MICHIGAN
ROSS SCHOOL OF BUSINESS

Working Paper

Near-Optimal Bisection Search for Nonparametric Dynamic Pricing with Inventory Constraint

Yanzhe Lei
Stephen M. Ross School of Business
University of Michigan

Stefanus Jasin
Stephen M. Ross School of Business
University of Michigan

Amitabh Sinha
Stephen M. Ross School of Business
University of Michigan

Ross School of Business Working Paper
Working Paper No. 1252
October 2014

Near-Optimal Bisection Search for Nonparametric Dynamic Pricing with Inventory Constraint

Murray (Yanzhe) Lei
Ross School of Business, University of Michigan, Ann Arbor, MI 48109,
leiyz@umich.edu

Stefanus Jasin
Ross School of Business, University of Michigan, Ann Arbor, MI 48109,
sjasin@umich.edu

Amitabh Sinha
Ross School of Business, University of Michigan, Ann Arbor, MI 48109,
amitabh@umich.edu

We consider a single-product revenue management problem with an inventory constraint and unknown, noisy, demand function. The objective of the firm is to dynamically adjust the prices to maximize total expected revenue. We restrict our scope to the nonparametric approach where we only assume some common regularity conditions on the demand function instead of a specific functional form. We propose a family of pricing heuristics that successfully balance the tradeoff between *exploration* and *exploitation*. The idea is to generalize the classic bisection search method to a problem that is affected both by stochastic noise and an inventory constraint. Our algorithm extends the bisection method to produce a sequence of pricing intervals that converge to the optimal static price with high probability. Using *regret* (the revenue loss compared to the deterministic pricing problem for a clairvoyant) as the performance metric, we show that one of our heuristics exactly matches the theoretical asymptotic lower bound that has been previously shown to hold for any feasible pricing heuristic. Although the results are presented in the context of revenue management problems, our analysis of the bisection technique for stochastic optimization with learning can be potentially applied to other application areas.

Key words: revenue management; pricing; nonparametric; learning; asymptotic analysis; bisection search.

1. Introduction. Dynamic pricing has became a common practice in many firms nowadays. It plays a central role in the revenue optimization of many industries including airlines, hotels, car rentals, and retails (Talluri and van Ryzin [36], Özer and Phillips [32]). In the typical dynamic pricing problem, firms adaptively adjust their prices in response to market demand and try to maximize their expected revenue. The success of this approach relies heavily on the firms' knowledge about the relationship between market demand and the posted price, which is characterized by a demand function. Although in reality firms may not know the exact demand function, firms can still dynamically price their products through a combination of active learning (e.g., price experimentation) and dynamic optimization. The challenge, however, is obvious: Given the limited time window of opportunity and the limited on-hand inventory, firms have to balance the effort spent on probing the true demand function (exploration) and generating near-optimal revenue (exploitation).

The literature on dynamic pricing with demand learning can be broadly divided into two categories: *parametric* and *nonparametric* models. (See den Boer [16] for a recent overview of the field.) In the parametric model, it is assumed that the firms know the functional form of the underlying demand function (e.g., linear, logit, etc.). The key challenges in such setting are to estimate the unknown demand parameters and to develop a price optimization scheme utilizing this estimate.

Some popular estimation procedures that have been studied in the literature include Bayesian method (Araman and Caldentey [2]; Farias and van Roy [21]; Harrison et al. [23]), Maximum Likelihood estimation (Broder and Rusmevichientong [11]; den Boer [17]; den Boer and Zwart [19]; den Boer and Zwart [18]; Chen et al. [13]), and Least Squares approach (Bertsimas and Perakis [4]; Keskin and Zeevi [27]). In contrast to parametric model, nonparametric model does not assume that the firms know the functional form of the demand function; instead, it only assumes a certain set of mild regularity conditions such as the decreasing property of demand as a function of price, the boundedness of the first and second derivatives of the demand function, and the unimodality of the revenue function. In such setting, the firms' tasks are further complicated by the fact that there is no explicit function to optimize. Current literature (e.g., Chen et al. [13] and Wang et al. [38]) suggest that parametric approaches outperform nonparametric approaches, at least asymptotically. Given that parametric approach assumes a precise knowledge of the functional form of the underlying demand function, this observation is hardly surprising. The question is whether a parametric approach is always applicable. To illustrate, suppose that the underlying demand function is actually a logit function. What will happen if we mistakenly assume a linear function instead of a logit function when estimating the demand parameters? As shown in Besbes and Zeevi [8], although model mis-specification is not always detrimental, it can lead to sub-optimal prices, which yield a large loss in revenue. It remains an open research problem whether there is a way to make parametric approach more robust with respect to model mis-specification. This leaves the firms in a quandary of having to choose between a parametric approach, with the risk of model mis-specification, or a nonparametric approach, with a weaker performance guarantee. The purpose of this paper is to address this issue. In particular, we will consider a nonparametric approach and study a scheme that will be shown to match the theoretical performance guarantee of the best known parametric approach in the single product setting.

The proposed heuristics. A good pricing policy must balance the tradeoff between demand learning (exploration) and revenue maximization (exploitation) while also successfully dealing with the dynamics caused by stochastic demands and inventory constraints. Our heuristics achieve these objectives by generating a sequence of shrinking intervals that converge to the optimal static price calculated via a deterministic relaxation of the original dynamic pricing problem. More specifically, we generalize the standard bisection search algorithm to stochastic and constrained setting. (Our heuristics actually generalize the trisection search. However, for consistency with the existing optimization literature, we will simply call it a bisection instead of a trisection.) We use empirical mean of the observed demands as an estimate of the true demand rate to shrink the intervals accordingly. The sampling frequencies are chosen carefully: If they are too small, the resulting estimates are not very accurate; if, on the other hand, they are too large, we spend too much time on the sub-optimal prices, which incurs a large revenue loss. For the single-product problem, the implementation of our heuristics can be essentially divided into two phases: the exploration phase and the exploitation phase. Since it is known in this setting that the optimal static price can be written as the maximum of the unconstrained maximizer and the clearance price (see Gallego and van Ryzin [22]), the purpose of the exploration phase is to determine the identity of the optimal static price via bisection search. We show that it is possible to distinguish this identity quickly with a very high probability. During the exploitation phase, we apply another bisection search to more efficiently shrink the intervals according to the identity of the optimal price.

Let $\theta > 0$ denote the relative size of the problem (i.e., the amount of initial inventory). It can be shown that the asymptotic revenue loss of the proposed heuristic is $O(\sqrt{\theta}\log\theta)$, which is very close to the known $\Omega(\sqrt{\theta})$ theoretical lower bound on the performance of any feasible pricing policy in the setting of unknown demand function. Moreover, the performance guarantee of this heuristic also dominates the performance of the best known nonparametric scheme for single-product problem in the literature, which is $O(\sqrt{\theta}\log^{4.5}\theta)$ (Wang et al. [38]). Can we further reduce the $O(\sqrt{\theta}\log\theta)$

revenue loss? It turns out that it is possible: If we use Stochastic Approximation algorithms (i.e., Kiefer-Wolfowitz and Robbins-Monro, see Broadie et al. [10]) during the exploitation phase instead of another bisection search, then the resulting revenue loss is exactly $\Theta(\sqrt{\theta})$. Thus, we have provided an "optimal" nonparametric pricing heuristic for the setting of a single-product problem with inventory constraint. (In the case where the firms know the functional form of the demand function, i.e., parametric model, the $\Omega(\sqrt{\theta})$ lower bound has been repeatedly shown to be tight. For example, in the setting without inventory constraints, Keskin and Zeevi [27], den Boer and Zwart [19], and Broder and Rusmevichientong [11], each proposes a parametric pricing heuristic that guarantees a revenue loss of the order of $O(\sqrt{\theta})$. As for the setting with inventory constraints, recently Chen et al. [13] propose a heuristic that exactly matches this lower bound. Their result holds for a general parametric model with an arbitrary set of inventory constraints. Thus, they have resolved the parametric dynamic pricing problem with inventory constraints.)

Related literature. Apart from the standard parametric and nonparametric approaches, there are also works in the literature that consider robust optimization approach. Lim and Shanthikumar [31] study a robust formulation of the classic single-product pricing problem where nature adversarially chooses the distribution governing the demand realization. They use a conservative max-min formulation that does not involve real-time demand learning and bears no closed form solution in general. Eren and Maglaras [20] also study the robust setting and use a competitive ratio formulation. However, they only deal with the setting without inventory constraint and assume deterministic demand. Perakis and Roels [33] adopt both the maximin and minimax formulation. Their focus is on deriving structural insights instead of proving a performance bound. As has been noted in Cohen et al. [14], the robust optimization literature mainly focuses on static problems and the previously realized uncertainty is not utilized to adjust the pricing decision; this may result in a rather conservative pricing decision. Cohen et al. [14] try to bridge the gap between robust approach and data-driven optimization by proposing algorithms that utilize the realized demands and converges to the optimal robust solution. However, there is no theoretical guarantee on the convergence rate of their algorithm. Rusmevichientong et al. [35] also adopt a data-driven approach. They provide a bound on the number of samples required to guarantee a near-optimal revenue if one uses the empirical optimal price under general consumer choice model. Their approach is restricted to static setting, i.e., the pricing decision does not depend on the previously realized demand uncertainties. Therefore, there is no trade-off between revenue earning and demand learning.

On the technical side, our work is also related to three other streams of literature. The first one is the continuum-armed bandit literature (e.g., Agrawal [1]; Auer et al. [3]; Cope [15]; and Kleinberg [29]). While there are some high-level connections between our approach and the bandit approach, the presence of an inventory constraint in our problems clearly distinguishes our work from theirs. Another stream of related literature is the study of bisection search. Despite its long history and broad prevalence, there is little work that studies its generalization into stochastic setting. To the best of our knowledge, Waeber et al. [37] is the only work that attempts to generalize the deterministic bisection search into a stochastic setting. However, the scope of their application is restricted to a root-finding problem. Thus, compared to the existing studies on bisection search method, our work is the only one that combines the challenge of stochastic setting and constrained optimization. These distinctions do not allow any direct comparison to the existing literature. Finally, our work is also related to the Online Convex Optimization (OCO) literature (see Cesa-Bianchi et al. [12] for a review). OCO considers a setting where at each time period, after a decision has been mand, *nature* choose a cost function adversarially. The performance of a given policy is then compared to the policy that uses the best *static* action in hindsight. Although there are some similarities in the problem formulation, the vast majority of the OCO literature restricts its scope to convex cost functions and unconstrained setting; this clearly differentiates our work from OCO.

Remainder of this paper. The remainder of the paper is organized as follows. In Section 2, we introduce the problem formulation. In Sections 3 and 4, we discuss our heuristics and prove their asymptotic bounds. Section 5 summarizes the paper and potential future research directions. Unless otherwise noted, the details of the proofs can be found in the Appendix.

2. Problem Formulation.

In this section, we first describe the problem setting and discuss general modeling assumptions. We then introduce the deterministic analog of the original stochastic pricing problem and discuss our performance metric.

2.1. Model setting.

We consider a monopolist selling a single product with C units of initial inventory. The selling horizon is discrete and divided into T periods. Without loss of generality, we assume that at most one customer arrives during each period. At the beginning of period t, the firm first posts the price p_t and in turn induces a stochastic demand $D_t(p_t)$ with rate $\lambda(p_t) = \mathbb{E}[D_t(p_t)]$. Note that, since at most one customer arrives during each period, the term $\lambda(p_t)$ can be interpreted as the probability of a request during period t given p_t. Demands across different periods are assumed to be independent. Let $r(p) = p\lambda(p)$ denote the revenue rate and p^u its unique maximizer. Also, let Ω_p and Ω_λ denote the convex set of feasible prices and demand rates, respectively. We make the following assumptions on the underlying demand and revenue rate functions:

Modeling Assumptions

A1. The function $\lambda(\cdot) : \Omega_p \to \Omega_\lambda$ is invertible and twice-differentiable. Moreover, $\lambda(p)$ is strictly decreasing in p, i.e., there exists a constant $L > 0$ such that $|\lambda'(p)| \geq L$. We will use $p(\cdot) : \Omega_\lambda \to \Omega_p$ to denote the inverse of $\lambda(\cdot)$.

A2. The function $r(p)$ is strictly unimodal. In addition, $r(\lambda) := p(\lambda)\lambda$ is strictly concave in λ. (By abuse of notation, we will often write $r(\lambda)$ instead of $r(p)$ to denote the direct dependency of revenue on demand rate instead of price.)

A3. $\lambda(p)$ and $p(\lambda)$ are Lipschitz continuous with a factor $K > 0$, i.e., $\forall p, p' \in \Omega_p, |\lambda(p) - \lambda(p')| \leq K|p - p'|$, and $\forall \lambda, \lambda' \in \Omega_\lambda, |p(\lambda) - p(\lambda')| \leq K|\lambda - \lambda'|$.

A4. There exists a "shut-off" price p^∞ such that if $\{p^k\}$ is any price sequence satisfying $p^k \to p^\infty$, then we have $\lambda(p^k) \to 0$.

A5. There exists positive constants $M_L < M_U$ such that $0 > -M_L \geq r''(\lambda) \geq -M_U$ and $M_L|p - p^u| \leq |r'(p)| \leq M_U|p - p^u|$.

Assumptions A1-A4, together with the first part of A5, are quite natural and have been repeatedly used in the literature (cf. Besbes and Zeevi [6], Wang et al. [38]). In particular, the existence of shut-off price allows the firm to effectively shut down the demand whenever desired. The second part of A5 is needed only for the analysis of Stochastic Approximation algorithms in Section 3.3. (They are standard assumptions in the Stochastic Approximation literature, e.g., Broadie et al. [10].)

2.2. The stochastic and deterministic pricing problems.

We say that a pricing policy $\pi := (p_t^\pi : 0 \leq t \leq T)$ is *non-anticipating* if the decision p_t^π at the beginning of period t only depends on past prices $\{p_s^\pi : 0 \leq s < t\}$ and past demand observations $\{D_s(p_s^\pi) : 0 \leq s < t\}$. Furthermore, we also say that a pricing policy π is *admissible* if $p_t^\pi \in \Omega_p$ for all t and π is non-anticipating. Let Π denote the set of all admissible pricing policies. The stochastic formulation of the dynamic pricing problem is given by

$$J^* = \max_{\pi\in\Pi} \mathbb{E}\left[\sum_{t=1}^{T} p_t^\pi \cdot D_t(p_t^\pi)\right] \text{ such that } \sum_{t=1}^{T} D_t(p_t^\pi) \le C \qquad a.s. \tag{1}$$

The deterministic analog of the above pricing problem is

$$J^D = \max_{p_t\in\Omega_p} \sum_{t=1}^{T} r(p_t) \text{ such that } \sum_{t=1}^{T} \lambda(p_t) \le C. \tag{2}$$

By assumption A1, the above deterministic problem can also be written as

$$J^D = \max_{\lambda_t\in\Omega_\lambda} \sum_{t=1}^{T} r(\lambda_t) \text{ such that } \sum_{t=1}^{T} \lambda_t \le C. \tag{3}$$

Let $\{p_t^D\}$ denote the unique optimal solution of (2); correspondingly, we also define $\lambda_t^D := \lambda(p_t^D)$. Since the demand function is time-homogeneous, it can be shown that $p_t^D = p^D$ for all t (see Gallego and van Ryzin [22] for proof). Thus, the optimal deterministic price is static. For analytical tractability, we will assume that both p^D and p^u lie in a proper interior of Ω_p. We state this assumption formally below.

A6. There exists $0 < \underline{p} < \bar{p}$ such that such that p^D, $p^u \in [\underline{p}, \bar{p}] \subset \Omega_p$.

2.3. Performance metric and asymptotic setting. Let J^π denote the expected revenue earned under pricing policy π. It is known that J^D is an upper bound for the expected revenue under any admissible policy, i.e., $J^D \ge J^\pi$ for all $\pi \in \Pi$ (see Gallego and van Ryzin [22] for proof, we omit the details). Thus, following the convention in the literature, as our performance metric, we will define the revenue loss of an admissible policy π as $\mathcal{R}^\pi = J^D - J^\pi$. Since it is typical for revenue management firms to sell a large inventory during a selling season, following the standard setting in the literature, in this paper we will consider a sequence of increasing problems where we scale both the size of the initial inventory level and the number of selling periods by a factor of $\theta > 0$. To be precise, the θ^{th} problem is parameterized by $(C_\theta, T_\theta) = (\theta C, \theta T)$. Let J_θ^D denote the optimal value of the deterministic problem (2) with scaling factor θ (it is not difficult to see that $J_\theta^D = \theta J^D$) and let J_θ^π denote the expected revenue under policy π for a problem with scaling factor θ. (Throughout this paper, the subscript θ will be consistently used as a reference to the problem with scaling factor θ.) Our objective is to study the asymptotic behavior of $\mathcal{R}_\theta^\pi = J_\theta^D - J_\theta^\pi$ as θ grows large. The scaling parameter θ can be interpreted as the size of the potential market, which is often large in the application of dynamic pricing. Ideally, we would expect that a good policy will have an expected revenue loss which grows relatively slowly with respect to θ. Notationwise, we will use $f(\theta) = O(g(\theta))$ to mean that $f(\theta) \le M_1 g(\theta)$ for some constant $M_1 > 0$ and for all large n. Likewise, $f(\theta) = \Theta(g(\theta))$ means that there exist constants $0 < M_2 < M_3$ such that $M_2 g(\theta) \le f(\theta) \le M_3 g(\theta)$ for large enough n and $f(\theta) = \Omega(g(\theta))$ means that there exists a constant $M_4 > 0$ such that $f(\theta) \ge M_4 g(\theta)$ for all large n. For notational simplicity, whenever there is no confusion, we will often suppress the dependency on θ.

3. Main Results. In this section, we first introduce a generalization of the standard bisection search heuristic to a stochastic and constrained problem. We then discuss two improvements of the basic bisection heuristic to further reduce the asymptotic revenue loss bound. (The proofs of these results can be found in Section 4.)

3.1. Preliminary ideas. The departure point for the construction of our heuristics is a structural property of the optimal solution of the deterministic problem (2). It is known (e.g., Gallego and van Ryzin [22]) that the optimal deterministic policy is a static price control where the firms apply the same price $p^D = \max\{p^u, p^c\}$ until stock-out, where $p^c = \text{argmin}_{p\in\Omega_p} |\lambda(p) - C/T|$. For analytical tractability, we will assume that $\lambda(\bar{p}) < C/T$, which implies $p^c = p(C/T)$. (This is the original static price control in Gallego and van Ryzin [22] and can be easily satisfied, for example, if the feasible set Ω_p is sufficiently large.) Intuitively, the static control prescribes that the firms apply the unconstrained optimal price if inventory is abundant, and the clearance price if inventory is scarce. If the firm knows p^D and applies it to the stochastic pricing problem until the inventory is depleted, then it incurs a revenue loss of order $O(\sqrt{\theta})$ (Gallego and van Ryzin [22]). Jasin [25] show that this bound cannot be improved in general, i.e., the revenue loss of static price policy is $\Theta(\sqrt{\theta})$. Motivated by the good performance of static price policy in the case where p^D is known, one fruitful idea that has been exploited in the literature (e.g., Besbes and Zeevi [6]; Wang et al. [38]) is to design an algorithm whose resulting price sequence converges to p^D in the long run. In this paper, we will follow the same strategy and try to efficiently estimate p^D.

3.2. Heuristic #1: Generalized Bisection Search. The key idea behind our first heuristic is to generalize the classical bisection search into a stochastic setting with constraint. Before presenting the complete algorithm for our heuristic, we first define a price experimentation subroutine that will be repeatedly used throughout the paper. We parametrize the subroutine with $I \subset [\underline{p}, \bar{p}]$ and $N \in \mathbb{R}$, where I denotes the sampling price range and N denotes the sampling frequency.

Bisection Sampling Subroutine. $BiSamp(I, N)$

a. Divide I into 3 intervals of equal length.
   Let $S := \{p_l, l = 1, 2, 3, 4\}$ be the resulting endpoints of each interval.

b. For each l, apply p_l for N consecutive periods.

c. Compute the empirical mean rates

$$\hat{r}(p_l) = \frac{\text{total revenue incurred by } p_l}{N} \quad \text{and}$$
$$\hat{\lambda}(p_l) = \frac{\text{total demand incurred by } p_l}{N}, \quad l = 1, 2, 3, 4$$

Note that $\hat{r}(\cdot)$ denotes the empirical revenue rate and $\hat{\lambda}(\cdot)$ denotes the empirical demand rate. The complete algorithm for our first heuristic is given below.

Bisection Dynamic Pricing Algorithm (BDPA).

Step 1: Initialization
Define $\underline{p}_1 = \underline{p}$, $\bar{p}_1 = \bar{p}$ and $I_1 = [\underline{p}_1, \bar{p}_1]$ to be the starting interval.

Step 2: Shrinking the Interval
For $k = 1, ..., \tau_\theta$, do:

a. Execute $BiSamp(I_k, N_{k,\theta})$ as long as the inventory level is still positive.
   If the inventory is depleted, apply p_∞ until time T_θ.

b. **If** $\hat{r}(p_{k,2}) < \hat{r}(p_{k,3})$, define $\underline{p}_{k+1} = p_{k,2}$, $\bar{p}_{k+1} = p_{k,4}$;
If $\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3})$ and $\hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta}$, define $\underline{p}_{k+1} = p_{k,1}$, $\bar{p}_{k+1} = p_{k,3}$;
If $\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3})$ and $\hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}$, define $\underline{p}_{k+1} = p_{k,2}$, $\bar{p}_{k+1} = p_{k,4}$;
If $\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3})$ and $|\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta}$, define $\underline{p}_{k+1} = p_{k,2}$, $\bar{p}_{k+1} = p_{k,4}$;

c. Define the price range for the next iteration $I_{k+1} = [\underline{p}_{k+1}, \bar{p}_{k+1}]$.

Step 3: Applying Near-Optimal Static Price
Apply $\hat{p}^D_\theta = \frac{1}{2}(\underline{p}_{\tau_\theta+1} + \bar{p}_{\tau_\theta+1})$ until the end of selling horizon. Apply p_∞ if inventory is depleted.

The above algorithm is defined by three groups of parameters: τ_θ, which is the total number of rounds of bisection search performed; $\Delta_{k,\theta}$, which serves as the tolerance level for stochastic error and will be elaborated in Section 4; and $N_{k,\theta}$, which denotes the sampling frequency. The value of these parameters must be carefully chosen. For example, if $N_{k,\theta}$ is too large, we would be spending too much time on sampling sub-optimal prices instead of converging to the optimal static price. If, on the other hand, $N_{k,\theta}$ is too small, we may not be able to accurately estimate the revenues and demand rates at different prices, which may lead to mis-identification of the optimal static price. If $\Delta_{k,\theta}$ is too large, we will not be able to know with a high enough probability whether certain price violates the capacity constraint; if $\Delta_{k,\theta}$ is too small, we will need to increase the sampling frequencies accordingly. Below, we provide an explicit choice of parameters that will be used in our analysis:

$$\tau_\theta = \sup\left\{ n \in \mathbb{N} \ : \ 4 \cdot \sum_{k=1}^{n} N_{k,\theta} \leq T_\theta \right\}, \ N_{k,\theta} = \left\lceil \left(\frac{3}{2}\right)^{4k} \log^2 T_\theta \right\rceil, \ \Delta_{k,\theta} = \left(\frac{2}{3}\right)^{2k} \log^{-1/4} T_\theta,$$

where $\lceil x \rceil = \inf\{y \geq x : \ y \in \mathbb{N}\}$. We make two observations: First, we define τ_θ to be the maximum number of full-rounds bisection search until the end of the selling season. Since the intervals generated by BDPA keep shrinking to the optimal static price with a high probability, such choice potentially has the smallest revenue loss. Second, the sampling frequencies $N_{k,\theta}$ are increasing in k, whereas the error tolerances $\Delta_{k,\theta}$ are decreasing in k. The reasoning behind these choices are intuitive: As the price interval shrinks, the revenue difference at two different prices within the interval decreases and yet the magnitude of stochastic noise does not change. Thus, more samples are needed to guarantee a more accurate estimate of the revenue rate, and smaller error tolerances are required. We state our first result below.

Theorem 1. *Under the aforementioned choice of parameters, we have:*

$$\mathcal{R}^{BDPA}_\theta = O(\theta^{3/4} \log^{1/2} \theta).$$

It is noteworthy that the performance guarantee in Theorem 1 is of the same order as the performance of nonparametric algorithm in Besbes and Zeevi [6]. This result, however, is not very satisfactory as there is still a big gap between the upper bound on the revenue loss and the theoretical lower bound of $\Omega(\theta)$. The reason behind this relatively poor performance is that BDPA tries to estimate p^u and p^c simultaneously and utilize the fact that p^D is the maximum of the two prices to estimate p^D. However, if we know the true identity of p^D, the original pricing problem can be simplified into either a unconstrained optimization problem (when $p^D = p^u$) or a root-finding problem (when $p^D = p^c$). Both problems can be solved in more loss-efficient manners than the original pricing problem. This enlightens us to first explore the identity of p^D, then exploit this identity using a more loss-efficient algorithm. The following two subsections are devoted to expanding this idea and achieve a better performance.

Remark 1. The iterative procedure in Step 2 helps us to shrink the size of price range while at the same time making sure that the new interval still contains the optimal static price. The key idea is to distinguish which of the three intervals does not contain p^u (or p^c) through revenue (or demand) rates comparison. To understand the reasoning behind the four scenarios in Step 2b, suppose that demand is deterministic and $p^D \in I_k$ for some $k \geq 1$. (In this case, the Bisection Sampling Routine gives us the true demand and revenue rate, i.e., $\hat{r}(p) = r(p)$, $\hat{\lambda}(p) = \lambda(p)$.) Now, if $r(p_{k,2}) < r(p_{k,3})$, by unimodality of $r(\cdot)$ we know that $p^u \geq p_{k,2}$. Then we know that $p^D = \max\{p^u, p^c\} \geq p_{k,2}$ and can safely delete $[p_{k,1}, p_{k,2})$ for the next round. This explains the intuition behind the first scenario. As for the second scenario, if $r(p_{k,2}) \geq r(p_{k,3})$, then $p^u \leq p_{k,3}$. Moreover, if $\lambda(p_{k,3}) < C/T - \Delta_{k,\theta}$, then $p^c \leq p_{k,3}$ (because $\lambda(\cdot)$ is decreasing). This implies that $p^D \leq p_{k,3}$ and, thus, we can safely delete $[p_{k,3}, p_{k,4})$ for the next round. If, on the other hand, $\lambda(p_{k,3}) \geq C/T - \Delta_{k,\theta}$, then for a sufficiently small $\Delta_{k,\theta}$, p^c belongs to a small region near $p_{k,3}$ such that $p^c \geq p_{k,2}$. Then we know $p^D = \max\{p^u, p^c\} \geq p_{k,2}$ and can safely delete $[p_{k,1}, p_{k,2})$ for the next round. This explains the intuition behind the third and fourth scenarios. If the demand observations are stochastic, as long as the empirical mean rates ($\hat{r}(\cdot)$ and $\hat{\lambda}(\cdot)$) are close enough to the true rates ($r(\cdot)$ and $\lambda(\cdot)$), we can infer the true order relationships with high probability. As an example, Figure 3.1 illustrates the intuition behind scenario 3. The black boxes in Figure 3.1(b) and (d) denote the ranges where $\hat{\lambda}(\cdot)$ and $\hat{r}(\cdot)$ fall with high probability, while Figure 3.1(a) and (c) show their respective deterministic counterparts. If N_k and Δ_k are well-chosen, the upper blue dotted line in Figure 3.1(b) will not cross the third box, and we can thus make correct prediction of the position of p^c. Also, in Figure 3.1(d), the prediction of the order relationship between $r(p_{k,2})$ and $r(p_{k,3})$ is correct as long as the middle two boxes do not overlap along the vertical axis. As a consequence, the shrinking strategy in stochastic setting (Figure 3.1(d)) is the same with those in deterministic setting (Figure 3.1(c)).

3.3. Heuristic #2: Double Bisection Search. It is important to note that, if $p^u \neq p^c$, then the functional behavior of $r(p)$ around p^u and p^c are different. To be precise, $r(p)$ is approximately quadratic around p^u and is approximately linear around p^c. This suggests that an efficient algorithm must take into account the distinction between p^u and p^c. Broadly speaking, our heuristics can be divided into two phases: (1) an *exploration* phase, during which we try to identify whether the optimal static price is p^u or p^c, and (2) an *exploitation* phase, during which we implement a more efficient search algorithm exploiting the identity of the optimal static price. For the exploration phase, we will use the generalized bisection search in BDPA. For the exploitation phase, we will use more efficient bisection search method depending on the identity of p^D distinguished by the exploration phase. The algorithm will accordingly generate a sequence of shrinking intervals that contain the optimal static price with a very high probability.

Double-Bisection Dynamic Pricing Algorithm (D-BDPA).

Step 1-2: Same as BDPA

Step 3: Identifying the Optimal Price
If $\hat{\lambda}(\underline{p}_{\tau_\theta+1}) < C/T - \Delta_{\tau_\theta,\theta}$, go to Step 4a; else, go to Step 4b.

Step 4a: Converge to p^u when $p^D = p^u > p^c$.
Define $I_1^u = [\underline{p}_1^u, \bar{p}_1^u] = I_{\tau_\theta+1}$. For $k = 1, ..., \tau_\theta^u$, do:

a. Execute $BiSamp(I_k^u, N_{k,\theta}^u)$.

b. **If** $\hat{r}(p_{k,2}^u) < \hat{r}(p_{k,3}^u)$, define $\underline{p}_{k+1}^u = p_{k,2}^u$, $\bar{p}_{k+1}^u = p_{k,4}^u$; else, define $\underline{p}_{k+1}^u = p_{k,1}^u$, $\bar{p}_{k+1}^u = p_{k,3}^u$.

c. Define the price range for next iteration $I_{k+1}^u = [\underline{p}_{k+1}^u, \bar{p}_{k+1}^u]$.

Apply $\hat{p}^D_\theta = \frac{1}{2}(\underline{p}^u_{\tau^u_\theta+1} + \bar{p}^u_{\tau^u_\theta+1})$ until the end of selling horizon. Apply p_∞ if inventory is depleted.

Step 4b: Converge to p^c when $p^D = p^c \geq p^u$.
Define $I^c_1 = [\underline{p}^c_1, \bar{p}^c_1] = I_{\tau_\theta+1}$. For $k = 1, ..., \tau^c_\theta$, do:

a. Execute $BiSamp(I^c_k, N^c_{k,\theta})$.

b. **If** $\hat{\lambda}(p^c_{k,2}) > C/T + \Delta^c_{k,\theta}$, define $\underline{p}^c_{k+1} = p^c_{k,2}$, $\bar{p}^c_{k+1} = p^c_{k,4}$; else, define $\underline{p}^c_{k+1} = p^c_{k,1}$, $\bar{p}^c_{k+1} = p^c_{k,3}$.

c. Define price range of next iteration $I^c_{k+1} = [\underline{p}^c_{k+1}, \bar{p}^c_{k+1}]$.

Apply $\hat{p}^D_\theta = \frac{1}{2}(\underline{p}^c_{\tau^c_\theta+1} + \bar{p}^c_{\tau^c_\theta+1})$ until the end of selling horizon. Apply p_∞ if inventory is depleted.

We introduce some more parameters: τ^u_θ, and τ^c_θ, which are the numbers of rounds of bisection search performed during exploitation phase (Step 4), respectively; $\Delta^c_{k,\theta}$, which serve as the tolerance level for stochastic error; and $N^c_{k,\theta}$ and $N^u_{k,\theta}$, which denote the sampling frequencies. As for the old parameters, we use the same $N_{k,\theta}$ and $\Delta_{k,\theta}$, but different τ_θ, since now the exploration phase only lasts for a few periods. Below, we provide an explicit choice of parameters which will be used in our analysis:

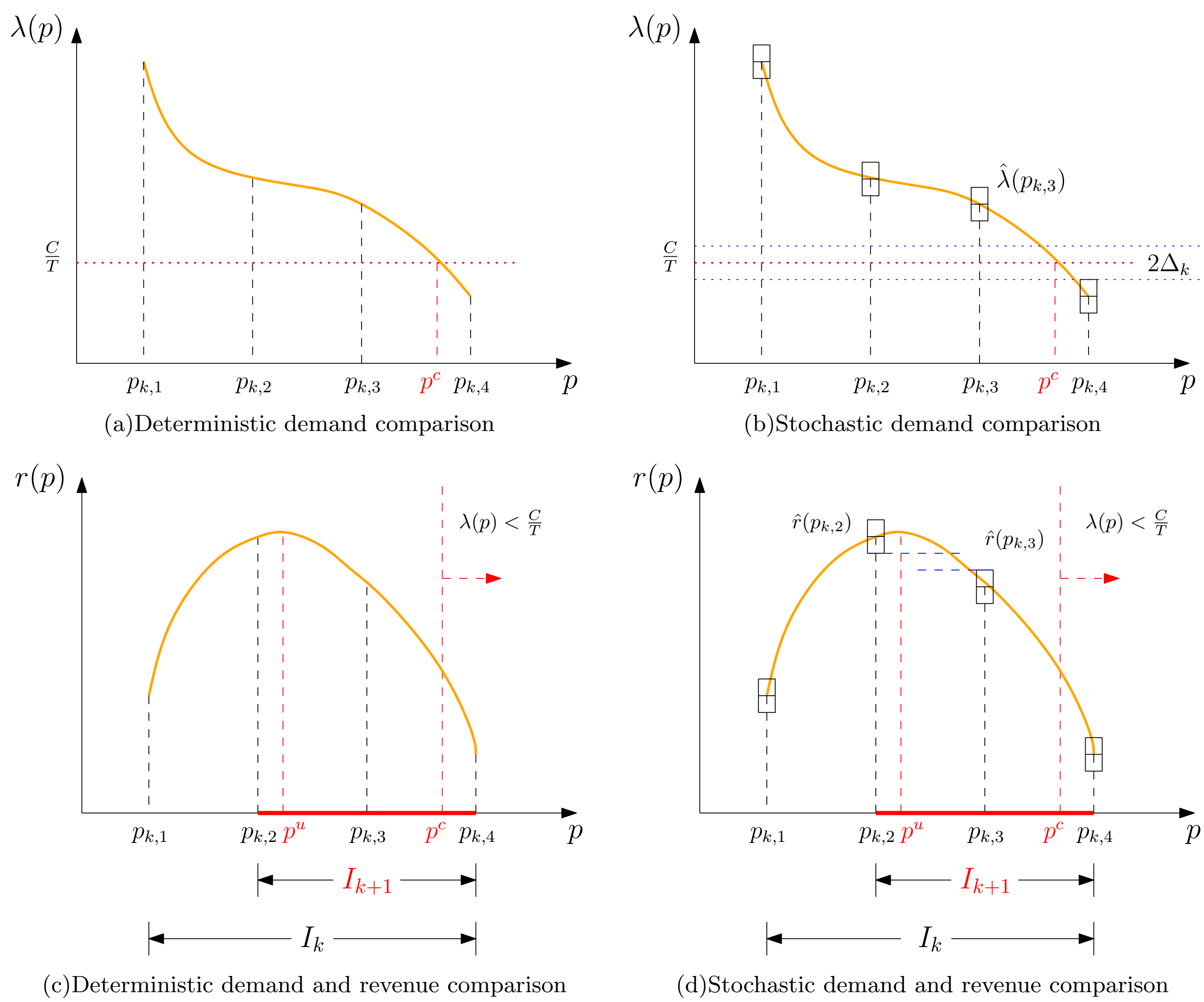

FIGURE 3.1. Convergence Behavior of Scenario 3, Step 2b in BDPA during k^{th} round.

$$\tau_\theta = \sup\left\{ n\in\mathbb{N} \ : \ 4\cdot\sum_{k=1}^{n} N_{k,\theta} \le \log^3 T_\theta \right\}$$
$$\tau_\theta^u = \sup\left\{ n\in\mathbb{N} \ : \ 4\cdot\sum_{k=1}^{n} N_{k,\theta}^u \le T_\theta - 4\cdot\sum_{k=1}^{n} N_{k,\theta} \right\},$$
$$\tau_\theta^c = \sup\left\{ n\in\mathbb{N} \ : \ 4\cdot\sum_{k=1}^{n} N_{k,\theta}^c \le T_\theta - 4\cdot\sum_{k=1}^{n} N_{k,\theta} \right\},$$
$$N_{k,\theta}^c = \left\lceil \left(\frac{3}{2}\right)^{2k} \log^2 T_\theta \right\rceil, \quad N_{k,\theta}^u = \left\lceil \left(\frac{3}{2}\right)^{4k} \log^3 T_\theta \right\rceil, \quad \Delta_{k,\theta}^c = \left(\frac{2}{3}\right)^{k} \log^{-3/8} T_\theta$$

We make several observations here. First, we set τ_θ such that the length of the exploration phase does not exceed $\log^3 T_\theta$, which is relatively short for large enough θ. This means that only a small number of price experimentations are needed to correctly identify (with a very high probability) whether $p^D = p^u$ or $p^D = p^c$. Secondly, the definitions of τ_θ^u and τ_θ^c follow from the fact that, during the exploration phase, we try to perform as many full-rounds of bisection search as possible until the end of the selling season. Thirdly, the sampling frequencies $(N_{k,\theta}^u, N_{k,\theta}^c)$ and tolerance of error $(\Delta_{k,\theta}^c)$ are different in exploitation phase comparing with those parameters in exploration phase$N_{k,\theta}$. These along with different shrinking strategy provide better performance. We state our result regarding the performance of D-BDPA below.

Theorem 2. *Under the aforementioned choice of parameters, we have:*

$$\mathcal{R}_\theta^{D-BDPA} = O(\sqrt{\theta}\log\theta).$$

Theorem 2 tells us that D-BDPA is asymptotically optimal. Moreover, its performance guarantee dominates the performance guarantee of any existing nonparametric algorithm in the literature, including the $O(\sqrt{\theta}\log^{4.5}\theta)$ of Wang et al. [38], and is very close to the known theoretical lower bound of $\Omega(\sqrt{\theta})$. In the next subsection we will show that if we replace the bisection search during the exploitation phase with Stochastic Approximation algorithm, then we can exactly match the lower bound.

3.4. Heuristic #3: Bisection Search and Stochastic Approximation. Stochastic Approximation refers to a class of iterative stochastic optimization algorithms. We refer to [30] for a comprehensive review. Broadly speaking, stochastic approximation algorithms can be divided into two different types: those that are try to solve a root-finding problem and those who try to stochastically estimate the maximum of a unimodal function. In this work, we consider the first and prototypical algorithms of this kind, i.e. Robbins-Monro (Robbins and Monro [34]) and Kiefer-Wolfowitz algorithms (Kiefer and Wolfowitz [28]). Let $R_t(p_t) = p_t \cdot D_t(p_t)$ denotes the realized revenue during period t under p_t, and define $P_X(x) = \arg\min_{y\in X} \ ||y-x||$ to be the geometric projection function. The complete description of the combined bisection search and Stochastic Approximation algorithm is given below.

SA-Bisection Dynamic Pricing Algorithm (SA-BDPA).

Steps 1 - 3: Same as D-BDPA

Step 4a: Converge to p^u when $p^u > p^c$. (Kiefer-Wolfowitz Scheme)
Let $p_1^u = \underline{p}_{\tau_\theta+1}$. For $k = 1, ..., \tau_\theta^u$, do:

a. Sample the revenue rate at price $p^u_k + c^u_k$ at period $4\sum_{k=1}^{\tau_\theta} N_k + 2k - 1$, and $p^u_k - c^u_k$ at period $4\sum_{k=1}^{\tau_\theta} N_k + 2k$ respectively; if inventory is depleted, apply p_∞.

b. Update the price according to

$$p^u_{k+1} = P_{I_{\tau_\theta+1}}\left[p^u_k + a^u_k \frac{R_k(p^u_k + c^u_k) - R_k(p^u_k - c^u_k)}{c^u_k}\right].$$

Step 4b: Converge to p^c when $p^c \geq p^u$. (Robbins-Monro Scheme)
Let $p^c_1 = \underline{p}_{\tau_\theta+1}$. For $k = 1, ..., \tau^c_\theta$, do:

a. Sample the revenue rate at price p^c_k for one period; if inventory is depleted, apply p_∞.

b. Update the price according to

$$p^c_{k+1} = P_{I_{\tau_\theta+1}}\left[p^c_k + a^c_k\left(\frac{C}{T} - D_k(p^c_k)\right)\right].$$

Note that SA-BDPA is parameterized by τ_θ, $\Delta_{k,\theta}$, $N_{k,\theta}$, a^u_k, a^c_k, and c^u_k. (The a^u_k, a^c_k, and c^u_k are standard parameters in Stochastic Approximation algorithm, see Broadie et al. (2011).) We state a theorem.

Theorem 3. *Under the same choice of τ_θ, $\Delta_{k,\theta}$, and $N_{k,\theta}$ as in Theorem 1 and a proper choice of a^u_k, a^c_k, and c^u_k, we have:*

$$\mathcal{R}^{SA-BDPA}_\theta = O(\sqrt{\theta}). \tag{4}$$

It is noteworthy that Besbes and Zeevi [6] also discuss a potential application of SA algorithms in their work. Specifically, they propose to apply the two types of SA schemes consecutively during the exploration phase to estimate p^u and p^c. At the end of the exploration phase, they propose that we choose the maximum of the two estimates and apply it during the remaining selling season until stock-out. The difference between their proposal and ours is obvious: They intend to use SA as an exploration algorithm while we use it as an exploitation algorithm. They conjecture that the revenue loss of their proposed SA-based dynamic pricing heuristic would be $O(\theta^{2/3})$, which is worse than ours.

4. Proof of Results. In this section, we will mainly discuss the proof of Theorem 2 and 3. We start by providing an outline of the proof in Section 4.1. The remaining details of the proof can be found in Sections 4.2 - 4.7 and in the Appendix at the end of this paper. As for the proof of Theorem 1, since it is very similar with proof of Theorem 2, we only discuss the outline briefly in Section 4.1.

4.1. Outline of the Proofs and Key Lemmas. We first discuss the outline of the proofs. For analytical convenience, we will consider a slightly modified pricing policy called *Modified D-BDPA* (MD-BDPA) and *Modified SA-BDPA* (MSA-BDPA), respectively, which operate exactly as D-BDPA and SA-BDPA with the exception that it does not apply p_∞ when the seller runs out of inventory. Under MD-BDPA and MSA-BDPA, any excess demand beyond the available inventory can be outsourced at a unit price of $2\bar{p}$. Since $p_t < 2\bar{p}$ for all $p_t \in [\underline{p}, \bar{p}]$, obviously, we have $J^{MD-BDPA} \leq J^{D-BDPA}$ and $J^{MSA-BDPA} \leq J^{SA-BDPA}$. Thus, in order to bound $J^* - J^{D-BDPA}$ and $J^* - J^{SA-BDPA}$, it suffices that we compute a bound for each $J^* - J^{MD-BDPA}$ and $J^* - J^{MSA-BDPA}$. The outline of the proof of Theorems 2 and 3 is as follows:

1. Bounding the Probability of Converging to p^D in Step 2

We compute a lower bound for the probability that the optimal deterministic price p^D lies in I_k for all k in Step 2. This is critical to ensure that the final interval in the exploration phase contains p^D with a high probability. Define $E_1 := \cap_{k=1}^{\tau_\theta+1}\{p^D \in I_k\}$. We state a lemma.

LEMMA 1. *Under the choice of parameters given in section 4.2, there exists a constant* $C_1 > 0$ *independent of* $\theta \geq 1$ *such that* $P(E_1) \geq 1 - C_1 \frac{(\log\log\theta)^2}{\theta}$.

The proof of Lemma 1 can be found in Section 4.2. It is not difficult to show that, after τ_θ rounds of bisection search in Step 2, the length of the remaining feasible price interval is of order $\log^{-1/4}\theta$ (see Section 3.2). So, Lemma 1 tells us that, by the end of the exploration phase, we are already sufficiently "close" to the optimal price (not close enough for us to ignore the exploitation phase and simply apply fixed price throughout the remaining selling horizon as in Besbes and Zeevi [6], but close enough for us to distinguish the identity of the optimal price).

2. Bounding the Probability of Distinguishing the Identity of p^D in Step 3

Once we guarantee that $p^D \in I_{\tau_\theta+1}$ with a high probability, we also need to guarantee that the action in Step 3 correctly distinguishes the identity of the optimal deterministic price with a high probability. If $p^D = p^u > p^c$, then we expect that the empirical demand rate at a point close to p^D will be much smaller than C/T. Similarly, if $p^D = p^c \geq p^u$, the empirical demand rate at a point close to p^D will be very close to C/T. Define $E_2 := \{\hat{\lambda}(\underline{p}_{\tau_\theta+1}) < C/T - \Delta_{\tau_\theta,\theta}\}$ if $p^u > p^c$ and $E_2 := \{\hat{\lambda}(\underline{p}_{\tau_\theta+1}) \geq C/T - \Delta_{\tau_\theta,\theta}\}$ otherwise. We state our second lemma.

LEMMA 2. *Under the choice of parameters given in section 3.2, there exists a constant* $C_2 > 0$ *independent of* $\theta \geq 1$ *such that* $P(E_1 \cap E_2) \geq 1 - C_2 \frac{(\log\log\theta)^2}{\theta}$.

The proof of Lemma 2 can be found in section 4.3.

3. Bounding the Revenue Loss in Step 4

After we know the identity of p^D, we can then properly bound the revenue loss incurred during the exploitation phase. Note that, by definition of τ_θ, the total revenue loss incurred during the exploitation phase is only $O(\log^3\theta)$. So, all that matters is the revenue loss incurred during the exploitation phase. In particular, by definition of $\pi \in \{\text{MD-BDPA, MSA-BDPA}\}$, we can write:

$$J_\theta^\pi = \mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] - 2\bar{p}\,\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right].$$

(Above, we suppress the notational dependency on π.) The bulk of the arguments in the rest of the analysis are in showing that

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] = r(p^D)\,T_\theta - O(\sqrt{\theta}\log\theta) \quad \text{(for Theorem 2)}$$

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] = r(p^D)\,T_\theta - O(\sqrt{\theta}) \quad \text{(for Theorem 3)}$$

$$\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] = O(\sqrt{\theta}) \quad \text{(for Theorems 2 and 3)},$$

which completes the proof. We now briefly explain how D-BDPA achieves this order of performance. (See section 4.4 and 4.5 for the parts regarding Theorem 2. We defer the proof of Theorem 3 in appendix since there are some similarities.) Assuming that the sequence of price intervals produced by D-BDPA converges to p^D, which happens with high probability. Since the exploration phase is relatively small, we can simply lower bound the collected revenue by zero. Now for the exploration phase, notice that if p^u is the optimal static price, the revenue function is relatively "flat" near p^u in the sense that it is approximately quadratic (see Lemma 3(i)). Hence, to correctly distinguish the order relationship of the demand rates at two different prices, we need to sample more, i.e. $N^u_{k,\theta} = \Theta((\frac{3}{2})^{4k}\log^2\theta)$. On the other hand, the convergence of revenue rate around p^u can be show to be quadratic (see Lemma 3(iii)). Now, assume without loss of generality that the selling season ends at the last period of the $(\tau^u_\theta)^{th}$ round of bisection search. Notice that $|I^u_k| = \Theta\left((\frac{2}{3})^{2k}\log^{-1/4}\theta\right)$ (see Section 4.4) and contains p^D with high probability, the revenue loss during Step 4a is of the order of

$$O\left(\sum_{k=1}^{\tau^u_\theta}\left(\frac{2}{3}\right)^{2k} N^u_{k,\theta}\right) = O\left(\sum_{k=1}^{\tau^u_\theta}\left(\frac{3}{2}\right)^{2k}\log^{3/2}\theta\right) = O\left(\left(\frac{3}{2}\right)^{2\tau^u_\theta}\log^{3/2}\theta\right) = O(\sqrt{\theta}\log\theta),$$

where the last inequality follows from Lemma 4. Now, if p^c is the optimal static price, the demand function is relatively "steep" near p^u in the sense that it is approximately linear (see Lemma 3(ii)). And accordingly we sample less frequently i.e. $N^c_{k,\theta} = \Theta((\frac{3}{2})^{2k}\log^2\theta)$. However, the convergence of revenue rate around p^c can be show to be linear (see Lemma 3(iii)), which is slower comparing with the case that $p^D = p^u$. Again, notice that $|I^c_k| = \Theta\left((\frac{2}{3})^{2k}\log^{-1/4}\theta\right)$ (see Section 4.4) and contains p^D with high probability, the revenue loss during Step 4n is of the order of

$$O\left(\sum_{k=1}^{\tau^c_\theta}\left(\frac{2}{3}\right)^{k} N^c_{k,\theta}\right) = O\left(\sum_{k=1}^{\tau^c_\theta}\left(\frac{3}{2}\right)^{k}\log^{7/4}\theta\right) = O\left(\left(\frac{3}{2}\right)^{\tau^c_\theta}\log^{7/4}\theta\right) = O(\sqrt{\theta}\log\theta),$$

where the last inequality follows from Lemma 4.

Building upon the intuition, we briefly explain the intuition behind the order of the performance guarantee of BDPA. Notice that BDPA executes bisection search without distinguishing the identity of p^D. As a consequence, it has to sample with higher frequency ($N_{k,\theta} = N^u_{k,\theta} > N_{k,\theta}$, since $r(p)$ is flat around p^u) without knowing if the revenue convergence rate is quadratic ($p^D = p^u$) or linear ($p^D = p^c$). Then, if the optimal price is p^c, BDPA will clearly suffer from oversampling. Quantitatively speaking, the revenue loss of BDPA is of the order of

$$O\left(\sum_{k=1}^{\tau^{BDPA}_\theta} N_{k,\theta}\left(\frac{2}{3}\right)^{k}\right) = O\left(\sum_{k=1}^{\tau^{BDPA}_\theta}\left(\frac{3}{2}\right)^{3k}\log^2\theta\right) = O\left(\left(\frac{3}{2}\right)^{3\tau^{BDPA}_\theta}\log^2\theta\right) = O(\theta^{3/4}\log^{1/2}\theta),$$

where $\tau^{BDPA}_\theta = \sup\{n\in\mathbb{N} : 4\sum_{k=1}^{n} N_{k,\theta} \le T_\theta\}$ is the rounds of bisection search performed in BDPA and satisfies $(\frac{3}{2})^{\tau^{BDPA}_\theta} = \Theta(\theta^{1/4}\log^{-1/2}\theta)$.

Below, we state two lemmas that will be repeatedly used in the proof.

Lemma 3. *(i) There exists a constant $K_u > 0$ such that for all $p_a, p_b \in [\underline{p}, \bar{p}]$, if $p^u > p_a > p_b$ (or $p_b > p_a > p^u$), then $r(p_a) - r(p_b) \geq K_u(p_a - p_b)^2$.*
(ii) For any $p_a, p_b \in [\underline{p}, \bar{p}]$, we have $|\lambda(p_a) - \lambda(p_b)| \geq L|p_a - p_b|$ for some positive constant L.
(iii) For any $p \in [\underline{p}, \bar{p}]$, we have $r(p^u) - r(p) \leq \frac{M_u K^2}{2}(p^u - p)^2$ and $r(p^c) - r(p) \leq (1 + 2K\bar{p})|p^c - p|$.

Lemma 4. *The following identities hold: $\tau_\theta = \Theta(\log\log\theta)$, $\tau_\theta^u = \Theta(\log\theta)$, and $\tau_\theta^c = \Theta(\log\theta)$. Moreover,*

$$\left(\frac{3}{2}\right)^{\tau_\theta} = \Theta\left(\log^{1/4}\theta\right), \quad \left(\frac{3}{2}\right)^{4\tau_\theta^u} = \Theta\left(\frac{\theta}{\log^3\theta}\right), \quad \textit{and} \quad \left(\frac{3}{2}\right)^{2\tau_\theta^c} = \Theta\left(\frac{\theta}{\log^2\theta}\right).$$

The first two parts of the first lemma tells us the "distinctiveness" of the revenue and demand function. They will provide useful guidelines for the choice of sampling frequencies. The third part of the first lemma provides upper bounds on the revenue loss depending on the identity of p^D. The second lemma quantifies the exact order of τ_θ, τ_θ^u, and τ_θ^c.

4.2. Proof of Lemma 1. By De Morgan's law and sub-additivity of probability measure, we have

$$P\left(\bar{E}_1\right) = P(\cup_{k=1}^{\tau_\theta+1}\{p^D \notin I_k\}) \leq \sum_{k=1}^{\tau_\theta+1} P(p^D \notin I_k),$$

where $\bar{E}$ is the complement of E. For $k > 1$, we can bound:

$$\begin{aligned} P(p^D \notin I_k) &= P(p^D \notin I_k | p^D \in I_{k-1})P(p^D \in I_{k-1}) + P(p^D \notin I_k | p^D \notin I_{k-1})P(p^D \notin I_{k-1}) \\ &\leq P(p^D \notin I_k,\, p^D \in I_{k-1}) + P(p^D \notin I_{k-1}) \\ &\leq \cdots \\ &\leq \sum_{j=1}^{k-1} P(p^D \notin I_{j+1},\, p^D \in I_j) \end{aligned}$$

where the last inequality follows from $P(p^D \notin I_1) = 0$. Substituting them back into the bound for $P(\bar{E}_1)$ and using the fact that $P(p^D \notin I_1) = 0$, we get:

$$P\left(\bar{A}_1\right) \leq \sum_{k=2}^{\tau_\theta+1}\sum_{j=1}^{k-1} P(p^D \notin I_{j+1},\, p^D \in I_j) = \sum_{k=1}^{\tau_\theta}(\tau_\theta - k + 1)P(p^D \notin I_{k+1},\, p^D \in I_k).$$

We will now proceed to bound the term $P(p^D \notin I_{k+1},\ p^D \in I_k)$ for $k = 1, ..., \tau_\theta$. Define five groups of events $B_k^1, ..., B_k^5$ as follows:

$$\begin{aligned} B_k^1 &= \{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}),\, p^u < p_{k,2}\}, \\ B_k^2 &= \{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}),\, p^u > p_{k,3}\}, \\ B_k^3 &= \{\hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta},\, p^c > p_{k,3}\}, \\ B_k^4 &= \{\hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta},\, p^c < p_{k,3}\}, \\ B_k^5 &= \{|\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta},\, p^c < p_{k,2}\}. \end{aligned}$$

We claim that:

$$P(p^D \notin I_{k+1},\, p^D \in I_k) \leq \sum_{l=1}^{5} P(B_k^l), \ \forall k \tag{5}$$

To prove this, first, note that, per the description of our algorithm, there are four different cases in Step 2(b) that we can enter in round k. So, we can bound:

$$\begin{aligned}P(p^D \notin I_{k+1},\, p^D \in I_k) \le\; & P(\hat r(p_{k,2}) < \hat r(p_{k,3}),\, p^D \notin I_{k+1},\, p^D \in I_k)\\ &+P(\hat r(p_{k,2}) \ge \hat r(p_{k,3})\;,\; \hat\lambda(p_{k,3}) < C/T-\Delta_{k,\theta},\, p^D \notin I_{k+1},\, p^D \in I_k)\\ &+P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}),\, \hat\lambda(p_{k,3}) > C/T+\Delta_{k,\theta},\, p^D \notin I_{k+1},\, p^D \in I_k)\\ &+P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}),\, |\hat\lambda(p_{k,3}) - C/T| \le \Delta_{k,\theta},\, p^D \notin I_{k+1},\, p^D \in I_k).\end{aligned}$$

Now, if $p^D = p^u > p^c$, we have:

$$\begin{aligned}&P\left(\hat r(p_{k,2}) < \hat r(p_{k,3}),\, p^u \notin I_{k+1},\, p^u \in I_k\right)\\ &\quad= P(\hat r(p_{k,2}) < \hat r(p_{k,3}),\, p^u \in [p_{k,1}, p_{k,2}),\, p^u \in I_k)\\ &\quad\le P(\hat r(p_{k,2}) < \hat r(p_{k,3}),\, p^u < p_{k,2},\, p^u \in I_k)\\ &\quad\le P(B_k^1).\end{aligned}$$

$$\begin{aligned}&P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}),\, \hat\lambda(p_{k,3}) < C/T-\Delta_{k,\theta},\, p^u \notin I_{k+1},\, p^u \in I_k)\\ &\quad\le P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), p^u \in (p_{k,3}, p_{k,4}],\, p^u \in I_k)\\ &\quad\le P(B_k^2).\end{aligned}$$

$$\begin{aligned}&P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), \hat\lambda(p_{k,3}) > C/T+\Delta_{k,\theta}, p^u \notin I_{k+1},\, p^u \in I_k)\\ &\quad= P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), \hat\lambda(p_{k,3}) > C/T+\Delta_{k,\theta}, p^u \in [p_{k,1}, p_{k,2}),\, p^u \in I_k)\\ &\quad\le P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), \hat\lambda(p_{k,3}) > C/T+\Delta_{k,\theta}, p^u < p_{k,2})\\ &\quad\le P(\hat\lambda(p_{k,3}) > C/T+\Delta_{k,\theta}, p^c < p_{k,2}) \quad \text{(because } p^D = p^u > p^c\text{)}\\ &\quad\le P(B_k^4).\end{aligned}$$

$$\begin{aligned}&P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), |\hat\lambda(p_{k,3}) - C/T| \le \Delta_{k,\theta}, p^u \notin I_{k+1},\, p^u \in I_k)\\ &\quad= P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), |\hat\lambda(p_{k,3}) - C/T| \le \Delta_{k,\theta}, p^u \in [p_{k,1}, p_{k,2}),\, p^u \in I_k)\\ &\quad\le P(|\hat\lambda(p_{k,3}) - C/T| \le \Delta_{k,\theta}, p^c < p_{k,2}) \quad \text{(because } p^D = p^u > p^c\text{)}\\ &\quad\le P(B_k^5).\end{aligned}$$

If, on the other hand, $p^D = p^c \ge p^u$, we have:

$$\begin{aligned}&P(\hat r(p_{k,2}) < \hat r(p_{k,3}), p^c \notin I_{k+1},\, p^c \in I_k)\\ &\quad= P(\hat r(p_{k,2}) < \hat r(p_{k,3}), p^c \in [p_{k,1}, p_{k,2}),\, p^c \in I_k)\\ &\quad\le P(\hat r(p_{k,2}) < \hat r(p_{k,3}), p^c < p_{k,2})\\ &\quad\le P(\hat r(p_{k,2}) < \hat r(p_{k,3}), p^u < p_{k,2}) \quad \text{(because } p^D = p^c \ge p^u\text{)}\\ &\quad= P(B_k^1).\end{aligned}$$

$$\begin{aligned}&P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), \hat\lambda(p_{k,3}) < C/T-\Delta_{k,\theta}, p^c \notin I_{k+1},\, p^c \in I_k)\\ &\quad\le P(\hat\lambda(p_{k,3}) < C/T-\Delta_{k,\theta}, p^c \in (p_{k,3}, p_{k,4}],\, p^c \in I_k)\\ &\quad\le P(B_k^3).\end{aligned}$$

$$\begin{aligned}&P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), \hat\lambda(p_{k,3}) > C/T+\Delta_{k,\theta}, p^c \notin I_{k+1},\, p^c \in I_k)\\ &\quad= P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), \hat\lambda(p_{k,3}) > C/T+\Delta_{k,\theta}, p^c \in [p_{k,1}, p_{k,2}),\, p^c \in I_k)\\ &\quad\le P(\hat\lambda(p_{k,3}) > C/T+\Delta_{k,\theta}, p^c < p_{k,2})\\ &\quad= P(B_k^4).\end{aligned}$$

$$\begin{aligned}&P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), |\hat\lambda(p_{k,3}) - C/T| \le \Delta_{k,\theta}, p^c \notin I_{k+1},\, p^c \in I_k)\\ &\quad= P(\hat r(p_{k,2}) \ge \hat r(p_{k,3}), |\hat\lambda(p_{k,3}) - C/T| \le \Delta_{k,\theta}, p^c \in [p_{k,1}, p_{k,2}),\, p^c \in I_k)\\ &\quad\le P(|\hat\lambda(p_{k,3}) - C/T| \le \Delta_{k,\theta}, p^c < p_{k,2})\\ &\quad= P(B_k^5).\end{aligned}$$

Thus, in either case (i.e., $p^D = p^u > p^c$ or $p^D = p^c \geq p^u$), the bound in (5) holds. Put this together with our earlier bound for $P\left(\bar{A}_1\right)$, we get:

$$P\left(\bar{A}_1\right) \leq \sum_{k=1}^{\tau_\theta} (\tau_\theta - k + 1) \left[\sum_{l=1}^{5} P(B_k^l)\right].$$

To complete the proof of Lemma 1, it suffices that we compute a bound for $P(B_k^l)$ for $k = 1, ..., \tau_\theta$, $l = 1, ..., 5$, which is our remaining focus.

Upper bound for $\mathbf{P}(B_k^1)$ and $\mathbf{P}(B_k^2)$

The probabilities $P(B_k^1)$ and $P(B_k^2)$ can be bounded in a similar manner. So, we will only show how to bound $P(B_k^1)$. Fix $k \in \{1, ..., \tau_\theta\}$. Note that $p^u < p_{k,2} < p_{k,3}$ on B_1^k. Then by Lemma 3 part (ii), on B_k^1, $r(p_{k,2}) - r(p_{k,3}) \geq K_u(p_{k,3} - p_{k,2})^2 = K_u(\frac{|I|}{3})^2(\frac{2}{3})^{2(k-1)}$. Since $|\hat{r}(p_{k,l}) - r(p_{k,l})| < \frac{1}{4}K_u(\frac{|I|}{3})^2(\frac{2}{3})^{2(k-1)}$ for $l \in \{2,3\}$ implies

$$\begin{aligned}
\hat{r}(p_{k,2}) - \hat{r}(p_{k,3}) &= (r(p_{k,2}) - r(p_{k,3})) + (\hat{r}(p_{k,2}) - r(p_{k,2})) - (\hat{r}(p_{k,3}) - r(p_{k,3})) \\
&\geq (r(p_{k,2}) - r(p_{k,3})) - |\hat{r}(p_{k,2}) - r(p_{k,2})| - |\hat{r}(p_{k,3}) - r(p_{k,3})| \\
&> K_u \left(\frac{|I|}{3}\right)^2 \left(\frac{2}{3}\right)^{2(k-1)} - \frac{2}{4} K_u \left(\frac{|I|}{3}\right)^2 \left(\frac{2}{3}\right)^{2(k-1)} > 0,
\end{aligned}$$

by Hoeffding's inequality (Hoeffding [24]), we can bound

$$\begin{aligned}
P(B_k^1) &\leq P\left(|\hat{r}(p_{k,l}) - r(p_{k,l})| \geq \frac{1}{4} K_u \left(\frac{|I|}{3}\right)^2 \left(\frac{2}{3}\right)^{2(k-1)} \quad \text{for some } l \in \{2,3\}\right) \\
&\leq \sum_{l=2}^{3} P\left(|\hat{r}(p_{k,l}) - r(p_{k,l})| \geq \frac{1}{4} K_u \left(\frac{|I|}{3}\right)^2 \left(\frac{2}{3}\right)^{2(k-1)}\right) \\
&\leq 4 \exp\left(-2 \frac{N_{k,\theta} \ \frac{1}{4^2} \ K_u^2 \ \left(\frac{|I|}{3}\right)^4 \left(\frac{2}{3}\right)^{4(k-1)}}{\bar{p}^2}\right).
\end{aligned}$$

By definition, $N_{k,\theta} = \Theta(\left(\frac{3}{2}\right)^{4k} \log^2 \theta)$. So, for all sufficiently large θ, $P(B_k^1) \leq \frac{4}{\theta}$. The same bound also holds for $P(B_k^2)$.

Upper bound for $\mathbf{P}(B_k^3)$ and $\mathbf{P}(B_k^4)$

The probabilities $P(B_k^3)$ and $P(B_k^4)$ can be bounded in a similar manner. So, we will only show how to bound $P(B_k^3)$. Note that $p^c > p_{k,3}$ implies $\lambda(p_{k,3}) > C/T$. So,

$$\begin{aligned}
P(B_k^3) &\leq P\left(\hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta} \ , \lambda(p_{k,3}) > C/T\right) \\
&\leq P\left(\hat{\lambda}(p_{k,3}) - \lambda(p_{k,3}) < -\Delta_{k,\theta}\right) \\
&\leq P\left(|\hat{\lambda}(p_{k,3}) - \lambda(p_{k,3})| > \Delta_{k,\theta}\right).
\end{aligned}$$

Again, by Hoeffding's inequality, since $\Delta_{k,\theta} = \Theta((\frac{2}{3})^{2k} \log^{-1/4} \theta)$ and $N_{k,\theta} = \Theta(\left(\frac{3}{2}\right)^{4k} \log^2 \theta)$, for all large θ, we have $P(|\hat{\lambda}(p_{k,3}) - \lambda(p_{k,3})| \geq \Delta_{k,\theta}) \leq 2\exp(-2N_{k,\theta}\Delta_{k,\theta}^2) \leq \frac{2}{\theta}$. The same bound also holds for $P(B_k^4)$.

Upper bound for $\mathbf{P}(B_k^5)$

By the decreasing property of demand function, $p^c < p_{k,2}$ implies $\lambda(p_{k,2}) \le C/T$. By Lemma 3 part (i), $\lambda(p_{k,2}) - \lambda(p_{k,3}) \ge L|p_{k,2} - p_{k,3}| = L\frac{|I|}{3}(\frac{2}{3})^{k-1}$. So, on B_k^5,

$$\begin{aligned}\lambda(p_{k,3}) - \hat{\lambda}(p_{k,3}) &\le \lambda(p_{k,2}) - L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1} - \left(\frac{C}{T} - \Delta_{k,\theta}\right)\\ &\le \frac{C}{T} - L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1} - \left(\frac{C}{T} - \Delta_{k,\theta}\right)\\ &\le -\frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1},\end{aligned}$$

where the last inequality follows because, by definition, $\Delta_{k,\theta} \le \frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}$ for all sufficiently large θ. Now, by similar arguments as above,

$$\begin{aligned}P(B_k^5) &\le P\left(\lambda(p_{k,3}) - \hat{\lambda}(p_{k,3}) < -\frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}\right)\\ &\le P\left(|\hat{\lambda}(p_{k,3}) - \lambda(p_{k,3})| > \frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}\right)\\ &\le 2\exp\left(-2N_k\left[\frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}\right]^2\right) \le \frac{2}{\theta} \quad \text{(for all sufficiently large } \theta) .\end{aligned}$$

Put all the bounds together, we have

$$P\left(\bar{E}_1\right) \le \sum_{k=1}^{\tau_\theta}(\tau_\theta - k + 1)\left[\sum_{l=1}^{5} P(B_k^l)\right] \le \frac{\tau_\theta(\tau_\theta+1)}{2}\cdot\frac{4}{\theta}\cdot 5 = \frac{10\tau_\theta(\tau_\theta+1)}{\theta}.$$

Since $\tau_\theta = \Theta(\log\log\theta)$ (see Lemma 4), we conclude that there exists a constant C_1 such that

$$P(E_1) = 1 - P\left(\bar{E}_1\right) \ge 1 - C_1\frac{(\log\log\theta)^2}{\theta}. \qquad \square$$

4.3. Proof of Lemma 2. The proof is similar to that of Lemma 1. We will analyze the two cases (i.e., $p^D = p^u > p^c$ and $p^D = p^c \ge p^u$) separately.

Case 1: $p^D = p^c \ge p^u$

If $p^u \le p^c$, then the optimal deterministic price p^D equals p^c. On E_1, we know that $p^c = p^D \in [\underline{p}_{\tau_\theta+1}, \bar{p}_{\tau_\theta+1}]$. This implies $\lambda(\underline{p}_{\tau_\theta+1}) \ge \lambda(p^c) = C/T$. So, we can bound:

$$\begin{aligned}1 - P(E_1 \cap E_2) &= 1 - P(E_1) + P(E_1) - P(E_1 \cap E_2)\\ &= P\left(\bar{E}_1\right) + P\left(E_1 \cap \bar{E}_2\right)\\ &\le P\left(\bar{E}_1\right) + P\left(p^c \in [\underline{p}_{\tau_\theta+1}, \bar{p}_{\tau_\theta+1}],\, \hat{\lambda}(\underline{p}_{\tau_\theta+1}) < C/T - \Delta_{\tau_\theta,\theta}\right)\\ &\le P\left(\bar{E}_1\right) + P\left(\lambda(\underline{p}_{\tau_\theta+1}) \ge C/T,\ \hat{\lambda}(\underline{p}_{\tau_\theta+1}) < C/T - \Delta_{\tau_\theta,\theta}\right)\\ &\le P\left(\bar{E}_1\right) + P\left(\hat{\lambda}(\underline{p}_{\tau_\theta+1}) - \lambda(\underline{p}_{\tau_\theta+1}) < -\Delta_{\tau_\theta,\theta}\right)\\ &\le P\left(\bar{E}_1\right) + P\left(|\hat{\lambda}(\underline{p}_{\tau_\theta+1}) - \lambda(\underline{p}_{\tau_\theta+1})| > \Delta_{\tau_\theta,\theta}\right)\\ &\le P\left(\bar{E}_1\right) + 2\exp\left(-2N_{\tau_\theta,\theta}\ \Delta^2_{\tau_\theta,\theta}\right) \quad \text{(by Hoeffding's inequality)}\\ &\le C_1\frac{(\log\log\theta)^2}{\theta} + \frac{2}{\theta} \quad \text{(by Lemma 1)},\end{aligned}$$

where the last inequality holds for all sufficiently large θ.

Case 2: $p^D = p^u > p^c$

If $p^u > p^c$, then $p^D = p^u$ and $\lambda(p^u) < \lambda(p^c) = C/T$. By definition of τ_θ, $|I_{\tau_\theta+1}|$ and $\Delta_{\tau_\theta,\theta}$ decrease to zero as $\theta \to \infty$. Since we always have $p^u = p^D \in [\underline{p}_{\tau_\theta+1}, \bar{p}_{\tau_\theta+1}]$ on A_1, it must also hold for all sufficiently large θ on A_1 that $p^c < \underline{p}_{\tau_\theta+1} < p^u$, $\lambda(\underline{p}_{\tau_\theta+1}) - \lambda(p^u) \le (\lambda(p^c) - \lambda(p^u))/4$, and $\Delta_{\tau_\theta,\theta} \le (\lambda(p^c) - \lambda(p^u))/4$. Arguing as in case 1, for all large θ, we can bound:

$$
\begin{aligned}
1 - P(E_1 \cap E_2) &= P\left(\bar{E}_1\right) + P\left(E_1 \cap \bar{E}_2\right) \\
&\le P\left(\bar{E}_1\right) + P\left(\max\left\{\Delta_{\tau_\theta,\theta}, \lambda(\underline{p}_{\tau_\theta+1}) - \lambda(p^u)\right\} \le (\lambda(p^c) - \lambda(p^u))/4,\ \hat{\lambda}(\underline{p}_{\tau_\theta+1}) \ge C/T - \Delta_{\tau_\theta,\theta}\right) \\
&\le P\left(\bar{E}_1\right) + P\left(\max\left\{\Delta_{\tau_\theta,\theta}, \lambda(\underline{p}_{\tau_\theta+1}) - \lambda(p^u)\right\} \le (\lambda(p^c) - \lambda(p^u))/4,\ \hat{\lambda}(\underline{p}_{\tau_\theta+1}) \ge \lambda(p^c) - \Delta_{\tau_\theta,\theta}\right) \\
&\le P\left(\bar{E}_1\right) + P\left(\hat{\lambda}(\underline{p}_{\tau_\theta+1}) - \lambda(\underline{p}_{\tau_\theta+1}) \ge (\lambda(p^c) - \lambda(p^u))/2\right) \\
&\le C_1 \frac{(\log\log\theta)^2}{\theta} + \frac{1}{\theta},
\end{aligned}
$$

where the last inequality follows by Lemma 1 and Hoeffding's inequality (for sufficiently large θ).

Put the bounds from case 1 and case 2 together, we conclude that there exists a constant $C_2 > 0$ independent of $\theta \ge 1$ such that $P(E_1 \cap E_2) \ge 1 - C_2 \frac{(\log\log\theta)^2}{\theta}$. $\square$

4.4. Bounding the Revenue Loss of D-BDPA Upon Entering Step 4a. Since $p^D = p^u > p^c$, for all sufficiently large θ, the following two conditions must hold: (i) $p^c \notin I_1^u$ and (ii) $r(p)$ is strictly concave in $I_1^u = I_{\tau_\theta+1}$. The first condition holds because p^u is strictly larger than p^c and the interval $I_{\tau_\theta+1}$ can be arbitrarily small for large θ. The second condition follows from the fact that $r(p)$ is locally strictly concave in the neighborhood of p^u (see Lemma 3 part (i)).

Let $E_u := \cap_{k=1}^{\tau_\theta^u}\{p^u \in I_k^u\}$. The following lemma is analogous to Lemma 1.

Lemma 5. *There exists a constant $C_3 > 0$ such that $P(E_1 \cap E_2 \cap E_u) \ge 1 - C_3 \frac{(\log\theta)^2}{\theta}$.*

We defer the proof of Lemma 5 to the appendix. Per our discussions in Section 4.1, the net revenue of MD-BDPA is the direct revenue minus the penalty, i.e.,

$$J_\theta^{MD-BDPA} = \mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] - 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right].$$

We will now proceed to bound the two expectations separately.

Step 1: Lower Bound for Direct Revenue Collected by MD-BDPA

We claim that there exists a constant $\tilde{C}_1 > 0$ such that

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \ge\ r(p^u)T_\theta - \tilde{C}_1\sqrt{\theta}\log\theta.$$

We focus our analysis on the sample path in $E_1 \cap E_2 \cap E_u$. Define $\tilde{T}_{\theta,1}^u = \sum_{k=1}^{\tau_\theta} 4N_{k,\theta}$ and $\tilde{T}_{\theta,2}^u = \sum_{k=1}^{\tau} 4N_{k,\theta} + \sum_{k=1}^{\tau_\theta} 4N_{k,\theta}^u$. The collected revenue can be lower bounded by two components as follows:

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \ge \mathbb{E}\left[\sum_{t=1+\tilde{T}_{\theta,1}^u}^{\tilde{T}_{\theta,2}^u} p_t D_t(p_t)\mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right]$$

$$\geq \mathbb{E}\left[\sum_{k=1}^{\tau_\theta^u}\sum_{l=1}^{4} N_{k,\theta}^u \ \hat{r}(p_{k,l}^u)\mathbf{1}\{E_1\cap E_2\cap E_u\}\right] + \mathbb{E}\left[\left(T_\theta - \tilde{T}_{\theta,2}^u\right)\hat{r}(\hat{p}^D)\mathbf{1}\{E_1\cap E_2\cap E_u\}\right]. \qquad (6)$$

For the first term, note that

$$\mathbb{E}\left[\sum_{k=1}^{\tau_\theta^u}\sum_{l=1}^{4} N_{k,\theta}^u \ \hat{r}(p_{k,l}^u) \ \middle| \ E_1\cap E_2\cap E_u\right] = \sum_{k=1}^{\tau_\theta^u}\sum_{l=1}^{4} N_{k,\theta}^u \ \mathbb{E}\left[r(p_{k,l}^u) | \ E_1\cap E_2\cap E_u\right].$$

Since on event $E_1\cap E_2\cap E_u$, $|p_{k,l}^u - p^u| \leq |I_1^u|(\frac{2}{3})^{k-1}$, then by Lemma 3(iii) we know that $r(p_{k,l}^u) \geq r(p^u) - \frac{9M_UK^2}{8}|I_1^u|^2\left(\frac{2}{3}\right)^{2k}$. Put this together with Lemma 5 and the fact that $\sum_{k=1}^{\tau_\theta^u} 4N_{k,\theta}^u \geq \tilde{T}_{\theta,2}^u - \log^3 T_\theta$, we have

$$\begin{aligned}
&\sum_{k=1}^{\tau_\theta^u}\sum_{l=1}^{4} N_{k,\theta}^u \mathbb{E}\left[r(p_{k,l}^u)|E_1\cap E_2\cap E_u\right] P(E_1\cap E_2\cap E_u)\\
&\geq \left[\sum_{k=1}^{\tau_\theta^u} 4N_{k,\theta}^u\left(r(p^u) - \frac{9M_UK^2}{8}|I_1^u|^2\left(\frac{2}{3}\right)^{2k}\right)\right]\left(1 - C_3\frac{(\log\theta)^2}{\theta}\right)\\
&\geq r(p^u)\left(\tilde{T}_{\theta,2}^u - \log^3 T_\theta\right) - C_3 \ \bar{p} \ \frac{\log^2\theta}{\theta}\left(\sum_{k=1}^{\tau_\theta^u} 4N_{k,\theta}^u\right) - \frac{9M_UK^2}{8}|I_1^u|^2\left[\sum_{k=1}^{\tau_\theta^u} 4N_{k,\theta}^u\left(\frac{2}{3}\right)^{2k}\right]\\
&\geq r(p^u)\tilde{T}_{\theta,2}^u - \bar{p}\log^3 T_\theta - C_3\bar{p}T\log^2\theta - \frac{81}{10}M_UK^2|I_1^u|^2\log^3 T_\theta\left(\frac{3}{2}\right)^{2(\tau_\theta^u+1)}\\
&\geq r(p^u)\tilde{T}_{\theta,2}^u - \Theta(\sqrt{\theta}\log\theta),
\end{aligned}$$

where the last inequality follows because $|I_1^u| = \Theta(\log^{-1/4}\theta)$ and $\left(\frac{3}{2}\right)^{4\tau_\theta^u} = \Theta\left(\frac{\theta}{\log^3\theta}\right)$ (see Lemma 4).

As for the second term in the RHS of (6), by the same arguments as above,

$$\begin{aligned}
&\mathbb{E}\left[\left(T_\theta - \tilde{T}_{\theta,2}^u\right)\hat{r}(\hat{p}^D)\mathbf{1}\{E_1\cap E_2\cap E_u\}\right]\\
&\geq \left(T_\theta - \tilde{T}_{\theta,2}^u\right)\left(r(p^u) - \frac{9M_UK^2}{8}|I_1^u|^2\left(\frac{2}{3}\right)^{2\tau_\theta^u}\right)\left(1 - C_3\frac{(\log\theta)^2}{\theta}\right)\\
&\geq r(p^u)\left(T_\theta - \tilde{T}_{\theta,2}^u\right) - C_3 \ \bar{p} \ T\log^2\theta - \frac{9M_UK^2}{8}|I_1^u|^2 \ T_\theta\left(\frac{2}{3}\right)^{2\tau_\theta^u}\\
&\geq r(p^u)\left(T_\theta - \tilde{T}_{\theta,2}^u\right) - \Theta(\sqrt{\theta}\log\theta),
\end{aligned}$$

where the last inequality follows because $|I_1^u| = \Theta(\log^{-1/4}\theta)$ and $\left(\frac{3}{2}\right)^{4\tau_\theta^u} = \Theta\left(\frac{\theta}{\log^3\theta}\right)$. Put the bounds for the two terms together proves our initial claim.

Step 2: Upper Bound for Total Penalty Incurred by Capacity Violation

We claim that there exists a constant $\tilde{C}_2 > 0$ such that

$$2\bar{p} \ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \leq \tilde{C}_2\sqrt{\theta}.$$

We first analyze the sample path on $E_1 \cap E_2 \cap E_u$. We know that

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] \\
&\le \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] + \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} \lambda(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] \\
&\le \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \mathbb{E}\left[\left(\sum_{t=1}^{\tilde{T}^u_{\theta,1}} \lambda(p_t) - \tilde{T}^u_{\theta,1}\frac{C}{T}\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] \\
&\quad + \mathbb{E}\left[\left(\sum_{k=1}^{\tau^u_\theta}\sum_{l=1}^{4} N^u_{k,\theta}\lambda(p^u_{k,l}) - \left(\tilde{T}^u_{\theta,2} - \tilde{T}^u_{\theta,1}\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] \\
&\quad + \mathbb{E}\left[\left(\sum_{t=\tilde{T}^u_{\theta,2}+1}^{T_\theta} \lambda(p_t) - \left(T_\theta - \tilde{T}^u_{\theta,2}\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] \\
&\le \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \tilde{T}^u_{\theta,1} + \mathbb{E}\left[\left(\sum_{k=1}^{\tau^u_\theta}\sum_{l=1}^{4} N^u_{k,\theta}\left(\lambda(p^u_{k,l}) - \frac{C}{T}\right)\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] \\
&\quad + \mathbb{E}\left[\left(\sum_{t=\tilde{T}^u_{\theta,2}+1}^{T_\theta} \lambda(p_t) - \left(T_\theta - \tilde{T}^u_{\theta,2}\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] \\
&\le \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \log^3 T_\theta + \sum_{k=1}^{\tau^u_\theta}\sum_{l=1}^{4} N^u_{k,\theta}\;\mathbb{E}\left[\left(\lambda(p^u_{k,l}) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] \\
&\quad + \mathbb{E}\left[\left(T_\theta - \tilde{T}^u_{\theta,2}\right)\left(\lambda\left(\hat{p}^D\right) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right],
\end{aligned}
$$

where the first and second inequalities follow from Jensen's Inequality; the third inequality follows from the boundedness of demand observation and the definition of $\tilde{T}^u_{\theta,1}$, $\tilde{T}^u_{\theta,2}$; the last inequality follows from Jensen's Inequality and the definition of τ_θ. Basically, we break the capacity violation into four parts: stochastic randomness, and the capacity violation during Step 2, during bisection search in Step 4 and applying $\hat{p}^D$ in Step 4.

By Cauchy-Schwarz's inequality and the boundedness of demand observation, the first term can be easily bounded as follows:

$$
\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D(p_t) - \lambda(p_t)\right)^+\right] \le \left\{\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D(p_t) - \lambda(p_t)\right)^2\right]\right\}^{1/2} = \left\{\sum_{t=1}^{T_\theta}\mathbb{E}\left[\left(D(p_t) - \lambda(p_t)\right)^2\right]\right\}^{1/2} \le \sqrt{T_\theta}.
$$

As for the third term, since $p^u > p^c$, which implies $\lambda(p^u) < \lambda(p^c) = C/T$, and $p^c \notin I^u_k$ for all k (for all large θ), we always have $\lambda(p^u_{k,l}) < \lambda(p^c) = C/T$. So, $(\lambda(p^u_{k,l}) - C/T)^+ = 0$ for all k and l. Similarly, since $\hat{p}^D \in I^u_{\tau^u_\theta+1}$, we have $\lambda(\hat{p}^D) < C/T$ for all large θ. So, the last term also equals to 0. Put the bounds together we have:

$$
\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_u\}\right] = O(\sqrt{\theta}).
$$

Thus, the total penalty for capacity violation satisfies

$$2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right]$$
$$= 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] + 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{\overline{E_1 \cap E_2 \cap E_u}\}\right]$$
$$\le 2\bar{p}\ O\left(\sqrt{\theta}\right) + 2\bar{p}\ T_\theta P\left(\overline{E_1 \cap E_2 \cap E_u}\right) = O(\sqrt{\theta}),$$

where the last inequality follows the boundedness of demand observation.

Finally, combining our results from **Steps 1** and **2** above we conclude that

$$J_\theta^{MD-BDPA} \ \ge\ r(p^u)T_\theta - \tilde{C}_1\sqrt{\theta}\log\theta - \tilde{C}_2\sqrt{\theta} \ =\ r(p^u)T_\theta - O(\sqrt{\theta}\log\theta). \quad \square$$

4.5. Bounding the Revenue Loss of D-BDPA Upon Entering Step 4b. The proof is similar to those in section 4.2. Let $E_c = \cap_{k=1}^{\tau_\theta^c}\{p_c \in I_k^c\}$. The following lemma is the analog of Lemma 5.

LEMMA 6. *There exists a constant $C_4 > 0$ such that $P\left(\ \overline{E_1 \cap E_2 \cap E_c}\ \right) \ge 1 - C_4\frac{(\log\theta)^2}{\theta}$.*

We defer the proof of Lemma 6 to the appendix. We again consider MD-BDPA. The net revenue generated by MD-BDPA is given by:

$$J_\theta^{MD-BDPA} \ge \mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] - 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right].$$

Step 1: Lower Bound for Direct Revenue Collected by MD-BDPA

We claim that there exists a constant $\tilde{C}_3 > 0$ such that

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \ge r(p^c)T_\theta - \tilde{C}_3\sqrt{\theta}\log\theta.$$

The proof is similar to Step 1 in section 4.1. We break up the revenue on the sample path of $E_1 \cap E_2 \cap E_c$ into two parts:

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \ge \mathbb{E}\left[\sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c\ \hat{r}(p_{k,l}^c)\mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right] + \mathbb{E}\left[\left(T_\theta - \tilde{T}_{\theta,2}^c\right)\hat{r}(\hat{p}^D)\mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right], \tag{7}$$

where $\tilde{T}_{\theta,1}^c = \sum_{k=1}^{\tau_\theta} 4N_{k,\theta}$ and $\tilde{T}_{\theta,2}^c = \sum_{k=1}^{\tau_\theta} 4N_{k,\theta} + \sum_{k=1}^{\tau_\theta^c} 4N_{k,\theta}^c$. For the first term, note that

$$\mathbb{E}\left[\sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c \hat{r}(p_{k,l}^c)\middle|\, E_1 \cap E_2 \cap E_c\right] \ =\ \sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c\ \mathbb{E}\left[r(p_{k,l}^c)\middle|\, E_1 \cap E_2 \cap E_c\right].$$

Since on $E_1 \cap E_2 \cap E_c$, $|p^c_{k,l} - p^c| \le \frac{3|I_1^c|}{2}(\frac{2}{3})^k$, by Lemma 3(iii) we know that $r(p^c) - r(p^c_{k,l}) \le \frac{3}{2}(1 + 2K\bar{p})|I_1^c|\left(\frac{2}{3}\right)^k$. Put this together with Lemma 6 and the fact that $\sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N^c_{k,\theta} \ge \tilde{T}^c_{\theta,2} - \log^3 T_\theta$ we have

$$\begin{aligned}
&\sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N^c_{k,\theta}\ \mathbb{E}\left[r(p^c_{k,l})\middle|\, E_1 \cap E_2 \cap E_c\right] P(E_1 \cap E_2 \cap E_c) \\
&\quad \ge \sum_{k=1}^{\tau_\theta^c} 4N^c_{k,\theta}\left[r(p^c) - \frac{3}{2}(1+2K\bar{p})|I_1^c|\left(\frac{2}{3}\right)^k\right]\left[1 - C_4\frac{\log^2\theta}{\theta}\right] \\
&\quad \ge r(p^c)(\tilde{T}^c_{\theta,2} - \log^3 T_\theta) - \bar{p}\ C_4\frac{\log^2\theta}{\theta}\left(\sum_{k=1}^{\tau_\theta^c} 4N^c_{k,\theta}\right) - \frac{3}{2}(1+2K\bar{p})|I_1^c|\left[\sum_{k=1}^{\tau_\theta^c} 4N^c_{k,\theta}\left(\frac{2}{3}\right)^k\right] \\
&\quad \ge r(p^c)\tilde{T}^c_{\theta,2} - \bar{p}\log^3 T_\theta - \bar{p}TC_4\log^2\theta - 18(K\bar{p}+1)|I_1^c|\log^2 T_\theta\left(\frac{3}{2}\right)^{\tau_\theta^c} \\
&\quad = r(p^c)\tilde{T}^c_{\theta,2} - O(\sqrt{\theta}\log\theta),
\end{aligned}$$

where the last inequality follows since $|I_1^c| = \Theta(\log^{-1/4}\theta)$ and $\left(\frac{3}{2}\right)^{2\tau_\theta^c} = \Theta\left(\frac{\theta}{\log^2\theta}\right)$, or equivalently $\left(\frac{3}{2}\right)^{\tau_\theta^c} = \Theta\left(\frac{\sqrt{\theta}}{\log\theta}\right)$. (See Lemma 4)

As for the second term in the RHS of (7), by the same argument as above,

$$\begin{aligned}
&\mathbb{E}\left[\left(T_\theta - \tilde{T}^c_{\theta,2}\right)\hat{r}(\hat{p}^D)\mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right] \\
&\quad \ge \left(T_\theta - \tilde{T}^c_{\theta,2}\right)\left[r(p^c) - \frac{3}{2}(1+2K\bar{p})|I_1^c|\left(\frac{2}{3}\right)^{\tau_\theta^c}\right]\left(1 - C_4\frac{\log^2\theta}{\theta}\right) \\
&\quad \ge r(p^c)\left(T_\theta - \tilde{T}^c_{\theta,2}\right) - C_4\ \bar{p}\ T\log^2\theta - \frac{3}{2}(1+2K\bar{p})|I_1^c|\ T_\theta\left(\frac{2}{3}\right)^{\tau_\theta^c} \\
&\quad \ge r(p^c)\left(T_\theta - \tilde{T}^u_{\theta,2}\right) - O(\sqrt{\theta}\log\theta),
\end{aligned}$$

where the last inequality follows since $|I_1^c| = \Theta(\log^{-1/4}\theta)$ and $\left(\frac{3}{2}\right)^{2\tau^c} = \Theta\left(\frac{\theta}{\log^2\theta}\right)$. Put the bounds for the two terms in together proves the initial claim.

Step 2: Upper Bound for Total Penalty Incurred by Capacity Violation

We claim that there exists a constant $\tilde{C}_4 > 0$ such that

$$2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \le \tilde{C}_4\sqrt{\theta}\log\theta. \tag{8}$$

We first analyze the sample path on $E_1 \cap E_2 \cap E_c$. We break the amount of capacity violation into several different parts. Following the same arguments as in Step 2 in section 4.2,

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right] \\
&\quad \le \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \log^3 T_\theta + \sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N^c_{k,\theta}\ \mathbb{E}\left[\left(\lambda(p^c_{k,l}) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right]
\end{aligned}$$

$$+\ \mathbb{E}\left[\left(T_\theta - \tilde{T}^c_{\theta,2}\right)\left(\lambda\left(\hat{p}^D\right)-\frac{C}{T}\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_c\}\right].$$

By Cauchy-Schwarz's inequality again, the first term can be upper bounded by $\sqrt{T_\theta}$. Then since for the sample paths on event $E_1\cap E_2\cap E_c$, $|\lambda(p^c_{k,l})-\lambda(p^c)|\le \frac{3}{2}K|I^c_1|(\frac{2}{3})^k$ for all k and l and $\lambda(p^c)=C/T$, we can bound

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t)-C_\theta\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_c\}\right]\\
&\le \sqrt{T_\theta}+\log^3 T_\theta+\sum_{k=1}^{\tau^c_\theta} 4\cdot N^c_{k,\theta}\cdot\frac{3}{2}K|I^c_1|\left(\frac{2}{3}\right)^k+\left(T_\theta-\tilde{T}^c_{\theta,2}\right)\cdot\frac{3}{2}K|I^c_1|\left(\frac{2}{3}\right)^{\tau^c_\theta}\\
&\le \sqrt{T_\theta}+\log^3 T_\theta+18K|I^c_1|\left(\frac{3}{2}\right)^{\tau^c_\theta}\log^2 T_\theta+\frac{3}{2}KT_\theta|I^c_1|\left(\frac{2}{3}\right)^{\tau^c_\theta}\\
&= O(\sqrt{\theta}\log\theta),
\end{aligned}$$

where the last inequality the same argument as in **Step 1** above.

Then, the total penalty for capacity violation satisfies

$$\begin{aligned}
&2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t)-C_\theta\right)^+\right]\\
&=2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t)-C_\theta\right)^+ \mathbf{1}\{E_1\cap E_2\cap E_c\}\right]+2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t)-C_\theta\right)^+ \mathbf{1}\{\overline{E_1\cap E_2\cap E_c}\}\right]\\
&\le O\left(\sqrt{\theta}\log\theta\right)+2\bar{p}\ T_\theta P(\overline{E_1\cap E_2\cap E_c})=O(\sqrt{\theta}\log\theta).
\end{aligned}$$

Finally, combining our results from **Step 1** and **2** above we have

$$J^{MD-BDPA}_\theta \ \ge\ r(p^u)T_\theta-\tilde{C}_3\sqrt{\theta}\log\theta-\tilde{C}_4\sqrt{\theta}\log\theta\ =\ r(p^c)T_\theta-O(\sqrt{\theta}\log\theta).\quad\square$$

5. Conclusion and Future Work. This paper presents a scheme of nonparametric dynamic pricing with demand learning. Our scheme generalizes the classical bisection search algorithm into a stochastic setting with a constraint. We show that the performance of one of our heuristics exactly matches the theoretical lower bound for any feasible pricing policy. Thus, we have closed the gap (in asymptotic sense) between the performance of parametric approach and nonparametric approach for the single product problem.

There are several possible extensions of this work. One important direction is a generalization to the multiproduct setting. Although we have focused our analysis in the paper only on the single product setting, it is an open question whether our bisection search heuristic can also be applied to multiproduct problem. There are at least two challenges for such an extension: First, it is not immediately clear how to do bisection in high dimensional spaces. To the best of our knowledge, there is no existing literature on applying bisection search to multidimensional constrained optimization problem, even in the deterministic setting. Second, in multiproduct setting, nonparametric approach might suffers from *curse of dimensionality*, since it has to estimate a multidimensional function. In fact, the the order of the revenue loss of the best known nonparametric scheme for multiproduct setting depends on the number of products in a non-trivial way (cf. Besbes and Zeevi [7]). It is curious to see whether applying bisection search algorithm to multiproduct setting can reduce the curse of dimensionality on revenue loss.

Additionally, throughout the paper, we have assumed that the demand function is stationary, i.e., it does not vary with time. In reality however, this assumption might not hold, which suggests that a good pricing heuristic should ideally take into account this possibility in its learning algorithm. The challenge, however, is obvious. For dynamic pricing with non-stationary demand, it is no longer true that the optimal solution to the deterministic problem is static pricing. This limits the ability to exploit the structure of the optimal solution, as we did in this paper. Actually, all of the works in non-stationary setting (Besbes et al. [5], Keskin and Zeevi [26]) consider only the problem without inventory constraint. Moreover, it is not clear how one can generalize the bisection search heuristic to non-stationary setting. Obviously this is an important research topic; we leave this as future research project.

Appendix A: Proof of Theorem 3.

A.1. Bounding the Revenue Loss in SA-BDPA Upon Entering Step 4a. Following the same arguments as in the proof of Theorem 2, we know that

$$J_\theta^{MSA-BDPA} \geq \mathbb{E}\left[\sum_{k=1}^{\tau_\theta^u}\left[R_k(p_k^u+c_k^u)+R_k(p_k^u-c_k^u)\right]\mathbf{1}\left\{E_1\cap E_2\right\}\right] - 2\bar{p}\,\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t)-C_\theta\right)^+\right] \tag{9}$$

where $\tau_\theta^u := \left\lfloor \frac{T_\theta - 4\sum_{k=1}^{\tau} N_k}{2}\right\rfloor$ and E_1 and E_2 are as defined in Lemma 1 and Lemma 2, respectively.

We start with bounding the first term, which is the direct revenue incurred by MSA-DPA. Note that, for all p, we have $r(p^u)-r(p)\leq \frac{M_U K^2}{2}(p^u-p)^2$. So,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{k=1}^{\tau_\theta^u}\left[R_k(p_k^u+c_k^u)+R_k(p_k^u-c_k^u)\right]\mathbf{1}\left\{E_1\cap E_2\right\}\right]\\
&\geq \sum_{k=1}^{\tau_\theta^u}\mathbb{E}\left[r(p_k^u+c_k^u)\mathbf{1}\left\{E_1\cap E_2\right\}\right]+\sum_{k=1}^{\tau_\theta^u}\mathbb{E}\left[r(p_k^u-c_k^u)\mathbf{1}\left\{E_1\cap E_2\right\}\right]\\
&\geq \sum_{k=1}^{\tau_\theta^u}\mathbb{E}\left[r(p^u)-\frac{M_UK^2}{2}(p^u-p_k^u-c_k^u)^2\mathbf{1}\left\{E_1\cap E_2\right\}\right]+\sum_{k=1}^{\tau_\theta^u}\mathbb{E}\left[r(p^u)-\frac{M_UK^2}{2}(p^u-p_k^u+c_k^u)^2\mathbf{1}\left\{E_1\cap E_2\right\}\right]\\
&\geq 2\ \tau_\theta^u\ r(p^u)-2M_UK^2\left[\sum_{k=1}^{\tau_\theta^u}E\left[(p^u-p_k^u)^2\mathbf{1}\left\{E_1\cap E_2\right\}\right]+(c_k^u)^2\right]\\
&\geq r(p^u)T_\theta-\bar{p}\ (2+\log^3 T_\theta)-2M_UK^2\left[\sum_{k=1}^{\tau_\theta^u}E\left[(p^u-p_k^u)^2\mathbf{1}\left\{E_1\cap E_2\right\}\right]+(c_k^u)^2\right],
\end{aligned}$$

where the last inequality follows because, by definition of τ_θ and τ_θ^u, we have $2\tau_\theta^u \geq T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta} - 2$. As for the second term in (9), which is the total penalty incurred by capacity violation, similar to the arguments in Step 2 in section 4.2, for sample paths on $E_1\cap E_2$, we can bound

$$\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t)-C_\theta\right)^+\mathbf{1}\{E_1\cap E_2\}\right]$$

$$\leq \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \mathbb{E}\left[\left(\sum_{t=1}^{4\sum_{k=1}^{\tau_\theta} N_{k,\theta}} \lambda(p_t) - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right]$$
$$+\mathbb{E}\left[\left(\sum_{t=4\sum_{k=1}^{\tau_\theta} N_{k,\theta}+1}^{T_\theta} \lambda(p_t) - \left(T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right]$$
$$\leq \sqrt{T_\theta} + \log^3 T_\theta + 0 = O(\sqrt{\theta}),$$

where the third inequality follows from Cauchy-Schwarz inequality, the definition of τ_θ, and the fact that p^c is to the left of I_k^u for all k.

Then, the total penalty for capacity violation satisfies

$$2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right]$$
$$= 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] + 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{\overline{E_1 \cap E_2}\}\right]$$
$$\leq\ O(\sqrt{\theta}) + 2\bar{p}\ T_\theta P(\overline{E_1 \cap E_2}) = O(\sqrt{\theta}).$$

Finally, combining the bounds for the two terms in (9), we get

$$J_\theta^{MSA-BDPA} \geq r(p^u)T_\theta - 2M_U K^2 \left[\sum_{k=1}^{\tau_\theta^u} \mathbb{E}\left[(p^u - p_k^u)^2 \mathbf{1}\{E_1 \cap E_2\}\right] + (c_k^u)^2\right] - O(\sqrt{\theta}).$$

Applying the standard result in Stochastic Approximation (e.g. Proposition 1 in Broadie et al. [10]), there exists positive constants C_a^u and C_c^u such that if $a_k^u = C_a/k$ and $c_k^u = C_c/k^{1/4}$ we have $\mathbb{E}[(p^u - p_k^u)^2 \mathbf{1}\{E_1 \cap E_2\}] \leq C_u/\sqrt{k}$, for all $k \geq 1$, where $C_u > 0$ is also a constant. Substitute this into the above bound, we get

$$\begin{aligned} J_\theta^{MSA-BDPA} &\geq r(p^u)T_\theta - 2M_U K^2 \sum_{k=1}^{\tau_\theta^u}\left(\frac{C_u}{\sqrt{k}} + \frac{C_c^2}{\sqrt{k}}\right) - O(\sqrt{\theta}) \\ &\geq r(p^u)T_\theta - 2M_U K^2 (C_u + C_c^2)\sqrt{\tau_\theta^u} - O(\sqrt{\theta}) \\ &\geq r(p^u)T_\theta - O(\sqrt{\theta}). \quad \square \end{aligned}$$

A.2. Bounding the Revenue Loss in SA-BDPA Upon Entering for Step 4b. Following the same arguments as in Step 1 in section 4.2, we know that

$$J_\theta^{MSA-BDPA}\ \geq\ \mathbb{E}\left[\sum_{k=1}^{\tau_\theta^c} R_k(p_k^c)\mathbf{1}\{E_1 \cap E_2\}\right] - 2\bar{p}\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \tag{10}$$

where $\tau_\theta^c := T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}$. For the first term in (10), note that $r(p^c) - r(p_k^c) \leq (1 + K\bar{p})|p^c - p_k^c|$. So, we can bound

$$\mathbb{E}\left[\sum_{k=1}^{\tau_\theta^c} R_k(p_k^c)\mathbf{1}\{E_1 \cap E_2\}\right]\ = \sum_{k=1}^{\tau_\theta^c} \mathbb{E}\left[r(p_k^c)\mathbf{1}\{E_1 \cap E_2\}\right]$$

$$
\begin{aligned}
&\geq \sum_{k=1}^{\tau_\theta^c} \mathbb{E}\left[\{r(p^c) - (1+K\bar{p})|p^c - p_k^c|\}\ \mathbf{1}\{E_1 \cap E_2\}\right] \\
&\geq r(p^c)T_\theta - \bar{p}\log^3 T_\theta - (1+K\bar{p})\sum_{k=1}^{\tau_\theta^c} \mathbb{E}\left[|p^c - p_k^c|\,\mathbf{1}\{E_1 \cap E_2\}\right] \\
&\geq r(p^c)T_\theta - \bar{p}\log^3 T_\theta - (1+K\bar{p})\sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(p_k^c - p^c)^2\,\mathbf{1}\{E_1 \cap E_2\}\right]},
\end{aligned}
$$

where the second inequality follows by definition of τ_θ^c and the last inequality follows from Jensen's inequality. As for the second term in (10), following the same arguments as in Step 2 in section 4.2, we know that for the sample paths on $E_1 \cap E_2$,

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] \\
&\quad \leq \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \mathbb{E}\left[\left(\sum_{t=1}^{4\sum_{k=1}^{\tau_\theta} N_{k,\theta}} \lambda(p_t) - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] \\
&\qquad + \mathbb{E}\left[\left(\sum_{t=4\sum_{k=1}^{\tau_\theta} N_{k,\theta}+1}^{T_\theta} \lambda(p_t) - \left(T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] \\
&\quad \leq \sqrt{T_\theta} + \log^3 T_\theta + \sum_{k=1}^{\tau_\theta^c} \mathbb{E}\left[\left(\lambda(p_k^c) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] \\
&\quad \leq O(\sqrt{\theta}) + \sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(\lambda_k(p_k^c) - \lambda(p^c))^2\,\mathbf{1}\{E_1 \cap E_2\}\right]} \\
&\quad \leq O(\sqrt{\theta}) + K\sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(p_k^c - p^c)^2\,\mathbf{1}\{E_1 \cap E_2\}\right]}.
\end{aligned}
$$

Thus, the total penalty for capacity violation satisfies

$$
\begin{aligned}
&2\bar{p}\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \\
&\quad = 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] + 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{\overline{E_1 \cap E_2}\}\right] \\
&\quad \leq O(\sqrt{\theta}) + 2\bar{p}\ K\sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(p_k^c - p^c)^2\,\mathbf{1}\{E_1 \cap E_2\}\right]} + 2\bar{p}\ T_\theta P(\overline{E_1 \cap E_2}) \\
&\quad = O(\sqrt{\theta}) + 2\bar{p}\ K\sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(p_k^c - p^c)^2\,\mathbf{1}\{E_1 \cap E_2\}\right]}.
\end{aligned}
$$

Combining the results above we get

$$
J_\theta^{MSA-BDPA} \ \geq\ r(p^c)T_\theta - O(\sqrt{\theta}) - (1+3K\bar{p})\sum_{k=1}^{\tau_\theta^c} \sqrt{E\left[(p_k^c - p^c)^2\,\mathbf{1}\{E_1 \cap E_2\}\right]}.
$$

Applying the established convergence result for Robbins-Monro type of Stochastic Approximation, by Theorem 1 in the electronic companion in Broadie et al. [10], we know that there exists positive constant C_a^c such that when $a_k^c = C_a^c/k$, we have $E[(p_k^c - p^c)^2 \mathbf{1}\{E_1 \cap E_2\}] \le C_c/k$, for all $k \ge 1$, where $C_c > 0$ is also a constant. Substitute this back into the previous bound, we have

$$\begin{aligned} J_\theta^{MSA-BDPA} &\ge r(p^c)T_\theta - \Theta(\sqrt{\theta}) - (1+3K\bar{p})\sum_{k=1}^{\tau_\theta^c}\sqrt{C_c/k} \\ &\ge r(p^c)T_\theta - \Theta(\sqrt{\theta}) - (1+3K\bar{p})\sqrt{C_c\tau_\theta^c} = r(p^c)T_\theta - \Theta(\sqrt{\theta}). \quad \square \end{aligned}$$

Appendix B: Proof of Key Lemmas.

B.1. Proof of Lemma 3 (i) We assume without loss of generality that $p_b > p_a > p^u$. Let $\lambda_a = \lambda(p_a)$, $\lambda_b = \lambda(p_b)$, and we have $\lambda_b < \lambda_a < \lambda^u$ since demand is decreasing in price. Now, by Assumption A5, we know that (see Boyd and Vandenberghe [9])

$$\begin{aligned} r(p_a) - r(p_b) &= r(\lambda_a) - r(\lambda_b) \ge \frac{M_l}{2}(\lambda_b - \lambda_a)^2 - r'(\lambda_a)(\lambda_b - \lambda_a) \\ &\ge \frac{M_l L^2}{2}(p_b - p_a)^2 - r'(\lambda_a)(\lambda_b - \lambda_a) \\ &\ge \frac{M_l L^2}{2}(p_b - p_a)^2, \end{aligned}$$

where the first inequality follows from Assumption A5x, second inequality follows from Lemma 3 part (ii), and the third inequality follows from Assumption A2. Setting $K_u = \frac{M_l L^2}{2}$ completes the proof of part (i).

(ii) Follows directly from Assumption A1.

(iii) Let us denote $\lambda = \lambda(p)$. Notice that $r(\lambda)$ is strictly concave in λ, by Taylor's expansion, there exists $\xi \in [\lambda, \lambda^u]$ (or possibly $[\lambda^u, \lambda)$) such that

$$r(p) = r(\lambda) = r(\lambda^u) + r'(\lambda^u)(\lambda - \lambda^u) + \frac{r''(\xi)}{2}(\lambda - \lambda^u)^2 \ge r(p^u) - \frac{M_U}{2}(\lambda - \lambda^u)^2 \ge r(\lambda^u) - \frac{M_U K^2}{2}(p - p^u)^2,$$

where the first and the second inequalities follow by Assumptions A2 and A4, respectively.

As for the second part, we know that

$$\begin{aligned} r(p^c) - r(p) &= r(p^c) - (p^c + p - p^c)[\lambda(p^c) + \lambda(p) - \lambda(p^c)] \\ &= \lambda(p^c)(p^c - p) + p\ (\lambda(p^c) - \lambda(p)) - (p - p^c)(\lambda(p) - \lambda(p^c)) \\ &\le |p^c - p| + K\bar{p}\ |p^c - p| + K|p^c - p|^2 \\ &\le (1 + 2K\bar{p})|p^c - p|, \end{aligned}$$

where the first inequality follows from the boundedness of demand and price and Assumption A3. $\square$

B.2. Proof of Lemma 4. We start with τ_θ. Define:

$$t_1 = \left\lceil \frac{1}{4}\log_{3/2}\left(\frac{1}{6}\log T_\theta + 1\right)\right\rceil - 3 \quad \text{and} \quad t_2 = \left\lceil \frac{1}{4}\log_{3/2}\left(\frac{65}{324}\log T_\theta + 1\right)\right\rceil + 1.$$

Note that $t_1 < t_2$ when θ is large and they are both $\Theta(\log\log\theta)$. Moreover, we also have

$$\begin{aligned} 4\cdot\sum_{k=1}^{t_2} N_{k,\theta} &\ge 4\sum_{k=1}^{t_2}\left(\frac{3}{2}\right)^{4k}\log^2 T_\theta \;=\; \frac{324}{65}\left[\left(\frac{3}{2}\right)^{4t_2} - 1\right]\log^2 T_\theta \;>\; \log^3 T_\theta \quad \text{and} \\ 4\cdot\sum_{k=1}^{t_1+1} N_{k,\theta} &< \left[4\sum_{k=1}^{t_1+1}\left(\frac{3}{2}\right)^{4k}\log^2 T_\theta\right] + 4t_1 \;<\; \left(\frac{3}{2}\right)^{4(t_1+2)}\log^2 T_\theta + 4t_1 \\ &\le \frac{1}{6}\log^3 T_\theta + \log^2 T_\theta + \Theta(\log\log T_\theta) \;<\; \log^3 T_\theta \quad (\text{for all large } \theta). \end{aligned}$$

Since $\sum_{k=1}^{t} N_{k,\theta}$ is increasing in t, we must have $t_1 < \tau_\theta < t_2$. We conclude that $\tau_\theta = \Theta(\log\log\theta)$ and $(2/3)^{\tau_\theta} = \Theta(\log^{-1/4}\theta)$. We now calculate the order of τ_θ^u. Define:

$$t_1^u = \left\lceil \frac{1}{4}\log_{3/2}\left(\frac{65T_\theta}{648\log^3 T_\theta}+1\right)\right\rceil - 1 \text{ and } t_2^u = \left\lceil \frac{1}{4}\log_{3/2}\left(\frac{65T_\theta}{162\log^3 T_\theta}+1\right)\right\rceil.$$

By definition of τ_θ^u and $N_{k,\theta}^u$, for all large enough θ, we have

$$\begin{aligned}
4\sum_{k=1}^{t_1^u} N_{k,\theta}^u &\le 4\sum_{k=1}^{t_1^u}\left[\left(\frac{3}{2}\right)^{4k}\log^3 T_\theta + 1\right] \le 4t_1^u + \frac{324}{65}\left[\left(\frac{3}{2}\right)^{4t_1^u} - 1\right]\log^3 T_\theta \\
&\le \frac{1}{2}T_\theta + \Theta\left(\log\left(\frac{T_\theta}{\log^3 T_\theta}\right)\right) \le T_\theta - \log^3 T_\theta \le T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta} \qquad \text{and} \\
4\sum_{k=1}^{t_2^u} N_{k,\theta}^u &\ge 4\sum_{k=1}^{t_2^u}\left[\left(\frac{3}{2}\right)^{4k}\log^3 T_\theta - 1\right] \ge \frac{324}{65}\log^3 T_\theta\left[\left(\frac{3}{2}\right)^{4t_2^u} - 1\right] - 4t_2^u \\
&\ge 2T_\theta - \Theta\left(\log\left(\frac{T_\theta}{\log^3 T_\theta}\right)\right) \ge T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta},
\end{aligned}$$

which implies that $t_1^u \le \tau_\theta^u \le t_2^u$. Since t_1^u and t_2^u are both $\Theta(\log\theta)$, we conclude that $\tau_\theta^u = \Theta(\log\theta)$. Moreover, $(2/3)^{4\tau_\theta^u} = \Theta(\theta^{-1}\log^3\theta)$. Finally, we calculate τ_θ^c. Define:

$$t_1^c = \left\lceil \frac{1}{2}\log_{3/2}\left(\frac{5T_\theta}{72\log^2 T_\theta}+1\right)\right\rceil - 1 \text{ and } t_2^c = \left\lceil \frac{1}{2}\log_{3/2}\left(\frac{5T_\theta}{18\log^2 T_\theta}+1\right)\right\rceil.$$

By definition of τ_θ^c and $N_{k,\theta}^c$, for all large enough θ, we have

$$\begin{aligned}
4\sum_{k=1}^{t_1^c} N_{k,\theta}^c &\le 4\sum_{k=1}^{t_1^c}\left[\left(\frac{3}{2}\right)^{2k}\log^2 T_\theta + 1\right] \le 4t_1^c + \frac{36}{5}\left[\left(\frac{3}{2}\right)^{2t_1^c} - 1\right]\log^2 T_\theta \\
&\le \frac{1}{2}T_\theta + \Theta\left(\log\left(\frac{T_\theta}{\log^2 T_\theta}\right)\right) \le T_\theta - \log^3 T_\theta \le T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta} \qquad \text{and} \\
4\sum_{k=1}^{t_2^c} N_{k,\theta}^c &\ge 4\sum_{k=1}^{t_2^c}\left[\left(\frac{3}{2}\right)^{2k}\log^2 T_\theta - 1\right] \ge \frac{36}{5}\left[\left(\frac{3}{2}\right)^{2t_2^c} - 1\right]\log^2 T_\theta - 4\tau_\theta^c \\
&\ge 2T_\theta - \Theta\left(\log\left(\frac{T_\theta}{\log^2 T_\theta}\right)\right) \ge T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}
\end{aligned}$$

which implies $t_1^c \le \tau_\theta^c \le t_2^c$. Since t_1^c and t_2^c are both $\Theta(\log\theta)$, we conclude that $\tau^c = \Theta(\log\theta)$. Moreover, $(2/3)^{2\tau_\theta^c} = \Theta(\theta^{-1}\log^2\theta)$. This completes the proof. □

B.3. Proof of Lemma 5. By the same arguments as in the proof of Lemma 1, $P\left(\bar{E}_u | E_1 \cap E_2\right) \le \sum_{k=1}^{\tau_\theta^u}(\tau_\theta^u - k + 1)P(p^u \notin I_{k+1}^u,\ p^u \in I_k^u)$. So, we can bound

$$\begin{aligned}
P\left(\ \overline{E_1 \cap E_2 \cap E_u}\ \right) &\le P\left(\ \overline{E_1 \cap E_2}\ \right) + P(E_1 \cap E_2 \cap \bar{E}_u) \\
&\le P\left(\ \overline{E_1 \cap E_2}\ \right) + \sum_{k=1}^{\tau_\theta^u}(\tau_\theta^u - k + 1)P(p^u \notin I_{k+1}^u, p^u \in I_k^u).
\end{aligned}$$

The remaining task then is to bound the term $P(p^u \notin I_{k+1}^u,\ p^u \in I_k^u)$ for $k = 1, ..., \tau_\theta^u$. Define:

$$B_{k,1}^u = \{\hat{r}(p_{k,2}^u) < \hat{r}(p_{k,3}^u), p^u < p_{k,2}^u\} \quad \text{and} \quad B_{k,2}^u = \{\hat{r}(p_{k,2}^u) \ge \hat{r}(p_{k,3}^u), p^u > p_{k,3}^u\}.$$

Observe that, for all k, we have

$$\begin{aligned}
P(p^u \notin I^u_{k+1}, p^u \in I^u_k) &\le P\left(\hat r(p^u_{k,2}) < \hat r(p^u_{k,3}),\, p^u \notin I^u_{k+1},\, p^u \in I^u_k\right) \\
&\quad + P\left(\hat r(p^u_{k,2}) \ge \hat r(p^u_{k,3}),\, p^u \notin I^u_{k+1},\, p^u \in I^u_k\right) \\
&= P(\hat r(p^u_{k,2}) < \hat r(p^u_{k,3}),\, p^u \in [p^u_{k,1}, p^u_{k,2}),\, p^u \in I^u_k) \\
&\quad + P(\hat r(p^u_{k,2}) \ge \hat r(p^u_{k,3}),\, p^u \in (p^u_{k,3}, p^u_{k,4}],\, p^u \in I^u_k) \\
&\le P(\hat r(p^u_{k,2}) < \hat r(p^u_{k,3}),\, p^u < p^u_{k,2},\, p^u \in I^u_k) \\
&\quad + P(\hat r(p^u_{k,2}) \ge \hat r(p^u_{k,3}),\, p^u > p^u_{k,3},\, p^u \in I^u_k) \\
&\le P\left(B^u_{k,1}\right) + P\left(B^u_{k,2}\right).
\end{aligned}$$

By Lemma 3 part (i), we have

$$r(p^u_{k,2}) - r(p^u_{k,3}) \ge K_u (p^u_{k,2} - p^u_{k,3})^2 = K_u \frac{|I^u_1|^2}{9}\left(\frac{2}{3}\right)^{2(k-1)} = \frac{1}{4} K_u |I^u_1|^2 \left(\frac{2}{3}\right)^{2k}.$$

Arguing as in the proof of Lemma 1, if $|\hat r(p_{k,l}) - r(p_{k,l})| < \frac{1}{8} K_u |I^u_1|^2 (\frac{2}{3})^{2k}$ for all k and $l \in \{2,3\}$, then we can correctly predict whether $r(p^u_{k,2}) \ge r(p^u_{k,3})$ or $r(p^u_{k,2}) < r(p^u_{k,3})$. (This guarantees that the deleted segment does not contain p^u.) So, applying Hoeffding's inequality together with the facts that $\hat r(p_{k,l}) < \bar p$ and $|I^u_1| = |I| \left(\frac{2}{3}\right)^{\tau_\theta} = \Theta(\log^{-1/4}\theta)$ (see Lemma 4), we can bound $P(B^u_{k,l})$ as follows:

$$\begin{aligned}
P(B^u_{k,l}) &\le P\left(|\hat r(p_{k,j}) - r(p_{k,j})| \ge \frac{1}{8} K_u |I^u_1|^2 \left(\frac{2}{3}\right)^{2k} \text{ for some } j \in \{2,3\}\right) \\
&\le \sum_{j=2}^{3} P\left(|\hat r(p_{k,j}) - r(p_{k,j})| \ge \frac{1}{8} K_u |I^u_1|^2 \left(\frac{2}{3}\right)^{2k}\right) \\
&\le 4 \cdot \exp\left(-2 \frac{N^u_{k,\theta}\, [\frac{1}{8} K_u\, |I^u_1|^2 \left(\frac{2}{3}\right)^{2k}]^2}{\bar p^2}\right) \\
&\le 4 \cdot \exp(-\log\theta) = \frac{4}{\theta}, \quad \text{for } l = 1,2 \text{ and sufficiently large } \theta.
\end{aligned}$$

Since it can be shown that $\tau^u_\theta = \Theta(\log\theta)$, put the above bounds together with our earlier bound for $P\left(\,\overline{E_1 \cap E_2 \cap E_u}\,\right)$ and $P\left(\,\overline{E_1 \cap E_2}\,\right)$ (from Lemma 2), we conclude that

$$P\left(\,\overline{E_1 \cap E_2 \cap E_u}\,\right) \le P\left(\,\overline{E_1 \cap E_2}\,\right) + \sum_{k=1}^{\tau^u_\theta} (\tau^u_\theta - k + 1) \sum_{l=1}^{2} P(B^u_{k,l}) = \Theta\left(\frac{(\log\theta)^2}{\theta}\right). \quad \square$$

B.4. Proof of Lemma 6 . Define two events:

$$B^c_{k,1} = \{\hat\lambda(p^c_{k,2}) > C/T + \Delta^c_{k,\theta},\ p^c < p^c_{k,2}\} \text{ and } B^c_{k,2} = \{\hat\lambda(p^c_{k,2}) \le C/T + \Delta^c_{k,\theta},\ p^c > p^c_{k,3}\}.$$

By similar arguments as in the proof of Lemma 5, we know that $P(\bar E_c \,|\, E_1 \cap E_2) \le \sum_{k=1}^{\tau^c_\theta} (\tau^c_\theta - k + 1) \left[\sum_{l=1}^{2} P(B^c_{k,l})\right]$. For event $B^c_{k,1}$, note that $p^c < p^c_{k,2}$ implies $\lambda(p^c_{k,2}) < C/T$. So,

$$P(B^c_{k,1}) \le P\left(\hat\lambda(p^c_{k,2}) > C/T + \Delta^c_{k,\theta}\ ,\ \lambda(p^c_{k,2}) < C/T\right) \le P\left(\hat\lambda(p^c_{k,3}) - \lambda(p^c_{k,3}) > \Delta^c_{k,\theta}\right).$$

Since $N^c_{k,\theta} = \Theta(\left(\frac{3}{2}\right)^{2k} \log^2\theta)$ and $\Delta^c_{k,\theta} = \Theta(\left(\frac{2}{3}\right)^{k} \log^{-3/8}\theta)$, by Hoeffding's inequality,

$$P(B^c_{k,1}) \le P\left(\hat\lambda(p^c_{k,2}) - \lambda(p^c_{k,2}) > \Delta^c_{k,\theta}\right) \le \exp\left(-2N^c_{k,\theta}(\Delta^c_{k,\theta})^2\right) \le \exp(-\log\theta) = \frac{1}{\theta}.$$

As for event $B^c_{k,2}$, note that $p^c > p^c_{k,3}$ implies $\lambda(p^c_{k,3}) > C/T$. By Lemma 3 part (ii), $\lambda(p^c_{k,2}) - \lambda(p^c_{k,3}) \geq L\cdot|p^c_{k,2} - p^c_{k,3}| = L\frac{|I^c_1|}{3}(\frac{2}{3})^{k-1}$. So, for the sample path in $B^c_{k,2}$, we have:

$$\begin{aligned}\lambda(p^c_{k,2}) - \hat{\lambda}(p^c_{k,2}) &\geq \lambda(p^c_{k,3}) + L\frac{|I^c_1|}{3}\left(\frac{2}{3}\right)^{k-1} - \left(\frac{C}{T} + \Delta^c_{k,\theta}\right)\\ &> \frac{C}{T} + L\frac{|I^c_1|}{3}\left(\frac{2}{3}\right)^{k-1} - \left(\frac{C}{T} + \Delta^c_{k,\theta}\right)\\ &= L\frac{|I^c_1|}{3}\left(\frac{2}{3}\right)^{k-1} - \Delta^c_{k,\theta} \;>\; \frac{1}{2}L\frac{|I^c_1|}{3}\left(\frac{2}{3}\right)^{k-1}\end{aligned}$$

where the last inequality follows because $|I^1_c| = |I|\left(\frac{2}{3}\right)^{\tau_\theta} = \Theta(\log^{-1/4}\theta)$ and so $\Delta^c_{k,\theta} < \frac{1}{2}L\frac{|I^c_1|}{3}\left(\frac{2}{3}\right)^{k-1}$ for all large θ. By similar argument as above,

$$\begin{aligned}P(B^c_{k,2}) &\leq P\left(\lambda(p^c_{k,2}) - \hat{\lambda}(p^c_{k,2}) > \frac{1}{2}L\frac{|I^c_1|}{3}\left(\frac{2}{3}\right)^{k-1}\right)\\ &\leq \exp\left(-2\frac{N^c_{k,\theta}\;[\frac{1}{2}L\frac{|I^c_1|}{3}\left(\frac{2}{3}\right)^{k-1}]^2}{\bar{p}^2}\right) \;\leq\; \exp(-\log\theta) = \frac{1}{\theta}.\end{aligned}$$

Since it can be shown that $\tau^c_\theta = \Theta(\log\theta)$, put the above bounds together with our earlier bound for $P\left(\,\overline{E_1\cap E_2\cap E_c}\,\right)$ and $P\left(\,\overline{E_1\cap E_2}\,\right)$ (from Lemma 2), we conclude that

$$P\left(\,\overline{E_1\cap E_2\cap E_c}\,\right) \leq P\left(\,\overline{E_1\cap E_2}\,\right) + \sum_{k=1}^{\tau^c_\theta}(\tau^c_\theta - k + 1)\sum_{l=1}^{2}P(B^c_{k,l}) \;=\; \Theta\left(\frac{(\log\theta)^2}{\theta}\right). \quad \square$$

References

[1] Agrawal, R. 1995. The continuum-armed bandit problem. *SIAM J. Control Optim.* **33**(6) 1926–1951.

[2] Araman, V. F., R. Caldentey. 2009. Dynamic pricing for nonperishable products with demand learning. *Oper. Res.* **57**(5) 1169–1188.

[3] Auer, P., R. Ortner, C. Szepesvári. 2007. Improved rates for the stochastic continuum-armed bandit problem. *Learning Theory*. Springer, 454–468.

[4] Bertsimas, D., G. Perakis. 2006. *Dynamic Pricing: A Learning Approach*. Springer.

[5] Besbes, O., Y. Gur, A. Zeevi. 2013. Non-stationary stochastic optimization. Working paper, Columbia Business School, Columbia University.

[6] Besbes, O., A. Zeevi. 2009. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Oper. Res.* **57**(6) 1407–1420.

[7] Besbes, O., A. Zeevi. 2012. Blind network revenue management. *Oper. Res.* **60**(6) 1537–1550.

[8] Besbes, O., A. Zeevi. 2013. On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. Working paper, Columbia Business School, Columbia University.

[9] Boyd, S., L. Vandenberghe. 2009. *Convex Optimization*. Cambridge university press.

[10] Broadie, M., D. Cicek, A. Zeevi. 2011. General bounds and finite-time improvement for the kiefer-wolfowitz stochastic approximation algorithm. *Oper. Res.* **59**(5) 1211–1224.

[11] Broder, J., P. Rusmevichientong. 2012. Dynamic pricing under a general parametric choice model. *Oper. Res.* **60**(4) 965–980.

[12] Cesa-Bianchi, N., G. Lugosi, et al. 2006. *Prediction, Learning, and Games*, vol. 1. Cambridge University Press Cambridge.

[13] Chen, G., S. Jasin, I. Duenyas. 2013. Analysis of self-adjusting controls for dynamic pricing with unknown demand functions. Working paper, Ross School of Business, University of Michigan.

[14] Cohen, M., R. Lobel, G. Perakis. 2013. Dynamic pricing though data sampling. Working paper, Operations Research Center, MIT.

[15] Cope, E. W. 2009. Regret and convergence bounds for a class of continuum-armed bandit problems. *IEEE Transactions on Automatic Control* **54**(6) 1243–1253.

[16] den Boer, A. V. 2013. Dynamic pricing and learning: historical origins, current research, and new directions. Working paper, Department of Applied Mathematics, University of Twente.

[17] den Boer, A. V. 2014. Dynamic pricing with multiple products and partially specified demand distribution. *Math. Oper. Res.* **39**(3) 863 – 888.

[18] den Boer, A. V., B. Zwart. 2013. Dynamic pricing and learning with finite inventories. Working paper, Department of Applied Mathematics, University of Twente.

[19] den Boer, A. V., B. Zwart. 2013. Simultaneously learning and optimizing using controlled variance pricing. *Management Sci.* **60**(3) 770–783.

[20] Eren, S. S., C. Maglaras. 2010. Monopoly pricing with limited demand information. *J. Revenue & Pricing Management* **9**(1) 23–48.

[21] Farias, V. F., B. van Roy. 2010. Dynamic pricing with a prior on market response. *Oper. Res.* **58**(1) 16–29.

[22] Gallego, G., G. van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Sci.* **40**(8) 999–1020.

[23] Harrison, J. M., N. B. Keskin, A. Zeevi. 2012. Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. *Management Sci.* **58**(3) 570–586.

[24] Hoeffding, W. 1963. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* **58**(301) 13–30.

[25] Jasin, S. 2014. Reoptimization and self-adjusting price control for network revenue management. *Oper. Res.* Article in Advance, June 13, http://dx.doi.org/10.1287/opre.2014.1297.

[26] Keskin, N. B., A. Zeevi. 2013. Chasing demand: Learning and earning in a changing environment. Working paper, University of Chicago.

[27] Keskin, N. B., A. Zeevi. 2013. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. Working paper, University of Chicago.

[28] Kiefer, J., J. Wolfowitz. 1952. Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.* **23**(3) 462–466.

[29] Kleinberg, R. D. 2004. Nearly tight bounds for the continuum-armed bandit problem. *Advances in Neural Information Processing Systems*. 697–704.

[30] Kushner, H. J., G. Yin. 2003. *Stochastic Approximation and Recursive Algorithms and Applications*, vol. 35. Springer.

[31] Lim, A. E. B., J. G. Shanthikumar. 2007. Relative entropy, exponential utility, and robust dynamic pricing. *Oper. Res.* **55**(2) 198–214.

[32] Özer, Ö., R. Phillips. 2012. *The Oxford Handbook of Pricing Management*. Oxford University Press.

[33] Perakis, G., G. Roels. 2010. Robust controls for network revenue management. *Manufacturing Service Oper. Management* **12**(1) 56–76.

[34] Robbins, H., S. Monro. 1951. A stochastic approximation method. *Ann. Math. Statist.* 400–407.

[35] Rusmevichientong, P., B. van Roy, P. W. Glynn. 2006. A nonparametric approach to multiproduct pricing. *Oper. Res.* **54**(1) 82–98.

[36] Talluri, K. T., G. J. van Ryzin. 2006. *The Theory and Practice of Revenue Management*, vol. 68. springer.

[37] Waeber, R., P. I. Frazier, S. G. Henderson. 2013. Bisection search with noisy responses. *SIAM J. Control Optim.* **51**(3) 2261–2279.

[38] Wang, Z., S. Deng, Y. Ye. 2014. Close the gaps: A learning-while-doing algorithm for single-product revenue management problems. *Oper. Res.* **62**(2) 318–331.