that all finer scales together change the value distribution by only $\mathcal{O}(\Delta_{k+1})$ in squared Hellinger distance, which is invisible in the first T_k rounds.

For algorithms using posted prices, due to needles, we have $\Omega(T^{(d+1)/(d+2)})$ regret for infinitely many T almost surely, as each needle scale succeeds with probability at least $\frac{1}{2}$ using fresh randomness. For those using infinite menus, the needles are not obstacles: Full feedback reveals a needle within $\widetilde{\mathcal{O}}(\Delta_k^{-1})$ rounds (as if all width-Δ_k intervals receive threshold queries simultaneously). Instead, those tilt scales force $\Omega(T^{d/(d+1)})$ regret for infinitely many T almost surely by an identical argument. □

4.5 Dynamic Pricing Compared to Lipschitz Bandits or Experts

We conclude this section by comparing dynamic pricing to Lipschitz bandits or experts [KSU19]. For Lipschitz bandits over $[0, 1]$ (that is, each round's action is some $x \in [0, 1]$, and each round's reward function f_t is i.i.d. with a 1-Lipschitz expectation), the zooming algorithm attains $\widetilde{\mathcal{O}}(T^{(d+1)/(d+2)})$ regret in terms of the zooming dimension d [KSU19]. This rate cannot be improved for $d \in (0, 1)$, even *i)* in terms of our interval dimension that is always no smaller, *ii)* for dynamic pricing problems alone, and *iii)* on a fixed value distribution F independent of T (see Theorem 4.6). So in terms of zooming or interval dimensions, dynamic pricing is one of the *hardest* Lipschitz bandits.

With full feedback or menus, however, dynamic pricing is strictly easier than Lipschitz experts. Our $T^{d/(d+1)}$ is $o(\sqrt{T})$ whenever $d < 1$, whereas Lipschitz experts suffer $\Omega(\sqrt{T})$ regret even if $d = 0$ (all Δ-optimal arms lie in an interval of length 4Δ; Lemma D.18). Indeed, comparing two arms whose expected payoffs differ by Δ takes about Δ^{-2} times their difference-in-payoff's second moment. Lipschitz experts only require the expected payoffs to be Lipschitz, so two very close arms can still have perfectly negatively correlated payoffs, making this second moment a constant. In pricing, the payoffs are determined by the same v_t, so for any two Δ-optimal prices $p < p'$, Eq. (1) gives

$$\mathop{\mathbb{E}}_{v\sim F}\Big[\big(p\mathbb{1}[v \geq p] - p'\mathbb{1}[v \geq p']\big)^2\Big] = p^2 \Pr_{v\sim F}\{p \leq v < p'\} + (p'-p)^2 \Pr_{v\sim F}\{v \geq p'\} \leq 3\Delta \cdot N(\Delta).$$

Hence comparing the Δ-optimal prices only brings an $N(\Delta)$ regret overhead, rather than the Δ^{-1} in Lipschitz experts. We summarize the comparison in Table 3, where d is the interval dimension in the first row (Definition 1) and the zooming dimension in the second row [KSU19].

5 Extension: Adversarial Dynamic Pricing Using Size-3 Menus

We now consider *adversarial dynamic pricing*, where an adaptive adversary chooses each value $v_t \in [0,1]$ based on the history. With a single posted price, the optimal regret against the best fixed price is $\widetilde{\Theta}(T^{2/3})$ [KL03]. Should v_t be revealed after each round (e.g., via infinite menus), one can achieve regret $\widetilde{\mathcal{O}}(\sqrt{T})$. Therefore, regret rates with size-1 menus and size-∞ menus both match their i.i.d. counterparts, and we ask whether it is still possible to attain $\widetilde{\mathcal{O}}(\sqrt{T})$ regret using small menus. And indeed, we achieve this with size-3 menus, leveraging the approach from Section 3.2.

Preliminaries. Adversarial dynamic pricing keeps most of the setup from the i.i.d. case: in each round t, the seller's algorithm offers a menu M_t based on the history $\mathcal{H}_{t-1}$, the customer with value $v_t \in [0,1]$ chooses an option in M_t maximizing their utility, and the seller observes the chosen option. The revenue in round t, revenue_t, is the expected payment in the chosen menu option. The only difference is that each value v_t is now chosen by an *adaptive adversary*, i.e., arbitrarily, possibly based on history $\mathcal{H}_{t-1}$ and the previous values $v_1, v_2, \ldots, v_{t-1}$, possibly using its own randomness, but without observing the internal randomness of the algorithm in round t. Note that the adversary could tailor its choices to the algorithm's specification. The algorithm maximizes the total expected revenue in the worst case over all adversaries.

This is a special case of *adversarial bandits* [ACFS02], a standard bandit framework in which the arms' rewards in each round are chosen by an adaptive adversary (defined similarly). We adopt a (standard) notion of regret against a given price sequence $p^*_{\texttt{seq}} = (p^*_1, \ldots, p^*_T)$ and adversary $\texttt{ADV}$:

$$\text{Reg}(T \mid p^*_{\texttt{seq}}, \texttt{ADV}) := \mathbb{E}\left[\textstyle\sum_{t\in[T]} \left(p^*_t\, \mathbb{1}[v_t \geq p^*_t] - \text{revenue}_t\right) \;\mid\; \texttt{ADV}\right].$$

We are interested in the worst case across all adversaries and all price sequences in a given class $\mathcal{C}$.

$$\text{Reg}(T) := \sup_{\texttt{ADV},\, p^*_{\texttt{seq}}\in\mathcal{C}} \text{Reg}(T \mid p^*_{\texttt{seq}}, \texttt{ADV}).$$

The basic notion is against the best fixed price, i.e., when $\mathcal{C}$ consists of all fixed-price sequences. Competing against *all* price sequences (a.k.a. *dynamic regret*), results in $\Omega(T)$ regret even with full feedback [Zin03]. We focus on the class $\mathcal{C}_S$ of sequences that "switch" (in the sense of $p^*_t \neq p^*_{t+1}$) at most S times for some known integer $S \geq 0$; in particular, $\mathcal{C}_0$ consists of all fixed-price sequences.

We build on known results from adversarial bandits with full feedback, a.k.a. *online learning*. For intuition on how they might help: one approach to achieve $\widetilde{\mathcal{O}}(\sqrt{T})$ regret for adversarial dynamic pricing with full feedback is to take Hedge [FS97], an algorithm for online learning, and run it over the price grid $\mathcal{S} := \{0, \frac{1}{T}, \frac{2}{T}, \dots, 1\}$, i.e., with prices in $\mathcal{S}$ as actions.

5.1 Adversarial Dynamic Pricing Algorithm using Online Learning

We follow the approach of Section 3.2, which has two key ingredients. First, by Lemma 3.1 (which is applicable to any sequence of realized values), a size-3 menu can exploit one price while asking an interval query via the other two. Second, a segment tree over the prices makes it suffice to estimate the probabilities of its tree intervals. The estimator in Section 3.2, however, fails against adversarial values. We instead estimate all tree intervals at once by importance weighting.

We run an online learning algorithm $\mathcal{A}^{\mathrm{OL}}$ over the grid $\mathcal{S} = \{0, \frac{1}{T}, \frac{2}{T}, \dots, 1\}$. In each round, we use *i)* one menu option for the "exploitation price" suggested by $\mathcal{A}^{\mathrm{OL}}$; and *ii)* two other menu options for an interval query on two "exploration prices", based on which we estimate the revenue of every price in $\mathcal{S}$. Using Lemma 3.1, we put a probability mass $1 - 2\gamma$ (with $\gamma = 1/T$) on the exploitation price, and γ on each of the exploration prices, such that our algorithm's expected revenue is only $\mathcal{O}(1)$ different from always using prices suggested by $\mathcal{A}^{\mathrm{OL}}$. We start with our requirements for $\mathcal{A}^{\mathrm{OL}}$.

Online Learning Algorithm. We allow any $\mathcal{A}^{\mathrm{OL}}$ with the following second-order guarantee:

Definition 3 (Online Learning Algorithm)**.** Let $\mathcal{C} \subseteq [K]^T$ be a class of comparator sequences. We say $\mathcal{A}^{\mathrm{OL}}$ is suitable for Algorithm 3, if for some $\boldsymbol{w}_t \in \triangle[K]$ and $\boldsymbol{q}_t \in [0,1]^K$ that are both functions of $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_{t-1}$, for every sequence of loss vectors $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T$ with $\ell_{t,i} \geq -1$ for all t and i, it ensures

$$\sum_{t=1}^{T} \left(\mathop{\mathbb{E}}_{i \sim \boldsymbol{w}_t} [\ell_{t,i}] - \ell_{t,i_t^*} \right) \leq \textsc{Penalty}(\mathcal{C}, T, K, \eta) + \eta \sum_{t=1}^{T} \mathop{\mathbb{E}}_{i \sim \boldsymbol{w}_t} \left[q_{t,i} \ell_{t,i}^2 \right]$$

for every $(i_1^*, i_2^*, \dots, i_T^*) \in \mathcal{C}$, where $\textsc{Penalty}$ is a function that only depends on $\mathcal{C}$, T, K, and η.

For example, with $\eta \leq 1$, Hedge with $w_{t,i} \propto \exp(-\eta \sum_{\tau<t} \ell_{\tau,i})$ [FS97] satisfies Definition 3 with $\mathcal{C} = \mathcal{C}_0$ and $\textsc{Penalty} = \eta^{-1} \log K$. Fixed-Share [HW98], which mixes each update of Hedge with the uniform distribution at rate T^{-1}, satisfies Definition 3 with the class $\mathcal{C}_{S-1}$ of sequences that switch at most $(S-1)$ times and $\textsc{Penalty} = \eta^{-1}(S \log(KT) + 1)$, for every $S \geq 1$. Both algorithms have $\boldsymbol{q}_t \equiv 1$, but this need not be the case for dynamic regret or small-loss algorithms. Equipped with $\mathcal{A}^{\mathrm{OL}}$ which takes care of exploitation, we then explain how to do exploration in Algorithm 3.

Segment Tree Decomposition. Similar to Section 3.2 and Algorithm 2, we build a segment tree $\mathcal{T}$ over $\mathcal{S}$ such that all $[0, p)$'s are partitioned into $\mathcal{O}(\log T)$ tree intervals. In order to estimate the revenues of all $p \in \mathcal{S}$, it suffices to estimate $\mathbb{1}_t(I) := \mathbb{1}[v_t \in I]$ for every tree interval $I \in \mathcal{T}$, because[10]

$$p\mathbb{1}[v_t \geq p] = p \left(1 - \sum_{I \in \mathcal{I}_{\mathcal{T}}[0,p)} \mathbb{1}_t(I) \right), \quad \forall p \in \mathcal{S}. \tag{7}$$

Importance-Weighted Estimator. We maintain an "exploration distribution" $\boldsymbol{\mu}_t$ over all tree intervals in $\mathcal{T}$, whose construction is specified later, and explore (the two endpoints of) a random

[10] We write $\mathbb{1}[v_t \geq p] = 1 - \sum_{I \in \mathcal{I}_{\mathcal{T}}[0,p)} \mathbb{1}_t(I)$ instead of $\mathbb{1}[v_t \geq p] = \sum_{I \in \mathcal{I}_{\mathcal{T}}[p,1]} \mathbb{1}_t(I)$ for a technical reason: Online learning algorithms like Hedge mostly require losses to be lower bounded by -1 (but allowed to be arbitrarily positive); see Definition 3. As we will discuss at the end of this section, this is also why Algorithm 3 fails to use size-2 menus.

tree interval $I_t \sim \boldsymbol{\mu}_t$ in each round. Based on the interval query outcome $\mathbb{1}_t(I_t)$, from standard importance sampling properties, the estimator $\widehat{\mathbb{1}}_t(I) = \mathbb{1}[I_t = I]\mathbb{1}_t(I_t)/\mu_t(I)$ (see Eq. (10)) satisfies

$$\mathop{\mathbb{E}}_{t}[\widehat{\mathbb{1}}_t(I)] = \mathbb{1}_t(I), \quad \mathop{\mathbb{E}}_{t}[\widehat{\mathbb{1}}_t^2(I)] = \frac{\mathbb{1}_t(I)}{\mu_t(I)}, \quad \forall I \in \mathcal{T} \text{ with } \mu_t(I) > 0,$$

where $\mathbb{E}_t$ conditions on v_t and the history before round t. Plugging $\widehat{\mathbb{1}}_t$ into Eq. (7) then gives the unbiased estimator $\widehat{R}_t(p)$ in Eq. (11), which we feed into $\mathcal{A}^{\mathrm{OL}}$ as the revenue of every price $p \in \mathcal{S}$.

Lemma 5.1 (Importance-Weighted Estimator). *For every round $t \in [T]$, for any price $p \in \mathcal{S}$ such that $w_t(p) > 0$, we have $\mathbb{E}_t[\widehat{R}_t(p)] = p\mathbb{1}[v_t \geq p]$. Meanwhile, $\widehat{R}_t(p) \leq 1$ is bounded from above.*

Exploration Distribution. It remains to choose the exploration distribution $\boldsymbol{\mu}_t$. We now apply Definition 3: With $-\widehat{R}_t(p)$ as the loss for price $p \in \mathcal{S}$ in round $t \in [T]$, taking expectations,

$$\mathbb{E}\left[\sum_{t=1}^{T}\left(\widehat{R}_t(p_t^*) - \mathop{\mathbb{E}}_{p\sim \boldsymbol{w}_t}[\widehat{R}_t(p)]\right)\right] \leq \textsc{Penalty}(\mathcal{C}, T, T+1, \eta) + \eta\,\mathbb{E}\left[\sum_{t=1}^{T}\sum_{p\in\mathcal{S}} w_t(p)q_t(p)\widehat{R}_t(p)^2\right]. \quad (8)$$

As $q_t(p) \leq 1$ (recall Definition 3), for each $t \in [T]$, using the definition of $\widehat{R}_t(p)$ and $\widehat{\mathbb{1}}_t(I)$, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{t}\left[\sum_{p\in\mathcal{S}} w_t(p)\widehat{R}_t^2(p)\right] &\overset{(a)}{\leq} \mathop{\mathbb{E}}_{t}\left[\sum_{p\in\mathcal{S}} w_t(p)p^2\left(2 + 2\sum_{I\in\mathcal{I}_{\mathcal{T}}[0,p)}\widehat{\mathbb{1}}_t(I)^2\right)\right] \\
&\overset{(b)}{=} 2\sum_{p\in\mathcal{S}} w_t(p)p^2 + 2\sum_{p\in\mathcal{S}} w_t(p)p^2 \sum_{I\in\mathcal{I}_{\mathcal{T}}[0,p)} \frac{\mathbb{E}_t[\mathbb{1}[I_t = I]]\mathbb{1}_t(I)}{\mu_t(I)^2} \\
&\overset{(c)}{\leq} 2\sum_{p\in\mathcal{S}} w_t(p)p^2 + 2\sum_{I\in\mathcal{T}} \frac{\mathbb{1}_t(I)}{\mu_t(I)} \sum_{p:\mathcal{I}_{\mathcal{T}}[0,p)\ni I} w_t(p)p^2,
\end{aligned}$$

where (a) uses the definition of $\widehat{R}_t(p)$ in Eq. (11) and $(a+b)^2 \leq 2a^2 + 2b^2$, together with $\widehat{\mathbb{1}}_t(I) = 0$ for all $I \neq I_t$ (so that the cross terms vanish); (b) uses the definition of $\widehat{\mathbb{1}}_t(I)$ in Eq. (10) and $\mathbb{1}[\cdot]^2 = \mathbb{1}[\cdot]$; (c) uses $\mathbb{E}_t[\mathbb{1}[I_t = I]] = \mu_t(I)$ and swaps the summations.

In Eq. (9), we thus set $\mu_t(I)$ proportional to $\sum_{p:\mathcal{I}_{\mathcal{T}}[0,p)\ni I} w_t(p)p^2$ so they cancel (only with a normalization Z_t remaining). Then the last term becomes $2Z_t\sum_{I\in\mathcal{T}}\mathbb{1}_t(I)$, and we further have *i)* $Z_t \leq 2(\log_2 T + 1)$ because each $\mathcal{I}_{\mathcal{T}}[0,p)$ contains at most two tree intervals of each depth; and *ii)* $\sum_{I\in\mathcal{T}}\mathbb{1}_t(I) \leq \log_2 T + 1$ because v_t lies in at most one tree interval of each depth. To summarize,

Lemma 5.2 (Second Moment). *For every round t, we have $\mathbb{E}_t\left[\sum_{p\in\mathcal{S}} w_t(p)\widehat{R}_t^2(p)\right] \leq 2+4(\log_2 T+1)^2$.*

This consequently gives the following regret bound for adversarial dynamic pricing (for a comparator price p with $w_t(p) = 0$, we still have $\mathbb{E}_t[\widehat{R}_t(p)] \geq p\mathbb{1}[v_t \geq p]$ by Lemma E.3).

Theorem 5.3 (Adversarial Dynamic Pricing). *For any comparator class $\mathcal{C} \subseteq \{0, \frac{1}{T}, \ldots, 1\}^T$, for any $\mathcal{A}^{OL}$ satisfying Definition 3 with $K = T+1$ and $\mathcal{C}$ translated to $[K]$, Algorithm 3 ensures*

$$\mathbb{E}\left[\sum_{t=1}^{T} p_t^*\mathbb{1}[v_t \geq p_t^*]\right] - \mathbb{E}\left[\sum_{t=1}^{T} revenue_t\right] \leq \textsc{Penalty}(\mathcal{C}, T, T+1, \eta) + \mathcal{O}\big(\eta T\log^2 T + 1\big),\ \forall (p_1^*, \ldots, p_T^*) \in \mathcal{C}.$$

In particular, Hedge [FS97] with $\eta = (T\log T)^{-1/2}$ gives $\mathcal{O}\big(\sqrt{T\log^3 T}\big)$ regret against the best fixed price $\mathcal{C}_0$, and Fixed-Share [HW98] with $\eta = \sqrt{S/(T\log T)}$ gives $\mathcal{O}\Big(\sqrt{ST\log^3 T}\Big)$ against $\mathcal{C}_{S-1}$.

Algorithm 3 Adversarial Dynamic Pricing Using Size-3 Menus

Input: Time horizon T, which is WLOG assumed to be a power of 2. Online learning algorithm $\mathcal{A}^{\mathrm{OL}}$

Build a segment tree by $(\mathcal{T}, \mathcal{I}_{\mathcal{T}}) \leftarrow \textsc{BuildSegTree}(0, \frac{1}{T}, \frac{2}{T}, \dots, 1)$ ▷ *see Subroutine 4*

Initialize an online learning algorithm $\mathcal{A}^{\mathrm{OL}}$ on the $T+1$ actions in $\mathcal{S} = \{0, \frac{1}{T}, \frac{2}{T}, \dots, 1\}$

for $t = 1, 2, \dots, T$ **do**

For any $p \in \mathcal{S}$, let $w_t(p)$ be the probability of $\mathcal{A}^{\mathrm{OL}}$ playing p

Calculate the exploration probability for all intervals $I \in \mathcal{T}$ (uniform if $Z_t = 0$):

$$\mu_t(I) := \frac{1}{Z_t} \sum_{p : \mathcal{I}_{\mathcal{T}}[0,p) \ni I} w_t(p) p^2, \ \forall I \in \mathcal{T}; \quad \text{where } Z_t := \sum_{I \in \mathcal{T}} \sum_{p : \mathcal{I}_{\mathcal{T}}[0,p) \ni I} w_t(p) p^2 \tag{9}$$

Sample exploitation price $p_t \sim \boldsymbol{w}_t$ and exploration interval $I_t \sim \boldsymbol{\mu}_t$

Offer the size-3 menu built from putting $1 - 2\gamma$ mass on p_t and γ mass on each endpoint of I_t, where $\gamma = \frac{1}{T}$. Observe the interval feedback $\mathbb{1}_t(I_t) := \mathbb{1}[v_t \in I_t]$

Construct the importance-weighted estimator for each tree interval $I \in \mathcal{T}$:

$$\widehat{\mathbb{1}}_t(I) := \frac{\mathbb{1}[I_t = I]\mathbb{1}_t(I_t)}{\mu_t(I)}, \quad \forall I \in \mathcal{T}. \tag{10}$$

For each candidate price $p \in \mathcal{S}$, feed $-\widehat{R}_t(p)$ as their round-t loss to $\mathcal{A}^{\mathrm{OL}}$, where

$$\widehat{R}_t(p) := p \left(1 - \sum_{I \in \mathcal{I}_{\mathcal{T}}[0,p)} \widehat{\mathbb{1}}_t(I)\right) \in (-\infty, 1], \quad \forall p \in \mathcal{S}. \tag{11}$$

Hence size-3 menus attain $\widetilde{\mathcal{O}}(\sqrt{T})$ regret against the best fixed price, even for adaptive adversaries, which is optimal up to logarithmic factors even with i.i.d. values and infinite menus (Theorem 4.1).

It remains open whether size-2 menus can attain $\widetilde{\mathcal{O}}(\sqrt{T})$ regret against adversarial values. Unless it gives up the revenue of the round, a size-2 menu can only explore an interval with the exploitation price p_t as an endpoint. How to estimate $\mathbb{1}[v_t \ge p] = \mathbb{1}[v_t \ge p_t] + \mathbb{1}[p \le v_t < p_t]$ for a $p < p_t$? If we use importance-weighting for both terms, the resulting estimators are unbounded from *both* above and below (cf. Lemma 5.1), and Definition 3 no longer applies. Such both-side unbounded losses have long been a challenge in online learning. While there are ad-hoc solutions for specific problems [e.g., DLWZ23, and the discussions therein], we are unaware of any that applies to our Algorithm 3.

We also briefly discuss why adversarial CDF estimation algorithms [MN21; OGKG23], where one estimates the empirical distribution F_T of adversarial values $v_1, \dots, v_T$ via threshold queries, cannot be used as a black-box. Suppose that, for every prefix $t \in [T]$, we are given an estimate $\widehat{F}_t$ that's ϵ_t-close to the current empirical distribution F_t. First, these existing works only give $\epsilon_t = \widetilde{\mathcal{O}}(t^{-1/3})$ using threshold queries, and it is unclear whether we can get $\epsilon_t = \widetilde{\mathcal{O}}(t^{-1/2})$ by interval queries (note that, our Algorithms 2 and 5 both rely on i.i.d. values). Even if we can, critically, an ϵ_t guarantee on the empirical distribution only translates to an $\epsilon_t t$ guarantee on estimating the *cumulative* revenue function. However, online learning algorithms (e.g., Hedge) require estimating the *per-round revenue*. To see why a cumulative guarantee is insufficient, consider a delayed online learning problem [WO02]: Suppose that the true value v_t is fully observed after $\epsilon_t t$ rounds, which also gives $\epsilon_t t$ close cumulative revenues (and stronger than the black-box guarantee), one must still suffer $\Omega\left(\sqrt{\sum_{t=1}^{T} \epsilon_t t}\right)$ regret. So even if $\epsilon_t = \widetilde{\mathcal{O}}(t^{-1/2})$, a black-box algorithm likely only gives $\widetilde{\mathcal{O}}(T^{3/4})$ regret. We finally remark that, while the estimator by Meister and Nietert [MN21] is unbiased, it has a second moment the same order as the number of candidate prices (namely $|\mathcal{S}|$), thus also only giving $\widetilde{\mathcal{O}}(T^{2/3})$ regret.

6 Further Discussion of Related Work

Dynamic Pricing. The study of regret-minimizing dynamic pricing was initiated in [KL03] and attracted much attention afterwards. In all this work (with a notable exception of [Zuo26], discussed below), each customer is offered a single posted price. Minimax-optimal regret rates scale as $\widetilde{\Theta}(T^{2/3})$ for all problem instances [KL03] and $\widetilde{\Theta}(\sqrt{T})$ for regular distributions [SW24; LSTW23b] or under a certain smoothness condition ([KL03], see Footnote 7). Subsequent work makes additional assumptions such as parametric structure [BR12; KZ14] or finitely many values [CCP19]. Various extensions were considered, most notably: adversarial or drifting values [KL03; Sli14; LSTW21; PS25] and limited supply [BZ09; BDKS15; WDY14; BKS18].

Dynamic pricing was also studied *beyond* the framework of one firm and one type of customers. For example, when multiple pricing agents interact with each other, competing [e.g., BV96; BV00] and/or colluding [e.g., CCDP20; Kle21; HMP21]; or when the chosen price(s) connect buyers and sellers, namely the bilateral trade [e.g., Ces+24; Ces+23; CJLZ26b].

Lipschitz Bandits (see Footnote 2) are closely related to dynamic pricing with a single posted price. Lipschitz bandits with real-valued actions, a.k.a. continuum-armed bandits [Agr95; Kle04], have the same minimax-optimal regret rate as dynamic pricing, $\widetilde{\Theta}(T^{2/3})$, and this rate is obtained by the same approach: uniform discretization of the action space. *Adaptive* discretization of the action space [e.g., KSU19; BMSS11; Sli14; PS25], a major theme in Lipschitz bandits, seamlessly carries over to dynamic pricing as well. This approach leads to improved guarantees in terms of *zooming dimension*, an instance-dependent quantity which (for $\mathbb{R}^d$-valued actions) is d in the worst case and gets smaller if near-optimal actions comprise a low-dimensional region of the action space. More background on Lipschitz bandits can be found in [Sli19, Ch. 4].

Our notion of "interval dimension" is a variant of zooming dimension which is more stringent but retains instance-dependence. Our algorithm performs adaptive discretization, albeit in a very different way. Interestingly, even for our minimax result for general distributions, we still need adaptive discretization to move from size-3 menus to size-2 ones. (Recall that for the corresponding worst-case result in Lipschitz bandits, uniform discretization suffices.)

Bandits Lower Bounds. Regret lower bounds in bandit problems, including those for dynamic pricing, must deal with the fundamental issue that they cannot consider a single problem instance (since they need to rule out an algorithm tailored to this problem instance). The most common approach, starting from [ACFS02], fixes the time horizon T and takes the worst case over some instance class.[11] Some results for dynamic pricing and Lipschitz bandits refine this approach and construct problem instances that are difficult for infinitely many T's at once, ruling out algorithms that attain a favorable regret rate after some initial time period [KL03; KSU19].

A standard approach for *per-instance* lower bounds, starting from [LR85], focuses on a particular problem instance but further requires non-trivial performance (e.g., $o(T)$ regret) on several hard "decoy" instances. The quantity being lower-bounded is *not* the dependence on T (like in minimax lower bounds) but instead the "leading constant": the C in a regret bound of the form $C \cdot f(T) + o(f(T))$. [MCP14] obtain such a result for Lipschitz bandits. Moreover, this approach fails under full-feedback problems, as detailed in Section 4.2.

Our per-instance lower bounds, on the other hand, follow a different template (called *local-minimax*)

[11] Various lower bounds of this type were proved for dynamic pricing [KL03; BDKS15] and Lipschitz bandits [e.g., Kle04; Sli14], as well as for other bandit structures such as linear bandits [e.g., DHK07; DHK08].

whereby the decoys are vanishingly close to the original problem instance and are subject to the same lower bound on regret. While this template was previously applied in [CL15; ZCDL16] to estimating convex functions from samples, their techniques do not apply to our setting.

Mechanism Design with Menus. Closest to us, [Zuo26] shows that for a buyer with a *fixed but unknown value* (i.e., without randomness), menus avoid the $\Theta(\log \log T)$ regret of posted prices, and size-2 menus attain the optimal regret even if the buyer is strategic. While they also consider size-2 menus, their fixed value setup is very different from our dynamic pricing problem. Han, Albert, and Xu [HAX24] also learn fixed type agents with menus: In online principal-agent problems, the agent's choice from a menu identifies the type in fewer rounds than a single strategy, e.g., in a single round for pricing with an infinite menu. In other mechanism problems, menus can sometimes help elicit more revenue even if the agent's type distribution is known, e.g., when selling multiple goods [HN19; BGN22] or in contract design [GSWZ23; CMG22; GHWX24]. In our single-item problem, however, no menu earns more than posting the monopoly price when F is known [Mye81], and therefore menus are only helpful in our online dynamic pricing setup where types must be learned.

Sample and Query Complexity in Pricing. Following the framework of Cole and Roughgarden [CR14], Huang, Mansour, and Roughgarden [HMR15] characterize the sample complexity of pricing a single buyer, e.g., $\widetilde{\Theta}(\epsilon^{-1.5})$ for MHR distributions. Later works develop refined multi-buyer bounds and extend to posted pricing [GHZ19; GHTZ21; JKMS24]. Leme, Sivan, Teng, and Worah [LSTW23b] introduce threshold (i.e., pricing) queries and separate them from samples, e.g., $\widetilde{\Theta}(\epsilon^{-3})$ threshold queries versus $\mathcal{O}(\epsilon^{-2})$ samples for a general CDF; Leme, Sivan, Teng, and Worah [LSTW23a] learn regular CDFs with $\widetilde{\mathcal{O}}(\epsilon^{-2.5})$ threshold queries via a $\widetilde{\Theta}(\epsilon^{-1/2})$-bit description; Chen, Jin, Lu, and Zhang [CJLZ26a] show that regularity no longer helps threshold queries for uniform pricing with multiple buyers; and Tang, Wang, and Zhang [TWZ26] study multiplicative approximations. With adversarial values, the median and the empirical CDF can be estimated with one threshold query per value [MN21; OGKG23]; see the end of Section 5 for more discussions. We introduce interval queries, which lie between samples and threshold queries, and show that they are as good as samples for the CDF and for the monopoly price of all four classes.

7 Conclusion and Future Research

We show that for dynamic pricing under unknown demands, a second offer is almost as good as infinite offers. This holds not only in the minimax sense over several distribution classes, but also locally around every value distribution. We also study the query-complexity version, where we show that between the extremes of threshold queries and value queries, the interval queries are almost as powerful as value queries. We also generalize to adversarial dynamic pricing, where our worst-case regret bound, attained by size-3 menus, is again as good as the one for infinite menus.

An immediate next-step question is whether *size-2* menus are also as powerful as infinite menus in adversarial dynamic pricing (whose challenges are outlined at the end of Section 5). Another direction is to extend beyond a single myopic customer per round: e.g., to multiple simultaneous customers with heterogeneous types, or to long-lived customers strategically optimizing their long-term payoffs.

AI Methodology Statement

The problem formulation, the idea of using menus as interval queries (or infinite menus as full feedback), the main algorithmic framework, instance-optimality, and extensions to adversarial

dynamic pricing are all by the authors. The authors used Claude Opus 4.8–5.5 and GPT-5.5–6.1 to help develop various intermediate drafts, verify many algorithmic ideas, help with lower bound constructions, and perform literature search. Claude Opus 5.5 is used to write the proofs in the appendices, especially those minimax lower bounds using standard techniques. The authors have manually proofread every proof and take full responsibility. The main text is fully human-written and iteratively polished by all the authors.

References

[ACFS02] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. "The Nonstochastic Multiarmed Bandit Problem". In: *SIAM Journal on Computing* 32.1 (2002), pp. 48–77.

[Agr95] Rajeev Agrawal. "The continuum-armed bandit problem". In: *SIAM J. Control and Optimization* 33.6 (1995), pp. 1926–1951.

[BDKS15] Moshe Babaioff, Shaddin Dughmi, Robert Kleinberg, and Aleksandrs Slivkins. "Dynamic Pricing with Limited Supply". In: *ACM Transactions on Economics and Computation* 3.1 (2015).

[BGN22] Moshe Babaioff, Yannai A. Gonczarowski, and Noam Nisan. "The Menu-Size Complexity of Revenue Approximation". In: *Games and Economic Behavior* 134 (2022), pp. 281–307.

[BKS18] Ashwinkumar Badanidiyuru, Robert Kleinberg, and Aleksandrs Slivkins. "Bandits with Knapsacks". In: *Journal of the ACM (JACM)* 65.3 (2018), pp. 1–55.

[BMP63] Richard E Barlow, Albert W Marshall, and Frank Proschan. "Properties of Probability Distributions with Monotone Hazard Rate". In: *The Annals of Mathematical Statistics* (1963), pp. 375–389.

[BMSS11] Sébastien Bubeck, Rémi Munos, Gilles Stoltz, and Csaba Szepesvári. "X-Armed Bandits". In: *Journal of Machine Learning Research* 12.5 (2011).

[BR12] Josef Broder and Paat Rusmevichientong. "Dynamic Pricing under a General Parametric Choice Model". In: *Operations Research* 60.4 (2012), pp. 965–980.

[BV00] Dirk Bergemann and Juuso Välimäki. "Experimentation in Markets". In: *The Review of Economic Studies* 67.2 (2000), pp. 213–234.

[BV96] Dirk Bergemann and Juuso Välimäki. "Learning and Strategic Pricing". In: *Econometrica* 64.5 (1996), pp. 1125–1149.

[BZ09] Omar Besbes and Assaf Zeevi. "Dynamic Pricing Without Knowing the Demand Function: Risk Bounds and Near-Optimal Algorithms". In: *Operations Research* 57.6 (2009), pp. 1407–1420.

[CCDP20] Emilio Calvano, Giacomo Calzolari, Vincenzo Denicolo, and Sergio Pastorello. "Artificial Intelligence, Algorithmic Pricing, and Collusion". In: *American Economic Review* 110.10 (2020), pp. 3267–3297.

[CCP19] Nicolo Cesa-Bianchi, Tommaso Cesari, and Vianney Perchet. "Dynamic Pricing with Finitely Many Unknown Valuations". In: *Algorithmic Learning Theory*. PMLR. 2019, pp. 247–273.

[Ces+23] Nicolò Cesa-Bianchi, Tommaso R. Cesari, Roberto Colomboni, Federico Fusco, and Stefano Leonardi. "Repeated Bilateral Trade against a Smoothed Adversary". In: *Proceedings of the Thirty Sixth Conference on Learning Theory*. PMLR. 2023, pp. 1095–1130.

[Ces+24] Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Federico Fusco, and Stefano Leonardi. "Bilateral Trade: A Regret Minimization Perspective". In: *Mathematics of Operations Research* 49.1 (2024), pp. 171–203.

[CJLZ26a] Houshuang Chen, Yaonan Jin, Pinyan Lu, and Chihao Zhang. "The Query Complexity of Uniform Pricing". In: *Proceedings of the ACM Web Conference 2026*. 2026, pp. 237–248.

[CJLZ26b] Houshuang Chen, Yaonan Jin, Pinyan Lu, and Chihao Zhang. "Tight Regret Bounds for Fixed-Price Bilateral Trade". In: *53rd International Colloquium on Automata, Languages, and Programming (ICALP 2026)*. Vol. 374. 2026, 57:1–57:21.

[CL15] T Tony Cai and Mark G Low. "A Framework for Estimation of Convex Functions". In: *Statistica Sinica* (2015), pp. 423–456.

[CMG22] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. "Designing Menus of Contracts Efficiently: The Power of Randomization". In: *Proceedings of the 23rd ACM Conference on Economics and Computation*. 2022, pp. 705–735.

[CR14] Richard Cole and Tim Roughgarden. "The Sample Complexity of Revenue Maximization". In: *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*. 2014, pp. 243–252.

[DHK07] Varsha Dani, Thomas P. Hayes, and Sham Kakade. "The Price of Bandit Information for Online Optimization". In: *20th Advances in Neural Information Processing Systems (NIPS)*. 2007.

[DHK08] Varsha Dani, Thomas P. Hayes, and Sham Kakade. "Stochastic Linear Optimization under Bandit Feedback". In: *21st Conf. on Learning Theory (COLT)*. 2008, pp. 355–366.

[DKW56] Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz. "Asymptotic Minimax Character of the Sample Distribution Function and of the Classical Multinomial Estimator". In: *The Annals of Mathematical Statistics* (1956), pp. 642–669.

[DLWZ23] Yan Dai, Haipeng Luo, Chen-Yu Wei, and Julian Zimmert. "Refined Regret for Adversarial MDPs with Linear Function Approximation". In: *International Conference on Machine Learning*. PMLR. 2023, pp. 6726–6759.

[FJ25] Yiding Feng and Yaonan Jin. "Beyond Regularity: Simple versus Optimal Mechanisms, Revisited". In: *2025 IEEE 66th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE. 2025, pp. 303–342.

[FS97] Yoav Freund and Robert E Schapire. "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting". In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119–139.

[GHTZ21] Chenghao Guo, Zhiyi Huang, Zhihao Gavin Tang, and Xinzhi Zhang. "Generalizing Complex Hypotheses on Product Distributions: Auctions, Prophet Inequalities, and Pandora's Problem". In: *Conference on Learning Theory*. PMLR. 2021, pp. 2248–2288.

[GHWX24] Jiarui Gan, Minbiao Han, Jibang Wu, and Haifeng Xu. "Generalized Principal-Agency: Contracts, Information, Games and Beyond". In: *Proceedings of the 20th Conference on Web and Internet Economics*. 2024.

[GHZ19] Chenghao Guo, Zhiyi Huang, and Xinzhi Zhang. "Settling the Sample Complexity of Single-Parameter Revenue Maximization". In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. 2019, pp. 662–673.

[GSWZ23] Guru Guruganesh, Jon Schneider, Joshua Wang, and Junyao Zhao. "The Power of Menus in Contract Design". In: *Proceedings of the 24th ACM Conference on Economics and Computation*. 2023, pp. 818–848.

[HAX24] Minbiao Han, Michael Albert, and Haifeng Xu. "Learning in Online Principal-Agent Interactions: The Power of Menus". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 38. 16. 2024, pp. 17426–17434.

[HH39] R. L. Hall and C. J. Hitch. "Price Theory and Business Behaviour". In: *Oxford Economic Papers* 2 (1939), pp. 12–45.

[HMP21] Karsten T. Hansen, Kanishka Misra, and Mallesh M. Pai. "Frontiers: Algorithmic Collusion: Supra-Competitive Prices via Independent Algorithms". In: *Marketing Science* 40.1 (2021), pp. 1–12.

[HMR15] Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. "Making the Most of Your Samples". In: *Proceedings of the Sixteenth ACM Conference on Economics and Computation*. 2015, pp. 45–60.

[HN19] Sergiu Hart and Noam Nisan. "Selling Multiple Correlated Goods: Revenue Maximization and Menu-Size Complexity". In: *Journal of Economic Theory* 183 (2019), pp. 991–1029.

[HW98] Mark Herbster and Manfred K Warmuth. "Tracking the Best Expert". In: *Machine Learning* 32.2 (1998), pp. 151–178.

[JKMS24] Billy Jin, Thomas Kesselheim, Will Ma, and Sahil Singla. "Sample Complexity of Posted Pricing for a Single Item". In: *Advances in Neural Information Processing Systems* 37 (2024), pp. 82296–82317.

[KL03] Robert Kleinberg and Tom Leighton. "The Value of Knowing a Demand Curve: Bounds on Regret for Online Posted-Price Auctions". In: *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*. 2003, p. 594.

[Kle04] Robert Kleinberg. "Nearly Tight Bounds for the Continuum-Armed Bandit Problem." In: *18th Advances in Neural Information Processing Systems (NIPS)*. 2004.

[Kle21] Timo Klein. "Autonomous Algorithmic Collusion: Q-Learning under Sequential Pricing". In: *The RAND Journal of Economics* 52.3 (2021), pp. 538–558.

[KSU19] Robert Kleinberg, Aleksandrs Slivkins, and Eli Upfal. "Bandits and Experts in Metric Spaces". In: *Journal of the ACM (JACM)* 66.4 (2019). Merged and revised version of conference papers in *ACM STOC 2008* and *ACM-SIAM SODA 2010*, pp. 1–77.

[KZ14] N Bora Keskin and Assaf Zeevi. "Dynamic Pricing with an Unknown Demand Model: Asymptotically Optimal Semi-Myopic Policies". In: *Operations Research* 62.5 (2014), pp. 1142–1167.

[LR85] Tze Leung Lai and Herbert Robbins. "Asymptotically Efficient Adaptive Allocation Rules". In: *Advances in Applied Mathematics* 6.1 (1985), pp. 4–22.

[LS20] Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge, UK: Cambridge University Press, 2020.

[LSTW21] Renato Paes Leme, Balasubramanian Sivan, Yifeng Teng, and Pratik Worah. "Learning to Price against a Moving Target". In: *International Conference on Machine Learning*. PMLR. 2021, pp. 6223–6232.

[LSTW23a] Renato Paes Leme, Balasubramanian Sivan, Yifeng Teng, and Pratik Worah. "Description Complexity of Regular Distributions". In: *Proceedings of the 24th ACM Conference on Economics and Computation*. 2023, p. 959.

[LSTW23b] Renato Paes Leme, Balasubramanian Sivan, Yifeng Teng, and Pratik Worah. "Pricing Query Complexity of Revenue Maximization". In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2023, pp. 399–415.

[MCP14] Stefan Magureanu, Richard Combes, and Alexandre Proutiere. "Lipschitz Bandits: Regret Lower Bound and Optimal Algorithms". In: *27th Conf. on Learning Theory (COLT)*. 2014, pp. 975–999.

[MN21] Michela Meister and Sloan Nietert. "Learning with Comparison Feedback: Online Estimation of Sample Statistics". In: *Algorithmic Learning Theory*. PMLR. 2021, pp. 983–1001.

[MP09] Andreas Maurer and Massimiliano Pontil. "Empirical Bernstein Bounds and Sample-Variance Penalization". In: *Conference on Learning Theory*. 2009.

[Mye81] Roger B Myerson. "Optimal Auction Design". In: *Mathematics of Operations Research* 6.1 (1981), pp. 58–73.

[OGKG23] Princewill Okoroafor, Vaishnavi Gupta, Robert Kleinberg, and Eleanor Goh. "Non-Stochastic CDF Estimation Using Threshold Queries". In: *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM. 2023, pp. 3551–3572.

[PS25] Chara Podimata and Aleksandrs Slivkins. *Adaptive Discretization against an Adversary: Lipschitz bandits, Dynamic Pricing, and Auction Tuning*. Working paper. 2025. Preliminary version at Conf. on Learning Theory (COLT), 2021.

[Sli14] Aleksandrs Slivkins. "Contextual bandits with similarity information". In: *J. of Machine Learning Research (JMLR)* 15.1 (2014). Preliminary version in *COLT 2011*, pp. 2533–2568.

[Sli19] Aleksandrs Slivkins. "Introduction to Multi-Armed Bandits". In: *Foundations and Trends® in Machine Learning* 12.1-2 (Nov. 2019). Published with *Now Publishers* (Boston, MA, USA). Also available at `https://arxiv.org/abs/1904.07272`, pp. 1–286.

[SW24] Sahil Singla and Yifan Wang. "Bandit Sequential Posted Pricing via Half-Concavity". In: *Proceedings of the 25th ACM Conference on Economics and Computation*. 2024, pp. 922–939.

[Swe39] Paul M Sweezy. "Demand under Conditions of Oligopoly". In: *Journal of Political Economy* 47.4 (1939), pp. 568–573.

[TWZ26] Wei Tang, Yifan Wang, and Mengxiao Zhang. "Pricing Query Complexity of Multiplicative Revenue Approximation". In: *arXiv Preprint arXiv:2602.10483* (2026).

[WDY14] Zizhuo Wang, Shiming Deng, and Yinyu Ye. "Close the Gaps: A Learning-While-Doing Algorithm for Single-Product Revenue Management Problems". In: *Operations Research* 62.2 (2014), pp. 318–331.

[WO02] Marcelo J Weinberger and Erik Ordentlich. "On Delayed Prediction of Individual Sequences". In: *IEEE Transactions on Information Theory* 48.7 (2002), pp. 1959–1976.

[ZCDL16] Yuancheng Zhu, Sabyasachi Chatterjee, John Duchi, and John Lafferty. "Local Minimax Complexity of Stochastic Convex Optimization". In: *Advances in Neural Information Processing Systems* 29 (2016).

[Zin03] Martin Zinkevich. "Online Convex Programming and Generalized Infinitesimal Gradient Ascent". In: *Proceedings of the 20th International Conference on Machine Learning*. 2003, pp. 928–936.

[Zuo26] Shiliang Zuo. "Learning Is Revelation in Disguise: Optimal Regret and Equivalence Results for Dynamic Pricing". In: *arXiv Preprint arXiv:2604.24093* (2026).

Technical Appendices

A Warm-Up: Infinite Menus **30**

B Query Complexities **33**

B.1 Definitions: Sample, Interval-Query, and Threshold-Query Complexities 33

B.2 Relations and Reductions Between Complexities 34

B.3 Two-Point and Needle-in-a-Haystack Methods 37

B.4 Results on Sample Complexities 38

B.5 Results on Interval-Query Complexities 41

B.6 Results on Threshold-Query Complexities 42

C Omitted Proofs in Section 3 (Upper Bounds) **44**

C.1 Empirical Bernstein and Segment Trees 44

C.2 Proof of Theorem 3.2 (Revenue Maximization with Size-3 Menus) 46

C.3 ScaleSearch for MHR Distributions (Line 1 in Algorithm 2) 48

C.4 Proof of Lemma 3.3 (Number of Rounds) 50

C.5 Proofs of Theorems 3.4 and 3.5 (Upper Bounds with Size-2 Menus) 53

C.6 Examples of Interval Dimensions 55

D Omitted Proofs in Section 4 (Lower Bounds) **57**

D.1 Two-Point Method for Regret 57

D.2 Proof of Theorem 4.1 (Minimax Lower Bounds) 58

D.3 Proofs of Theorem 4.2 and Corollary 4.3 (Local Minimax Lower Bound) 59

D.4 Proof of Theorem 4.4 (Near Instance-Optimality via Interval Dimensions) 64

D.5 Proof of Theorem 4.5 (Near Instance-Optimality for MHR Distributions) 66

D.6 Proof of Theorem 4.6 (Anytime Lower Bounds) 67

D.7 Lower Bound for Lipschitz Experts 70

E Omitted Proofs in Section 5 (Adversarial Dynamic Pricing) **71**

A Warm-Up: Infinite Menus

This appendix proves our claims on infinite menus. The infinite menu built from a price distribution earns the same revenue as posting a random price from it, and reveals the value if this distribution has full support on $[0,1]$ (Lemma A.1), which also proves Lemma 3.1 as a special case. Conversely, no menu earns more than a random posted price (Lemma A.3). Hence infinite menus are as powerful as full feedback (Theorem A.4), and their minimax regret is governed by the sample complexity (Reduction A.5), which gives Theorem A.6.

Lemma A.1 (Infinite Menus as Randomized Prices)**.** *Let P be a price distribution over $[0,1]$, and let $q(x) := \Pr_{p\sim P}\{p \le x\}$ for $x \in [0,1]$. Consider the following infinite menu*

$$M := \{(q(x), \pi(x)) \mid x \in [0,1]\}, \quad \text{where } \pi(x) := q(x)\cdot x - \int_0^x q(y)\, dy.$$

Then a customer with value v chooses the option with $x = v$, which gives the highest payoff ("optimal"), and has (one of) the largest allocation probability among all optimal options. Hence:

- *If P has full support on $[0,1]$, the seller infers v from the chosen option;*
- *If P has a finite support $p_1 < p_2 < \cdots < p_m$, the seller observes which of $[0,p_1), [p_1,p_2), \ldots, [p_m,1]$ contains v.*
- *The customer gets the item with probability $q(v) = \Pr_{p\sim P}\{v \ge p\}$ and pays $\pi(v) = \mathbb{E}_{p\sim P}[p\mathbb{1}[v \ge p]]$, the same as posting a random price $p \sim P$.*

Proof. We first show that the option with $x = v$ is optimal, then that it wins the tie-breaking, and finally compute what the seller observes and what the customer pays.

Optimality. For any $x, v \in [0,1]$, the utility $U(q', \pi'; v) = q'v - \pi'$ of an option (q', π') satisfies

$$U(q(v),\pi(v);v) - U(q(x),\pi(x);v) = \int_0^v q(y)\,dy - \int_0^x q(y)\,dy - q(x)(v-x) = \int_x^v (q(y)-q(x))\,dy \ge 0,$$

where the inequality holds as q is non-decreasing, so $q(y) - q(x)$ is non-negative for $y > x$ and non-positive for $y < x$. Moreover, $U(q(v),\pi(v);v) = \int_0^v q(y)\,dy \ge 0 = U(0,0;v)$, so the option with $x = v$ is also at least as good as the outside option.

Tie-Breaking. Suppose that the option with some x is also optimal. If $q(x) = q(v)$, then q is constant between x and v, so $\pi(x) = \pi(v)$ and the two options are the same. Otherwise, the integral above is 0, so q is constant on the open interval between x and v, which (by the right-continuity of q) requires $x < v$, so $q(x) < q(v)$. Hence the option with $x = v$ has the largest allocation probability among all optimal options, and the customer chooses it.

Observation and Payment. If P has full support on $[0,1]$, then q is strictly increasing, so $q(v)$ determines v. If P is supported on $p_1 < p_2 < \cdots < p_m$, then q is constant on each of $[0,p_1), [p_1,p_2), \ldots, [p_m,1]$ and takes different values on different intervals, so $q(v)$ tells which of them contains v. Finally, $q(v) = \Pr_{p\sim P}\{v \ge p\}$ by definition, and by Fubini's theorem, $\int_0^v q(y)\,dy = \mathbb{E}_{p\sim P}[(v-p)\mathbb{1}[v \ge p]]$, so $\pi(v) = q(v)\cdot v - \mathbb{E}_{p\sim P}[(v-p)\mathbb{1}[v \ge p]] = \mathbb{E}_{p\sim P}[p\mathbb{1}[v \ge p]]$. □

Lemma A.2 (Restatement of Lemma 3.1)**.** *Fix a price distribution P with discrete support $0 \le p_1 < \cdots < p_m \le 1$. For $i \in [m]$, let w_i be the probability of p_i under P. Construct a size-m menu $M = \left\{\left(\sum_{j\le i} w_j,\ \sum_{j\le i} w_j p_j\right) \mid i \in [m]\right\}$.*

Given M, a customer with value $v \in [p_i, p_{i+1})$ chooses the i-th option, where $p_{m+1} := \infty$, and one with $v < p_1$ chooses the outside option. Hence, the seller observes which of the $(m+1)$ intervals, $[0, p_1), [p_1, p_2), \ldots, [p_m, 1]$, contains v; and the customer gets the item with probability $\Pr_{p\sim P}\{v \geq p\}$ and pays $\mathbb{E}_{p\sim P}[p\,\mathbb{1}[v \geq p]]$ in expectation, the same as posting a random price $p \sim P$.

Proof. Apply Lemma A.1 to $P = \sum_{i=1}^{m} w_i \delta_{p_i}$, for which $q(x) = \sum_{j:p_j \leq x} w_j$. By Lemma A.1, $\pi(x) = \mathbb{E}_{p\sim P}[p\mathbb{1}[x \geq p]]$, so for $x \in [p_i, p_{i+1})$, the option indexed by x is $(\sum_{j\leq i} w_j, \sum_{j\leq i} w_j p_j)$, i.e., the i-th option of M, and for $x < p_1$, it is the outside option $(0, 0)$. So M together with the outside option is the infinite menu in Lemma A.1, and Lemma A.1 gives all claims, as P is supported on $p_1 < p_2 < \cdots < p_m$. □

Conversely, no menu earns more than a random posted price, value by value:

Lemma A.3 (Menu Revenue)**.** *For every menu M, there is a price distribution P over $[0, 1]$, which only depends on M, such that for every $v \in [0, 1]$, a customer with value v pays at most $\mathbb{E}_{p\sim P}[p\mathbb{1}[v \geq p]]$ under M, i.e., at most the expected payment when a random price $p \sim P$ is posted. Hence, for a value $v \sim G$ and a (possibly random) menu independent of v, with P averaged over the randomness of the menu, the expected revenue is at most $\mathbb{E}_{p\sim P}[R_G(p)]$, and the regret of the menu is at least $\mathbb{E}_{p\sim P}[R_G(p^*_G) - R_G(p)]$.*

Proof. It suffices to consider a deterministic menu M, as for a random menu, we can average P over its randomness. We write the payment via the utility, and then compare it with the menu in Lemma A.1.

Let $q(v)$ and $\pi(v)$ be the allocation probability and the payment of the option chosen by a customer with value v, and let $U(v) := q(v)v - \pi(v)$ be their utility. As $U(v)$ is the maximum of $q'v - \pi'$ over the options (q', π') in M and the outside option $(0, 0)$, it is convex and 1-Lipschitz, with $q(v)$ as a subgradient at v. Hence q is non-decreasing, $U(v) = U(0) + \int_0^v q(y)\,dy$, and $U(0) \geq 0$. Let $\widetilde{q}$ be the right-continuous version of q, and let P be the price distribution over $[0, 1]$ with $\Pr_{p\sim P}\{p \leq x\} = \widetilde{q}(x) + 1 - \widetilde{q}(1)$ for every $x \in [0, 1]$, which puts the remaining probability $1 - \widetilde{q}(1)$ on price 0 and only depends on M. As $\widetilde{q} \geq q$, and they differ on at most countably many points,

$$\pi(v) = q(v)v - U(0) - \int_0^v q(y)\,dy \leq \widetilde{q}(v)v - \int_0^v \widetilde{q}(y)\,dy = \mathop{\mathbb{E}}_{p\sim P}[p\mathbb{1}[v \geq p]]\,,$$

where the equality is the payment $\pi(v)$ in Lemma A.1 for this P, in which the constant $1-\widetilde{q}(1)$ cancels. Finally, by Fubini's theorem, $\mathbb{E}_{v\sim G}[\mathbb{E}_{p\sim P}[p\mathbb{1}[v \geq p]]] = \mathbb{E}_{p\sim P}[p \Pr_{v\sim G}\{v \geq p\}] = \mathbb{E}_{p\sim P}[R_G(p)]$, so the regret is at least $R_G(p^*_G) - \mathbb{E}_{p\sim P}[R_G(p)] = \mathbb{E}_{p\sim P}[R_G(p^*_G) - R_G(p)]$. □

Theorem A.4 (Infinite Menus, Full Feedback, and Both)**.** *Consider three kinds of dynamic pricing algorithms: (a) those offering infinite menus, which observe the chosen options; (b) those posting a single random price in each round t and observing v_t afterwards, i.e., full feedback; and (c) those offering infinite menus and also observing full feedback. Let $\mathcal{A}$ be an algorithm of one kind, and let $\gamma \in (0, 1]$ be an arbitrarily small but fixed constant. Then each of the other two kinds has an algorithm $\mathcal{A}'$ such that, for every T and every sequence of realized values $v_1, v_2, \ldots, v_T \in [0, 1]$ (not necessarily i.i.d.), the expected revenue of $\mathcal{A}'$ in the first T rounds is at least that of $\mathcal{A}$ minus γ.*

Proof. Algorithms of kinds (a) and (b) are also of kind (c), as an algorithm may ignore v_t, and a posted price is a size-1 menu. So it suffices to turn an algorithm $\mathcal{A}$ of kind (c) into one of kind (b) without loss, and into one of kind (a) with loss at most γ.

Fix the values $v_1, v_2, \ldots$. In round t, let M_t be the menu of $\mathcal{A}$, which depends on $v_1, v_2, \ldots, v_{t-1}$ and the internal randomness of $\mathcal{A}$, and let P_t be the price distribution given by Lemma A.3 for M_t, so that the payment of a customer with value v under M_t is at most $\mathbb{E}_{p\sim P_t}[p\mathbb{1}[v \geq p]]$ for every v. Both algorithms below compute M_t in round t, and after round t, they compute the option that v_t chooses from M_t and feed it to $\mathcal{A}$ together with v_t. So $\mathcal{A}$ receives the same feedback as when it runs by itself, and its menus have the same distribution.

- Kind (b): In round t, post a price drawn from P_t using fresh randomness. Conditioning on M_t, its expected revenue in round t is $\mathbb{E}_{p\sim P_t}[p\mathbb{1}[v_t \geq p]]$, which is at least that under M_t.
- Kind (a): In round t, offer the menu in Lemma A.1 built from $(1-\gamma_t)P_t + \gamma_t \operatorname{Unif}[0,1]$, where $\gamma_t := \gamma 2^{-t}$. As this price distribution has full support on $[0,1]$, the chosen option reveals v_t. Its revenue in round t is at least $(1-\gamma_t)\,\mathbb{E}_{p\sim P_t}[p\mathbb{1}[v_t \geq p]]$, which is at least the payment under M_t minus γ_t.

Summing over the first T rounds, the first algorithm loses nothing compared with $\mathcal{A}$, and the second one loses at most $\sum_t \gamma_t \leq \gamma$. □

As Theorem A.4 holds on every sequence of values, it also holds for i.i.d. values from every value distribution. Hence every regret bound for one kind, whether minimax over a class, local over a pair $\{G_+, G_-\}$, or on a fixed value distribution, also holds for the other two kinds up to an arbitrarily small additive constant. In particular, for every class $\mathcal{F}$, the minimax regret $\operatorname{Reg}^{(\infty)}_{\mathcal{F}}(T)$ of infinite menus is also that of full feedback, with or without menus.

Consequently, the minimax regret of infinite menus is governed by the sample complexity (defined in Appendix B.1), as we show in Reduction A.5.

Reduction A.5 (Revenue Maximization with Infinite Menus ↔ Sample Complexity)**.** *Let $\mathcal{F}$ be a class of value distributions.*

(a) *If for every $n \geq 1$, some estimator $\widehat{p}$ from n samples has $\mathbb{E}[R(p^*) - R(\widehat{p})] \leq \epsilon_n$ under every $F \in \mathcal{F}$, then $\operatorname{Reg}^{(\infty)}_{\mathcal{F}}(T) \leq 2 + \sum_{n=1}^{T-1} \epsilon_n$.*

(b) *Conversely, if an algorithm with infinite menus has regret at most ϵT in T rounds under every $F \in \mathcal{F}$, then $\textsc{Sample}_{\mathcal{F}}(\epsilon) \leq T - 1$.*

In particular, for $c \geq 1$, $\epsilon_n = \widetilde{\mathcal{O}}(n^{-1/c})$ gives $\operatorname{Reg}^{(\infty)}_{\mathcal{F}}(T) = \widetilde{\mathcal{O}}(T^{1-1/c})$, and $\operatorname{Reg}^{(\infty)}_{\mathcal{F}}(T) = \widetilde{\mathcal{O}}(T^{1-1/c})$ gives $\textsc{Sample}_{\mathcal{F}}(\epsilon) = \widetilde{\mathcal{O}}(\epsilon^{-c})$.

Proof. (a) Samples to Regret. Consider the algorithm with full feedback that posts an arbitrary price in round 1 and, in each round $t \geq 2$, posts the estimator $\widehat{p}_t$ from $v_1, v_2, \ldots, v_{t-1}$. Its regret is at most $1 + \sum_{t=2}^{T} \mathbb{E}[R(p^*) - R(\widehat{p}_t)] \leq 1 + \sum_{n=1}^{T-1} \epsilon_n$, so Theorem A.4 with $\gamma = 1$ gives an algorithm with infinite menus and regret at most $2 + \sum_{n=1}^{T-1} \epsilon_n$.

(b) Regret to Samples. Given $T-1$ samples, we run the algorithm for T rounds, using the t-th sample as v_t to compute the option chosen in each round $t \leq T-1$. As the menu in each round $t \leq T$ only depends on $v_1, v_2, \ldots, v_{t-1}$ and the internal randomness, this gives the menus of all T rounds with the same distribution as in a real run. Draw $t \in [T]$ uniformly at random, and output $\widehat{p} \sim P_t$, where P_t is given by Lemma A.3 for the menu in round t. As this menu is independent of v_t, Lemma A.3 bounds the expected revenue in round t by $\mathbb{E}[R(\widehat{p}) \mid t]$. Averaging over t, $\mathbb{E}[R(\widehat{p})]$ is at least $\frac{1}{T}$ times the expected revenue in the T rounds, which is at least $R(p^*) - \epsilon$. Hence $\mathbb{E}[R(p^*) - R(\widehat{p})] \leq \epsilon$, i.e., $\textsc{Sample}_{\mathcal{F}}(\epsilon) \leq T - 1$.

In particular. We have $\sum_{n<T} \widetilde{\mathcal{O}}(n^{-1/c}) = \widetilde{\mathcal{O}}(T^{1-1/c})$ for $c \geq 1$. If $\mathrm{Reg}_{\mathcal{F}}^{(\infty)}(T) = \widetilde{\mathcal{O}}(T^{1-1/c})$, then some $T = \widetilde{\mathcal{O}}(\epsilon^{-c})$ has $\mathrm{Reg}_{\mathcal{F}}^{(\infty)}(T) \leq \epsilon T$, so (b) gives $\textsc{Sample}_{\mathcal{F}}(\epsilon) = \widetilde{\mathcal{O}}(\epsilon^{-c})$. □

Theorem A.6 (Revenue Maximization with Infinite Menus)**.** *With infinite menus, the minimax regret, the sample complexity of the monopoly price, and that of the CDF are as follows:*

(a) *General and regular distributions:* $\mathrm{Reg}_{\mathrm{all}}^{(\infty)}(T), \mathrm{Reg}_{\mathrm{regular}}^{(\infty)}(T) = \Theta(\sqrt{T})$, $\textsc{Sample}_{\mathrm{all}}, \textsc{Sample}_{\mathrm{regular}} = \Theta(\epsilon^{-2})$, *and* $\textsc{Sample}_{\mathrm{all}}^{\mathrm{CDF}}, \textsc{Sample}_{\mathrm{regular}}^{\mathrm{CDF}} = \Theta(\epsilon^{-2})$.

(b) *MHR distributions:* $\mathrm{Reg}_{\mathrm{MHR}}^{(\infty)}(T) = \widetilde{\Theta}(T^{1/3})$, $\textsc{Sample}_{\mathrm{MHR}} = \widetilde{\Theta}(\epsilon^{-1.5})$, *and* $\textsc{Sample}_{\mathrm{MHR}}^{\mathrm{CDF}} = \Theta(\epsilon^{-2})$.

(c) *Distributions with curvature* $\kappa \in (0, 2)$*:* $\mathrm{Reg}_{\mathrm{curv}(\kappa)}^{(\infty)}(T) = \widetilde{\Theta}(\kappa^{-1/3}T^{1/3})$, $\textsc{Sample}_{\mathrm{curv}(\kappa)} = \widetilde{\Theta}(\kappa^{-1/2}\epsilon^{-1.5})$, *and* $\textsc{Sample}_{\mathrm{curv}(\kappa)}^{\mathrm{CDF}} = \Theta(\epsilon^{-2})$.

Proof. All the sample complexities are in Theorem B.10, so it remains to bound the regret.

Upper Bounds. Those for $\mathcal{F}_{\mathrm{all}}$ also hold for $\mathcal{F}_{\mathrm{regular}}$. By Lemma B.12, some estimator from n samples has $\mathbb{E}[R(p^*) - R(\widehat{p})] = \mathcal{O}(n^{-1/2})$ under every value distribution, so Reduction A.5(a) gives $\mathrm{Reg}_{\mathrm{all}}^{(\infty)}(T) = \mathcal{O}(\sqrt{T})$. For MHR and curvature distributions, as infinite menus include size-2 menus, $\mathrm{Reg}_{\mathcal{F}}^{(\infty)}(T) \leq \mathrm{Reg}_{\mathcal{F}}^{(2)}(T)$, so Theorem 3.4 gives the upper bounds.

Lower Bounds. Those for $\mathcal{F}_{\mathrm{regular}}$ also hold for $\mathcal{F}_{\mathrm{all}}$. They are in Theorem 4.1, where Lemma D.3 gives the one for $\mathcal{F}_{\mathrm{regular}}$. □

B Query Complexities

This appendix collects all our results on query complexities. We define the complexities in Table 2 (Appendix B.1) and relate them to each other and to the regret (Appendix B.2). We then state the two lower-bound methods, used throughout this section when proving lower bounds (Appendix B.3). We then prove the bounds in Table 2 for samples (Appendix B.4), interval queries (Appendix B.5), and threshold queries (Appendix B.6).

B.1 Definitions: Sample, Interval-Query, and Threshold-Query Complexities

We define query complexity in a unified way. In each round t, a new customer arrives, with value $v_t \sim F$ like in dynamic pricing. The algorithm issues a *query* to the customer, chosen from some class of feasible queries, and learns something about v_t. The goal is to learn a property of interest: either the CDF of value distribution F, or just the monopoly price p_F^*. The *query complexity* is the smallest number of queries guaranteed to learn the property within a given distance $\epsilon > 0$ (ϵ-*learn* it) whenever F belongs to a given class $\mathcal{F}$ of value distributions.

The precise meaning of "ϵ-learn" depends on the property. For monopoly-price estimation, it means to find a (randomized) price $\widehat{p}$ whose expected revenue is ϵ-optimal:[12]

$$R(p^*) - \mathbb{E}[R(\widehat{p})] \leq \epsilon, \quad \text{where } p^* \text{ is the monopoly price in Eq. (2)}, \tag{12}$$

[12]Similar to Huang, Mansour, and Roughgarden [HMR15], we define the complexities of the monopoly price via a randomized price whose expected revenue is near-optimal (where the expectation is over both the samples, or queries, and the randomness of $\widehat{p}$). By Markov's inequality, Eq. (12) implies $\Pr\{R(p^*) - R(\widehat{p}) \leq 4\epsilon\} \geq \frac{3}{4}$, a guarantee with constant success probability as the version of Leme, Sivan, Teng, and Worah [LSTW23b]. This is for the ease of reduction to regret.

and for CDF estimation, it means to construct a value distribution $\widehat{F}$ that is, with probability $\frac{3}{4}$, ϵ-close to F in Lévy distance, i.e.,

$$D_F(x+\epsilon) - \epsilon \le D_{\widehat{F}}(x) \le D_F(x-\epsilon) + \epsilon, \quad \forall x \in [0,1]. \tag{13}$$

We consider three query classes: *threshold queries* that reveal $\mathbb{1}[v_t \ge p]$ for a given threshold p, *value queries* that directly reveal v_t, and an intermediate class of *interval queries* that reveal whether v_t lies in a given interval $[a,b)$. The corresponding query complexity is denoted as, resp., $\textsc{Threshold}_{\mathcal{F}}(\epsilon)$, $\textsc{Sample}_{\mathcal{F}}(\epsilon)$, $\textsc{Interval}_{\mathcal{F}}(\epsilon)$ for monopoly-price estimation, and $\textsc{Threshold}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon)$, $\textsc{Sample}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon)$, $\textsc{Interval}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon)$ for CDF estimation.[13]

Finally, we present the following fact justifying the range $\kappa \in (0,2)$ in Section 2:

Lemma B.1 (Range of Curvature)**.** *The class $\mathcal{F}_{\mathrm{curv}(\kappa)}$ is empty for $\kappa > 2$, and every $F \in \mathcal{F}_{\mathrm{curv}(2)}$ has $p^* = \frac{1}{2}$, so $\textsc{Sample}_{\mathrm{curv}(\kappa)}(\epsilon) = 0$ for $\kappa = 2$.*

Proof. Let $F \in \mathcal{F}_{\mathrm{curv}(\kappa)}$. Taking $p \in \{0,1\}$ in the definition of $\mathcal{F}_{\mathrm{curv}(\kappa)}$ gives $\kappa(p^*)^2 \le R(p^*)$ and $\kappa(1-p^*)^2 \le R(p^*) - R(1) \le R(p^*)$, where $R(p^*) \le p^*$ (so $p^* > 0$). Hence $\kappa \le \min\{\frac{1}{p^*}, \frac{p^*}{(1-p^*)^2}\}$, which is less than 2 unless $p^* = \frac{1}{2}$, as the first term is decreasing and the second is increasing in p^*, and both are 2 at $p^* = \frac{1}{2}$. So $\mathcal{F}_{\mathrm{curv}(\kappa)}$ is empty for $\kappa > 2$, and every $F \in \mathcal{F}_{\mathrm{curv}(2)}$ has $p^* = \frac{1}{2}$, in which case outputting $\widehat{p} = \frac{1}{2}$ needs no sample. □

B.2 Relations and Reductions Between Complexities

Samples reveal more than interval queries, which reveal more than threshold queries; an accurate CDF gives a near-optimal price; an algorithm with small regret gives a near-optimal price; and conversely, an estimator of the monopoly price or of the CDF gives an algorithm with small regret. These together give the black-box reductions in Figure 1, formalized as follows.

Theorem B.2 (Relations Between Complexities)**.** *For every class $\mathcal{F}$ and $\epsilon \in (0,1]$:*

(a) $\textsc{Sample}_{\mathcal{F}}(\epsilon) \le \textsc{Interval}_{\mathcal{F}}(\epsilon) \le \textsc{Threshold}_{\mathcal{F}}(\epsilon)$, *and* $\textsc{Sample}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon) \le \textsc{Interval}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon) \le \textsc{Threshold}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon)$.

(b) $\textsc{Sample}_{\mathcal{F}}(6\epsilon) \le 11 \log \frac{2}{\epsilon} \cdot \textsc{Sample}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon)$, *and the same holds for interval and threshold queries.*

(c) *If an algorithm using size-m menus has T-round regret at most $\frac{\epsilon T}{2}$ under every $F \in \mathcal{F}$, where $T \ge \log \frac{2}{\epsilon}$, then the interval-query complexity $\textsc{Interval}_{\mathcal{F}}(\epsilon) \le 4(m+1)T$. Moreover, if an algorithm using posted prices has T-round regret at most ϵT under every $F \in \mathcal{F}$, then the threshold-query complexity $\textsc{Threshold}_{\mathcal{F}}(\epsilon) \le T$.*

(d) *If $\textsc{Interval}_{\mathcal{F}}(\epsilon) \le n(\epsilon)$ for every $\epsilon \in (0,1]$, where n is a non-increasing function, then $\mathrm{Reg}^{(3)}_{\mathcal{F}}(T) \le 2 + 2\sum_{j \ge 1} 2^{-j} \min\{n(2^{-j}), T\}$. The same holds for $\textsc{Threshold}_{\mathcal{F}}$ and $\mathrm{Reg}^{(2)}_{\mathcal{F}}(T)$ in place of $\textsc{Interval}_{\mathcal{F}}$ and $\mathrm{Reg}^{(3)}_{\mathcal{F}}(T)$.*

(e) *If $\textsc{Interval}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon) \le n(\epsilon)$ for every $\epsilon \in (0,1]$, where n is a non-increasing function, then $\mathrm{Reg}^{(3)}_{\mathcal{F}}(T) \le 2 + 2\sum_{j \ge 1} 2^{-j} \min\{11 \log(12 \cdot 2^j) \cdot n(2^{-j}/6), T\}$.*

Proof. (a) is Lemma B.3, (b) is Reduction B.4, (c) is Reduction B.6, and (d) is Lemma B.7. For (e), (b) gives $\textsc{Interval}_{\mathcal{F}}(\epsilon) \le 11 \log \frac{12}{\epsilon} \cdot n(\frac{\epsilon}{6})$ for every $\epsilon \in (0,1]$, which is non-increasing in ϵ, so (d) applies. □

[13] Recall that *sample complexity* refers to query complexity with value queries.

Lemma B.3 (Samples, Interval Queries, and Threshold Queries)**.** *For every class $\mathcal{F}$ and $\epsilon \in (0,1]$, $\textsc{Sample}_{\mathcal{F}}(\epsilon) \le \textsc{Interval}_{\mathcal{F}}(\epsilon) \le \textsc{Threshold}_{\mathcal{F}}(\epsilon)$, and $\textsc{Sample}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon) \le \textsc{Interval}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon) \le \textsc{Threshold}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon)$.*

Proof. A threshold query at p is answered by the interval query on $[0,p)$, as $\mathbb{1}[v \ge p] = 1 - \mathbb{1}[v \in [0,p)]$, and an interval query on $[a,b)$ is answered by a sample v, as $\mathbb{1}[v \in [a,b)]$ is a function of v. So an algorithm using n threshold queries can be run with n interval queries, and one using n interval queries can be run with n samples, without changing the distribution of its output. □

Reduction B.4 (Monopoly Price → CDF)**.** *Suppose that $\widehat{F}$ is ϵ-close to F in Lévy distance (Eq. (13)), and let $\widehat{p} := \max\{\widehat{p}^* - \epsilon, 0\}$, where $\widehat{p}^*$ maximizes $\widehat{R}(p) := p \Pr_{v\sim\widehat{F}}\{v \ge p\}$. Then $R(p^*) - R(\widehat{p}) \le 5\epsilon$. Consequently, $\textsc{Sample}_{\mathcal{F}}(6\epsilon) \le 11 \log \frac{2}{\epsilon} \cdot \textsc{Sample}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon)$ for every class $\mathcal{F}$ and $\epsilon \in (0,1]$, and the same holds for interval and threshold queries.*

Proof. We show that the estimated revenue $\widehat{R}(\widehat{p}^*)$ is at least $R(p^*) - 3\epsilon$, and that lowering $\widehat{p}^*$ by ϵ keeps at least $\widehat{R}(\widehat{p}^*) - 2\epsilon$ of the true revenue. Here $\widehat{p}^*$ exists, as $\widehat{R}$ is upper semi-continuous.

If $p^* < 2\epsilon$, then $R(p^*) \le p^* < 2\epsilon$ and the claim is trivial, so we assume $p^* \ge 2\epsilon$. Using Lévy distance, $D_{\widehat{F}}(p^* - 2\epsilon) \ge D_F(p^* - \epsilon) - \epsilon \ge D_F(p^*) - \epsilon$, so

$$\widehat{R}(\widehat{p}^*) \ge \widehat{R}(p^* - 2\epsilon) = (p^* - 2\epsilon) D_{\widehat{F}}(p^* - 2\epsilon) \ge (p^* - 2\epsilon)\left(D_F(p^*) - \epsilon\right) \ge R(p^*) - 3\epsilon.$$

If $\widehat{p}^* < \epsilon$, then $R(p^*) \le \widehat{R}(\widehat{p}^*) + 3\epsilon \le \widehat{p}^* + 3\epsilon < 4\epsilon$ and the claim is again trivial, so we assume $\widehat{p}^* \ge \epsilon$, which gives $\widehat{p} = \widehat{p}^* - \epsilon$. Again using Lévy distance, $D_F(\widehat{p}) = D_F(\widehat{p}^* - \epsilon) \ge D_{\widehat{F}}(\widehat{p}^*) - \epsilon$, so

$$R(\widehat{p}) \ge \widehat{p}\left(D_{\widehat{F}}(\widehat{p}^*) - \epsilon\right) \ge \widehat{p}^* D_{\widehat{F}}(\widehat{p}^*) - 2\epsilon = \widehat{R}(\widehat{p}^*) - 2\epsilon.$$

Therefore, $R(\widehat{p}) \ge R(p^*) - 5\epsilon$.

To complete the proof, we use Lemma B.5 stated below to "boost" a CDF estimator (which only succeeds with probability $\frac{3}{4}$). Specifically, let k be the smallest odd integer at least $8 \log \frac{2}{\epsilon}$, so that $k \le 11 \log \frac{2}{\epsilon}$ and $e^{-k/8} \le \frac{\epsilon}{2}$. Run the CDF estimator k times on fresh samples or queries, and output the above $\widehat{p}$ for the median $\widetilde{F}$ of its k outputs, defined in Lemma B.5. Lemma B.5 then guarantees that $\widetilde{F}$ is ϵ-close to F with probability $1 - \frac{\epsilon}{2}$. We hence get $\mathbb{E}[R(p^*) - R(\widehat{p})] \le 5\epsilon + \frac{\epsilon}{2} \le 6\epsilon$. □

Lemma B.5 (Boosting CDF Estimates)**.** *Let k be odd, and let $\widehat{F}_1, \widehat{F}_2, \ldots, \widehat{F}_k$ be independent value distributions, each ϵ-close to F in Lévy distance with probability at least $\frac{3}{4}$. Then their median $\widetilde{F}$, the value distribution with $D_{\widetilde{F}}(x) := \mathrm{median}\{D_{\widehat{F}_1}(x), D_{\widehat{F}_2}(x), \ldots, D_{\widehat{F}_k}(x)\}$, is ϵ-close to F with probability at least $1 - e^{-k/8}$.*

Proof. By Hoeffding's inequality, more than $\frac{k}{2}$ of the $\widehat{F}_i$'s are ϵ-close to F with probability at least $1 - e^{-2k(1/4)^2} = 1 - e^{-k/8}$. On this event, for every $x \in [0,1]$, more than half of $D_{\widehat{F}_1}(x), \ldots, D_{\widehat{F}_k}(x)$ lie in $[D_F(x+\epsilon) - \epsilon, D_F(x-\epsilon) + \epsilon]$, and hence so does their median. Finally, $\widetilde{F}$ is a value distribution, i.e., $D_{\widetilde{F}}$ is non-increasing and left-continuous with $D_{\widetilde{F}}(0) = 1$, as the median of k numbers is non-decreasing and continuous in each of them. □

Reduction B.6 (Interval-Query and Threshold-Query Complexities → Revenue Maximization with Finite Menus)**.** *Suppose that an algorithm using size-m menus has T-round regret at most $\frac{\epsilon T}{2}$ under*

every $F \in \mathcal{F}$, where m is a positive integer and $T \geq \log \frac{2}{\epsilon}$. Then the interval-query complexity $\textsc{Interval}_{\mathcal{F}}(\epsilon) \leq 4(m+1)T$. *Moreover, if an algorithm using posted prices ($m = 1$) has T-round regret at most ϵT under every $F \in \mathcal{F}$, then the threshold-query complexity* $\textsc{Threshold}_{\mathcal{F}}(\epsilon) \leq T$.

Proof. We first simulate the algorithm for T rounds by queries, as detailed at the end of this proof. Then we draw $t \in [T]$ uniformly at random and output $\widehat{p} \sim P_t$, where P_t is given by Lemma A.3 for the menu in round t. As this menu is independent of v_t, Lemma A.3 bounds the expected revenue in round t by $\mathbb{E}[R(\widehat{p}) \mid t]$, so averaging over t, $\mathbb{E}[R(p^*) - R(\widehat{p})]$ is at most the regret divided by T if all T rounds are simulated. It remains to simulate the option chosen in each round by a few queries.

Options as Intervals. As the allocation probability of the chosen option is non-decreasing in the value, and ties are broken towards larger allocation probabilities, the values choosing each of the at most $m+1$ options of a menu (including the outside option) form an interval $[a, b)$ or $[a, 1]$.

Rejection Sampling. If $m = 1$, these are $[0, a)$ and $[a, 1]$, so a threshold query at a reveals the chosen option. Then the algorithm takes exactly T threshold queries, and $\mathbb{E}[R(p^*) - R(\widehat{p})] \leq \epsilon$. In general, we simulate each round by rejection sampling: We repeatedly pick one of these $m+1$ intervals uniformly at random and query it on a fresh sample until the answer is 1, where querying $[a, 1]$ means negating the query on $[0, a)$, and then feed the option of the last queried interval to the algorithm. Each attempt stops at an interval I with probability $\frac{D_F(I)}{m+1}$, so the fed option has the same distribution as the option chosen by a fresh value, and each attempt succeeds with probability exactly $\frac{1}{m+1}$, independently of the others. Hence the simulation does not change the distribution of the algorithm.

Stopping. We stop after $4(m+1)T$ queries, and output an arbitrary price if the T rounds are not all simulated by then. This only happens if fewer than T of the $4(m+1)T$ attempts succeed, which by the Chernoff bound has probability at most $e^{-(3/4)^2 \cdot 4T/2} \leq e^{-T} \leq \frac{\epsilon}{2}$, as $T \geq \log \frac{2}{\epsilon}$. Therefore, $\mathbb{E}[R(p^*) - R(\widehat{p})] \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. □

Conversely, an estimator using interval (or threshold) queries gives an algorithm using size-3 (or size-2) menus, as Lemma 3.1 lets each menu ask a query while exploiting the previous estimate:

Lemma B.7 (Revenue Maximization → Interval-Query Complexity)**.** *Let $\mathcal{F}$ be a class, and let $n : (0, 1] \to [0, \infty)$ be non-increasing. If* $\textsc{Interval}_{\mathcal{F}}(\epsilon) \leq n(\epsilon)$ *for every $\epsilon \in (0, 1]$, then for every $T \geq 1$,*

$$\mathrm{Reg}^{(3)}_{\mathcal{F}}(T) \leq 2 + 2 \sum_{j \geq 1} 2^{-j} \min\left\{n(2^{-j}), T\right\}.$$

If $\textsc{Threshold}_{\mathcal{F}}(\epsilon) \leq n(\epsilon)$ *for every $\epsilon \in (0, 1]$, then the same bound holds for* $\mathrm{Reg}^{(2)}_{\mathcal{F}}(T)$.

Proof. For $j = 1, 2, \ldots$, let $\epsilon_j := 2^{-j}$, and fix an algorithm that uses at most $n(\epsilon_j)$ interval queries and outputs a $\widehat{p}_j$ with $\mathbb{E}[R(p^*) - R(\widehat{p}_j)] \leq \epsilon_j$ under every $F \in \mathcal{F}$. Consider the algorithm using size-3 menus that runs these algorithms one after another, in phases $j = 1, 2, \ldots$ until the T rounds end, and answers each interval query $[a, b)$ of phase j in one round by the menu built from the price distribution $\gamma\delta_a + \gamma\delta_b + (1 - 2\gamma)\delta_{\widehat{p}_{j-1}}$, where $\gamma := \frac{1}{T}$ and $\widehat{p}_0 := 0$. By Lemma 3.1, this menu reveals whether the value lies in $[a, b)$, so each phase runs its algorithm on fresh values, and the menu earns at least $(1 - 2\gamma)R(\widehat{p}_{j-1}) \geq R(\widehat{p}_{j-1}) - 2\gamma$ in expectation. As $\widehat{p}_{j-1}$ only depends on the earlier phases, each round of phase j has expected regret at most $\epsilon_{j-1} + 2\gamma = 2\epsilon_j + 2\gamma$, where $\epsilon_0 := 1$ bounds the gap of $\widehat{p}_0$. As phase j has at most $\min\{n(\epsilon_j), T\}$ rounds, the regret is at most

$2\gamma T + \sum_j 2\epsilon_j \min\{n(\epsilon_j), T\}$, which is the claimed bound. For threshold queries, the same argument applies to the size-2 menu built from $\gamma\delta_a + (1-\gamma)\delta_{\widehat{p}_{j-1}}$ for a threshold query at a. □

B.3 Two-Point and Needle-in-a-Haystack Methods

Throughout, for a value distribution G, its demand is $D_G(p) := \Pr_{v\sim G}\{v \geq p\}$, and $R_G(p^*_G) - R_G(p)$ is the *gap* of a price p under G. For $G \ll G'$ (i.e., G's support is a subset of that of G'), let $\chi^2(G\|G') := \int (\frac{dG}{dG'} - 1)^2\, dG'$, so that $H^2(G, G') \leq \frac{1}{2}\chi^2(G\|G')$. As $1 - H^2$ is multiplicative over independent samples, n samples have total variation distance at most $\sqrt{2nH^2(G, G')}$ between G and G'.

A two-point lower bound takes two value distributions that are close but share no near-optimal price:

Lemma B.8 (Two-Point Method for Complexities)**.** *Let G_+ and G_- be two value distributions in a class $\mathcal{F}$. If every price has gap at least 4ϵ either under G_+ or under G_-, that is,*

$$\max\left\{R_{G_+}(p^*_{G_+}) - R_{G_+}(p), R_{G_-}(p^*_{G_-}) - R_{G_-}(p)\right\} \geq 4\epsilon, \quad \forall p \in [0,1],$$

then $\textsc{Sample}_{\mathcal{F}}(\epsilon) \geq \frac{1}{8H^2(G_+,G_-)}$*. If no value distribution is ϵ-close to both G_+ and G_- in Lévy distance, then* $\textsc{Sample}^{\text{CDF}}_{\mathcal{F}}(\epsilon) \geq \frac{1}{8H^2(G_+,G_-)}$*.*

Proof. Prove by contradiction. Suppose that $n < \frac{1}{8H^2(G_+,G_-)}$ samples suffice, so that these samples have total variation distance at most $\sqrt{2nH^2(G_+, G_-)} < \frac{1}{2}$ between G_+ and G_-. Let $\widehat{p}$ be the randomized price found in Eq. (12), and let A be the event that $\widehat{p}$ has gap less than 4ϵ under G_+. By Markov's inequality, Eq. (12) gives $\Pr_{G_+}\{A\} \geq \frac{3}{4}$, and as every price in A has gap at least 4ϵ under G_-, it also gives $\Pr_{G_-}\{A\} \leq \frac{1}{4}$. Then $\Pr_{G_+}\{A\} - \Pr_{G_-}\{A\} \geq \frac{1}{2}$, a contradiction. Similarly, for CDF estimation, let $\widehat{F}$ be the value distribution found in Eq. (13), and let A be the event that $\widehat{F}$ is ϵ-close to G_+. Then A has probability at least $\frac{3}{4}$ under G_+ and at most $\frac{1}{4}$ under G_-, giving the same contradiction. □

We now introduce the following needle-in-a-haystack method [KSU19], which says that if one perturbs a distribution G into J decoys, each differing from G only on a small interval (disjoint across decoys), it is hard for any algorithm using few threshold queries to tell G apart from most of these decoys.

Lemma B.9 (Needle-in-a-Haystack Method)**.** *Let $G, F_1, F_2, \ldots, F_J$ be value distributions and $I_1, I_2, \ldots, I_J$ be disjoint sets, such that for every j, $D_{F_j} = D_G$ outside I_j, whereas $|D_{F_j} - D_G| \leq \eta$ and $D_{F_j}, D_G \in [\sigma, 1-\sigma]$ on I_j for some $\sigma \in (0, \frac{1}{2}]$. If the observations of an algorithm using n threshold queries have total variation distance at least $\frac{1}{3}$ between G and F_j for at least $\frac{3J}{4}$ indices j, then $n \geq \frac{\sigma J}{36\eta^2}$.*

Proof. A threshold query at p observes a Bernoulli variable with mean $D_G(p)$ or $D_{F_j}(p)$, whose KL divergence is 0 if $p \notin I_j$, and at most $\frac{(D_G(p) - D_{F_j}(p))^2}{D_{F_j}(p)(1-D_{F_j}(p))} \leq \frac{2\eta^2}{\sigma}$ otherwise. By the chain rule, the observations have KL divergence at most $\frac{2\eta^2}{\sigma} n_j$ between G and F_j, where n_j is the expected number of queries in I_j under G. As $\sum_j n_j \leq n$, at most $\frac{J}{4}$ indices have $n_j > \frac{4n}{J}$, so some j in the assumption has $n_j \leq \frac{4n}{J}$, and Pinsker's inequality gives $\frac{1}{3} \leq \sqrt{\frac{\eta^2 n_j}{\sigma}} \leq \sqrt{\frac{4\eta^2 n}{\sigma J}}$. □

B.4 Results on Sample Complexities

Several constructions below use the uniform distribution over $[\frac{1}{2}, 1]$, denoted by $U := \mathrm{Unif}[\frac{1}{2}, 1]$. Its demand is 1 on $[0, \frac{1}{2}]$ and $2 - 2p$ on $[\frac{1}{2}, 1]$, so that its revenue is $R_U(p) = \min\{p, 2p(1-p)\}$, maximized at $\frac{1}{2}$ with

$$\frac{1}{2} - R_U(p) \geq 2\left(p - \frac{1}{2}\right)^2 \qquad \text{for all } p \in [0,1]. \tag{14}$$

Hence $U \in \mathcal{F}_{\mathrm{curv}(2)} \subseteq \mathcal{F}_{\mathrm{curv}(\kappa)}$ for every $\kappa \in (0, 2]$, and so is every distribution F whose revenue function R_F satisfies $R_F(\frac{1}{2}) = R_U(\frac{1}{2})$, and $R_F(p) \leq R_U(p)$ for all $p \in [0, 1]$.

Theorem B.10 (Sample Complexities)**.** *For every $\kappa \in (0, 2)$, as $\epsilon \to 0$:*

(a) $\textsc{Sample}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon) = \Theta(\epsilon^{-2})$ *for each* $\mathcal{F} \in \{\mathcal{F}_{\mathrm{all}}, \mathcal{F}_{\mathrm{regular}}, \mathcal{F}_{\mathrm{MHR}}, \mathcal{F}_{\mathrm{curv}(\kappa)}\}$*;*
(b) $\textsc{Sample}_{\mathrm{all}}(\epsilon) = \Theta(\epsilon^{-2})$ *and* $\textsc{Sample}_{\mathrm{regular}}(\epsilon) = \Theta(\epsilon^{-2})$*;*
(c) $\textsc{Sample}_{\mathrm{MHR}}(\epsilon) = \widetilde{\Theta}(\epsilon^{-1.5})$*;*
(d) $\textsc{Sample}_{\mathrm{curv}(\kappa)}(\epsilon) = \widetilde{\Theta}(\kappa^{-1/2}\epsilon^{-1.5})$.

Proof. Upper Bounds. Those for $\mathcal{F}_{\mathrm{all}}$ also hold for its subclasses. Lemma B.11 gives (a), and Lemma B.12 gives (b). (c) is due to Huang, Mansour, and Roughgarden [HMR15], and (d) is Lemma B.13.

Lower Bounds. Those for $\mathcal{F}_{\mathrm{regular}}$ also hold for $\mathcal{F}_{\mathrm{all}}$. Lemma B.14 gives (b). For (a), Lemma B.16 gives the bound for MHR distributions, and hence for regular and general ones, as $\mathcal{F}_{\mathrm{MHR}} \subseteq \mathcal{F}_{\mathrm{regular}} \subseteq \mathcal{F}_{\mathrm{all}}$, and Lemma B.17 gives the bound for curvature distributions. (c) is due to Huang, Mansour, and Roughgarden [HMR15], and (d) is Lemma B.18. □

Lemma B.11 (Empirical CDF)**.** *Let $\widehat{F}_n$ be the empirical distribution of n samples from F, and $\mathrm{err}_n := \sup_x |\Pr_{v\sim\widehat{F}_n}\{v \leq x\} - \Pr_{v\sim F}\{v \leq x\}|$. Then $\widehat{F}_n$ is err_n-close to F in Lévy distance, and $\Pr\{\mathrm{err}_n > x\} = \mathcal{O}(e^{-2nx^2})$ for every $x > 0$. In particular, $\textsc{Sample}^{\mathrm{CDF}}_{\mathcal{F}}(\epsilon) = \mathcal{O}(\epsilon^{-2})$ for every class $\mathcal{F}$.*

Proof. For every $x \in [0, 1]$, taking left limits gives $|D_{\widehat{F}_n}(x) - D_F(x)| \leq \mathrm{err}_n$, so $D_F(x + \mathrm{err}_n) - \mathrm{err}_n \leq D_F(x) - \mathrm{err}_n \leq D_{\widehat{F}_n}(x) \leq D_F(x) + \mathrm{err}_n \leq D_F(x - \mathrm{err}_n) + \mathrm{err}_n$, so $\widehat{F}_n$ is err_n-close to F. The tail bound is the Dvoretzky–Kiefer–Wolfowitz inequality [DKW56], so some $n = \mathcal{O}(\epsilon^{-2})$ has $\Pr\{\mathrm{err}_n > \epsilon\} \leq \frac{1}{4}$. □

Lemma B.12 (Empirical Revenue)**.** *Let $\widehat{p}$ maximize the empirical revenue $\widehat{R}(p) := p \Pr_{v\sim\widehat{F}_n}\{v \geq p\}$ of n samples from F. Then $R(p^*) - R(\widehat{p}) \leq 2\mathrm{err}_n$, where err_n is as in Lemma B.11, so $\mathbb{E}[R(p^*) - R(\widehat{p})] = \mathcal{O}(n^{-1/2})$. In particular, $\textsc{Sample}_{\mathrm{all}}(\epsilon) = \mathcal{O}(\epsilon^{-2})$.*

Proof. For every $p \in [0, 1]$, $|\widehat{R}(p) - R(p)| \leq p|\Pr_{v\sim\widehat{F}_n}\{v < p\} - \Pr_{v\sim F}\{v < p\}| \leq \mathrm{err}_n$, as err_n also bounds the difference of the left limits. Hence $R(p^*) - R(\widehat{p}) \leq \widehat{R}(p^*) - \widehat{R}(\widehat{p}) + 2\mathrm{err}_n \leq 2\mathrm{err}_n$. By Lemma B.11, $\mathbb{E}[\mathrm{err}_n] = \int_0^\infty \Pr\{\mathrm{err}_n > x\}\, dx = \mathcal{O}(n^{-1/2})$, so some $n = \mathcal{O}(\epsilon^{-2})$ has $\mathbb{E}[R(p^*) - R(\widehat{p})] \leq \epsilon$. □

Lemma B.13 (Monopoly Price of Curvature Distributions via Samples)**.** *For every $\kappa \in (0, 2)$, $\textsc{Sample}_{\mathrm{curv}(\kappa)}(\epsilon) = \widetilde{\mathcal{O}}(\kappa^{-1/2}\epsilon^{-1.5})$.*

Proof. By Theorem 3.4(c), Algorithm 2, which uses size-2 and hence infinite menus, has regret at most $c(\kappa^{-1/3}T^{1/3}+\log T)\log^4 T$ under every $F\in\mathcal{F}_{\mathrm{curv}(\kappa)}$ for an absolute constant c. As $\kappa\le 2$, this is at most ϵT for some $T=\widetilde{\mathcal{O}}(\kappa^{-1/2}\epsilon^{-1.5})$, so Reduction A.5(b) gives $\textsc{Sample}_{\mathrm{curv}(\kappa)}(\epsilon)\le T-1$. □

Lemma B.14 (Sample Lower Bound for Regular Distributions)**.** *For every $\epsilon\in(0,\frac{1}{8}]$, $\textsc{Sample}_{\mathrm{regular}}(\frac{\epsilon}{25})\ge\frac{1}{32\epsilon^2}$, even for the class of value distributions that are both regular and quasi-MHR.*

Proof. By Lemma B.15 below, the two value distributions $F_\pm$ there are regular and quasi-MHR, every price has gap at least $\frac{\epsilon}{6}>4\cdot\frac{\epsilon}{25}$ under F_+ or under F_-, and $H^2(F_+,F_-)\le 4\epsilon^2$. So Lemma B.8 gives $\textsc{Sample}_{\mathrm{regular}}(\frac{\epsilon}{25})\ge\frac{1}{32\epsilon^2}$, even for the smaller class. □

The pair tilts the revenue plateau of a regular distribution to the left or to the right, which is also used for the regret (Lemma D.3):

Lemma B.15 (A Regular Pair)**.** *For $\epsilon\in(0,\frac{1}{8}]$, let $F_\pm$ be the value distributions with*

$$D_{F_\pm}(p)=1 \text{ on } [0,\tfrac12],\quad D_{F_\pm}(p)=\frac{1}{2p}\left(1\pm\epsilon\left(\frac{1}{p}-2\right)\right) \text{ on } [\tfrac12,\tfrac34],\quad D_{F_\pm}(p)=\frac{8}{3}\left(1\mp\frac{2\epsilon}{3}\right)(1-p) \text{ on } [\tfrac34,1].$$

Then F_+ and F_- are regular and quasi-MHR, $R_{F_\pm}(p^*_{F_\pm})\ge\frac12$, every price has gap at least $\frac{\epsilon}{6}$ under F_+ or under F_-, and $H^2(F_+,F_-)\le 4\epsilon^2$.

Proof. With $\epsilon=0$, this is a regular distribution F_0 whose revenue is $\frac12$ on the plateau $[\frac12,\frac34]$, and $F_\pm$ tilt this plateau in opposite directions, as $R_{F_\pm}=\frac12\pm\frac{\epsilon}{2}(\frac1p-2)$ on $[\frac12,\frac34]$. Their densities are $f_\pm=f_0(1\pm\epsilon g)$, where $f_0=\frac{1}{2p^2}$ and $g=\frac2p-2\in[\frac23,2]$ on $(\frac12,\frac34)$, and $f_0=\frac83$ and $g=-\frac23$ on $(\frac34,1)$. We check the class, the gaps, and the distance.

- *Regular.* The virtual values of $F_\pm$ are $\frac{\pm\epsilon}{1\pm\epsilon g}$ on $(\frac12,\frac34)$, which are non-decreasing as g is decreasing, and at most $2\epsilon\le\frac14$ in absolute value, while they are $2p-1\ge\frac12$ on $(\frac34,1)$.
- *Quasi-MHR*, i.e., $\frac1p\ln\frac{1}{D_{F_\pm}(p)}$ is non-decreasing [FJ25]: It is 0 on $(0,\frac12]$ and continuous, and on $(\frac12,1)$, its derivative has the sign of $\frac{pf_\pm}{D_{F_\pm}}-\ln\frac{1}{D_{F_\pm}}$, which is positive, as $\frac{pf_\pm}{D_{F_\pm}}\ge\frac{1-2\epsilon}{1+2\epsilon/3}>\frac23>\ln\frac{18}{11}\ge\ln\frac{1}{D_{F_\pm}}$ on $(\frac12,\frac34)$, and $\frac{pf_\pm}{D_{F_\pm}}=\frac{p}{1-p}\ge\ln\frac{1}{1-p}>\ln\frac{1}{D_{F_\pm}}$ on $(\frac34,1)$.
- *Gaps.* On $[\frac12,\frac34]$, R_{F_+} decreases from its maximum $\frac12$ and R_{F_-} increases to its maximum $\frac12+\frac{\epsilon}{3}$, so as $R_{F_+}+R_{F_-}=1$ there, the two gaps of every price there sum to $\frac{\epsilon}{3}$. Below $\frac12$, $R_{F_-}=p<\frac12$, and above $\frac34$, $R_{F_+}=(1-\frac{2\epsilon}{3})R_{F_0}\le\frac12-\frac{\epsilon}{3}$. Hence every price has gap at least $\frac{\epsilon}{6}$ under F_+ or under F_-, and both maxima are at least $\frac12$.
- *Distance.* $H^2(F_+,F_-)\le\epsilon^2\int f_0g^2\le 4\epsilon^2$. □

Lemma B.16 (CDF Sample Lower Bound for MHR Distributions)**.** *For every $\epsilon\in(0,\frac{1}{22}]$, $\textsc{Sample}^{\mathrm{CDF}}_{\mathrm{MHR}}(\epsilon)\ge\frac{1}{968\epsilon^2}$.*

Proof. We compare Unif$[0,1]$ with a slightly tilted distribution. Let $\eta:=22\epsilon\in(0,1]$, and let F_1 and F_2 be the value distributions with demands $D_{F_1}(p)=1-p$ and $D_{F_2}(p)=1-p^{1+\eta}$ on $[0,1]$.

- *MHR.* Both are MHR, as $\frac{(1+\eta)x^\eta}{1-x^{1+\eta}}$ is non-decreasing.
- *Lévy separation.* At $x=e^{-1}$, we have $D_{F_2}(x)-D_{F_1}(x)=e^{-1}(1-e^{-\eta})\ge\frac{\eta}{2e}>4\epsilon$, so $D_{F_2}(x)>D_{F_1}(x-2\epsilon)+2\epsilon$, i.e., F_1 and F_2 are not 2ϵ-close in Lévy distance, and no value distribution is ϵ-close to both.

- *Distance.* As $\frac{1+\eta}{(1+\eta/2)^2} = 1 - \frac{\eta^2}{(2+\eta)^2}$, we have $H^2(F_1, F_2) = 1 - \int_0^1 \sqrt{(1+\eta)x^\eta}\, dx = 1 - \frac{\sqrt{1+\eta}}{1+\eta/2} \le \frac{\eta^2}{4}$.

So Lemma B.8 gives $\textsc{Sample}^{\text{CDF}}_{\text{MHR}}(\epsilon) \ge \frac{1}{2\eta^2} = \frac{1}{968\epsilon^2}$. □

Lemma B.17 (CDF Sample Lower Bound for Curvature Distributions)**.** *For every $\kappa \in (0, 2]$ and $\epsilon \in (0, \frac{1}{33}]$, $\textsc{Sample}^{\text{CDF}}_{\text{curv}(\kappa)}(\epsilon) \ge \frac{1}{2200\epsilon^2}$.*

Proof. We map the pair in Lemma B.16 into $[\frac{1}{2}, 1]$. Let $\eta := 33\epsilon \in (0, 1]$, let F_1 and F_2 be the value distributions with demands $1-p$ and $1-p^{1+\eta}$ on $[0, 1]$, and let G_i be the distribution of $1 - \frac{v}{2}$ for $v \sim F_i$.

- *Curvature.* $G_1 = U$, and $D_{G_2}(p) = \min\{1, (2-2p)^{1+\eta}\} \le D_{G_1}(p)$ with equality at $\frac{1}{2}$, so $R_{G_2} \le R_U$ with equality at $\frac{1}{2}$. Hence $G_1, G_2 \in \mathcal{F}_{\text{curv}(2)} \subseteq \mathcal{F}_{\text{curv}(\kappa)}$ by Eq. (14).
- *Lévy separation.* At $x := 1 - \frac{1}{2e}$, we have $D_{G_1}(x) - D_{G_2}(x) = e^{-1}(1 - e^{-\eta}) > 6\epsilon$, whereas $D_{G_1}(x + 2\epsilon) = D_{G_1}(x) - 4\epsilon$, so $D_{G_2}(x) < D_{G_1}(x + 2\epsilon) - 2\epsilon$, and no value distribution is ϵ-close to both.
- *Distance.* As the map $v \mapsto 1 - \frac{v}{2}$ is injective, $H^2(G_1, G_2) = H^2(F_1, F_2) = 1 - \frac{\sqrt{1+\eta}}{1+\eta/2} \le \frac{\eta^2}{4}$.

So Lemma B.8 gives $\textsc{Sample}^{\text{CDF}}_{\text{curv}(\kappa)}(\epsilon) \ge \frac{1}{2\eta^2} \ge \frac{1}{2200\epsilon^2}$. □

Lemma B.18 (Sample Lower Bound for Curvature Distributions)**.** *For every $\kappa \in (0, 2)$ and $\epsilon \le \frac{1}{1024}\min\{\kappa, (2-\kappa)^2\}$, we have $\textsc{Sample}_{\text{curv}(\kappa)}(\epsilon) \ge \frac{1}{1.1 \cdot 10^5 \sqrt{\kappa}\, \epsilon^{1.5}}$.*

Proof. We use the two pairs in Lemmas B.19 and B.20 below, depending on κ.

- If $\kappa \le \frac{1}{8}$, let $s := \sqrt{4\epsilon/\kappa}$, which is at most $\frac{1}{16}$ as $\epsilon \le \frac{\kappa}{1024}$. Under the pair $G_\pm$ in Lemma B.19, every price has gap at least $2\kappa s^2 = 8\epsilon > 4\epsilon$ under G_+ or under G_-, and $H^2(G_+, G_-) \le 1600\kappa^2 s^3$. So Lemma B.8 gives $\textsc{Sample}_{\text{curv}(\kappa)}(\epsilon) \ge \frac{1}{12800\kappa^2 s^3} \ge \frac{1}{1.1 \cdot 10^5 \sqrt{\kappa}\, \epsilon^{1.5}}$, as $\kappa^2 s^3 = 8\kappa^{1/2}\epsilon^{1.5}$.
- If $\kappa > \frac{1}{8}$, let $s := \sqrt{10\epsilon}$, which is at most $\frac{2-\kappa}{8}$ as $\epsilon \le \frac{(2-\kappa)^2}{640}$. Under the pair U, G_3 in Lemma B.20, every price has gap at least $\frac{s^2}{2} = 5\epsilon > 4\epsilon$ under U or under G_3, and $H^2(U, G_3) \le 24s^3$. So Lemma B.8 gives $\textsc{Sample}_{\text{curv}(\kappa)}(\epsilon) \ge \frac{1}{192 s^3} \ge \frac{1}{6100\epsilon^{1.5}}$, which is at least $\frac{1}{1.1 \cdot 10^5 \sqrt{\kappa}\, \epsilon^{1.5}}$ as $\kappa > \frac{1}{8}$. □

Lemma B.19 (A Pair with Small Curvature)**.** *Let $\kappa \in (0, \frac{1}{8}]$ and $s \in (0, \frac{1}{16}]$. Let $Q(p) := \frac{1}{4} - 2\kappa(p - \frac{1}{2})^2$, and let ψ be the odd function that equals x on $[0, s]$, $2s - x$ on $[s, 2s]$, and 0 on $[2s, \infty)$. Then there are value distributions $G_\pm$ with revenue functions $R_{G_\pm}(p) = \min\{p, Q(p) \pm 4\kappa s \psi(p - \frac{1}{2})\}$, such that $G_\pm \in \mathcal{F}_{\text{curv}(\kappa)}$ with monopoly prices $\frac{1}{2} \pm s$, every price has gap at least $2\kappa s^2$ under G_+ or under G_-, and $H^2(G_+, G_-) \le 1600\kappa^2 s^3$.*

Proof. These are valid revenue functions, i.e., $D_{G_\pm} = \frac{R_{G_\pm}}{p}$ is non-increasing, as $\frac{Q}{p}$ is decreasing where $Q < p$, and on $J := (\frac{1}{2} - 2s, \frac{1}{2} + 2s) \subseteq (\frac{3}{8}, \frac{5}{8})$, the derivative of $Q \pm 4\kappa s \psi(p - \frac{1}{2})$ is at most $8\kappa s + 4\kappa s \le \frac{3}{32} < \frac{R_{G_\pm}}{p}$.

- *Curvature.* For $x := p - \frac{1}{2}$, $R_{G_+}(\frac{1}{2} + s) - R_{G_+}(p) \ge 2\kappa(s^2 + x^2 - 2s\psi(x)) \ge \kappa(x - s)^2$, where the last step is checked on $|x| \le s$, $s < |x| \le 2s$, and $|x| > 2s$ separately. So $G_+ \in \mathcal{F}_{\text{curv}(\kappa)}$ with monopoly price $\frac{1}{2} + s$, and symmetrically, $G_- \in \mathcal{F}_{\text{curv}(\kappa)}$ with monopoly price $\frac{1}{2} - s$.
- *Gaps.* Both maxima are $\frac{1}{4} + 2\kappa s^2$, whereas $R_{G_+} + R_{G_-} = 2Q \le \frac{1}{2}$ on J and $R_{G_\pm} \le \frac{1}{4}$ elsewhere, so every price has gap at least $2\kappa s^2$ under G_+ or under G_-.

- *Distance.* $G_\pm$ only differ on J, where their densities are $f_0 \mp 4\kappa s(\frac{\psi(p-1/2)}{p})'$ with $f_0 := \frac{Q}{p^2} - \frac{Q'}{p} \geq \frac{2}{5}$ and $|(\frac{\psi(p-1/2)}{p})'| \leq \frac{1}{p} + \frac{s}{p^2} \leq \frac{28}{9}$. As $(\sqrt{a}-\sqrt{b})^2 \leq \frac{(a-b)^2}{a+b}$, we have $H^2(G_+, G_-) \leq \frac{1}{2} \cdot 4s \cdot \frac{(8\kappa s \cdot 28/9)^2}{4/5} \leq 1600\kappa^2 s^3$. □

For larger curvature, we shift the kink of R_U to the right, as in Huang, Mansour, and Roughgarden [HMR15, Theorem 4.12]:

Lemma B.20 (A Pair with Large Curvature)**.** *Let $\kappa \in (\frac{1}{8}, 2)$ and $s \in (0, \frac{2-\kappa}{8}]$. Let G_3 multiply the density of U by $1-4s$ on $[\frac{1}{2}, \frac{1}{2}+s)$ and by $1+\theta$ on $[\frac{1}{2}+s, 1]$, where $\theta := \frac{4s^2}{1/2-s} \leq 16s^2$ keeps the total mass 1. Then $U, G_3 \in \mathcal{F}_{\mathrm{curv}(\kappa)}$, every price has gap at least $\frac{s^2}{2}$ under U or under G_3, and $H^2(U, G_3) \leq 24s^3$.*

Proof. R_{G_3} is p on $[0, \frac{1}{2}]$, $p(1-2(1-4s)(p-\frac{1}{2}))$ on $[\frac{1}{2}, \frac{1}{2}+s]$, and $2(1+\theta)p(1-p)$ on $[\frac{1}{2}+s, 1]$.

- *Curvature.* R_{G_3} increases on $[0, \frac{1}{2}+s]$, as its left derivative at $\frac{1}{2}+s$ is $16s^2 > 0$, and it is quadratic with second derivative at most $-4(1-4s)$ on both sides of $\frac{1}{2}+s$, so $R_{G_3}(\frac{1}{2}+s) - R_{G_3}(p) \geq 2(1-4s)(p-\frac{1}{2}-s)^2 \geq \kappa(p-\frac{1}{2}-s)^2$ for $p \geq \frac{1}{2}$. For $p \leq \frac{1}{2}$, the same inequality holds as $R_{G_3}(\frac{1}{2}+s) - p - \kappa(\frac{1}{2}+s-p)^2$ is concave in p, and non-negative at $p \in \{0, \frac{1}{2}\}$ as $\kappa \leq 2(1-4s) \leq \frac{2}{(1+2s)^2}$. So $G_3 \in \mathcal{F}_{\mathrm{curv}(\kappa)}$, and so is U by Eq. (14).
- *Gaps.* Every $p \geq \frac{1}{2} + \frac{s}{2}$ has gap at least $\frac{s^2}{2}$ under U, and as R_{G_3} increases on $[0, \frac{1}{2}+s]$, every $p < \frac{1}{2} + \frac{s}{2}$ has gap at least $R_{G_3}(\frac{1}{2}+s) - R_{G_3}(\frac{1}{2}+\frac{s}{2}) \geq \frac{s^2}{2}$ under G_3.
- *Distance.* $H^2(U, G_3) = s(1-\sqrt{1-4s})^2 + (\frac{1}{2}-s)(\sqrt{1+\theta}-1)^2 \leq 16s^3 + 32s^4 \leq 24s^3$. □

B.5 Results on Interval-Query Complexities

Theorem B.21 (Interval-Query Complexities)**.** *For every $\kappa \in (0, 2)$, as $\epsilon \to 0$:*

(a) $\textsc{Interval}_{\mathcal{F}}^{\mathrm{CDF}}(\epsilon) = \widetilde{\Theta}(\epsilon^{-2})$ *for each* $\mathcal{F} \in \{\mathcal{F}_{\mathrm{all}}, \mathcal{F}_{\mathrm{regular}}, \mathcal{F}_{\mathrm{MHR}}, \mathcal{F}_{\mathrm{curv}(\kappa)}\}$*;*
(b) $\textsc{Interval}_{\mathrm{all}}(\epsilon) = \widetilde{\Theta}(\epsilon^{-2})$ *and* $\textsc{Interval}_{\mathrm{regular}}(\epsilon) = \widetilde{\Theta}(\epsilon^{-2})$*;*
(c) $\textsc{Interval}_{\mathrm{MHR}}(\epsilon) = \widetilde{\Theta}(\epsilon^{-1.5})$*;*
(d) $\textsc{Interval}_{\mathrm{curv}(\kappa)}(\epsilon) = \widetilde{\Theta}(\kappa^{-1/2}\epsilon^{-1.5})$.

That is, interval queries are as powerful as samples up to logarithmic factors (Theorem B.10).

Proof. By Theorem B.2(a), every interval-query complexity is at least the corresponding sample complexity, so the lower bounds follow from Theorem B.10. For the upper bounds, Lemmas B.22 and B.23 give (a) and (b) for $\mathcal{F}_{\mathrm{all}}$, and hence for its subclasses, and Lemmas B.24 and B.25 give (c) and (d). □

Lemma B.22 (CDF Estimation via Interval Queries)**.** *For every $\epsilon \in (0, 1]$, $\textsc{Interval}_{\mathrm{all}}^{\mathrm{CDF}}(\epsilon) = \mathcal{O}(\epsilon^{-2}\log^4\frac{2}{\epsilon})$.*

Proof. Run Algorithm 5 with $\delta := \epsilon^4$, and stop it after $160H^3\iota\epsilon^{-2} = \mathcal{O}(\epsilon^{-2}\log^4\frac{2}{\epsilon})$ queries (so that the number of queries won't exceed our budget), where H and ι are as in Lemma C.5. By Lemma C.5, with probability at least $1 - \mathcal{O}(\epsilon\log^3\frac{2}{\epsilon})$, it is not stopped, and it outputs an $\widehat{F}$ that is ϵ-close to F in Lévy distance. This probability is at least $\frac{3}{4}$ when ϵ is below an absolute constant, so the bound holds for all $\epsilon \in (0, 1]$, as $\textsc{Interval}_{\mathrm{all}}^{\mathrm{CDF}}(\epsilon)$ is non-increasing in ϵ. □

The same algorithm also estimates the monopoly price in expectation, as its failure probability can be made small:

Lemma B.23 (Monopoly Price of General Distributions via Interval Queries)**.** *For every* $\epsilon \in (0,1]$, $\textsc{Interval}_{\text{all}}(\epsilon) = \mathcal{O}(\epsilon^{-2}\log^4\frac{2}{\epsilon})$.

Proof. Run Algorithm 5 with accuracy $\frac{\epsilon}{6}$ and $\delta := \epsilon^5$, and stop it after $160H^3\iota(\frac{\epsilon}{6})^{-2} = \mathcal{O}(\epsilon^{-2}\log^4\frac{2}{\epsilon})$ queries, where H and ι are as in Lemma C.5 for accuracy $\frac{\epsilon}{6}$. Output the price $\widehat{p}$ given by Reduction B.4 for its output, or an arbitrary price if stopped. By Lemma C.5, with probability at least $1 - \mathcal{O}(\epsilon^2\log^3\frac{2}{\epsilon})$, it is not stopped and its output is $\frac{\epsilon}{6}$-close to F, in which case $R(p^*) - R(\widehat{p}) \le \frac{5\epsilon}{6}$. As every gap is at most 1, $\mathbb{E}[R(p^*) - R(\widehat{p})] \le \frac{5\epsilon}{6} + \mathcal{O}(\epsilon^2\log^3\frac{2}{\epsilon}) \le \epsilon$ when ϵ is below an absolute constant, and the bound holds for all $\epsilon \in (0,1]$ as $\textsc{Interval}_{\text{all}}(\epsilon)$ is non-increasing in ϵ. □

Lemma B.24 (Monopoly Price of MHR Distributions via Interval Queries)**.** $\textsc{Interval}_{\text{MHR}}(\epsilon) = \widetilde{\mathcal{O}}(\epsilon^{-1.5})$.

Proof. By Theorem 3.4(b), Algorithm 2 uses size-2 menus and has regret at most $cT^{1/3}\log^4 T$ under every $F \in \mathcal{F}_{\text{MHR}}$ for an absolute constant c. This is at most $\frac{\epsilon T}{2}$ for some $T = \widetilde{\mathcal{O}}(\epsilon^{-1.5})$ with $T \ge \log\frac{2}{\epsilon}$, so Reduction B.6 with $m = 2$ gives $\textsc{Interval}_{\text{MHR}}(\epsilon) \le 12T = \widetilde{\mathcal{O}}(\epsilon^{-1.5})$. □

Lemma B.25 (Monopoly Price of Curvature Distributions via Interval Queries)**.** *For every* $\kappa \in (0,2)$, $\textsc{Interval}_{\text{curv}(\kappa)}(\epsilon) = \widetilde{\mathcal{O}}(\kappa^{-1/2}\epsilon^{-1.5})$.

Proof. By Theorem 3.4(c), Algorithm 2 uses size-2 menus and has regret at most $c(\kappa^{-1/3}T^{1/3} + \log T)\log^4 T$ under every $F \in \mathcal{F}_{\text{curv}(\kappa)}$ for an absolute constant c. As $\kappa \le 2$, this is at most $\frac{\epsilon T}{2}$ for some $T = \widetilde{\mathcal{O}}(\kappa^{-1/2}\epsilon^{-1.5})$ with $T \ge \log\frac{2}{\epsilon}$, so Reduction B.6 with $m = 2$ gives $\textsc{Interval}_{\text{curv}(\kappa)}(\epsilon) \le 12T = \widetilde{\mathcal{O}}(\kappa^{-1/2}\epsilon^{-1.5})$. □

B.6 Results on Threshold-Query Complexities

For threshold queries, we prove extra lower bounds for the curvature distributions. All remaining upper and lower bounds in Table 2 are from [LSTW23a; LSTW23b; KSU19].

Theorem B.26 (Threshold-Query Lower Bounds for Curvature Distributions)**.** *For every* $\kappa \in (0,2)$, *as* $\epsilon \to 0$ *and* $T \to \infty$*:*

(a) $\textsc{Threshold}^{\text{CDF}}_{\text{curv}(\kappa)}(\epsilon) = \Omega(\epsilon^{-3})$*;*

(b) $\textsc{Threshold}_{\text{curv}(\kappa)}(\epsilon) = \Omega((2-\kappa)^{3/2}\epsilon^{-2.5})$*;*

(c) $\text{Reg}^{(1)}_{\text{curv}(\kappa)}(T) = \Omega((2-\kappa)^{3/5}T^{3/5})$*, i.e., posted prices suffer this regret.*

Compared with Theorem B.21, interval queries are thus polynomially more powerful than threshold queries for curvature distributions, for both properties.

Proof. (a) is Lemma B.27, (b) is Lemma B.28, and (c) is Lemma B.29. □

Lemma B.27 (CDF Threshold-Query Lower Bound for Curvature Distributions)**.** *For every* $\kappa \in (0,2]$ *and* $\epsilon \in (0, \frac{1}{480}]$, $\textsc{Threshold}^{\text{CDF}}_{\text{curv}(\kappa)}(\epsilon) \ge \frac{1}{6\cdot 10^6\epsilon^3}$.

Proof. Using needle-in-a-haystack, we hide a needle at one of $\Theta(\epsilon^{-1})$ locations of U, each of which a threshold query can only probe separately. Let $\eta := 8\epsilon$ and $J := \lfloor \frac{1}{30\eta} \rfloor \geq \frac{1}{60\eta}$. For $j \in [J]$, let $p_j := \frac{3}{5} + 6\eta j - 3\eta$, $I_j := (p_j - 3\eta, p_j + 3\eta)$, and $\phi_j(p) := \max\{0, 1 - \frac{|p-p_j|}{3\eta}\}$, so that $I_1, I_2, \ldots, I_J \subseteq (\frac{3}{5}, \frac{4}{5})$ are disjoint. Let F_j have $D_{F_j} := D_U - \eta\phi_j$, i.e., F_j moves mass η from the right half of I_j to its left half.

- *Curvature.* As $D_{F_j} \leq D_U$ with equality at $\frac{1}{2}$, we have $R_{F_j} \leq R_U$ with equality at $\frac{1}{2}$, so $F_j \in \mathcal{F}_{\mathrm{curv}(2)} \subseteq \mathcal{F}_{\mathrm{curv}(\kappa)}$ by Eq. (14).
- *Lévy separation.* An $\widehat{F}$ satisfying Eq. (13) for F_j has $D_{\widehat{F}}(p_j) \leq D_{F_j}(p_j - \epsilon) + \epsilon = 2 - 2p_j - \eta + \frac{10\epsilon}{3}$, whereas one for U has $D_{\widehat{F}}(p_j) \geq D_U(p_j + \epsilon) - \epsilon = 2 - 2p_j - 3\epsilon$, so no $\widehat{F}$ satisfies both.

Hence, if an algorithm outputs an $\widehat{F}$ satisfying Eq. (13) with probability at least $\frac{3}{4}$ under U and under every F_j, then its observations have total variation distance at least $\frac{3}{4} - \frac{1}{4} = \frac{1}{2}$ between U and every F_j. As $D_U, D_{F_j} \in [\frac{1}{5}, \frac{4}{5}]$ on I_j, Lemma B.9 with $\sigma = \frac{1}{5}$ gives $n \geq \frac{J}{180\eta^2} \geq \frac{1}{6 \cdot 10^6 \epsilon^3}$. □

Lemma B.28 (Threshold-Query Lower Bound for Curvature Distributions)**.** *For every $\kappa \in (0, 2)$ and all sufficiently small $\epsilon > 0$ (depending on κ),* $\textsc{Threshold}_{\mathrm{curv}(\kappa)}(\epsilon) \geq \frac{(2-\kappa)^{3/2}}{2 \cdot 10^7 \epsilon^{2.5}}$.

Proof. We hide a needle near the smooth peak of a uniform distribution, at one of $\Theta(\epsilon^{-1/2})$ locations within distance $\Theta(\sqrt{\epsilon})$ of the peak. Let $c_0 := \frac{6+\kappa}{4} \in (\kappa, 2)$, $w := \frac{2-c_0}{4} = \frac{2-\kappa}{16}$, and $\delta := \sqrt{(2-\kappa)/8}$. Let $F_0 := \mathrm{Unif}[1 - \frac{1}{c_0}, 1]$, so that $R_{F_0}(p) = c_0 p(1-p)$ for $p \geq 1 - \frac{1}{c_0}$ and $R_{F_0}(p) = p$ otherwise, and hence $\frac{c_0}{4} - R_{F_0}(p) \geq c_0(p - \frac{1}{2})^2$ for all $p \in [0, 1]$. Unlike U, whose revenue has a kink at its monopoly price, R_{F_0} has a smooth peak at $\frac{1}{2}$ with curvature $c_0 > \kappa$. Let $\eta := 8\epsilon$ and $J := \lfloor \frac{\delta}{3\sqrt{\eta}} \rfloor$, and for $j \in [J]$, let $p_j := \frac{1}{2} - \delta\sqrt{\eta} + 6\eta j - 3\eta$, $I_j := (p_j - 3\eta, p_j + 3\eta)$, and $\phi_j(p) := \max\{0, 1 - \frac{|p-p_j|}{3\eta}\}$. Let F_j have $R_{F_j} := R_{F_0} + \eta\phi_j$. For all sufficiently small ϵ, $D_{F_j} = \frac{R_{F_j}}{p}$ is decreasing, and:

- $F_j \in \mathcal{F}_{\mathrm{curv}(\kappa)}$ *with monopoly price* p_j. As $|p_j - \frac{1}{2}| \leq \delta\sqrt{\eta}$, every $p \notin I_j$ has gap at least $(1 - c_0\delta^2)\eta + c_0(p - \frac{1}{2})^2 \geq \kappa(p - p_j)^2$, where the last step uses $\kappa(p - p_j)^2 \leq c_0(p - \frac{1}{2})^2 + \frac{c_0\kappa}{c_0 - \kappa}\delta^2\eta$ and $\frac{c_0^2\delta^2}{c_0 - \kappa} = \frac{(6+\kappa)^2}{96} < 1$. Every $p \in I_j$ has gap $c_0(p - p_j)(p + p_j - 1) + \frac{|p-p_j|}{3} \geq \frac{|p-p_j|}{4} \geq \kappa(p - p_j)^2$.
- *Gaps.* Every price outside I_j has gap at least $(1 - c_0\delta^2)\eta \geq \frac{\eta}{2} = 4\epsilon$ under F_j, so by Markov's inequality, Eq. (12) under F_j gives $\widehat{p} \in I_j$ with probability at least $\frac{3}{4}$.
- *Needles.* On I_j, $|D_{F_j} - D_{F_0}| \leq \frac{\eta}{p} \leq 3\eta$, and $D_{F_0}, D_{F_j} \in [\frac{1}{2}, 1 - w]$, as $D_{F_0}(\frac{1}{2}) = \frac{c_0}{2} = 1 - 2w$.

As $\sum_j \Pr_{F_0}\{\widehat{p} \in I_j\} \leq 1$, at least $\frac{3J}{4}$ indices j have $\Pr_{F_0}\{\widehat{p} \in I_j\} \leq \frac{4}{J} \leq \frac{1}{3}$, and hence total variation distance at least $\frac{3}{4} - \frac{1}{3} > \frac{1}{3}$ between F_0 and F_j. So Lemma B.9 with $\sigma = w$ and 3η in place of η gives $n \geq \frac{wJ}{324\eta^2} \geq \frac{(2-\kappa)^{3/2}}{2 \cdot 10^7 \epsilon^{2.5}}$. □

By the online-to-batch conversion, this bound also forces posted prices to suffer more regret than menus (Theorem 4.1):

Lemma B.29 (Posted Prices for Curvature Distributions)**.** *For every $\kappa \in (0, 2)$ and all sufficiently large T (depending on κ),* $\mathrm{Reg}^{(1)}_{\mathrm{curv}(\kappa)}(T) = \Omega((2 - \kappa)^{3/5} T^{3/5})$.

Proof. If an algorithm using posted prices had regret at most $\gamma(2-\kappa)^{3/5}T^{3/5}$ under every $F \in \mathcal{F}_{\mathrm{curv}(\kappa)}$, then Reduction B.6 with $m = 1$ would give $\textsc{Threshold}_{\mathrm{curv}(\kappa)}(\gamma(2 - \kappa)^{3/5}T^{-2/5}) \leq T$, which contradicts Lemma B.28 for a small enough absolute constant $\gamma > 0$ and large enough T. □

C Omitted Proofs in Section 3 (Upper Bounds)

We first state the two tools shared by all our algorithms (Appendix C.1), and then prove Theorem 3.2 (Appendix C.2), Lemma 3.3 (Appendix C.4), ScaleSearch and the properties of MHR distributions (Appendix C.3), and Theorems 3.4 and 3.5 (Appendix C.5). Appendix C.6 lists examples of interval dimensions and characterizes when $V_T(F)$ tends to infinity. The resulting query complexities are collected in Appendix B.

C.1 Empirical Bernstein and Segment Trees

All our algorithms estimate the probability $D(I) = \Pr_{v\sim F}\{v \in I\}$ of interval I by Subroutine 1, and combine these estimates via segment trees. This subsection states the properties of these two tools.

Empirical Bernstein. Subroutine 1 enters our analysis through two facts: its confidence radius r_K is valid, and it shrinks as $\sqrt{D(I)/K}$ (even without knowing $D(I)$) rather than $\sqrt{1/K}$.

Lemma C.1 (Confidence Radius Shrinks with the Mean)**.** *Let $X_1, X_2, \ldots$ be i.i.d. Bernoulli random variables with mean μ, and for $K \geq 2$, let $\widehat{\mu}_K$ and r_K be their empirical mean and the confidence radius in Eq. (3) after the first K of them, where $\iota = \log\frac{2}{\delta}$ for some $\delta \in (0, \frac{1}{2}]$. For any $K_0 \geq 2$, with probability at least $1 - 3K_0\delta$, every $K \in \{2, 3, \ldots, K_0\}$ satisfies*

$$|\widehat{\mu}_K - \mu| \leq r_K \leq \sqrt{\frac{2\iota\mu}{K}} + \frac{15\iota}{2K}. \tag{15}$$

In particular, this applies to a call of Subroutine 1 on an interval I with $\mu = D(I)$, if its queries use fresh i.i.d. values $v \sim F$ given the history before the call.

Proof. We apply two bounds of Maurer and Pontil [MP09], and then bound the sample variance by the mean. By Eq. (3), $r_K = \sqrt{2\iota V_K/K} + \frac{7\iota}{3(K-1)}$, where $V_K := \frac{1}{K(K-1)}\sum_{k<k'\leq K}(X_k - X_{k'})^2$ is the sample variance. For each $K \in \{2, 3, \ldots, K_0\}$:

- $|\widehat{\mu}_K - \mu| \leq r_K$ with probability at least $1 - 2\delta$, by [MP09, Theorem 4] applied to both $\{X_k\}$ and $\{1 - X_k\}$, which have the same V_K;
- $\sqrt{V_K} \leq \sqrt{\mu(1-\mu)} + \sqrt{\frac{2\iota}{K-1}}$ with probability at least $1 - \delta$, by [MP09, Theorem 10].

By a union bound, both hold for all such K with probability at least $1 - 3K_0\delta$. Plugging the second one into r_K,

$$r_K \overset{(a)}{\leq} \sqrt{\frac{2\iota\mu}{K}} + \sqrt{\frac{2\iota}{K}\cdot\frac{4\iota}{K}} + \frac{14\iota}{3K} \overset{(b)}{\leq} \sqrt{\frac{2\iota\mu}{K}} + \frac{15\iota}{2K},$$

where (a) uses $\mu(1-\mu) \leq \mu$ and $K - 1 \geq \frac{K}{2}$, and (b) uses $\sqrt{8} + \frac{14}{3} \leq \frac{15}{2}$. This proves Eq. (15). Finally, a call of Subroutine 1 on I observes a prefix of such a sequence with $\mu = D(I)$, as each query uses a fresh value. □

Both terms in Eq. (15) are at most $\frac{\epsilon}{2}$ once $K \geq 8\iota\mu\epsilon^{-2} + 15\iota\epsilon^{-1}$. Thus, intervals with smaller probabilities need fewer queries:

Lemma C.2 (Fewer Queries for Smaller Probability Intervals)**.** *Consider a call of Subroutine 1 on an interval I with precision $\epsilon \in (0, 1]$ and failure probability $\delta \in (0, \frac{1}{2}]$, whose queries use fresh i.i.d. values $v \sim F$ given the history before the call, and let $\iota := \log\frac{2}{\delta}$. With probability at least $1 - 78\iota\epsilon^{-2}\delta$, it stops within $8\iota D(I)\epsilon^{-2} + 15\iota\epsilon^{-1} + 2$ queries and outputs an estimate $\widehat{D}(I)$ with $|\widehat{D}(I) - D(I)| \leq \epsilon$.*

Subroutine 4 BUILDSEGTREE(endpoints $p_0 < p_1 < \cdots < p_N$)

Initialize the tree $\mathcal{T}$ as a single root interval, namely $\mathcal{T} := \{[p_0, p_N)\}$
for tree interval $I = [p_i, p_j) \in \mathcal{T}$ **do**
  if $j - i \geq 2$ **then** ▷ *Otherwise, $j = i + 1$ and then $[p_i, p_j)$ is a leaf*
    Add its two children $[p_i, p_k)$ and $[p_k, p_j)$ to $\mathcal{T}$, where $k = \lfloor \frac{i+j}{2} \rfloor$
Initialize the tree decomposition for the empty interval as $\mathcal{I}_{\mathcal{T}}(\varnothing) = \varnothing$
for all $0 \leq i < j \leq N$ in the increasing order of $(j - i)$ **do**
  if $[p_i, p_j)$ belongs to $\mathcal{T}$, i.e., it is a tree interval **then**
    The tree decomposition $\mathcal{I}_{\mathcal{T}}[p_i, p_j) := \{[p_i, p_j)\}$ is itself
  else
    Let $I \in \mathcal{T}$ be the smallest tree interval containing $[p_i, p_j)$. Let its two children be I_L and I_R, both belonging to $\mathcal{T}$. Then $\mathcal{I}_{\mathcal{T}}[p_i, p_j) := \mathcal{I}_{\mathcal{T}}([p_i, p_j) \cap I_L) \cup \mathcal{I}_{\mathcal{T}}([p_i, p_j) \cap I_R)$
return segment tree $\mathcal{T}$ and the associated tree decomposition $\mathcal{I}_{\mathcal{T}}$

Proof. Apply Lemma C.1 with $K_0 := \lceil 25\iota\epsilon^{-2} \rceil$, which is at most $26\iota\epsilon^{-2}$ as $\iota\epsilon^{-2} \geq 1$, so it fails with probability at most $3K_0\delta \leq 78\iota\epsilon^{-2}\delta$. On its event, $r_K \leq \epsilon$ for every $K \geq K^* := \max\{2, \lceil 8\iota D(I)\epsilon^{-2} + 15\iota\epsilon^{-1} \rceil\}$, where $K^* \leq 8\iota D(I)\epsilon^{-2} + 15\iota\epsilon^{-1} + 2 \leq 25\iota\epsilon^{-2} \leq K_0$ as $D(I), \epsilon \leq 1 \leq \iota$. So the call stops at some $K \leq K^*$, and its output $\widehat{\mu}_K$ satisfies $|\widehat{\mu}_K - D(I)| \leq r_K \leq \epsilon$. □

Segment Trees. Given endpoints $p_0 < p_1 < \cdots < p_N$, Subroutine 4 builds a segment tree $\mathcal{T}$ by recursively halving $[p_0, p_N)$ at its middle endpoint, and decomposes every $[p_i, p_j)$ into tree intervals $\mathcal{I}_{\mathcal{T}}[p_i, p_j)$. We use the properties in Lemma C.3.

Lemma C.3 (Segment Tree Decomposition)**.** *The segment tree $\mathcal{T}$ built by Subroutine 4 on $p_0 < p_1 < \cdots < p_N$ has $2N - 1$ tree intervals in at most $\lceil \log_2 N \rceil + 1$ depths, and tree intervals of the same depth are disjoint. Moreover, for any $0 \leq i < j \leq N$, $\mathcal{I}_{\mathcal{T}}[p_i, p_j)$ partitions $[p_i, p_j)$ into tree intervals, at most two of each depth.*

Proof. Each non-leaf tree interval is the disjoint union of its two children. Hence any two tree intervals are either nested or disjoint, tree intervals of the same depth are disjoint, and as the leaves are the N intervals $[p_{i-1}, p_i)$, there are $2N - 1$ tree intervals. As each child of $[p_i, p_j)$ contains at most $\lceil \frac{j-i}{2} \rceil$ leaves, a tree interval of depth h contains at most $\lceil N/2^h \rceil$ leaves. A tree interval with children contains at least 2 leaves, so its depth is less than $\log_2 N$, and hence all depths are at most $\lceil \log_2 N \rceil$.

For decomposition, if $[p_i, p_j)$ is not a tree interval, let I be the smallest tree interval containing it, with children I_L and I_R. By the minimality of I, neither child contains $[p_i, p_j)$, so $[p_i, p_j) \cap I_L$ and $[p_i, p_j) \cap I_R$ are non-empty and partition $[p_i, p_j)$. By induction on $j - i$, $\mathcal{I}_{\mathcal{T}}[p_i, p_j)$ thus partitions $[p_i, p_j)$ into tree intervals. To count them, note that $[p_i, p_j) \cap I_L$ shares the right endpoint of I_L, and $[p_i, p_j) \cap I_R$ shares the left endpoint of I_R. Hence it suffices to show the following claim, as each of the two parts then contributes at most one tree interval of each depth:

If $S \subseteq J$ shares the right endpoint of a tree interval J, then $\mathcal{I}_{\mathcal{T}}(S)$ contains at most one tree interval of each depth, all of which are deeper than J unless $S = J$. The same holds if S shares the left endpoint of J.

We prove the claim by induction on the number of leaves in S (the case of left endpoints is symmetric). If S is a tree interval, then $\mathcal{I}_{\mathcal{T}}(S) = \{S\}$, and S is deeper than J unless $S = J$. Otherwise, let I' be the smallest tree interval containing S, so that $S \subsetneq I' \subseteq J$ and S shares the right endpoint of I',

Algorithm 5 CDF Estimation via Interval Queries

Input: accuracy $\epsilon \in (0,1]$, failure probability $\delta \in (0, \frac{1}{2}]$

1: Let $N = 2^{\lceil \log_2 \epsilon^{-1} \rceil}$. Let $(\mathcal{T}, \mathcal{I}_\mathcal{T}) \leftarrow \textsc{BuildSegTree}(0, \frac{1}{N}, \frac{2}{N}, \dots, 1)$ ▷ *see Subroutine 4*

2: **for** each interval I in the segment tree $\mathcal{T}$ **do**

3: Estimate $D(I)$ by $\widehat{D}(I) \leftarrow \textsc{EmpiricalBernstein}(I, \frac{\epsilon}{2\log_2(2N)}, \delta)$ ▷ *see Subroutine 1*

4: For $i = 0, 1, \dots, N$, let $\widehat{d}_i := 1 - \sum_{I \in \mathcal{I}_\mathcal{T}[0, \frac{i}{N})} \widehat{D}(I)$, which estimates $D(\frac{i}{N})$

5: **return** the value distribution $\widehat{F}$ with

$$D_{\widehat{F}}(x) := \max\left\{0, \min_{0 \le i \le \lceil Nx \rceil} \widehat{d}_i\right\}, \quad \forall x \in [0,1]. \tag{16}$$

and let I'_L and I'_R be the children of I'. As $I'_R \not\supseteq S$ and both end at the same point, we have $S \supsetneq I'_R$, so $\mathcal{I}_\mathcal{T}(S) = \mathcal{I}_\mathcal{T}(S \cap I'_L) \cup \{I'_R\}$, where $S \cap I'_L \neq I'_L$ shares the right endpoint of I'_L. By induction, $\mathcal{I}_\mathcal{T}(S \cap I'_L)$ contains at most one tree interval of each depth, all deeper than I'_L and hence than its sibling I'_R. So $\mathcal{I}_\mathcal{T}(S)$ contains at most one tree interval of each depth, all at least as deep as I'_R, which is deeper than J as $I' \subseteq J$. This proves the claim and hence the lemma. □

C.2 Proof of Theorem 3.2 (Revenue Maximization with Size-3 Menus)

Algorithm 5 is the CDF estimation algorithm sketched in Section 3.2. With endpoints $0, \frac{1}{N}, \frac{2}{N}, \dots, 1$, Subroutine 4 builds the segment tree $\mathcal{T}$ described there, whose tree intervals of depth h are $[\frac{k}{2^h}, \frac{k+1}{2^h})$ (for $k = 0, 1, \dots, 2^h - 1$), and $\mathcal{I}_\mathcal{T}[0, \frac{i}{N})$ decomposes $[0, \frac{i}{N})$ into $\mathcal{O}(\log_2 N)$ tree intervals (Lemma C.3). As the estimates $\widehat{d}_i$ may not be monotone in i (and thus fail to induce a CDF), Eq. (16) takes their running minimum and truncates it at 0, so that $D_{\widehat{F}}$ is non-increasing and left-continuous with $D_{\widehat{F}}(0) = \widehat{d}_0 = 1$.

Theorem C.4 (Restatement of Theorem 3.2; CDF Estimation via Interval Queries and Size-3 Menus)**.** $\textsc{Interval}_{\text{all}}^{\text{CDF}}(\epsilon) = \mathcal{O}(\epsilon^{-2} \log^4 \frac{2}{\epsilon})$ *and* $\text{Reg}_{\text{all}}^{(3)}(T) = \mathcal{O}(\sqrt{T} \log^3 T)$.

Proof. The first claim is Lemma B.22, which runs Algorithm 5 with $\delta := \epsilon^4$ and applies Lemma C.5 below. The second claim is Lemma C.6 below, which runs Algorithm 5 in phases using size-3 menus. □

Lemma C.5 (CDF Estimation via a Segment Tree)**.** *For any $F \in \mathcal{F}_{\text{all}}$, $\epsilon \in (0,1]$, and $\delta \in (0, \frac{1}{2}]$, let $H := \lceil \log_2 \epsilon^{-1} \rceil + 1$ and $\iota := \log \frac{2}{\delta}$. With probability at least $1 - \mathcal{O}(H^2 \iota \epsilon^{-3} \delta)$, Algorithm 5 uses at most $160 H^3 \iota \epsilon^{-2}$ interval queries, and its output $\widehat{F}$ in Eq. (16) is ϵ-close to F in Lévy distance.*

Proof. We show that all tree intervals are estimated accurately, then count the queries, and finally bound the Lévy distance. As $N = 2^{H-1}$, Algorithm 5 calls Subroutine 1 on each of the $2N - 1$ tree intervals with precision $\eta := \frac{\epsilon}{2H}$.

Accurate Estimates. Each query uses a fresh value $v \sim F$, so Lemma C.2 applies to each call given the history before it. As $2N - 1 < 4\epsilon^{-1}$, a union bound shows that, with probability at least $1 - 4\epsilon^{-1} \cdot 78 \iota \eta^{-2} \delta = 1 - \mathcal{O}(H^2 \iota \epsilon^{-3} \delta)$, every tree interval I satisfies $|\widehat{D}(I) - D(I)| \le \eta$ and takes at most $8\iota D(I) \eta^{-2} + 15 \iota \eta^{-1} + 2$ queries. We assume this event from now on.

Number of Queries. By Lemma C.3, the tree intervals of each of the H depths are disjoint, so their probabilities sum to at most 1, and there are at most $2N \le 4\epsilon^{-1}$ tree intervals. Hence the number

Algorithm 6 Revenue Maximization using Size-3 Menus

Input: Time horizon $T \geq 4$

Let $\gamma := T^{-2}$ and $\widehat{p}_0 := 0$

for phase $j = 1, 2, \ldots$ until the T rounds end **do**

  Run Algorithm 5 with $\epsilon_j = 2^{-j}$ and $\delta = T^{-5}$, answering each of its interval queries $[a, b)$ by the size-3 menu built from the price distribution $\gamma\delta_a + \gamma\delta_b + (1 - 2\gamma)\delta_{\widehat{p}_{j-1}}$ via Lemma 3.1

  Let $\widehat{p}_j$ be the price given by Reduction B.4 for the output of Algorithm 5 in this phase

of queries is at most

$$8H\iota\eta^{-2} + 4\epsilon^{-1}\left(15\iota\eta^{-1} + 2\right) = 32H^3\iota\epsilon^{-2} + 120H\iota\epsilon^{-2} + 8\epsilon^{-1} \leq 160H^3\iota\epsilon^{-2}.$$

Lévy Distance. For each $i \in [N]$, Lemma C.3 shows that $\mathcal{I}_{\mathcal{T}}[0, \frac{i}{N})$ partitions $[0, \frac{i}{N})$ into tree intervals, at most two of each depth, so $\widehat{d}_i$ in Eq. (16) satisfies

$$\left|\widehat{d}_i - D_F\left(\tfrac{i}{N}\right)\right| \leq \sum_{I \in \mathcal{I}_{\mathcal{T}}[0, \frac{i}{N})} \left|\widehat{D}(I) - D(I)\right| \leq 2H\eta = \epsilon.$$

This also holds for $i = 0$, as $\widehat{d}_0 = 1 = D_F(0)$. As D_F is non-increasing, the minimum over $j \leq i$ and the truncation at 0 in Eq. (16) keep these bounds, so $|D_{\widehat{F}}(\frac{i}{N}) - D_F(\frac{i}{N})| \leq \epsilon$. For any $x \in (\frac{i-1}{N}, \frac{i}{N}]$, we have $D_{\widehat{F}}(x) = D_{\widehat{F}}(\frac{i}{N})$ and $D_F(x + \epsilon) \leq D_F(\frac{i}{N}) \leq D_F(x - \epsilon)$ as $\frac{1}{N} \leq \epsilon$; for $x = 0$, we have $D_{\widehat{F}}(0) = 1$. Therefore, $D_F(x + \epsilon) - \epsilon \leq D_{\widehat{F}}(x) \leq D_F(x - \epsilon) + \epsilon$ for all $x \in [0, 1]$, i.e., $\widehat{F}$ is ϵ-close to F in Lévy distance. □

The second claim of Theorem 3.2 runs Algorithm 5 in phases of halving accuracies, exploiting in each phase the price estimated in the previous one (Algorithm 6):

Lemma C.6 (Size-3 Menus via Phased CDF Estimation)**.** *Let $T \geq 4$. The regret of Algorithm 6 is $\mathcal{O}(\sqrt{T}\log^3 T)$ under every $F \in \mathcal{F}_{\text{all}}$.*

Proof. By Lemma 3.1, each menu reveals whether $v_t \in [a, b)$, so every phase runs Algorithm 5 faithfully, and it earns the expected revenue of a random price drawn from the price distribution, which is at least $(1 - 2\gamma)R(\widehat{p}_{j-1})$ in phase j. We first show that every phase succeeds with high probability, and then bound the regret of each phase.

Number of Phases. Phase i estimates $2^{i+1} - 1$ tree intervals, each with at least 2 queries, so it takes at least 2^{i+1} rounds. Hence, if phase j is reached, the earlier phases took at least $\sum_{i<j} 2^{i+1} = 2^{j+1} - 4$ rounds, so $2^{j+1} - 4 \leq T$. As $T \geq 4$, at most $\log_2 T$ phases are reached, and all of them have $\epsilon_j \geq \frac{1}{T}$.

Good Event. Each phase uses fresh values, so Lemma C.5 applies to each phase given the history before it, with $H_j := j + 1 \leq \log_2 T + 1$ and $\iota := \log(2T^5)$. By a union bound over the at most $\log_2 T$ phases, with probability at least $1 - \mathcal{O}(\log^4 T/T^2)$, every completed phase j outputs an $\widehat{F}$ (Eq. (16)) that is ϵ_j-close to F, and every phase j lasts at most $160H_j^3\iota\epsilon_j^{-2}$ rounds. Its complement contributes at most $T \cdot \mathcal{O}(\log^4 T/T^2) = \mathcal{O}(1)$ to the regret.

Regret of Each Phase. On this event, Reduction B.4 gives $R(p^*) - R(\widehat{p}_j) \leq 5\epsilon_j$ for every completed phase j, which also holds for $j = 0$ as $\epsilon_0 = 1$. So each round of phase j has regret at most

$5\epsilon_{j-1} + 2\gamma = 10\epsilon_j + 2T^{-2}$, and phase j has regret at most

$$10\epsilon_j \min\left\{T, 160H_j^3\iota\epsilon_j^{-2}\right\} + 2 \le 10\sqrt{160H_j^3\iota T} + 2,$$

where we used $\min\{a,b\} \le \sqrt{ab}$. Summing over the at most $\log_2 T$ phases, the regret is $\mathcal{O}(\sqrt{T}\log^3 T)$. □

C.3 ScaleSearch for MHR Distributions (Line 1 in Algorithm 2)

As explained in Footnote 9, running the layers of Algorithm 2 directly on F is wasteful whenever $\Delta \ge R_F(p_F^*)$, at which every price is Δ-optimal. We resolve this by rescaling the prices: Algorithm 2 first calls ScaleSearch (Line 1) to find a scale $\overline{p} \ge p_F^*$, and then runs on the rescaled distribution

$$G := \text{the distribution of } \min\left\{\frac{v}{\overline{p}}, 1\right\} \text{ for } v \sim F, \quad \text{i.e., } D_G(p) = D_F(\overline{p}p) \text{ for all } p \in [0,1]. \tag{17}$$

Whenever Algorithm 2 posts p to G, we instead post $\overline{p}p$ to F. As $R_G(p) = \frac{R_F(\overline{p}p)}{\overline{p}}$, the monopoly price $p_G^* = \frac{p_F^*}{\overline{p}}$ of G lies in $[0,1]$, and every regret under G is multiplied by $\overline{p} \le 1$ under F. In this way, instead of suffering the $R_F(p_F^*)^{-1}$ dependency in Footnote 9, we only suffer $R_G(p_G^*)^{-1}\overline{p} = \frac{\overline{p}^2}{R_F(p_F^*)}$, as $R_G(p_G^*) = \frac{R_F(p_F^*)}{\overline{p}}$. In this section, we prove that for MHR distributions, ScaleSearch ensures $\overline{p} < 10p_F^*$ (Lemma C.7), and $R_F(p_F^*) \ge \frac{p_F^*}{e}$ (Lemma C.8), so $R_G(p_G^*)^{-1} = \frac{\overline{p}}{R_F(p_F^*)} \le 10e$.

ScaleSearch searches the geometric grid $x_j = 2^{-j}$, which must contain a price $x_{j^*} \in (\frac{p_F^*}{2}, p_F^*]$ earning at least $\frac{R_F(p_F^*)}{2}$ (recall the one-sided Lipschitzness Eq. (1)). It must never rule out x_{j^*}, but must also rule out the prices far above p_F^*, so it combines pessimism and optimism:

- *Pessimism.* The largest lower confidence bound among this grid, namely L in Subroutine 7, implies that $R_F(p_F^*) \ge L$ with high probability.
- *Optimism.* A grid price is kept if its upper confidence bound is at least $\frac{L}{2}$. Because x_{j^*} earns at least $\frac{R_F(p_F^*)}{2} \ge \frac{L}{2}$, it is always kept, so twice the largest kept price is at least p_F^*.

ScaleSearch thus runs empirical Bernstein on all these grid prices simultaneously (i.e., in each round, ask all prices once). It waits until all confidence radii drop below $\frac{L}{16}$, so every kept price indeed earns a constant fraction of $R_F(p_F^*)$. For MHR distributions, the revenue decays exponentially above p_F^*, so all kept prices are $\mathcal{O}(p_F^*)$. This also does not bring a significant regret overhead: Posting x_j gives a revenue sample with variance $x_j R_F(x_j) \le R_F(p_F^*)$, so $\widetilde{\mathcal{O}}(R_F(p_F^*)^{-1})$ rounds per grid price suffice, each with regret at most $R_F(p_F^*)$, and the total regret is $\mathcal{O}(\log^2 T)$.

This intuition is summarized in Lemma C.7 below. Further utilizing the tail and shape of MHR distributions in Lemmas C.8 and C.9 respectively, we then derive the MHR bound in Theorem 3.4(b).

Lemma C.7 (Scale Search). *Let $J := \lceil \log_2 T \rceil$. With probability at least $1 - 3(J+1)T^{-2}$:*

(a) *Subroutine 7 has regret $\mathcal{O}(\log^2 T)$.*
(b) *If $p_F^* \ge 2^{-J}$, its output $\overline{p}$ satisfies $p_F^* \le \overline{p} \le 1$.*
(c) *Additionally, if F is MHR, then the output $\overline{p} < 10p_F^*$.*

Proof. ScaleSearch posts $1, \frac{1}{2}, \ldots, 2^{-J}$ in turn until all revenue estimates are accurate compared with the best lower confidence bound L, and returns twice the largest price that may earn at least $\frac{L}{2}$.

Subroutine 7 ScaleSearch(time horizon T)

Let $J := \lceil \log_2 T \rceil$ and $x_j := 2^{-j}$ for $0 \le j \le J$
for $K = 1, 2, \ldots$ **do**
  Post each of $x_0, x_1, \ldots, x_J$ once using the size-1 menu $\{(1, x_j)\}$
  For each j, let $\widehat{R}(x_j)$ be the average of $x_j \mathbb{1}[v \ge x_j]$ over the K rounds posting x_j
  For each j, let $r_{j,K}$ be the confidence radius r_K in Eq. (3) for the outcomes $\mathbb{1}[v \ge x_j]$ of these K rounds, with $\delta = T^{-3}$
  Let $L := \max_j(\widehat{R}(x_j) - x_j r_{j,K})$. If $K \ge 2$ and $\max_j x_j r_{j,K} \le \frac{L}{16}$, **return** $\overline{p} := \min\{1, 2\max\{x_j : \widehat{R}(x_j) + x_j r_{j,K} \ge \frac{L}{2}\}\}$

Let $\iota := \log(2T^3)$. The outcomes $\mathbb{1}[v \ge x_j]$ are i.i.d. Bernoulli with mean $\frac{R_F(x_j)}{x_j}$. So by Lemma C.1 with $K_0 := T$ for each j and a union bound, with probability at least $1 - 3(J+1)T^{-2}$, every j and every $K \ge 2$ satisfy

$$|\widehat{R}(x_j) - R_F(x_j)| \le x_j r_{j,K} \le \sqrt{\frac{2\iota x_j R_F(x_j)}{K}} + \frac{15\iota x_j}{2K}.$$

We assume this good event from now on. If $p_F^* < 2^{-J}$, then $R_F(p_F^*) < \frac{1}{T}$ and the regret is less than 1, so assume $p_F^* \ge 2^{-J}$, and let $x_{j^*} \in (\frac{p_F^*}{2}, p_F^*]$, so that $R_F(x_{j^*}) \ge x_{j^*} D_F(p_F^*) \ge \frac{R_F(p_F^*)}{2}$. We always have $L \le \max_j R_F(x_j) \le R_F(p_F^*)$.

- *Regret.* As $x_j R_F(x_j) \le R_F(p_F^*)$ and $x_j \le 1$, once every x_j has been posted $K \ge 10908\iota R_F(p_F^*)^{-1}$ times, every $x_j r_{j,K}$ is at most $\frac{R_F(p_F^*)}{36}$. Then $L \ge R_F(x_{j^*}) - 2x_{j^*} r_{j^*,K} \ge \frac{4}{9} R_F(p_F^*) \ge 16 \max_j x_j r_{j,K}$, so ScaleSearch stops. Hence it takes $\mathcal{O}((J+1)\iota R_F(p_F^*)^{-1})$ rounds, each with regret at most $R_F(p_F^*)$, so its regret is $\mathcal{O}(\log^2 T)$.
- *Output.* When ScaleSearch stops, $L \ge R_F(x_{j^*}) - 2x_{j^*} r_{j^*,K} \ge \frac{R_F(p_F^*)}{2} - \frac{L}{8}$, so $L \ge \frac{4}{9} R_F(p_F^*)$. As $\widehat{R}(x_{j^*}) + x_{j^*} r_{j^*,K} \ge R_F(x_{j^*}) \ge \frac{R_F(p_F^*)}{2} \ge \frac{L}{2}$, we have $\overline{p} \ge \min\{1, 2x_{j^*}\} \ge p_F^*$. Moreover, every j with $\widehat{R}(x_j) + x_j r_{j,K} \ge \frac{L}{2}$ has $R_F(x_j) \ge \frac{L}{2} - 2x_j r_{j,K} \ge \frac{3L}{8} \ge \frac{R_F(p_F^*)}{6}$. If F is MHR, Lemma C.8 gives $R_F(p_F^* + u) \le R_F(p_F^*)(1+x)e^{-x}$ with $x := \frac{u}{p_F^*}$, which is less than $\frac{R_F(p_F^*)}{6}$ if $x \ge 4$, as $5e^{-4} < \frac{1}{6}$. So all such x_j are less than $5p_F^*$, and $\overline{p} < 10p_F^*$. □

Lemma C.8 (MHR Tails)**.** *Let F be MHR, and let $q^* := D_F(p_F^*)$. Then $q^* \ge e^{-1}$, so $R_F(p_F^*) \ge \frac{p_F^*}{e}$, and every $u \ge 0$ satisfies $R_F(p_F^* + u) \le R_F(p_F^*)(1+x)e^{-x}$ with $x := \frac{u}{p_F^*}$.*

Proof. Let $h := \frac{f}{D_F}$ be the hazard rate, so that $R_F'(p) = D_F(p)(1 - ph(p))$ for almost every p. As h is non-decreasing and p_F^* maximizes R_F, we have $h(p) \le \frac{1}{p_F^*}$ for $p < p_F^*$ and $h(p) \ge \frac{1}{p_F^*}$ for $p > p_F^*$. Hence $q^* = \exp(-\int_0^{p_F^*} h) \ge e^{-1}$, and $D_F(p_F^* + u) = q^* \exp(-\int_{p_F^*}^{p_F^*+u} h) \le q^* e^{-x}$, so $R_F(p_F^* + u) \le (p_F^* + u) q^* e^{-x} = R_F(p_F^*)(1+x)e^{-x}$. □

Lemma C.9 (MHR Shape)**.** *Let F be MHR and $\Delta \le \frac{R_F(p_F^*)}{4}$. Then every Δ-optimal price p satisfies $|p - p_F^*| \le 4\sqrt{\Delta}$.*

Proof. We lower bound the gap $R_F(p_F^*) - R_F(p)$ on both sides of p_F^*, with h and q^* as in Lemma C.8.

- *Left side.* For $s \in [0, p_F^*]$, as $h \le \frac{1}{p_F^*}$ below p_F^*, $R_F'(p_F^* - s) \ge q^*(1 - \frac{p_F^* - s}{p_F^*}) = \frac{q^* s}{p_F^*}$. Integrating, $R_F(p_F^*) - R_F(p_F^* - u) \ge \frac{q^* u^2}{2p_F^*} \ge \frac{u^2}{2e}$, which exceeds Δ if $u > \sqrt{2e\Delta}$.
- *Right side.* By Lemma C.8, $R_F(p_F^*) - R_F(p_F^* + u) \ge R_F(p_F^*)(1 - (1+x)e^{-x})$ with $x := \frac{u}{p_F^*}$. Since $1 - (1+x)e^{-x} = \int_0^x y e^{-y}\, dy$, it is at least $\frac{x^2}{2e}$ for $x \le 1$ and at least $1 - \frac{2}{e} > \frac{1}{4}$ for $x \ge 1$. So $p_F^* + u$ is not Δ-optimal if $x \ge 1$; otherwise, its gap is at least $\frac{R_F(p_F^*)u^2}{2e(p_F^*)^2} = \frac{q^* u^2}{2ep_F^*} \ge \frac{u^2}{2e^2}$, which exceeds Δ if $u > e\sqrt{2\Delta}$.

As $\sqrt{2e} < e\sqrt{2} < 4$, every Δ-optimal price is within $4\sqrt{\Delta}$ of p_F^*. □

C.4 Proof of Lemma 3.3 (Number of Rounds)

Throughout, we analyze Algorithm 2 after Line 1, i.e., on the rescaled distribution G in Eq. (17), so that R_G, p_G^*, and D_G below are the revenue, monopoly price, and demand of G. The candidates $\mathcal{S}_\Delta$ of layer Δ are initialized in Algorithm 2 and then shrink in its Eliminate step, and layer $\frac{\Delta}{2}$ starts from the survivors. We use the one-sided Lipschitzness Eq. (1) in the following form: For any $p < p'$,

$$pD_G([p, p')) = R_G(p) - R_G(p') + (p' - p)D_G(p') \le R_G(p) - R_G(p') + (p' - p). \tag{18}$$

Lemma C.10 (Restatement of Lemma 3.3; Number of Rounds)**.** (a) *For each $\Delta = 2^{-k}$, Algorithm 2 spends $\widetilde{\mathcal{O}}((p_N - p_0 + \Delta)/\Delta^2)$ rounds where p_0, p_N are both 4Δ-optimal prices.*
(b) *For each $\Delta = 2^{-k}$, Algorithm 2 incurs $\mathcal{O}(\Delta)$ regret in each of these rounds.*
(c) *Consequently, for any given T, the regret of Algorithm 2 is upper bounded as*

$$\widetilde{\mathcal{O}}\left(\sup_{\Delta = 2^{-k}, k \ge 0} \min\left\{T\Delta, \frac{\operatorname{diam}\{p \in [0,1] \mid R(p) \ge R(p^*) - \Delta\}}{\Delta}\right\}\right),$$

where $\operatorname{diam}\{A\} := \sup\{A\} - \inf\{A\}$ *is the diameter of a set A.*

Proof. By Lemma C.11, all estimates are accurate except with probability $3T^{-1}$, which contributes $\mathcal{O}(1)$ regret, so we assume this event $\mathcal{E}$. Fix a layer Δ with candidates $\mathcal{S}_{2\Delta} = \{p_0 < p_1 < \cdots < p_N\}$, which are all 4Δ-optimal by Lemma C.13 below for the previous layer (or trivially for $\Delta = \frac{1}{2}$).

Claim (b): Regret per Round. Every menu in layer Δ puts probability at least $1 - T^{-2}$ on a price in $\mathcal{S}_{2\Delta}$, and Line 8 posts $p_N \in \mathcal{S}_{2\Delta}$. As these prices are 4Δ-optimal, every round has regret $4\Delta + T^{-2}$.

Claim (a): Number of Rounds. By Lemma C.11, a call estimating $D_G(I)$ with precision ϵ takes at most $8\iota D_G(I)\epsilon^{-2} + 15\iota\epsilon^{-1} + 2$ rounds, where $\iota := \log(2T^3)$. We bound the first term by telescoping Eq. (18) over disjoint intervals, and the other two by the number of calls. If $N = 0$, layer Δ only makes the calls in Lines 8 and 12, which take $\mathcal{O}(\iota\Delta^{-1})$ rounds by the last two bullets below, so we assume $N \ge 1$ and let $c_N := 80(\log_2 N + 2)$.

- *One interval.* Every interval in Lines 7, 11, and 12 is some $I = [a, b)$ with $a \in \mathcal{S}_{2\Delta}$, estimated with precision $\epsilon = \frac{\Delta}{c(a+\Delta)}$, where $c = c_N$ for tree intervals and $c = 40$ otherwise. As $(a+\Delta)^2 \le 2(a+\Delta^2)$ and $a + \Delta \le 2$, this call takes at most $16\iota c^2\Delta^{-2}(aD_G(I) + \Delta^2 D_G(I)) + 30\iota c\Delta^{-1} + 2$ rounds.
- *Telescoping across depth.* The tree intervals of each of the at most $\log_2 N + 2$ depths are disjoint subintervals of $[p_0, p_N)$ (Lemma C.3), and so are the intervals $[p_i, p_i + \frac{\Delta}{2})$ with $i < N$, as the candidates are Δ apart. For each such family of disjoint intervals $[a, b)$, summing Eq. (18) over a

partition of $[p_0, p_N)$ that contains them, the revenues telescope; as p_0 and p_N are 4Δ-optimal and the probabilities sum to at most 1,

$$\sum_{[a,b)} a\, D_G([a,b)) \le p_N - p_0 + R_G(p_0) - R_G(p_N) \le p_N - p_0 + 4\Delta \quad \text{and} \quad \sum_{[a,b)} \Delta^2 D_G([a,b)) \le \Delta.$$

As $\mathcal{T}$ has $2N-1$ tree intervals (Lemma C.3), the tree intervals take at most $16\iota c_N^2 \Delta^{-2}(\log_2 N + 2)(p_N - p_0 + 5\Delta) + (2N-1)(30\iota c_N \Delta^{-1} + 2)$ rounds.

- *All $[p_i, p_i + \frac{\Delta}{2})$ except for $i = N$.* The calculation in the two bullets above carries over, giving that all $[p_i, p_i + \frac{\Delta}{2})$ with $i < N$ take at most $16 \cdot 40^2 \iota \Delta^{-2}(p_N - p_0 + 5\Delta) + N(1200\iota\Delta^{-1} + 2)$ rounds.
- *The last $I = [p_N, p_N + \frac{\Delta}{2})$*, which is not inside $[p_0, p_N)$. If $D_G(I) \ge \frac{4\Delta}{p_N}$, the lower confidence bound $\widehat{\mu}_K - r_K \ge D_G(I) - 2r_K$ exceeds $\frac{2\Delta}{p_N}$ once $r_K \le \frac{D_G(I)}{5}$, so by Eq. (15), Line 12 stops this call within $\mathcal{O}(\iota D_G(I)^{-1}) = \mathcal{O}(\iota\Delta^{-1})$ rounds. Otherwise, $p_N D_G(I) + \Delta^2 D_G(I) \le 5\Delta$, so it takes $\mathcal{O}(\iota\Delta^{-1})$ rounds by the first bullet.
- *Rounds when posting p_N.* Line 8 estimates $D_G(p_N) \le 1$ up to $\frac{\Delta}{8(p_N - p_0 + \Delta)}$, which takes $\mathcal{O}(\iota(p_N - p_0 + \Delta)^2\Delta^{-2}) = \mathcal{O}(\iota(p_N - p_0 + \Delta)\Delta^{-2})$ rounds, as $p_N - p_0 + \Delta \le 2$.

As the candidates in $\mathcal{S}_{2\Delta}$ are Δ apart, $N \le \frac{p_N - p_0}{\Delta}$, and as every call takes at least two rounds, $N \le T$. Hence layer Δ takes $\mathcal{O}(\iota(p_N - p_0 + \Delta)\Delta^{-2}(\log_2 N + 2)^3) = \mathcal{O}((p_N - p_0 + \Delta)\Delta^{-2}\log^4 T)$ rounds, which proves claim (a).

Claim (c). Let S_Δ be the set of Δ-optimal prices of F, let $s(\Delta) := \frac{\operatorname{diam} S_\Delta}{\Delta}$, and let

$$S := \sup_{k \ge 0} \min\left\{T2^{-k}, s(2^{-k})\right\} = \sup_{k \ge 0} \min\left\{T2^{-k}, \frac{\operatorname{diam} S_{2^{-k}}}{2^{-k}}\right\} \tag{19}$$

be the supremum in claim (c), so that $S \ge \min\{T, s(1)\} = 1$ as every price is 1-optimal. The failure event in Lemma C.7 contributes $\mathcal{O}(1)$ regret, and if $p_F^* < 2^{-J}$ with $J := \lceil \log_2 T \rceil$, then $R_F(p_F^*) < \frac{1}{T}$ and the regret is less than 1. Otherwise, ScaleSearch has regret $\mathcal{O}(\log^2 T)$.

For each Δ, as $R_G(p) = \frac{R_F(\overline{p}p)}{\overline{p}}$ by Eq. (17), G's Δ-optimal prices are $\frac{1}{\overline{p}} S_{\overline{p}\Delta}$. So in layer Δ, the candidates satisfy $p_N - p_0 \le \frac{1}{\overline{p}} \operatorname{diam} S_{4\overline{p}\Delta}$, and by (a) and (b), the layer has regret $\mathcal{O}((p_N - p_0 + \Delta)\Delta^{-1}\log^4 T)$ under G. As every regret under G is multiplied by $\overline{p} \le 1$ under F, the layer has regret $\mathcal{O}((s(4\overline{p}\Delta) + 1)\log^4 T)$ under F, and each of its rounds has regret at most $\overline{p}(4\Delta + T^{-2})$.

For $i \ge 0$ and $\Delta \in (2^{-i-1}, 2^{-i}]$, as $S_\Delta \subseteq S_{2^{-i}}$, we have $s(\Delta) \le 2s(2^{-i})$ and $T\Delta \le T2^{-i}$, so $\min\{T\Delta, s(\Delta)\} \le 2S$. For $\Delta > 1$, $s(\Delta) \le \frac{1}{\Delta} < S$.

Let $\Delta_j := 4\overline{p}2^{-j}$. For every $k \ge 0$, by the bound for one layer above, the layers 2^{-j} with $1 \le j \le k$ have regret $\mathcal{O}(\log^4 T \sum_{j=1}^k (s(\Delta_j) + 1))$ in total, and all later rounds have regret $\mathcal{O}(T\Delta_k + 1)$ in total. Let $k_T := \lceil \log_2 T \rceil + 2$, so that $T\Delta_{k_T} \le 1$. If $s(\Delta_j) \le 2S$ for all $1 \le j \le k_T$, then $k = k_T$ gives regret $\mathcal{O}(S\log^5 T)$. Otherwise, let $k^* \le k_T$ be the smallest $j \ge 1$ with $s(\Delta_j) > 2S$. By the previous paragraph, $T\Delta_{k^*} \le 2S$, so $k = k^* - 1$ gives regret $\mathcal{O}(\log^4 T \cdot k^*(2S+1) + 2T\Delta_{k^*} + 1) = \mathcal{O}(S\log^5 T)$.

Adding the $\mathcal{O}(\log^2 T)$ regret of ScaleSearch, the regret remains $\mathcal{O}(S\log^5 T)$, which proves (c). □

Lemma C.11 (All Estimates Are Accurate)**.** *Let $\iota := \log(2T^3)$. With probability at least $1 - 3T^{-1}$, every call of Subroutine 1 in Algorithm 2, say on an interval I, satisfies Eq. (15) with $\mu = D_G(I)$ for every number $K \ge 2$ of queries it makes. On this event, which we denote by $\mathcal{E}$, every call with precision ϵ outputs $\widehat{D}(I)$ with $|\widehat{D}(I) - D_G(I)| \le \epsilon$, and it stops within $8\iota D_G(I)\epsilon^{-2} + 15\iota\epsilon^{-1} + 2$ queries unless the T rounds end or Line 12 stops it earlier.*

Proof. Each call uses fresh values and makes at most T queries, so by Lemma C.1 with $K_0 := T$ and $\delta = T^{-3}$, it violates Eq. (15) with probability at most $3T^{-2}$. As there are at most T calls, a union bound proves the first claim. On $\mathcal{E}$, a call stops once $r_K \le \epsilon$, which holds for every $K \ge 8\iota D_G(I)\epsilon^{-2} + 15\iota\epsilon^{-1}$ by Eq. (15), and then its output $\widehat{\mu}_K$ satisfies $|\widehat{\mu}_K - D_G(I)| \le r_K \le \epsilon$. □

Lemma C.12 (Revenue Estimation Error)**.** *On $\mathcal{E}$, for every layer Δ and every two candidates $p < p'$ in $\mathcal{S}_\Delta$ before the Eliminate step, the estimates in Eq. (5) satisfy*

$$\left|\left(\widehat{R}(p') - \widehat{R}(p)\right) - \left(R_G(p') - R_G(p)\right)\right| \le \frac{\Delta}{4}.$$

Proof. The key is that the error of $\widehat{D}(p_N)$ enters $\widehat{R}(p)$ multiplied by p, so it only contributes $(p' - p)$ times this error to the difference. Formally, fix a candidate p, and let $p_i := \max\{p'' \in \mathcal{S}_{2\Delta} \mid p'' \le p\}$ as in Eq. (5), so that $p \in \{p_i, p_i + \frac{\Delta}{2}\}$. As $\mathcal{I}_\mathcal{T}[p_i, p_N)$ partitions $[p_i, p_N)$ (Lemma C.3),

$$R_G(p) = pD_G(p) = p\left(D_G(p_N) + \sum_{I \in \mathcal{I}_\mathcal{T}[p_i, p_N)} D_G(I) - D_G([p_i, p))\right),$$

where $D_G([p_i, p_i)) := 0$. Comparing with Eq. (5), $\widehat{R}(p) - R_G(p) = p(e_{\text{tail}} + e_p)$, where e_{tail} is the error of $\widehat{D}(p_N)$ in Line 8, and e_p collects the errors of the tree intervals in $\mathcal{I}_\mathcal{T}[p_i, p_N)$ (Line 7) and of $\widehat{D}([p_i, p))$ (Lines 11 and 12).

- *The term pe_p.* $\mathcal{I}_\mathcal{T}[p_i, p_N)$ is empty if $N = 0$, and otherwise has at most $2(\log_2 N + 2)$ tree intervals (Lemma C.3), each starting at a price at least p_i, so on $\mathcal{E}$ each of their errors is at most $\frac{\Delta}{80(\log_2 N+2)(p_i+\Delta)}$, and the error of $\widehat{D}([p_i, p))$ is at most $\frac{\Delta}{40(p_i+\Delta)}$. As $p \le p_i + \Delta$, we get $|pe_p| \le \frac{\Delta}{40} + \frac{\Delta}{40} = \frac{\Delta}{20}$.
- *The term pe_{tail}.* On $\mathcal{E}$, $|e_{\text{tail}}| \le \frac{\Delta}{8(p_N - p_0 + \Delta)}$. As $p' - p \le p_N + \frac{\Delta}{2} - p_0$, we get $(p' - p)|e_{\text{tail}}| \le \frac{\Delta}{8}$.

The error of $\widehat{R}(p') - \widehat{R}(p)$ is $(p' - p)e_{\text{tail}} + p'e_{p'} - pe_p$, whose absolute value is thus at most $\frac{\Delta}{8} + \frac{\Delta}{20} + \frac{\Delta}{20} < \frac{\Delta}{4}$. □

Lemma C.13 (Candidates Always Approximate the Monopoly Price)**.** *On $\mathcal{E}$, after the Eliminate step of every layer Δ, the candidates $\mathcal{S}_\Delta$ contain a $\widehat{p}$ with $p^*_G - \frac{\Delta}{2} \le \widehat{p} \le p^*_G$ and $R_G(\widehat{p}) \ge R_G(p^*_G) - \frac{\Delta}{2}$, and every $p \in \mathcal{S}_\Delta$ satisfies $R_G(p) > R_G(p^*_G) - 2\Delta$.*

Proof. We use induction on the layers. For $\mathcal{S}_1 = \{0, \frac{1}{2}, 1\}$, some $\widehat{p} \in \mathcal{S}_1$ has $p^*_G - \frac{1}{2} \le \widehat{p} \le p^*_G$. Now consider layer Δ, and suppose that some $p' \in \mathcal{S}_{2\Delta}$ has $p^*_G - \Delta \le p' \le p^*_G$.

- *A candidate close to p^*_G.* The Refine step adds p' to $\mathcal{S}_\Delta$, and also $p' + \frac{\Delta}{2}$ if $p' + \frac{\Delta}{2} \le 1$. One of them, say $\widehat{p}$, satisfies $p^*_G - \frac{\Delta}{2} \le \widehat{p} \le p^*_G$: namely $p' + \frac{\Delta}{2}$ if $p' + \frac{\Delta}{2} \le p^*_G$ (which is then added, as $p^*_G \le 1$), and p' otherwise. By Eq. (1), $R_G(\widehat{p}) \ge R_G(p^*_G) - (p^*_G - \widehat{p}) \ge R_G(p^*_G) - \frac{\Delta}{2}$.
- *It survives Line 12.* This line only removes $p_N + \frac{\Delta}{2}$, so it only matters if $\widehat{p} = p_N + \frac{\Delta}{2}$, i.e., $p' = p_N$ and $p_N + \frac{\Delta}{2} \le p^*_G$. Then Eq. (18) and $R_G(p_N) \le R_G(p^*_G)$ give $p_N D_G([p_N, p^*_G)) \le p^*_G - p_N \le \Delta$. So $D_G([p_N, p_N + \frac{\Delta}{2})) \le \frac{\Delta}{p_N}$, and on $\mathcal{E}$, its lower confidence bound never exceeds $\frac{2\Delta}{p_N}$.
- *It survives the Eliminate step.* Let $\widehat{p}^* := \operatorname{argmax}_{p \in \mathcal{S}_\Delta} \widehat{R}(p)$. By Lemma C.12, $\widehat{R}(\widehat{p}^*) - \widehat{R}(\widehat{p}) \le R_G(\widehat{p}^*) - R_G(\widehat{p}) + \frac{\Delta}{4} \le R_G(p^*_G) - R_G(\widehat{p}) + \frac{\Delta}{4} \le \frac{3}{4}\Delta < \Delta$.

This proves the first claim for $\mathcal{S}_\Delta$, which is also the induction hypothesis for the next layer. For the second claim, every survivor p has $\widehat{R}(\widehat{p}) - \widehat{R}(p) \le \widehat{R}(\widehat{p}^*) - \widehat{R}(p) < \Delta$, so by Lemma C.12 again,

$$R_G(p_G^*) - R_G(p) \le (R_G(p_G^*) - R_G(\widehat{p})) + \left(\widehat{R}(\widehat{p}) - \widehat{R}(p)\right) + \frac{\Delta}{4} < \frac{\Delta}{2} + \Delta + \frac{\Delta}{4} < 2\Delta. \qquad \square$$

C.5 Proofs of Theorems 3.4 and 3.5 (Upper Bounds with Size-2 Menus)

We first reduce Theorem 3.4 to the covering numbers of the rescaled distribution G in Eq. (17):

Lemma C.14 (Reduction to Covering Numbers)**.** *Let F be a value distribution with $R_F(p_F^*) > 0$. Except on events that contribute $\mathcal{O}(1)$ to the regret, either the regret of Algorithm 2 on F is $\mathcal{O}(\log^2 T)$, or* ScaleSearch *outputs some $\overline{p} \in [p_F^*, 1]$, and the regret is at most*

$$\mathcal{O}(\log^2 T) + \mathcal{O}\left(\overline{p} W_G \log^4 T\right), \qquad \text{where } W_G := \min_{k \ge 0} \left\{ T 2^{-k} + \sum_{j=0}^{k} N_G(2^{-j}) \right\}, \tag{20}$$

and N_G is the covering number of the rescaled distribution G in Eq. (17), which satisfies

$$N_G(\Delta) \le N_F(\overline{p}\Delta) \quad \text{and} \quad N_G(\Delta) \le \frac{2}{\Delta} \qquad \text{for all } \Delta \in (0, 1]. \tag{21}$$

Proof. Let $J := \lceil \log_2 T \rceil$. The failure events of Lemmas C.7 and C.11 contribute $\mathcal{O}(1)$ regret, and if $p_F^* < 2^{-J} \le \frac{1}{T}$, the regret is at most $T R_F(p_F^*) < 1$. Otherwise, by Lemma C.7, ScaleSearch has regret $\mathcal{O}(\log^2 T)$, which is the whole regret if the T rounds end before it stops; otherwise, it outputs $\overline{p} \in [p_F^*, 1]$. The rest of Algorithm 2 runs on G in Eq. (17): posting $p \in [0, 1]$ to G corresponds to posting $\overline{p}p$ to F, and $R_G(p) = \frac{R_F(\overline{p}p)}{\overline{p}}$ is maximized at $\frac{p_F^*}{\overline{p}} \in [0, 1]$, so every regret under G is multiplied by $\overline{p}$ under F. Hence by Lemma C.18 below for G, the total regret is at most Eq. (20). Finally, if p is Δ-optimal for G, then $\overline{p}p$ is $\overline{p}\Delta$-optimal for F, so Lemma C.17 below for G gives Eq. (21). $\square$

Compared with $V_T(F)$ in Theorem 3.5 and with the supremum in Lemma 3.3 (which has $\frac{\mathrm{diam}}{\Delta}$ in place of N and only dyadic Δ), W_G differs in two ways. First, W_G sums over the scales, whereas the other two only take the largest scale; the sum is at most $\mathcal{O}(\log T)$ times the largest one (claim (c) of Lemma 3.3), which brings the extra $\log T$ in Theorem 3.5, while summing directly keeps $\log^4 T$ in Theorem 3.4. Second, W_G is for the rescaled G rather than F, which matters for MHR distributions: $\overline{p} W_G = \mathcal{O}(T^{1/3})$ for every MHR distribution (see the proof below), whereas $V_T(F)$ itself can be $\Theta(\sqrt{T})$ (Footnote 9).

Theorem C.15 (Restatement of Theorem 3.4; Upper Bounds)**.** *For dynamic pricing with size-2 menus, Algorithm 2 ensures:*

(a) $\mathrm{Reg}^{(2)}_{\mathrm{all}}(T) = \mathcal{O}(\sqrt{T} \log^4 T)$;
(b) $\mathrm{Reg}^{(2)}_{\mathrm{MHR}}(T) = \mathcal{O}(T^{1/3} \log^4 T)$;
(c) $\mathrm{Reg}^{(2)}_{\mathrm{curv}(\kappa)}(T) = \mathcal{O}((\kappa^{-1/3} T^{1/3} + \log T) \log^4 T)$ *for every curvature $\kappa \in (0, 2)$.*

Proof. If $R_F(p_F^*) = 0$, then $v = 0$ a.s. and every algorithm has regret 0. Otherwise, by Lemma C.14 above, it suffices to bound $\overline{p} W_G$ in Eq. (20) for each class, recalling $\overline{p} \in [p_F^*, 1]$, the rescaled G in Eq. (17), and Eq. (21).

(a) *General distributions.* Taking k with $2^k \in [\sqrt{T}, 2\sqrt{T})$, Eq. (21) gives $\overline{p}W_G \le T2^{-k} + \sum_{j\le k} 2^{j+1} \le \sqrt{T} + 2^{k+2} \le 9\sqrt{T}$.

(b) *MHR distributions.* By Lemma C.7, $\overline{p} < 10p_F^*$, and by Lemma C.8, $R_F(p_F^*) \ge \frac{p_F^*}{e}$, so $\frac{R_F(p_F^*)}{4\overline{p}} \ge \frac{1}{40e}$. As Lemma C.9 gives $N_F(\Delta) \le 1 + 8\Delta^{-1/2}$ for $\Delta \le \frac{R_F(p_F^*)}{4}$, Eq. (21) gives $N_G(\Delta) \le 1 + 8(\overline{p}\Delta)^{-1/2}$ for $\Delta \le \frac{1}{40e}$. If $\overline{p}T^2 < 1$, then $R_F(p_F^*) \le \overline{p} < T^{-2}$ and the regret is less than 1. Otherwise, take k with $2^k \in [\overline{p}^{1/3}T^{2/3}, 2\overline{p}^{1/3}T^{2/3})$. In W_G, the terms with $2^{-j} > \frac{1}{40e}$ sum to at most $\sum_{j:2^j<40e} 2^{j+1} \le 160e$ by Eq. (21), and the other terms sum to at most $\sum_{j\le k}(1 + 8(\overline{p}2^{-j})^{-1/2}) \le k + 1 + 28\overline{p}^{-1/2}2^{k/2}$. Hence $\overline{p}W_G \le \overline{p}(T2^{-k} + 160e + k + 1 + 28\overline{p}^{-1/2}2^{k/2}) = \mathcal{O}(\overline{p}^{2/3}T^{1/3} + \log T) = \mathcal{O}(T^{1/3})$.

(c) *Distributions with curvature κ.* G has curvature $\kappa\overline{p}$, as $R_G(\frac{p_F^*}{\overline{p}}) - R_G(p) = \frac{R_F(p_F^*) - R_F(\overline{p}p)}{\overline{p}} \ge \kappa\overline{p}(p - \frac{p_F^*}{\overline{p}})^2$. So every Δ-optimal price of G is within $(\Delta/(\kappa\overline{p}))^{1/2}$ of $\frac{p_F^*}{\overline{p}}$, and $N_G(\Delta) \le 1 + 2(\kappa\overline{p}\Delta)^{-1/2}$. If $\kappa\overline{p}T^2 < 1$, the regret is at most $TR_F(p_F^*) \le T\overline{p} < \frac{1}{\kappa T}$, which is at most $\kappa^{-1/3}T^{1/3}$ if $\kappa \ge T^{-2}$, while $\kappa^{-1/3}T^{1/3} > T$ otherwise. Otherwise, take k with $2^k \in [(\kappa\overline{p})^{1/3}T^{2/3}, 2(\kappa\overline{p})^{1/3}T^{2/3})$, so that $k = \mathcal{O}(\log T)$ as $\kappa \le 2$ (Lemma B.1), and $\overline{p}W_G \le \overline{p}(T2^{-k} + k + 1 + 7(\kappa\overline{p})^{-1/2}2^{k/2}) = \mathcal{O}(\kappa^{-1/3}T^{1/3} + \log T)$.

Plugging these bounds into Eq. (20) proves the theorem. □

Theorem C.16 (Restatement of Theorem 3.5; Instance-Dependent Regret Upper Bound)**.** *For any value distribution F and any $T \ge 2$, the regret of Algorithm 2 is $\mathcal{O}(V_T(F)\log^5 T)$, where*

$$V_T(F) := \sup_{\Delta\in(0,1]} \min\{T\Delta, N_F(\Delta)\},$$

which implies regret $\mathcal{O}(CT^{d/(d+1)}\log^5 T)$ on any F with interval dimension d and multiplier C.

Proof. By claim (c) of Lemma 3.3, where $\widetilde{\mathcal{O}}$ hides $\log^5 T$, it suffices to show that its supremum in Eq. (19) satisfies

$$S = \sup_{k\ge 0} \min\left\{T2^{-k}, \frac{\operatorname{diam} S_{2^{-k}}}{2^{-k}}\right\} \le V_T(F),$$

where S_Δ is the set of Δ-optimal prices of the original value distribution F. Indeed, for every $\Delta = 2^{-k}$, Lemma C.17 gives $\frac{\operatorname{diam} S_\Delta}{\Delta} = \frac{u_\Delta - \ell_\Delta}{\Delta} \le N_F(\Delta)$, so $\min\{T\Delta, \frac{\operatorname{diam} S_\Delta}{\Delta}\} \le \min\{T\Delta, N_F(\Delta)\} \le V_T(F)$.

For the second claim, it suffices to show $V_T(F) \le CT^{d/(d+1)}$. Indeed, $C \ge N_F(1) = 1$, and every $\Delta \in (0,1]$ has either $\Delta \le T^{-1/(d+1)}$, so $T\Delta \le T^{d/(d+1)}$, or $\Delta > T^{-1/(d+1)}$, so $N_F(\Delta) \le C\Delta^{-d} < CT^{d/(d+1)}$. □

We now prove the lemmas used above, in order of first use.

Lemma C.17 (Covering Numbers)**.** *For any value distribution F and $\Delta > 0$, the Δ-optimal prices have a smallest one ℓ_Δ and a largest one u_Δ, where $u_\Delta > \ell_\Delta$, and the covering number $N_F(\Delta) = \lceil \frac{u_\Delta - \ell_\Delta}{\Delta} \rceil$ (defined in Section 3.4 for $\Delta \le 1$, and by the same formula for $\Delta > 1$) satisfies*

$$\max\left\{1, \frac{u_\Delta - \ell_\Delta}{\Delta}\right\} \le N_F(\Delta) \le 1 + \frac{u_\Delta - \ell_\Delta}{\Delta}.$$

In particular, $N_F(\Delta) = 1$ for $\Delta \ge 1$, and $N_F(\Delta) \le \frac{2}{\Delta}$ for $\Delta \in (0,1]$.

Proof. As $D_F(p)$ is non-increasing and left-continuous in p, R_F is upper semi-continuous, so the set of Δ-optimal prices is closed, and it has a smallest and a largest element. They differ: If $\Delta \geq R_F(p_F^*)$, then every price is Δ-optimal as $R_F \geq 0$; otherwise, $p_F^* \geq R_F(p_F^*) > \Delta$, and every $p \in [p_F^* - \Delta, p_F^*]$ is Δ-optimal by Eq. (1). As $u_\Delta - \ell_\Delta > 0$, the display follows from the definition of $N_F(\Delta)$. Finally, every price is 1-optimal as $R_F \leq 1$, so $u_\Delta - \ell_\Delta \leq 1$ for all Δ, which gives the last two claims. □

Lemma C.18 (Regret of the Layers via Covering Numbers)**.** *Run Algorithm 2 without Line 1, i.e., with $\overline{p} = 1$, on a value distribution G with covering number N_G. Its regret is $\mathcal{O}(W_G \log^4 T)$, where $W_G := \min_{k \geq 0}\{T2^{-k} + \sum_{j=0}^{k} N_G(2^{-j})\}$.*

Proof. We bound the regret of each layer by its covering number, and the regret of all rounds after a given layer by its scale. The failure event in Lemma C.11 contributes $\mathcal{O}(1)$ regret. Otherwise, fix a layer Δ and write $\mathcal{S}_{2\Delta} = \{p_0 < p_1 < \cdots < p_N\}$. By Lemma C.13, all of $p_0, \ldots, p_N$ are 4Δ-optimal, so each round in this layer has regret at most $4\Delta + T^{-2}$, and $p_N - p_0 \leq u_{4\Delta} - \ell_{4\Delta} \leq 4\Delta N_G(4\Delta)$ by Lemma C.17. As $N_G(4\Delta) \geq 1$, Lemma 3.3 shows that this layer takes $\mathcal{O}(N_G(4\Delta)\Delta^{-1}\log^4 T)$ rounds, so its regret is $\mathcal{O}(N_G(4\Delta)\log^4 T)$. Now fix any $k \geq 0$. The layers $\Delta = 2^{-i}$ with $1 \leq i \leq k+2$ have $4\Delta = 2^{-(i-2)}$, so their total regret is $\mathcal{O}(\log^4 T \cdot (N_G(2) + \sum_{j=0}^{k} N_G(2^{-j})))$, where $N_G(2) = 1$. Each round in a later layer has regret at most $4 \cdot 2^{-k-3} + T^{-2} \leq 2^{-k} + T^{-2}$. Hence the regret is $\mathcal{O}(\log^4 T \cdot (T2^{-k} + \sum_{j=0}^{k} N_G(2^{-j})))$ for every k, and minimizing over k proves the lemma. □

C.6 Examples of Interval Dimensions

We list value distributions with each interval dimension d, on which Algorithm 2 beats the worst-case $\widetilde{\mathcal{O}}(\sqrt{T})$ regret whenever $d < 1$. Most of them follow from the following lemma.

Lemma C.19 (Polynomial Drops Give Interval Dimensions)**.** *Let $\alpha \geq 1$ and $c > 0$. If a value distribution F satisfies $R(p^*) - R(p) \geq c|p - p^*|^\alpha$ for all $p \in [0, 1]$, then F has interval dimension $1 - \frac{1}{\alpha}$ with multiplier $2c^{-1/\alpha} + 1$.*

Proof. Let $\Delta \in (0, 1]$. Every Δ-optimal p has $c|p - p^*|^\alpha \leq \Delta$, i.e., $|p - p^*| \leq (\frac{\Delta}{c})^{1/\alpha}$, so $u_\Delta - \ell_\Delta \leq 2(\frac{\Delta}{c})^{1/\alpha}$. By Lemma C.17, $N(\Delta) \leq 1 + 2c^{-1/\alpha}\Delta^{1/\alpha - 1} \leq (2c^{-1/\alpha} + 1)\Delta^{-(1-1/\alpha)}$, as $\Delta^{1/\alpha - 1} \geq 1$. □

Example 4 (Interval Dimension 0)**.** In each of the following, R has a unique maximizer p^* and drops at least linearly around it, i.e., $R(p^*) - R(p) \geq c|p - p^*|$ for some $c > 0$ and all $p \in [0, 1]$. So by Lemma C.19 with $\alpha = 1$, $N(\Delta) \leq \frac{2}{c} + 1$ for all $\Delta \in (0, 1]$, i.e., F has interval dimension 0.

- *Finitely many values* [CCP19], with a unique optimal price p^*. Below the smallest value and between two consecutive values, the demand $D(p)$ is constant, so R increases linearly there; above the largest value, $R = 0$. Hence $p^* > 0$ is one of the values, R is linear with slope $D(p^*) > 0$ on $(v', p^*]$, where v' is the previous value (or 0), and the other local maxima of R are at the finitely many other values, so $R \leq R(p^*) - g$ outside $(v', p^*]$ for some $g > 0$. As $|p - p^*| \leq 1$, $c := \min\{D(p^*), g\}$ works.
- *Kinked demands*, i.e., at a unique maximizer p^*, R has a positive left derivative and a negative right derivative. This models the kinked-demand theory of oligopoly [Swe39; HH39]: The rivals of a firm match its price cuts but not its price increases, so its demand barely grows when it cuts its price below the prevailing price p^* (i.e., the price elasticity $\varepsilon(p) := \frac{pf(p)}{D(p)}$ is less than 1, so $R'(p) = D(p)(1 - \varepsilon(p)) > 0$), but quickly drops when it raises its price above p^* (i.e., $\varepsilon(p) > 1$, so $R'(p) < 0$). The one-sided derivatives give some $c', r > 0$ with $R(p^*) - R(p) \geq c'|p - p^*|$ for

$|p - p^*| < r$. As R is upper semi-continuous and p^* is its unique maximizer, $R \le R(p^*) - g$ for some $g > 0$ on the compact set $\{p \in [0,1] \mid |p - p^*| \ge r\}$. So $c := \min\{c', g\}$ works.

- *Pareto distributions* with shape $\beta > 1$ and scale $p_0 < 1$, with values above 1 rounded down to 1, i.e., $D(p) = \min\{1, (\frac{p_0}{p})^\beta\}$. Then $R(p) = p$ below p_0, and $R(p) = p_0^\beta p^{1-\beta}$ above p_0, which is convex and decreasing. So $p^* = p_0$, $R(p_0) - R(p) = p_0 - p$ for $p < p_0$, and by convexity, $R(p_0) - R(p) \ge \frac{R(p_0)-R(1)}{1-p_0}(p - p_0)$ for $p > p_0$. So $c := \min\{1, \frac{p_0 - p_0^\beta}{1-p_0}\}$ works.
- *Uniform distributions* $\mathrm{Unif}[a,b]$ with $\frac{b}{2} < a < b \le 1$, another case where the monopoly price is the lowest value. Here $R(p) = p$ below a, $R(p) = \frac{p(b-p)}{b-a}$ on $[a,b]$, which is concave with right derivative $\frac{b-2a}{b-a} < 0$ at a, and $R = 0$ above b. So $p^* = a$, and $R(a) - R(p) \ge \frac{2a-b}{b-a}(p - a)$ for $p \in [a,b]$ by concavity. As $R(a) - R(p) = a$ for $p > b$, $c := \min\{a, \frac{2a-b}{b-a}\}$ works.

Example 5 (Interval Dimension $\frac{1}{2}$)**.** Each of the following has interval dimension $\frac{1}{2}$.

- *MHR distributions* with $R(p^*) > 0$. By Lemma C.9, $N(\Delta) \le 1 + 8\Delta^{-1/2} \le 9\Delta^{-1/2}$ for $\Delta \le \frac{R(p^*)}{4}$, and by Lemma C.17, $N(\Delta) \le \frac{2}{\Delta} < 4R(p^*)^{-1/2}\Delta^{-1/2}$ for $\Delta > \frac{R(p^*)}{4}$. So F has interval dimension $\frac{1}{2}$ with multiplier $\max\{9, 4R(p^*)^{-1/2}\}$.
- *Distributions with curvature* κ, i.e., $F \in \mathcal{F}_{\mathrm{curv}(\kappa)}$, which satisfies $R(p^*) - R(p) \ge \kappa(p - p^*)^2$ for all p by definition. So Lemma C.19 with $\alpha = 2$ gives interval dimension $\frac{1}{2}$ with multiplier $2\kappa^{-1/2} + 1$.
- *Smooth around a unique maximizer.* If F has a density that is continuously differentiable around a unique maximizer $p^* \in (0,1)$ of R with $R''(p^*) < 0$, which satisfies the "curvature and smooth" assumption of Kleinberg and Leighton [KL03, Theorem 1.2] (see Footnote 7), then $F \in \mathcal{F}_{\mathrm{curv}(\kappa)}$ for some $\kappa \in (0,2)$ by Kleinberg and Leighton [KL03, Lemma 3.11], so the previous bullet applies. Examples include the uniform, exponential, logistic (i.e., logit demands), and normal distributions, rescaled or truncated to $[0,1]$ with an interior maximizer.

Example 6 (Interval Dimension 1)**.** Every value distribution always has interval dimension 1 with multiplier 2, because $N(\Delta) \le \frac{2}{\Delta}$ (Lemma C.17). In each of the following, we verify that $N(\Delta) \ge \frac{c'}{\Delta}$ for some $c' > 0$ and all $\Delta \in (0,1]$, so it has no interval dimension $d < 1$.

- *The equal-revenue distribution* $D(p) = \min\{1, \frac{1}{2p}\}$, i.e., unit-elastic demand. Its revenue is $R(p) = \frac{1}{2}$ for all $p \in [\frac{1}{2}, 1]$, so all these prices are optimal, and $N(\Delta) \ge \frac{1}{2\Delta}$.
- *Two optimal prices* where R has another maximizer $p' \ne p^*$, e.g., finitely many values with two optimal prices. As p' and p^* are Δ-optimal for every Δ, $N(\Delta) \ge \frac{|p' - p^*|}{\Delta}$.

Example 7 (Every Interval Dimension $d \in [0,1)$)**.** We now give a family of distributions that can have any interval dimension $d \in [0,1)$, which will be useful for the construction of our lower bounds.

- *The family* F_α. For $\alpha \ge 1$, let F_α be given by $D_{F_\alpha}(p) = \frac{1-(2p-1)^\alpha}{2p}$ for $p \in [\frac{1}{2}, 1]$ and $D_{F_\alpha}(p) = 1$ for $p < \frac{1}{2}$. It is non-increasing, as its revenue function $R(p) = \frac{1-(2p-1)^\alpha}{2}$ is non-increasing on $[\frac{1}{2}, 1]$, and $R(p) = p$ below $\frac{1}{2}$. Its Δ-optimal prices form $[\frac{1}{2} - \Delta, \frac{1}{2} + \frac{1}{2}(2\Delta)^{1/\alpha}]$ for $\Delta \le \frac{1}{2}$, so it has interval dimension $1 - \frac{1}{\alpha}$ with multiplier 3. It interpolates from F_1 with $D_{F_1}(p) = \frac{1}{p} - 1$ ($d = 0$), through $F_2 = \mathrm{Unif}[\frac{1}{2}, 1]$ ($d = \frac{1}{2}$), to the equal-revenue distribution as $\alpha \to \infty$ ($d \to 1$).
- *Random distributions.* The random value distributions in Theorem 4.6 also have every interval dimension $d \in (0,1)$.

When $V_T(F)$ Tends to Infinity. We also discuss when $V_T(F)$, the instance-dependent quantity defined in Eq. (6), tends to infinity. We claim that it stays bounded only for value distributions with interval dimension 0, which are exactly those whose revenue drops at least linearly as in Example 4:

Lemma C.20 (When $V_T(F)$ Tends to Infinity)**.** *Let F be a value distribution. Then $V_T(F)$ is non-decreasing in T, and $\lim_{T\to\infty} V_T(F) = \sup_{\Delta\in(0,1]} N_F(\Delta)$. Moreover, the following are equivalent:*

1. *$V_T(F)$ is bounded, i.e., $\sup_{T\geq 1} V_T(F) < \infty$.*
2. *F has interval dimension 0.*
3. *There is $c > 0$ such that $R(p^*) - R(p) \geq c|p - p^*|$ for all $p \in [0, 1]$.*

In particular, $V_T(F) \to \infty$ as $T \to \infty$ if R has another maximizer $p' \neq p^*$; or if $p^* \in (0, 1)$ and R is differentiable at p^*, e.g., if F has a density around p^* that is continuous at p^*.

Proof. Write $R^* := R(p^*)$, $N := N_F$, and $N^* := \sup_{\Delta\in(0,1]} N(\Delta) \in [0, \infty]$. For $T \leq T'$ and every Δ, we have $\min\{T\Delta, N(\Delta)\} \leq \min\{T'\Delta, N(\Delta)\} \leq N(\Delta)$, so $V_T(F)$ is non-decreasing in T and at most N^*. Conversely, for every $\Delta \in (0, 1]$ and $T \geq \frac{N(\Delta)}{\Delta}$, we have $V_T(F) \geq \min\{T\Delta, N(\Delta)\} = N(\Delta)$. Hence $\lim_{T\to\infty} V_T(F) = N^*$, so (1) holds if and only if $N^* < \infty$. This is equivalent to (2): If $N^* < \infty$, then F has interval dimension 0 with multiplier N^*, and if F has interval dimension 0 with multiplier C, then $N^* \leq C$.

For (2) $\Rightarrow$ (3), let $N(\Delta) \leq C$ for all $\Delta \in (0, 1]$. Fix $p \in [0, 1]$, and let $\Delta := R^* - R(p) \in [0, 1]$. For every $\Delta' \in (0, 1]$ with $\Delta' \geq \Delta$, both p and p^* are Δ'-optimal, so $|p-p^*| \leq u_{\Delta'} - \ell_{\Delta'} \leq \Delta' N(\Delta') \leq C\Delta'$ by Lemma C.17. If $\Delta > 0$, then taking $\Delta' = \Delta$ gives $R^* - R(p) \geq \frac{1}{C}|p - p^*|$. If $\Delta = 0$, then letting $\Delta' \to 0$ gives $p = p^*$. Hence (3) holds with $c = \frac{1}{C}$.

For (3) $\Rightarrow$ (2), every Δ-optimal price p satisfies $|p - p^*| \leq \frac{R^*-R(p)}{c} \leq \frac{\Delta}{c}$, so $u_\Delta - \ell_\Delta \leq \frac{2\Delta}{c}$ and $N(\Delta) \leq \lceil \frac{2}{c} \rceil$ for all $\Delta \in (0, 1]$. That is, F has interval dimension 0 with multiplier $\lceil \frac{2}{c} \rceil$, and $V_T(F) \leq \lceil \frac{2}{c} \rceil$ for all T.

Finally, (3) fails in both cases of the last claim. If $p' \neq p^*$ is another maximizer, then $R^* - R(p') = 0 < c|p' - p^*|$ for every $c > 0$. Indeed, as p' and p^* are Δ-optimal for every Δ, we have $N(\Delta) \geq \frac{|p'-p^*|}{\Delta}$, so taking $\Delta = \sqrt{|p' - p^*|/T}$ gives $V_T(F) \geq \sqrt{|p' - p^*|T}$ for $T \geq |p' - p^*|$. If $p^* \in (0, 1)$ and R is differentiable at p^*, then $R'(p^*) = 0$ as p^* maximizes R, so $\frac{R^*-R(p)}{|p-p^*|} \to 0$ as $p \to p^*$. Moreover, if F has a density f around p^* that is continuous at p^*, then $D(p)$ is differentiable at p^* with derivative $-f(p^*)$, so $R(p) = pD(p)$ is differentiable at p^* as well. □

D Omitted Proofs in Section 4 (Lower Bounds)

All our regret lower bounds compare two close value distributions, which we state as the two-point method for regret in Appendix D.1. We then prove the minimax lower bounds (Appendix D.2), the local minimax lower bounds (Appendices D.3, D.4, and D.5), the anytime lower bounds (Appendix D.6), and a lower bound for Lipschitz experts (Appendix D.7).

D.1 Two-Point Method for Regret

As in Appendix B.3, for a value distribution G, its demand at a price p is $D_G(p) := \Pr_{v\sim G}\{v \geq p\}$, the probability of an interval I is $D_G(I) := \Pr_{v\sim G}\{v \in I\}$, and $R_G(p^*_G) - R_G(p)$ is the *gap* of a price p under G. For $G \ll G'$ (i.e., G's support is a subset of that of G'), let $\chi^2(G\|G') := \int (\frac{dG}{dG'} - 1)^2 \, dG'$, so that $H^2(G, G') \leq \frac{1}{2}\chi^2(G\|G')$. As $1 - H^2$ is multiplicative over independent samples, n samples have total variation distance at most $\sqrt{2nH^2(G, G')}$ between G and G'. As in Lemma B.8, a two-point lower bound takes two value distributions that are close but share no near-optimal price.

Lemma D.1 (Two-Point Method for Regret). *Let $\Delta > 0$, and let G_+, G_- be two value distributions such that every price has gap at least Δ under G_+ or under G_-. If for some $T' \le T$, the first $T'-1$ values have total variation distance at most $\frac{1}{2}$ between G_+ and G_-, e.g., if $H^2(G_+, G_-) \le \frac{1}{8T'}$, then any algorithm, even with infinite menus and full feedback, suffers regret at least $\frac{\Delta T'}{4}$ on G_+ or on G_-.*

Proof. We show that, in each round $t \le T'$, the regrets under G_+ and under G_- sum to at least $\frac{\Delta}{2}$. Fix such a round t, and let P be the price distribution in Lemma A.3 for the menu in round t, which only depends on the first $t-1$ values and the internal randomness. Let $B_\pm$ be the set of prices with gap at least Δ under $G_\pm$, and $\Phi_\pm := \Pr_{p\sim P}\{p \in B_\pm\}$. As $B_+ \cup B_- = [0,1]$, we have $\Phi_+ + \Phi_- \ge 1$. By Lemma A.3, the round-t regret under $G_\pm$ is at least $\Delta\,\mathbb{E}_{G_\pm}[\Phi_\pm]$. As the first $t-1$ values have total variation distance at most $\frac{1}{2}$, the sum of the two is at least $\Delta(\mathbb{E}_{G_+}[\Phi_+] + \mathbb{E}_{G_+}[\Phi_-] - \frac{1}{2}) \ge \frac{\Delta}{2}$. Summing over $t \le T'$, the regrets in the first T' rounds under G_+ and under G_- sum to at least $\frac{\Delta T'}{2}$, so one of them is at least $\frac{\Delta T'}{4}$. As every round has non-negative regret by Lemma A.3, the same holds for the T-round regret. □

D.2 Proof of Theorem 4.1 (Minimax Lower Bounds)

Theorem D.2 (Restatement of Theorem 4.1; Minimax Lower Bounds). *For dynamic pricing with infinite menus and full feedback:*

(a) $\mathrm{Reg}^{(\infty)}_{\mathrm{all}}(T) = \Omega(\sqrt{T})$;
(b) $\mathrm{Reg}^{(\infty)}_{\mathrm{MHR}}(T) = \Omega(T^{1/3})$;
(c) $\mathrm{Reg}^{(\infty)}_{\mathrm{curv}(\kappa)}(T) = \Omega(\kappa^{-1/3}T^{1/3})$ *for every curvature* $\kappa \in (0,2)$;
(d) *the minimax regret over value distributions with interval dimension* $d \in (0,1)$ *is* $\Omega(T^{d/(d+1)})$.

Proof. Each bound takes two value distributions in the class that are close in Hellinger distance but share no near-optimal price, and applies Lemma D.1:

- (a) General and regular distributions: Lemma D.3 below uses the regular pair in Lemma B.15, which tilts the revenue plateau of a regular distribution to the left or to the right.
- (b) MHR distributions: Theorem 4.5 applies to $F = \mathrm{Unif}[0,1]$, which is MHR with monopoly price $\frac{1}{2}$, and whose hazard rate $\frac{1}{1-v}$ has derivative in $[1,16]$ on $(0, \frac{3}{4})$.
- (c) Curvature κ: Lemma D.4 below uses the pairs in Lemmas B.19 and B.20, which shift a quadratic revenue peak or the kink of R_U.
- (d) Interval dimension $d \in (0,1)$: Let F be F_α in Appendix C.6 with $\alpha = \frac{1}{1-d}$, which has interval dimension d with multiplier 3, and $N(\Delta) \ge \frac{\Delta^{-d}}{2}$ for $\Delta \le \frac{1}{2}$, as its Δ-optimal prices form $[\frac{1}{2} - \Delta, \frac{1}{2} + \frac{1}{2}(2\Delta)^{1/\alpha}]$. Using Theorem 4.4, there exists a constant $c > 0$ and, for every $T \ge \frac{1}{c}$, a G such that F and G both have interval dimension d with multiplier $\frac{3}{c}$, and any algorithm suffers regret at least $cT^{d/(d+1)}$ on one of them. □

Lemma D.3 (Regret Lower Bound for Regular Distributions). *For every $T \ge 2$, $\mathrm{Reg}^{(\infty)}_{\mathrm{regular}}(T) \ge \frac{\sqrt{T}}{136}$, even if the seller observes v_t in each round, and even for the class of value distributions that are both regular and quasi-MHR.*

Proof. Let $\epsilon := (32T)^{-1/2} \le \frac{1}{8}$, and let $F_\pm$ be the pair in Lemma B.15, which are regular and quasi-MHR. Then $H^2(F_+, F_-) \le 4\epsilon^2 = \frac{1}{8T}$, and every price has gap at least $\frac{\epsilon}{6}$ under F_+ or under F_-. So Lemma D.1 with $\Delta = \frac{\epsilon}{6}$ and $T' = T$ gives $\mathrm{Reg}^{(\infty)}_{\{F_+,F_-\}}(T) \ge \frac{\epsilon T}{24} \ge \frac{\sqrt{T}}{136}$. □

Remark 8 (Learning a Uniform Reserve for Quasi-MHR Buyers)**.** Feng and Jin [FJ25] introduce quasi-MHR distributions, i.e., those with a non-decreasing *conditional expected hazard rate* $\frac{1}{v}\ln\frac{1}{D_F(v)}$, which generalize MHR distributions. Among many problems, they study the sample complexity of the *empirical uniform reserve*, i.e., learning from samples a uniform reserve price whose revenue is a $(1-\epsilon)$-approximation of the optimal one. The tight bound for quasi-MHR buyers is left open, which they show to be between $\widetilde{\Omega}(\epsilon^{-1.5})$ and $\widetilde{\mathcal{O}}(\epsilon^{-2})$. We close this by showing their $\widetilde{\mathcal{O}}(\epsilon^{-2})$ upper bound is tight up to logarithmic factors: Even for a single buyer, for whom a uniform reserve is a posted price, the two value distributions in Lemma B.15 are regular and quasi-MHR, with optimal revenues at least $\frac{1}{2}$, and any estimator needs $\Omega(\epsilon^{-2})$ samples (Lemma B.14). The main reason is that, while quasi-MHR distributions keep the tail bounds of MHR distributions, the faster rates for MHR distributions (Theorem 3.4 and [HMR15]) rely on the quadratic drop of R around p^* (Lemma C.9), which is lost for quasi-MHR distributions: The revenues of the pair above are almost flat on $[\frac{1}{2},\frac{3}{4}]$.

Lemma D.4 (Regret Lower Bound for Curvature Distributions)**.** *For every $\kappa\in(0,2)$ and $T\geq 3\max\{\kappa^{-2},(2-\kappa)^{-3}\}$, $\mathrm{Reg}^{(\infty)}_{\mathrm{curv}(\kappa)}(T)\geq\frac{\kappa^{-1/3}T^{1/3}}{1100}$, even if the seller observes v_t in each round.*

Proof. We use the two pairs in Lemmas B.19 and B.20, depending on κ.

- If $\kappa\leq\frac{1}{8}$, let $s:=(12800\kappa^2T)^{-1/3}$, which is at most $\frac{1}{16}$ as $T\geq\kappa^{-2}$, and let $G_\pm$ be the pair in Lemma B.19. Then $H^2(G_+,G_-)\leq 1600\kappa^2s^3=\frac{1}{8T}$, so Lemma D.1 with $\Delta=2\kappa s^2$ and $T'=T$ gives $\mathrm{Reg}^{(\infty)}_{\{G_+,G_-\}}(T)\geq\frac{\kappa s^2T}{2}\geq\frac{\kappa^{-1/3}T^{1/3}}{1100}$.
- If $\kappa>\frac{1}{8}$, let $s:=(192T)^{-1/3}$, which is at most $\frac{2-\kappa}{8}$ as $T\geq 3(2-\kappa)^{-3}$, and let U,G_3 be the pair in Lemma B.20. Then $H^2(U,G_3)\leq 24s^3=\frac{1}{8T}$, so Lemma D.1 with $\Delta=\frac{s^2}{2}$ and $T'=T$ gives $\mathrm{Reg}^{(\infty)}_{\{U,G_3\}}(T)\geq\frac{s^2T}{8}\geq\frac{T^{1/3}}{270}$, which is at least $\frac{\kappa^{-1/3}T^{1/3}}{540}$ as $\kappa>\frac{1}{8}$. □

D.3 Proofs of Theorem 4.2 and Corollary 4.3 (Local Minimax Lower Bound)

Throughout Appendices D.3 and D.4, we fix an F with $R^*:=R(p^*)>0$, and let the smallest and the largest Δ-optimal prices of F be $\ell_\Delta<u_\Delta$. Let $m(\Delta):=D_F([\ell_\Delta,u_\Delta))$ be the probability between them. As $N(\Delta)$ may not be monotone in Δ, we use its non-increasing envelope $\overline{N}(\Delta):=\sup_{\Delta'\geq\Delta}N(\Delta')$, so that $1\leq N\leq\overline{N}$. Note that $1\leq V_T(F)\leq\sqrt{2T}$, as $N(1)=1$ and $N(\Delta)\leq\frac{2}{\Delta}$. Furthermore, this envelope keeps V_T:

Lemma D.5 (Envelope Keeps V_T)**.** *If a value distribution G satisfies $N_G(\delta)\leq K\overline{N}(\delta)$ for all $\delta\in(0,1]$ with some $K\geq 1$, then $V_T(G)\leq KV_T(F)$.*

Proof. Fix $\delta\in(0,1]$. As N is integer-valued and at most $\frac{2}{\delta}$ on $[\delta,1]$, and $N=1$ on $[1,\infty)$ by Lemma C.17, the supremum in $\overline{N}(\delta)$ is attained at some $\delta'\in[\delta,1]$. So $\min\{T\delta,\overline{N}(\delta)\}\leq\min\{T\delta',N(\delta')\}\leq V_T(F)$, and hence $\min\{T\delta,N_G(\delta)\}\leq K\min\{T\delta,\overline{N}(\delta)\}\leq KV_T(F)$. Taking the supremum over δ proves the lemma. □

By the two-point method (Lemma D.1), it suffices to find two decoy instances that are close to F, are not much harder than F, and share no price with a small gap:

Definition 9 (Decoy Pair)**.** Fix a value distribution F and a horizon T. For $\Delta>0$, two value distributions G_+ and G_- form a Δ-*decoy pair* if:

1. (Close) $H^2(F,G_\pm)\leq\frac{1}{64T}$;
2. (Not harder) $N_{G_\pm}(\delta)\leq\frac{325}{(R^*)^3}\overline{N}(\delta)$ for all $\delta\in(0,1]$;

3. (Separating) every price has gap at least Δ under G_+ or under G_-.

As sketched in Figure 2, we construct decoy pairs in two steps: Lemma D.6 moves mass between ℓ_Δ and u_Δ, which brings little Hellinger distance when $m(\Delta) \approx \Delta \cdot N(\Delta)$ is large, and Lemma D.7 then sharpens the new revenue peaks, so that $G_\pm$ are not much harder than F. The proof below applies this at a scale Δ^* where $V_T(F)$ is nearly attained (in the special case where $V_T(F)$ is a constant, use $\Delta = \Theta(\frac{1}{T})$ instead).

Perturbing by a Density Ratio. Both moving-mass lemmas (Lemma D.6 here and Lemma D.10 in Appendix D.4) perturb F as follows. For a function h with $\int h\, dF = 0$ and $|\theta h| \le \frac{1}{2}$, let

$$\frac{dG}{dF} := 1 + \theta h, \quad \Psi(p) := \int h(v) \mathbb{1}[v \ge p]\, dF(v).$$

Then

$$R_G(p) = R(p) + p \cdot \theta\Psi(p) \quad \text{and} \quad \chi^2(G\|F) = \theta^2 \int h^2\, dF. \tag{22}$$

That is, G moves mass from where $h < 0$ to where $h > 0$, and $\theta\Psi(p)$ is the net mass moved to above p, which changes the demand at p by $\theta\Psi(p)$. We always take $h = \pm(-\frac{\mathbb{1}_I}{D_F(I)} + \frac{\mathbb{1}_J}{D_F(J)})$ for two disjoint intervals I, J, which moves mass θ between I and J and gives $\chi^2(G\|F) = \theta^2(\frac{1}{D_F(I)} + \frac{1}{D_F(J)})$.

Moving mass between ℓ_Δ and u_Δ gives two value distributions close to F that share no price with gap less than Δ, with a squared Hellinger distance to F inversely proportional to $m(\Delta)$:

Lemma D.6 (Moving Mass)**.** *Let $\Delta \le \frac{(R^*)^2}{10}$ and $m := m(\Delta) \ge \frac{20\Delta}{R^*}$. There are two value distributions G_+ and G_- with $\frac{dG_\pm}{dF} \in [\frac{1}{2}, \frac{3}{2}]$, $H^2(F, G_\pm) \le \frac{38\Delta^2}{(R^*)^3 m}$, and $|R_{G_\pm}(p) - R(p)| \le \frac{5\Delta}{R^*}$ for all p, such that every price has gap at least Δ under G_+ or under G_-.*

Proof. Write $\ell = \ell_\Delta$ and $u = u_\Delta$, so that $\ell \le p^* \le u$, and let $G_\pm$ be given by Eq. (22) with $\pm\theta$. As $m = D_F([\ell, p^*)) + D_F([p^*, u))$, one of the following two cases holds.

Case 1: $D_F([p^*, u)) \ge \frac{m}{2}$. This is the case in Figure 2 (a). We move mass θ between $I := [p^*, u)$ and $[u, 1]$: G_+ lifts u above p^*, and G_- lowers the upper part of I. Formally, let $\beta := D_F(u) \ge R(u) \ge 0.9R^*$, $h := -\frac{\mathbb{1}_I}{D_F(I)} + \frac{\mathbb{1}_{[u,1]}}{\beta}$, and $\theta := \frac{4\Delta}{p^*} \le \frac{4\Delta}{R^*}$. Then $\Psi = 0$ on $[0, p^*]$, $\Psi(p) = \frac{D_F([p^*, p))}{D_F(I)} \in [0, 1]$ on $[p^*, u]$, and $\Psi(p) = \frac{D_F(p)}{\beta} \in [0, 1]$ on $[u, 1]$. As $\frac{\theta}{D_F(I)} \le \frac{8\Delta}{R^* m} \le 0.4$ and $\frac{\theta}{\beta} \le 0.45$, we have $|\theta h| \le \frac{1}{2}$.

- Under G_+, $R_{G_+}(u) \ge R^* - \Delta + \theta u$. Every $p \le p^*$ has $R_{G_+}(p) = R(p) \le R^*$, so its gap is at least $\theta u - \Delta \ge 3\Delta$; every $p \in (p^*, u)$ with $\Psi(p) \le \frac{1}{2}$ has $R_{G_+}(p) \le R^* + \frac{\theta u}{2}$, so its gap is at least $\frac{\theta u}{2} - \Delta \ge \Delta$.
- Under G_-, $R_{G_-}(p^*) = R^*$. Every $p \in [p^*, u]$ with $\Psi(p) \ge \frac{1}{2}$ has $R_{G_-}(p) \le R^* - \frac{\theta p^*}{2} = R^* - 2\Delta$; every $p > u$ has $R_{G_-}(p) \le R(p) < R^* - \Delta$.
- $\chi^2(G_\pm\|F) = \theta^2(\frac{1}{D_F(I)} + \frac{1}{\beta}) \le \frac{16\Delta^2}{(R^*)^2}(\frac{2}{m} + \frac{1.12}{R^*}) \le \frac{50\Delta^2}{(R^*)^3 m}$.

Case 2: $D_F([\ell, p^*)) \ge \frac{m}{2}$. The difference from Case 1 is that the probability mass to be moved now lies below p^*, so we move mass across p^* rather than across u. We include the full construction for completeness. We move mass θ between $I := [\ell, p^*)$ and $[p^*, 1]$: G_+ lifts p^* above ℓ, and G_- lowers the upper part of I and all prices above p^*. Formally, let $\beta := D_F(p^*) \ge R^*$, $h := -\frac{\mathbb{1}_I}{D_F(I)} + \frac{\mathbb{1}_{[p^*,1]}}{\beta}$, and $\theta := \frac{5\Delta}{R^*}$. Then $\Psi = 0$ on $[0, \ell]$, $\Psi(p) = \frac{D_F([\ell, p))}{D_F(I)} \in [0, 1]$ on $[\ell, p^*]$, and $\Psi(p) = \frac{D_F(p)}{\beta} \in [0, 1]$ on $[p^*, 1]$. As $\frac{\theta}{D_F(I)} \le \frac{10\Delta}{R^* m} \le \frac{1}{2}$ and $\frac{\theta}{\beta} \le \frac{5\Delta}{(R^*)^2} \le \frac{1}{2}$, we have $|\theta h| \le \frac{1}{2}$. Also, $\ell \ge R(\ell) \ge 0.9R^*$.

- Under G_+, $R_{G_+}(p^*) = R^* + \theta p^*$. Every $p \le \ell$ has gap at least $\theta p^* \ge 5\Delta$; every $p \in (\ell, p^*)$ with $\Psi(p) \le \frac{1}{2}$ has gap at least $\frac{\theta p^*}{2} \ge 2.5\Delta$.
- Under G_-, $R_{G_-}(\ell) = R(\ell) \ge R^* - \Delta$. Every $p \in [\ell, p^*]$ with $\Psi(p) \ge \frac{1}{2}$ has $R_{G_-}(p) \le R^* - \frac{\theta\ell}{2} \le R^* - 2.25\Delta$; every $p > p^*$ has $R_{G_-}(p) = R(p)(1 - \frac{\theta}{\beta}) \le R(p)(1-\theta)$, which is at most $R^* - 2.5\Delta$ if $R(p) \ge \frac{R^*}{2}$ and less than $\frac{R^*}{2} \le R^* - 5\Delta$ otherwise. So these prices have gap at least 1.25Δ.
- $\chi^2(G_\pm \| F) \le \frac{25\Delta^2}{(R^*)^2}(\frac{2}{m} + \frac{1}{R^*}) \le \frac{75\Delta^2}{(R^*)^3 m}$.

In both cases, the prices with gap at least Δ under G_+ and under G_- together cover $[0,1]$. Finally, $H^2(F, G_\pm) \le \frac{1}{2}\chi^2(G_\pm \| F)$, and $|R_{G_\pm}(p) - R(p)| = \theta p |\Psi(p)| \le \theta \le \frac{5\Delta}{R^*}$ by Eq. (22). □

Moving mass may lift a plateau of the revenue, creating many near-optimal prices. We therefore need to "sharpen" the new revenue peak as follows:

Lemma D.7 (Sharpening). *Let $K \ge 1$, $\Delta \in (0, \frac{R^*}{8K}]$, and G be a value distribution with $|R_G(p) - R(p)| \le K\Delta$ for all p. Then there is G' with $\frac{dG'}{dG} \in [\frac{1}{2}, \frac{3}{2}]$ and $\chi^2(G' \| G) \le \frac{102K\Delta}{(R^*)^3 \overline{N}(\Delta)}$ such that*

1. *no price has a smaller gap under G' than under G;*
2. *$N_{G'}(\delta) \le \frac{65K}{(R^*)^2}\overline{N}(\delta)$ for all $\delta \in (0,1]$.*

Proof. Plan. As sketched in Section 4.3, the first step moves a small fraction of the mass on the Δ-optimal region below p^*_G to above p^*_G, and the second step scales up the Δ-optimal region above p^*_G, taking the extra mass from the prices above this region. Then $R_{G'}$ decreases linearly fast in both directions of p^*_G throughout the Δ-optimal region, making only those prices close to p^*_G near-optimal.

Setup. Let $\Gamma(p) := R_G(p^*_G) - R_G(p)$ be the gap of p under G, and $\ell' \le p^*_G \le u'$ be the smallest and the largest prices with $\Gamma \le \Delta$. As $R_G(p^*_G) \ge R^* - K\Delta \ge \frac{7}{8}R^*$, every p with $\Gamma(p) \le \Delta$ has $\min\{p, D_G(p)\} \ge R_G(p) \ge \frac{3}{4}R^*$. Let $c := \min\{\frac{R^*}{24}, \frac{1}{\overline{N}(\Delta)}\}$.

Construction. We construct G' in two steps. Each step keeps p^*_G a maximizer and does not decrease the gap of any price.

- *Left.* Let $\frac{dG_1}{dG} := 1 - c$ on $[\ell', p^*_G)$, $1 + c_1$ on $[p^*_G, 1]$, and 1 elsewhere, where $c_1 := \frac{cD_G([\ell', p^*_G))}{D_G(p^*_G)}$. Then $R_{G_1} = (1 + c_1)R_G$ on $[p^*_G, 1]$, $R_{G_1} = R_G$ on $[0, \ell')$, and $R_{G_1}(p) = R_G(p) + cpD_G([\ell', p))$ on $[\ell', p^*_G)$. So no gap decreases, and for $p \in [\ell', p^*_G)$, by Eq. (1) for G, the gap of p under G_1 is at least
$$\Gamma(p) + cp^*_G D_G([p, p^*_G)) \ge \Gamma(p) + c\left((p^*_G - p)D_G(p^*_G) - \Gamma(p)\right) \ge \frac{3}{4}cR^*(p^*_G - p).$$
- *Right.* Let $\frac{dG'}{dG_1} := 1 + c$ on $[p^*_G, u')$, $1 - c_2$ on $[u', 1]$, and 1 elsewhere, where $c_2 := \frac{cD_{G_1}([p^*_G, u'))}{D_{G_1}([u', 1])}$. Then $R_{G'} = R_{G_1}$ on $[0, p^*_G]$, $R_{G'} = (1 - c_2)R_{G_1}$ on $(u', 1]$, and $R_{G'}(p) = R_{G_1}(p) - cpD_{G_1}([p^*_G, p))$ on $[p^*_G, u']$. So no gap decreases, and for $p \in (p^*_G, u']$, by Eq. (1) for G_1, $p^*_G D_{G_1}([p^*_G, p)) \ge (p - p^*_G)D_{G_1}([p, 1]) \ge (p - p^*_G)D_G(p)$, so the gap of p under G' is at least $c(p - p^*_G)D_G(p)$, which is at least $\frac{3}{4}cR^*(p - p^*_G)$ if $\Gamma(p) \le \Delta$.

The first claim therefore follows.

Covering Numbers. For the second claim, let p have gap at most δ under G', so $\Gamma(p) \le \delta$.

- If $\delta \ge \Delta$, then $R(p) \ge R_G(p) - K\Delta \ge R^* - (1 + 2K)\delta$, so $p \in [\ell_{(1+2K)\delta}, u_{(1+2K)\delta}]$, an interval of length at most $(1 + 2K)\delta N((1 + 2K)\delta) \le 3K\delta\overline{N}(\delta)$. Hence $N_{G'}(\delta) \le 1 + 3K\overline{N}(\delta) \le 4K\overline{N}(\delta)$.

- If $\delta < \Delta$, then $p \in [\ell', u']$, so the two steps give $|p - p^*_G| \le \frac{4\delta}{3cR^*}$. Hence $N_{G'}(\delta) \le 1 + \frac{8}{3cR^*}$, which is at most $1 + \frac{8}{3R^*}\overline{N}(\Delta) \le \frac{11}{3R^*}\overline{N}(\delta)$ if $c = \frac{1}{\overline{N}(\Delta)}$ (as $\delta < \Delta$), and at most $1 + \frac{64}{(R^*)^2} \le \frac{65}{(R^*)^2}\overline{N}(\delta)$ otherwise.

Distance. Every p with $\Gamma(p) \le \Delta$ also has $R(p) \ge R_G(p) - K\Delta \ge R^* - (1+2K)\Delta$, so

$$u' - \ell' \le u_{(1+2K)\Delta} - \ell_{(1+2K)\Delta} \le (1+2K)\Delta N((1+2K)\Delta) \le 3K\Delta\overline{N}(\Delta).$$

By Eq. (1) for G, $\ell' D_G([\ell', u')) \le u' - \ell' + \Delta$, so $m' := D_G([\ell', u')) \le \frac{16K}{3R^*}\Delta\overline{N}(\Delta)$ as $\overline{N}(\Delta) \ge 1$. Finally, as $D_G(p^*_G) \ge \frac{7}{8}R^*$ and $D_G(u') \ge \frac{3}{4}R^*$, we have $c_1 \le \frac{8c}{7R^*}m'$ and $c_2 = \frac{cD_G([p^*_G, u'))}{D_G(u')} \le \frac{4c}{3R^*}m'$. As $c \le \frac{R^*}{24}$ and $m' \le 1$, all of c, c_1, c_2 are at most $\frac{1}{12}$, so $\frac{dG'}{dG} \in [\frac{1}{2}, \frac{3}{2}]$. Moreover, $|\frac{dG'}{dG} - 1|$ is c on $[\ell', p^*_G)$, at most $c + \frac{13}{12}c_1 \le \frac{3c}{R^*}$ on $[p^*_G, u')$, at most $c_1 + c_2 \le \frac{3c}{R^*}m'$ on $[u', 1]$, and 0 elsewhere. Hence

$$\chi^2(G'\|G) \le c^2 m' + \frac{9c^2}{(R^*)^2}m' + \frac{9c^2}{(R^*)^2}m'^2 \le \frac{19c^2m'}{(R^*)^2} \le \frac{102Kc^2\Delta\overline{N}(\Delta)}{(R^*)^3} \le \frac{102K\Delta}{(R^*)^3\overline{N}(\Delta)}. \qquad \square$$

We now prove Theorem 4.2 from Lemmas D.6 and D.7.

Theorem D.8 (Restatement of Theorem 4.2; Local Minimax Lower Bound)**.** *Let F be such that $R_F(p^*_F) > 0$ and $T \ge 2$, and let $c_1 := \frac{R_F(p^*_F)^{12}}{2^{45}}$, $c_2 := \frac{325}{R_F(p^*_F)^3}$, and $c_3 := \frac{2^{20}}{R_F(p^*_F)^6}$. There are value distributions G_+, G_- such that:*

(a) *Any algorithm, even with infinite menus, suffers regret at least $c_1 V_T(F)$ on G_+ or on G_-;*
(b) *$V_T(G_\pm) \le c_2 V_T(F)$; and*
(c) *$H^2(F, G_\pm) \le \frac{c_3}{V_T(F)^2}$ where $H^2(F, G) := \frac{1}{2}\int(\sqrt{dF} - \sqrt{dG})^2$ is the squared Hellinger distance.*

Proof. Write $V := V_T(F)$ and $c_0 := \frac{(R^*)^6}{2^{22}}$. We first combine Lemmas D.6 and D.7 into a decoy pair at every scale Δ where there is enough probability mass to move, namely, any $\Delta \le \frac{(R^*)^2}{40}$ such that $m(\Delta) \ge \frac{20\Delta}{R^*}$. Let $G_\pm$ be given by Lemma D.6, and let $G'_\pm$ be the results of Lemma D.7 with $K = \frac{5}{R^*}$ applied to them, which is applicable as $|R_{G_\pm} - R| \le \frac{5\Delta}{R^*}$. Then every price has gap at least Δ under G'_+ or under G'_-, as this holds for $G_\pm$ (Lemma D.6) and sharpening decreases no gap (Lemma D.7 1). Also $N_{G'_\pm} \le \frac{65K}{(R^*)^2}\overline{N} = \frac{325}{(R^*)^3}\overline{N}$ (Lemma D.7 2). Moreover, by the triangle inequality for H and $H^2 \le \frac{1}{2}\chi^2$, and then by the χ^2 bounds in Lemmas D.6 and D.7, we have

$$H^2(F, G'_\pm) \le 2H^2(F, G_\pm) + \chi^2(G'_\pm\|G_\pm) \le \frac{76\Delta^2}{(R^*)^3 m(\Delta)} + \frac{510\Delta}{(R^*)^4\overline{N}(\Delta)}. \tag{23}$$

So G'_+ and G'_- indeed form a Δ-decoy pair (Definition 9) for every horizon T' for which the right-hand side is at most $\frac{1}{64T'}$. It only remains to pick an appropriate Δ. Consider a $\Delta^* \in (0, 1]$ with $\min\{T\Delta^*, N(\Delta^*)\} \ge V/2$. Except for the case where V is a constant (Case 2 below), we apply the above construction to this Δ^*.

Case 1: Large $N(\Delta^*) \ge \frac{1}{c_0}$. We use a Δ^*-decoy pair for a shorter horizon $T' \le T$ (i.e., indistinguishable during the first T' rounds, which already forces regret $\Omega(\Delta^* T')$). Let

$$T' := \min\left\{T, \left\lfloor \frac{c_0 N(\Delta^*)}{\Delta^*} \right\rfloor\right\}, \qquad \text{where } \frac{c_0 N(\Delta^*)}{\Delta^*} \ge c_0 N(\Delta^*) \ge 1.$$

- *The pair.* We move mass and then sharpen at Δ^*, where $\Delta^* \le 2c_0 \le \frac{(R^*)^2}{40}$ as $N(\Delta^*) \le \frac{2}{\Delta^*}$. There is enough probability mass to move, as far apart near-optimal prices have large probability mass in between (as discussed in Section 4.3): Writing $\ell := \ell_{\Delta^*}$ and $u := u_{\Delta^*}$,

$$m(\Delta^*) \overset{(a)}{\ge} \ell\, m(\Delta^*) \overset{(b)}{=} (u-\ell)D_F(u) - (R(u)-R(\ell)) \overset{(c)}{\ge} \frac{R^*}{2}\left(N(\Delta^*)-1\right)\Delta^* - \Delta^* \overset{(d)}{\ge} \frac{R^*}{4}\Delta^* N(\Delta^*),$$

  where (a) uses $\ell \le 1$, (b) uses Eq. (1), (c) uses $D_F(u) \ge R(u) \ge R^* - \Delta^* \ge \frac{R^*}{2}$, $u - \ell \ge (N(\Delta^*)-1)\Delta^*$ (Lemma C.17), and $R(u) - R(\ell) \le \Delta^*$, and (d) uses $N(\Delta^*) \ge \frac{1}{c_0} \ge 2 + \frac{4}{R^*}$. In particular, $m(\Delta^*) \ge \frac{20\Delta^*}{R^*}$. The pair is close: As $\overline{N} \ge N$ and $\frac{\Delta^*}{N(\Delta^*)} \le \frac{c_0}{T'}$ by the choice of T', the right-hand side of Eq. (23) at Δ^* is at most

$$\frac{304\Delta^*}{(R^*)^4 N(\Delta^*)} + \frac{510\Delta^*}{(R^*)^4 N(\Delta^*)} \le \frac{2^{10}c_0}{(R^*)^4 T'} \le \frac{1}{64T'}.$$

  So moving mass and then sharpening gives a Δ^*-decoy pair G_+, G_- for the horizon T'.
- *The regret.* As $H^2(F, G_\pm) \le \frac{1}{64T'}$, the first $T'-1$ values have total variation distance at most $\sqrt{2T'H^2(F,G_\pm)} \le \frac{1}{4}$ between F and $G_\pm$, and hence at most $\frac{1}{2}$ between G_+ and G_-. So Lemma D.1 gives regret at least $\frac{\Delta^* T'}{4}$ on G_+ or on G_-, where, as $\lfloor x \rfloor \ge \frac{x}{2}$ for $x \ge 1$ and $c_0 \le 1$,

$$\Delta^* T' \ge \min\left\{T\Delta^*, \frac{c_0 N(\Delta^*)}{2}\right\} \ge \frac{c_0}{2}\min\{T\Delta^*, N(\Delta^*)\} \ge \frac{c_0 V}{4}.$$

- *The distance.* As $N(\Delta^*) \ge \frac{V}{2}$ and $N(\Delta^*) \le \frac{2}{\Delta^*}$, we have $\frac{c_0 N(\Delta^*)}{\Delta^*} \ge \frac{c_0}{2}N(\Delta^*)^2 \ge \frac{c_0 V^2}{8}$, so $T' \ge \min\{T, \frac{c_0 V^2}{16}\}$, and hence $H^2(F, G_\pm) \le \frac{1}{64T'} \le \max\{\frac{1}{64T}, \frac{1}{4c_0 V^2}\}$.

Case 2: Constant V, namely $N(\Delta^*) < \frac{1}{c_0}$ (which implies $V \le 2N(\Delta^*) < \frac{2}{c_0}$ is at most a constant depending on R^*). We then take a tiny scale $\Delta := \frac{4c_0}{T} = \frac{(R^*)^6}{2^{20}T}$ to force regret $\Omega(1)$. We construct the pair depending on whether F has enough probability mass between its Δ-optimal prices to move:

- *Enough probability mass to move, i.e., $m(\Delta) \ge \frac{20\Delta}{R^*}$.* We move mass and then sharpen at Δ, which is possible as $\Delta \le \frac{(R^*)^2}{40}$. As $\overline{N} \ge 1$, the right-hand side of Eq. (23) is at most $\frac{4\Delta}{(R^*)^2} + \frac{510\Delta}{(R^*)^4} \le \frac{2^{10}\Delta}{(R^*)^4} \le \frac{1}{64T}$, so this gives a Δ-decoy pair.
- *Little probability mass to move, i.e., $m(\Delta) < \frac{20\Delta}{R^*}$.* Then we cannot move mass, but this also implies that under F, lowering the price from p^* loses revenue quickly. So we keep $G_+ := F$, and let G_- add a small mass just below p^*, which makes a price just below p^* better than all prices above the added mass. Formally, let $\eta := \frac{128\Delta}{(R^*)^3} = \frac{(R^*)^3}{2^{13}T}$, $x := \frac{\eta p^*}{2}$, $y := p^* - x$, and $\xi \in (0, \frac{x}{2})$, and let

$$G := (1-\eta)F + \eta\,\mathrm{Unif}[y, y+\xi],$$

  so that $H^2(F, G) \le \eta$ by the convexity of H^2, and $|R_G - R| \le \eta$. Then:
  (a) Under G, every price $p \ge y + \xi$ has gap at least Δ, as the added buyers buy at y but not at p. Indeed, $R(y) \ge R^* - x$ by Eq. (1), so $R_G(y) \ge (1-\eta)(R^* - x) + \eta(p^* - x) = (1-\eta)R^* + \frac{\eta p^*}{2}$, whereas $R_G(p) = (1-\eta)R(p) \le (1-\eta)R^*$, and $\frac{\eta p^*}{2} \ge \frac{\eta R^*}{2} \ge \Delta$.
  (b) Under F, no price $p < y + \xi$ is Δ-optimal, as p is more than $\frac{x}{2}$ below p^* and F has little probability mass in between. Indeed, otherwise $\ell_\Delta \le p < p^* \le u_\Delta$, so by Eq. (1), the gap of p under F would be

$$(p^* - p)D_F(p^*) - pD_F([p, p^*)) > \frac{x}{2}R^* - m(\Delta) > \frac{32\Delta}{R^*} - \frac{20\Delta}{R^*} > \Delta,$$

  where we used $D_F(p^*) \ge R^*$, $pD_F([p, p^*)) \le m(\Delta)$, and $\frac{x}{2}R^* \ge \frac{\eta(R^*)^2}{4} = \frac{32\Delta}{R^*}$.

Adding mass may also lift a plateau of the revenue, so we sharpen G_- by applying Lemma D.7 to G with η in place of Δ and $K = 1$. Then every price has gap at least Δ under $G_+ = F$ by (b) or under G_- by (a), as sharpening decreases no gap. Moreover, $N_{G_+} = N \le \overline{N}$, $N_{G_-} \le \frac{65}{(R^*)^2}\overline{N} \le \frac{325}{(R^*)^3}\overline{N}$, $H^2(F, G_+) = 0$, and $H^2(F, G_-) \le 2H^2(F, G) + \chi^2(G_-\|G) \le 2\eta + \frac{102\eta}{(R^*)^3} \le \frac{1}{64T}$. So G_+, G_- indeed form a Δ-decoy pair.

Therefore, in both cases, the regret on G_+ or on G_- is at least $\min\{\frac{c_0}{16}, \frac{c_0^2}{2}\}V = \frac{(R^*)^{12}}{2^{45}}V$, which proves claim (a). By Lemma D.5, every decoy pair has $V_T(G_\pm) \le \frac{325}{(R^*)^3}V$, which gives claim (b). As $V^2 \le 2T$, we have $\max\{\frac{1}{64T}, \frac{1}{4c_0V^2}\} = \frac{1}{4c_0V^2} = \frac{2^{20}}{(R^*)^6V^2}$, which proves claim (c). □

Together with Theorem 3.5 and Lemma C.20, this gives the near instance-optimality of Algorithm 2:

Corollary D.9 (Restatement of Corollary 4.3; Near Instance-Optimality of Algorithm 2)**.** *Under the conditions of Theorem 4.2,*

1. *Algorithm 2 has regret $\mathcal{O}(V_T(F)\log^5 T)$ on F and $\mathcal{O}(R_F(p_F^*)^{-3}V_T(F)\log^5 T)$ on G_+ and on G_-.*
2. *Any algorithm, even with infinite menus, suffers $\Omega(R_F(p_F^*)^{12}V_T(F))$ regret on G_+ or on G_-.*
3. *If F does not have interval dimension 0, then $H^2(F, G_\pm) \to 0$ as $T \to \infty$ (see Lemma C.20).*

Proof. Let $G_\pm$ be given by Theorem 4.2. For the first claim, Theorem 3.5 gives regret $\mathcal{O}(V_T(F)\log^5 T)$ on F and $\mathcal{O}(V_T(G_\pm)\log^5 T)$ on $G_\pm$, where $V_T(G_\pm) \le \frac{325}{R_F(p_F^*)^3}V_T(F)$ by Theorem 4.2(b). The second claim is Theorem 4.2(a). For the third claim, if F does not have interval dimension 0, then $V_T(F) \to \infty$ as $T \to \infty$ by Lemma C.20, so Theorem 4.2(c) gives $H^2(F, G_\pm) \to 0$. □

D.4 Proof of Theorem 4.4 (Near Instance-Optimality via Interval Dimensions)

In the special case where F has an interval dimension, we keep F itself as one of the two instances. We make the proof sketch in Section 4.3 precise. Besides explicit constants, the only new case is when the wide side is the lower one, $[\ell_\Delta, \ell_{\Delta'})$, where we move mass down instead of up (symmetric to the case we sketched in Section 4.3; Case 2 in the proof below).

Lemma D.10 (One-Sided Moving Mass)**.** *Let $0 < \Delta' \le \Delta \le \frac{(R^*)^2}{16}$ be such that $\lambda := \max\{u_\Delta - u_{\Delta'}, \ell_{\Delta'} - \ell_\Delta\} \ge \frac{32\Delta}{(R^*)^2}$. Then there is a value distribution G with $\chi^2(G\|F) \le \frac{96\Delta^2}{(R^*)^3\lambda}$ and $|R_G(p) - R(p)| \le \frac{4\Delta}{R^*}$ for all p, such that every price has gap more than Δ' under F or at least Δ' under G.*

Proof. As in Lemma D.6, G moves mass θ via Eq. (22), but only on the wide side: In Case 1, from $[b, u)$ up to $[u, 1]$, which lifts u above every price up to b; in Case 2, from $[a, 1]$ down to $[\ell, a)$, which lowers every price from a on below ℓ. Write $\ell = \ell_\Delta$, $u = u_\Delta$, $a = \ell_{\Delta'}$, and $b = u_{\Delta'}$, so $\ell \le a \le p^* \le b \le u$, and every $p \notin [a, b]$ has gap more than Δ' under F. As $\Delta \le \frac{R^*}{2}$, each $p \in \{\ell, a, b, u\}$ has $R(p) = pD_F(p) \ge R^* - \Delta \ge \frac{R^*}{2}$, so $p \ge \frac{R^*}{2}$ and $D_F(p) \ge \frac{R^*}{2}$.

Case 1: $u - b = \lambda$. Let $I := [b, u)$, $\beta := D_F(u)$, $h := -\frac{\mathbb{1}_I}{D_F(I)} + \frac{\mathbb{1}_{[u,1]}}{\beta}$ (where $\mathbb{1}_I(v) := \mathbb{1}[v \in I]$), and $\theta := \frac{\Delta+\Delta'}{u} \le \frac{4\Delta}{R^*}$. Then $\Psi = 0$ on $[0, b]$, $\Psi(p) = \frac{D_F([b,p))}{D_F(I)} \in [0, 1]$ on $[b, u]$, and $\Psi(p) = \frac{D_F(p)}{\beta} \in [0, 1]$ on $[u, 1]$. So every $p \le b$ has $R_G(p) = R(p) \le R^*$, whereas $R_G(u) = R(u) + \theta u \ge R^* + \Delta'$, and hence p has gap at least Δ' under G. Applying Eq. (1) to $b < u$ gives $bD_F(I) = (u-b)\beta - (R(u) - R(b)) \ge \frac{R^*}{2}\lambda - \Delta$, so $D_F(I) \ge \frac{R^*\lambda}{4}$ as $b \le 1$ and $\lambda \ge \frac{4\Delta}{R^*}$.

Case 2: $a - \ell = \lambda$. Let $I := [\ell, a)$, $\beta := D_F(a)$, $h := \frac{\mathbb{1}_I}{D_F(I)} - \frac{\mathbb{1}_{[a,1]}}{\beta}$, and $\theta := \frac{\Delta+\Delta'}{R^*} \le \frac{2\Delta}{R^*}$. Then $\Psi = 0$ on $[0, \ell]$, $\Psi(p) = -\frac{D_F([\ell,p))}{D_F(I)} \in [-1, 0]$ on $[\ell, a]$, and $\Psi(p) = -\frac{D_F(p)}{\beta} \in [-1, 0]$ on $[a, 1]$. So $R_G(\ell) = R(\ell) \ge R^* - \Delta$, whereas every $p \ge a$ has $R_G(p) = R(p)(1 - \frac{\theta}{\beta}) \le R^*(1 - \theta) = R^* - \Delta - \Delta'$ as $\beta \le 1$, and hence p has gap at least Δ' under G. Applying Eq. (1) to $\ell < a$ gives $\ell D_F(I) = (a - \ell)\beta - (R(a) - R(\ell)) \ge \frac{R^*}{2}\lambda - \Delta$, so $D_F(I) \ge \frac{R^*\lambda}{4}$.

In both cases, $\theta \le \frac{4\Delta}{R^*}$, $\beta \ge \frac{R^*}{2}$, and $D_F(I) \ge \frac{R^*\lambda}{4}$, so $\frac{\theta}{D_F(I)} \le \frac{16\Delta}{(R^*)^2\lambda} \le \frac{1}{2}$ and $\frac{\theta}{\beta} \le \frac{8\Delta}{(R^*)^2} \le \frac{1}{2}$, i.e., $|\theta h| \le \frac{1}{2}$. Moreover, $|R_G(p) - R(p)| = \theta p|\Psi(p)| \le \theta \le \frac{4\Delta}{R^*}$, and as $\lambda \le 1$,

$$\chi^2(G\|F) = \theta^2\left(\frac{1}{D_F(I)} + \frac{1}{\beta}\right) \le \frac{16\Delta^2}{(R^*)^2}\left(\frac{4}{R^*\lambda} + \frac{2}{R^*}\right) \le \frac{96\Delta^2}{(R^*)^3\lambda}. \qquad \square$$

Theorem D.11 (Restatement of Theorem 4.4; Near Instance-Optimality via Interval Dimensions)**.** *If F has interval dimension $d \in (0, 1)$ with multiplier C, there is a constant $c > 0$ only depending on $R_F(p_F^*), C$, and d, such that: For any $T \ge \frac{1}{c}$ satisfying $N_F(\Delta) \ge \frac{\Delta^{-d}}{C}$ with $\Delta = T^{-1/(d+1)}$, there is a value distribution G that satisfies:*

(a) *Any algorithm, even with infinite menus, suffers regret at least $cT^{d/(d+1)}$ on F or on G;*
(b) *G also has interval dimension d with multiplier C/c; and*
(c) *$H^2(F, G) \le \frac{1}{cT}$.*

Proof. First, we claim that $R^* > 0$. Indeed, as $R \ge 0$, every price is δ-optimal for $\delta \ge R^*$, so $\lceil \frac{1}{\delta} \rceil = N(\delta) \le C\delta^{-d}$ for all $\delta \in (0, 1]$ with $\delta \ge R^*$, which forces $R^* \ge C^{-1/(1-d)} > 0$ as $d < 1$. Pick

$$c := \min\left\{\frac{(R^*)^4}{2^{20}C(4C^2)^{1/(1-d)}}, \left(\frac{(R^*)^2}{128C}\right)^{(d+1)/d}\right\},$$

where the first term makes the regret and the distance large and small enough, respectively, and the second one makes $N(\Delta)$ large. Let $T \ge \frac{1}{c}$ and $\Delta := T^{-1/(d+1)}$ with $N(\Delta) \ge \frac{\Delta^{-d}}{C}$, and let $\Delta' := (4C^2)^{-1/(1-d)}\Delta \le \Delta$ be the smaller scale at which we separate F from the decoy instance. As $C \ge N(1) = 1$, we have $T \ge \frac{2^{20}}{(R^*)^4} \ge (\frac{32}{(R^*)^2})^{d+1}$, so $\Delta \le \frac{(R^*)^2}{32}$, and $N(\Delta) \ge \frac{T^{d/(d+1)}}{C} \ge \frac{128}{(R^*)^2}$.

Step 1: Δ'-Optimal Prices Are Narrow. As F has interval dimension d with multiplier C,

$$u_{\Delta'} - \ell_{\Delta'} \le \Delta' N(\Delta') \le C\Delta'^{1-d} = \frac{\Delta^{1-d}}{4C} \le \frac{1}{4}\Delta N(\Delta).$$

As $u_\Delta - \ell_\Delta \ge \Delta(N(\Delta) - 1)$ and $N(\Delta) \ge 4$, the width $\lambda := \max\{u_\Delta - u_{\Delta'}, \ell_{\Delta'} - \ell_\Delta\}$ of the wider side satisfies $\lambda \ge \frac{1}{2}(\Delta(N(\Delta) - 1) - \frac{1}{4}\Delta N(\Delta)) \ge \frac{1}{4}\Delta N(\Delta)$, which is at least $\frac{32\Delta}{(R^*)^2}$ and at least $\frac{\Delta^{1-d}}{4C}$.

Step 2: Decoy Instances. Let G_0 be the value distribution in Lemma D.10, and let G be the sharpened version from Lemma D.7 with $K = \frac{4}{R^*}$ applied to G_0, which is valid as $\Delta \le \frac{(R^*)^2}{32} = \frac{R^*}{8K}$. Then every price has gap at least Δ' under F or under G, and $N_G \le \frac{260}{(R^*)^3}\overline{N}$ on $(0, 1]$. Moreover, as $\overline{N}(\Delta) \ge N(\Delta) \ge \frac{\lambda}{\Delta}$,

$$H^2(F, G) \le 2H^2(F, G_0) + \chi^2(G\|G_0) \le \frac{96\Delta^2}{(R^*)^3\lambda} + \frac{408\Delta}{(R^*)^4\overline{N}(\Delta)} \le \frac{504\Delta^2}{(R^*)^4\lambda}.$$

Step 3: Horizon. Let $c_1 := \frac{(R^*)^4}{32256}$ and $T' := \min\{T, \lfloor \frac{c_1\lambda}{\Delta^2} \rfloor\}$, so that $H^2(F,G) \le \frac{\Delta^2}{64c_1\lambda} \le \frac{1}{64T'}$. As $\frac{\lambda}{\Delta^2} \ge \frac{N(\Delta)}{4\Delta} \ge \frac{\Delta^{-(d+1)}}{4C} = \frac{T}{4C}$, and $\frac{c_1 T}{4C} \ge \frac{2^{20}(4C^2)^{1/(1-d)}}{129024} \ge 1$ as $T \ge \frac{1}{c}$, we have $T' \ge \frac{c_1 T}{8C}$, using $\lfloor x \rfloor \ge \frac{x}{2}$ for $x \ge 1$.

Step 4: Regret. The first $T'-1$ values have total variation distance at most $\sqrt{2T'H^2(F,G)} \le \frac{1}{4}$ between F and G. So Lemma D.1 with Δ' in place of Δ shows that any algorithm suffers regret at least $\frac{\Delta' T'}{4}$ on F or on G, where

$$\frac{\Delta' T'}{4} \ge \frac{c_1 \Delta' T}{32C} = \frac{(R^*)^4}{1032192\, C(4C^2)^{1/(1-d)}} T^{d/(d+1)} \ge cT^{d/(d+1)}.$$

Step 5: Distance and Interval Dimension. We have $H^2(F,G) \le \frac{1}{64T'} \le \frac{C}{8c_1 T} = \frac{4032C}{(R^*)^4 T} \le \frac{1}{cT}$. Finally, $\overline{N}(\delta) \le C\delta^{-d}$ for all $\delta \in (0,1]$, as δ^{-d} is non-increasing and $N(\delta) = 1 \le C$ for $\delta \ge 1$. So $N_G(\delta) \le \frac{260C}{(R^*)^3}\delta^{-d} \le \frac{C}{c}\delta^{-d}$ for all $\delta \in (0,1]$, i.e., G has interval dimension d with multiplier $\frac{C}{c}$. Hence G satisfies all three claims. □

D.5 Proof of Theorem 4.5 (Near Instance-Optimality for MHR Distributions)

As sketched in Section 4.3, the decoy instance G multiplies the demand of F by e^{ϕ}, i.e., adds a ramp ϕ to the log-demand, which raises the revenue of $p^* + s$ above that of p^* while keeping the hazard rate non-decreasing.

Theorem D.12 (Restatement of Theorem 4.5; Near Instance-Optimality for MHR Distributions)**.** *Let F be MHR. If on an open interval containing p^*_F, its hazard rate function $h(v) = \frac{f(v)}{D_F(v)}$ is differentiable with h' bounded by some $[c, C]$ with $0 < c < C$, then for every $T > 0$, there is an MHR distribution G with $H^2(F,G) \le \frac{1}{64T}$, such that any algorithm, even with infinite menus, suffers regret $\mathrm{Reg}^{(\infty)}_{\{F,G\}}(T) = \Omega(T^{1/3})$.*

Proof. Write $D(p) := \Pr_{v\sim F}\{v \ge p\}$ and $R^* := R(p^*)$, and take a closed interval $I \subseteq (0,1)$ with p^* in its interior and $0 < c \le C$ such that $c \le h' \le C$ on I. As $R'(p) = D(p)(1 - ph(p))$ is continuous there and p^* maximizes R, we have $p^* h(p^*) = 1$. Let $C_1 := \frac{1}{p^*} + 2C$.

For $s > 0$, let $L := \frac{2C_1 s}{cR^*}$, $a := p^* - L$, $b := p^* + s$, and $\phi(p) := \frac{c}{2}(\min\{p,b\} - a)_+^2$, which is 0 below a, grows quadratically on $[a,b]$, and is constant above b. Let G be the value distribution with $D_G(p) = D(p)e^{\phi(p)}$, so that $R_G = Re^{\phi}$. Let $s_0 > 0$ be such that every $s \le s_0$ satisfies

$$\text{(i) } [a,b] \subseteq I, \qquad \text{(ii) } L \le \frac{1}{2Cp^*}, \qquad \text{(iii) } L + s \le 1, \qquad \text{(iv) } \frac{c}{2}(L+s)^2 \le \frac{1}{2}, \qquad \text{(v) } s^2 \le 2e^2 R^*,$$

where (i) is used for the MHR property, (ii)–(iv) for the gaps and the distance, and (v) for the gaps under F. Fix any $s \le s_0$.

- *G is MHR.* Its hazard rate is $h - \phi'$, which equals h outside $[a,b]$, is non-decreasing on $[a,b]$ as $\phi'' = c \le h'$, jumps up at b, and is at least $h(a) \ge 0$.
- *Gaps under F.* By Lemma C.8, the gap of $p^* + u$ under F is at least $R^*(1 - (1+x)e^{-x})$ with $x := \frac{u}{p^*}$. Since $1 - (1+x)e^{-x} = \int_0^x ye^{-y}\,dy$ is at least $\frac{x^2}{2e}$ for $x \le 1$ and at least $1 - \frac{2}{e} > \frac{1}{4}$ for $x \ge 1$, and $\frac{R^*}{(p^*)^2} = \frac{D(p^*)}{p^*} \ge e^{-1}$, this gap is at least $\min\{\frac{u^2}{2e^2}, \frac{R^*}{4}\}$. So by (v), every $p \ge p^* + \frac{s}{2}$ has gap at least $\frac{s^2}{8e^2}$ under F.

- *Gaps under G.* Every $p < p^* + \frac{s}{2}$ has gap at least $\frac{s^2}{4}$ under G. Indeed, for $x \in [0, s]$, we have $(p^*+x)h(p^*+x)-1 \le (p^*+x)(h(p^*)+Cx)-1 \le C_1 x$, so $R'(p^*+x) \ge -C_1 x$ and $R(b) \ge R^* - \frac{C_1 s^2}{2}$. As ϕ is non-decreasing and $\phi(b) - \phi(p^* + \frac{s}{2}) \ge \frac{cLs}{2} = \frac{C_1 s^2}{R^*}$, where $\frac{C_1 s^2}{R^*} \le \phi(b) \le \frac{1}{2}$, we have

$$\frac{R_G(b) - R_G(p)}{e^{\phi(p^*+s/2)}} \ge \left(R^* - \frac{C_1 s^2}{2}\right)\left(1 + \frac{C_1 s^2}{R^*}\right) - R^* = \frac{C_1 s^2}{2}\left(1 - \frac{C_1 s^2}{R^*}\right) \ge \frac{C_1 s^2}{4} \ge \frac{s^2}{4}.$$

- *Distance.* $\chi^2(G\|F) \le \frac{702 C_1^4 s^3}{c(R^*)^3}$. Indeed, $\frac{dG}{dF}$ is 1 below a, $e^{\phi(b)}$ above b, and $(1 - \frac{\phi'}{h})e^{\phi}$ on $[a, b]$, where $h \ge h(p^*) - CL \ge \frac{1}{2p^*}$ and $f \le h \le C_1$. As $\phi \le \frac{1}{2}$ and $\phi' \le c(L + s)$, $|\frac{dG}{dF} - 1|$ is at most $2\phi + \frac{2\phi'}{h} \le 5c(L + s)$ on $[a, b]$, and at most $2\phi(b) \le c(L + s)^2$ above b. So $\chi^2(G\|F) \le 25 C_1 c^2 (L + s)^3 + c^2(L + s)^4 \le 26 C_1 c^2 (L + s)^3$, where $L + s \le \frac{3C_1 s}{cR^*}$ as $C_1 \ge 2C \ge 2c$.

Therefore, every price has gap at least $\frac{s^2}{8e^2}$ under F or under G. Take $s := \min\{s_0, (\frac{c(R^*)^3}{22464 C_1^4 T})^{1/3}\}$, so that $H^2(F, G) \le \frac{1}{2}\chi^2(G\|F) \le \frac{1}{64T}$. Then the first $T - 1$ values have total variation distance at most $\sqrt{2TH^2(F, G)} \le \frac{1}{2}$ between F and G, and Lemma D.1 with $\Delta = \frac{s^2}{8e^2}$ shows that $\mathrm{Reg}^{(\infty)}_{\{F,G\}}(T) \ge \frac{s^2 T}{32e^2}$. As $s = \Omega(T^{-1/3})$ with constants only depending on F, this gives the claimed $\Omega(T^{1/3})$ bound. □

D.6 Proof of Theorem 4.6 (Anytime Lower Bounds)

Definition 10 (Plateau)**.** For a revenue function R with maximum $L := \max_{p\in[0,1]} R(p)$, a closed interval $P = [x - w, x + w]$ with $w > 0$ is the *plateau* of R, with center x, half-width w, and height L, if $R = L$ on P and $R < L$ outside P. That is, the set of maximizers of R is the interval P.

We make the sketch in Section 4.4 precise. Starting from a revenue function R_0 with a plateau P_0 around $\frac{1}{2}$, each scale $k = 1, 2, \ldots$ raises a part of the current plateau P_{k-1} by Δ_k, and the top of this part is the next plateau P_k. Odd scales are *tilts*, which raise a random half of P_{k-1} and fool infinite menus; even scales are *needles*, which raise a random short piece of P_{k-1} and fool posted prices. The heights shrink super-exponentially, so that the finer scales are invisible at the horizon of each scale.

Parameters. Fix $d \in (0, 1)$. Let $q := \max\{5, \frac{2}{1-d}\}$, let $\Delta_1 > 0$ be small enough, and let

$$\Delta_{k+1} := \Delta_k^q \quad \text{and} \quad w_k := \Delta_{k+1}^{1-d} \qquad \text{for } k \ge 0.$$

Here Δ_k is the height of scale k, and w_k is the half-width of the plateau P_k (so P_k has width $2w_k$). The exponent $1 - d$ yields interval dimension d, and the large exponent q makes the finer scales negligible. As Δ_1 is small, we have $w_0 \le 0.1$, and for all $k \ge 1$,

$$w_k \le \min\left\{\frac{w_{k-1}}{4}, \Delta_k\right\}, \qquad \frac{10\Delta_k}{w_{k-1}} \le \frac{1}{2}, \qquad \text{and} \qquad \Delta_{>k} := \sum_{m>k} \Delta_m \le 1.01\Delta_{k+1}. \tag{24}$$

In words, the plateaus shrink by a factor of at least 4, and each half-width w_k is at most the height Δ_k, which the construction below uses to fit the bumps into P_{k-1}; each height is small compared with the previous plateau, which keeps the tilts gentle (Lemma D.16); and all finer scales together are about as high as the next one (Lemma D.17).

Construction. Let $x_0 := \frac{1}{2}$, $L_0 := \frac{1}{4}$, and $R_0(p) := \min\{p, \frac{1}{4} - \frac{1}{4}(|p - x_0| - w_0)_+\}$, so that $R_0 = L_0$ on the plateau $P_0 := [x_0 - w_0, x_0 + w_0]$ and $R_0 < L_0$ elsewhere. For $k \ge 1$, given R_{k-1} with $R_{k-1} = L_{k-1}$ on the plateau $P_{k-1} := [x_{k-1} - w_{k-1}, x_{k-1} + w_{k-1}]$ and $R_{k-1} < L_{k-1}$ elsewhere, we draw an independent ξ_k, and let $R_k := R_{k-1} + \Delta_k \phi_k$ and $L_k := L_{k-1} + \Delta_k$, where the bump ϕ_k is as follows; scale k is a tilt scale if k is odd and a needle scale otherwise.

- *Tilt.* ξ_k is uniform over $\{-1, 1\}$, and ϕ_k is 0 outside the left (if $\xi_k = -1$) or the right (if $\xi_k = 1$) half of P_{k-1}, rises linearly from both ends of this half, and equals 1 on its middle part. Formally, $x_k := x_{k-1} + \frac{\xi_k w_{k-1}}{2}$ and $\phi_k(p) := \min\{1, \frac{(w_{k-1}/2 - |p - x_k|)_+}{w_{k-1}/2 - w_k}\}$. The slopes of $\Delta_k \phi_k$ are $\pm\sigma_k$ with $\sigma_k := \frac{\Delta_k}{w_{k-1}/2 - w_k} \in [\frac{2\Delta_k}{w_{k-1}}, \frac{4\Delta_k}{w_{k-1}}]$.
- *Needle.* ξ_k is uniform over $\{1, 2, \ldots, J_k\}$, where $J_k := \lfloor \frac{w_{k-1}}{4\Delta_k} \rfloor \geq 8$, and ϕ_k is a narrow bump at the ξ_k-th of J_k disjoint positions in P_{k-1}. Formally, $x_k := x_{k-1} - w_{k-1} + 4\Delta_k(2\xi_k - 1)$ and $\phi_k(p) := \min\{1, \frac{(w_k + 3\Delta_k - |p - x_k|)_+}{3\Delta_k}\}$, which is supported in $[x_k - 4\Delta_k, x_k + 4\Delta_k] \subseteq P_{k-1}$, where these supports are disjoint for different ξ_k. The slopes of $\Delta_k \phi_k$ are $\pm\sigma_k$ with $\sigma_k := \frac{1}{3}$.

In both cases, $\phi_k = 1$ exactly on $P_k := [x_k - w_k, x_k + w_k] \subseteq P_{k-1}$, so $R_k = L_k$ on P_k and $R_k < L_k$ elsewhere. Let F_k be the value distribution with $D_{F_k}(p) = \frac{R_k(p)}{p}$, let $R := \lim_k R_k$, let F be the value distribution with $D_F(p) = \frac{R(p)}{p}$, which depends on $\xi := (\xi_1, \xi_2, \ldots)$, and let μ be the distribution of F when ξ is drawn as above.

Horizons. Scale k forces regret at the horizon

$$T_k := \left\lfloor \frac{w_{k-1}}{2720\Delta_k^2} \right\rfloor \text{ on a tilt scale,} \qquad T_k := \left\lfloor \frac{J_k}{512\Delta_k^2} \right\rfloor \text{ on a needle scale,}$$

up to which the two halves (with infinite menus) or the J_k positions (with posted prices) cannot be told apart.

Theorem D.13 (Restatement of Theorem 4.6; Anytime Lower Bounds)**.** *For every $d \in (0, 1)$, there is a constant $c > 0$ and a probability distribution μ over value distributions with interval dimension d and multiplier 11, such that for every algorithm, the following holds with probability 1 over the randomized $F \sim \mu$:*

1. *Even with infinite menus, its regret on F is at least $cT^{d/(d+1)}$ for infinitely many T.*
2. *When using posted prices, its regret on F is at least $cT^{(d+1)/(d+2)}$ for infinitely many T.*

Proof. By Lemma D.14 below, μ is supported on value distributions with interval dimension d and multiplier 11. Fix an algorithm, and consider the first claim; the second claim is the same with needle scales in place of tilt scales. For a tilt scale k, let B_k be the event that the regret on F_k in T_k rounds is at least $\frac{T_k \Delta_k}{4}$. It only depends on $\xi_1, \ldots, \xi_k$, and by Lemma D.15, its conditional probability given $\xi_1, \ldots, \xi_{k-1}$ is at least $\frac{1}{2}$. So for tilt scales $k_1 < \cdots < k_n$, conditioning on $\xi_1, \ldots, \xi_{k_n - 1}$ gives $\Pr\{B_{k_1}^c \cap \cdots \cap B_{k_n}^c\} \leq \frac{1}{2} \Pr\{B_{k_1}^c \cap \cdots \cap B_{k_{n-1}}^c\} \leq \cdots \leq 2^{-n}$. Hence almost surely, B_k holds for infinitely many tilt scales k, and for each of them, the regret on F in T_k rounds is at least $\frac{T_k \Delta_k}{8}$ by Lemma D.17. Finally, $T_k \geq \frac{w_{k-1}}{5440\Delta_k^2} = \frac{\Delta_k^{-(1+d)}}{5440}$ on a tilt scale, so $T_k \Delta_k \geq 5440^{-1/(1+d)} T_k^{d/(d+1)}$; on a needle scale, $J_k \geq \frac{w_{k-1}}{8\Delta_k}$ and $T_k \geq \frac{J_k}{1024\Delta_k^2} \geq \frac{\Delta_k^{-(2+d)}}{8192}$, so $T_k \Delta_k \geq 8192^{-1/(2+d)} T_k^{(1+d)/(2+d)}$. □

We now prove the lemmas used above, in order of first use.

Lemma D.14 (Structure)**.** *For every ξ, F is a value distribution with revenue function R, whose maximum $L_\infty := \lim_k L_k$ is attained only at $p^* = \lim_k x_k$. For $k \geq 1$ and $p \in P_{k-1}$, the gap $L_\infty - R(p)$ is $\Delta_k + \Delta_{>k}$ if $\phi_k(p) = 0$, and $\sigma_k \operatorname{dist}(p, P_k) + \Delta_{>k}$ if $0 < \phi_k(p) < 1$; for $p \notin P_0$, it is at least $\frac{1}{4}(|p - \frac{1}{2}| - w_0) + \Delta_{>0}$. Moreover, F has interval dimension d with multiplier 11.*

Proof. All ϕ_k are supported in $P_0 \subseteq [0.4, 0.6]$, where $R_k \in [\frac{1}{4}, 0.26]$ and hence $\frac{R_k(p)}{p} \in [0.41, 0.66]$. On P_0, R_k is $\frac{1}{3}$-Lipschitz, as the slopes of different scales lie on the disjoint sets $P_{k-1} \setminus P_k$ and $\sigma_k \leq \frac{1}{3}$, so $(\frac{R_k}{p})' = \frac{1}{p}(R_k' - \frac{R_k}{p}) < 0$ almost everywhere on P_0. Off P_0, $R_k = R_0$, and $\frac{R_0}{p}$ is non-increasing: It is 1 where $R_0 = p$; left of P_0 where $R_0 < p$, its derivative is $-\frac{1+2w_0}{8p^2}$; right of P_0, its numerator decreases. So each $\frac{R_k}{p}$ is a continuous survival function, and so is its uniform limit $\frac{R}{p}$, as $0 \leq R_{k+1} - R_k \leq \Delta_{k+1}$.

As the scales after k are supported in P_k, we have $R = R_k = L_{k-1} + \Delta_k \phi_k$ on $P_{k-1} \setminus P_k$, where $\Delta_k(1 - \phi_k(p)) = \sigma_k \operatorname{dist}(p, P_k)$ if $\phi_k(p) > 0$; this gives the gaps on P_{k-1}. Off P_0, $R = R_0 \leq \frac{1}{4} - \frac{1}{4}(|p - \frac{1}{2}| - w_0)$. Every $p \neq p^*$ lies in some $P_{k-1} \setminus P_k$ or off P_0, so its gap is positive.

For the interval dimension, let $\delta \in [\Delta_{k+1}, \Delta_k)$ for some $k \geq 1$. By the gaps above, the prices in P_{k-1} with $\phi_k = 0$ and the prices off P_{k-1} have gaps larger than δ, and the δ-optimal prices with $0 < \phi_k < 1$ lie within $\frac{\delta}{\sigma_k}$ of P_k. So $N_F(\delta) \leq \frac{2w_k}{\delta} + \frac{2}{\sigma_k} + 1$, where $\frac{w_k}{\delta} \leq \frac{\Delta_{k+1}^{1-d}}{\delta} \leq \delta^{-d}$, and $\frac{2}{\sigma_k}$ is 6 on a needle scale and at most $\frac{w_{k-1}}{\Delta_k} = \Delta_k^{-d} \leq \delta^{-d}$ on a tilt scale, so $N_F(\delta) \leq 10\delta^{-d}$. If $\delta \geq \Delta_1$, then every δ-optimal price is within $w_0 + 4\delta$ of $\frac{1}{2}$, so $N_F(\delta) \leq \frac{2w_0}{\delta} + 9 \leq 11\delta^{-d}$. □

Lemma D.15 (One Scale)**.** *Fix $k \geq 1$ and $\xi_1, \ldots, \xi_{k-1}$.*

1. *If k is a tilt scale, then every algorithm, even with infinite menus and even if the seller observes v_t in each round, has regret at least $\frac{T_k \Delta_k}{4}$ on F_k in T_k rounds for some $\xi_k \in \{-1, 1\}$.*
2. *If k is a needle scale, then every algorithm using posted prices has regret at least $\frac{T_k \Delta_k}{2}$ on F_k in T_k rounds for at least $\frac{J_k}{2}$ values of ξ_k.*

Proof. Write $G := F_{k-1}$ and $w := w_{k-1}$. Both claims compare F_k with G, under which all prices in P_{k-1} have the same revenue L_{k-1}: On a tilt scale, the two choices of ξ_k form a two-point pair; on a needle scale, the J_k choices form a needle in a haystack. For every ξ_k, the prices outside the support of ϕ_k have revenue at most L_{k-1} under F_k, whereas $R_{F_k}(p^*_{F_k}) = L_{k-1} + \Delta_k$, so their gaps are at least Δ_k.

For the first claim, let $F^{\pm}$ be F_k with $\xi_k = \pm 1$. By Lemma D.16 below and the triangle inequality for H, $H^2(F^+, F^-) \leq \frac{340\Delta_k^2}{w}$, so the first $T_k - 1$ values have total variation distance at most $\sqrt{2T_k H^2(F^+, F^-)} \leq \frac{1}{2}$ between F^+ and F^-. The supports of ϕ_k for $\xi_k = \pm 1$ only share x_{k-1}, where both vanish, so every price has gap at least Δ_k under F^+ or under F^-, and Lemma D.1 with the horizon T_k gives the first claim.

For the second claim, let $F^{(j)}$ be F_k with $\xi_k = j$, let I_j be the support of ϕ_k in this case, and let n_j be the number of rounds whose price lies in I_j. Under G and $F^{(j)}$, posting p observes Bernoulli variables with means $\frac{R_G(p)}{p}$ and $\frac{R_{F^{(j)}}(p)}{p}$, which coincide for $p \notin I_j$, and otherwise differ by at most $\frac{\Delta_k}{p} \leq 2.5\Delta_k$ and lie in $[0.41, 0.67]$, so their KL divergence is at most $\frac{(2.5\Delta_k)^2}{0.22} \leq 32\Delta_k^2$. By the chain rule, the observations in T_k rounds have KL divergence at most $32\Delta_k^2 \mathbb{E}_G[n_j]$ between G and $F^{(j)}$. As $\sum_j \mathbb{E}_G[n_j] \leq T_k$, at least $\frac{J_k}{2}$ values of j have $\mathbb{E}_G[n_j] \leq \frac{2T_k}{J_k}$, for which this KL divergence is at most $\frac{64T_k\Delta_k^2}{J_k} \leq \frac{1}{8}$, and the total variation distance is at most $\frac{1}{4}$ by Pinsker's inequality. Then $\mathbb{E}_{F^{(j)}}[n_j] \leq \frac{2T_k}{J_k} + \frac{T_k}{4} \leq \frac{T_k}{2}$, so the regret on $F^{(j)}$ is at least $\Delta_k \mathbb{E}_{F^{(j)}}[T_k - n_j] \geq \frac{T_k\Delta_k}{2}$. □

Lemma D.16 (Each Scale Is Close in Hellinger Distance)**.** *For every $k \geq 1$ and every ξ, $H(F_k, F_{k-1}) \leq 2.2\sqrt{\Delta_k}$. Moreover, on a tilt scale k, $|\frac{dF_k}{dF_{k-1}} - 1| \leq \frac{10\Delta_k}{w_{k-1}} \leq \frac{1}{2}$, and $H^2(F_k, F_{k-1}) \leq \frac{85\Delta_k^2}{w_{k-1}}$.*

Proof. Write $G := F_{k-1}$, $w := w_{k-1}$, and $L := L_{k-1} \in [\frac{1}{4}, 0.26]$. The densities of F_k and G coincide outside the support of ϕ_k, which lies in $P_{k-1} \subseteq [0.4, 0.6]$. On this support, G has density $\frac{L}{p^2} \le 1.7$, and F_k has density $\frac{L+\Delta_k\phi_k}{p^2} - \frac{\Delta_k\phi_k'}{p}$.

- On a tilt scale, $|\phi_k'| \le \frac{4}{w}$, so as $p \le 0.6$ and $w \le 0.1$,

$$\left|\frac{dF_k}{dG} - 1\right| \le \frac{\Delta_k}{L} + \frac{0.6 \cdot 4\Delta_k}{wL} \le \frac{10\Delta_k}{w} \le \frac{1}{2}.$$

  As the support of ϕ_k has length w and hence D_G at most $1.7w$, we get $H^2(F_k, G) \le \frac{1}{2}\chi^2(F_k\|G) \le \frac{85\Delta_k^2}{w}$. As $w = \Delta_k^{1-d}$ and Δ_k is small, this gives $H(F_k, G) \le \sqrt{85}\,\Delta_k^{(1+d)/2} \le 2.2\sqrt{\Delta_k}$.
- On a needle scale, $|\phi_k'| \le \frac{1}{3\Delta_k}$, so $\frac{dF_k}{dG} \in [0.19, 1.81]$ on the support of ϕ_k, which has length $8\Delta_k$ and hence D_G at most $14\Delta_k$. So $H^2(F_k, G) \le \frac{1}{2} \cdot 0.81^2 \cdot 14\Delta_k \le 4.6\Delta_k$, i.e., $H(F_k, G) \le 2.2\sqrt{\Delta_k}$. □

Lemma D.17 (Finer Scales Do Not Affect the Revenue)**.** *For every ξ and $k \ge 1$, the regrets of an algorithm in T_k rounds on F and on F_k differ by at most $\frac{T_k\Delta_k}{8}$ if k is a tilt scale, and by at most $\frac{T_k\Delta_k}{4}$ if k is a needle scale and the algorithm uses posted prices.*

Proof. The finer scales change the revenue by at most $\Delta_{>k}$, and the value distribution by a squared Hellinger distance $\mathcal{O}(\Delta_{k+1})$, both negligible within T_k rounds. Formally, by construction, $L_\infty - L_k = \Delta_{>k}$ and $0 \le R - R_k \le \Delta_{>k}$.

On a tilt scale, the total revenue in T_k rounds lies in $[0, T_k]$ and is a function of the values and the internal randomness, so the regrets differ by at most $T_k\Delta_{>k} + T_k \cdot \sqrt{2T_k}\,H(F, F_k)$. By Lemma D.16, $H(F_m, F_{m-1}) \le 2.2\sqrt{\Delta_m}$ for every m, so as H is a metric and $F_m \to F$, we have $H(F, F_k) \le 2.3\sqrt{\Delta_{k+1}}$. So the regrets differ by at most $T_k(1.01\Delta_{k+1} + 3.3\sqrt{T_k\Delta_{k+1}}) \le \frac{T_k\Delta_k}{8}$, as $T_k \le \frac{\Delta_k^{-(1+d)}}{2720}$ and $\Delta_{k+1} = \Delta_k^q$ with $q > 3 + d$.

On a needle scale, the regret of a posted-price algorithm on a value distribution G is $\mathbb{E}_G[\sum_t (R_G(p_G^*) - R_G(p_t))]$, so the regrets differ by at most $T_k\Delta_{>k} + T_k \cdot \mathrm{TV}$, where TV is the total variation distance between the observations under F and under F_k. These Bernoulli variables have means that differ by at most $\frac{\Delta_{>k}}{p} \le 2.5\Delta_{>k}$, and lie in $[0.41, 0.67]$ where they differ, so each has KL divergence at most $\frac{(2.5\Delta_{>k})^2}{0.22} \le 29\Delta_{>k}^2$, and by the chain rule and Pinsker's inequality, $\mathrm{TV} \le \sqrt{\frac{29T_k\Delta_{>k}^2}{2}} \le 3.9\sqrt{T_k}\Delta_{>k}$. So the regrets differ by at most $5T_k^{3/2}\Delta_{>k} \le \frac{T_k\Delta_k}{4}$, as $T_k \le \frac{\Delta_k^{-(2+d)}}{2048}$ and $q > \frac{4+d}{2}$. □

D.7 Lower Bound for Lipschitz Experts

Definition 11 (Lipschitz Experts)**.** In the *Lipschitz experts* problem on $[0, 1]$ [KSU19], in each round t, an algorithm picks an arm $x_t \in [0, 1]$, then a payoff function $f_t : [0, 1] \to [0, 1]$ is drawn independently from an unknown distribution P whose expected payoff $\mu(x) := \mathbb{E}_{f\sim P}[f(x)]$ is 1-Lipschitz, and the algorithm receives $f_t(x_t)$ and observes f_t. Its regret is $T\max_x \mu(x) - \mathbb{E}[\sum_{t=1}^T \mu(x_t)]$.

The separation from Lipschitz experts in Section 4.5 uses the following lower bound, where the near-optimal arms concentrate as much as possible:

Lemma D.18 (Lipschitz Experts)**.** *For every $T \ge 1$, there are two payoff distributions P_+, P_-, under each of which the Δ-optimal arms lie in an interval of length 4Δ for every $\Delta > 0$ (i.e., interval dimension 0), such that any Lipschitz expert algorithm has regret at least $\frac{\sqrt{T}}{64}$ under P_+ or under P_-.*

Proof. Let $\epsilon := \frac{1}{8\sqrt{T}}$, $\mu_\pm(x) := \frac{1}{2} - \frac{1}{2}|x - \frac{1}{2} \mp \epsilon|$, which is 1-Lipschitz. Let $g := \frac{\mu_+ + \mu_-}{2} \geq \frac{1}{4}$ and $h := \frac{\mu_+ - \mu_-}{4\epsilon} \in [-\frac{1}{4}, \frac{1}{4}]$. Let P_+ and P_- draw every round t's payoff function f_t as follows:

$$P_+ : \ f_t = \begin{cases} g + h & \text{w.p. } \frac{1+2\epsilon}{2}, \\ g - h & \text{w.p. } \frac{1-2\epsilon}{2}, \end{cases} \qquad P_- : \ f_t = \begin{cases} g + h & \text{w.p. } \frac{1-2\epsilon}{2}, \\ g - h & \text{w.p. } \frac{1+2\epsilon}{2}, \end{cases}$$

so that their expected payoffs are $g \pm 2\epsilon h = \mu_\pm$. Under $P_\pm$, the Δ-optimal arms form an interval of length at most 4Δ, and as the gaps of every arm under P_+ and under P_- sum to at least ϵ, every arm has gap at least $\frac{\epsilon}{2}$ under one of them. Each payoff function only reveals which of $g \pm h$ is drawn, a coin whose squared Hellinger distance between P_+ and P_- is $1 - \sqrt{1 - 4\epsilon^2} \leq 4\epsilon^2$, so the first $T - 1$ payoff functions have total variation distance at most $\sqrt{8T\epsilon^2} \leq \frac{1}{2}$. In each round t, the arm x_t only depends on the first $t - 1$ payoff functions and the internal randomness, so as every arm has gap at least $\frac{\epsilon}{2}$ under P_+ or under P_-, the expected gaps of x_t under P_+ and under P_- sum to at least $\frac{\epsilon}{2}(1 - \frac{1}{2}) = \frac{\epsilon}{4}$. Summing over t, one of the two regrets is at least $\frac{\epsilon T}{8} = \frac{\sqrt{T}}{64}$. □

E Omitted Proofs in Section 5 (Adversarial Dynamic Pricing)

We first verify that Hedge and Fixed-Share satisfy Definition 3, then prove the two lemmas on the estimates (Lemmas 5.1 and 5.2), and finally prove Theorem 5.3. Throughout, $\mathbb{E}_t$ conditions on v_t and the history before round t, so that it only averages over the draws $p_t \sim \boldsymbol{w}_t$ and $I_t \sim \boldsymbol{\mu}_t$. If $\mu_t(I) = 0$, then $I_t \neq I$ almost surely, and we let $\widehat{\mathbb{1}}_t(I) := 0$ in Eq. (10).

Definition 12 (Hedge and Fixed-Share)**.** Let $\eta > 0$ be the learning rate. Given the weights $\boldsymbol{w}_t \in \triangle[K]$ and the losses $\boldsymbol{\ell}_t$ in round t, let

$$\widetilde{w}_{t+1,i} := \frac{w_{t,i} \exp(-\eta \ell_{t,i})}{\sum_{j \in [K]} w_{t,j} \exp(-\eta \ell_{t,j})}, \quad \forall i \in [K].$$

Both algorithms start from $w_{1,i} = \frac{1}{K}$ for all $i \in [K]$. *Hedge* [FS97] sets $w_{t+1,i} = \widetilde{w}_{t+1,i}$, i.e., $w_{t,i} \propto \exp(-\eta \sum_{\tau < t} \ell_{\tau,i})$. *Fixed-Share* [HW98] mixes each update with the uniform distribution at rate $\frac{1}{T}$, i.e., $w_{t+1,i} = (1 - \frac{1}{T})\widetilde{w}_{t+1,i} + \frac{1}{KT}$.

Lemma E.1 (Hedge and Fixed-Share Are Suitable)**.** *Let $\eta \leq 1$. Hedge (Definition 12) satisfies Definition 3 with the class $\mathcal{C}_0$ and $\textsc{Penalty} = \frac{\log K}{\eta}$. Fixed-Share satisfies it with the class $\mathcal{C}_{S-1}$ and $\textsc{Penalty} = \frac{S \log(KT) + 1}{\eta}$ for every $S \geq 1$. Both have $\boldsymbol{q}_t \equiv 1$.*

Proof. We bound the regret in each round by the change of $\log w_{t,i}$ for the comparator action i, and then sum these changes over the rounds.

One Round. Fix t. As $\eta \leq 1$ and $\ell_{t,j} \geq -1$, we have $\eta \ell_{t,j} \geq -1$, so $e^{-\eta \ell_{t,j}} \leq 1 - \eta \ell_{t,j} + \eta^2 \ell_{t,j}^2$ by $e^{-x} \leq 1 - x + x^2$ for $x \geq -1$. Together with $\log y \leq y - 1$,

$$\log \sum_{j \in [K]} w_{t,j} \exp(-\eta \ell_{t,j}) \leq -\eta \mathop{\mathbb{E}}_{j \sim \boldsymbol{w}_t} [\ell_{t,j}] + \eta^2 \mathop{\mathbb{E}}_{j \sim \boldsymbol{w}_t} [\ell_{t,j}^2].$$

As $\log \frac{\widetilde{w}_{t+1,i}}{w_{t,i}} = -\eta \ell_{t,i} - \log \sum_j w_{t,j} \exp(-\eta \ell_{t,j})$, for every $i \in [K]$,

$$\eta \left(\mathop{\mathbb{E}}_{j \sim \boldsymbol{w}_t} [\ell_{t,j}] - \ell_{t,i} \right) \leq \log \frac{\widetilde{w}_{t+1,i}}{w_{t,i}} + \eta^2 \mathop{\mathbb{E}}_{j \sim \boldsymbol{w}_t} [\ell_{t,j}^2]. \tag{25}$$

Summing over Rounds. Fix $(i_1^*, i_2^*, \dots, i_T^*) \in \mathcal{C}_{S-1}$, and apply Eq. (25) with $i = i_t^*$ in each round t. Regrouping the logarithms by rounds,

$$\sum_{t=1}^{T} \log \frac{\widetilde{w}_{t+1,i_t^*}}{w_{t,i_t^*}} = \log \frac{1}{w_{1,i_1^*}} + \sum_{t=1}^{T-1} \log \frac{\widetilde{w}_{t+1,i_t^*}}{w_{t+1,i_{t+1}^*}} + \log \widetilde{w}_{T+1,i_T^*},$$

where $\log \frac{1}{w_{1,i_1^*}} = \log K$ and $\log \widetilde{w}_{T+1,i_T^*} \le 0$. After dividing by η, Definition 3 holds with $\boldsymbol{q}_t \equiv 1$, as long as PENALTY bounds $\frac{1}{\eta}$ times this sum, which we now do for each algorithm.

Hedge. Every sequence in $\mathcal{C}_0$ has $i_{t+1}^* = i_t^*$, and $w_{t+1,i} = \widetilde{w}_{t+1,i}$, so the middle sum is 0, and $\text{PENALTY} = \frac{\log K}{\eta}$.

Fixed-Share. As $w_{t+1,i} \ge (1 - \frac{1}{T})\widetilde{w}_{t+1,i}$, each of the at most $T-1$ terms in the middle sum with $i_{t+1}^* = i_t^*$ is at most $\log \frac{T}{T-1} \le \frac{1}{T-1}$, so these terms sum to at most 1. As $w_{t+1,i} \ge \frac{1}{KT}$ and $\widetilde{w}_{t+1,i} \le 1$, each of the at most $S-1$ other terms is at most $\log(KT)$. Hence $\text{PENALTY} = \frac{\log K + (S-1)\log(KT) + 1}{\eta} \le \frac{S\log(KT)+1}{\eta}$. □

Lemma E.2 (Restatement of Lemma 5.1; Importance-Weighted Estimator)**.** *For every round $t \in [T]$, for any price $p \in \mathcal{S}$ such that $w_t(p) > 0$, we have $\mathbb{E}_t[\widehat{R}_t(p)] = p\mathbb{1}[v_t \ge p]$. Meanwhile, $\widehat{R}_t(p) \le 1$ is bounded from above.*

Proof. By Lemma C.3, $\mathcal{I}_{\mathcal{T}}[0,p)$ partitions $[0,p)$ into tree intervals (and is empty if $p = 0$), so $\mathbb{1}[v_t \ge p] = 1 - \sum_{I \in \mathcal{I}_{\mathcal{T}}[0,p)} \mathbb{1}_t(I)$. For every tree interval I with $\mu_t(I) > 0$, as $\Pr_t\{I_t = I\} = \mu_t(I)$, Eq. (10) gives $\mathbb{E}_t[\widehat{\mathbb{1}}_t(I)] = \mathbb{1}_t(I)$. If $w_t(p) > 0$ and $p > 0$, then every $I \in \mathcal{I}_{\mathcal{T}}[0,p)$ has $\mu_t(I) > 0$ by Eq. (9), as p contributes $w_t(p)p^2 > 0$ to the sum defining $\mu_t(I)$. Hence, by Eq. (11),

$$\mathop{\mathbb{E}}_t\left[\widehat{R}_t(p)\right] = p\left(1 - \sum_{I \in \mathcal{I}_{\mathcal{T}}[0,p)} \mathop{\mathbb{E}}_t\left[\widehat{\mathbb{1}}_t(I)\right]\right) = p\left(1 - \sum_{I \in \mathcal{I}_{\mathcal{T}}[0,p)} \mathbb{1}_t(I)\right) = p\mathbb{1}[v_t \ge p],$$

which also holds if $p = 0$, as both sides are then 0. Finally, as $\widehat{\mathbb{1}}_t \ge 0$, we have $\widehat{R}_t(p) \le p \le 1$ for every $p \in \mathcal{S}$. □

The proof of Theorem 5.3 also compares with prices that $\mathcal{A}^{\text{OL}}$ may not play, for which the estimate is still optimistic in expectation:

Lemma E.3 (Optimistic Estimates for All Prices)**.** *For every round t and every price $p \in \mathcal{S}$, we have $\mathbb{E}_t[\widehat{R}_t(p)] \ge p\mathbb{1}[v_t \ge p]$.*

Proof. For every tree interval I, as $\Pr_t\{I_t = I\} = \mu_t(I)$, Eq. (10) gives $\mathbb{E}_t[\widehat{\mathbb{1}}_t(I)] = \mathbb{1}_t(I)$ if $\mu_t(I) > 0$, and $\widehat{\mathbb{1}}_t(I) = 0$ otherwise, so $\mathbb{E}_t[\widehat{\mathbb{1}}_t(I)] \le \mathbb{1}_t(I)$. As $\mathcal{I}_{\mathcal{T}}[0,p)$ partitions $[0,p)$ by Lemma C.3, and $p \ge 0$,

$$\mathop{\mathbb{E}}_t\left[\widehat{R}_t(p)\right] = p\left(1 - \sum_{I \in \mathcal{I}_{\mathcal{T}}[0,p)} \mathop{\mathbb{E}}_t\left[\widehat{\mathbb{1}}_t(I)\right]\right) \ge p\left(1 - \sum_{I \in \mathcal{I}_{\mathcal{T}}[0,p)} \mathbb{1}_t(I)\right) = p\mathbb{1}[v_t \ge p].$$ □

Lemma E.4 (Restatement of Lemma 5.2; Second Moment)**.** *For every round t, we have*

$$\mathop{\mathbb{E}}_t\left[\sum_{p \in \mathcal{S}} w_t(p)\widehat{R}_t^2(p)\right] \le 2 + 4(\log_2 T + 1)^2.$$

Proof. Fix a round t. As only I_t can have $\widehat{\mathbb{1}}_t(I) \neq 0$, we have $(\sum_{I\in\mathcal{I}_\mathcal{T}[0,p)} \widehat{\mathbb{1}}_t(I))^2 = \sum_{I\in\mathcal{I}_\mathcal{T}[0,p)} \widehat{\mathbb{1}}_t^2(I)$, and $\mathbb{E}_t[\widehat{\mathbb{1}}_t^2(I)]$ is $\frac{\mathbb{1}_t(I)}{\mu_t(I)}$ if $\mu_t(I) > 0$ and 0 otherwise. Let $\mathcal{T}_t := \{I \in \mathcal{T} \mid \mu_t(I) > 0\}$. Then

$$\begin{aligned}\mathop{\mathbb{E}}_t\left[\sum_{p\in\mathcal{S}} w_t(p)\widehat{R}_t^2(p)\right] &\overset{(a)}{\le} 2\sum_{p\in\mathcal{S}} w_t(p)p^2 + 2\sum_{p\in\mathcal{S}} w_t(p)p^2 \sum_{I\in\mathcal{I}_\mathcal{T}[0,p)\cap\mathcal{T}_t} \frac{\mathbb{1}_t(I)}{\mu_t(I)} \\ &\overset{(b)}{=} 2\sum_{p\in\mathcal{S}} w_t(p)p^2 + 2\sum_{I\in\mathcal{T}_t} \frac{\mathbb{1}_t(I)}{\mu_t(I)} \sum_{p:\mathcal{I}_\mathcal{T}[0,p)\ni I} w_t(p)p^2 \overset{(c)}{\le} 2 + 2Z_t\sum_{I\in\mathcal{T}} \mathbb{1}_t(I),\end{aligned}$$

where (a) uses $\widehat{R}_t^2(p) \le 2p^2 + 2p^2(\sum_{I\in\mathcal{I}_\mathcal{T}[0,p)} \widehat{\mathbb{1}}_t(I))^2$; (b) swaps the summations; and (c) uses $\sum_p w_t(p)p^2 \le 1$ and that, by Eq. (9), the inner sum is $Z_t\mu_t(I)$ if $Z_t > 0$ and is 0 if $Z_t = 0$. By Lemma C.3 with $N = T$, the tree $\mathcal{T}$ has at most $\log_2 T + 1$ depths, each $\mathcal{I}_\mathcal{T}[0,p)$ has at most two tree intervals of each depth, and v_t lies in at most one tree interval of each depth. So $Z_t = \sum_p w_t(p)p^2|\mathcal{I}_\mathcal{T}[0,p)| \le 2(\log_2 T + 1)$ and $\sum_I \mathbb{1}_t(I) \le \log_2 T + 1$, which proves the lemma. □

Theorem E.5 (Restatement of Theorem 5.3; Adversarial Dynamic Pricing)**.** *For any comparator class $\mathcal{C} \subseteq \{0, \frac{1}{T}, \dots, 1\}^T$, for any $\mathcal{A}^{OL}$ satisfying Definition 3 with $K = T + 1$ and $\mathcal{C}$ translated to $[K]$, Algorithm 3 ensures*

$$\mathbb{E}\left[\sum_{t=1}^T p_t^*\mathbb{1}[v_t \ge p_t^*]\right] - \mathbb{E}\left[\sum_{t=1}^T \mathit{revenue}_t\right] \le \textsc{Penalty}(\mathcal{C}, T, T+1, \eta) + \mathcal{O}\left(\eta T\log^2 T + 1\right),\ \forall (p_1^*, \dots, p_T^*) \in \mathcal{C}.$$

In particular, Hedge [FS97] with $\eta = (T\log T)^{-1/2}$ gives $\mathcal{O}\left(\sqrt{T\log^3 T}\right)$ regret against the best fixed price $\mathcal{C}_0$, and Fixed-Share [HW98] with $\eta = \sqrt{S/(T\log T)}$ gives $\mathcal{O}\left(\sqrt{ST\log^3 T}\right)$ against $\mathcal{C}_{S-1}$.

Proof. The first claim is Lemma E.6 below, whose $\mathcal{O}(\cdot)$ term is $\eta T(2 + 4(\log_2 T + 1)^2) + 2$. The rates compare with prices in $[0, 1]$ rather than in $\mathcal{S}$, which Lemma E.7 below handles. As the rates are asymptotic, we assume $T \ge 3$; for Fixed-Share, we also assume $S \le T$, as otherwise $\sqrt{ST\log^3 T} \ge T$ exceeds the trivial regret bound T. Then $\eta \le 1$ in both cases below, and $\log(T(T + 1)) \le 3\log T$.

- Hedge: By Lemma E.1, it satisfies Definition 3 with the class $\mathcal{C}_0$ and $\textsc{Penalty} = \frac{\log(T+1)}{\eta}$. So for $\eta = (T\log T)^{-1/2}$, Lemma E.7 with $S = 0$ bounds the regret against every fixed price in $[0, 1]$ by
$$\log(T + 1)\sqrt{T\log T} + \sqrt{\frac{T}{\log T}}\left(2 + 4(\log_2 T + 1)^2\right) + 3 = \mathcal{O}\left(\sqrt{T\log^3 T}\right).$$
- Fixed-Share: By Lemma E.1, it satisfies Definition 3 with the class $\mathcal{C}_{S-1}$ and $\textsc{Penalty} = \frac{S\log(T(T+1))+1}{\eta}$. So for $\eta = \sqrt{S/(T\log T)}$, Lemma E.7 with $S - 1$ in place of S bounds the regret against every sequence in $\mathcal{C}_{S-1}$ by
$$(3S\log T + 1)\sqrt{\frac{T\log T}{S}} + \sqrt{\frac{ST}{\log T}}\left(2 + 4(\log_2 T + 1)^2\right) + 3 = \mathcal{O}\left(\sqrt{ST\log^3 T}\right). \qquad \square$$

Lemma E.6 (Theorem 5.3 with Explicit Constants)**.** *Let $\mathcal{C} \subseteq \mathcal{S}^T$, and let $\mathcal{A}^{OL}$ satisfy Definition 3 with $K = T + 1$ and $\mathcal{C}$ translated to $[K]$. Then for every $(p_1^*, p_2^*, \dots, p_T^*) \in \mathcal{C}$, Algorithm 3 ensures*

$$\mathbb{E}\left[\sum_{t=1}^T p_t^*\mathbb{1}[v_t \ge p_t^*]\right] - \mathbb{E}\left[\sum_{t=1}^T \mathit{revenue}_t\right] \le \textsc{Penalty}(\mathcal{C}, T, T + 1, \eta) + \eta T\left(2 + 4(\log_2 T + 1)^2\right) + 2.$$

Proof. We compare the revenue of the menus with that of the exploitation prices (Step 1), apply Definition 3 to the estimated revenues on every sample path (Step 2), and take expectations using Lemmas 5.1, 5.2, and E.3 (Step 3).

Step 1: Exploration Loses at Most 2γ Revenue per Round. Fix a round t, and let $I_t = [a, b)$. The menu in round t is built from the price distribution P_t that puts mass $1 - 2\gamma$ on p_t and mass γ on each of a and b. As all prices are in $[0, 1]$,

$$\text{revenue}_t = \mathop{\mathbb{E}}_{p\sim P_t}[p\mathbb{1}[v_t \geq p]] \geq (1-2\gamma)p_t\mathbb{1}[v_t \geq p_t] \geq p_t\mathbb{1}[v_t \geq p_t] - 2\gamma.$$

As $p_t \sim \boldsymbol{w}_t$ under $\mathbb{E}_t$, we get $\mathbb{E}_t[\text{revenue}_t] \geq \sum_{p\in\mathcal{S}} w_t(p)p\mathbb{1}[v_t \geq p] - 2\gamma$.

Step 2: Online Learning on Every Sample Path. By Lemma 5.1, the loss $-\widehat{R}_t(p)$ fed to $\mathcal{A}^{\text{OL}}$ is at least -1. As Definition 3 holds for every sequence of such losses, and $\boldsymbol{q}_t \leq 1$, on every sample path,

$$\sum_{t=1}^{T}\widehat{R}_t(p_t^*) - \sum_{t=1}^{T}\sum_{p\in\mathcal{S}} w_t(p)\widehat{R}_t(p) \leq \textsc{Penalty}(\mathcal{C}, T, T+1, \eta) + \eta\sum_{t=1}^{T}\sum_{p\in\mathcal{S}} w_t(p)\widehat{R}_t^2(p).$$

Step 3: Taking Expectations. As p_t^* is fixed in advance and $\boldsymbol{w}_t$ is determined by the history before round t, Lemma E.3 gives $\mathbb{E}_t[\widehat{R}_t(p_t^*)] \geq p_t^*\mathbb{1}[v_t \geq p_t^*]$, and Lemma 5.1 gives $\mathbb{E}_t[\sum_{p\in\mathcal{S}} w_t(p)\widehat{R}_t(p)] = \sum_{p\in\mathcal{S}} w_t(p)p\mathbb{1}[v_t \geq p]$, as only prices with $w_t(p) > 0$ contribute. Taking expectations in Step 2 and bounding its last term by Lemma 5.2,

$$\mathbb{E}\left[\sum_{t=1}^{T} p_t^*\mathbb{1}[v_t \geq p_t^*]\right] - \mathbb{E}\left[\sum_{t=1}^{T}\sum_{p\in\mathcal{S}} w_t(p)p\mathbb{1}[v_t \geq p]\right] \leq \textsc{Penalty}(\mathcal{C}, T, T+1, \eta) + \eta T\left(2 + 4(\log_2 T + 1)^2\right).$$

By Step 1, replacing the second expectation by $\mathbb{E}[\sum_t \text{revenue}_t]$ adds at most $2\gamma T = 2$. □

The rates in Theorem 5.3 compare with prices in $[0, 1]$ rather than in $\mathcal{S}$. Rounding them down to $\mathcal{S}$ adds no switches and loses at most $\frac{1}{T}$ revenue per round:

Lemma E.7 (Rounding Comparators to the Grid)**.** *Let $S \geq 0$, and let $\mathcal{A}^{OL}$ satisfy Definition 3 with $K = T + 1$ and the class $\mathcal{C}_S$ translated to $[K]$. Then for every $(p_1^*, p_2^*, \ldots, p_T^*) \in \mathcal{C}_S$, i.e., every sequence of prices in $[0, 1]$ that switches at most S times, Algorithm 3 ensures*

$$\mathbb{E}\left[\sum_{t=1}^{T} p_t^*\mathbb{1}[v_t \geq p_t^*]\right] - \mathbb{E}\left[\sum_{t=1}^{T} \textit{revenue}_t\right] \leq \textsc{Penalty}(\mathcal{C}_S, T, T+1, \eta) + \eta T\left(2 + 4(\log_2 T + 1)^2\right) + 3.$$

Proof. Let $\overline{p}_t := \frac{\lfloor Tp_t^*\rfloor}{T} \in \mathcal{S}$. If $p_t^* = p_{t+1}^*$, then $\overline{p}_t = \overline{p}_{t+1}$, so $(\overline{p}_1, \overline{p}_2, \ldots, \overline{p}_T)$ also switches at most S times, and Lemma E.6 applies to it. It remains to show that rounding loses at most $\frac{1}{T}$ in each round t:

$$\overline{p}_t\mathbb{1}[v_t \geq \overline{p}_t] \overset{(a)}{\geq} \overline{p}_t\mathbb{1}[v_t \geq p_t^*] \overset{(b)}{\geq} \left(p_t^* - \frac{1}{T}\right)\mathbb{1}[v_t \geq p_t^*] \geq p_t^*\mathbb{1}[v_t \geq p_t^*] - \frac{1}{T},$$

where (a) uses $\overline{p}_t \leq p_t^*$ and $\overline{p}_t \geq 0$, and (b) uses $p_t^* - \overline{p}_t < \frac{1}{T}$. □