Dynamic Pricing under Operational Frictions

by

Qi Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in the University of Michigan
2017

Doctoral Committee:

Professor Izak Duenyas, Co-Chair
Assistant Professor Stefanus Jasin, Co-Chair
Assistant Professor Eric M. Schwartz
Assistant Professor Cong Shi

Qi Chen
georgeqc@umich.edu
ORCID iD: 0000-0002-6026-9103

Dedication

This dissertation is dedicated to my wife, Xu Zhang, for all her love and unwavering support; and to our daughter, Zoe Chen, for all the joys she brings to us.

Acknowledgements

First of all, I would like to thank my advisors, Professor Izak Duenyas and Professor Stefanus Jasin, for their years of mentorship and help. They brought me into the academic world, helped me develop the skills to become a researcher. I am greatly inspired by their work ethics. I could not have achieved what I have achieved without them. I would also like to thank my dissertation committee members, Professor Cong Shi and Professor Eric Schwartz for their helpful comments and feedback.

Moreover, I would like to express my thanks to the faculty members and fellow Ph.D. students at the Technology and Operations group at the Ross School of Business for all the enlightening seminars, individual discussions and support throughout my Ph.D. program. In particular, I would like to thank Professor Damian Beil, for a lot of invaluable research and career advice, and my Ph.D. coordinators, Professor Amitabh Sinha and Professor Stephen Leider, for many years of help.

Finally, I would like to thank my parents for their love and support.

Contents

List of Tables

List of Figures

Abstract

This dissertation investigates the tactical dynamic pricing decisions in industries where sellers sell multiple types of capacity-constrained products/services to their customers. Motivated by operational frictions posed by business considerations, I develop dynamic pricing heuristics that have both provably good revenue performance and nice features which can address these operational frictions. The first essay studies how to do effective dynamic pricing without too many price changes. In practice, many sellers have concerns about dynamic pricing due to the computational complexity of frequent re-optimizations, the negative perception of excessive price adjustments, and the lack of flexibility caused by existing business constraints. To address these concerns, I develop a pricing heuristic which is computationally easy to implement and only needs to adjust a small number of prices and do so infrequently to guarantee a strong revenue performance. In addition, when not all products are equally admissible to price adjustment, my heuristic can replace the price adjustment of some products by their similar products and maintain an equivalent revenue performance. These features allow the sellers to achieve most of the benefit of dynamic pricing with much fewer price changes and provide extra flexibility to manage prices. While the first essay assumes that the sellers know the underlying demand function, this information is sometimes unavailable to the sellers in practice. The second and the third essays study how to jointly learn the demand and dynamically price the products to minimize revenue loss compared to a standard revenue upper bound in the literature. The second essay addresses the parametric case where the seller knows the functional form of the demand but not the parameters; the third essay addresses the nonparametric case where the seller does not even know the functional form of the demand. There is a considerable gap between the revenue loss lower bound under any pricing policy and the performance bound of the best known heuristic in the literature. To close the gap, in my second essay, I propose a heuristic that exactly match the lower bound for the parametric case, and show that under a demand separation condition, a much sharper revenue loss bound can be obtained; in my third essay, I propose a heuristic whose performance is arbitrarily close to the lower bound for the nonparametric case. All the proposed heuristics are computationally very efficient and can be used as a baseline for developing more sophisticated heuristics for large-scale problems.

Chapter 1 Introduction

Revenue Management practice has become a crucial component in firms' operations in many industries such as airlines, hospitality, fashion and ground transportation. In these industries, sellers sell multiple types of capacity-constrained products/services to their customers during a finite selling season. Often times, the capacity is fixed during the selling season, unused capacity has little salvage value, and the customer demand is uncertain. In such scenarios where the sellers need to match fixed supply with uncertain demand, dynamic pricing is very useful: by adjusting the prices, the sellers can effectively influence how fast the demand arrives over time during the selling season to better match supply with demand.

Although the idea of dynamic pricing is straightforward, implementing it effectively is not an easy task. Indeed, for large-scale real-world applications where the sellers need to sell many different types of products which may share common capacity constrained resources, finding the optimal dynamic pricing policy is computationally untractable. Moreover, despite the potential benefit of dynamic pricing, in practice some sellers still adopt a static pricing policy due to the following operational frictions: (1) business considerations and constraints may disallow too many price changes; (2) a dynamic pricing policy may be less useful than a more static one when demand function is not perfectly estimated. Motivated by these concerns of using dynamic pricing, in my dissertation, I develop effective dynamic pricing heuristics to address these operational frictions.

The first essay studies how to do effective dynamic pricing without too many price changes. In practice, many business considerations can make excessive price adjustment undesirable or infeasible. For one thing, a seller may be unwilling to adjust prices too frequently because excessive price adjustment may leave customers with negative perceptions of the seller and affect long-term profitability. For another, a seller may not even be able to adjust the prices for some of the product/services he offers due to contractual agreement with certain segments of his customers. In this essay, I develop a family of pricing heuristics called *Linear Price Correction* (LPC) to address how a seller should dynamically price his products/services with price change restrictions. LPC is computationally easy to implement; it requires only a single optimization at the beginning of the selling season and automatically adjusts the prices over time, taking into account the restrictions on when and what price can be adjusted. Moreover, under LPC, to guarantee a strong revenue performance, it is sufficient to adjust the prices of a small number of products and do so infrequently. This property helps the seller focus his effort on the prices of the most important products instead of all products. In addition, in the case where not all products are equally admissible to price adjustment due to existing business constraints, LPC can immediately substitute the price adjustment of the original products with the price adjustment of similar products and maintain an equivalent revenue performance. This property provides the seller with extra flexibility in managing his prices.

The remainder of my dissertation addresses how to do effective dynamic pricing when the underlying demand function is unknown and needs to be estimated. Since sellers may have different levels of knowledge of the underlying demand function, I consider two cases separately in my second and third essays. The second essay addresses the parametric case where the seller knows the functional form of the underlying demand function but not the exact parameters; the third essay addresses the nonparametric case where the seller does not even know the functional form of the demand function. In both cases, the seller needs to jointly learn the demand and dynamically price the products/services. Developing effective pricing policy is hard since the seller has to deal with the tension between learning the demand function (exploration) and using an effective dynamic pricing policy based on the estimated demand function (exploitation). Following the convention, my goal is to find pricing heuristics that minimize sellers' expected *revenue loss* compared to a standard revenue upper bound in the literature. The lower bound of the revenue loss of any pricing policy under either case is well-known in the literature, but there is a considerable gap between this lower bound and the performance bound of the best known heuristic in the literature. To close the gap, I develop several self-adjusting heuristics with strong performance bounds.

In my second essay, I develop a heuristic called *Parametric Self-adjusting Control* (PSC) which combines maximum likelihood estimation and a self-adjusting pricing scheme. I show that the revenue loss under PSC exactly matches the revenue loss lower bound. In addition, I show that if the parametric demand function family further satisfies a separation condition, the seller can learn the demand function much faster due to a property which I call *passive learning*. To tap into this nice property, I develop another heuristic called *Accelerated Parametric Self-adjusting Control* (APSC), which combines a doubling re-estimation scheme with a self-adjusting pricing scheme. I show that this heuristic can attain a much sharper revenue loss bound when the separation condition holds.

In my third essay, I propose a heuristic called *Non-parametric Self-adjusting Control* (NSC) which combines spline estimation, functional approximations and a self-adjusting pricing scheme for the nonparametric case. I show that as long as the underlying demand function is sufficiently smooth, the revenue loss of NSC can be arbitrarily close to the revenue loss lower bound. My results suggest that in terms of performance, the nonparametric approach can be as robust as the parametric approach, at least asymptotically. All the proposed heuristics are computationally very efficient and can be used as a baseline for developing more sophisticated heuristics for large-scale problems.

Chapter 2 Pricing with Minimal and Flexible Price Adjustment

2.1 Abstract

I study a standard dynamic pricing problem where the seller (a monopolist) possesses a finite amount of inventories and attempts to sell the products during a finite selling season. Despite the potential benefits of dynamic pricing, many sellers still adopt a static pricing policy due to (1) the complexity of frequent re-optimizations, (2) the negative perception of excessive price adjustments, and (3) the lack of flexibility caused by existing business constraints. In this essay, I develop a family of pricing heuristics that can be used to address all these challenges. My heuristic is computationally easy to implement; it requires only a single optimization at the beginning of the selling season and automatically adjusts the prices over time. Moreover, to guarantee a strong revenue performance, the heuristic only needs to adjust the prices of a *small* number of products and do so infrequently. This property helps the seller focus his effort on the prices of the most important products instead of all products. In addition, in the case where not all products are equally admissible to price adjustment (due to existing business constraints such as contractual agreement, strategic product positioning, etc.), my heuristic can immediately substitute the price adjustment of the original products with the price adjustment of similar products and maintain an equivalent revenue performance. This property provides the seller with extra flexibility in managing his prices.

2.2 Introduction

Nowadays, Revenue Management (RM) practice has become very prevalent in many industries such as airlines, hospitality, fashion, ground transportation, and many others. (See Talluri and van Ryzin (2005, chap.10) for more examples.) In a typical RM setting, the seller possesses a finite amount of inventories and attempts to maximize his revenue by selling a collection of products during a finite selling season. Often times, replenishment of inventory is not viable during the selling season and the leftovers have little salvage value (e.g., empty hotel rooms). There are two types of RM commonly found in practice: *quantity*-based RM and *price*-based RM. In the first category, prices are fixed over the selling season and the focus is on making a dynamic resource allocation. As for the second category, prices become the key decision variables and the seller adjusts his prices as often as he wishes and sells all products until stock-out. Although the two types of RM are not mutually exclusive, market context and the seller's value proposition may dictate which of the two is more appropriate. In this dissertation I am primarily interested in price-based RM. (For a review of quantity-based RM, see Talluri and van Ryzin (2005, chap.2).)

Pricing is, without doubt, one of the most important decisions that affect the seller's profitability. According to a study by McKinsey & Company, "Pricing right is the fastest and most effective way for managers to increase profits" (Marn et al. 2003). The study argues that a 1% price increase in a typical S&P 1500 company would generate an 8% increase in operating profit, an impact which is almost 50% greater than that of a 1% reduction in variable cost and more than three times greater than that of a 1% increase in volume. Perhaps more strikingly, an annual report of the operating profit for airlines and rental car companies in the US during 2009 reveals that a 1% increase in average price improved total operating profit by up to 67% and 30%, respectively (Sen 2013). (Although a 67% improvement in profit is arguably rather unusual, a moderate $8\% - 25\%$ increase via dynamic pricing is not uncommon (Sahay 2007).) And yet, despite its apparent benefit, dynamic pricing still poses several serious challenges. First, the *complexity* of the required large-scale optimization leads to prohibitive computational burden. To illustrate, a typical major US airline operates thousands of flights daily and posts fares several months into the future. Accounting for the number of different booking classes per flight, this can easily translate into daily pricing decisions for *millions* of itineraries. Hotel industry is no exception. Koushik et al. (2012) reports that a single run of price optimization at the InterContinental Hotels Group (excluding the estimation time) takes about four hours to complete. Similarly, Pekgun et al. (2013) also reveals that it takes about six hours for the Carlson Rezidor Hotel Group to complete its price optimization *once*. Given the increased competition in many industries where the prices of some products are now being adjusted even *hourly* (Rigby et al. 2012), this begs the question whether there exists a scalable pricing heuristic which can be easily implemented in real-time.

Second, dynamic pricing typically involves *frequent* price adjustments of *many* products, which may not be desirable for the firms. For one thing, even when full-scale dynamic pricing tools are readily available, the seller may want to intentionally avoid excessive price adjustments due to brand positioning and customer relationship considerations. Widely accepted as it is in the airline industry, dynamic pricing suffers a considerable setback in some other industries due to negative customers' perception. For example, in hotel industry, the most common criticism of dynamic pricing is that it treats customers unequally and unfairly (Ramasastry 2005), and lab experiments confirm the unfairness perception of price discrimination (Haws and Bearden 2006). Aside from the customers' perception issue, frequent price adjustments of many products may also not be feasible due to existing *business constraints*, i.e., the seller may not have the flexibility to adjust the prices of some products because of existing regulations and contractual agreement. For example, hotels often face customers from the so-called *negotiated segment* and provide *fixed* corporate rates for large travel buyers such as IBM and HP (Koushik et al. 2012). Thus, hotels are practically forced to provide a fixed price that cannot be adjusted over time to the negotiated segment while at the same time are free to dynamically adjust the prices for other customer segments. This situation is not unique to hotel industry alone. The practice of *selective* dynamic pricing, which combines dynamic pricing of some products with fixed pricing of other products, is not uncommon and can be found in many industries (e.g., with the exception of Sears, Amazon.com, and Kmart, most retailers only change their prices daily on less than 10% of their assortments (Rigby et al. 2012)). And yet, despite its common practice, I am not aware of any work in the academic literature that rigorously analyzes the feasibility and effectiveness of such approach.

The preceding discussions lead to several important research questions: (1) Can we construct a pricing heuristic that is easy to implement and does not require frequent price adjustments? (2) Can we adjust the price of only a *small* number of products in order to mitigate customers' negative perception while at the same time maintaining a decent revenue performance? If such minimal price adjustment is possible, (3) how should we pick the set of products whose prices are to be updated? Is there a simple rule that can be used as a guidance? Moreover, in the case where the seller's business constraints disallow him to dynamically adjust the prices of some products, (4) can he still maintain an equivalent revenue performance by dynamically adjusting the price of other products? If yes, which *other* products should be used? In this essay I address all these questions. In particular, I will construct a family of real-time heuristics which, depending on the firm's need, can be used to address any of the aforementioned issues.

Static price control and re-optimization. There is a rich operations management (OM) literature on dynamic pricing. (For overviews, see Bitran and Caldentey (2003) and Elmaghraby and Keskinocak (2003).) In the RM context, motivated by the well-known curse of dimensionality of Dynamic Program (DP), many existing works have focused on the construction of easy-to-use heuristics. There are two popular approaches that can be found in the literature. The first is based on the so-called *Approximate Dynamic Programming* (ADP).

Some works along this line are Erdelyi and Topaloglu (2011) and Kunnumkal and Topaloglu (2010). The second approach, which is closer to my work in this essay, is based on solving a *deterministic* analog of the original stochastic problem. One of the seminal works on this approach is Gallego and van Ryzin (1997). The trade-off between the two approaches is obvious. On the one hand, the sophisticated ADP requires more computational power than the deterministic approach. On the other hand, while the former yields an "adaptive" price sequence, which depends on sales realization, the latter only results in a deterministic (static) price. The good news is that static price control is *asymptotically* optimal (Gallego and van Ryzin 1997). This may partly explain its decent performance, hence its wide adoption, in many industries. Yet, a considerable amount of revenue is still lost. As noted earlier, the main drawback of static pricing is that it completely ignores the observed demand realizations and the remaining inventory levels. One potential way of utilizing this progressively revealed information is to periodically *re-optimize* the aforementioned deterministic optimization. The impact of re-optimization in quantity-based RM has been extensively studied in the literature (e.g., see Chen and de Mello (2010), Reiman and Wang (2008), Secomandi (2008), Ciocan and Farias (2012), Jasin and Kumar (2012, 2013)). As for price-based RM, Maglaras and Meissner (2006) is the first to show that re-optimizing static price control guarantees at least the same asymptotic performance as static price without re-optimization. Thus, although re-optimization does not necessarily result in a monotonically increasing revenue, it cannot severely degrade revenue either. This is in contrast to the potentially negative impact of re-optimization in quantity-based RM (Jasin and Kumar 2013). Chen and Farias (2013) analyze the impact of re-optimization in the presence of imperfect forecast for a single product RM. They show that a combination of re-optimization and re-estimation yields a significant improvement in revenue. The paper that is perhaps closest to ours is Jasin (2014). The author provides a tighter bound for the expected revenue loss of the re-optimized static price control studied in Maglaras and Meissner (2006). This confirms the theoretical benefit of re-optimization for a very general class of multi-product and multi-resource RM. In addition, the author also proposes a simple pricing heuristic that can be implemented in real-time. (See Section 2.5 for further discussions on this.) A parallel but independent work by Atar and Reiman (2012) studies a continuous time version of the same problem and shows that the problem can be reduced to a diffusion control problem whose optimal solution is a Brownian bridge. The Brownian bridge structure motivates them to develop a diffusion-scale dynamic pricing heuristic that has similar error correction terms as the simple heuristic developed in Jasin (2014).

Although re-optimization is intuitively appealing and enjoys a good theoretical guarantee, unfortunately, it is not always practically feasible. As previously discussed, even a single optimization of a large-scale real problem instance can take hours to complete (Pekgun et al. 2013). This obviously serves as a bottleneck for the number of re-optimizations that can be implemented in one day. A recent work by Golrezaei et al. (2014) in the context of assortment optimization also highlights the same issue. The problem being re-optimized in their setting is a linear program, which is considered by many as one of the most tractable family of optimization problems. And yet, their simulation shows that the running time of frequent re-optimizations can be 800 times larger than that of a single optimization. While the resulting time-lag due to re-optimization may not be too detrimental for brick-and-mortar stores who update their prices less frequently, it is clearly less feasible for online retailers with more frequent price adjustments. In such settings, any proposed control must ideally be implementable in real-time without unnecessarily invoking large-scale re-optimization.

The proposed heuristic. In this essay, I introduce a new family of dynamic pricing heuristics, which I call *Linear Price Correction* (LPC). LPC only requires a single deterministic optimization at the beginning of the selling season and can be implemented in real-time. In addition, LPC only needs to adjust the price of a small number of products, admits a general asynchronous update schedule, and allows update substitution among "similar" products. Needless to say, it is also possible to couple LPC with occasional re-optimizations to further improve its performance. All these properties taken together allow the seller to enjoy the benefit of dynamic pricing while at the same time reducing the computational burden of re-optimization and mitigating the negative effect of frequent price changes on customers' perception.

The remainder of the essay is organized as follows. Section 2.3 describes the problem setting and the asymptotic approach I take to analyze the performance of any dynamic pricing heuristic. The proposed heuristic LPC is formally introduced in Section 2.4 where I also discuss its minimal and asynchronous price adjustment properties which allow LPC to achieve good performance by adjusting the prices of only a small number of products and do so infrequently. In Section 2.5, I show the flexibility of LPC in choosing the prices of which products to adjust by demonstrating how to achieve equivalent revenue performances by adjusting prices of different sets of products that are "equivalent". Section 2.6 uses numerical experiments to show the strong performance of LPC and its modifications, and to illustrate the managerial insights drawn from previous sections. Finally, Section 2.7 concludes. The proofs of all my results are deferred to Appendix A.1.

2.3 Problem Formulation

Consider a multi-period and multi-product pricing problem where the seller sells a catalog of n products (indexed by j), each of which is made up of a combination of m types of resources (indexed by i) whose initial inventory levels are given by $C \in \mathbb{R}^m$. As is usually the case, the number of products is much larger than the number of resources. Denote by $A = [A_{ij}]$ the *consumption matrix*, whose element A_{ij} indicates the amount of resource i required by one unit of product j. Without loss of generality, I assume that the rows of A are linearly independent. The selling season is finite and divided into T periods. At the beginning of period t, the seller posts the price $p_t = (p_{t,j})$. The price then induces a demand $D_t(p_t) = (D_{t,j}(p_t))$ with rate $\lambda(p_t) = \mathbf{E}[D_t(p_t)]$. As is common in the literature, I allow at most one customer arrival per period. Hence, the function $\lambda(p_t)$ can also be interpreted as the arrival probability in period t. Let $r(p_t) := p_t' \lambda(p_t)$ denote the revenue rate in period t, where p_t' indicates the transpose of p_t. Let Ω_p and Ω_λ denote the convex set of feasible prices and demand rates, respectively. I make the following assumptions:

(A1) The demand function $\lambda(p_t) : \Omega_p \to \Omega_\lambda$ is invertible, twice differentiable, monotonically decreasing in its individual argument, and bounded from above by $\bar{\lambda}$.

(A2) The revenue function $r(p_t) = p_t' \lambda(p_t) = \lambda_t' p(\lambda_t) = r_t(\lambda_t)$ is continuous, strictly jointly concave in λ_t, and bounded from above by $\bar{r}$.

(A3) For each product j, there exists a turn-off price p_j^∞ such that if $\{p^k\}$ is any price sequence satisfying $p_j^k \to p_j^\infty$, then I have $\lambda_j(p^k) \to 0$.

(A4) The absolute eigenvalues of $\nabla^2 \lambda_j(p_t)$ and $\nabla^2 r(p_t)$ are bounded from above by $\bar{v}$.

Assumptions (A1) - (A3) are similar to the standard regularity conditions in the literature (Gallego and van Ryzin 1997). (A1) is a mild assumption to ensure basic analytical properties of the demand rate. (A2) follows from the invertibility assumption in (A1) and is needed to guarantee that the function $r(.)$ has a unique, bounded optimizer. The revenue functions under a vast class of demand models such as linear and logit demand satisfy these assumptions. As for (A3), the existence of turn-off prices allows a seller to effectively shut down the demand for any product whenever desirable. (A4) is easily satisfied in general, especially for compact Ω_p. The constants $\bar{\lambda}$, $\bar{r}$ and $\bar{v}$ are independent of t.

The RM pricing problem. The optimal stochastic pricing problem can be written as:

$$\text{(SPP):} \qquad J^{Stoc} \;=\; \max_{\pi\in\Pi_p} \; \mathbf{E}\left[\sum_{t=1}^{T}(p_t^{\pi})'\,D_t(p_t^{\pi})\right] \quad \text{s.t.} \quad A\left[\sum_{t=1}^{T} D_t(p_t^{\pi})\right] \le C,$$

where Π_p is the set of all non-anticipating pricing policies and the constraints must hold almost surely. Alternatively, by the invertibility of demand function, I can also use $\{\lambda_t\}$ as the decision variables and replace p_t and $D_t(p_t)$ with $p_t(\lambda_t)$ and $D_t(\lambda_t)$ respectively. I then replace the random variables in SPP by their mean and obtain a more tractable deterministic formulation below.

$$\text{(DPP):} \qquad J^{Det} \;=\; \max \sum_{t=1}^{T} r(\lambda_t) \quad \text{s.t.} \quad \sum_{t=1}^{T} A\lambda_t \le C \quad \text{and} \quad \lambda_t \in \Omega_\lambda, \;\; \forall t.$$

Let $\{\lambda_t^D\}$ denote the unique optimal solution to DPP. Correspondingly, I define $p_t^D := p(\lambda_t^D)$. Since demand is time-homogeneous, it can be shown that $\lambda_t^D = \lambda_1^D := \lambda^D$ and $p_t^D = p_1^D := p^D$ for all t. This explains the name *static* pricing. For analytical tractability, I will assume that λ^D lies in the interior of Ω_λ. I formally state this assumption below.

(A5) There exist strictly positive constants ϕ_L and ϕ_U such that $[\lambda^D - \phi_L\mathbf{e},\, \lambda^D + \phi_U\mathbf{e}] \subseteq \Omega_\lambda$.

Assumption (A5) essentially says that *all* products matter. It implies the optimal deterministic price is neither so low that it induces too many requests nor so high that it completely shuts down the demand of some products. As a practical rule of thumb, if some products are not profitable (i.e. $\lambda_j^D = 0$ for some j), they can be discarded from the catalog and one can re-run the optimization. This helps the seller to focus on the products that matter. Hence, (A5) is *not* restrictive at all.

Performance measure and asymptotic regime. Ideally, one would like to define *revenue loss* of any control π as the difference between the revenue earned under the optimal pricing policy and the revenue earned under the control. Since the former is not easy to compute, I resort to using an upper bound as an approximation. It is known that $J^{Stoc} \le J^{Det}$. (This is a standard result in the literature and is an immediate consequence of Jensen's inequality. I omit its proof.) Let R_π denote total revenue earned under heuristic π throughout the selling season. The expected revenue loss of heuristic π is then defined as: $\text{RL}_\pi = J^{Det} - \mathbf{E}[R_\pi]$. Following Gallego and van Ryzin (1997), in this essay I consider a sequence of increasing problems parameterized by $\theta > 0$. To be precise, in the θ^{th} problem, I scale both the length of selling season and the initial inventory levels by a factor of θ while keeping all the other parameters unchanged. Let $T(\theta)$ and $C(\theta)$ denote the length of the selling season and initial inventory levels in the θ^{th} problem, respectively. Then $T(\theta) = \theta T$ and $C(\theta) = \theta C$. One may interpret the parameter θ as the *scale*, or *relative size*, of the problem. (If C is normalized to 1, then θ has an immediate interpretation as the size of initial inventory levels. Alternatively, if T is normalized to 1, the scale θ can be interpreted as the size of potential demands.) Notationwise, I will simply attach (θ) as a reference to the θ^{th} problem. Observe that the optimal solution of the scaled deterministic problem is the same as the optimal solution of the unscaled one (i.e., $\lambda_t^D(\theta) = \lambda^D$ and $p_t^D(\theta) = p^D$), so I have $J^{Det}(\theta) = \theta J^{Det}$.

2.4 Minimal and Asynchronous Price Adjustments

In this section, I will develop a pricing heuristic that adjusts the prices of only a small number of products and admits a general asynchronous update schedule. I show that my heuristic guarantees a strong asymptotic performance despite the fact that it only adjusts the prices of a small number of products. This has an obvious managerial significance. For example, at Chicago O'Hare airport, United Airlines operates more than forty routes to and from the North East and another thirty or so routes to and from the West Coast and the Mountain Area (see www.united.com). Assuming one fare class per flight, the company needs to price approximately $40 \times 30 = 1,200$ itineraries from the North East to the West Coast and the Mountain Area that make one stop at O'Hare airport. My result suggests that United only needs to dynamically price $40 + 30 = 70$ itineraries instead of $1,200$. Moreover, the price of these 70 itineraries can be adjusted asynchronously instead of simultaneously.

To introduce my heuristic, I start with a notion of a *base*. (This is the set of products whose prices are to be adjusted under the heuristic. I will allow more adjustable prices in Section 2.5.) A subset of products $\mathcal{B}$ is said to be a *base* if (1) it contains exactly m products and (2) the products in $\mathcal{B}$ *span the resource space*, meaning the columns of matrix $A\nabla\lambda(p^D)$ that correspond to the products in $\mathcal{B}$ (by the same index) span the whole $\mathbb{R}^m$. Note that, since the rows of A are linearly independent and $\nabla\lambda(p^D)$ is invertible, the rank of $A\nabla\lambda(p^D)$ is m. So, there always exists at least one base. Let H be a real n by m matrix satisfying $AH = I$, where I is an m by m identity matrix. I call H a *projection* matrix and say that a projection matrix H *selects the base* $\mathcal{B}$ if the rows of $\nabla p(\lambda^D)H$ (by the same index) that correspond to the products *not* in $\mathcal{B}$ are all zero vectors. As will be evident shortly, a proper choice of matrix H is important to ensure that only the prices of the base products are dynamically adjusted while the prices of the non-base products are never changed. The following lemma establishes the existence of a projection matrix for any given base.

Lemma 2.4.1 *For any base $\mathcal{B}$, there exists a unique projection matrix H that selects it.*

The heuristic. Fix a base $\mathcal{B}$ and assume without loss of generality that $\mathcal{B} = \{1, \ldots, m\}$. For each $j \in \mathcal{B}$, define $\gamma_j = \{t_l^j : 1 \le l \le K_j\}$ to be the updating schedule for product j. (An updating schedule can be viewed as a business constraint that prescribes when the price of a given product is adjustable.) In particular, the l^{th} updating time is denoted by t_l^j and the number of updates is K_j. For convenience, I will write $t_0^j = 1$ and $t_{K_j+1}^j = T+1$. Let $k_t^j = \max\left\{k : t_k^j \le t\right\}$ denote the number of price updates for product j by time t. This setting is very general: *I allow the price of each product in the base to be updated asynchronously (i.e., independently of the other products in the base)*. Let H be a projection matrix that selects $\mathcal{B}$. For any set $\mathcal{A} \subseteq \{1, \ldots, n\}$, let $E^{\mathcal{A}}$ denote an n by n diagonal matrix with $E_{ii}^{\mathcal{A}} = 1$ if $i \in \mathcal{A}$ and 0 otherwise. (This matrix helps select a set of rows of another matrix when it is left-multiplied, e.g., $E^j \nabla p(\lambda^D) H$ is a matrix whose j^{th} row is the same as the j^{th} row of $\nabla p(\lambda^D) H$ and all its other rows are zeros.) Define $\Delta_t(p_t) := D_t(p_t) - \mathbf{E}[D_t(p_t)] = D_t(p_t) - \lambda(p_t)$ and $\tilde{\Delta}_l^j := \sum_{s=t_{l-1}^j}^{t_l^j - 1} \Delta_s(p_s)$, $l = 1, \ldots, K_j+1$. The term $\Delta_t(p_t)$ can be interpreted as demand error during period t and the term $\tilde{\Delta}_l^j$ can be interpreted as *cumulative* demand errors between two subsequent updating times for product j. (For brevity, whenever there is no confusion, I will often suppress notational dependency on p_t and simply write Δ_t, D_t, and λ_t.) Let C_t denote the remaining inventory levels at the end of period t. The definition of my heuristic is given below.

Linear Price Correction (LPC)

1. During period 1, set $p_1 = p^D$.
2. At the beginning of period $t > 1$, do:
   a. First compute $\hat{p}_t \;=\; p^D - \sum_{j=1}^{m} E^j \, \nabla p(\lambda^D) H \left[\sum_{l=1}^{k_t^j} \frac{A \tilde{\Delta}_l^j}{T - t_l^j + 1} \right]$.
   b. Set the price according to the following rule:
      (1) If $C_{t-1} \ge A^j$ for all j, and $\hat{p}_s \in \Omega_p$ for all $s \le t$, set $p_{t,j} = \hat{p}_{t,j}$;
      (2) Otherwise, set $p_{t,j} = p_j^{\infty}$.

The idea behind LPC is to use static price p^D as baseline prices and apply real-time adjustment to only the prices of m products in the chosen base. The proposed adjustment has an intuitive interpretation: If past demand realization is higher than expected (i.e., the term $\tilde{\Delta}$'s are positive), then LPC immediately increases future prices; if, on the other hand, past demand realization is lower than expected, then LPC immediately decreases future prices. To see that the given update formula only adjusts prices of base products, define $\tilde{\xi}_l^j \mathbf{e}_j \;:=\; E^j \, \nabla p(\lambda^D) H A \tilde{\Delta}_l^j$ and $\xi_s^j \mathbf{e}_j \;:=\; E^j \, \nabla p(\lambda^D) H A \Delta_s$, where $\mathbf{e}_j$ is a vector with proper size whose j^{th} element equals one and any of its other elements equals zero. Note that I can write $\hat{p}_t$ as:

$$\begin{bmatrix} \hat{p}_{t,1} \\ \vdots \\ \hat{p}_{t,m} \\ \\ \hat{p}_{t,m+1} \\ \vdots \\ \hat{p}_{t,n} \end{bmatrix} = \begin{bmatrix} p_1^D - \sum_{l=1}^{k_t^1} \frac{\tilde{\xi}_l^1}{T-t_l^1+1} \\ \vdots \\ p_m^D - \sum_{l=1}^{k_t^m} \frac{\tilde{\xi}_l^m}{T-t_l^m+1} \\ \\ p_{m+1}^D \\ \vdots \\ p_n^D \end{bmatrix}.$$

Obviously, only the prices of the first m products are adjusted. Moreover, for each $j \in \mathcal{B}$, if the current period t is such that $t_{l-1}^j < t < t_l^j$ for some l, then $p_{t,j} = p_{t-1,j}$. So, the price of product $j \in \mathcal{B}$ in the periods between two subsequent updating times does not change. To help the reader better understand the mechanism of this pricing heuristic, I give an example below.

Example 1. Consider a network RM with 3 products and 2 resources. Without loss of generality, I assume that $\mathcal{B} = \{1, 2\}$ is a base. Suppose that $\gamma_1 = \{2, 5, ...\}$ and $\gamma_2 = \{4, 5, ...\}$ (i.e., I want to adjust the price of product 1 in periods 2, 5, etc. and the price of product 2 in periods 4, 5, etc.). Assuming no stock-out, the price formula for the first five periods, are given by:

$$\begin{bmatrix} p_{1,1} \\ \\ p_{1,2} \\ \\ p_{1,3} \end{bmatrix} = \begin{bmatrix} p_1^D \\ \\ p_2^D \\ \\ p_3^D \end{bmatrix}, \quad \begin{bmatrix} p_{2,1} \\ \\ p_{2,2} \\ \\ p_{2,3} \end{bmatrix} = \begin{bmatrix} p_1^D - \frac{\xi_1^1}{T-1} \\ \\ p_2^D \\ \\ p_3^D \end{bmatrix}, \quad \begin{bmatrix} p_{3,1} \\ \\ p_{3,2} \\ \\ p_{3,3} \end{bmatrix} = \begin{bmatrix} p_1^D - \frac{\xi_1^1}{T-1} \\ \\ p_2^D \\ \\ p_3^D \end{bmatrix},$$

$$\begin{bmatrix} p_{4,1} \\ \\ p_{4,2} \\ \\ p_{4,3} \end{bmatrix} = \begin{bmatrix} p_1^D - \frac{\xi_1^1}{T-1} \\ \\ p_2^D - \frac{\xi_1^2+\xi_2^2+\xi_3^2}{T-3} \\ \\ p_3^D \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} p_{5,1} \\ \\ p_{5,2} \\ \\ p_{5,3} \end{bmatrix} = \begin{bmatrix} p_1^D - \left(\frac{\xi_1^1}{T-1} + \frac{\xi_2^1+\xi_3^1+\xi_4^1}{T-4}\right) \\ \\ p_2^D - \left(\frac{\xi_1^2+\xi_2^2+\xi_3^2}{T-3} + \frac{\xi_4^2}{T-4}\right) \\ \\ p_3^D \end{bmatrix}.$$

General performance bound. I will now discuss the performance of LPC. I first provide a general bound that can be applied to arbitrary updating schedule and then I discuss its implication for several specific schedules. For the sake of generality, I will allow the choice of updating schedule to also depend on θ, i.e., $\gamma_j(\theta) = \{t_l^j(\theta) : 1 \le l \le K_j(\theta)\}$, $j \in \mathcal{B}$. Let $R_{H,\gamma_{\mathcal{B}}}(\theta)$ denote the total revenue earned under LPC with projection matrix H and updating schedules $\gamma_{\mathcal{B}} := \{\gamma_j(\theta)\}_{j\in\mathcal{B}}$. Let $||.||_2$ denote the usual spectral norm of a matrix, i.e., $||X||_2^2$ equals the maximum eigenvalue of $X'X$. I state my result below.

Theorem 2.4.1 *There exist positive constants* Ψ *and* $\bar{\Psi}$ *independent of* $\theta \ge 1$*, the projection matrix* H *that selects* $\mathcal{B}$*, and the choice of updating schedules* $\{\gamma_j(\theta)\}_{j\in\mathcal{B}}$ *such that*

$$\begin{aligned} J^{Det}(\theta) - \mathbf{E}\left[R_{H,\gamma_{\mathcal{B}}}(\theta)\right] \ \le \ \Psi \quad & + \ \bar{\Psi} \sum_{j\in\mathcal{B}} \sum_{t=1}^{T(\theta)-1} \min\left\{1, \left|\left|\nabla p(\lambda^D) H A\right|\right|_2^2 \, U_1^j(T(\theta), t)\right\} \\ & + \ \bar{\Psi} \sum_{j\in\mathcal{B}} \sum_{t=1}^{T(\theta)-1} \min\left\{1, \left|\left|\nabla p(\lambda^D) H A\right|\right|_2^2 \, U_2^j(T(\theta), t)\right\}, \end{aligned}$$

where the terms $U_1^j(T,t)$ *and* $U_2^j(T,t)$ *are defined as*

$$U_1^j(T,t) \ = \ \frac{t - t_{k_t^j}^j + 1}{(T-t)^2} \ + \ \sum_{l=1}^{k_t^j} \frac{t_l^j - t_{l-1}^j}{\left(T - t_l^j + 1\right)^2} \quad \text{and} \quad U_2^j(T,t) \ = \ \frac{1}{T-t} \sum_{s=1}^{t} \sum_{l=1}^{k_s^j} \frac{t_l^j - t_{l-1}^j}{\left(T - t_l^j + 1\right)^2}.$$

I want to stress: The above bound is *very* general. It characterizes the performance of LPC for *any* given base and *any* given updating schedule[1], either synchronous or asynchronous. (The implications of Theorem 2.4.1 for specific schedules will be discussed below.) Note that the bound is *separable* over the products in the base. This suggests that the seller cannot compensate the lack of updating of one product in the base by applying more frequent updates to the remaining product(s) in the base. If there exist multiple feasible bases, the bound in Theorem 2.4.1 suggests that I use the base $\mathcal{B}$ and the corresponding projection matrix H that minimizes $||\nabla p(\lambda^D) H A||_2$. Although, in general, it is not possible to explicitly characterize the "optimal" base products chosen by this selection rule, it turns out that I can provide a very intuitive characterization of the "optimal" base product for the case of single-resource RM.

[1] In the setting of quantity-based RM, Jasin and Kumar (2012) also provide a bound for revenue loss which depends on a general choice of updating schedule. However, they assume that the admission control for *all* products must be *simultaneously* updated at the same time. In contrast, LPC allows *each* product to have its own updating schedule. This level of generality, together with the non-linearity of the objective function and capacity constraints, introduces non-trivial analytical subtleties which do not previously exist in the analysis of Jasin and Kumar (2012).

Lemma 2.4.2 *Suppose that $m = 1$. Among all projection matrices that select a base, the projection matrix H^* that achieves the smallest $||\nabla p(\lambda^D)HA||_2$ selects the base that consists of product $j^* = \arg\max_{j=1,\ldots,n} \left|(A\nabla\lambda(p^D))_j\right|$.*

The intuition of the above lemma is most easily explained if one considers a special case of single-resource RM with $A = [1, \ldots, 1]$ and separable demands (i.e., $\lambda_j(p)$ only depends on p_j). In this setting, $A\nabla\lambda(p^D)$ becomes a row vector whose j^{th} element equals the demand sensitivity of product j with respect to its own price, $\lambda'_j(p_j^D)$. Thus, under LPC, the optimal projection matrix selects the most price-sensitive product into the base. This can be intuitively explained as follows: Among all products, product j^* needs the smallest price perturbation to correct the same demand error. Since I am using the deterministic model as my performance benchmark, ideally, I would want to have a price trajectory that stays as close as possible to the baseline price p^D. This can be achieved by adjusting the product that requires the smallest perturbation. As for the more general case of single-resource RM with general demand and general capacity consumption matrix A, a similar intuition also holds: one wants to pick the product whose price adjustment has the largest impact on capacity consumption.

Special updating schedules. I will now apply the result of Theorem 2.4.1 to derive an explicit performance bound for several special updating schedules that only adjust the prices of base products and draw some managerial insights. I start with the most commonly used update schedule where prices are being adjusted periodically according to some frequencies.

Corollary 2.4.1 (h-Periodic Schedule) *Fix $h(\theta) \geq 1$ and define $t_l^j(\theta) = t_l(\theta) = l\,h(\theta) + 1$ for all $j \in \mathcal{B}$. There exist positive constants Ψ, $\hat{\Psi}$, and $\bar{\Psi}$ independent of $\theta \geq 1$ and $h(\theta) \geq 1$ such that the expected revenue loss of LPC is bounded by $\Psi + \hat{\Psi}\sqrt{h(\theta)} + \bar{\Psi}\log^2\theta$.*

Two comments are in order. First, if $h(\theta) = T(\theta)$, then the periodic schedule reduces to static pricing and the revenue loss is $\mathcal{O}(\sqrt{\theta})$. This bound is consistent with the result in Gallego and van Ryzin (1997). If, on the other hand, $h(\theta) = 1$, the revenue loss is reduced to $\mathcal{O}(\log^2\theta)$. Since LPC requires only one optimization followed by simple price updates, it provides a significant improvement[2] over static pricing with negligible computational effort. Second, although Corollary 2.4.1 assumes a synchronous schedule, it is not difficult to derive a bound for an asynchronous periodic update schedule because the bound is separable in

[2]Since θ represents the size of the problem, the percentage revenue loss under LPC is approximately $\frac{\log^2\theta}{\theta} \times 100\%$ whereas the percentage revenue loss under static pricing is about $\frac{\sqrt{\theta}}{\theta}$ x 100%. Numerically, for a problem instance with initial inventory levels equal to 100, as in a typical airplane with 100 seats, my experiments in Section 2.6 show a 2% improvement in revenue, which is quite significant for typical RM applications.

individual product. For example, one plausible asynchronous schedule would be to adjust the prices of base products on weekly basis, but on different days of the week. The asymptotic performance bound will remain the same as in Corollary 2.4.1. One caveat of periodic schedule is that, in order to reduce the revenue loss to $\mathcal{O}(\log^2 \theta)$, a very frequent updates of the prices of all base products (roughly $\Theta(\theta)$ times) is required. But, per my discussions in Section 2.2, this may not be practically feasible – or even if it is, it may not be strategically desirable due to customers' perception issue. To address this, below I propose two schedules that still guarantee $\mathcal{O}(\log^2 \theta)$ revenue loss albeit with much fewer price updates.

Corollary 2.4.2 (α-Power Schedule) *Fix $\alpha \geq 1$. For all $j \in \mathcal{B}$, let $t_0^j(\theta) = t_0(\theta) = 1$ and define $t_l^j(\theta) = t_l(\theta) = \left\lceil T(\theta) - \sum_{s=1}^{K(\theta)-l+1} s^\alpha \right\rceil$ for $1 \leq l \leq K(\theta)$, where $K(\theta) := \{k : \sum_{s=1}^{k} s^\alpha < T(\theta), \sum_{s=1}^{k+1} s^\alpha \geq T(\theta)\}$. Then $K(\theta) \leq ((\alpha+1)T(\theta))^{1/(\alpha+1)}$ and there exist positive constants Ψ and $\bar{\Psi}$ independent of $\theta \geq 1$ such that the expected revenue loss of LPC is bounded by $\Psi + \bar{\Psi} \log^2 \theta$.*

Corollary 2.4.3 (β-Geometric Schedule) *Fix $\beta > 1$. For all $j \in \mathcal{B}$, let $t_0^j(\theta) = t_0(\theta) = 1$, and for $l \geq 1$, iteratively define $t_l^j(\theta) = t_l(\theta) = \left\lceil \frac{(\beta-1)T(\theta) + t_{l-1}(\theta)}{\beta} \right\rceil$ as long as $t_{l-1}(\theta) < T(\theta)$. Let $K(\theta)$ be such that $t_{K(\theta)}^j(\theta) = T(\theta)$. Then, $K(\theta) \leq 1 + \log_\beta T(\theta)$, and there exist positive constants Ψ and $\bar{\Psi}$ independent of $\theta \geq 1$ such that the expected revenue loss of LPC is bounded by $\Psi + \bar{\Psi} \log^2 \theta$.*

Corollaries 2.4.2 and 2.4.3 offer two interesting insights. First, by carefully choosing the update times, one can use a small number of updates (only about $\theta^{\frac{1}{\alpha+1}}$ updates with power schedule and $\log_\beta \theta$ updates with geometric schedule) to guarantee a $\mathcal{O}(\log^2 \theta)$ revenue loss.[3] Second, for both schedules, most of the updates happen near the end of the selling season. This implies that the crucial moments for dynamic pricing is near the end of the selling season instead of at the beginning, which suggests that the seller can perhaps apply static price at the beginning of the season and only switch to dynamic pricing later. Needless to say, although Corollaries 2.4.2 and 2.4.3 assume synchronous schedules, it is also possible to use asynchronous schedules. For example, the prices of some base products can be updated using power schedule and the prices of other base products can be updated using geometric schedule. Again, since the bound in Theorem 2.4.1 is separable over the products in the base, the $\mathcal{O}(\log^2 \theta)$ bound still holds.

[3] My simulations show that the non-asymptotic performance of *1-Power* schedule is almost the same as that of *1-Periodic* schedule. This is very impressive since when $\theta = 500$, *1-Power* needs 44 adjustments while *1-Periodic* requires 500 adjustments. For larger θ, the difference is even bigger.

The impact of adjusting the prices of fewer, or more, than m products. Since adjusting the price of all products may not be desirable, or even feasible, it is important to understand the impact of restricting the number of adjustable products on revenue. Corollaries 2.4.1 - 2.4.3 partially answer this question by showing a surprising result that adjusting the prices of only m products (in the base) is sufficient to guarantee a $\mathcal{O}(\log^2 \theta)$ revenue loss.[4] This is a powerful result because, in most RM applications, the number of resources m is typically much smaller than the number of products n. In particular, it provides an important managerial insight that the seller does not need to aggressively adjust the prices of all products to benefit from dynamic pricing. The result on minimal price adjustment, however, leads to two interesting questions. First, can one still guarantee the $\mathcal{O}(\log^2 \theta)$ revenue loss by adjusting the prices of fewer than m products? The answer is unfortunately negative and the revenue loss under such scenario is of order $\sqrt{\theta}$ in general. To understand why this is so, consider the case where demands are separable and $A = I$ is an m by m identity matrix. Since this corresponds to an aggregate of m independent problems (e.g., m independent one-stop flights), if one only dynamically adjusts the price of $m' < m$ products, then it is equivalent to applying static price control to the remaining $m - m'$ problems, which is known to have $\Theta(\sqrt{\theta})$ revenue loss in general (Jasin 2014). Second, what is the incremental benefit of adjusting the prices of more than m products? To answer this, I again consider the case of a single-resource RM. (By minimal price adjustment property, I already know that one only needs to adjust the price of *one* product to guarantee a significant improvement over static pricing. The question is whether adjusting the prices of more products has a significant impact on performance.) Let $b = (A \nabla \lambda(p^D))'$ and denote by $b_{(i)}$ the i^{th} largest element (in absolute value) of b. For $k \geq 1$, let Π_k denote the set of all non-anticipating pricing policies that adjust the price of at most k products in each period. (If the price of product j is not adjusted in period t under $\pi \in \Pi_k$, then $p^{\pi}_{t,j} = p^{\pi}_{t-1,j}$.) Then, the following result holds.

Theorem 2.4.2 *Suppose that $m = 1$. There exist positive constants Ψ and $\bar{\Psi}$ independent of $\theta \geq 1$ and $1 \leq k \leq n$ such that*

$$\min_{\pi \in \Pi_k} \left\{ J^{Det}(\theta) - \mathbf{E}[R_\pi(\theta)] \right\} \leq \Psi + \frac{\bar{\Psi}}{\sum_{i=1}^{k} b_{(i)}^2} \log^2 \theta.$$

[4]Since I only have m resources, it seems "intuitive" that I should be able to perform well by adjusting the prices of only m products. However, since adjusting the prices of only m products also affects the demands for the other $n - m$ products whose prices are not adjusted, it is not immediately clear what impact this would have on revenue. My result is different from the so-called *action-space reduction* discussed in pg. 220 of Talluri and van Ryzin (2005). Under the action-space reduction scenario, one first computes the optimal *aggregate* decision variable and then *disaggregates* this variable to recover the optimal price for each product. However, there is no guarantee that this disaggregation will result in the adjustment of only the prices of m products. In contrast, under my scenario, the prices of $n - m$ products are never changed.

The above performance bound suggests that the incremental benefit of adjusting the price of an additional product decreases as the number of the adjustable products increases. To see this, suppose that $A = [1, \dots, 1]$ and demands are separable and identical across different products with $\lambda_j(.) = \lambda_1(.)$ for all j. This implies $p_j^D = p_1^D$ for all j and $b_{(i)} = \lambda_1'(p_1^D)$ for all i. Then, the bound in Theorem 2.4.2 is of order $\frac{\log^2 \theta}{k}$. Since the function $1/k$ drops quickly for small k and slowly for large k, this suggests that it is not necessary for the seller to adjust the prices of too many products to get most of the potential revenue. (See Section 2.6 for numerical evidence of this observation in the multi-resource case. My results show that the revenue improvement of adjusting the price of m products over static pricing is about $80-90\%$ of the revenue improvement of adjusting the price of all n products, in most cases. Moreover, in terms of revenue loss, while adjusting the price of m products reduces the revenue loss of static pricing by about $1-1.2\%$, adjusting the price of n products only further reduces the revenue loss by an additional $0.1-0.2\%$ in most cases. (See Table 2.2 in Section 2.8.) Given that the average margins in RM industries are typically very small, only about 3% (Irvine 2014), this highlights the practical significance of minimal adjustments for real-world implementation.) In particular, if the seller wishes to adjust the prices of more than m products to further increase revenue, then s/he only needs to consider adjusting the prices of a few more products instead of all.

2.5 Equivalent Performance via Adjusting the Prices of Other Products

Corollaries 2.4.2 and 2.4.3 in the previous section provide an important managerial insight: Managers need to update the prices of only a small subset of their products, and do so sufficiently rarely, to guarantee a strong revenue performance. Those results, however, assume that only the prices of the same m products are updated throughout the selling season. Can we do better? For example, why should we update the price of one product ten times and the other products not at all if a major concern of some practitioners is that customers get upset by frequent price changes? Can we reduce the number of price updates per product by somehow *distributing* the required adjustments across different products over different time periods (e.g., one price update per product for ten different products instead of ten price updates for one product)? Also, what if the seller dictates that the price of some products should not, or cannot, be changed either due to existing business constraints or contractual agreements? Can we somehow re-assign the scheduled update for these products to other "similar products"? As discussed in Section 2.2, although these questions have significant practical relevance and are faced by many sellers, I am not aware of any existing work in the literature addressing these issues. In this section, I will discuss a generalization of LPC that partially addresses these issues. My proposed heuristic provides important practical insights on how to do *equivalent* pricing via adjusting the prices of similar products. To illustrate the basic idea, I start with two examples.

Example 2. Consider a single flight RM with n types of ticket. I assume that each ticket only requires one seat and demands are separable. Note that $\nabla\lambda(p^D)$ is a diagonal matrix. As Corollary 2.4.3 indicates, it is sufficient to adjust the price of only one type of ticket $\Theta(\log_2\theta)$ times to obtain $\mathcal{O}(\log^2\theta)$ revenue loss. If I evenly distribute these adjustments to all n tickets, the number of price updates per ticket is about $\lceil(\log_2\theta)/n\rceil$. It turns out that this still guarantees $\mathcal{O}(\log^2\theta)$ revenue loss. Thus, dynamically adjusting one type of ticket $\Theta(\log_2\theta)$ times is equivalent to dynamically adjusting n types of tickets $\Theta((\log_2\theta)/n)$ times for each. This has an important managerial implication. As an illustration, consider economy seats. There are usually about 13 different fare classes for economy seats. Since a typical US passenger flight has fewer than 500 seats and $\log_2(500) = 8.96$, by my previous arguments, I can either adjust the price of one fare class nine times or the price of *any* nine fare classes once during the selling season.

Example 3. Consider a network RM problem with 3 resources and 6 products and suppose that

$$A\nabla\lambda(p^D) = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & -2 \\ 0 & -1 & 0 & -2 & -2 & 0 \\ 0 & 0 & -1 & -2 & 0 & -1 \end{bmatrix}.$$

Obviously, $\mathcal{B} = \{1,2,3\}$ forms a base. Suppose that the previously prescribed schedule for $\mathcal{B}$ is $\gamma_1 = \{2,3,5\}$, $\gamma_2 = \{3,4,5\}$, and $\gamma_3 = \{4,6\}$. Unlike in the previous example where I can arbitrarily pick any nine products, here, the choice of "similar products" is more subtle. A new set of products is similar to the original set of products if its corresponding columns (by the same index) in $A\nabla\lambda(p^D)$ can linearly represent the columns in $A\nabla\lambda(p^D)$ that correspond to the original set of products. In my example, this means that I can replace updating $\{2,3\}$ in period 4 with $\{4,5\}$, or replace updating $\{3\}$ in period 6 with $\{4,5\}$. I cannot directly replace the price adjustment of product 3 in period 4 with product 4 because column 4 is not parallel to column 3. But, since product 2 will be adjusted in period 4 under both the original schedule and the new schedule, I can achieve an equivalent revenue by bundling the price adjustment of product 2 and 3 in period 4 and substituting it with the price adjustment of $\{2,4\}$.

Equivalent pricing control. I now formally state the idea behind the preceding examples. For clarity, I assume that $\mathcal{B} = \{1, \ldots, m\}$ is a base and H is a projection matrix that selects $\mathcal{B}$. Let $\gamma_{\mathcal{B}} := \{\gamma_j(\theta)\}_{j=1}^m$ denote the *existing* updating schedule for base products. I will show in this section that, for any *equivalent schedule* of $\gamma_{\mathcal{B}}$ (to be formally defined below), I can construct a pricing heuristic that guarantees the same asymptotic performance as LPC under $\gamma_{\mathcal{B}}$. In other words, if the seller wants to modify the current price updating schedules to a new one for strategic considerations, then I can provide a new pricing control that guarantees an equivalent performance as long as the new updating schedule is equivalent to the current updating schedule. Before introducing equivalent schedule, I first introduce the concept of *equivalent set*: A set of products $\mathcal{G} \subseteq \{1, \ldots, n\}$ is said to be *equivalent* to the set $\mathcal{S} \subseteq \mathcal{B}$ (mathematically, I write: $\mathcal{G} \sim_{\mathcal{B}} \mathcal{S}$) if the columns in $A\nabla\lambda(p^D)$ that correspond to the products in $\mathcal{S}$ can be written as a linear combination of the columns in $A\nabla\lambda(p^D)$ that correspond to products in $\mathcal{G}$. (Note that, by my definition, $\mathcal{G} \sim_{\mathcal{B}} \mathcal{S}$ does not imply $\mathcal{S} \sim_{\mathcal{B}} \mathcal{G}$.) Let $\mathcal{S}_t \subseteq \mathcal{B}$ be a subset of products that are adjusted in period t under $\gamma_{\mathcal{B}}$. Let $\mathcal{G}_t$ be one of the (possibly) many sets that are equivalent to $\mathcal{S}_t$. I say that a price updating schedule γ is an *equivalent schedule* of $\gamma_{\mathcal{B}}$ if in each period t only products in $\mathcal{G}_t$ are adjusted under γ. Let $\mathbf{\Gamma}(\gamma_{\mathcal{B}})$ denote the set of all equivalent schedules of $\gamma_{\mathcal{B}}$. I now define an equivalent pricing control for any $\gamma \in \mathbf{\Gamma}(\gamma_{\mathcal{B}})$. Let $\mathcal{G}_t \sim_{\mathcal{B}} \mathcal{S}_t$ and denote by S_t and G_t the submatrices of $A\nabla\lambda(p^D)$ whose columns correspond to the products in $\mathcal{S}_t$ and $\mathcal{G}_t$, respectively. By definition of equivalent set, there exists a $|\mathcal{G}_t|$ by $|\mathcal{S}_t|$ matrix Y_t such that $S_t = G_t Y_t$. For any such $\mathcal{G}_t$, $\mathcal{S}_t$ and Y_t, I can construct a unique n by n matrix $Q_t = Q(Y_t, \mathcal{G}_t, \mathcal{S}_t)$ as follows: its submatrix with rows and columns not in $\mathcal{G}_t \cup \mathcal{S}_t$ equals an identity matrix, its submatrix with rows in $\mathcal{G}_t$ and columns in $\mathcal{S}_t$ equals Y_t, and any of its other elements equals 0. I call $Q(Y_t, \mathcal{G}_t, \mathcal{S}_t)$ a *transformation matrix* because, from its construction, it uses the matrix Y_t to transform the price adjustment for products in $\mathcal{S}_t$ into price adjustment for products in $\mathcal{G}_t$. The following lemma provides some important properties of $Q(Y_t, \mathcal{G}_t, \mathcal{S}_t)$.

Lemma 2.5.1 *For any $\mathcal{G}_t \sim_{\mathcal{B}} \mathcal{S}_t$ and any Y_t such that $S_t = G_t Y_t$, let $Q_t = Q(Y_t, \mathcal{G}_t, \mathcal{S}_t)$. Then, the following holds:*

(1) $A\nabla\lambda(p^D)Q_t E^{\mathcal{B}} = A\nabla\lambda(p^D)E^{\mathcal{B}} = A\nabla\lambda(p^D)E^{\mathcal{G}_t \cup (\mathcal{B} - \mathcal{S}_t)}Q_t$;

(2) There exists a projection matrix H_t such that $\nabla p(\lambda^D)H_t = Q_t \nabla p(\lambda^D)H$ and the rows in $\nabla p(\lambda^D)H_t$ that correspond to products not in $\mathcal{G}_t \cup (\mathcal{B} - \mathcal{S}_t)$ are zeros;

(3) The rows in $Q_t E^{\mathcal{S}_t} \nabla p(\lambda^D)H$ that correspond to products not in $\mathcal{G}_t$ are zeros.

Define $\mathcal{Q}_t(\gamma) := \arg\min_Q\{||Q||_2 : Q = Q(Y, \mathcal{G}_t, \mathcal{S}_t), S_t = G_t Y\}$ in each period t. (This optimization problem turns out to be a convex optimization with linear constraints and can be efficiently solved off-line.) I am now ready to introduce the concept of *equivalent pricing.* Let γ be an equivalent schedule of the existing schedule $\gamma_{\mathcal{B}}$. Then, a pricing control π with schedule γ is said to be *equivalent* to an existing LPC with updating schedule $\gamma_{\mathcal{B}}$ if, in Step 2a in the definition of LPC, it uses the following update formula:

$$\hat{p}_t = p^D - \sum_{j=1}^{m} \sum_{l=1}^{k_t^j} Q_{t_l^j} E^j \nabla p(\lambda^D) H \frac{A\tilde{\Delta}_l^j}{T - t_l^j + 1}$$

for some $Q_t \in \mathcal{Q}_t(\gamma)$[5] in each period t. (In light of part (3) of Lemma 2.5.1, the above update formula guarantees that only adjustable products under γ are adjusted in each period.)

Example 2 (cont'd). Consider again the single flight problem described in Example 2. Suppose that $n = 3$ and assume, without loss of generality, that $\mathcal{B} = \{1\}$ with the corresponding projection matrix $H = (1, 0, 0)'$. Suppose that the seller originally plans to periodically adjust the price of only product 1 at the beginning of every period using the following update formula:

$$\begin{bmatrix} p_{t,1} \\ p_{t,2} \\ p_{t,3} \end{bmatrix} = p^D - \sum_{s=1}^{t-1} \nabla p(\lambda^D) H \frac{A\Delta_s}{T-s} = \begin{bmatrix} p_1^D - \sum_{s=1}^{t-1} p_1'(\lambda_1^D) \frac{\Delta_s}{T-s} \\ p_2^D \\ p_3^D \end{bmatrix}.$$

To develop an equivalent pricing control, which alternates among the three products such that the price of only one product is being adjusted in every period, I construct a sequence of transformation matrices $\{Q_{t_l}\}$ for each update time t_l as follows. Let Q^1 be a 3 by 3 identify matrix. For $j \in \{2, 3\}$, denote by Q^j the transformation matrix that transforms the price adjustment of product 1 into price adjustment of product j. In particular, by the construction of transformation matrix

$$Q^2 = \begin{bmatrix} 0 & 0 & 0 \\ \frac{p_2'(\lambda_2^D)}{p_1'(\lambda_1^D)} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{p_3'(\lambda_3^D)}{p_1'(\lambda_1^D)} & 0 & 0 \end{bmatrix}.$$

[5] Note that, given $\gamma \in \mathbf{\Gamma}(\gamma_{\mathcal{B}})$, $\mathcal{Q}_t(\gamma)$ may not be a singleton. However, as can be seen in the proof of Theorem 2.5.1, the performance bound of an equivalent pricing control under γ depends on Q_t only via its spectral norm $||Q_t||_2$. In particular, the smaller the norm, the smaller the revenue loss bound. This observation motivates my definition of $\mathcal{Q}_t(\gamma)$ where $||Q_t||_2$ is minimized. Since all matrices in $\mathcal{Q}_t(\gamma)$ have the same spectral norm, my performance bound does not depend on the particular selection of Q_t within $\mathcal{Q}_t(\gamma)$.

For all l satisfying $l \equiv j \pmod 3$, set $Q_{t_l} = Q^j$. The resulting equivalent pricing control is then given by $\hat{p}_t = p^D - \sum_{s=1}^{t-1} Q_s \nabla p(\lambda^D) H \frac{A\Delta_s}{T-s}$. Assuming no stock-out, the explicit formulae of the price of all three products for the first five periods are:

$$\begin{bmatrix} p_{1,1} \\ \\ p_{1,2} \\ \\ p_{1,3} \end{bmatrix} = \begin{bmatrix} p_1^D \\ \\ p_2^D \\ \\ p_3^D \end{bmatrix}, \quad \begin{bmatrix} p_{2,1} \\ \\ p_{2,2} \\ \\ p_{2,3} \end{bmatrix} = \begin{bmatrix} p_1^D - p_1'(\lambda_1^D)\frac{\Delta_1}{T-1} \\ \\ p_2^D \\ \\ p_3^D \end{bmatrix}, \quad \begin{bmatrix} p_{3,1} \\ \\ p_{3,2} \\ \\ p_{3,3} \end{bmatrix} = \begin{bmatrix} p_1^D - p_1'(\lambda_1^D)\frac{\Delta_1}{T-1} \\ \\ p_2^D - p_2'(\lambda_2^D)\frac{\Delta_2}{T-2} \\ \\ p_3^D \end{bmatrix},$$

$$\begin{bmatrix} p_{4,1} \\ \\ p_{4,2} \\ \\ p_{4,3} \end{bmatrix} = \begin{bmatrix} p_1^D - p_1'(\lambda_1^D)\frac{\Delta_1}{T-1} \\ \\ p_2^D - p_2'(\lambda_2^D)\frac{\Delta_2}{T-2} \\ \\ p_3^D - p_3'(\lambda_3^D)\frac{\Delta_3}{T-3} \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} p_{5,1} \\ \\ p_{5,2} \\ \\ p_{5,3} \end{bmatrix} = \begin{bmatrix} p_1^D - p_1'(\lambda_1^D)\frac{\Delta_1}{T-1} - p_1'(\lambda_1^D)\frac{\Delta_4}{T-4} \\ \\ p_2^D - p_2'(\lambda_2^D)\frac{\Delta_2}{T-2} \\ \\ p_3^D - p_3'(\lambda_3^D)\frac{\Delta_3}{T-3} \end{bmatrix}.$$

Thus, in this example, I have shown how to adjust the prices of three products $T/3$ times each instead of adjusting the price of one product T times using equivalent pricing.

Performance result. For any updating schedule $\gamma \in \mathbf{\Gamma}(\gamma_{\mathcal{B}})$, let $\mathbf{Q} \in \mathcal{Q}(\gamma) := \{\{Q_t\}_{t=1}^T : Q_t \in \mathcal{Q}_t(\gamma)\}$ denote a sequence of transformation matrices that correspond to γ and let $R^{\mathbf{Q}}_{H,\gamma_{\mathcal{B}},\gamma}$ denote the resulting revenue. The following theorem provides a uniform performance bound for equivalent pricing control under any updating schedule γ that is equivalent to $\gamma_{\mathcal{B}}$.

Theorem 2.5.1 *There exist positive constants Ψ and $\bar{\Psi}$ independent of $\theta \geq 1$, the projection matrix H that selects $\mathcal{B}$, and the choice of updating schedules $\gamma_{\mathcal{B}}$ such that*

$$\begin{aligned} \sup_{\gamma \in \mathbf{\Gamma}(\gamma_{\mathcal{B}})} \sup_{\mathbf{Q} \in \mathcal{Q}(\gamma)} \left\{ J^{Det}(\theta) - \mathbf{E}\left[R^{\mathbf{Q}}_{H,\gamma_{\mathcal{B}},\gamma}(\theta) \right] \right\} \;\leq\; \Psi \quad &+ \bar{\Psi} \sum_{j \in \mathcal{B}} \sum_{t=1}^{T(\theta)-1} \min\left\{ 1, ||\nabla p(\lambda^D) H A||_2^2 U_1^j(T(\theta), t) \right\} \\ &+ \bar{\Psi} \sum_{j \in \mathcal{B}} \sum_{t=1}^{T(\theta)-1} \min\left\{ 1, ||\nabla p(\lambda^D) H A||_2^2 U_2^j(T(\theta), t) \right\}, \end{aligned}$$

where the terms $U_1^j(T,t)$ and $U_2^j(T,t)$ are defined as in Theorem 2.4.1.

Observe that the bound in Theorem 2.5.1 is similar to the bound in Theorem 2.4.1. This shows that, for any schedule γ that is equivalent to the base schedule $\gamma_{\mathcal{B}}$, the seller can use equivalent pricing to guarantee the same asymptotic performance as the LPC under the base schedule $\gamma_{\mathcal{B}}$. This result provides the seller with an extra flexibility to manage his prices.

LPC with synchronous price adjustment of more than m products. Although the LPC discussed in Section 2.4 allows for arbitrary asynchronous price adjustment, it is restricted to adjust the price of *exactly* m products. Generalizing LPC to the case of *arbitrary* asynchronous price adjustment of more than m products is not a trivial task and beyond the scope of this essay. It is, however, possible to use equivalent pricing to develop a version of LPC that *synchronously* adjusts the prices of $k \geq m$ products. To illustrate how to use equivalent pricing to do synchronous price adjustment for $k \geq m$ products, consider the LPC discussed in Section 2.4 where the base is $\mathcal{B}$ and $\gamma_j(\theta) = \gamma_1(\theta)$ for all $j \in \mathcal{B}$. Let $\mathcal{G}$ denote a set of $k \geq m$ products that span the resource space (i.e., the set of products whose corresponding columns (by the same index) in $A\nabla\lambda(p^D)$ span $\mathbb{R}^m$). Since $\mathcal{G} \sim_{\mathcal{B}} \mathcal{B}$, I can construct a transformation matrix Q as described above and apply equivalent pricing with $Q_t = Q$ for all t. The resulting price update formula is given by

$$\hat{p}_t \;=\; p^D - \sum_{l=1}^{k_t^1} Q\nabla p(\lambda^D)H\frac{A\tilde{\Delta}_l^1}{T - t_l^1 + 1} \;=\; p^D - \sum_{l=1}^{k_t^1} \nabla p(\lambda^D)\tilde{H}\frac{A\tilde{\Delta}_l^1}{T - t_l^1 + 1}\,,$$

where the second equality follows from the second part of Lemma 2.5.1 with $\tilde{H}$ being a projection matrix such that the rows in $\nabla p(\lambda^D)\tilde{H}$ that correspond to products *not* in $\mathcal{G}$ are zeros. Note that such pricing control has a practical implication: It provides the seller with an extra flexibility to trade off the negative impact of excessive price adjustment with the incremental improvement in revenue due to adjusting the price of more products. (See Theorem 2.4.2 and numerical experiments in Section 2.6 for further discussions.)

The difference between LPC and LRC. As briefly mentioned in Section 2.2, Jasin (2014) has developed a dynamic pricing heuristic which he calls *Linear Rate Correction* (LRC), and it adjusts the price in period t using the update formula $\hat{p}_t \;=\; p\left(\lambda^D - H\sum_{s=1}^{t-1}\frac{A\Delta_s}{T-s}\right)$, where H is a projection matrix. To see the difference between LPC and LRC, first, note that, since $p(\cdot)$ is not always separable, the prices of *all* n products under LRC must be *simultaneously* updated at the same time. (Even if the projection matrix H is chosen to select a certain base, there is no guarantee that LRC will adjust the price of only the products in the base.) Thus, minimal price adjustment of only m products is, in general, not possible with LRC. Second, since $p(\cdot)$ is not always separable, there is no analog of the general LPC update formula for LRC. This means that neither asynchronous update nor equivalent pricing is possible with LRC, which may limit the applicability of LRC for real-world implementation (e.g., due to existing business constraints). Indeed, aside from the fact that LRC and LPC are examples of *linear control*[6], they are close only in the special case where the prices

[6] Linear control has been widely studied in engineering (Ben-Tal et al. 2009) and finance (Calafiore 2009),

of all products are updated at the same time (e.g., the synchronous *1-Periodic* schedule). In that special case, the update formula of LPC can be viewed as a linearization of the update formula of LRC. (The generic asynchronous LPC, however, is *not* a linearization of any form of LRC.)

2.6 Numerical Experiments

In this section, I run several experiments to illustrate the theoretical results in Sections 2.4 and 2.5 as well as to highlight the applicability of my heuristic in practice and its managerial implications. For my simulations, I use a multinomial logit demand with 10 products and 4 resources. (See Appendix A.2 for more detail.) I use $T = 1$ and $C_i = 0.1$ for each resource i. Note that, per my definition, the actual number of selling periods and initial inventory levels are given by θT and θC, respectively. For example, $\theta = 1,000$ corresponds to a problem instance with 1,000 selling periods and initial inventory levels equal to 100. I compare the expected revenue loss under different heuristics for a wide range of θ's. In particular, since typical RM firms sell about 100-1,000 inventories per season (e.g., mid-size airplanes have about 100-500 seats and large-size hotels can easily have more than 1,000 rooms), I use θ ranging from 500-10,000.

I denote by *Static* the static price control developed in Gallego and van Ryzin (1997), and by *LRC* the linear rate control developed in Jasin (2014). As for my heuristics, I denote by *LPC-k* the LPC that simultaneously adjusts the prices of $k \geq m$ products in every period. (Recall that to ensure LPC adjusts at most k prices, I only need to find a proper transformation matrix. I select the transformation matrix following the proposed guideline in Section 2.4.) Correspondingly, I use *RSC-k* to denote the heuristic that adjusts the prices of the same k products as in *LPC-k* via exact re-optimization of DPP in every period, with an additional constraint that the prices of the unadjustable products remain the same as the static price. In addition to the said heuristics, I also test two simple modifications of *LPC-k* that only adjust the same k prices and can improve the *non-asymptotic* performance of the vanilla *LPC-k*. The first one is a projection-based LPC where, in each period, I apply LPC update formula followed by a projection into $[(1 - \alpha\%)p^D, (1 + \alpha\%)p^D]$; I denote the

and has only been recently studied in operations management (Bertsimas et al. 2010, Atar and Reiman 2012, Jasin 2014). In general, a linear control assumes the form of a baseline control plus a linear combination of past system perturbations. (This explains the forms of LRC and LPC.) While most existing literature on linear control focuses on finding a way to compute the optimal control parameters, my work explicitly constructs a particular form of linear control, which has certain desirable properties, and proposes a particular choice of parameters values that yields a strong performance guarantee. Needless to say, once the form is assumed, it may be possible to apply standard techniques in the literature to optimize the parameters of LPC. However, this is beyond the scope of this essay.

resulting heuristic by *Proα-k*. If α is small, *Proα-k* is very similar to static price control; if α is large, *Proα-k* is very similar to *LPC-k*. Per my discussions in Section 2.4, since I am using static price as my benchmark, I would ideally like to have a heuristic whose price trajectory stays as close as possible to the static price. However, since demands are random, I must also allow some room for price adjustments to account for demand variability. This motivates the use of projection as a way to control the intensity of price fluctuation. The second modification of *LPC-k* is a re-optimization-based LPC, denoted by *Hybβ-k*, where I re-optimize DPP at the first β updating times of the *2-Geometric* schedule and apply LPC in the remaining periods.

Experiment 1: Performance of LPC. Figure 2.1 illustrates the performance of *LPC-10* and other existing heuristics. Consistent with my asymptotic results, *LPC-10* performs much better than *Static*.[7] Figure 2.1 also shows that *LPC-10* performs slightly worse than *LRC* and *RSC-10*, which is not surprising because both *LRC* and *RSC-10* are known to have a slightly stronger performance guarantee of $\mathcal{O}(\log\theta)$ than LPC (Jasin 2014). I want to stress that although *RSC-10* performs very well, it is also very time-consuming (see Table 2.1). In contrast, *LPC-10* is computationally very fast. Admittedly, there is still a revenue gap between the "ideal but not implementable" *RSC-10* and *LPC-10*. The question is whether there is a cheap way to improve the performance of *LPC-10* without resorting to heavy frequent re-optimizations. It turns out that I can significantly narrow the gap between *RSC-10* and *LPC-10* by simple modifications of *LPC-10*. The first plot in Figure 2.2 shows that *Pro30-10*, which enforces the prices of LPC to fluctuate within a 30% band around the static price, can reduce the revenue loss gap by almost a half. This indicates that a simple projection can have a significant impact on revenue. (In general, I can also use product-dependent α parameters and optimize them by running an off-line Monte-Carlo optimization.) The second plot in Figure 2.2 further shows that *Hyb8-10*, which combines LPC with *only* 8 optimizations, can reduce the revenue loss gap by more than 75%. This is fairly impressive considering the fact that, even for small $\theta = 500$, *RSC-10* already requires 500 re-optimizations. It highlights the versatility of LPC for practical implementation; in particular, I can use LPC in combination with occasional re-optimizations in the case where frequent re-optimizations is clearly not feasible.

[7] It is interesting to note that not all linear price controls are guaranteed to perform well. For example, under 1-Periodic schedule, one intuitively appealing linear price control is $\hat{p}_t = p^D - \sum_{j=1}^{m} E^j \nabla p(\lambda^D) H \sum_{s=1}^{t-1} A\Delta_s$. Similar to LPC, this heuristic also adjusts prices to compensate for randomness in demand realizations. But, in contrast to LPC, this heuristic adjusts the price in a myopic manner; it attempts to fully correct the errors made in the previous period in the next period. Although this heuristic appears reasonable at first sight, my numerical experiments suggest that it is not even asymptotically optimal. This highlights that developing a linear price control that has strong performance is not a trivial task.

Experiment 2: Minimal price adjustment. In this experiment, I test the minimal adjustment property discussed in Section 2.4. The plots in Figure 2.3 show the comparison between *LPC-4* and *RSC-4*, as well as the two types of modified LPC with the same projection matrix as *LPC-4*. All these heuristics adjust the prices of the same $m = 4$ products. (Note that *LRC* cannot be included in this comparison because it cannot adjust prices of fewer than $n = 10$ products.) Similar to experiment 1, while *RSC-4* performs very well, it requires a lot of re-optimizations, which may not be feasible in practice. The two simple modifications of *LPC-4*, *Pro30-4* and *Hyb8-4*, which are computationally much cheaper, can attain a similar performance as *RSC-4*.

At the end of Section 2.4, I discussed the impact of increasing the number of adjustable products on revenue performance. Figure 2.4 illustrates my theoretical results. (See also Table 2.2.) The first plot in Figure 2.4 shows that, in comparison to *Static* that adjusts no prices at all, allowing $m = 4$ adjustable products yields a significant reduction in revenue loss. This is due to the minimal adjustment property of LPC. Beyond the initial four products, although allowing more adjustable products further decreases the revenue loss, its incremental benefit becomes much smaller. In particular, the plot shows that the impact of allowing two additional adjustable products (see the gap between *LPC-4* and *LPC-6*) captures almost half of the benefit of allowing six more adjustable products (see the gap between *LPC-4* and *LPC-10*). I observe the same phenomenon in the second plot in Figure 2.4 for *Hyb8* heuristics. This suggests that the managerial insights drawn from Theorem 2.4.2 still hold in network setting: If the seller wishes to adjust the prices of more than m products to increase revenue, then adjusting a few more products is sufficient to capture pretty much all the potential benefit of adjusting all products.

Experiment 3: Equivalent pricing with business constraints. In this experiment, I study a case where the seller has additional constraints on when and what prices to adjust. I assume that (1) the prices of products 5, 8 and 9 cannot be adjusted, (2) the prices of products 2, 3, 4 can *only* be adjusted in the second half of the selling season, and (3) the prices of products 6, 7, 10 can only be adjusted in the first half of the selling season. These are plausible constraints motivated by practical applications. For example, products 5, 8 and 9 can be viewed as corporate rate rooms that cannot be adjusted over time. Products 2-4 and 6, 7, 10 can be viewed as special rate rooms for certain events (e.g., conference) whose prices cannot be adjusted in a certain time window. Based on my discussions in Section 2.5, LPC can be automatically adapted to this setting via equivalent pricing with an original base of $\mathcal{B} = \{1, 2, 3, 4\}$; I denote this heuristic simply as *LPC*. Similar to previous experiments, I can apply re-optimized static price control with the additional constraints that certain prices cannot be adjusted in particular periods; I denote the resulting heuristic simply as *RSC*. It is also possible to use the modified LPC, which I denote as *Pro*α and *Hybk*, accordingly. Figure 2.5 shows that simple modifications of LPC, which is computationally easy, can attain a similar performance as *RSC* which requires frequent re-optimizations and may not be implementable in practice. This highlights the versatility of LPC for practical implementation in the presence of business constraints.

2.7 Closing Remarks

In this essay, I consider a standard dynamic pricing problem and propose a new family of pricing heuristics, which I call LPC. I show that LPC provides a strong improvement over static pricing: The revenue loss is reduced from $\mathcal{O}(\sqrt{\theta})$ to $\mathcal{O}(\log^2 \theta)$. In addition, it also has desirable features that can be used to address practical concerns. First, LPC only requires a single optimization and can be implemented in real-time, which makes it useful for solving large-scale problems where other computationally intensive heuristics are not viable. Second, LPC guarantees a strong revenue performance by adjusting the price of a few "important" products infrequently. This helps address the issue of acceptability of dynamic pricing in the eyes of customers due to excessive price adjustments. Third, LPC allows the seller to maintain an equivalent revenue performance via adjusting the prices of other products. This not only can be used to further reduce the number of required price changes per product, but also provides an extra flexibility for the sellers to manage his prices in the presence of various business constraints. My simulation results show that LPC not only has a good theoretical performance but also works well numerically. Furthermore, its performance can be further improved by simple modifications such as projection and occasional re-optimizations. To conclude, I believe that my work provides novel managerial insights that make dynamic pricing more applicable and practically appealing for real-world implementation.

2.8 Tables

Table 2.1: Simulation time for RSC-10, LPC-10 and Hyb8-10

θ	RSC-10	LPC-10	Hyb8-10
500	8305.0	13.3	209.7
5000	87552.4	86.2	212.3

Typical running time measured in milliseconds for a single simulation for selected heuristics.

Table 2.2: Comparison of revenue loss (R.L.) and revenue improvement (R.I.)

	% R.L. compared to revenue upper bound					% R.I. over static pricing control.				$\frac{\text{% R.I. of LPC-4}}{\text{% R.I. of LPC-10}}$
θ	Static	LPC-4	LPC-6	LPC-8	LPC-10	LPC-4	LPC-6	LPC-8	LPC-10	
500	5.94%	5.19%	4.59%	4.23%	4.15%	0.79%	1.43%	1.82%	1.90%	41.6%
1000	4.22%	3.00%	2.61%	2.60%	2.28%	1.28%	1.69%	1.69%	2.02%	63.3%
2000	2.99%	1.72%	1.57%	1.43%	1.37%	1.32%	1.46%	1.61%	1.67%	78.6%
3000	2.48%	1.25%	1.09%	1.05%	0.99%	1.27%	1.43%	1.47%	1.53%	82.8%
4000	2.13%	0.98%	0.86%	0.82%	0.77%	1.18%	1.30%	1.34%	1.40%	84.1%
5000	1.94%	0.81%	0.70%	0.69%	0.65%	1.15%	1.26%	1.28%	1.31%	88.0%
6000	1.81%	0.67%	0.61%	0.55%	0.58%	1.16%	1.22%	1.29%	1.25%	92.5%
7000	1.64%	0.59%	0.54%	0.50%	0.47%	1.07%	1.12%	1.16%	1.20%	89.1%
8000	1.58%	0.55%	0.48%	0.43%	0.40%	1.04%	1.12%	1.16%	1.20%	87.2%
9000	1.42%	0.50%	0.42%	0.41%	0.39%	0.94%	1.01%	1.02%	1.05%	89.3%
10000	1.34%	0.45%	0.42%	0.39%	0.35%	0.90%	0.93%	0.96%	1.00%	90.0%

2.9 Figures

Figure 2.1: Revenue loss under different heuristics

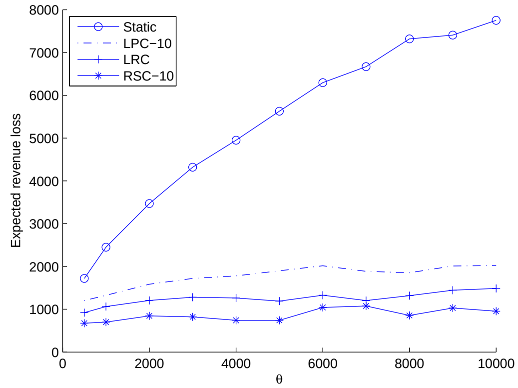
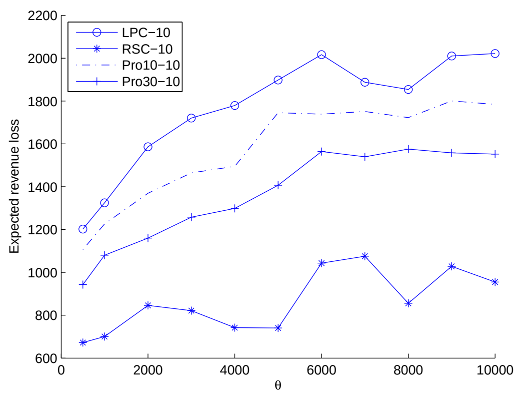


Figure 2.2: Improving LPC-10 using projection and occasional re-optimizations



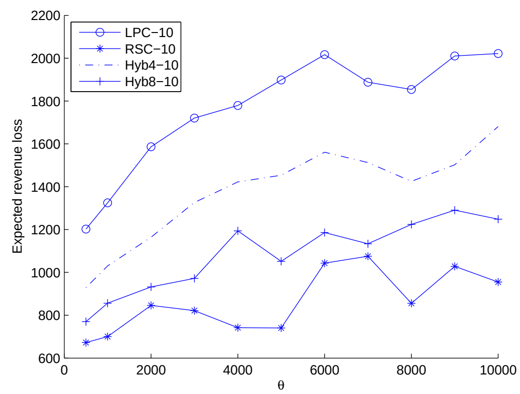
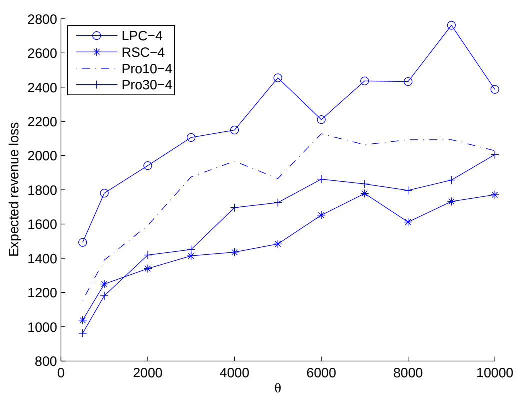


Figure 2.3: Improving LPC-4 using projection and occasional re-optimizations



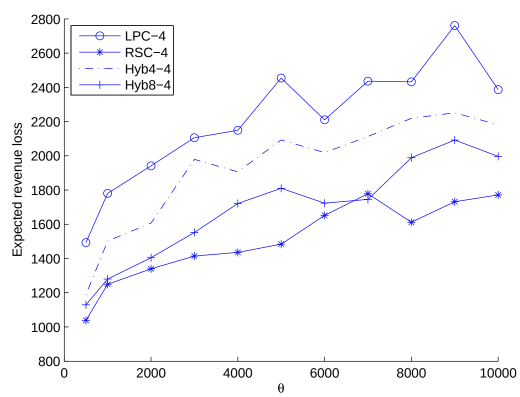


Figure 2.4: Revenue impact of the number of adjustable products for LPC and Hyb8

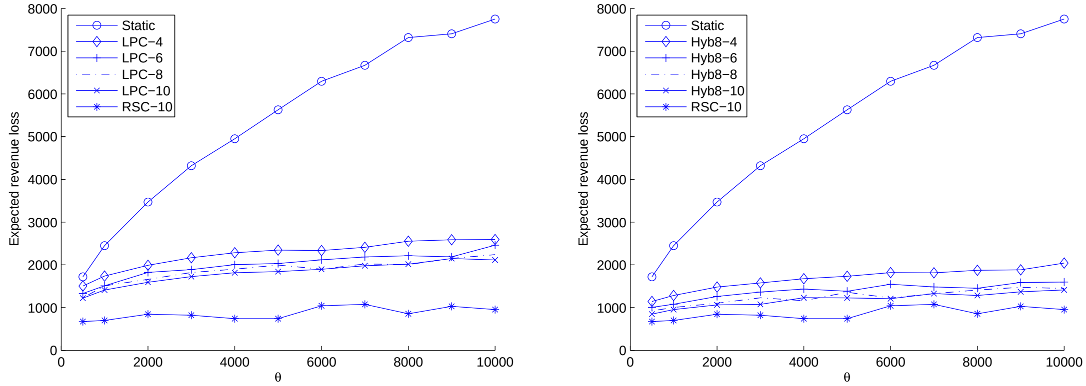
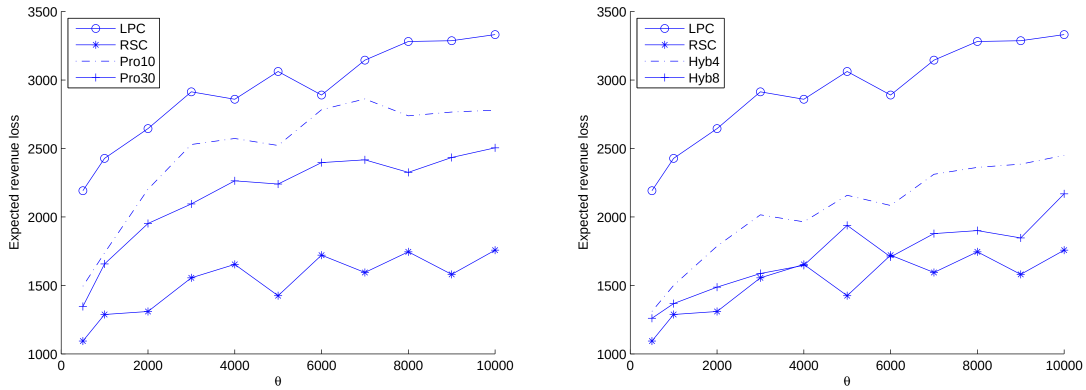


Figure 2.5: Improving LPC using projection and occasional re-optimizations



Chapter 3 Pricing with Unknown Demand: Parametric Case

3.1 Abstract

Consider a multi-period network revenue management (RM) problem where a seller sells multiple products made from multiple resources with finite capacity in an environment where the the demand function form is known but the exact parameters are unknown a priori (i.e., the parametric case). The objective of the seller is to jointly learn the demand and price the products to minimize his expected revenue loss. It is widely known in the literature that the revenue loss of any pricing policy in the parametric case is at least $\Omega(\sqrt{k})$. However, there is a considerable gap between this lower bound and the performance bound of the best known heuristic in the literature. To close the gap, I develop several self-adjusting heuristics with strong performance bound. For the general parametric case, my proposed Parametric Self-adjusting Control (PSC) attains a $\mathcal{O}(\sqrt{k})$ revenue loss, matching the theoretical lower bound. If the parametric demand function family further satisfies a well-separated condition, by taking advantage of passive learning, my proposed Accelerated Parametric Self-adjusting Control achieves a much sharper revenue loss of $\mathcal{O}(\log^2 k)$. All the proposed heuristics are computationally very efficient and can be used as a baseline for developing more sophisticated heuristics for large-scale problems.

3.2 Introduction

Revenue management (RM) has wide applications in many industries such as airlines, hotels, fashion goods, car rentals, etc.. (Talluri and van Ryzin 2005) The common trait in these industries is that the seller uses a fixed amount of resources, which cannot be replenished during the selling season, to produce products which are used to satisfy the uncertain demand. Any unused resources have little salvage value when the selling season ends. Since the uncertainty of demand only gradually resolves as customers with heterogenous willingness-to-pay arrive over time, the seller needs to make important operational decisions to ensure that the products are sold to the right customer at the right time at the right price to maximize the profit. One of such operational leverages that the sellers often use is dynamic pricing.

An effective use of dynamic pricing requires a good knowledge of the market response to price changes. This requirement poses a practical problem because usually sellers do not possess such information. Therefore, an important question that sellers need to address is how to use dynamic pricing when the demand function is not perfectly known a priori. Since demand estimation is always subject to estimation errors, one common approach the sellers use in practice is to rely on frequent re-optimization. The idea is that the inaccuracy of demand estimation can be mitigated out by frequent re-optimization. However, for most RM applications, the problem being re-optimized is very large in size which makes frequent re-optimization computationally difficult. Therefore, an important research question which is both of theoretical interest and practical implications is how sellers should learn the demand function and adjust price over time to achieve a good performance without necessarily incurring a lot of computational burden. In this essay, I develop a joint learning and pricing heuristic to address this question.

Dealing with computational burden. There is a large body of OM literature that studies this dynamic pricing problem. It turns out that even for the simpler case where the demand function is known, finding the optimal pricing policy is already computationally difficult due to the curse of dimensionality of Dynamic Programming (DP). As a result, most of the early work in this area focuses specifically on the simpler known demand case, and develop computationally implementable heuristics instead of solving for the optimal pricing control. (See Bitran and Caldentey (2003) and Elmaghraby and Keskinocak (2003) for a comprehensive literature review for dynamic pricing with known demand functions.) One popular approach is to develop heuristics using Approximate Dynamic Programming techniques. (See Erdelyi and Topaloglu (2011) and Kunnumkal and Topaloglu (2010).) This method can adaptively adjust the price over time according to the demand realization, but it is still computationally intensive and may not be amenable for large scale problems. Another stream of work is based on solving a deterministic counterpart of the original stochastic problem. Some seminar works along this line are Gallego and van Ryzin (1994, 1997). They propose the well-known static price control where the static price obtained by solving the deterministic problem at the beginning is used throughout the entire selling season. However, since this heuristic cannot adjust the price over time to account for the demand realization, there is room for improvement. Indeed, one straightforward extension is to *re-optimize* the deterministic problem over time to utilize the progressively revealed information on demand. Maglaras and Meissner (2006) is the first to show that re-optimized static price control has at least the same asymptotic performance as static price control without re-optimization. A recent work by Jasin (2014) proves a much tighter bound for the expected revenue for re-optimized static price control and also develops a re-optimization free self-adjusting control with comparable good performance. A follow-up work by Chen et al. (2016) develops another simple self-adjusting control heuristic with similar performance and shows that in order to get good performance, it is sufficient to adjust the prices of a small subset of all products.

Incorporating demand learning. Motivated by the fact that demand function is often unknown in practice, the more recent literature on dynamic pricing has shifted the focus to finding good dynamic pricing policies when the seller also needs to learn the underlying demand function. The central tension is learning the demand function (exploration) and using the optimal price based on the estimated demand function (exploitation). The more time the seller spends on learning demand function by price experimentation, the less opportunity there is for him to fully tap the knowledge of the demand function he has learned. However, if he spends too little time on exploration, he may end up with a poor estimation of the demand function which limits the revenue he can earn during exploitation.

Learning demand is a crucial part of this problem and researchers either take a non-parametric approach or a parametric approach for demand estimation. It has been proved in the literature that the *revenue loss* of any feasible pricing control (the absolute revenue loss to a clairvoyant optimal revenue) is at least $\Omega(\sqrt{k})$ where k is the size of the problem. (See Section 3.3 for more details on the asymptotic setting.) Therefore, a considerable size of literature focuses on developing heuristics to approach this lower bound. In the non-parametric case, the seller has no information about the structure of the demand function other than some regularity conditions. Besbes and Zeevi (2009) is one of the first to investigate this problem in the *single product* setting. They propose a non-parametric heuristic that attains a revenue loss of $\mathcal{O}(k^{3/4})$. A recent work by Wang et al. (2014) develops a more sophisticated heuristic that achieves an improved revenue loss of $\mathcal{O}(\sqrt{k}\log^{9/2} k)$. However, their bound only applies to *single product case* where some special structure of the problem can be utilized. It is not clear how to extend their heuristic to a more general setting where the seller sells multiple products made from multiple resources. The only heuristic that can be used in this more general setting is developed by Besbes and Zeevi (2012). However, the performance bound of their heuristic deteriorates as the number of products increases.

The non-parametric case assumes no prior information about the demand function, but in reality, sellers may have some useful knowledge of the structure of the demand function based on their past experience and historical demand data. Motivated by this observation, in this essay, I take the parametric approach in which the underlying demand function belongs to a family of parameterized functions. The seller knows the functional form of the demand function but doesn't know the underlying parameters. Intuitively, the revenue loss bound of parametric case can be no worse than the non-parametric case since all non-parametric heuristics can be used in a parametric setting. Surprisingly, the best revenue loss bound for parametric approach in the literature is no sharper than that for non-parametric approach. This gives rise to two important theoretical questions. Is there a heuristic that *exactly* achieves the revenue loss lower bound $\Theta(\sqrt{k})$ in the parametric case? Is the structural information of the demand functional form useful at all?

Other related literature. Although the performance result for the parametric case is limited, there is a stream of literature that discovers some nice results for a more stylized *uncapacitated* problem where there is no inventory constraints. Broder and Rusmevichientong (2012) study the general parametric demand family and propose a heuristic based on Maximum Likelihood Estimation (MLE) which achieves the lowest possible revenue loss bound $\mathcal{O}(\sqrt{k})$. They also show that if the demand function family satisfies the so-called *well-separated* condition, the bound reduces to $\mathcal{O}(\log k)$. Other researchers study linear or generalized linear demand families and develop various kinds of heuristics based on Least Square Estimation approach. Their work highlights the impact of information accumulation rate on the expected revenue. (See den Boer and Zwart (2014), Keskin and Zeevi (2014), den Boer (2014)). However, all these papers essentially ignore the inventory constraints which is present in most RM applications. From a technical point of view, finding the optimal price is much more straightforward in the uncapacitated problems than the capacitated problems I study where decisions in different periods affect each other via the inventory constraints. Note also that if the initial inventory is very large, the capacitated problem reduces to an uncapacitated problem. Therefore, the problem I study is indeed more general than theirs. It is worth noting that there are other approaches that tackle this uncapacitated problem. For example, both Harrison et al. (2012) and Farias and van Roy (2010) take a Bayesian approach to model the uncertainty of the underlying demand function. This approach can be traced back to earlier economics literature where the importance price experimentation is investigated. (See Rothschild (1974), McLennan (1984), Easley and Kiefer (1988)).

Although in reality the demand function may change over time, I do not address this issue in this essay and only focus on the case when the underlying demand function is stationary. There is some recent work in OM that looks into the demand learning problem in a non-stationary environment (e.g., Besbes et al. (2015), Keskin and Zeevi (2016)). There is also a fast growing literature of Convex Online Optimization in computer science literature that studies a similar problem. I refer the readers to Hazan et al. (2016) for a literature review.

Proposed heuristic and contributions. For the general parametric case, I develop a heuristic called *Parametric Self-adjusting Control* (PSC) that combines Maximum Likelihood (ML) estimation with self-adjusting price updates, and derive an analytical performance bound. To the best of my knowledge, this is the first essay that develops a joint learning and pricing heuristic in the network RM setting with parametric demand model. I show that PSC is rate-optimal. To be precise, the revenue loss of PSC is $\mathcal{O}(\sqrt{k})$ (Theorem 3.4.1), which matches the theoretical lower bound. In addition, I also show that if the parametric demand function family satisfies the so-called well-separated condition which was first introduced by Broder and Rusmevichientong (2012), then there exists heuristics that can outperform the $\Omega(\sqrt{k})$ lower bound. I develop an *Accelerated Parametric Self-adjusting Control* (APSC), a variation of PSC, that attains a much sharper performance bound of $\mathcal{O}(\log^2 k)$ (Theorem 3.5.1).

My contribution is two-fold. On the theoretical end, I develop a heuristic for the most general joint learning and pricing problem for the parametric case that achieves the best revenue loss rate among all feasible controls. On the practical end, my re-optimization free heuristics are computationally tractable, overturning the common impression that re-optimization is a necessity to achieve good performance when the underlying demand is unknown. It also highlights the applicability of self-adjusting idea in dynamic pricing problems, which can provide guidelines for companies to develop more sophisticated dynamic pricing controls.

The remainder of the essay is organized as follows. I first formulate the problem in Section 3.3. I then introduce PSC and evaluate its performance for the general parametric in Section 3.4. Next, I introduce the well-separated condition and APSC, and evaluate the performance of APSC in Section 3.5. Finally, I conclude the essay in Section 3.6. All the proofs of the results can be found in Appendix A.3.

3.3 Problem Formulation

Notation. The following notation will be used throughout the essay. (Other notation will be introduced when necessary.) Denote by $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{R}_{++}$ the set of real, nonnegative real, and positive real numbers respectively. For column vectors $a = (a_1; \ldots ; a_n) \in \mathbb{R}^n, b = (b_1; \ldots ; b_n) \in \mathbb{R}^n$, denote by $a \succeq b$ if $a_i \geq b_i$ for all i, and by $a \succ b$ if $a_i > b_i$ for all i. Similarly, denote by $\mathbb{Z}$, $\mathbb{Z}_+$, and $\mathbb{Z}_{++}$ the set of integers, nonnegative integers, and positive integers respectively. Denote by $\cdot$ the inner product of two vectors and by $\otimes$ the tensor product of sets or linear spaces. I use a prime to denote the transpose of a vector or a matrix, an I to denote an identity matrix with a proper dimension, and an $\mathbf{e}$ to denote a vector of ones with a proper dimension. For any vector $v = [v_j] \in \mathbb{R}^n$, $||v||_p := (\sum_{j=1}^n |v_j|^p)^{1/p}$ is its p-norm ($1 \leq p \leq \infty$) and, for any real matrix $M = [M_{ij}] \in \mathbb{R}^{n\times n}$, $||M||_p := \sup_{||v||_p=1} ||Mv||_p$ is its induced p-norm. For example, $||M||_1 = \max_{1\leq j\leq n} \sum_{i=1}^n |M_{ij}|$, $||M||_2 =$ the largest eigenvalue of $M'M$, and $||M||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^n |M_{ij}|$. (Note that $||M||_1 = ||M'||_\infty$.) For any function $f : X \to Y$, denote by $||f(.)||_\infty := \sup_{x\in X} ||f(x)||_\infty$ the infinity-norm of f. I use ∇ to denote the usual derivative operator and use a subscript to indicate the variables with respect to which this operation is applied to. (No subscript ∇ means that the derivative is applied to all variables.) If $f : \mathbb{R}^n \to \mathbb{R}$, then $\nabla_x f = (\frac{\partial f}{\partial x_1}; \ldots; \frac{\partial f}{\partial x_n})$; if, on the other hand, $f = (f_1; \ldots; f_n) : \mathbb{R}^n \to \mathbb{R}^n$, then

$$\nabla_x f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}.$$

The model. I consider the problem of a monopolist selling his products to incoming customers during a finite selling season and aiming to maximize his total expected revenue. There are n types of products, each of which is made up of a combination of a subset of m types of resources. For example, in the airline setting, a product refers to a multi-flight itinerary and a resource refers to a seat in a single-leg flight; in the hotel setting, a product refers to a multi-day stay and a resource refers to a one-night stay at a particular room. Denote by $A = [A_{ij}] \in \mathbb{R}^{m\times n}$ the *resource consumption matrix*, which characterizes the types and amounts of resources needed by each product. To be precise, a single unit of product j requires A_{ij} units of resource i. Without loss of generality, I assume that the matrix A has full row rank. (If this is not the case, then one can apply the standard row elimination procedure to delete the redundant rows. See Jasin (2014).) Denote by $C \in \mathbb{R}^m$ the vector of initial capacity levels of all resources at the beginning of the selling season. Since, in many industries (e.g., hotels and airlines), replenishment of resources during the selling season is either too costly or simply not feasible, following the standard model in the literature (Gallego and van Ryzin 1997), I will assume that the seller has no opportunity to procure additional units of resources during the selling season. In addition, I also assume without loss of generality that the remaining resources at the end of the selling season have zero salvage value.

Consider a discrete-time model with T *decision* periods, indexed by $t = 1, 2, ..., T$. At the beginning of period t, the seller first decides the price $p_t = (p_{t,1}; \ldots; p_{t,n})$ for his products, where p_t is chosen from a convex and compact set $\mathcal{P} = \otimes_{l=1}^n [\underline{p}_l, \bar{p}_l] \subseteq \mathbb{R}^n$ of feasible price vectors. The posted price p_t, in turn, induces a demand, or sale, for one of the products with a certain probability. Here, I implicitly assume that at most one sale for one product occurs in each period. This is without loss of generality since I can always slice the selling season fine enough to guarantee that at most one customer arrives in each period. Let $\Delta^{n-1} := \{(x_1; \ldots; x_n) \in \mathbb{R}^n | \sum_{i=1}^n x_i \leq 1$, and $x_i \geq 0$ for all $i\}$ denote the standard $(n-1)$-

simplex. Contrary to most existing RM literature where it is assumed that the seller knows the purchasing probability vector under any price a priori, in this essay, I simply assume that it can be estimated using statistical learning approaches. Specifically, let $\lambda(.;.) : \mathcal{P} \times \Theta \to \Delta^{n-1}$ denote the family of *demand functions* where Θ is a compact subset of $\mathbb{R}^q$ and $q \in \mathbb{Z}_{++}$ is the number of unknown parameters. I denote by θ^* the true parameters for the underlying demand function[1]. Under the parametric demand case, the seller knows the functional form of demand $\lambda(.;\theta)$ for any $\theta \in \Theta$, but he does not know θ^*. Let $\Lambda_\theta := \{\lambda(p;\theta) : p \in \mathcal{P}\}$ denote the set of feasible demand rates under some parameter vector $\theta \in \Theta$. I assume that Λ_θ is convex. (It can be shown that, under the most commonly used parametric function families such as linear, logit, and exponential demand, Λ_θ is convex for all $\theta \in \Theta$.)

Let $D_t(p_t) = (D_{t,1}(p_t); \ldots; D_{t,n}(p_t))$ denote the vector of realized demand in period t under price p_t. It should be noted that, although demands for different products in the same period are not necessarily independent, demands over different periods are assumed to be independent (i.e., D_t only depends on the posted price p_t in period t). By definition, I have $D_t(p_t) \in \mathcal{D} := \{D \in \{0,1\}^n : \sum_{j=1}^n D_j \leq 1\}$ and $\mathbf{E}_{\theta^*}[D_t(p_t)] = \lambda(p_t;\theta^*)$. This allows me to write $D_t(p_t) = \lambda(p_t;\theta^*) + \Delta_t(p_t)$, where $\Delta_t(p_t)$ is a zero-mean random vector. For notational simplicity, whenever it is clear from the context which price p_t is being used, I will simply write $D_t(p_t)$ and $\Delta_t(p_t)$ as D_t and Δ_t respectively. The sequence $\{\Delta_t\}_{t=1}^T$ will play an important role in my analysis later. The one-period expected revenue function under θ is given by the revenue function defined as $r(p;\theta) := p \cdot \lambda(p;\theta)$. I assume that for all $\theta \in \Theta$, $\lambda(p;\theta)$ is invertible (see parametric family assumptions below); so, by abuse of notation, I can write $r(p;\theta) = p \cdot \lambda(p;\theta) = \lambda \cdot p(\lambda;\theta) = r(\lambda;\theta)$. I make the following regularity assumptions about $\lambda^*(.)$ and $r^*(.)$. (These are all standard assumptions in the literature and are immediately satisfied by commonly used demand function families such as linear, logit and exponential.)

Parametric Family Assumptions. *There exist positive constants $\bar{r}, \underline{v}, \bar{v}, \omega, \underline{v}, \bar{v}$ such that for all $p \in \mathcal{P}$ and for all $\theta \in \Theta$:*

P1. *$\lambda(.;\theta) : \mathcal{P} \to \Lambda_\theta$ is in $\mathsf{C}^2(\mathcal{P})$ and it has an inverse function $p(.;\theta) : \Lambda_\theta \to \mathcal{P}$ that is in $\mathsf{C}^2(\Lambda_\theta)$. $\lambda(p;.) : \Theta \to \Delta^{n-1}$ is in $\mathsf{C}^1(\Theta)$. For all $\lambda, \lambda' \in \Lambda_\theta$, $||p(\lambda;\theta) - p(\lambda';\theta)||_2 \leq \omega||\lambda - \lambda'||_2$.*

[1]Although I implicitly assume that the demand function is stationary, my heuristics can be extended to accommodate some time-varying demand scenarios if the time-dependence of demand function has certain structural form. For example, in the fashion industry, irrespective of the condition of the market, the seller usually knows the fractions of the total sales that will be realized at multiple milestones over the selling season. This can be captured by incorporating in my demand model additively a time factor which is a known time-dependent fraction of an unknown total market size. Note that this model can be handled under the current stationary estimation framework by treating the unknown total market size as an additional parameter.

P2. *For all* $1 \le i, j \le n$, $||\lambda(p;\theta) - \lambda(p;\theta^*)||_2 \le \omega||\theta - \theta^*||_2$, $|\frac{\partial \lambda_j}{\partial p_i}(p;\theta) - \frac{\partial \lambda_j}{\partial p_i}(p;\theta^*)| \le \omega||\theta - \theta^*||_2$.

P3. $||r(.;\theta)||_\infty \le \bar{r}$ *and* $r(.;\theta)$ *is strongly concave in* λ, *i.e.,* $-\bar{v}I \preceq \nabla^2_{\lambda\lambda} r(\lambda;\theta) \preceq -\underline{v}I$ *for all* $\lambda \in \Lambda_\theta$.

P4. *There exists a set of turn-off prices* $p_j^\infty \in \mathbb{R} \cap \{\infty\}$ *for* $j = 1, \ldots, n$ *such that for any* $p = (p_1; \ldots; p_n)$, $p_j = p_j^\infty$ *implies that* $\lambda_j(p;\theta) = 0$ *for all* $\theta \in \Theta$.

Assumptions P1 and P2 are fairly natural and are easily satisfied by many demand functions, e.g., linear demand, logit demand, and exponential demand. As for Assumption P3, the boundedness of $r(.;\theta)$ follows from the compactness of Θ and Λ_θ and the smoothness of $r(.;\theta)$. The strong concavity of $r(.;\theta)$ as a function of λ is a standard assumption in the literature and is satisfied by many commonly used demand functions such as linear, exponential, and logit functions. It should be noted that although some of these functions, such as logit, do not naturally correspond to a concave revenue function when viewed as a function of p, they are nevertheless concave when viewed as a function of λ. This highlights the benefit of treating revenue as a function of demand rate instead of as a function of price. Assumption P4 is common in the literature. (See Besbes and Zeevi (2009) and Wang et al. (2014).) In particular, the existence of turn-off prices p_j^∞ allows the seller to effectively shut down the demand for any product whenever needed, e.g., in the case of stock-out. Additional assumptions will be provided later.

Admissible controls and the induced probability measures. Let $D_{1:t} := (D_1, D_2, \ldots, D_t)$ and $p_{1:t} := (p_1, p_2, \ldots, p_t)$ denote respectively the observed vectors of demand and price realizations up to and including period t. Let $\mathcal{H}_t$ denote the σ-field generated by $D_{1:t}$ and $p_{1:t}$. I define a *control* π as a sequence of functions $\pi = (\pi_1, \pi_2, \ldots, \pi_T)$, where π_t is a $\mathcal{H}_{t-1}$-measurable real function that maps the history $D_{1:t-1}$ to $\mathcal{P} \cup \{p^\infty\}$. This class of controls is often referred to as *non-anticipating controls* because the decision in each period depends only on the accumulated observations up to the beginning of the period. Under policy π, the seller sets the price in period t equal to $p_t^\pi = \pi_t(D_{1:t-1}; p_{1:t-1})$ almost surely (a.s.). Let Π_θ denote the set of all *admissible controls*. That is,

$$\Pi_\theta := \left\{ \pi : \sum_{t=1}^{T} AD_t(p_t^\pi;\theta) \le C \;\text{ and }\; p_t^\pi = \pi_t(\mathcal{H}_{t-1}) \;\; a.s. \right\}.$$

(Although the true underlying parameter is θ^*, I define above the set of admissible controls for any $\theta \in \Theta$.) Note that I require the capacity constraint to hold almost surely for all

$\pi \in \Pi_\theta$, which can be satisfied by using turn-off prices p^∞ in case of stock-out. Let $\mathbb{P}_t^{\pi,\theta}$ denote the induced probability measure of $D_{1:t}$ under admissible control $\pi \in \Pi_\theta$. For any realization $D_{1:t} = d_{1:t} := (d_1, d_2, \ldots, d_t)$, where $d_s = (d_{s,j}) \in \mathcal{D}$, $s = 1, \ldots, t$, I have:

$$\mathbb{P}_t^{\pi,\theta}(d_{1:t}) \;=\; \prod_{s=1}^{t}\left[\left(1 - \sum_{j=1}^{n} \lambda_j(p_s^\pi;\theta)\right)^{\left(1-\sum_{j=1}^{n} d_{s,j}\right)} \prod_{j=1}^{n} \lambda_j(p_s^\pi;\theta)^{d_{s,j}}\right],$$

where $p_s^\pi = \pi_s(d_{1:s-1})$. (By definition of $\lambda(p;\theta)$, the term $1-\sum_{j=1}^{n}\lambda_j(p_s^\pi;\theta)$ can be interpreted as the probability of no-purchase in period s under price p_s^π.) For notational simplicity, I will write $\mathbb{P}_\theta^\pi := \mathbb{P}_T^{\pi,\theta}$ and denote by $\mathbf{E}_\theta^\pi$ the expectation with respect to probability measure $\mathbb{P}_\theta^\pi$. Total expected revenue under $\pi \in \Pi_\theta$ is then given by:

$$R_\theta^\pi \;=\; \mathbf{E}_\theta^\pi\left[\sum_{t=1}^{T} (p_t^\pi)' D_t(p_t^\pi;\theta)\right].$$

Whenever it is clear that the prices $p_{1:t} \in \mathcal{P}^t$ are generated by an admissible control π, it is also convenient to write $\mathbb{P}_t^{p_{1:t},\theta}(d_{1:t}) = \prod_{s=1}^{t}[(1-\sum_{j=1}^{n}\lambda_j(p_s;\theta))^{(1-\sum_{j=1}^{n} d_{s,j})}\prod_{j=1}^{n}\lambda_j(p_s;\theta)^{d_{s,j}}]$.

Maximum likelihood estimator. As noted earlier, the seller does not know the true parameter vector θ^*. But, he can estimate this parameter vector using statistical methods. In this essay, I will focus primarily on *Maximum Likelihood* (ML) estimation. (The analysis of other statistical methods is beyond the scope of this essay.) The behavior of ML estimator has been intensively studied in the statistics literature. It not only has certain desirable theoretical properties, but is also widely used in practice. To guarantee the regular behavior of ML estimator, certain statistical conditions need to be satisfied. To formalize these conditions, it is convenient to first consider the distribution of a sequence of demand realizations when a sequence of $\tilde{q} \in \mathbb{Z}_{++}$ fixed price vectors $\tilde{p} = (\tilde{p}^{(1)}, \tilde{p}^{(2)}, \ldots, \tilde{p}^{(\tilde{q})}) \in \mathcal{P}^{\tilde{q}}$ have been applied. For all $d_{1:\tilde{q}} \in \mathcal{D}^{\tilde{q}}$, I define the distribution $\mathbb{P}^{\tilde{p},\theta}$ as follows:

$$\mathbb{P}^{\tilde{p},\theta}(d_{1:\tilde{q}}) \;=\; \prod_{s=1}^{\tilde{q}}\left[\left(1 - \sum_{j=1}^{n} \lambda_j(\tilde{p}^{(s)};\theta)\right)^{\left(1-\sum_{j=1}^{n} d_{s,j}\right)} \prod_{j=1}^{n} \lambda_j(\tilde{p}^{(s)};\theta)^{d_{s,j}}\right].$$

Let $\mathbf{E}_\theta^{\tilde{p}}$ denote the expectation with respect to $\mathbb{P}^{\tilde{p},\theta}$. The PSC and APSC that I will develop later use a set of "exploration prices" $\tilde{p}$ in the first L periods and then use maximum likelihood estimation to estimate the demand parameters. The exploration prices that I use need to satisfy the following conditions to guarantee the regular behavior of ML estimator:

Statistical Conditions on Exploration Prices. *There exist constants* $0 < \lambda_{\min} < \lambda_{\max} < 1$, $c_f > 0$, *and a sequence of prices* $\tilde{p} = (\tilde{p}^{(1)}, \ldots, \tilde{p}^{(\tilde{q})}) \in \mathcal{P}^{\tilde{q}}$ *such that:*

S1. $\mathbb{P}^{\tilde{p},\theta}(.) \neq \mathbb{P}^{\tilde{p},\theta'}(.)$ *whenever* $\theta \neq \theta'$;

S2. *For all* $\theta \in \Theta$, $1 \leq k \leq \tilde{q}$ *and* $1 \leq j \leq n$, $\lambda_j(\tilde{p}^{(k)}; \theta) \geq \lambda_{\min}$ *and* $\sum_{j=1}^{n} \lambda_j(\tilde{p}^{(k)}; \theta) \leq \lambda_{\max}$.

S3. *For all* $\theta \in \Theta$, $\mathcal{I}(\tilde{p}, \theta) \succeq c_f I$ *where* $\mathcal{I}(\tilde{p}, \theta) := [\mathcal{I}_{i,j}(\tilde{p}, \theta)] \in \mathbb{R}^{q \times q}$ *is a* q *by* q *matrix defined as*

$$\mathcal{I}_{i,j}(\tilde{p}, \theta) \quad = \quad \mathbf{E}_{\theta}^{\tilde{p}} \left[-\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log \mathbb{P}^{\tilde{p},\theta}(D_{1:\tilde{q}}) \right].$$

I call $\tilde{p}$ the *exploration prices*. Some comments are in order. S1 and S2 are crucial to guarantee that the estimation problem is well-defined, i.e., the seller is able to identify the true parameter vector by observing sufficient demand realizations under the exploration prices $\tilde{p}$. (If this is not the case, then the estimation problem is ill-defined and there is no hope for learning the true parameter vector.) The symmetric matrix $\mathcal{I}(\tilde{p}, \theta)$ defined in S3 is known as the *Fisher information* matrix in the literature, and it captures the amount of information that the seller obtains about the true parameter vector using the exploration prices $\tilde{p}$. S3 requires the Fisher matrix to be strongly positive definite; this is needed to guarantee that the seller's information about the underlying parameter vector strictly increases as he observes more demand realizations under $\tilde{p}$. All the results in this section require assumptions P1-P4 and S1-S3 to hold.

Remark 3.3.1 *I want to point out that, given the demand function family, it is easy to find such exploration prices. For example, for linear and exponential demand function families, any* $\tilde{q} = n + 1$ *price vectors* $\tilde{p}^{(1)}, \ldots, \tilde{p}^{(n+1)}$ *constitute a set of exploration prices if (a) they are all in the interior of* $\mathcal{P}$ *and (b) the vectors* $(1; \tilde{p}^{(1)})$, $\ldots$, $(1; \tilde{p}^{(n+1)}) \in \mathbb{R}^{n+1}$ *are linearly independent. For logit demand function family, any* $\tilde{q} = 2$ *price vectors* $\tilde{p}_1, \tilde{p}_2$ *constitute a set of exploration prices if (a) they are both in the interior of* $\mathcal{P}$ *and (b)* $\tilde{p}_i^{(1)} \neq \tilde{p}_i^{(2)}$ *for all* $i = 1, \ldots, n$. *The choice of exploration prices is related to the literature of optimum experimental design. Although it is possible to "optimally" choose the exploration prices using techniques in optimal experiment design, it is beyond the scope of this essay. Interested readers are referred to Pzman (2013) for more details.*

The deterministic formulation and performance measure. It is common in the literature to consider the deterministic analog of the dynamic pricing problem. Specifically, for any $\theta \in \Theta$, define:

$$(\mathrm{P}(\theta)) \qquad J_\theta^D := \max_{p \in \mathcal{P}} \left\{ \sum_{t=1}^{T} r(p_t;\theta) : \ \sum_{t=1}^{T} A\lambda(p_t;\theta) \preceq C \right\},$$

$$\text{or equivalently,} \quad (\mathrm{P}_\lambda(\theta)) \qquad J_\theta^D := \max_{\lambda_t \in \Lambda_\theta} \left\{ \sum_{t=1}^{T} r(\lambda_t;\theta) : \ \sum_{t=1}^{T} A\lambda_t \preceq C \right\}.$$

By assumption P3, $\mathrm{P}_\lambda(\theta)$ is a convex program and it computationally easy to solve. (To avoid triviality, I assume that for all $\theta \in \Theta$, $\mathrm{P}(\theta)$ is feasible.) It can be shown that J_θ^D is in fact an upper bound for the expected revenue of any admissible control. That is, $R_\theta^\pi \leq J_\theta^D$ for all $\pi \in \Pi_\theta$. (See Besbes and Zeevi (2012) for proof.) This allows me to use J_θ^D as a benchmark to quantify the performance of any admissible pricing control. In this essay, I follow the common convention and define the expected revenue loss of an admissible control $\pi \in \Pi_{\theta^*}$ as $\rho^\pi := J_{\theta^*}^D - R_{\theta^*}^\pi$. Denote by $p^D(\theta)$ (resp. $\lambda^D(\theta)$) the optimal solution of $\mathrm{P}(\theta)$ (resp. $\mathrm{P}_\lambda(\theta)$). In addition, denote by $\mu^D(\theta)$ the optimal dual solution corresponding to the capacity constraints of $\mathrm{P}(\theta)$. (Note that $\mu^D(\theta)$ is also the optimal dual solution corresponding to the capacity constraints of $\mathrm{P}_\lambda(\theta)$.) Observe that $\mathrm{P}(\theta^*)$ is equivalent to P defined in the sense that $\lambda^D(\theta^*) = \lambda^D, p^D(\theta^*) = p^D, \mu^D(\theta^*) = \mu^D$, and $J_{\theta^*}^D = J^D$. Let $\mathrm{Ball}(x, r)$ denote a Euclidean ball centered at x with radius r. I state my last parametric assumption below:

P5. (Interior Assumption) *There exists* $\phi > 0$ *such that* $\mathrm{Ball}(p^D(\theta^*), \phi) \subseteq \mathcal{P}$.

Assumption P5 is sufficiently mild and is satisfied by most problem instances. Intuitively, it states that the static price should neither be too low that it attracts too much demand nor too high that it induces no demand. The same interior assumption has also been made in Jasin (2014) and Chen et al. (2016).

Asymptotic setting. As briefly discussed in Section 3.2, most revenue management industries can be categorized as either moderate or large size. Thus, following the standard convention in the literature (e.g., Besbes and Zeevi (2009) and Wang et al. (2014)), I will consider a sequence of increasing problems where the length of the selling season and the initial inventory levels are both scaled by a factor of $k > 0$. (One can interpret k as the *size* of the problem. For example, $k = 500$ could correspond to a flight with capacity 500 seats and $k = 5,000$ could correspond to a large hotel with capacity 5,000 rooms.) To be precise, in the k^{th} problem, the length of selling season and the initial inventory levels are given by kT and kC respectively whereas the optimal deterministic solution is still $\lambda^D(\theta)$ and the optimal dual solution is still $\mu^D(\theta)$. Let $\rho^\pi(k)$ denote the expected revenue loss under admissible control $\pi \in \Pi_{\theta^*}$ for the problem with scaling factor k. I am primarily interested in identifying the order of $\rho^\pi(k)$ for large k. Intuitively, one would expect that a better-performing control should have a revenue loss which grows relatively slowly with respect to k. The following notation will be used throughout the remainder of the essay. For any two functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, I write $f(k) = \Omega(g(k))$ if there exists $M > 0$ independent of k such that $f(k) \geq Mg(k)$. Similarly, I also write $f(k) = \Theta(g(k))$ if there exists $M, K > 0$ independent of k such that $Mg(k) \leq f(k) \leq Kg(k)$, and write $f(k) = \mathcal{O}(g(k))$ if there exists $M > 0$ independent of k such that $f(k) \leq Mg(k)$.

3.4 General Demand Function Family

I am now ready to discuss my heuristic for the general family of parametric demand. My main result in this section is to show that PSC is *rate-optimal*, i.e., it attains the performance lower bound. It has been repeatedly shown in the literature (e.g., Besbes and Zeevi (2012), Broder and Rusmevichientong (2012), Wang et al. (2014)) that, in the most general setting, no admissible pricing control can have a better performance than $\Omega(\sqrt{k})$, i.e., $\rho^\pi(k) = \Omega(\sqrt{k})$ for all $\pi \in \Pi$. This obviously poses a fundamental limitation on the performance of any pricing control that I could hope for. An important question of both theoretical and practical interest is whether this lower bound is actually tight and whether there exists an easily implementable pricing control that guarantees a $\mathcal{O}(\sqrt{k})$ revenue loss. In the general parametric setting with only a *single* product and *without* capacity constraints (i.e., the uncapacitated setting), this question has been answered by Broder and Rusmevichientong (2012). If, on the other hand, the resources have limited capacity (i.e., the capacitated setting), Lei et al. (2014) recently propose a hybrid heuristic that guarantees a $\mathcal{O}(\sqrt{k})$ revenue loss. Thus, the question of the attainability of the lower bound in the single-product setting has been completely resolved. As for the general parametric setting with multiple products and capacity constraints, I am not aware of any result that guarantees a $\mathcal{O}(\sqrt{k})$ revenue loss. The heuristics analyzed in Wang et al. (2014) and Lei et al. (2014) are not easily generalizable to multiproduct setting. (This is because their heuristics exploit the structure of the optimal deterministic solution in the single-product setting. Unfortunately, no analogs of such structures exist in the multiproduct setting.) Moreover, the analysis of multiproduct setting with capacity constraints introduce new subtleties that do not previously exist in the uncapacitated setting. A family of self-adjusting controls, i.e., *Linear Rate Correction* (LRC), has been shown to perform very well in the capacitated multiproduct setting when the demand function is *known* to the seller (Jasin (2014)). Motivated by this result, I will adapt LRC and develop a family of self-adjusting controls called *Parametric Self-adjusting Control* (PSC) that can be employed in the unknown demand setting. I will show that PSC attains the best achievable revenue loss bound for the joint learning and pricing problem. I explain PSC below.

Parametric Self-adjusting Control. The idea behind PSC is to divide the selling season into two stages: the *exploration* stage, where I do price experimentations using the exploration prices, and the *exploitation* stage, where I apply LRC using the parameter estimate computed at the end of the exploration stage. The exploration stage lasts for L periods (L itself is a decision variable to be optimized) while the exploitation stage lasts for $T-L$ periods. Let $Q \in \mathbb{R}^{n\times n}$ be a real matrix satisfying $AQ = A$ and let $\hat{\theta}_L$ denote the ML estimate of θ^* computed at the end of the exploration stage. For all $t \geq L+1$, define $\hat{\Delta}_t := D_t - \lambda(p_t; \hat{\theta}_L)$. Let C_t denote the remaining capacity at the *end* of period t. The complete PSC procedure is given below.

Parametric Self-adjusting Control (PSC)

Tuning Parameter: L

Stage 1 (Exploration)

a. Set exploration prices $\{\tilde{p}^{(1)}, \tilde{p}^{(2)}, ..., \tilde{p}^{(\tilde{q})}\}$. (See below.)

b. For $t = 1$ to L, do:

- If $C_{t-1} \succ 0$, apply price $p_t = \tilde{p}^{(\lfloor (t-1)\tilde{q}/L \rfloor + 1)}$ in period t.
- Otherwise, for product $j = 1$ to n, do:
  - If product j requires any resource that has been depleted, set $p_{t,j} = p_j^{\infty}$.
  - Otherwise, set $p_{t,j} = p_{t-1,j}$.

Stage 2 (Exploitation)

a. Compute the ML estimate $\hat{\theta}_L$ given $p_{1:L}$ and $D_{1:L}$.

b. Solve the deterministic optimization $\mathrm{P}_\lambda(\hat{\theta}_L)$.

c. For $t = L+1$ to T, do:

- If $C_{t-1} \succ 0$, apply the following price in period t

$$p_t = p\left(\lambda^D(\hat{\theta}_L) - \sum_{s=L+1}^{t-1} \frac{Q\hat{\Delta}_s}{T-s}\,; \hat{\theta}_L\right).$$

- Otherwise, for product $j = 1$ to n, do:
  - If product j requires any resource that has been depleted, set $p_{t,j} = p_j^{\infty}$.
  - Otherwise, set $p_{t,j} = p_{t-1,j}$.

Please note that in the PSC the exploration prices that satisfy conditions S1-S3 are set as described in Remark 3.3.1 and, as I will show below, an optimal tuning parameter for L is to set $L = \lceil\sqrt{kT}\rceil$. In comparison to the original LRC, which uses $p_t = p(\lambda^D(\theta^*) - \sum_{s=1}^{t-1} \frac{Q\Delta_s}{T-s}; \theta^*)^2$, since the underlying parameter vector θ^* is not known and the sequence $\{\Delta_s\}$ is not observable, I use $\hat{\theta}_L$ and $\{\hat{\Delta}_s\}$ as their substitute in PSC. Intuitively, one would expect that if $\hat{\theta}_L$ is sufficiently close to θ^*, then PSC should retain the strong performance of LRC. This intuition, however, is *not* immediately obvious. It should be noted that while LRC only deals with the impact of *natural* randomness due to demand fluctuations, as captured in $\{\Delta_s\}$, PSC also introduces a sequence of *systematic* biases due to estimation error as captured in $\{\hat{\Delta}_s\}$ (by definition, $\mathbf{E}^\pi[\hat{\Delta}_s] \neq 0$). Thus, despite the strong performance of LRC, it is not a priori clear whether linear rate adjustments alone, without re-optimizations *and* re-estimations, is sufficient to reduce the impact of estimation error on revenue loss. Interestingly, the answer is yes. In fact, PSC is rate-optimal.

Theorem 3.4.1 (Rate-Optimality of PSC) *Suppose that I use $L = \lceil\sqrt{kT}\,\rceil$. Then, there exists a constant $M_1 > 0$ independent of $k \geq 1$ such that $\rho^{PSC}(k) \leq M_1\sqrt{k}$ for all $k \geq 1$.*

As a comparison, if I apply the same static price $p_t = p^D(\hat{\theta}_L)$ throughout the exploitation stage, subject to capacity constraints, then the optimal length of exploration stage is of the order $k^{2/3}$ and the resulting revenue loss is $\mathcal{O}(k^{2/3}\log^{0.5} k)$ (Besbes and Zeevi 2009). This underscores an important point that a simple and autonomous price update is sufficient to reduce the revenue loss from $\mathcal{O}(k^{2/3}\log^{0.5} k)$ to $\mathcal{O}(k^{1/2})$. Let $E(t) := ||\theta^* - \hat{\theta}_t||_2$ and define $\epsilon(t) := \mathbf{E}^\pi[E(t)^2]^{1/2}$. The proof of Theorem 3.4.1 depends crucially on the following lemmas.

Lemma 3.4.1 (Continuity of the Optimal Solutions) *There exist constants $\kappa > 0$ and $\bar{\delta} > 0$ independent of $k > 0$, such that for all $\theta \in \mathsf{Ball}(\theta^*, \bar{\delta})$,*

a. $p^D(\theta) \in \mathsf{Ball}(p^D(\theta^*), \phi/2)$, $\mathsf{Ball}(p^D(\theta), \phi/2) \subseteq \mathcal{P}$ and $||\lambda^D(\theta^*) - \lambda^D(\theta)||_2 \leq \kappa||\theta^* - \theta||_2$,

b. $\mu^D(.) : \Theta \to \mathbb{R}^m_+$ is continuous at θ^*;

c. The capacity constraints of $P_\lambda(\theta)$ that correspond to the rows $\{i : \mu_i^D(\theta^*) > 0\}$ are binding.

Lemma 3.4.2 (Bounds for ML Estimator with I.I.D Observations) *There exist positive constants η_1, η_2, η_3 independent of $k > 0$, such that for all $\delta > 0$, I have $\mathbb{P}^\pi(E(L) > \delta) \leq \eta_1 \exp(-\eta_2 L\delta^2)$ and $\epsilon(L) \leq \eta_3/\sqrt{L}$.*

[2] Jasin (2014) uses $Q = HA$ for some H satisfying $AH = I$.

Lemma 3.4.3 (Exploitation revenue under PSC) *Let $\bar{\delta}$ be as defined in Lemma 3.4.1. Let $\hat{R}^{PSC}(k)$ denote the revenue under PSC during the exploitation stage. There exists a constant $M_0 > 0$ independent of $L > 0$ and $k \geq 3$ such that for all $k \geq 3$,*

$$\sum_{t=L+1}^{kT} r(\lambda^D(\theta^*);\theta^*) - \mathbf{E}^\pi\left[\hat{R}^{PSC}(k)\right] \leq M_0\left[\epsilon(L)^2 k + \frac{\log k}{1-\mathbb{P}^\pi(E(L)>\bar{\delta})} + L + \frac{1+k\,\mathbb{P}^\pi\left(E(L)>\bar{\delta}\right)}{1-\mathbb{P}^\pi(E(L)>\bar{\delta})}\right].$$

Some comments are in order. Lemma 3.4.1 means that the deterministic problem $\mathrm{P}(\hat{\theta}_L)$ is similar to the deterministic problem $\mathrm{P}(\theta^*)$ as long as the estimate $\hat{\theta}_L$ is sufficiently close to θ^*. In particular, the Lipschitz continuity of $\lambda^D(\theta)$ is useful to quantify the size of perturbation in the deterministic solution as a function of the estimation error. Lemma 3.4.2 is a typical statistical result that is needed to bound the size of the estimation error at the end of the exploration stage. Lemma 3.4.3 is the *key*. It characterizes the trade-off between exploration and exploitation by establishing the impact of the length of the exploration stage on the total revenue loss incurred during the exploitation stage; this, in turn, helps me to determine the optimal length of the exploration stage. I want to stress: The result of Lemma 3.4.3 is rather surprising. To see this, note that, if the true parameter vector is misestimated by a *small* error ϵ, then $\lambda^D(\hat{\theta}_L)$ is roughly ϵ away from $\lambda^D(\theta^*)$ as suggested by Lemma 3.4.1(a). If the seller simply uses the static price $p^D(\hat{\theta}_L)$ throughout the exploitation stage, then the one-period revenue loss is roughly $r(\lambda^D(\theta^*);\theta^*)-r(\lambda^D(\hat{\theta}_L);\theta^*) \approx \nabla_\lambda r(\lambda^D(\theta^*);\theta^*)\cdot(\lambda^D(\theta^*)-\lambda^D(\hat{\theta}_L)) \approx \Theta(\epsilon)$, which leads to a total revenue loss of $\mathcal{O}(\epsilon\, k)$. This is in contrast to the analysis in the *uncapacitated* setting where $\nabla_\lambda r(\lambda^D(\theta^*);\theta^*) = 0$ (because in this case $\lambda^D(\theta^*)$ is the global unconstrained optimizer of $r(\lambda;\theta^*)$), and thus a smaller revenue loss of order ϵ^2 is incurred in each period, which yields a total revenue loss of $\mathcal{O}(\epsilon^2 k)$ (see Broder and Rusmevichientong (2012)). This explains why the results in the uncapacitated setting are not directly applicable to the capacitated setting. In PSC, I use a feedback correction mechanism (i.e., the term $-\sum_{s=L+1}^{t-1}\frac{\hat{\Delta}_s}{T-s}$) that has the ability to mitigate the impact of systematic error ϵ on revenue loss. To further highlight the strength of self-adjusting price update, I report a numerical simulation in Table 3.1. Let *STA* denote the control that uses the deterministic price in the exploitation stage instead of adjusting prices using PSC's price update formula. (This control is the network RM version of the control in Besbes and Zeevi (2009).) Table 3.1 displays the revenue loss (RL) for *PSC* and *STA* and shows that *PSC* significantly outperforms *STA*. Finally, it should be noted that, although my analysis holds for all Q satisfying $AQ = A$, different choices of Q may lead to a different *non-asymptotic* performance. In particular, from the proof of Lemma 3.4.3, it can be seen that the constant M_0 is $\mathcal{O}(1+||Q||_2^2)$. Therefore, one approach to determine Q is to solve $\min\{||Q||_2 : s.t.\ AQ = A\}$. Note that this optimization is a convex program and A is known to the seller before the selling season; thus, the seller can solve the optimal Q off-line very efficiently.

3.5 Well-Separated Demand Function Family

The joint learning and pricing problem studied in Section 3.4 is very general: It allows both a general parametric demand form and an arbitrary number of unknown parameters. In this general case, the problem is naturally hard not only because active price experimentations are costly but also because, as it turns out, not all prices are equally informative. An example of the so-called *uninformative price* can be seen in Figure 3.1. Intuitively, if the seller experiments with an uninformative price, then he will not be able to statistically distinguish the true demand curve from the wrong one regardless of the choice of the estimation procedure. Indeed, as pointed out by Broder and Rusmevichientong (2012), this is the reason why one cannot improve on the $\Omega(\sqrt{k})$ lower bound for revenue loss in general. To guarantee a stronger performance bound than $\Theta(\sqrt{k})$, I need to impose additional assumptions on the demand model. One condition that has been studied in the literature is the so-called *well-separated* condition of the family of demand functions for a single product proposed by Broder and Rusmevichientong (2012) (see Figure 3.1). They show that, for the case of the uncapacitated single-product RM, if the demand function family is well-separated, the $\Omega(\sqrt{k})$ lower bound on revenue loss can be reduced to $\Omega(\log k)$. This is a significant improvement in terms of the potentially achievable performance of an admissible pricing control. It is not, however, a priori clear whether a similar result also holds in the more general network RM setting with *multiple products* and *capacity constraints*. In what follows, I first provide the definition of well-separated condition in multidimensional parameter space, and then I discuss a heuristic called *Accelerated Parametric Self-adjusting Control* (APSC), which is specifically designed to address this setting.

Well-separated demand. To formalize the definition of well-separated demand, it is convenient to first consider the distribution of a sequence of demand realizations $D_{1:t} = d_{1:t}$ under a sequence of prices $p^{\pi}_{1:t} \in \mathcal{P}^t$ generated by an admissible control π, which is defined as

$$\mathbb{P}^{\pi,\theta}_t(d_{1:t}) \;=\; \mathbb{P}^{p^{\pi}_{1:t},\theta}_t(d_{1:t}) \;=\; \prod_{s=1}^{t}\left[\left(1-\sum_{j=1}^{n}\lambda_j(p^{\pi}_s;\theta)\right)^{(1-\sum_{j=1}^{n} d_{s,j})}\prod_{j=1}^{n}\lambda_j(p^{\pi}_s;\theta)^{d_{s,j}}\right].$$

Define $\mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max}) := \{p \in \mathcal{P} : \sum_{j=1}^{n} \lambda_j(p;\theta) \le \tilde{\lambda}_{\max}, \lambda_j(p;\theta) \ge \tilde{\lambda}_{\min}, j = 1, \dots, n,$ for all $\theta \in$

$\Theta\}$, for some $0 < \tilde{\lambda}_{\min} < \tilde{\lambda}_{\max} < 1$. I state the well-separated assumptions below. All the results in this section require these additional assumptions to hold.

Well-Separated Assumptions. *For any* $0 < \tilde{\lambda}_{\min} < \tilde{\lambda}_{\max} < 1$*, there exists* $c_f > 0$ *such that:*

W1. *For all* $p \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$, $\mathbb{P}^{p,\theta}(.) \neq \mathbb{P}^{p,\theta'}(.)$ *whenever* $\theta \neq \theta'$;

W2. *For all* $\theta \in \Theta$, $p \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$, $I(p,\theta) \succeq c_f I$ *for* $I(p,\theta) := [I_{i,j}(p,\theta)] \in \mathbb{R}^{q\times q}$ *defined as*

$$[I(p,\theta)]_{i,j} = \mathbf{E}_\theta^p\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log\,\mathbb{P}^{p,\theta}(D)\right] = \mathbf{E}_\theta^p\left[-\frac{\partial}{\partial\theta_i}\log\,\mathbb{P}^{p,\theta}(D)\,\frac{\partial}{\partial\theta_j}\log\,\mathbb{P}^{p,\theta}(D)\right].$$

W3. *For any* $p_{1:t} = (p_1, \ldots, p_t) \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})^t$, $\log\,\mathbb{P}_t^{p_{1:t},\theta}(D_{1:t})$ *is concave in* θ *on* Θ.

Assumptions W1 and W2 are the multiproduct multiparameter analogs of the well-separated condition given in Broder and Rusmevichientong (2012). A necessary condition for W1 to hold is that there is no "redundancy". This means that the number of products must be at least as many as the number of the unknown parameters. If the number of products is strictly smaller than the number of unknown parameters (i.e. $n < q$), then I am essentially trying to solve a system of n equations with q unknowns, which may result in the non-uniqueness of θ. Note that W2 is analogous to condition S3 and it ensures that seller's information about the parameter vector strictly increases as he observes more demand realizations under any $p \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$. The last condition W3 requires the log-likelihood function to behave nicely. This is easily satisfied by many commonly used demand functions such as linear, logit, and exponential demand functions. Note that this well-separated condition is not overly restrictive as it permits, for example general demand functions with unknown additive market size (i.e., for each product j, its demand is $\lambda_j(p) = a_j + g_j(p)$ where the market size a_j is unknown and $g_j : \mathcal{P} \to [0,1]$ is a known function) and general demand functions with unknown multiplicative market size (i.e., for each product j, its demand is $\lambda_j(p) = a_j g_j(p)$ where the market size a_j is unknown and $g_j : \mathcal{P} \to [0,1]$ is a known function). For more examples of well-separated demand in the single-product/single-parameter setting, see Broder and Rusmevichientong (2012).

Passive learning with APSC. Estimating the unknown demand parameters from a family of well-separated candidate functions is considerably much easier than estimating the unknown parameters in the general setting. As discussed earlier, in the general parametric case, not all prices are equally informative. In contrast, under the well-separated condition, *all* prices are informative. This means that the demand data under *any* price will help improve the estimation, and the seller can continue to *passively* learn the demand parameter vector during the exploitation stage. The following result on ML estimation is the analog of Lemma 3.4.2 for non-i.i.d observations when the demand function family is well-separated.

Lemma 3.5.1 (Estimation Error of ML Estimator with Non-I.I.D Observations) *Fix some* $0 < \tilde{\lambda}_{\min} < \tilde{\lambda}_{\max} < 1$*. Suppose that an admissible control* π *satisfies* $p_s = \pi_s(D_{1:s-1}) \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ *for all* $1 \leq s \leq t$*. Then, under W1-W3, there exist constants* $\eta_4, \eta_5, \eta_6 > 0$*, such that* $\forall \delta > 0, \mathbb{P}^\pi(E(t) > \delta) \leq \eta_4 t^{q-1} \exp(-\eta_5 t \delta^2)$ *and* $\epsilon(t) \leq \eta_6 \sqrt{[(q-1)\log t + 1]/t}$*.*

Remark 3.5.1 *The result derived in Broder and Rusmevichientong (2012) (Theorem 4.7) can be viewed as a special case of ours. In particular, their result holds for the single product and single parameter setting whereas my result holds for a multidimensional setting with multiple products and multiple parameters. Although Hellinger distance and likelihood ratio are the common arguments used in deriving bounds in both results, I want to point out that the multidimensional parameter space is more complicated. To be precise, in the single dimension case, all candidate parameters lie on a line. Therefore, if ML estimator* $\hat{\theta}_t$ *is* δ *away from* θ^**, then there are only two possibilities: Either* $\hat{\theta}_t > \theta^* + \delta$ *or* $\hat{\theta}_t < \theta^* - \delta$*. Thus deriving the tail bound reduces to bounding the probability that, given the observations, the likelihood of* θ^* *is smaller than either of the* two *points:* $\theta^* - \delta$ *and* $\theta^* + \delta$*. In contrast, in the multidimensional parameter case, if ML estimation error is larger than* δ*, one needs to bound the probability that the likelihood of* θ^* *is smaller than any of an* infinite *number of points that lie on the boundary of a multidimensional ball. This makes my extension nontrivial. Another observation is that as the dimension of the parameter space increases, the bounds deteriorate. This results in the different orders of regret bounds for the single parameter and the multiple parameters cases. However, since the bounds do not deteriorate too much, I am still able to attain a sharp performance bound for APSC when multiple parameters need to be estimated.*

Accelerated Parametric Self-adjusting Control (APSC) divides the selling season into two stages similar to PSC: the initial exploration stage, which lasts L periods, and the exploitation stage, which lasts $T - L$ periods. However, unlike PSC, which stops learning the value of the underlying parameter vector once it exits the exploration stage, APSC continues to incorporate passive learning during its exploitation stage. To do this, APSC further divides the exploitation stage into small segments with increasing length (see Figure 3.2). Let

t_z, $z = 1, \ldots, Z+1$, be a sequence of *strictly* increasing integers satisfying $t_1 = L, t_2 = L+1$, $t_{Z+1} = T$, $t_z = \left\lceil \frac{t_{z+1}-L}{2} \right\rceil + L$ for all $z = 2, \ldots, Z$, and let segment z contains all the periods in $(t_z, t_{z+1}] := \{t_z + 1, t_z + 2, ..., t_{z+1}\}$. (Note that when T and L are given, the sequence of integers is *uniquely* determined. It is not difficult to see that Z, the number of segments obtained under the procedure mentioned above, satisfies $Z \leq \lceil \log_2(T-L+1) \rceil \leq \lceil \log_2 T \rceil$.) The idea is to re-estimate the parameter vector at the beginning of each segment and use the new estimate to update the deterministic solution over time. The re-estimation periods are spaced in a way that updates occur more frequently during the early part of the selling season, when my estimate is still highly inaccurate, and gradually phase out as the estimation accuracy improves. Once the parameter estimate is updated, ideally, the seller can update his deterministic solution by re-optimization. However, recall that frequent re-optimizations may still be computationally challenging for large-scale RM applications. To address this concern, I propose a re-optimization-free subroutine to update the deterministic solution at re-estimation points: (1) At the beginning of segment 1 (i.e., the beginning of period $L+1$), solve the deterministic optimization problem $\mathrm{P}(\hat{\theta}_1)$ to obtain the exact deterministic solution $\lambda^D(\hat{\theta}_1)$; (2) At the beginning of segment $z \geq 2$ (i.e., the beginning of period $t_z + 1$), use Newton's method (see more details below) to obtain an approximate solution of $\mathrm{P}(\hat{\theta}_z)$. Since this procedure involves some subtleties, I discuss this subroutine below before laying out the full description of APSC.

To better explain the intuition behind the subroutine, I first briefly review Newton's method for the multi-variate equality constrained problem. Let $\mathcal{X}$ be a convex set in $\mathbb{R}^n$, f be a strongly concave function, and F and G be a matrix and a vector, respectively, with a proper dimension. I write down a nonlinear programming (NP) problem with equality constraints and its Karush-Kuhn-Tucker (KKT) conditions below:

$$\text{(NP)} \quad \max_{x \in \mathcal{X}} \{f(x) : \;\; Fx = G\}, \qquad \text{(KKT)} \quad \{\nabla_x f(x^*) = F'\mu^*, Fx^* = G\},$$

where $(x^*; \mu^*)$ is the optimal pair of primal and dual solution. Since KKT conditions are both necessary and sufficient for the prescribed setting, to solve NP, I only need to solve the system of equations characterized by the KKT to which I will apply iterative Newton's method. To be precise, suppose that I have an approximate pair of primal and dual solution $(x_z; \mu_z)$. Then, my next pair of solution is given by $(x_{z+1}; \mu_{z+1}) = (x_z; \mu_z) + (\Delta_x; \Delta_\mu)$, where the *Newton steps* Δ_x and Δ_μ are characterized by the following:

$$\begin{array}{rcl} \nabla f(x_z + \Delta_x) & = & F'(\mu_z + \Delta_\mu) \\ F(x_z + \Delta_x) & = & G \end{array} \quad \approx \quad \begin{array}{rcl} \nabla f(x_z) + \nabla^2 f(x_z)\Delta_x & = & F'\mu_z + F'\Delta_\mu \\ Fx_z + F\Delta_x & = & G \end{array}$$

$$\Leftrightarrow \quad \begin{bmatrix} \Delta_x \\ \Delta_\mu \end{bmatrix} = \begin{bmatrix} -\nabla^2 f(x_z) & F' \\ F & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla f(x_z) - F'\mu_z \\ G - Fx_z \end{bmatrix}.$$

The key result for Newton's method is that it has a locally quadratic convergence rate, i.e., there exists some positive constants γ and ξ such that if $||x_z - x^*||_2 \leq \gamma$, then $||x_{z+1} - x^*||_2 \leq \xi ||x_z - x^*||_2^2$ (see Boyd and Vandenberghe (2004) for details). My idea is to tap into this locally quadratic convergence of Newton's method, coupled with the convergence result of ML estimator in Lemma 3.5.1, to develop a procedure for obtaining a sequence of solutions $\{\lambda_z^{NT}\}_{z=1}^Z$ that closely approximates $\{\lambda^D(\hat{\theta}_{t_z})\}_{z=1}^Z$. To implement this, I need to approximate $\mathrm{P}_\lambda(\hat{\theta}_{t_z})$ with an equality constrained problem $\mathrm{ECP}(\hat{\theta}_{t_z})$ (to be defined shortly) so that Newton's iteration can be properly applied. Let $C_i - (A\lambda^D(\hat{\theta}_{t_1}))_i$ denote the amount of slack for the i^{th} capacity constraint in $\mathrm{P}_\lambda(\hat{\theta}_{t_1})$ and define $\mathcal{B} := \{i : C_i/T - (A\lambda^D(\hat{\theta}_{t_1}))_i \leq \eta\}$ to be the set of *potential* binding constraints in $\mathrm{P}_\lambda(\theta^*)$, where η is a threshold level to be chosen by the seller. (Since I do not know which constraints are actually binding in $\mathrm{P}_\lambda(\theta^*)$, I use $\mathcal{B}$ as my estimate. It can be shown that the constraints in $\mathcal{B}$ coincide with the binding constraints in $\mathrm{P}_\lambda(\theta^*)$ with a very high probability as $k \to \infty$ if η is properly chosen. I address how η should be chosen in Theorem 3.5.1 below.) Let B and C_B denote the submatrix of A and subvector of C with rows corresponding to the indices in $\mathcal{B}$ respectively. Similarly, let N and C_N denote the submatrix of A and subvector of C with rows corresponding to the indices *not* in $\mathcal{B}$ respectively. Define the *Equality Constrained Problem* (ECP) as follows:

$$\mathrm{ECP}(\theta) \qquad \max_{x \in \mathbb{R}^n} \left\{ r(x;\theta) : \quad Bx = \frac{C_B}{T} \right\}$$

Denote by $x^D(\theta)$ the optimal solution of $\mathrm{ECP}(\theta)$. Note that if $\mathcal{B}$ coincides with the set of binding constraints of $\mathrm{P}_\lambda(\theta^*)$ at the optimal solution $\lambda^D(\theta^*)$, then not only $x^D(\theta^*)$ coincides with $\lambda^D(\theta^*)$, but also a stability result similar to Lemma 3.4.1(a) holds: there exist positive constants $\tilde{\delta}, \tilde{\kappa}$ such that for all $||\theta - \theta^*||_2 \leq \tilde{\delta}$, $||x^D(\theta) - \lambda^D(\theta^*)||_2 = ||x^D(\theta) - x^D(\theta^*)||_2 \leq \tilde{\kappa}||\theta - \theta^*||_2$. This means that $\mathrm{ECP}(\theta)$ closely approximates $\mathrm{P}_\lambda(\theta^*)$ when θ is close to θ^*. I define the Newton iteration for $\mathrm{ECP}(\hat{\theta}_{t_z})$ in segment z as follows:

$$\mathsf{Newton}_z(x,\mu) \;:=\; \begin{bmatrix} x+\Delta_x \\ \mu+\Delta_\mu \end{bmatrix} \;=\; \begin{bmatrix} x \\ \mu \end{bmatrix} + \begin{bmatrix} -R^{-1} & B' \\ B & O \end{bmatrix}^{-1} \begin{bmatrix} G-B'\mu \\ C_B-Bx \end{bmatrix}$$

$$= \begin{bmatrix} x \\ \mu \end{bmatrix} + \begin{bmatrix} -R+RB'S^{-1}BR & RB'S^{-1} \\ S^{-1}BR & S^{-1} \end{bmatrix} \begin{bmatrix} G-B'\mu \\ C_B-Bx \end{bmatrix}$$

where $R = [\nabla^2_{\lambda\lambda} r(x;\hat{\theta}_{t_z})]^{-1}$, $G = \nabla_\lambda r(x;\hat{\theta}_{t_z})$, and $S = BRB'$. (This formula is derived using the formula for Newton step in multi-variate equality constrained problem and the block matrix inversion formula.) Let $\mathcal{S}_z := \Lambda_{\hat{\theta}_{t_z}} \cap \{\lambda \in \mathbb{R}^n : N\lambda \le C_N, B\lambda = C_B\}$ for $z = 1,\ldots,Z$. I can now state the *Deterministic Price Update Procedure* (DPUP) below which will be a part of the APSC described later.

Deterministic Price Update Procedure

Tuning Parameter: η

For $z=1$, do:

a. Solve $\mathrm{P}_\lambda(\hat{\theta}_{t_1})$ and obtain $\lambda^D(\hat{\theta}_{t_1})$
b. Identify $\mathcal{B} := \{i : C_i/T - (A\lambda^D(\hat{\theta}_{t_1}))_i \le \eta\}$
c. Set $x_1^{NT} := \lambda^D(\hat{\theta}_{t_1})$, $\mu_1^{NT} = (BB')^{-1}B\,\nabla_\lambda r(x_1^{NT};\hat{\theta}_{t_1})$, and let $\lambda_1^{NT} := x_1^{NT}$.

For $z \ge 2$, do:

a. Set $(x_z^{NT};\mu_z^{NT}) := \mathsf{Newton}_z(x_{z-1}^{NT}, \mu_{z-1}^{NT})$
b. Let λ_z^{NT} be the projection of x_z^{NT} on $\mathcal{S}_z$, i.e., $\lambda_z^{NT} := \arg\min_{\lambda\in\mathcal{S}_z} ||x_z^{NT}-\lambda||_2$

I briefly explain the intuition behind DPUP. Recall that my goal is to obtain an approximate solution for each $\mathrm{P}_\lambda(\hat{\theta}_{t_z})$, $z=1,\ldots,Z$, without re-optimization. Since $\mathrm{ECP}(\hat{\theta}_{t_z})$ and $\mathrm{P}_\lambda(\hat{\theta}_{t_z})$ are similar, the projection of $x^D(\hat{\theta}_{t_z})$ on $\mathcal{S}_z$ should be a very good approximation of $\lambda^D(\hat{\theta}_{t_z})$. Therefore, if I can find a good approximation of $x^D(\hat{\theta}_{t_z})$, say x_z, then by projecting x_z on $\mathcal{S}_z$, I can attain a good feasible approximation of $\lambda^D(\hat{\theta}_{t_z})$. This is where I need to apply Newton's method to approximately solve each $\mathrm{ECP}(\hat{\theta}_{t_z})$. In particular, segment 1 carries out two objectives: (1) I want to find the set of potential binding constraints $\mathcal{B}$ and (2) I need to compute an initial pair of approximate primal and dual solution $(x_1^{NT};\mu_1^{NT})$ to $\mathrm{ECP}(\hat{\theta}_{t_1})$. I use $\lambda^D(\hat{\theta}_{t_1})$ as my initial primal solution x_1^{NT}. The approximate dual solution μ_1^{NT} is computed using the formula proposed in Boyd and Vandenberghe (2004). (Naturally, since $\nabla_\lambda r(x^D(\hat{\theta}_{t_1});\hat{\theta}_{t_1}) = B'\mu^D(\hat{\theta}_{t_1})$ must hold at the optimal primal and dual solution of $\mathrm{ECP}(\hat{\theta}_{t_1})$, this suggests that I use $\mu_1^{NT} = (BB')^{-1}B\nabla_\lambda r(x_1^{NT};\hat{\theta}_{t_1})$.) For any later segment $z > 1$, I first use $(x_{z-1}^{NT};\mu_{z-1}^{NT})$ as an initial feasible point for $\mathrm{ECP}(\hat{\theta}_{t_z})$ and apply a single iteration of Newton update to obtain a *much better* (due to the locally quadratic convergence of Newton's method) approximate solution $(x_z^{NT};\mu_z^{NT})$ of $\mathrm{ECP}(\hat{\theta}_{t_z})$. Then, I project

x_z^{NT} to $\mathcal{S}_z$ to obtain a feasible solution, λ_z^{NT}, to $\mathrm{P}_\lambda(\hat{\theta}_{t_z})$. By doing this, I manage to replace the full-scale re-optimization of $\mathrm{P}_\lambda(\hat{\theta}_{t_z})$ into one Newton update and one projection. It should be noted that, although it is theoretically possible to apply two (or more) iterations of Newton update, it is asymptotically unnecessary due to the locally quadratic convergence of Newton's method. Indeed, I show that $||x_z^{NT} - \lambda^D(\theta^*)||_2 = \Theta(||\hat{\theta}_{t_z} - \theta^*||_2)$. Thus, in light of Lemma 3.4.1(a), x_z^{NT} approximates $\lambda^D(\theta^*)$ as well as $\lambda^D(\hat{\theta}_{t_z})$ in terms of the order of approximation error. (See Figure 3.3 for an illustration of DPUP.) Below, I provide the full description of APSC heuristic.

Accelerated Parametric Self-adjusting Control (APSC)

Tuning Parameters: L, η

Stage 1 (Exploration)

a. Set exploration prices $\{\tilde{p}^{(1)}, \tilde{p}^{(2)}, ..., \tilde{p}^{(\tilde{q})}\}$. (See below.)

b. For $t = 1$ to L, do:

- If $C_{t-1} \succ 0$, apply price $p_t = \tilde{p}^{(\lfloor (t-1)\tilde{q}/L \rfloor + 1)}$ in period t,
- Otherwise, for product $j = 1$ to n, do:
  - If product j requires any resource that has been depleted, set $p_{t,j} = p_j^\infty$.
  - Otherwise, set $p_{t,j} = p_{t-1,j}$.

Stage 2 (Exploitation)

For time segment $z = 1$ to Z, do:

a. At the beginning of period $t_z + 1$, compute ML estimate $\hat{\theta}_{t_z}$

b. Use DPUP(η) to obtain λ_z^{NT}.

c. For $t = t_z + 1$ to t_{z+1}, do:

- If $C_{t-1} \succ 0$, apply the following price in period t

$$p_t := p\left(\lambda_z^{NT} - \sum_{s=t_1+1}^{t-1} \frac{Q\hat{\Delta}_s}{T-s}; \hat{\theta}_{t_z}\right),$$

- Otherwise, for product $j = 1$ to n, do:
  - If product j requires any resource that has been depleted, set $p_{t,j} = p_j^\infty$.
  - Otherwise, set $p_{t,j} = p_{t-1,j}$.

Please note that in APSC the exploration prices that satisfy conditions S1-S3 are set as described in Remark 3.3.1. Moreover, under the choice of L, η described in the following theorem, APSC has a strong revenue performance as stated in the theorem below.

Theorem 3.5.1 *Fix any $\epsilon > 0$. Suppose that I use $L = \lceil \log^{1+\epsilon}(kT) \rceil$ and $\eta = \log^{-\epsilon/4} k$. There exists a constant $M_2 > 0$ independent of $k \geq 3$ such that $\rho^{APSC}(k) \leq M_2 [\log^{1+\epsilon} k + (q-1)\log^2 k]$ for all $k \geq 3$.*

Remark 3.5.2 *Broder and Rusmevichientong (2012) has established that, under the well-separated case with one unknown parameter, the best achievable lower bound on the performance of any admissible pricing control in the* uncapacitated *single product case is* $\Omega(\log k)$ *and this bound is achievable by a heuristic called MLE-GREEDY. An open research question is whether this bound is also achievable in the more general case of capacitated network RM with well-separated demand. My result gives a partial answer. I show that the revenue loss of APSC is worse than* $\mathcal{O}(\log k)$ *by a factor of* $\log k$*. However, in the case where there is only one parameter to estimate, the revenue loss of APSC is* $\mathcal{O}(\log^{1+\epsilon} k)$*. Since* ϵ *can be chosen to be arbitrarily small, APSC almost attains the best achievable performance bound for the special case with a single unknown parameter.*

3.6 Closing remarks

I study the multi-product/multi-resource dynamic pricing problem in the presence of inventory constraint and unknown general parametric demand functions. I develop the PSC heuristic which first learns the demand function parameter by price experimentation and then adjusts the price over time according to the realized demand. The heuristic is computationally appealing since it only requires one estimation and one optimization. Surprisingly, the heuristic has a very strong performance as its revenue loss rate is only $\mathcal{O}(\sqrt{k})$, matching the best achievable revenue loss rate. This is the first heuristic for the parametric case that attains the exact revenue loss lower bound for capacitated network RM problems.

I also study the case when the family of the demand functions satisfies a so-called "well-separated" condition. Under this condition, demand parameter estimation becomes much easier and the seller can now spend more time on exploitation. Indeed, I develop the APSC heuristic, a modification of PSC, that reduces the revenue loss to $\mathcal{O}(\log^2 k)$. APSC conducts re-estimations according to a doubling schedule. Under APSC, the seller only needs to conduct one optimization and $\mathcal{O}(\log_2 k)$ estimations.

My results also suggest the wide applicability of self-adjusting idea in dynamic pricing problems. This self-adjusting idea can be used as a guideline for the companies to develop more sophisticated dynamic pricing policies in practice. Another surprising take-away it that even if the demand function is unknown, re-optimization is not indispensable to achieving a good performance.

3.7 Tables

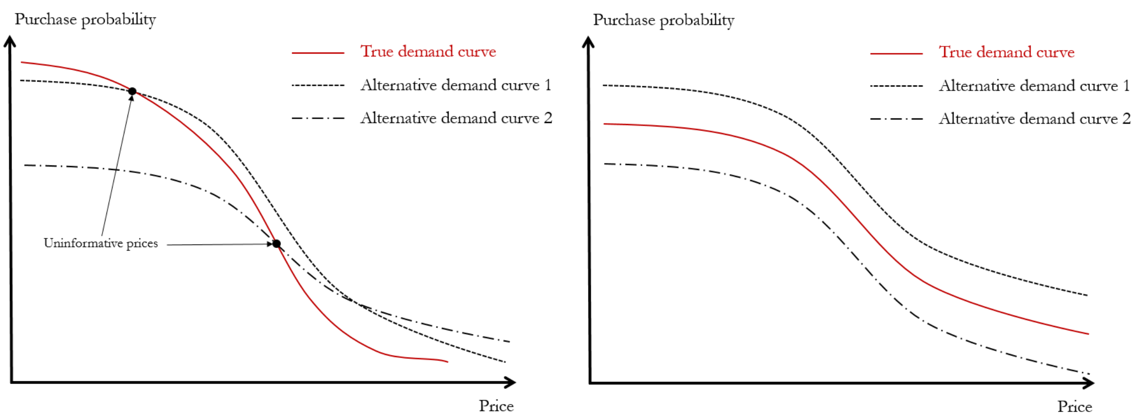
Table 3.1: Performance comparison of STA and PSC

	Revenue	STA		PSC	
k	upper bd.	RL(Std.)	% of RL	RL(Std.)	% of RL
100	24970	9876 (48)	39.5%	7711 (82)	30.9%
300	74911	20133 (169)	26.9%	14323 (205)	19.1%
1000	249702	45817 (443)	18.3%	29587 (437)	11.8%
3000	749107	97342 (1080)	13.0%	55633 (896)	7.4%
10000	2497023	223564 (2855)	9.0%	110542 (2012)	4.4%
30000	7491069	459024 (6274)	6.1%	205426 (4683)	2.7%
100000	24970230	1035790 (14572)	4.1%	371655 (9497)	1.5%
300000	74910689	2174142 (31567)	2.9%	702589 (21923)	0.9%

In this numerical example, I set $n = 2, m = 2,\ A = [1, 1; 0, 2], C = [1; 1]$. The demand model is a logit function, and $[\lambda_1(p_1, p_2); \lambda_2(p_1, p_2)] = (1 + \exp(4 - 0.015p_1) + \exp(8 - 0.02p_2))^{-1}\ [\exp(4 - 0.015p_1); \exp(8 - 0.02p_2)]$. For each heuristic, I vary the scale k from 100 to 300000 and run 1000 trials for each k.

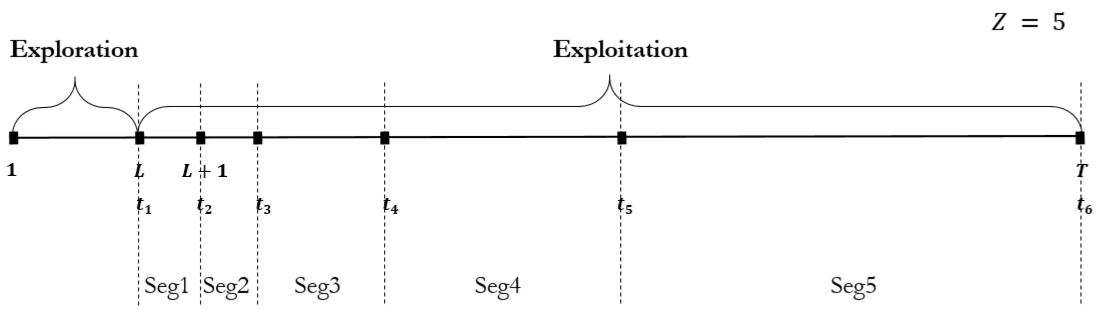
3.8 Figures

Figure 3.1: Uninformative prices (left) and well-separated demand family (right)



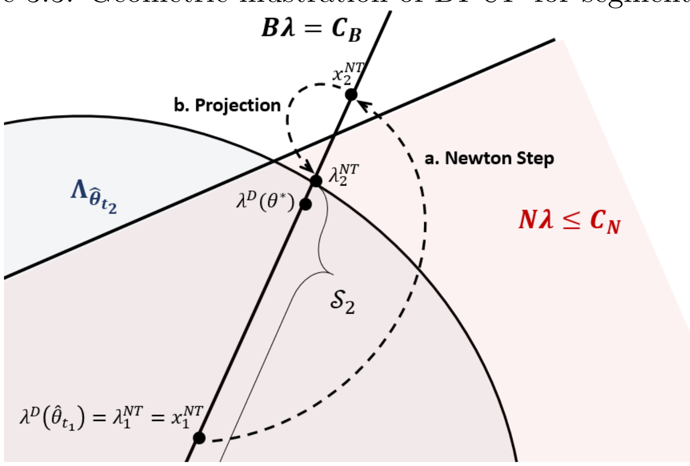
Note: For a general demand function family (left), there may be uninformative prices where the true demand curve and alternative demand curves intersect. The seller cannot statistically distinguish the true demand function from the alternatives when using those prices. This phenomenon does not occur in well-separated demand function family (right).

Figure 3.2: Illustration of APSC



Note: In this example, the first L periods are dedicated to exploration and the remaining periods are divided into five exploitation segments. The seller estimates the demand parameters and optimizes for the deterministic solution at the beginning of period $t_1 + 1$. The demand parameters are then re-estimated and the deterministic solution is updated accordingly at the beginning of periods $t_2 + 1, t_3 + 1, t_4 + 1, t_5 + 1$.

Figure 3.3: Geometric illustration of DPUP for segment $z = 2$



Note: In segment 2, step (a) is to apply Newton's method to the previous approximate solution x_1^{NT} to obtain a better solution to $\mathrm{ECP}(\hat{\theta}_{t_2})$, i.e., x_2^{NT}. This solution may not be feasible to $\mathrm{P}_\lambda(\hat{\theta}_{t_2})$, so in step (b), x_2^{NT} is projected on $\mathcal{S}_2$, which is a ray in this example, to obtain λ_2^{NT}.

Chapter 4 Pricing with Unknown Demand: Nonparametric Case

4.1 Abstract

I study a multi-period network revenue management problem where a seller sells multiple products, made from multiple resources with finite capacity, in an environment where the underlying demand function is a priori unknown (in the nonparametric sense). The objective of the seller is to simultaneously learn the unknown demand function and dynamically price his products to minimize the expected revenue loss. For the problem where the number of selling periods and initial capacity are scaled by $k > 0$, it is known that the expected revenue loss of any non-anticipating pricing policy is $\Omega(\sqrt{k})$. However, there is a considerable gap between this theoretical lower bound and the performance bound of the best known heuristic control in the literature. In this essay, I propose a *Nonparametric Self-adjusting Control* and show that its expected revenue loss is $\mathcal{O}(k^{1/2+\epsilon} \log k)$ for any arbitrarily small $\epsilon > 0$, provided that the underlying demand function is sufficiently smooth. This is the tightest bound of its kind for the problem setting that I consider in this essay and it significantly improves the performance bound of existing heuristic controls in the literature. In addition, my intermediate results on the large deviation bounds for spline estimation and nonparametric stability analysis of constrained optimization are of independent interest and are potentially useful for other applications.

4.2 Introduction

Revenue management (RM), which was first implemented in the 1960s by legacy airline companies to maintain their edge in the competitive airline market, has recently become widespread in many industries such as hospitality, fashion goods, and car rentals (for more detail, see Talluri and van Ryzin (2005)). The sellers in these industries face the common challenge of allocating a *fixed* capacity of perishable resources (e.g., seats in a jet, rooms in a hotel, etc.) to satisfy *volatile* demand of products or services. If the seller fails to satisfy demand appropriately, a considerable amount of profit is at stake either due to the zero salvage value of unused capacity or the loss of potential revenue. (For example, in the airline industry, it is known that the benefit of using RM is roughly comparable to the airline's annual margin, which is about 4-5% of total revenue (Talluri and van Ryzin 2005).) Given this, RM is aimed at helping the sellers to make optimal decisions such that the right product is sold to the right customer at the right time and at the right price. One type of operational leverage often employed by the sellers is *dynamic pricing*: By adjusting the prices over time, the seller can effectively control the rate at which demand arrives so he can better match demand with available resources.

Despite its known benefits (Talluri and van Ryzin 2005), the efficacy of dynamic pricing hinges upon the seller's knowledge of market's response to different prices, i.e., the underlying demand function. Unfortunately, in most (if not all) real-life applications, this underlying demand function is not easily accessible to the sellers. Although many sellers have adopted sophisticated statistical methods, the estimated demand function is inevitably subject to estimation error, which in turn affects the quality of the sellers' pricing decisions. The negative impact of inaccurate demand estimation is further magnified in practice because typical RM industries tend to have a large sales volume; thus, even small errors can lead to a significant revenue loss in absolute term. Given this limitation, one pressing issue faced by RM practitioners is *how to dynamically price their products when the underlying demand function is unknown a priori*. This essay studies joint learning and pricing problem in a general network RM setting with *multiple* products and *multiple* capacitated resources for the *nonparametric* demand case. (By *nonparametric*, I mean the case where the seller does not even know the functional form of demand. This is in contrast to the so-called *parametric* case where the seller a priori knows the form of demand function (e.g., linear, exponential, logit, etc.) and he only needs to estimate the unknown parameters (e.g., the intercept and the slope of a linear demand function)). In this essay, I construct a heuristic control that is not only easy to implement for large-scale problems but also has a provable analytical performance bound. My bound significantly improves the performance bound of existing heuristic controls in the literature.

Literature review. A large body of RM literature has investigated the canonical dynamic pricing problem where the seller knows the underlying demand function. The prevailing view is that, even in the case where learning is not in play, computing an optimal control is already challenging to do. This is so because the common technique for solving sequential decision problems, the so-called Dynamic Programming (DP), suffers from the well-known curse of dimensionality. This curse of dimensionality is exacerbated in many RM industries because the sellers typically have to manage the price of at least *thousands* of products on a daily basis. To illustrate, a typical major US airline operates more than a thousand flights daily, each of which has more than ten different booking classes that are characterized by different combinations of service level and purchase restriction. Since passengers book tickets in advance, the airline needs to price not only the tickets for the same-day flights but also those with departure dates several months in the future. All these factors put together can easily translate into a daily pricing decision for *millions* of itineraries. Due to this challenge, instead of finding the optimal pricing control, a considerable body of existing literature has focused on developing computationally implementable heuristic controls with provably good performance guarantee. (See Bitran and Caldentey (2003) and Elmaghraby and Keskinocak (2003) for a comprehensive review of the literature.)

Within the canonical RM literature, some works have focused on developing heuristic controls based on the solution of a deterministic pricing problem, i.e., the deterministic counterpart of the original stochastic control problem, which is computationally much easier to solve than the DP. This approach was first proposed by Gallego and van Ryzin (1994). They develop a static price control by first solving a convex optimization problem at the beginning of the selling season and then using its optimal solution as static price throughout the selling season, subject to available resources. Although the proposed heuristic control is easy to implement, its drawback is obvious: It ignores the observed demand realizations, which leaves room for further improvement. One intuitively appealing idea that has been studied in the literature involves frequent *re-optimization* of the deterministic pricing problem throughout the selling season. Maglaras and Meissner (2006) show that the re-optimized static price control (RSC) cannot perform worse than static price control without re-optimization (in asymptotic sense). However, it is not immediately clear from their analysis alone whether re-optimization actually guarantees a better performance (and if so, by how much). A recent work by Jasin (2014) answers this question in the affirmative by showing that RSC *does* significantly improve the performance of static price control (again, in asymptotic sense). While existing literature has shown frequent re-optimization to be beneficial, its implementation can be very time-consuming especially when applied to large-scale problems that often arise in practice; this has motivated the development for computationally much easier yet equally effective heuristic controls. For example, motivated by the optimal structure of the diffusion control problem of the continuous-time dynamic pricing problem, Atar and Reiman (2012) develop a re-optimization-free *bridge* pricing control that guarantees the same asymptotic performance as RSC. An equally effective heuristic control is also obtained in Jasin (2014). Motivated by the structure of the re-optimized prices under RSC, Jasin (2014) proposes a real-time control, called *Linear Rate Correction* (LRC), that has a similar structure as the bridge control and does not require any re-optimization at all. To be precise, LRC only requires a single optimization at the beginning of the selling season and automatically adjusts the price according to a pre-specified update rule throughout the remaining selling season. Inspired by the strong performance of bridge pricing control and LRC in the setting with known demand function, in this essay, I construct a nonparametric self-adjusting control akin to LRC and show that its asymptotic performance is very close to the theoretical lower bound on the performance of any feasible pricing control in the setting with unknown demand function. Below, I discuss the literature on joint learning and pricing.

There is a growing literature that studies joint demand learning and pricing problem. Most existing works have combined a particular statistical learning procedure (e.g., Maximum Likelihood, Least Squares, etc.) with a certain dynamic pricing control (most notably, the static price control). A central highlight in this literature is the trade-off between the cost of learning the demand function (exploration) and the reward of using the "optimal" price computed using the estimated demand function (exploitation). The longer the time the seller spends on learning the demand function, the less opportunity there is to exploit the knowledge of the newly estimated demand function. On the flip side, if the exploration time is too short, it will result in a poor estimation, which yields highly sub-optimal prices. What is the best performance that any non-anticipating pricing control can achieve in the setting with unknown demand function? Suppose that I scale the length of the selling season and the initial resource capacity by a factor of $k > 0$. (The constant k can be interpreted as the size of the problem. See Section 4.3 for more discussions on this.) One way to measure the performance of a feasible control is to study the *order of expected revenue loss* which is defined as the order (with respect to k) of the gap between the total expected revenue earned under this control and a well-established deterministic upper bound. (See Section 4.3 for more details on this performance metric.) It is widely known in the literature that the expected revenue loss of any feasible pricing control in general is $\Omega(\sqrt{k})$ (e.g., Besbes and Zeevi (2009), Broder and Rusmevichientong (2012), Keskin and Zeevi (2014)). For the case of uncapacitated RM, where there is no limit on the number of resources that can be used, this lower bound has been repeatedly shown to be tight (e.g., Broder and Rusmevichientong (2012), Keskin and Zeevi (2014)). As for the case of capacitated RM, most existing literature has primarily focused on the setting of a single-product and single-resource RM (often called *single-leg RM* due to the early application of RM in airline industry). Besbes and Zeevi (2009) is among the first to investigate this problem under both parametric and nonparametric cases. Their proposed heuristic control for the parametric case yields an expected revenue loss of $\mathcal{O}(k^{2/3} \log^{0.5} k)$ whereas their proposed heuristic control for the nonparametric case guarantees an expected revenue loss of $\mathcal{O}(k^{3/4} \log^{0.5} k)$. This suggests that there is a considerable gap between the performance of parametric and nonparametric approaches.

Recent works by Wang et al. (2014) and Lei et al. (2014) have managed to significantly shrink this gap; they develop sophisticated nonparametric heuristic controls that guarantee a $\mathcal{O}(\sqrt{k}\log^{4.5} k)$ and $\mathcal{O}(\sqrt{k})$ expected revenue loss, respectively. Thus, for the case of capacitated RM in single-leg setting, existing works in the literature have managed to not only completely close the gap between the performance of parametric and nonparametric approaches, at least in the asymptotic sense, but also show that the theoretical lower bound of $\Omega(\sqrt{k})$ is indeed tight.

The general network RM problem with multiple products and multiple limited resources is more difficult to analyze than the single-leg RM. (In Section 4.5, I will explain why the proofs and the arguments in the uncapacitated setting cannot be applied to the capacitated setting. Moreover, the arguments in Wang et al. (2014) and Lei et al. (2014) for the single-leg capacitated RM also cannot be applied to the more general network RM setting. This is so because both Wang et al. (2014) and Lei et al. (2014) heavily exploit the special structure in the single-leg RM. Unfortunately, no analogous structure is known for the network RM.) To the best of my knowledge, the only paper that addresses the joint learning and pricing problem for general network RM is Besbes and Zeevi (2012). They consider the nonparametric case only and show that the performance bound of their proposed heuristic control (i.e., so-called Algorithm 2 in their paper) is $\mathcal{O}(k^{(n+2)/(n+3)}\log^{0.5} k)$, where n is the number of products. Note that the fraction $(n+2)/(n+3)$ in the bound highlights the curse of dimensionality for network RM since the performance bound quickly deteriorates as the number of products n increases. If, however, the true demand function is sufficiently smooth (e.g., infinitely differentiable), this bound can be reduced; they propose another nonparametric heuristic control (Algorithm 3) that guarantees a $\mathcal{O}(k^{2/3+\epsilon}\log^{0.5} k)$ expected revenue loss for some $\epsilon > 0$ that can be arbitrarily small. As one can see, there is still a considerable gap between the lower bound of $\Omega(\sqrt{k})$ and the performance bound of $\mathcal{O}(k^{2/3+\epsilon}\log^{0.5} k)$. My proposed heuristic control in this essay significantly reduces this gap from $k^{2/3+\epsilon}$ to $k^{1/2+\epsilon}$ (up to logarithmic terms).

I would like to note here that all the results discussed above for the joint learning and pricing problem are derived for the setting where the seller can use *a continuum of prices* drawn from a certain convex and compact set. This distinction is crucial as the complexity of the problem changes as I switch from a continuum setting to the setting where the set of feasible prices is finite. Besbes and Zeevi (2012) have also considered this finite set setting and proposed a heuristic control (Algorithm 1) with a performance bound of $\mathcal{O}(k^{2/3}\log^{0.5} k)$. A recent work by Ferreira et al. (2016) improves this bound to $\mathcal{O}(\sqrt{k}\log k \log\log k)$ by using a Thompson sampling-based heuristic control. Note that although the best known performance bound that I am aware of for the network RM with finite price set setting is already close to $\sqrt{k}$, it is not clear that the theoretical lower bound for this setting is still $\Omega(\sqrt{k})$. In fact, a recent work by Flajolet and Jaillet (2015) show that the UCB-Simplex control they propose attains a logarithmic performance bound in a simpler setting where there is only one capacity constraint (besides time). This seems to suggest that the lower bound of the general network RM with finite price set setting could be as small as logarithmic. More broadly, when the set of feasible prices is finite, the joint learning and pricing problem is closely related to the *bandit problems* extensively studied in the Reinforcement Learning literature (e.g., Lai and Robbins (1985), Auer et al. (2002a), etc.). This stream of literature had not considered the inter-temporal constraints on actions over time (such as the capacity constraint in the network RM setting) until only very recently (e.g., Badanidiyuru et al. (2013), Badanidiyuru et al. (2015), Combes et al. (2015), etc.). Badanidiyuru et al. (2013) is among the first to consider the so-called *bandit with knapsack* (BwK) problem. In BwK, a decision maker has a fixed amount of resources and needs to accumulate rewards by sequentially selecting from a finite set of arms whose reward distributions are unknown. Pulling each arm stochastically depletes those resources according to an unknown consumption distribution of that arm, and the decision maker stops collecting rewards when he runs out of resources. Note that the general RM with finite price set and unknown demand fits into the BwK framework by treating each feasible price vector as a bandit and viewing time as a resource with deterministic depletion rate. Badanidiyuru et al. (2015) show that the performance lower bound of BwK is $\Omega(\sqrt{k})$ and propose two heuristic controls that match this lower bound up to logarithmic factors. While the network RM with finite number of feasible prices is a special case of the BwK problem, I would like to point out that the bounds derived in Badanidiyuru et al. (2015) cannot be directly compared to the bounds derived in all the RM papers discussed above since the quantification order used in evaluating the asymptotic bounds are different. To be precise, Badanidiyuru et al. (2015) allow the underlying demand distribution to vary in k while existing RM works assume that the underlying demand distribution does not vary as k scales; hence, both the lower and upper bounds in Badanidiyuru et al. (2015) have weakly larger asymptotic order. (This phenomena is not unique to Badanidiyuru et al. (2015) alone. For example, the gap in performance bounds due to the use of different quantification orders also arises in the traditional bandit setting such as in the logarithmic bounds of Auer et al. (2002a) versus the square-root bounds of Auer et al. (2002b)—see page 50 in Auer et al. (2002b) for more discussions.) Finally, I want to point out that, in order to derive the order of performance bound discussed above, the heuristic controls developed for finite price setting are typically compared with a revenue upper bound benchmark under the finite price setting while the heuristic controls developed for the continuum price setting are compared with a larger revenue upper bound benchmark under the continuum setting. Since the two upper bound benchmarks are different, with the former being smaller than the latter, the heuristic controls developed under the two settings are not easily comparable by simply looking at the order of their performance bounds as a function of k. Moreover, one also cannot simply extend existing heuristic controls developed in finite price setting to continuum price setting since the performance bounds of these heuristic controls (e.g., Badanidiyuru et al. (2013), Badanidiyuru et al. (2015), Ferreira et al. (2016)) deteriorate quickly as the number of the feasible prices increases. This is so because these heuristic controls do not exploit existing relationship between expected demand value at different price points and need to learn the expected demand value at most price points separately, which is not very efficient if the number of price points is large.

Apart from the RM and joint learning and optimization literature, my work is also closely related to *Spline Regression* in the statistics literature. A typical problem in statistics is to estimate the mean response as a function of some input variables. (The demand learning aspect of my problem is one such problem: my goal is to estimate the mean demand of each product as a function of the prices of all products.) Spline Regression generates an estimate in the form of a linear combination of *spline basis functions* (originally studied in Applied Mathematics to approximate *deterministic* functions) and uses Least Squares criterion to compute the corresponding coefficients. (See Gyorfi et al. (2002) for more details on Spline Regression.) To the best of my knowledge, most existing literature on Spline Regression is mostly concerned with the estimation accuracy of the response function; thus, the typical convergence result for Spline Regression is limited to only the estimation error of the response function itself. In my problem, estimation and optimization are intertwined and it is crucial to understand how the estimation error affects the subsequent optimization. This requires results on bounds of the error between higher order partial derivatives of the response function and its spline estimate. To derive these bounds, deviating from the Spline Regression approach, I generate my spline estimate by using a specific linear operator (i.e., $\mathcal{L}$ defined in Step 3 of Technical Details for Spline Approximation part (b) in Section 4.4.1) instead of using Least Squares criterion. (I call my estimation method *Spline Estimation* to differentiate my approach from Spline Regression. For more details on Spline Estimation, see Section 4.4.1. I want to emphasize here that although I choose Spline Estimation in combination with my optimization, this is only for the purpose of mathematical analysis. In general, I suspect that the seller can also use other estimation schemes such as Spline Regression, local polynomial approximation, etc. in lieu of Spline Estimation in my proposed heuristic control and still enjoys a strong performance.) This linear operator, also known as a *quasi-interpolant* in Schumaker (2007), is originally devised to analyze the error of using spline functions to approximate a *deterministic* function. I generalize the analysis of spline approximation for deterministic functions in Schumaker (2007) to estimating response functions using spline functions. I manage to not only derive large deviation bounds of the estimation error of the demand function itself but also its partial derivatives, which are very useful in analyzing my proposed heuristic control.

Contributions and the organization of the essay. My contributions in this essay can be summarized in the following two points:

1. I develop a nonparametric control called *Nonparametric Self-adjusting Control* (NSC) that can be applied to the general network RM setting with multiple products and multiple limited resources. I show that if the underlying demand function is sufficiently smooth, the expected revenue loss of NSC is $\mathcal{O}(k^{1/2+\epsilon}\log k)$ for some $\epsilon > 0$ that can be arbitrarily small (see Theorem 4.5.1 in Section 4.5). This is the *tightest* bound of its kind for the setting that I am considering (i.e., the continuum price setting): It significantly improves the $\mathcal{O}(k^{2/3+\epsilon}\log^{0.5} k)$ bound of Besbes and Zeevi (2012) and is only slightly worse than the theoretical lower bound of $\Omega(\sqrt{k})$. In a nutshell, NSC is a combination of four elements: (1) Spline Estimation of the underlying demand function, (2) linear approximation of the estimated demand function, (3) quadratic approximation of the estimated revenue function, and (4) approximate self-adjusting control akin to the one developed in Jasin (2014). In this essay, I show that although each of these elements introduces its own error (and some of them are *not* even unbiased), under properly selected tuning parameters, the cumulative impact of these errors is asymptotically only slightly larger than $\Theta(\sqrt{k})$. This makes it possible to prove the strong performance guarantee of NSC. Note that, per my discussions above, the first element (i.e., using Spline Estimation to estimate the underlying demand function) is only for the purpose of the analysis; in practice, the seller can still use other estimation schemes in combination with the other three elements. Although the proper selection of the tuning parameters will depend on the specific estimation scheme being used, I suspect that the resulting heuristic control still enjoys a similar strong performance guarantee.

2. In addition to contributing to the RM literature, my intermediate results in this essay also contribute to the more general statistics and optimization literature. For the nonparametric estimation, I generalize the analysis of spline approximation for deterministic functions in Schumaker (2007) to the setting with noisy observations and derive large deviation bounds for the estimation error of the function itself and its higher order partial derivatives (see Lemma 4.4.1 in Section 4.4). These bounds seem

to be new—although spline functions have been used in statistics, I am not aware of existing large deviation bound for higher order partial derivatives of the estimate—and are particularly useful for my analysis because the resulting spline estimate is ultimately used in the subsequent optimization phase in my heuristic control. Moreover, the bound for partial derivatives also facilitates the stability analysis of the optimal solution. Aside from the statistical error bound, for the analysis of NSC, I also need to derive a nonparametric Lipschitz-type stability result for the optimal solution of a perturbed optimization problem (see Lemma 4.4.2 in Section 4.4). This result also seems to be new—although *parametric* stability analysis of optimization problem has been intensively studied in the literature (see Bonnans and Shapiro (2000)), nonparametric stability analysis is very rare. As of the writing of this essay, I am not aware of any existing result on nonparametric stability analysis that can be directly used for my purpose. Aside from its use in the analysis of NSC, my stability result is of independent interest and is potentially applicable to other optimization problems.

The reminder of this essay is organized as follows. I first formulate the problem in Section 4.3. My nonparametric approach is discussed in Sections 4.4 - 4.6. In particular, Section 4.4 provides some preliminaries on spline approximation and nonparametric stability analysis; Section 4.5 describes the proposed NSC and its performance bound (Theorem 4.5.1); Section 4.6 provides the proof of Theorem 4.5.1. Finally, I conclude the essay in Section 4.7. Unless otherwise noted, all the extra details of the proofs can be found in Appendix A.4.

4.3 Problem formulation

In this section I describe the problem setting, modeling assumptions, and the asymptotic regime.

Notation. The following notation will be used throughout this essay. (Additional notation will be introduced when necessary.) Denote by $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{R}_{++}$ the set of real, non-negative real, and positive real numbers respectively. For column vectors $a = (a_1; \dots; a_n) \in \mathbb{R}^n, b = (b_1; \dots; b_n) \in \mathbb{R}^n$, denote by $a \succeq b$ if $a_i \geq b_i$ for all i, and by $a \succ b$ if $a_i > b_i$ for all i. Similarly, denote by $\mathbb{Z}$, $\mathbb{Z}_+$, and $\mathbb{Z}_{++}$ the set of integers, non-negative integers, and positive integers respectively. For any $a, b \in \mathbb{Z}$ with $a \leq b$, let $\overline{[a,b]} := \{a, a+1, \dots, b-1, b\}$. I denote by $\cdot$ the inner product of two vectors and by $\otimes$ the tensor product of sets or function spaces. I use a prime to denote the transpose of a vector or a matrix, an I to denote an identity matrix with a proper dimension, and an $\mathbf{e}$ to denote a vector of ones with a proper dimension. For any vector $v = [v_j] \in \mathbb{R}^n$, $||v||_p := (\sum_{j=1}^n |v_j|^p)^{1/p}$ denote its p-norm ($1 \leq p \leq \infty$) and, for any real matrix $M = [M_{ij}] \in \mathbb{R}^{n \times n}$, $||M||_p := \sup_{||v||_p=1} ||Mv||_p$

denote its induced p-norm. For example, $||M||_1 = \max_{1\leq j\leq n}\sum_{i=1}^n |M_{ij}|$, $||M||_2 =$ the largest eigenvalue of $M'M$, and $||M||_\infty = \max_{1\leq i\leq n}\sum_{j=1}^n |M_{ij}|$. (Note that $||M||_1 = ||M'||_\infty$.) For any function $f : \mathcal{X} \to \mathcal{Y}$, I denote by $||f(.)||_\infty := \sup_{x\in\mathcal{X}} ||f(x)||_\infty$ the infinity-norm of f. I use ∇ to denote the usual derivative operator and a subscript to indicate the variables with respect to which this operation is being applied to. (No subscript ∇ means that the derivative is applied to all variables.) If $f : \mathbb{R}^n \to \mathbb{R}$, then $\nabla_x f = (\frac{\partial f}{\partial x_1}; \ldots; \frac{\partial f}{\partial x_n})$; if, on the other hand, $f = (f_1; \ldots; f_n) : \mathbb{R}^n \to \mathbb{R}^n$, then

$$\nabla_x f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}.$$

Denote by $\mathsf{C}^s(\mathcal{S})$ the set of functions whose first s^{th} order partial derivatives are continuous on its domain $\mathcal{S}$, and by $\mathsf{P}^s([a, b])$ the set of single variate polynomial functions with degree s on an interval $[a, b] \subseteq \mathbb{R}$, e.g., $\mathsf{P}^1([0, 1])$ is the set of all linear functions on the interval $[0, 1]$.

The model. I consider the setting of a monopolist selling his products to incoming customers during a finite selling season, aiming to maximize his total expected revenue. There are n types of products, each of which is made up of a combination of a subset of m types of resources. For example, in the airline setting, a product refers to a multi-flight itinerary and a resource refers to a seat in a jet of a single-leg flight; in the hotel setting, a product refers to a multi-day stay and a resource refers to a one-night stay at a particular room. I denote by $A = [A_{ij}] \in \mathbb{R}^{m\times n}$ the *resource consumption matrix*, which characterizes the types and amounts of resources needed by each product (i.e., a single unit of product j requires A_{ij} units of resource i). Without loss of generality, I assume that the matrix A has full row rank. (If this is not the case, then I can first apply the standard row elimination procedure to delete the redundant rows. See Jasin (2014).) Denote by $C \in \mathbb{R}^m$ the vector of initial capacity levels of all resources at the beginning of the selling season. Since, in many industries (e.g., hotels and airlines), replenishment of resources during the selling season is either too costly or simply not feasible, following the standard model in the literature (Gallego and van Ryzin 1997), I will assume that the seller has no opportunity to procure additional units of resources during the selling season. In addition, I also assume without loss of generality that the remaining resources at the end of the selling season have zero salvage value.

The selling season is divided into T discrete periods, indexed by $t = 1, 2, ..., T$. At the beginning of period t, the seller first decides the price $p_t = (p_{t,1}; \ldots; p_{t,n})$ for his products, where p_t is chosen from a convex and compact set $\mathcal{P} = \otimes_{l=1}^{n}[\underline{p}_l, \bar{p}_l] \subseteq \mathbb{R}^n$ of feasible price vectors. The posted price p_t, in turn, induces a demand, or sale, for one of the products with a certain probability. Here, I implicitly assume that at most one sale for one product occurs in each period. (I have made this assumption and chosen to focus on discrete time model to simplify the presentation of the analysis. My analysis can be easily extended to either a discrete-time model with bounded demand arrivals in each period or continuous-time model with Poisson arrivals.) Let $\mathbf{\Delta}^{n-1} := \{(x_1; \ldots; x_n) \in \mathbb{R}^n | \sum_{i=1}^{n} x_i \leq 1, \text{ and } x_i \geq 0 \text{ for all } i\}$ denote the standard $(n-1)$-simplex. Let $\lambda^*(.) : \mathcal{P} \to \mathbf{\Delta}^{n-1}$ denote the induced *demand rate* or *purchase probability* vector; I also call $\lambda^*(.)$ the underlying *demand function*. Contrary to most existing RM literature where it is assumed that the seller knows $\lambda^*(.)$ a priori, in this essay, I simply assume that this function can be estimated using statistical learning procedures. Let $\Lambda_{\lambda^*} := \{\lambda^*(p) : p \in \mathcal{P}\}$ denote the convex and compact set of feasible demand rates and let $D_t(p_t) = (D_{t,1}(p_t); \ldots; D_{t,n}(p_t))$ denote the vector of realized demand in period t under price p_t. It should be noted that, although demands for different products in the same period are not necessarily independent, demands over different periods are assumed to be independent (i.e., D_t only depends on the posted price p_t in period t). By definition, I have $D_t(p_t) \in \mathcal{D} := \{D : \sum_{j=1}^{n} D_j \leq 1, D_j \in \{0,1\} \text{ for all } j\}$ and $\mathbf{E}[D_t(p_t)] = \lambda^*(p_t)$. This allows me to write $D_t(p_t) = \lambda^*(p_t) + \Delta_t(p_t)$, where $\Delta_t(p_t)$ is a zero-mean random vector. For notational simplicity, whenever it is clear from the context which price p_t is being used, I will simply write $D_t(p_t)$ and $\Delta_t(p_t)$ as D_t and Δ_t respectively. The sequence $\{\Delta_t\}_{t=1}^{T}$ will play an important role in my analysis later. Define the revenue function $r^*(p) := p \cdot \lambda^*(p)$ to be the one-period expected revenue that the seller can earn under price p. It is typically assumed in the literature that $\lambda^*(.)$ is invertible (see the regularity assumptions below). I can then write $r^*(p) = p \cdot \lambda^*(p) = p^*(\lambda) \cdot \lambda = r^*_\lambda(\lambda)$ to emphasize the dependency of revenue on demand rate instead of on price. I make the following regularity assumptions on $\lambda^*(.)$, $r^*(.)$ and $r^*_\lambda(.)$:

Regularity Assumptions. *There exists positive constants $\bar{r}$, $\underline{v} < \bar{v}$ such that:*

R1. $\lambda^*(.) : \mathcal{P} \to \Lambda_{\lambda^*}$ *is in* $\mathsf{C}^2(\mathcal{P})$ *with Lipschitz continuous second order partial derivatives, and it has an inverse function* $p^*(.) : \Lambda_{\lambda^*} \to \mathcal{P}$ *that is in* $\mathsf{C}^2(\Lambda_{\lambda^*})$;

R2. *There exists a set of turn-off prices* $p_j^\infty \in [\underline{p}_j, \bar{p}_j]$ *for* $j = 1, \ldots, n$ *such that for any* $p = (p_1; \ldots; p_n)$, $p_j = p_j^\infty$ *implies that* $\lambda_j^*(p) = 0$.

R3. $||r^*_\lambda(.)||_\infty \leq \bar{r}$, $r^*_\lambda(.)$ *is strongly concave, and all the eigenvalues of* $\nabla^2 r^*_\lambda(\lambda)$ *are between* $-\bar{v}$ *and* $-\underline{v}$ *for all* $\lambda \in \Lambda_{\lambda^*}$.

Assumption R1 is fairly natural and is easily satisfied by many popular demand functions such as linear, logit, and exponential functions. Assumption R2 is common in the literature. (See Besbes and Zeevi (2009) and Wang et al. (2014) for similar assumptions.) Its purpose is to allow the seller to effectively shut down the demand for any product whenever needed, e.g., in the case of stock-out. (The existence of such turn-off prices follows naturally when one considers truncated demand functions. It is also possible to consider an unbounded set of feasible prices instead of the compact set I assume above, with a potentially infinite turn-off price; in such setting, my results still hold.) As for Assumption R3, the boundedness of $r^*_\lambda(.)$ follows from the compactness of Λ_{λ^*} and the continuity of $r^*_\lambda(.)$. The strong concavity of $r^*_\lambda(.)$ is a standard assumption in the literature and is satisfied by many commonly used demand functions such as linear, exponential, and logit functions. It should be noted that although some of these functions, such as logit, do not naturally correspond to a concave revenue function when viewed as a function of p, they are nevertheless concave when viewed as a function of λ. This highlights the benefit of treating revenue as a function of demand rate instead of as a function of price.

In addition, following Besbes and Zeevi (2012) and the literature on nonparametric estimation, I will assume that the function $\lambda^*(.)$ has a certain level of *smoothness*. Let $\bar{s}$ denote the largest integer such that $\left|\partial^{a_1,\ldots,a_n}\lambda^*_j(p)/\partial p_1^{a_1}\ldots\partial p_n^{a_n}\right|$ is uniformly bounded for all $j \in \overline{[1,n]}$ and $0 \leq a_1,\ldots,a_n \leq \bar{s}$. I call $\bar{s}$ the *smoothness index*. I make the following smoothness assumptions:

Nonparametric Function Smoothness Assumptions.

N1. $\bar{s} \geq 2$.

N2. *There exists a constant* $W > 0$ *such that for all* $j \in \overline{[1,n]}$ *and* $p \in \mathcal{P}$ *and integers* $0 \leq a_1,\ldots,a_n \leq \bar{s}$, *I have* $\left|\frac{\partial^{a_1,\ldots,a_n}\lambda^*_j(p)}{\partial p_1^{a_1}\ldots\partial p_n^{a_n}}\right| \leq W$.

The above assumptions are fairly mild and are satisfied by most commonly used demand functions, including linear, polynomial with higher degree, logit, and exponential with a bounded domain of feasible prices. I note that very similar assumptions are also made in Besbes and Zeevi (2012). More broadly, this type of smoothness assumptions are commonly made in the nonparametric estimation literature in statistics (see, for example, Gyorfi et al. (2002)). The smoothness index $\bar{s}$ indicates the level of difficulty in estimating the corresponding demand function: The larger the value of $\bar{s}$, the smoother the demand function, and the easier it is to estimate its shape because its value cannot have a drastic local change.

Admissible controls and the induced probability measures. Let $D_{1:t} := (D_1, D_2, \ldots, D_t)$ and $p_{1:t} := (p_1, p_2, \ldots, p_t)$ denote respectively the observed vectors of demand and price realizations up to and including period t. Let $\mathcal{H}_t$ denote the σ-field generated by

$D_{1:t}$ and $p_{1:t}$. I define a *control* π as a sequence of functions $\pi = (\pi_1, \pi_2, \ldots, \pi_T)$, where π_t is a $\mathcal{H}_{t-1}$-measurable real function that maps the history $D_{1:t-1}$ and $p_{1:t-1}$ to $\otimes_{j=1}^{n}[\underline{p}_j, \bar{p}_j]$. This class of controls is referred to as *non-anticipating controls* because the decision in each period depends only on the accumulated information up to the beginning of the period. Under control π, the seller sets the price in period t equal to $p_t^\pi = \pi_t(D_{1:t-1}, p_{1:t-1})$. Let Π denote the set of all *admissible controls*. That is,

$$\Pi := \left\{ \pi : \sum_{t=1}^{T} AD_t(p_t^\pi) \preceq C \;\; a.s., \;\; \text{and} \;\; p_t^\pi = \pi_t(\mathcal{H}_{t-1}) \right\}.$$

Note that, even though the seller does not know the underlying demand function, the existence of turn-off prices $p_1^\infty, \ldots, p_n^\infty$ guarantees that the capacity constraints can be satisfied if the seller applies p_j^∞ for product j as soon as the remaining capacity is not sufficient to produce one more unit of product j. Let $\mathbb{P}_{\lambda^*,t}^\pi$ denote the induced probability measure under an admissible control $\pi \in \Pi$, i.e.,

$$\mathbb{P}_{\lambda^*,t}^\pi(d_{1:t}) \;=\; \mathbb{P}_{\lambda^*,t}^\pi(D_{1:t} = d_{1:t}) \;=\; \prod_{s=1}^{t} \left[\left(1 - \sum_{j=1}^{n} \lambda_j^*(p_s^\pi)\right)^{\left(1-\sum_{j=1}^{n} d_{s,j}\right)} \prod_{j=1}^{n} \lambda_j^*(p_s^\pi)^{d_{s,j}} \right],$$

where $p_s^\pi = \pi_s(d_{1:s-1})$ and $d_s = [d_{s,j}] \in \mathcal{D}$ for all $s = 1, \ldots, t$. (By definition, the term $1 - \sum_{j=1}^{n} \lambda_j^*(p_s^\pi)$ can be interpreted as the probability of no-purchase in period s under price p_s^π.) For notational simplicity, I will simply write $\mathbb{P}_{\lambda^*,T}^\pi$ as $\mathbb{P}_{\lambda^*}^\pi$ and denote by $\mathbf{E}_{\lambda^*}^\pi$ the expectation with respect to the probability measure $\mathbb{P}_{\lambda^*}^\pi$. The total expected revenue under $\pi \in \Pi$ is $R^\pi = \mathbf{E}_{\lambda^*}^\pi[\sum_{t=1}^{T} p_t^\pi \cdot D_t(p_t^\pi)]$.

The deterministic formulation and performance metric. The following optimization is the deterministic analog of the original stochastic pricing problem:

$$J^D \;:=\; \max_{p_t \in \mathcal{P}} \left\{ \sum_{t=1}^{T} r^*(p_t) : \;\; \sum_{t=1}^{T} A\lambda^*(p_t) \preceq C \right\},$$

$$\text{or equivalently,} \quad J^D \;:=\; \max_{\lambda_t \in \Lambda_{\lambda^*}} \left\{ \sum_{t=1}^{T} r_\lambda^*(\lambda_t) : \;\; \sum_{t=1}^{T} A\lambda_t \preceq C \right\}.$$

By assumption R3, the second optimization above is a convex program and can be efficiently solved. (To avoid triviality, I assume that both optimizations are feasible.) It can be shown that J^D is in fact an upper bound for the total expected revenue under any admissible control. That is, $R^\pi \leq J^D$ for all $\pi \in \Pi$. (See Besbes and Zeevi (2012) for more details.)

This allows me to use J^D as a benchmark to quantify the performance of any admissible pricing control. In this essay, I follow the standard convention and define the *expected revenue loss* of an admissible control $\pi \in \Pi$ as $\rho^\pi := J^D - R^\pi$. Since $r^*_\lambda(.)$ is strongly concave, by Jensen's inequality, it can be shown that the optimal solutions of J^D are static, i.e., $p_t = p^D$ and $\lambda_t = \lambda^D$ for all t, where p^D and λ^D can be obtained by solving the following "one-period" optimizations, respectively:

$$(\mathbf{P}) \qquad r^D \quad := \quad \max_{p \in \mathcal{P}} \left\{ r^*(p) : \;\; A\lambda^*(p) \preceq \frac{C}{T} \right\},$$

$$\text{and,} \quad (\mathbf{P}_\lambda) \quad r^D \quad := \quad \max_{\lambda \in \Lambda_{\lambda^*}} \left\{ r^*_\lambda(\lambda) : \;\; A\lambda \preceq \frac{C}{T} \right\}.$$

Note that $p^D = p^*(\lambda^D)$ and $Tr^D = J^D$. Moreover, the optimal dual variables that correspond to the capacity constraints in $\mathbf{P}$ are the same as the optimal dual variables that correspond to the capacity constraints in $\mathbf{P}_\lambda$; I denote these dual variables as μ^D. Let $\mathsf{Ball}(x, r)$ denote a closed Euclidean ball centered at x with radius r. I state my fourth regularity assumption below:

R4. (Interior Assumption) *There exists* $\phi > 0$ *such that* $\mathsf{Ball}(p^D, \phi) \subseteq \mathcal{P}$.

Assumption R4 is sufficiently mild. Intuitively, it states that the static price should neither be too low that it attracts too much demand nor too high that it induces no demand. A similar interior assumption has also been made in Jasin (2014) and Chen et al. (2016).

Asymptotic setting. Following the convention in the literature (e.g., Besbes and Zeevi (2009) and Wang et al. (2014)), in this essay, I will consider a sequence of increasing problems where the length of the selling season and the initial resource capacity are scaled by a factor of $k > 0$. To be precise, in the k^{th} problem, the length of the selling season and the initial capacity are given by kT and kC, respectively. (One can interpret k as the *size* of the problem. For example, in single-leg setting, $C = 1$ and $k = 50$ could correspond to a small jet with capacity 50 seats and $k = 500$ could correspond to a large jet with capacity 500 seats.) The optimal solutions for $\mathbf{P}$ and $\mathbf{P}_\lambda$ in the k^{th} problem are still p^D and λ^D; the optimal dual solution corresponding to the capacity constraints in $\mathbf{P}$ and $\mathbf{P}_\lambda$ is still μ^D. But, the deterministic upper bound becomes $J^D(k) = kTr^D = kJ^D$. Let $\rho^\pi(k)$ denote the expected revenue loss under an admissible control $\pi \in \Pi$ for the problem with scaling factor k. I am primarily interested in identifying the order of $\rho^\pi(k)$ for large k. (Intuitively, one would expect that a better heuristic control will have an expected revenue loss that grows more slowly with respect to k.) The following notations will be used throughout the remainder of the essay. For any two functions $f : \mathbb{Z}_{++} \to \mathbb{R}$ and $g : \mathbb{Z}_{++} \to \mathbb{R}_+$, I write $f(k) = \Omega(g(k))$ if there exists $M > 0$ independent of k such that $f(k) \geq Mg(k)$. Similarly, I write $f(k) = \mathcal{O}(g(k))$ if there exists $K > 0$ independent of k such that $f(k) \leq Kg(k)$.

4.4 Supporting technical results

In this section, I present some technical results on spline estimation and nonparametric stability analysis of a perturbed optimization problem, and I will also introduce a quadratic programming approximation of **P**. I will use these technical results in the analysis of NSC in Section 4.5. As noted in Section 4.2, many of these results are of independent interest and can potentially be used in different application areas.

4.4.1 Spline approximation

I first describe the problem of approximating a deterministic function from noiseless observations using spline approximation and then I will discuss the problem of estimating a function from noisy observations. Spline functions have been widely used in engineering to approximate complicated functions, and their popularity is primarily due to their flexibility in effectively approximating complex curve shapes (Schumaker 2007). This flexibility lies in the piecewise nature of spline functions—a spline function is constructed by attaching piecewise polynomial functions with a certain degree, and the coefficients of these polynomials are computed in such a way that a sufficiently high degree of smoothness is ensured in the places where the polynomials are connected. More formally, for all $l \in \overline{[1,n]}$, let $\underline{p}_l = x_{l,0} < x_{l,1} \cdots < x_{l,d} < x_{l,d+1} = \bar{p}_l$ be a partition that divides $[\underline{p}_l, \bar{p}_l]$ into $d+1$ sub-intervals of equal length where $d \in \mathbb{Z}_{++}$. Let $\mathcal{G} := \otimes_{l=1}^{n} \mathcal{G}_l$ denote a set of grid points, where $\mathcal{G}_l = \{x_{l,i}\}_{i=0}^{d+1}$. I define the function space of *tensor-product polynomial splines of order* $(s; \ldots; s) \in \mathbb{R}^n$ with a set of grid points $\mathcal{G}$ as $\mathsf{S}(\mathcal{G}, s) := \otimes_{l=1}^{n} \mathsf{S}_l(\mathcal{G}_l, s)$, where $\mathsf{S}_l(\mathcal{G}_l, s) := \{f \in \mathsf{C}^{s-2}([\underline{p}_l, \bar{p}_l]) : f$ is a single-variate polynomial of degree $s-1$ on each sub-interval $[x_{l,i}, x_{l,i+1})$, for all $i \in \overline{[0, d-1]}$ and $[x_{l,d}, x_{l,d+1}]\}$. One of the key questions that spline approximation theory addresses is the following: Given an arbitrary function f that satisfies N1-N2, find a spline function $g^* \in \mathsf{S}(\mathcal{G}, s)$ that approximates f well. Among the various approaches, one of the most popular approximations is using the so-called *tensor-product B-Spline basis functions* (Schumaker 2007). This approach is based on the key observation that $\mathsf{S}(\mathcal{G}, s)$ is actually a linear space of dimension $(s+d)^n$. This means that there exists a set of $(s+d)^n$ basis functions (this set is not necessarily unique) such that any function in $\mathsf{S}(\mathcal{G}, s)$ can be represented by a linear combination of these basis functions. I choose to use tensor-product B-Spline basis functions, denoted by $\{N_{i_1,\ldots,i_n}(x_1, \ldots, x_n)\}_{i_1=1,\ldots,i_n=1}^{s+d,\ldots,s+d}$, as the set of basis functions. These functions are defined formally in the Technical Details part (a) below. Given the basis functions, for any spline function $g \in \mathsf{S}(\mathcal{G}, s)$, there exists a set of coefficients $\{c_{i_1,\dots,i_n}\}_{i_1=1,\dots,i_n=1}^{s+d,\dots,s+d}$ such that $g(x) = \sum_{i_1=1}^{s+d} \cdots \sum_{i_n=1}^{s+d} c_{i_1,\dots,i_n} N_{i_1,\dots,i_n}(x)$ for all $x \in \mathcal{P}$. Therefore, the problem of finding g^* is reduced to the problem of computing the coefficients for representing g^*, which I address below in the Technical Details part (b). For a more comprehensive discussion of this approach, see Schumaker (2007).

Technical Details for Spline Approximation: The B-Spline Approach

---

(a) Tensor-product B-Spline Basis Functions.

Step 1: For each $l \in \overline{[1,n]}$, define an *extended partition* $\mathcal{G}_l^e := \{y_{l,i}\}_{i=1}^{2s+d}$, where

$$y_{l,1} = \cdots = y_{l,s} = x_{l,0},\ y_{l,s+1} = x_{l,1}, \dots, y_{l,s+d} = x_{l,d},\ y_{l,s+d+1} = \cdots = y_{l,2s+d} = x_{l,d+1}.$$

Step 2: For $i_l \in \overline{[1, s+d]}, l \in \overline{[1,n]}$, define the *tensor-product B-Spline basis function* as $N_{i_1,\dots,i_n}(x_1,\dots,x_n) = \prod_{l=1}^{n} N_{l,i_l}^s(x_l)$, where

$$N_{l,i}^s(x_l) = \begin{cases} (-1)^s (y_{l,i+s} - y_{l,i})[y_{l,i},\dots,y_{l,i+s}](x_l - y)_+^{s-1}, & \text{if } y_{l,i} \le x_l < y_{l,i+s} \\ \\ 0, & \text{otherwise} \end{cases}$$

for all $x_l \in [\underline{p}_l, \bar{p}_l]$ for all $l \in \overline{[1,n]}$ and for all $i \in \overline{[1, s+d]}$, where $(x_l - y)_+ = \max\{0, x_l - y\}$, and $[t_1,\dots,t_{r+1}]f(y)$ is the r^{th} *order divided difference* of a single variate real function f over the points $t_1 < t_2 < \cdots < t_r < t_{r+1}$ defined as follows (see Definition 2.49 and Theorem 2.50 in Schumaker (2007) for more discussion):

$$[t_1,\dots,t_{r+1}]f(y) := \sum_{i=1}^{r+1} \frac{f(t_i)}{\prod_{j=1, j\neq i}^{r+1}(t_i - t_j)}.$$

(b) Calculating the Linear Coefficients.

Step 1: For $l \in \overline{[1,n]}, i \in \overline{[1, s+d]}$, let

$$\tau_{l,i,j} = y_{l,i} + (y_{l,i+s} - y_{l,i})\frac{j-1}{s-1} \quad \text{and} \quad \beta_{l,i,j} = \sum_{v=1}^{j} \frac{(-1)^{v-1}}{(s-1)!} \phi_{l,i,s}^{(s-v)}(0) \psi_{l,i,j}^{(v-1)}(0), \quad \text{for } j \in \overline{[1,s]},$$

where $\phi_{l,i,s}(t) = \prod_{r=1}^{s-1}(t - y_{l,i+r})$, $\psi_{l,i,j}(t) = \prod_{r=1}^{j-1}(t - \tau_{l,i,r}), \psi_{l,i,1}(t) \equiv 1$.

Step 2: For any $f \in \mathsf{C}^0(\mathcal{P})$, define a set of linear functionals $\{\gamma_{l,i} : \mathsf{C}^0([\underline{p}_l, \bar{p}_l]) \to \mathbb{R}\}_{l=1,i=1}^{n,s+d}$ as:

$$\gamma_{l,i} f := \sum_{j=1}^{s} \beta_{l,i,j}[\tau_{l,i,1}, \ldots, \tau_{l,i,j}]f = \sum_{j=1}^{s} \beta_{l,i,j} \sum_{r=1}^{j} \frac{f(x_1, \ldots, x_{l-1}, \tau_{l,i,r}, x_{l+1}, \ldots, x_n)}{\prod_{s=1, s\neq r}^{j}(\tau_{l,i,r} - \tau_{l,i,s})},$$

where f is viewed as a single variate function of x_l here, and the second equality follows by Theorem 2.50 in Schumaker (2007) (note that for any given l and i, $\tau_{l,i,1}, \ldots, \tau_{l,i,s}$ are pairwise distinct). Define another set of linear functionals $\{\gamma_{i_1,\ldots,i_n}\}_{i_1=1,\ldots,i_n=1}^{s+d,\ldots,s+d}$ such that

$$\gamma_{i_1,\ldots,i_n} f = \gamma_{1,i_1} \circ \gamma_{2,i_2} \cdots \circ \gamma_{n,i_n} f,$$

where γ_{l,i_l} is understood as being applied to f as a function of x_l. By the construction of γ_{l,i_l}, basic algebra yields:

$$\gamma_{i_1,\ldots,i_n} f = \sum_{j_1=1}^{s} \sum_{r_1=1}^{j_1} \cdots \sum_{j_n=1}^{s} \sum_{r_n=1}^{j_n} \frac{f(\tau_{1,i_1,r_1}, \ldots, \tau_{n,i_n,r_n}) \prod_{l=1}^{n} \beta_{l,i_l,j_l}}{\prod_{l=1}^{n} \prod_{s_l=1, s_l \neq r_l}^{j_l} (\tau_{l,i_l,r_l} - \tau_{l,i_l,s_l})}.$$

Step 3: Define a linear operator $\mathcal{L}_l : \mathsf{C}^0([\underline{p}_l, \bar{p}_l]) \to \mathsf{S}_l(\mathcal{G}_l, s)$ as
$\mathcal{L}_l f(x_l) = \sum_{i=1}^{s+d} (\gamma_{l,i} f) N_{l,i}^s(x_l)$, for all $l \in \overline{[1, n]}$. Similarly, define a linear operator
$\mathcal{L} : \mathsf{C}^0(\mathcal{P}) \to \mathsf{S}(\mathcal{G}, s)$ as

$$\mathcal{L} f(x_1, \ldots, x_n) = \sum_{i_1=1}^{s+d} \cdots \sum_{i_n=1}^{s+d} (\gamma_{i_1,\ldots,i_n} f) N_{i_1,\ldots,i_n}(x_1, \ldots, x_n).$$

Note that $\mathcal{L} = \mathcal{L}_1 \circ \mathcal{L}_2 \circ \cdots \circ \mathcal{L}_n$, where this composition of linear operators is understood as $\mathcal{L}_l$ being applied to a function of x_l.

Step 4: Set $g^* = \mathcal{L} f$.

Spline approximation with noisy observations. I now discuss the estimation of demand function $\lambda^*(.)$ using spline approximation under noisy observations. Let $\tilde{\mathcal{G}} := \{(\tau_{1,i_1,j_1}; \ldots; \tau_{n,i_n,j_n}) : i_l \in \overline{[1, s+d]}, j_l \in \overline{[1, s]} \text{ for all } l \in \overline{[1, n]}\}$. Note that the constants $\{\gamma_{i_1,\ldots,i_n} \lambda_j^*\}_{i_1=1,\ldots,i_n=1}^{s+d,\ldots,s+d}$ depend on $\lambda_j^*(.)$ only via $\lambda_j^*(p), p \in \tilde{\mathcal{G}}$. So, if the seller could observe the demand rate of product j under prices in $\tilde{\mathcal{G}}$, he could construct an approximation of $\lambda_j^*(.)$ using a linear combination of tensor-product B-splines. In my problem, the seller cannot observe $\lambda_j^*(p)$ for $p \in \tilde{\mathcal{G}}$, but only its noisy observation $D_j(p) = \lambda_j^*(p) + \Delta_j$. To address this, I use empirical mean as a surrogate for $\lambda_j^*(p)$ and propose the following *Spline Estimation* algorithm to estimate the demand.

Spline Estimation

Input Parameters: L_0, n, s; **Tuning Parameter:** d

Algorithm:

Step 1: Estimate $\lambda^*(p)$ at points $p \in \tilde{\mathcal{G}}$. For each $p \in \tilde{\mathcal{G}}$

a. Apply price p L_0 times

b. Let $\tilde{\lambda}(p)$ be the sample mean of the L_0 observations.

Step 2: Construct spline approximation.

a. For all $j \in \overline{[1,n]}$ and $i_l \in \overline{[1,s+d]}$, $l \in \overline{[1,n]}$, calculate coefficients $c^j_{i_1,\ldots,i_n}$ as:

$$c^j_{i_1,\ldots,i_n} = \sum_{j_1=1}^{s} \sum_{r_1=1}^{j_1} \cdots \sum_{j_n=1}^{s} \sum_{r_n=1}^{j_n} \frac{\tilde{\lambda}_j(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n}) \ \prod_{l=1}^{n} \beta_{l,i_l,j_l}}{\prod_{l=1}^{n} \prod_{s_l=1,s_l\neq r_l}^{j_l} (\tau_{l,i_l,r_l} - \tau_{l,i_l,s_l})}.$$

b. Construct a tensor-product spline function $\tilde{\lambda}(p) = (\tilde{\lambda}_1(p);\ldots;\tilde{\lambda}_n(p))$, where

$$\tilde{\lambda}_j(p) = \sum_{i_1=1}^{s+d} \cdots \sum_{i_n=1}^{s+d} c^j_{i_1,\ldots,i_n} N_{i_1,\ldots,i_n}(p).$$

Note that the algorithm above conducts $\tilde{L}_0 := L_0(s+d)^n s^n$ samples. The idea of *Spline Estimation* is as follows. In Step 1, I apply each $p \in \tilde{\mathcal{G}}$ as many as L_0 times and calculate its empirical mean $\tilde{\lambda}(p)$. In Step 2, I approximate the underlying demand function $\lambda^*(.)$ using a spline function. In particular, I use a modified version of B-Spline approach by replacing the actual function value $\lambda^*(p)$ ($p \in \tilde{\mathcal{G}}$) with its empirical mean $\tilde{\lambda}(p)$. Note that my estimation approach is different from the so-called *Spline Regression* (see Gyorfi et al. (2002)). While Spline Regression uses Least Squares to compute the linear coefficients for each of the spline basis function, I use the empirical means at sample points and a specific linear operator (originally devised and analyzed in the deterministic approximation theory of spline functions, see Schumaker (2007)) to compute the linear coefficients. I choose to use Spline Estimation in my heuristic instead of Spline Regression because it allows me to use existing results on Spline Approximation Theory to derive the large deviation bounds for Spline Estimation in Lemma 4.4.1. I suspect that similar results also hold for Spline Regression. Let $a \wedge b = \min\{a,b\}$. The following lemma shows how well $\tilde{\lambda}(.)$ approximates $\lambda^*(.)$.

Lemma 4.4.1 *Set $d = \lceil (L_0^{1/2}/\log k)^{1/(s+n/2)} \rceil$ and let $L_0 \geq \log^3 k$ be a positive integer that may depend on k. Suppose that $s \geq 2$. There exist positive constants Ψ_r for each $r \in \overline{[0, (s-2) \wedge \bar{s}]}$ and K independent of $k \geq 1$ such that for all $j \in \overline{[1, n]}$ and $r_l \in \mathbb{Z}_+, l \in \overline{[1, n]}$ satisfying $\sum_{l=1}^{n} r_l = r$, the following holds:*

$$\mathbb{P}^{\pi}\left(\left\| \frac{\partial^r (\lambda_j^*(.) - \tilde{\lambda}_j(.))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}} \right\|_{\infty} \geq \Psi_r \left(\frac{\log k}{\sqrt{L_0}} \right)^{\frac{s \wedge \bar{s} - r}{s + n/2}} \right) \;\leq\; K \exp(-\log^2 k). \tag{4.1}$$

The condition $L_0 \geq \log^3 k$ implies $(\log k/\sqrt{L_0})^{\frac{s \wedge \bar{s} - r}{s+n/2}} \to 0$ as $k \to \infty$. This means that the difference between the r^{th} order partial derivatives of the underlying demand function and the spline approximation is uniformly small with a high probability for large k. When $r = 0$, my bound becomes a large deviation bound for the function estimate itself and is similar to the known bound for Spline Regression. (Suppose I set $s = \bar{s}$ and re-write my bound in a Hoeffding-type form, see Remark 4.4.1 below. Integrating the right hand side with respect to x over $\mathbb{R}_+$, I obtain the ∞-risk of Spline Estimation which is of order $(1/\sqrt{\tilde{L}_0})^{2\bar{s}/(2\bar{s}+n)}$. This is to be compared with the well-known 2-risk of Spline Regression which is of order $(\sqrt{\log \tilde{L}_0/\tilde{L}_0})^{2\bar{s}/(2\bar{s}+n)}$, see Corollary 15.1 in Gyorfi et al. (2002).) I want to stress that the large deviation bound for the function estimate itself is *not* sufficient for my purpose. Specifically, I need additional large deviation bounds for the first and second order partial derivatives of the estimated demand function, as in Lemma 4.4.1, in order to conduct a stability analysis of the deterministic optimization problem **P** in my analysis later.

Remark 4.4.1 (Interpreting (4.1) as a Hoeffding-type Error Bound) *Hoeffding-type error bounds commonly appear in statistical estimations. Informally, they relate a measure of estimation error (e.g., 2-norm of the parameter estimation error in parametric models) with the number of samples L_0 in the following way:*

$$\mathbb{P}(\mathit{Error} \geq x) \leq C_1 \exp(-C_2 L_0 x^2)\,,$$

for some constants C_1 and C_2 that are independent of x and L_0. Note that, in Hoeffding-type of inequality, the right hand side converges to zero as x tends to zero and the variable x shows up as a quadratic term in the exponent. In contrast, when I write (4.1) into a similar form, I obtain

$$\mathbb{P}^{\pi}\left(\left\| \frac{\partial^r (\lambda_j^*(.) - \tilde{\lambda}_j(.))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}} \right\|_{\infty} \geq x \right) \;\leq\; K \exp(-\bar{\Psi}_r L_0 x^{\frac{2s+n}{s \wedge \bar{s} - r}})\,,$$

where $\bar{\Psi}_r = \Psi_r^{-\frac{2s+n}{s\wedge\bar{s}-r}}$. *Due to the well-known curse of dimensionality in nonparametric function estimation, the right hand side of my inequality does* not *tend to zero as fast as a typical Hoeffding-type inequality (i.e.,* $x^2 > x^{\frac{2s+n}{s\wedge\bar{s}-r}}$ *when* x *is small). Moreover, the convergence rate on the right hand side also depends on model parameters. In particular, it is decreasing in* n *and* r*, and is increasing in* $\bar{s}$*. This makes intuitive sense: as the problem dimension* n *increases, estimation becomes more difficult; as the order of derivative* r *decreases or as the smoothness index* $\bar{s}$ *increases, the underlying demand function (or the partial derivative of the underlying demand function) becomes smoother and is easier to estimate. The convergence rate is increasing in* s *when* $s \le \bar{s}$ *because higher* s *allows more flexibility in spline approximation. Interestingly, when* $s > \bar{s}$*, the convergence rate actually decreases in* s*. This is possibly due to the fact that, when* s *is higher than smoothness index* $\bar{s}$*, the extra flexibility introduces unnecessary complexity (i.e., redundant linear coefficients to be estimated), which leads to more sampling.*

4.4.2 Stability analysis

In this subsection, I first present a nonparametric stability result for a class of optimization problems, and then apply this result to the perturbation analysis of my deterministic optimization **P**. Consider the following non-linear optimization problems:

$$(\mathbf{NP}) \quad \max_{x\in\mathcal{X}} \ \{f(x): \ Ug(x) \preceq V\} \quad \text{and} \quad (\tilde{\mathbf{NP}}(\delta)) \quad \max_{x\in\mathcal{X}} \ \left\{\tilde{f}(x): \ U\tilde{g}(x) \preceq V-\delta\right\}.$$

where $\mathcal{X}$ is a convex compact subset of $\mathbb{R}^n$, $f : \mathcal{X} \to \mathbb{R}$ and $g : \mathcal{X} \to \mathbb{R}^n$ are both twice continuously differentiable functions, $\tilde{f} : \mathcal{X} \to \mathbb{R}$ and $\tilde{g} : \mathcal{X} \to \mathbb{R}^n$ are continuously differentiable approximations of f and g, $\delta \in \mathbb{R}^m$, $V \in \mathbb{R}^m$, and U is an m by n non-negative matrix that has full row rank. Let x^* and $\tilde{x}_\delta$ denote the optimal solution of **NP** and $\tilde{\mathbf{NP}}(\delta)$ respectively (i.e., if they are feasible). I state a useful stability result.

Proposition 4.4.1 *Suppose that the following conditions hold:*

(i) $g(.)$ *has a twice continuously differentiable inverse function* $g^{-1}(.) : \mathcal{Y} \to \mathcal{X}$ *where* $\mathcal{Y} := g(\mathcal{X})$ *is a convex compact subset of* $\mathbb{R}^n$*;*

(ii) $f(g^{-1}(.)) : \mathcal{Y} \to \mathbb{R}^n$ *is strongly concave;*

(iii) **NP** *is feasible;*

(iv) x^* *is in the interior of* $\mathcal{X}$*.*

Then, there exist $\bar{\delta} > 0$ *and* $K > 0$ *such that for all* δ, $\tilde{f}(.)$ *and* $\tilde{g}(.)$ *satisfying* $||Ug(.) - U\tilde{g}(.) + \delta||_\infty \leq \bar{\delta}$, $\tilde{\mathbf{NP}}(\delta)$ *is feasible and*

$$||x^* - \tilde{x}_\delta||_2 \;\leq\; K\left(||(\nabla f(.) - \nabla\tilde{f}(.))'||_\infty + ||g(.) - \tilde{g}(.)||_\infty + ||\delta||_\infty\right).$$

The above result can be viewed as a Lipschitz-type stability result for a family of nonparametric optimization problems. Per my discussions in Section 4.2, although stability analysis of parametric optimization problems has been intensively studied in the literature (e.g., Bonnans and Shapiro (2000)), stability results for nonparametric optimization problems are very rare. (See Remark 4.4.2 for a brief discussion on the relationship between Proposition 4.4.1 and existing results on parametric stability analysis.) In my case, since the original unperturbed optimization can be transformed into a convex optimization, I can use a convexity argument to establish Proposition 4.4.1.

I now apply Proposition 4.4.1 to my deterministic optimization problem $\mathbf{P}$. Using the spline approximate $\tilde{\lambda}(p)$ derived in Section 4.4.1, I can formulate an approximate optimization of $\mathbf{P}$ as follows:

$$(\tilde{\mathbf{P}}) \quad \tilde{r}^D := \max_{p\in\mathcal{P}} \left\{\tilde{r}(p): \;\; A\tilde{\lambda}(p) \preceq \frac{C}{T}\right\}$$

where $\tilde{r}(p) = p\cdot\tilde{\lambda}(p)$. Let $\tilde{p}^D$ denote an optimal solution of $\tilde{\mathbf{P}}$ if it is feasible and let $\tilde{\lambda}^D = \tilde{\lambda}(\tilde{p}^D)$. The following lemma follows directly from Proposition 4.4.1 and provides a characterization of $\tilde{p}^D$.

Lemma 4.4.2 *Suppose that* $s \geq 3$. *Then, there exist constants* $\bar{\delta} > 0$ *and* $K > 0$ *such that if* $||\lambda^*(.) - \tilde{\lambda}(.)||_\infty \leq \bar{\delta}$, $\tilde{\mathbf{P}}$ *is feasible and* $||p^D - \tilde{p}^D||_2 \leq K(||\lambda^*(.) - \tilde{\lambda}(.)||_\infty + ||(\nabla\lambda^*(.) - \nabla\tilde{\lambda}(.))'||_\infty)$.

Lemma 4.4.2 means that if demand estimation error is small, $\tilde{\mathbf{P}}$ is feasible and its optimal solution $\tilde{p}^D$ lies in close proximity of p^D. This observation is crucial for my analysis later.

Remark 4.4.2 (On Proposition 4.4.1 and Existing Parametric Stability Results) *The existing Lipschitz stability result of the optimal solution of a parameterized optimization problem (e.g., Theorem 5.53 part (a) in Bonnans and Shapiro (2000)) can be viewed as a special case of my nonparametric Lipschitz-type stability result in Proposition 4.4.1. Let* $\mathcal{U} \subseteq \mathbb{R}^q, q \in \mathbb{Z}_{++}$, *be a compact parameter set. Suppose that the objective functions* f *and* $\tilde{f}$ *come from a family of parameterized functions* $\{f(.;u)\}_{u\in\mathcal{U}}$ *where* $f(.) = f(.;u_0)$ *and*

$\tilde{f}(.) = f(.;v)$ for $u_0, v \in \mathcal{U}$. Also, suppose that the constraint functions g and $\tilde{g}$ come from a family of parameterized functions $\{g(.;u)\}_{u\in\mathcal{U}}$ where $g(.) = g(.;u_0)$ and $\tilde{g}(.) = g(.;v)$ for $u_0, v \in \mathcal{U}$. For simplicity, assume $\delta = \mathbf{0}$. In the perturbation analysis of parametric optimization problems, $f(.;.)$ and $g(.;.)$ are typically assumed to be twice continuously differentiable, which means that $||(\nabla f(.;u_0) - \nabla f(.;v))'||_\infty = \mathcal{O}(||u_0 - v||_\infty)$ and $||g(.;u_0) - g(.:v)||_\infty = \mathcal{O}(||u_0 - v||_\infty)$. Applying Proposition 4.4.1 to this setting immediately yields Theorem 5.53 part (a) in Bonnans and Shapiro (2000).

4.4.3 An approximate quadratic program

In this subsection, I introduce a quadratic program approximation of **P**. (This will be useful when I discuss my heuristic in Section 4.5.) The idea is simple: I approximate the objective of **P** with a quadratic function and its constraints with linear functions. My objective here is to show that if the parameters of the quadratic and linear functions are correctly chosen, the resulting quadratic program will have the same optimal solution as **P** and it will possess some very useful stability properties. To begin with, I first linearize the constraints of **P**. Since the capacity constraints form an affine transformation of the demand function, I will simply linearize the demand function. For any $a \in \mathbb{R}^n, B \in \mathbb{R}^{n\times n}$, let $B_1, \ldots, B_n$ be the columns of B and define $\theta_\iota = (a; B_1; \ldots; B_n) \in \mathbb{R}^{n^2+n}$, where the subscript ι stands for *linear demand*. I denote a linear demand function by $\lambda(p;\theta_\iota) = a + B'p$. Next, I discuss a quadratic approximation for the objective of **P**. For any $E \in \mathbb{R}, F \in \mathbb{R}^n, G \in \mathbb{R}^{n\times n}$, let $G_1, \ldots, G_n$ denote the columns of G and define $\theta_o = (E; F; G_1; \ldots; G_n) \in \mathbb{R}^{n^2+n+1}$ where the subscript o stands for *objective*. I denote the resulting quadratic function by $q(p;\theta_o) = E + F'p + \frac{1}{2}p'Gp$. Finally, let $\theta = (\theta_o; \theta_\iota) \in \mathbb{R}^{2n^2+2n+1}$. For any $\theta \in \mathbb{R}^{2n^2+2n+1}, \delta \in \mathbb{R}^m$, I can define a quadratic program $\mathbf{QP}(\theta;\delta)$ as follows:

$$(\mathbf{QP}(\theta;\delta)) \qquad \max_{p\in\mathcal{P}} \left\{ q(p;\theta_o): \;\; A\lambda(p;\theta_\iota) \preceq \frac{C}{T} - \delta \right\}.$$

If I choose the parameters θ and δ carefully, $\mathbf{QP}(\theta;\delta)$ can be a very good approximation of **P**. Specifically, let $\theta^*_\iota = (a^*; B^*_1; \ldots; B^*_n)$, where $B^* := \nabla\lambda^*(p^D)$ and $a^* := \lambda^D - (B^*)'p^D$. Define an n by n symmetric matrix $H^* := B^*\nabla^2 r^*_\lambda(\lambda^D)(B^*)' - B^* - (B^*)'$. Then, one can verify that

$$H^*_{ij} = -(u^*_{ij})'(B^*)^{-1}\lambda^D, \text{ where } u^*_{ij} = \left[\frac{\partial^2\lambda^*_1(p^D)}{\partial p_i \partial p_j}; \ldots; \frac{\partial^2\lambda^*_n(p^D)}{\partial p_i \partial p_j}\right]. \tag{4.2}$$

(See Appendix A.4.4 for derivation.) Let $\theta_o^* = (E^*; F^*; G_1^*; \dots; G_n^*)$ where

$$E^* := \frac{1}{2}(p^D)' H^* p^D, \;\; F^* := a^* - H^* p^D, \;\; G^* := B^* + (B^*)' + H^*,$$

and let $\theta^* := (\theta_o^*; \theta_\iota^*)$. Note that $\mathbf{QP}(\theta^*; \mathbf{0})$ is a very intuitive approximation of $\mathbf{P}$ since the function $\lambda(p; \theta_\iota^*) = a^* + (B^*)'p = \lambda^D + (B^*)'(p - p^D)$ can be viewed as a linearization of $\lambda^*(.)$ at p^D. (Since $\nabla\lambda^*(p^D)$ is invertible as implied by R1 and R4, I can write $p(\lambda; \theta_\iota^*) = p^D + ((B^*)')^{-1}(\lambda - \lambda^D)$ as the inverse demand function). Note also that the gradients of the objective function and the constraints in $\mathbf{QP}(\theta^*; \mathbf{0})$ at p^D coincide with those in $\mathbf{P}$. By Karush-Kuhn-Tucker (KKT) optimality conditions, it can be shown that the optimal solution of $\mathbf{QP}(\theta^*; \mathbf{0})$ is the same as the optimal solution of $\mathbf{P}$. I formally state these results in Lemma 4.4.3 below. Let $p_\delta^D(\theta)$ and $\mu_\delta^D(\theta)$ denote the optimal primal and dual solutions of $\mathbf{QP}(\theta; \delta)$ respectively (if they exist), and let $\lambda_\delta^D(\theta) = \lambda(p_\delta^D(\theta); \theta_\iota)$.

Lemma 4.4.3 *There exist constants $\kappa > 0, \omega > 0$ and $\bar{\delta} > 0$ such that, for all $\theta_\iota \in \mathsf{Ball}(\theta_\iota^*, \bar{\delta})$, $\theta_o \in \mathsf{Ball}(\theta_o^*, \bar{\delta})$ and $\delta \in \mathsf{Ball}(\mathbf{0}, \bar{\delta})$, the following results hold:*

(a) *B is invertible and $||(B')^{-1}||_2 \leq \omega$;*

(b) *For all $p \in \mathcal{P}$ and for all $i, j \in \overline{[1, n]}$, $||\lambda(p; \theta_\iota) - \lambda(p; \theta_\iota^*)||_2 \leq \omega||\theta_\iota - \theta_\iota^*||_2$ and $|\frac{\partial \lambda_j}{\partial p_i}(p; \theta_\iota) - \frac{\partial \lambda_j}{\partial p_i}(p; \theta_\iota^*)| \leq \omega||\theta_\iota - \theta_\iota^*||_2$;*

(c) *For all $\lambda, \lambda' \in \lambda(\mathcal{P}; \theta_\iota)$, $||p(\lambda; \theta_\iota) - p(\lambda'; \theta_\iota)||_2 \leq \omega||\lambda - \lambda'||_2$;*

(d) *$q(p(.; \theta_\iota); \theta_o)$ is strongly concave.*

(e) *$p^D = p_{\mathbf{0}}^D(\theta^*)$, $\lambda^D = \lambda_{\mathbf{0}}^D(\theta^*)$, $\mu^D = \mu_{\mathbf{0}}^D(\theta^*)$;*

(f) *$\mathbf{QP}(\theta; \delta)$ is feasible and has a unique optimal solution. Moreover, $p_\delta^D(\theta) \in \mathsf{Ball}(p_{\mathbf{0}}^D(\theta^*), \phi/2)$, $\mathsf{Ball}(p_\delta^D(\theta), \phi/2) \subseteq \mathcal{P}$, $||p_{\mathbf{0}}^D(\theta^*) - p_\delta^D(\theta)||_2 \leq \kappa(||\theta^* - \theta||_2 + ||\delta||_2)$, $||\lambda_{\mathbf{0}}^D(\theta^*) - \lambda_\delta^D(\theta)||_2 \leq \kappa(||\theta^* - \theta||_2 + ||\delta||_2)$, and the constraints of $\mathbf{QP}(\theta; \delta)$ that correspond to the rows $\{i : \mu_{\mathbf{0},i}^D(\theta^*) > 0\}$ are binding.*

Note that Lemma 4.4.3 part (f) not only establishes Lipschitz continuity of the optimal solution, but also provides additional results regarding the properties of the capacity constraints at the optimal solution. These play an important role in deriving a sharp performance bound of my heuristic.

Remark 4.4.3 (On the Quadratic Revenue Function Approximation) *Note that, as the equation below shows, $q(p;\theta_o^*)$ can be viewed as the revenue function under the approximate linear demand function plus an additional correction term:*

$$\begin{aligned} q(p;\theta_o^*) &= \frac{1}{2}(p^D)'H^*p^D + p'(a^* - H^*p^D) + \frac{1}{2}p'(B^* + (B^*)' + H^*)p \\ &= p'(a^* + (B^*)'p) + \frac{1}{2}(p - p^D)'H^*(p - p^D) \\ &= r(p;\theta_\iota^*) + \frac{1}{2}(p - p^D)'H^*(p - p^D), \end{aligned}$$

where $r(p;\theta_\iota^*) := p \cdot \lambda(p;\theta_\iota^*)$ is the natural revenue function under the approximate linear demand function. I add a correction term above in order to ensure that if I do a change of variables $p = p(\lambda;\theta_\iota^*)$ to change the pricing decision to demand rate decision, the resulting objective function is actually the second order Taylor's expansion of r_λ^*:

$$\begin{aligned} &q(p(\lambda;\theta_\iota^*);\theta_o^*) \\ &= r(p(\lambda;\theta_\iota^*);\theta_\iota^*) + \frac{1}{2}(p(\lambda;\theta_\iota^*) - p^D)'H^*(p(\lambda;\theta_\iota^*) - p^D) \\ &= \lambda'(p^D + ((B^*)')^{-1}(\lambda - \lambda^D)) + \frac{1}{2}(\lambda - \lambda^D)'(B^*)^{-1}H^*((B^*)')^{-1}(\lambda - \lambda^D) \\ &= \lambda'(p^D + ((B^*)')^{-1}(\lambda - \lambda^D)) - (\lambda - \lambda^D)'((B^*)')^{-1}(\lambda - \lambda^D) + \frac{1}{2}(\lambda - \lambda^D)'\nabla^2 r_\lambda^*(\lambda^D)(\lambda - \lambda^D) \\ &= \lambda'p^D + (\lambda - \lambda^D)'(B^*)^{-1}\lambda^D + \frac{1}{2}(\lambda - \lambda^D)'\nabla^2 r_\lambda^*(\lambda^D)(\lambda - \lambda^D) \\ &= r_\lambda^*(\lambda^D) + (\lambda - \lambda^D)'(p^D + (B^*)^{-1}\lambda^D) + \frac{1}{2}(\lambda - \lambda^D)'\nabla^2 r_\lambda^*(\lambda^D)(\lambda - \lambda^D) \\ &= r_\lambda^*(\lambda^D) + \nabla r_\lambda^*(\lambda^D)'(\lambda - \lambda^D) + \frac{1}{2}(\lambda - \lambda^D)'\nabla^2 r_\lambda^*(\lambda^D)(\lambda - \lambda^D). \end{aligned} \tag{4.3}$$

*Hence, in light of R3, $q(p(\lambda;\theta_\iota^*);\theta_o^*)$ is strongly concave in λ. This observation is important because it allows me to use the general result in Proposition 4.4.1 to derive perturbation result for the optimal primal and dual solutions of **QP**$(\theta;\delta)$ (see condition (ii) in Proposition 4.4.1).*

4.5 Main result

I am now ready to describe *Nonparametric Self-adjusting Control* (NSC) and discuss its asymptotic performance; the proof of this result is given in Section 4.6. NSC consists of an exploration procedure and an exploitation procedure. The exploration procedure uses the Spline Estimation algorithm discussed in Section 4.4.1 to construct a spline approximation $\tilde{\lambda}(.)$ of the underlying demand function $\lambda^*(.)$. This function $\tilde{\lambda}(.)$ is then used to construct a linear function $\lambda(.;\hat{\theta}_\iota)$ that closely approximates $\lambda(.;\theta_\iota^*)$ in the neighborhood of p^D and a quadratic program that closely approximates $\mathbf{P}$. During the exploitation phase, I use the optimal solution of the approximate quadratic program as baseline control and automatically adjust the price according to a pre-determined price update rule. Further detail will be provided below. Recall that $\tilde{L}_0$ is the duration of the Spline Estimation algorithm. Let C_t denote the remaining capacity at the *end* of period t. Let $\hat{\theta} := (\hat{\theta}_o;\hat{\theta}_\iota)$, where $\hat{\theta}_\iota := (\hat{a};\hat{B}_1;\ldots;\hat{B}_n), \hat{\theta}_o := (\hat{E};\hat{F};\hat{G}_1;\ldots;\hat{G}_n)$ for

$$\begin{aligned}
&\hat{B} := \nabla\tilde{\lambda}(\tilde{p}^D),\ \hat{a} := \tilde{\lambda} - \hat{B}'\tilde{p}^D,\ \hat{E} := \tfrac{1}{2}(\tilde{p}^D)'\hat{H}\tilde{p}^D,\ \hat{F} := \hat{a} - \hat{H}\tilde{p}^D,\\
&\hat{G} := \hat{B} + \hat{B}' + \hat{H},\ \text{and } \hat{H} = [\hat{H}_{ij}]\\
&\text{where } \hat{H}_{ij} := -\hat{u}_{ij}'\hat{B}^{-1}\tilde{\lambda}^D \text{ and } \hat{u}_{ij} := \left[\frac{\partial^2\tilde{\lambda}_1(\tilde{p}^D)}{\partial p_i \partial p_j};\ldots;\frac{\partial^2\tilde{\lambda}_n(\tilde{p}^D)}{\partial p_i \partial p_j}\right].
\end{aligned}$$

(From Section 4.4.2, $\tilde{p}^D$ is an optimal solution of $\tilde{\mathbf{P}}$.) My proposed NSC heuristic is given below.

Nonparametric Self-adjusting Control (NSC)

Input parameters: $n, s,$ **Tuning Parameters:** d, L_0

Stage 1 (Exploration Phase 1 - Spline Estimation)

a. For $t = 1$ to $\tilde{L}_0 \wedge T$:
  - If $C_{t-1} \prec A_j$ for some $j = 1,\ldots,n$, set $p_{t,j} = p_j^\infty$ for all $j = 1,\ldots,n$.
  - Otherwise, follow Step 1 in *Spline Estimation* algorithm.

b. At the end of period $\tilde{L}_0 \wedge T$, do:
  - If $\tilde{L}_0 \geq T$, terminate NSC.
  - If $\tilde{L}_0 < T$ and $C_{\tilde{L}_0} \prec A_j$ for some $j = 1,\ldots,n$:
    - For all $t > \tilde{L}_0$, set $p_{t,j} = p_j^\infty$ for all $j = 1,\ldots,n$.
    - Terminate NSC.
  - If $\tilde{L}_0 < T$ and $C_{\tilde{L}_0} \succeq A_j$ for all $j = 1,\ldots,n$:
    - Follow Step 2 in *Spline Estimation* algorithm to get $\tilde{\lambda}(.)$.
    - Go to Stage 2 below.

Stage 2 (Exploration Phase 2 - Function Approximation)

a. Solve $\tilde{\mathbf{P}}$ and obtain the optimizer $\tilde{p}^D$.

b. Let $\delta := C/T - C_{\tilde{L}_0}/(T - \tilde{L}_0)$.

c. Compute $\hat{a}, \hat{B}, \hat{E}, \hat{F}, \hat{G}, \hat{H}$ and $\hat{\theta} = (\hat{\theta}_o;\hat{\theta}_\iota)$.
  - If $\hat{B}$ is invertible, go to Stage 2(d) below.

- Otherwise, for $t = \tilde{L}_0 + 1$ to T:
  - If $C_{t-1} \succeq A_j$ for $j = 1, \ldots, n$, apply $p_t = \tilde{p}^D$.
  - Otherwise, for product $j = 1$ to n, do:
    - If $C_{t-1} \prec A_j$, set $p_{t,j} = p_j^{\infty}$.
    - Otherwise, set $p_{t,j} = p_{t-1,j}$.

d. Solve $\mathbf{QP}(\hat{\theta}; \delta)$ for its static price $p_\delta^D(\hat{\theta})$.

Stage 3 (Exploitation)

For $t = \tilde{L}_0 + 1$ to T:

- Compute: $\hat{p}_t = p_\delta^D(\hat{\theta}) - \nabla_\lambda p(\lambda_\delta^D(\hat{\theta}); \hat{\theta}_\iota) \cdot \sum_{s=\tilde{L}_0+1}^{t-1} \frac{\tilde{\Delta}_s}{T-s}$, where $\tilde{\Delta}_t := D_t - \lambda(p_t; \hat{\theta}_\iota)$.
- If $\hat{p}_t \in \mathcal{P}$ and $C_{t-1} \succeq A_j$ for $j = 1, \ldots, n$, apply $p_t = \hat{p}_t$.
- Otherwise, for product $j = 1$ to n, do:
  - If $C_{t-1} \prec A_j$, set $p_{t,j} = p_j^{\infty}$.
  - Otherwise, set $p_{t,j} = p_{t-1,j}$.

I now explain the main ideas behind NSC. The exploitation part (Stage 3) of NSC is motivated by LRC heuristic developed in Jasin (2014), which (roughly) uses

$$p_t \;=\; p^* \left(\lambda^D - \sum_{s=1}^{t-1} \frac{\Delta_s}{T-s} \right), \quad \text{where } \Delta_t = D_t(p_t) - \lambda^*(p_t)$$

and has a strong performance guarantee in the setting of known demand function. In my setting, the demand function $\lambda^*(.)$ is unknown (hence, the inverse demand function $p^*(.)$ is also unknown) *and* the sequence $\{\Delta_t\}_{t=1}^T$ is not observable. If I still wish to use LRC, an intuitive fix is to replace $\lambda^*(.)$ and $\{\Delta_t\}_{t=1}^T$ with their best estimates. This motivates the use of Spline Estimation in Stage 1 to compute an approximate demand function $\tilde{\lambda}(.)$. However, although $\tilde{\lambda}(.)$ can approximate $\lambda^*(.)$ well by tapping into the smoothness of $\lambda^*(.)$, the piecewise nature of spline functions and the shape of the spline basis functions imply that $\tilde{\lambda}(.)$ may not be invertible, i.e., $\tilde{\lambda}(.)$ may not admit a well-defined inverse demand function. But, this is crucial since LRC uses $p^*(.)$ to adjust the prices. This motivates me to use demand linearization in Stage 2. The objective of Stage 2 is to construct a linear function that closely approximates the linearization of the true demand function $\lambda^*(.)$ around p^D and construct a quadratic program that closely approximates $\mathbf{P}$ around its optimal solution p^D. I choose to use linear approximation of the demand function and quadratic approximation of the revenue function because, by Lemma 4.4.3 part (e), the optimal solution of the constructed approximate quadratic program coincides with the optimal solution of $\mathbf{P}$ if the parameters are chosen to be $(\theta^*, \mathbf{0})$ (see Section 4.4.3 for more discussions). Although θ^* is unknown to the seller, I can utilize the spline approximation $\tilde{\lambda}(.)$ to construct parameters $\hat{\theta}$ that closely approximate θ^*. To see why this is so, note that if L_0 is carefully selected, the spline estimation procedure yields a spline function $\tilde{\lambda}(.)$ that closely approximates $\lambda^*(.)$, together with its first and second order partial derivatives (by Lemma 4.4.1), with a very high probability; this in turn indicates that any optimizer $\tilde{p}^D$ of $\tilde{\mathbf{P}}$ lies in a close proximity of p^D (by Lemma 4.4.2). Since θ^* (resp. $\hat{\theta}$) can essentially be viewed as a function of p^D (resp. $\tilde{p}^D$), $\lambda^*(p^D)$ (resp. $\tilde{\lambda}(\tilde{p}^D)$) and its first and second order derivatives evaluated at p^D (resp. $\tilde{p}^D$), this suggests that $\hat{\theta} = (\hat{\theta}_o; \hat{\theta}_\iota)$ is a good approximation of $\theta^* = (\theta^*_o; \theta^*_\iota)$. It is worth stressing that Spline Estimation is crucial for determining reasonably good linear demand and quadratic revenue function approximations. As mentioned above, among all possible approximate linear demand functions, only those that are linearized at a point close to p^D are effective. (Similarly for the revenue functions.) To find a point that is close to p^D (i.e., $\tilde{p}^D$ in my NSC) via optimizing the approximate deterministic pricing problem, I need to use Spline Estimation to get an approximate function that *uniformly* approximates the underlying demand function well.

Finally, after obtaining $\lambda(.; \hat{\theta}_\iota)$, I replace $p^*(.)$ and Δ_t in LRC with $p(.; \hat{\theta}_\iota)$ and $\tilde{\Delta}_t$. This leads to the price update formula in Stage 3:

$$\hat{p}_t \; = \; p^D_\delta(\hat{\theta}) - \nabla_\lambda p(\lambda^D_\delta(\hat{\theta}); \hat{\theta}_\iota) \cdot \sum_{s=\tilde{L}_0+1}^{t-1} \frac{\tilde{\Delta}_s}{T-s}.$$

It is natural to expect that if $\lambda(.; \hat{\theta})$ approximates $\lambda^*(.)$ well, then NSC should retain the strong performance of LRC, this intuition is not immediately obvious and requires a mathematical justification. Note that, in addition to demand randomness, there are at least three sources of errors that affect the performance of NSC: (1) errors from functional estimation (i.e., due to estimating $\lambda^*(.)$ with $\tilde{\lambda}(.)$), (2) errors from function approximation (i.e., due to demand linearization and quadratic approximation), and (3) errors from *systematic* biases due to the terms $\{\tilde{\Delta}_t\}_{t=1}^T$. (In LRC, the perturbation term $\sum_{s=1}^{t-1} \Delta_s/(T-s)$ is unbiased because $\mathbf{E}^\pi_{\lambda^*}[\Delta_t] = 0$. In contrast, in NSC, the perturbation term $\sum_{s=\tilde{L}_0+1}^{t-1} \tilde{\Delta}_s/(T-s)$ is biased because $\mathbf{E}^\pi_{\lambda^*}[\tilde{\Delta}_t] \neq 0$. This means that I am systematically introducing new biases in each period. It is not a priori clear what kind of impact these biases will have on revenue performance.) Thus, despite the strong performance of LRC in the known demand function setting, it is not a priori clear whether self-adjusting alone, without re-optimizations *and* without re-estimations during Stage 3, is sufficient to reduce the impact of these errors on expected revenue loss. Interestingly, the following result states that the performance of NSC is close to the best achievable (asymptotic) performance bound.

Theorem 4.5.1 *Suppose that I use $s \geq 4$, $L_0 = \lceil (kT)^{(s+n/2)/(2s+n-2)} (\log(kT))^{(2s+n-4)/(2s+n-2)} \rceil$ and $d = \lceil (L_0^{1/2}/\log(kT))^{1/(s+n/2)} \rceil$. There exists a constant $M_1 > 0$ independent of $k > 3$ such that for all $s \geq 4$, I have*

$$\rho^{NSC}(k) \quad \leq \quad M_1 k^{\frac{1}{2}+\epsilon(n,s,\bar{s})} \log k, \quad \textit{where } \epsilon(n,s,\bar{s}) = \frac{1}{2}\left(\frac{2s-2(s\wedge\bar{s})+n+2}{2s+n-2}\right).$$

Some comments are in order. First, unlike the heuristic proposed in Besbes and Zeevi (2012), which requires the knowledge of $\bar{s}$, NSC does *not* require the knowledge of the smoothness index $\bar{s}$. This is practically appealing because it is usually difficult to guess the smoothness index of a function when the function itself is unknown. Second, since most commonly used demand functions such as polynomial with arbitrary degree, logit, and exponential are infinitely differentiable (i.e., $\bar{s}$ can be arbitrarily large), for any fixed $\epsilon > 0$, I can select integers $s \geq (n+2)/(4\epsilon) - (n-2)/2$ such that the performance under NSC is $\mathcal{O}(k^{1/2+\epsilon} \log k)$. Theoretically, this means that the asymptotic performance of NSC is very close to the best achievable performance lower bound of $\Omega(\sqrt{k})$. Third, despite the systematic biases it introduces, self-adjusting control in Stage 3 (surprisingly) plays a vital role in guaranteeing the stated performance bound. To illustrate, consider the case where $\bar{s}$ is arbitrarily large. Suppose that I only apply static price $p_t = p_\delta^D(\hat{\theta})$ throughout Stage 3, subject to capacity constraints. Then, under the optimally tuned L_0 and s, one can show that the resulting expected revenue loss is $\mathcal{O}(k^{2/3+\epsilon} \log k)$, which is significantly worse than the bound in Theorem 4.5.1. This underscores the importance of self-adjusting price update in reducing the expected revenue loss from $\mathcal{O}(k^{2/3+\epsilon} \log k)$ to $\mathcal{O}(k^{1/2+\epsilon} \log k)$. Finally, to further validate the theoretical result in Theorem 4.5.1, I conduct a simple numerical study with two types of products and two types of resources. Table 4.1 shows that NSC performs well: For problems with a wide range of k, its relative revenue loss (i.e., $\rho^\pi(k)/J^D(k)$) is about 3 - 8% lower than the relative revenue loss of Algorithm 3 in Besbes and Zeevi (2012). To implement NSC for large-scale problems, the main computational burden lies in solving the nonlinear optimization $\tilde{\mathbf{P}}$ because $\tilde{\lambda}(p)$ is stitched together by many (not necessarily concave) multinomial function. (In fact, Algorithm 3 in Besbes and Zeevi (2012) also suffers from this computational complexity. Moreover, I would also like to point out that, in contrast to local polynomial approximation used in Algorithm 3 in Besbes and Zeevi (2012), my spline approximation is globally differentiable and is more amenable to optimizations.) Thus, for problems with many different types of products and resources, one may want to optimize an approximation of $\tilde{\mathbf{P}}$ that is computationally more tractable. The question of which approximation should be used is an important and practically relevant one; however, it is beyond the scope of the current essay and I leave it for future research pursuit.

Remark 4.5.1 (On the analysis of uncapacitated vs. capacitated RM) *Per my discussions in Section 4.2, most existing literature on joint learning and pricing focus on the setting of uncapacitated RM where there is no limit on the number of resources that can be used. In such setting, it has been repeatedly shown in the literature that the $\Omega(\sqrt{k})$ lower bound is actually tight for both single product and multiple products settings (see Besbes and Zeevi (2009), Broder and Rusmevichientong (2012), Keskin and Zeevi (2014)). The presence of capacity constraints makes the problem significantly more challenging. To see this, note that, if I mis-calculate p^D by ϵ (i.e., I use $\tilde{p}^D = p^D + \epsilon$), by the strong concavity of $r^*_\lambda(.)$ and Lipschitz property of demand, $r^*(p^D) - r^*(\tilde{p}^D)$ is approximately on the order of ϵ^2 in the uncapacitated setting (because $\nabla r^*(p^D) = \mathbf{0}$ due to p^D being the unconstrained optimizer of $r^*(p)$). Thus, the expected revenue loss during T periods is $\mathcal{O}(T\epsilon^2)$. In contrast, in the capacitated setting, $\nabla r^*(p^D) \neq \mathbf{0}$ in general. This means that $r^*(p^D) - r^*(\tilde{p}^D)$ is on the order of ϵ, which implies that the expected revenue loss during T periods is $\mathcal{O}(T\epsilon)$. This is the reason why the analysis in uncapacitated RM is not directly applicable to capacitated RM.*

Remark 4.5.2 (Applying NSC in deterministic demand arrival case) *Although my NSC is designed for the stochastic demand case, it can be readily adapted and applied in the deterministic demand case as well. In this case, there is no random noise in demand observations, so one can simply set $L_0 = 1$ in the Spline Estimation subroutine. The other tuning parameter d needs to be adjusted accordingly. Specifically, given $L_0 = 1$, for any $s \geq 4$ and d, the estimation error of the demand function and its first order partial derivatives are in the order of $\epsilon := \mathcal{O}(d^{-(s\wedge\bar{s}-1)})$ by a similar analysis as in Step 1 in the proof of Lemma 4.4.1. The expected revenue loss during the exploration stages is in the order of the number of prices being tested, i.e., $\mathcal{O}(d^n)$, while the expected revenue loss during the exploitation stage is $\mathcal{O}(\epsilon^2 k)$. Hence, the expected revenue loss throughout the selling season is $\mathcal{O}(\epsilon^{-\frac{n}{s\wedge\bar{s}-1}} + \epsilon^2 k)$, which is minimized at $\epsilon = k^{-\frac{s\wedge\bar{s}-1}{2(s\wedge\bar{s})+n-2}}$. Thus, by setting $d = k^{\frac{1}{2(s\wedge\bar{s})+n-2}}$, the performance bound of NSC for deterministic demand is $\mathcal{O}(k^{\frac{n}{2(s\wedge\bar{s})+n-2}})$. This means that when the demand function is sufficiently smooth (i.e., $\bar{s} = \infty$), for any $\epsilon > 0$, I can choose s large enough so that the performance of NSC in the deterministic demand setting is $\mathcal{O}(k^\epsilon)$. This highlights the fact that stochastic and deterministic demand cases have different complexities.*

Remark 4.5.3 (On my demand linearization approach) *Although the estimated spline function is not used in the exploitation stage once the function approximations (i.e., linear* *demand approximation and quadratic revenue approximation) are conducted, I would like to re-iterate that Spline Estimation is crucial in NSC as it ensures that the demand function is linearized at a point that is sufficiently close to p^D so that the resulting function approximations are reasonably good. Other demand linearization approaches have been proposed in the literature as well. For example, in the* single product without capacity constraint setting*, by using the simple structure of the optimal price $\lambda(p^*) + p^*\lambda'(p^*) = 0$, Besbes and Zeevi (2015) propose a simpler and more direct demand linearization approach that works well in their setting. This approach is unlikely to work in my* multiple products with multiple capacity constraints *setting because the optimal solution p^D does not permit the same simple structure anymore; instead, it is characterized by the KKT conditions (i.e., one needs to compare a combinatorial number of KKT points to find p^D).*

4.6 Proof of Theorem 4.5.1

In this section, I provide a complete proof of Theorem 4.5.1. I first discuss an outline of the proof, together with the key ideas and key lemmas, in Section 4.6.1 and then I fill in the remaining details in Sections 4.6.2 - 4.6.4. Throughout this section, I fix $\pi=$ NSC and assume that $T = 1$ (this is without lost of generality).

4.6.1 Key ideas and outline of the proof

The proof of Theorem 4.5.1 uses a combination of large deviation arguments, stability analysis, and stopping time arguments. Below, I divide the proof into three parts.

Part 1

In this part, I argue that, if k is large, $||\theta^* - \hat{\theta}||_2$ is small with a very high probability. This result allows me to use the perturbation result in Lemma 4.4.3 when analyzing the revenue loss later (in Part 3). Let $\epsilon(L_0) = (\log k/\sqrt{L_0})^{((s\wedge\bar{s})-2)/(s+n/2)}$ and define $\mathcal{E} := \{||\theta^* - \hat{\theta}||_2 \leq M_2\epsilon(L_0)\}$, where M_2 is as defined in Lemma 4.6.1 below.

Lemma 4.6.1 *There exist constants $M_2, M_3 > 0$ independent of $k \geq 1$ and L_0 such that if $L_0 \geq \log^3 k$ and $s \geq 4$, then $\mathbb{P}^{\pi}_{\lambda^*}(||\theta^* - \hat{\theta}||_2 > M_2\epsilon(L_0)) \leq M_3/k$.*

The complete proof of Lemma 4.6.1 is given in Section 4.6.2. Here, I simply provide some basic intuition behind the proof. The proof uses Lemma 4.4.1 and Lemma 4.4.2. In particular, recall that Lemma 4.4.1 indicates that, with a very high probability, the spline function $\tilde{\lambda}(.)$ closely approximates the underlying demand function $\lambda^*(.)$ both in terms of the function value and its first and second order partial derivatives when $s \geq 4$ and $\bar{s} \geq 2$. This result together with the nonparametric perturbation result in Lemma 4.4.2 establishes that $\tilde{p}^D$ is very close to p^D with very high probability. But then, the first and second order derivatives of $\tilde{\lambda}(.)$ evaluated at $\tilde{p}^D$ also closely approximate those of $\lambda^*(.)$ evaluated at p^D. Note that by construction, θ^* (resp. $\hat{\theta}$) can essentially be viewed as a function of p^D (resp. $\tilde{p}^D$), $\lambda^*(p^D)$ (resp. $\tilde{\lambda}(\tilde{p}^D)$) and its first and second order derivatives evaluated at p^D (resp. $\tilde{p}^D$). Hence, $\hat{\theta}$ closely approximates θ^* with a very high probability.

In the remainder of this part, I first discuss some observations that follow from Lemma 4.6.1 and then use these observations to define a constant Ω_1, which will be used in Parts 2 and 3. (As will be clear later, the problem behaves nicely and the prerequisite of Lemma 4.4.3 is satisfied when $k \geq \Omega_1$. This sets the stage for the analysis in Parts 2 and 3.) Four observations are in order. First, note that, as k becomes large, the probability of $\mathcal{E}$ tends to one. Second, $\epsilon(L_0) \to 0$ as $k \to \infty$ under the condition that $L_0 > \log^3 k$, and this condition is satisfied for all sufficiently large k because my selection of L_0 implies

$$k^{\frac{s+n/2}{2s+n-2}}(\log k)^{\frac{2s+n-4}{2s+n-2}} \leq L_0 \leq 2k^{\frac{s+n/2}{2s+n-2}}(\log k)^{\frac{2s+n-4}{2s+n-2}}, \tag{4.4}$$

and $\log^3 k$ is smaller than the left hand side of (4.4) for large k. In light of Lemma 4.6.1, this observation means that if k is large, θ^* and $\hat{\theta}$ can be arbitrarily close with a very high probability. By (4.4), I have the following bounds for $\epsilon(L_0)$ as well:

$$2^{-\frac{s\wedge\bar{s}-2}{2s+n}}\left(\frac{\log k}{\sqrt{k}}\right)^{\frac{s\wedge\bar{s}-2}{2s+n-2}} \leq \epsilon(L_0) \leq \left(\frac{\log k}{\sqrt{k}}\right)^{\frac{s\wedge\bar{s}-2}{2s+n-2}}. \tag{4.5}$$

Third, by definition of $\tilde{L}_0$ and my choice of d in Theorem 4.5.1, I can bound

$$\tilde{L}_0 = s^n(s+d)^n L_0 \leq s^n(s+1)^n d^n L_0 \leq s^n(s+1)^n 2^{n+2} k^{\frac{s+n}{2s+n-2}}(\log k)^{\frac{2(s-2)}{2s+n-2}}, \tag{4.6}$$

where the second inequality follows from (4.4). The above inequality implies $\tilde{L}_0/k \to 0$ as $k \to \infty$. So, there exists a constant $\Omega_0 > 0$ such that for all $k > \Omega_0$, I have $\tilde{L}_0 \leq k/2$. Fourth, there exists a constant $M_4 > 0$ independent of $k \geq \Omega_0$ such that for all $k \geq \Omega_0$,

$$\begin{aligned} ||\delta||_2 &= \left|\left|C - \frac{C_{\tilde{L}_0}}{k-\tilde{L}_0}\right|\right|_2 = \left|\left|\frac{(kC - \tilde{L}_0 C) - (kC - A\sum_{s=1}^{\tilde{L}_0} D_s)}{k-\tilde{L}_0}\right|\right|_2 = \left|\left|\frac{A\sum_{s=1}^{\tilde{L}_0} D_s - \tilde{L}_0 C}{k-\tilde{L}_0}\right|\right|_2 \\ &\leq 2\left(||A\mathbf{e}||_2 + ||C||_2\right)\frac{\tilde{L}_0}{k} \leq 2\left(||A\mathbf{e}||_2 + ||C||_2\right) s^n(s+1)^n 2^{n+2}(\log k/\sqrt{k})^{\frac{2(s-2)}{2s+n-2}} \\ &\leq 2\left(||A\mathbf{e}||_2 + ||C||_2\right) s^n(s+1)^n 2^{n+2}(\log k/\sqrt{k})^{\frac{2(s\wedge\bar{s}-2)}{2s+n-2}} \leq M_4\epsilon(L_0)^2 \end{aligned}$$

where the first inequality follows because $\tilde{L}_0 \leq k/2$ for $k \geq \Omega_0$ and I have at most one arrival per period, the second inequality follows from (4.6), and the last inequality follows from (4.5).

Let $\bar{\delta}$ be defined as in Lemma 4.4.3. Putting together the four observations above, I conclude that there exists a constant $\Omega_1 > 0$ independent of k such that for all $k \geq \Omega_1$ the following holds:

$$\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{E}) \;\geq\; 1 - M_3/k \;\geq\; 1/2; \tag{4.7}$$

$$\text{Conditioning on } \mathcal{E}, ||\theta^* - \hat{\theta}||_2 \leq M_2\epsilon(L_0) \leq \bar{\delta}; \tag{4.8}$$

$$||\delta||_2 \leq M_4\epsilon(L_0)^2 \leq \bar{\delta}. \tag{4.9}$$

Inequality (4.7) indicates that I only need to focus on the revenue loss on the event $\mathcal{E}$. Inequalities (4.8) and (4.9) are crucial; they ensure that, for $k \geq \Omega_1$, conditioning on $\mathcal{E}$, the prerequisite of Lemma 4.4.3 is satisfied and the perturbation bounds therein can be used to analyze the performance of NSC.

Part 2

In this part, I define a stopping time τ and analyze its properties. This will be crucial for my analysis in Part 3. In particular, it helps me to quantify the amount of revenue loss under NSC during the exploitation phase. (Stopping time argument is also used in Jasin (2014). However, unlike the arguments in Jasin (2014), which assume *known demand function*, here I also need to deal with estimation errors, approximation errors, and systematic biases.) I first define τ and state its properties in Lemmas 4.6.2 and 4.6.3. For clarity, I delay the complete proof of these two lemmas in Section 4.6.3 and only discuss the intuition here. Let τ be the minimum of k and the first time $t \geq \tilde{L}_0 + 1$ such that the following condition (†) is violated:

$$(\dagger) \qquad \psi > S(t), \text{ where } \psi := \sqrt{\epsilon(L_0)},\; S(k) := \infty \text{ and } \forall t \in \overline{[1, k-1]},$$
$$S(t) := \left|\left|\sum_{s=\tilde{L}_0+1}^{t} \frac{\tilde{\Delta}_s}{k-s}\right|\right|_2 + \frac{1}{k-t}.$$

(Recall that $\tilde{\Delta}_s = \Delta_s + \lambda^*(p_s) - \lambda(p_s; \hat{\theta}_\iota)$.) The purpose of condition (†) is to guarantee that $\hat{p}_t$ is not too far way from $p^D_\delta(\hat{\theta})$ (see the pricing formula in Stage 3 of NSC) before τ and the cumulative deviation of the actual demand realizations from the target average demand is not too large before τ. Let Ω_1 and $\mathcal{E}$ be as defined in Part 1. I state two lemmas.

Lemma 4.6.2 *Suppose that $L_0 \geq \log^3 k$. There exists a constant $\Omega_2 > \Omega_1$ independent of $k \geq 1$ such that for all $k \geq \Omega_2$ and all sample paths on $\mathcal{E}$, $\hat{p}_t \in \mathcal{P}$ (i.e., $p_t = \hat{p}_t$) and $C_t \succeq A_j$ for all $j \in \overline{[1, n]}$ and $t \in \overline{[\tilde{L}_0 + 1, \tau - 1]}$.*

Lemma 4.6.3 *There exists a constant $M_5 > 0$ independent of $k \geq 1$ such that, for all $k \geq \Omega_2$, I have $\mathbf{E}^{\pi}_{\lambda^*}[k - \tau|\mathcal{E}] \leq M_5(\epsilon(L_0)^2 k + \epsilon(L_0)^{-1} \log k + \epsilon(L_0)^{-2})$.*

Lemma 4.6.2 essentially says that, when k is sufficiently large, everything behaves "nicely" before the stopping time τ on $\mathcal{E}$. As will be clear in Part 3, this enables me to explicitly characterize the cumulative revenue loss under NSC *before* τ. After τ, NSC may end up charging the turn-off prices (i.e., due to stock-out) and the characterization of p_t becomes less tractable. Fortunately, Lemma 4.6.3 implies that $\mathbf{E}^{\pi}_{\lambda^*}[k - \tau]$ is small for large k (i.e., τ is large). So, by regularity condition R3, I can simply bound the per period revenue loss after τ with $\bar{r}$.

The complete proof of Lemma 4.6.3 is deferred to Section 4.6.3. For now, I provide the main intuition and highlight how my argument differs from that in Jasin (2014). Note that, since $\mathbf{E}^{\pi}_{\lambda^*}[k - \tau|\mathcal{E}] = \sum_{t=1}^{k-1} \mathbb{P}^{\pi}_{\lambda^*}(\tau \leq t|\mathcal{E})$, the proof of Lemma 4.6.3 boils down to computing a bound (for each t) for the conditional probability $\mathbb{P}^{\pi}_{\lambda^*}(\tau \leq t|\mathcal{E})$. Roughly speaking, this is equivalent to analyzing the probability that $S(s)$ is smaller than the threshold ψ for $\tilde{L}_0 + 1 \leq s \leq t$. Note that $S(t)$ can be bounded as follows:

$$S(t) \leq \underbrace{\left|\left| \sum_{s=\tilde{L}_0+1}^{t} \frac{\Delta_s}{k-s} \right|\right|_2}_{\text{random noise}} + \underbrace{\left|\left| \sum_{s=\tilde{L}_0+1}^{t} \frac{\lambda^*(p_s) - \lambda(p_s; \hat{\theta}_\iota)}{k-s} \right|\right|_2}_{\text{systematic biases}} + \frac{1}{k-t},$$

where the *random noise* comes from the stochasticity of demand and the *systematic bias* comes from the estimation error due to Spline Estimation and demand linearization. The systematic biases term does not appear in Jasin (2014); here, it is the primary driving force of the order of $\mathbf{E}^{\pi}_{\lambda^*}[k-\tau|\mathcal{E}]$. In the proof, I use Markov's inequality and integration inequality to bound that term. Note that in order to derive a tight bound using Markov's inequality, I need to make sure that the order of $||\lambda^*(p_s) - \lambda(p_s; \hat{\theta}_\iota)||_2$ is small enough. It turns out that, for all $s < \tau$, the definitions of τ and ψ ensure that p_s is very close to p^D; moreover, since $\lambda(.; \theta^*_\iota)$ is a good approximation of $\lambda^*(.)$ in the neighborhood of p^D, $\lambda(p_s; \theta^*_\iota)$ is very close to $\lambda^*(p_s)$ as well. This observation together with Lemma 4.6.1 further implies that, conditioning on $\mathcal{E}$, for all $s < \tau$, the order of $||\lambda^*(p_s) - \lambda(p_s; \hat{\theta}_\iota)||_2 = \mathcal{O}(\epsilon(L_0))$ (see derivation in (4.44) for more details) is sufficiently small. However, for $s \geq \tau$, p_s is not guaranteed to be sufficiently close to p^D and the order of $||\lambda^*(p_s) - \lambda(p_s; \hat{\theta}_\iota)||_2$ could be as large as

$\Theta(1)$, which will blow up the Markov's bound I derive. (Although the spline estimate $\tilde{\lambda}(.)$ is *uniformly* close to $\lambda^*(.)$, its linear approximation $\lambda(.;\hat{\theta}_\iota)$ is not always close to $\lambda^*(.)$, except for prices that are sufficiently close to p^D, see (4.44).) This means that I cannot use Markov's inequality directly on τ as it is defined in (†). The culprit here is the term $S(t)$ which, by definition of $\tilde{\Delta}_s$, includes the summation of many systematic biases terms that may turn out to be very large. To address this, I introduce another stopping time $\tilde{\tau}$ as the minimum of k and the first time $t \geq \tilde{L}_0 + 1$ such that the following condition (††) is violated:

$$(\dagger\dagger) \qquad \psi > \tilde{S}(t), \quad \text{where } \tilde{S}(k) := \infty \text{ and } \forall t \in \overline{[1, k-1]},$$
$$\tilde{S}(t) := \left|\left| \sum_{s=\tilde{L}_0+1}^{t} \frac{\Delta_s}{k-s} + \sum_{s=\tilde{L}_0+1}^{t} \frac{\lambda^*(p_s) - \lambda(p_s;\hat{\theta}_\iota)}{k-s} \mathbf{1}_{\{s \leq \tau\}} \right|\right|_2 + \frac{1}{k-t}.$$

I prove in Lemma 4.6.3 (see Section 4.6.3) that $\tilde{\tau}$ actually equals τ on every sample path, but $\tilde{\tau}$ is easier to work with because the term $\tilde{S}(t)$ in the stopping criterion only involves one systematic bias term that may be large (i.e., $(\lambda^*(p_\tau) - \lambda(p_\tau;\hat{\theta}_\iota))/(k-\tau)$). The desired result can then be attained by Markov's inequality and integration inequality.

Part 3

Finally, I analyze the revenue loss of NSC as a function of k. Here, I collect the results from Parts 1 and 2 and use standard arguments to "count" the revenue loss incurred throughout the selling season (see, for example, Jasin (2014)). If $k = \mathcal{O}(1)$, the revenue loss can be bounded by a constant; if k is large, all the useful properties of τ and $\mathcal{E}$ derived above (Lemmas 4.6.1 - 4.6.3) hold and I can use them to analyze the revenue loss of NSC. I break down the revenue loss of NSC into three parts: (i) revenue loss incurred during the exploration stage, (ii) revenue loss incurred during the exploitation stage before τ, and (iii) revenue loss incurred during the exploitation stage after τ. Since the length of the exploration stage is $\tilde{L}_0$, by R3, I can bound (i) with $\tilde{L}_0\bar{r}$. As for (ii) and (iii), I derive an upper bound by conditioning on $\mathcal{E}$ and $\mathcal{E}^c$. Since $\mathbb{P}^\pi_{\lambda^*}(\mathcal{E}^c)$ is very small for large k, the majority of the revenue loss comes from the expected revenue loss conditioning on $\mathcal{E}$. This means that, roughly speaking, (iii) can be bounded by $\bar{r}\,\mathbf{E}^\pi_{\lambda^*}[k-\tau|\mathcal{E}]$. The remaining work is to carefully bound (ii) conditioning on $\mathcal{E}$ using Taylor's expansion.

Let $\Omega := \max\{\Omega_1, \Omega_2, \Omega_3\}$, where Ω_1 is as defined in Part 1, Ω_2 is as defined in Lemma 4.6.2, and Ω_3 is a constant independent of k such that $\epsilon(L_0) < 1$ for all $k \geq \Omega_3$. If $k < \Omega$, $\rho^\pi(k) < \bar{r}\Omega = \mathcal{O}(1)$. So, I can focus on the case $k \geq \Omega$. Let R^π_t denote the revenue earned in period t under policy π, and let $\hat{R}^\pi_{\lambda^*}(k) := \sum_{t=\tilde{L}_0+1}^{k} R^\pi_t$ denote the revenue earned during the exploitation stage. For notational brevity, I will simply write $\lambda_t = \lambda^*(p_t)$. Let $\bar{\Delta}_t := R^\pi_t - r^*(\lambda_t)$. Note that $\{\bar{\Delta}_t\}_{t=\tilde{L}_0+1}^{k-1}$ is a martingale difference sequence with respect to filtration $\{\mathcal{H}_t\}_{t=\tilde{L}_0+1}^{k-1}$. Thus, by R3 and Optional Stopping Time Theorem, I have $-\mathbf{E}_{\lambda^*}^{\pi}[\sum_{t=\tilde{L}_0+1}^{\tau-1}\bar{\Delta}_t] = -\mathbf{E}_{\lambda^*}^{\pi}[\sum_{t=\tilde{L}_0+1}^{\tau}\bar{\Delta}_t] + \mathbf{E}_{\lambda^*}^{\pi}[\bar{\Delta}_\tau] = \mathbf{E}_{\lambda^*}^{\pi}[\bar{\Delta}_\tau] \le \bar{r}$. Therefore, for $k \ge \Omega$, $\rho^\pi(k)$ can be bounded as follows:

$$
\begin{aligned}
\rho^\pi(k) &= \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=1}^{\tilde{L}_0}\left(r^D - R_t^\pi\right) + \sum_{t=\tilde{L}_0+1}^{\tau-1}\left(r^D - r_\lambda^*(\lambda_t) - \bar{\Delta}_t\right) + \sum_{t=\tau}^{k}\left(r^D - R_t^\pi\right)\right] \\
&\le \bar{r}\tilde{L}_0 + \bar{r} + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\left(r^D - r_\lambda^*(\lambda_t)\right) + \sum_{t=\tau}^{k}\left(r^D - R_t^\pi\right)\right] \\
&= \bar{r}(1+\tilde{L}_0) + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\left(r^D - r_\lambda^*(\lambda_t)\right) + \sum_{t=\tau}^{k}\left(r^D - R_t^\pi\right)\,\middle|\,\mathcal{E}^c\right]\mathbb{P}_{\lambda^*}^{\pi}(\mathcal{E}^c) \\
&\qquad + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\left(r^D - r_\lambda^*(\lambda_t)\right) + \sum_{t=\tau}^{k}\left(r^D - R_t^\pi\right)\,\middle|\,\mathcal{E}\right]\mathbb{P}_{\lambda^*}^{\pi}(\mathcal{E}) \\
&\le \bar{r}(1+\tilde{L}_0) + \bar{r}k\mathbb{P}_{\lambda^*}^{\pi}(\mathcal{E}^c) + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\left(r^D - r_\lambda^*(\lambda_t)\right) + \sum_{t=\tau}^{k}\left(r^D - R_t^\pi\right)\,\middle|\,\mathcal{E}\right] \\
&\le \bar{r}(1+\tilde{L}_0) + \bar{r}k\mathbb{P}_{\lambda^*}^{\pi}(\mathcal{E}^c) + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\left(r^D - r_\lambda^*(\lambda_t)\right)\,\middle|\,\mathcal{E}\right] + \bar{r}\,\mathbf{E}_{\lambda^*}^{\pi}[k-\tau+1\,|\,\mathcal{E}] \\
&\le \bar{r}\left(2 + M_3 + \tilde{L}_0 + \mathbf{E}_{\lambda^*}^{\pi}[k-\tau\,|\,\mathcal{E}]\right) + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\left(r^D - r_\lambda^*(\lambda_t)\right)\,\middle|\,\mathcal{E}\right], \qquad (4.10)
\end{aligned}
$$

where the last inequality follows because $k\mathbb{P}_{\lambda^*}^{\pi}(\mathcal{E}^c) \le M_3$ by Lemma 4.6.1. To bound the second term in (4.10), I use Taylor's expansion. Note that, by R3,

$$
r^D - r_\lambda^*(\lambda_t) = r_\lambda^*(\lambda^D) - r_\lambda^*(\lambda_t) \le \nabla r_\lambda^*(\lambda^D)\cdot(\lambda^D - \lambda_t) + \frac{\bar{v}}{2}||\lambda^D - \lambda_t||_2^2.
$$

I will show in Section 4.6.4 that there exist constants $M_6, M_7 > 0$ independent of $k \ge \Omega$ such that

$$
\mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\nabla r_\lambda^*(\lambda^D)\cdot(\lambda^D - \lambda_t)\,\middle|\,\mathcal{E}\right] \le M_6\left(1 + \epsilon(L_0)^2 k + \epsilon(L_0)^{-1}\log k + \epsilon(L_0)^{-2}\right); \qquad (4.11)
$$

$$
\frac{\bar{v}}{2}\,\mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}||\lambda^D - \lambda_t||_2^2\,\middle|\,\mathcal{E}\right] \le M_7\left(\log k + \epsilon(L_0)^2 k\right). \qquad (4.12)
$$

Combining (4.5)-(4.6) and (4.10)-(4.12) with Lemma 4.6.3, I conclude that there exist constants $M_8, M_9 > 0$ independent of $k > \Omega$ such that for all $k > \Omega$, I have:

$$\begin{aligned}
\rho^\pi(k) &\leq M_8\left(\epsilon(L_0)^2 k + \epsilon(L_0)^{-1}\log k + \epsilon(L_0)^{-2} + \bar{r}\tilde{L}_0\right) \\
&\leq M_8\left[k\left(\frac{\log^2 k}{k}\right)^{\frac{s\wedge\bar{s}-2}{2s+n-2}} + 2^{\frac{s\wedge\bar{s}-2}{2s+n}}\log k\left(\frac{\sqrt{k}}{\log k}\right)^{\frac{s\wedge\bar{s}-2}{2s+n-2}} + 2^{\frac{2(s\wedge\bar{s}-2)}{2s+n}}\left(\frac{k}{\log^2 k}\right)^{\frac{s\wedge\bar{s}-2}{2s+n-2}}\right] \\
&\quad + M_8\left[\bar{r}s^n(s+1)^n 2^{n+2} k^{\frac{s+n}{2s+n-2}}(\log k)^{\frac{2(s-2)}{2s+n-2}}\right] \leq M_9 k^{\frac{2s-s\wedge\bar{s}+n}{2s+n-2}}\log k.
\end{aligned}$$

Letting $M_1 = M_9 + \bar{r}\Omega$ completes the proof of Theorem 4.5.1.

4.6.2 Part 1: Proof of Lemma 4.6.1

Define $\mathcal{F} := \mathcal{F}_1 \cap \mathcal{F}_2$, where

$$\mathcal{F}_1 := \left\{||p^D - \tilde{p}^D||_2 \leq C_0\epsilon(L_0)\right\},$$

$$\mathcal{F}_2 := \left\{\left|\left|\frac{\partial^r(\lambda_j^*(.)-\tilde{\lambda}_j(.))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right|\right|_\infty < C_1\epsilon(L_0), \forall j \in \overline{[1,n]}, r \in \overline{[0,2]}, r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}, \textstyle\sum_{l=1}^n r_l = r\right\},$$

C_0 is a positive constant to be chosen later and $C_1 := \max\{\Psi_0, \Psi_1, \Psi_2\}$ (recall that Ψ_0, Ψ_1, Ψ_2 are the constants discussed in Lemma 4.4.1). Let $\Phi := \max\{\Phi_1, \Phi_2\}$, where $\Phi_1 > 3, \Phi_2 > 3$ are constants to be chosen later. I first derive an upper bound for $||\theta^* - \hat{\theta}||_2$ conditioning on $\mathcal{F}$ for $k \geq \Phi$. (Unless otherwise noted, in what follows, I will simply assume that $\mathcal{F}$ is satisfied and $k \geq \Phi$.)

By R1 (i.e., Lipschitz continuity of the second order partial derivatives of $\lambda^*(.)$), the compactness of $\mathcal{P}$, and the continuity of $\tilde{\lambda}(.)$ (note that $s \geq 4 > 2$ implies $\tilde{\lambda}(.) \in \mathsf{C}(\mathcal{P})$), there exists a constant $C_2 > 0$ such that, conditioning on $\mathcal{F}$, for all $r \in \overline{[0,2]}$, $r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}$ satisfying $\sum_{l=1}^n r_l = r$, $p \in \mathcal{P}$, and $j \in \overline{[1,n]}$, I have:

$$\left|\frac{\partial^r(\lambda_j^*(p^D) - \lambda_j^*(\tilde{p}^D))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right| \leq C_2||p^D - \tilde{p}^D||_2 \leq C_2 C_0\epsilon(L_0) \quad \text{and} \quad |\tilde{\lambda}_j(p)| \leq C_2. \tag{4.13}$$

So, the following two inequalities hold:

$$\begin{aligned}
||\lambda^*(p^D) - \lambda^*(\tilde{p}^D)||_2 &\leq \sqrt{n}\,||\lambda^*(p^D) - \lambda^*(\tilde{p}^D)||_\infty \\
&= \sqrt{n} \max_{j=1,\dots,n} |\lambda_j^*(p^D) - \lambda_j^*(\tilde{p}^D)| \leq \sqrt{n} C_2 C_0 \epsilon(L_0) \quad \text{and} \qquad (4.14)
\end{aligned}$$

$$\begin{aligned}
||u_{ij}^* - \hat{u}_{ij}||_2 &= \sqrt{\sum_{l=1}^{n} \left| \frac{\partial^2 \lambda_l^*(p^D)}{\partial p_i \partial p_j} - \frac{\partial^2 \lambda_l^*(\tilde{p}^D)}{\partial p_i \partial p_j} + \frac{\partial^2 \lambda_l^*(\tilde{p}^D)}{\partial p_i \partial p_j} - \frac{\partial^2 \tilde{\lambda}_l(\tilde{p}^D)}{\partial p_i \partial p_j} \right|^2} \\
&\leq \sqrt{\sum_{l=1}^{n} 2 \left| \frac{\partial^2 \lambda_l^*(p^D)}{\partial p_i \partial p_j} - \frac{\partial^2 \lambda_l^*(\tilde{p}^D)}{\partial p_i \partial p_j} \right|^2 + \sum_{l=1}^{n} 2 \left|\left| \frac{\partial^2 \lambda_l^*(.)}{\partial p_i \partial p_j} - \frac{\partial^2 \tilde{\lambda}_l(.)}{\partial p_i \partial p_j} \right|\right|_\infty^2} \\
&\leq \sqrt{2nC_0^2 C_2^2 \epsilon(L_0)^2 + 2nC_1^2 \epsilon(L_0)^2} = \epsilon(L_0)\sqrt{2nC_0^2C_2^2 + 2nC_1^2}. \quad (4.15)
\end{aligned}$$

Let $||.||_F$ denote the Frobenius norm. By similar arguments as above, there exists a constant $C_3 > 0$ independent of k such that:

$$\begin{aligned}
||B^* - \hat{B}||_F &= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \left| \frac{\partial \lambda_i^*(p^D)}{\partial p_j} - \frac{\partial \lambda_i^*(\tilde{p}^D)}{\partial p_j} + \frac{\partial \lambda_i^*(\tilde{p}^D)}{\partial p_j} - \frac{\partial \tilde{\lambda}_i(\tilde{p}^D)}{\partial p_j} \right|^2} \\
&\leq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} 2 \left| \frac{\partial \lambda_i^*(p^D)}{\partial p_j} - \frac{\partial \lambda_i^*(\tilde{p}^D)}{\partial p_j} \right|^2 + \sum_{i=1}^{n}\sum_{j=1}^{n} 2 \left|\left| \frac{\partial \lambda_i^*(.)}{\partial p_j} - \frac{\partial \tilde{\lambda}_i(.)}{\partial p_j} \right|\right|_\infty^2} \\
&\leq \sqrt{2n^2C_2^2C_0^2\epsilon(L_0)^2 + 2n^2C_1^2\epsilon(L_0)^2} \leq C_3\epsilon(L_0). \qquad (4.16)
\end{aligned}$$

I now derive a bound for $||H^* - \hat{H}||_2$. To do this, I need to first find a bound for $||\hat{B}^{-1}||_2$. Let $\sigma_{\max}(X)$ and $\sigma_{\min}(X)$ denote the maximum and the minimum eigenvalues of a symmetric real matrix X, respectively. Since $B^* = \nabla\lambda^*(p^D)$ is invertible, $B^*(B^*)'$ is positive definite; so, $\bar{\sigma}^* := \sigma_{\max}(B^*(B^*)') > 0$ and $\underline{\sigma}^* := \sigma_{\min}(B^*(B^*)') > 0$. Moreover, since $C_3\epsilon(L_0) \to 0$ as $k \to \infty$, by (4.16), there exists $\Phi_1 > 0$ such that, for all $k > \Phi_1$, $||B^* - \hat{B}||_2 \leq ||B^* - \hat{B}||_F \leq C_3\epsilon(L_0) \leq \underline{\sigma}^*/(4\sqrt{\bar{\sigma}^*})$. Therefore, for all $v \in \mathbb{R}^n$ with $||v||_2 = 1$,

$$\begin{aligned}
v'\hat{B}'\hat{B}v &= v'(\hat{B} - B^* + B^*)'(\hat{B} - B^* + B^*)v \\
&= v'(B^*)'B^*v + v'(B^*)'(\hat{B} - B^*)v + v'(\hat{B} - B^*)'B^*v + v'(\hat{B} - B^*)'(\hat{B} - B^*)v \\
&\geq \underline{\sigma}^* - 2||v||_2^2||B^*||_2||\hat{B} - B^*||_2 \;\geq\; \underline{\sigma}^* - 2\sqrt{\bar{\sigma}^*}\underline{\sigma}^*/(4\sqrt{\bar{\sigma}^*}) \;=\; \underline{\sigma}^*/2.
\end{aligned}$$

This means that $\sigma_{\min}(\hat{B}'\hat{B}) \geq \underline{\sigma}^*/2 > 0$. Since $(\hat{B}'\hat{B})^{-1} = \hat{B}^{-1}(\hat{B}^{-1})'$,

$$||\hat{B}^{-1}||_2 = \sqrt{\sigma_{\max}(\hat{B}^{-1}(\hat{B}^{-1})')} = \sqrt{\sigma_{\min}(\hat{B}'\hat{B})^{-1}} \leq \sqrt{2/\underline{\sigma}^*}. \tag{4.17}$$

By telescoping, I can bound

$$\begin{aligned} |H^*_{ij} - \hat{H}_{ij}| &= |(u^*_{ij})'(B^*)^{-1}\lambda^*(p^D) - \hat{u}'_{ij}\hat{B}^{-1}\tilde{\lambda}(\tilde{p}^D)| \\ &\leq |(u^*_{ij})'(B^*)^{-1}\lambda^*(p^D) - (u^*_{ij})'(B^*)^{-1}\tilde{\lambda}(\tilde{p}^D)| \\ &\quad + |(u^*_{ij})'(B^*)^{-1}\tilde{\lambda}(\tilde{p}^D) - (u^*_{ij})'\hat{B}^{-1}\tilde{\lambda}(\tilde{p}^D)| + |(u^*_{ij})'\hat{B}^{-1}\tilde{\lambda}(\tilde{p}^D) - \hat{u}'_{ij}\hat{B}^{-1}\tilde{\lambda}(\tilde{p}^D)| \\ &\leq ||u^*_{ij}||_2\,||(B^*)^{-1}||_2\,||\lambda^*(p^D) - \tilde{\lambda}(\tilde{p}^D)||_2 \\ &\quad + ||u^*_{ij}||_2\,||(B^*)^{-1}||_2\,||B^* - \hat{B}||_2\,||\hat{B}^{-1}||_2\,||\tilde{\lambda}(\tilde{p}^D)||_2 \\ &\quad + ||u^*_{ij} - \hat{u}_{ij}||_2\,||\hat{B}^{-1}||_2\,||\tilde{\lambda}(\tilde{p}^D)||_2 \end{aligned} \tag{4.18}$$

where the last inequality follows because $(B^*)^{-1} - \hat{B}^{-1} = (B^*)^{-1}(\hat{B} - B^*)\hat{B}^{-1}$. I now bound the three terms on the right hand side of (4.18) one by one. For the first term of (4.18), by (4.14) and the definition of $\mathcal{F}_2$, I have:

$$\begin{aligned} &||u^*_{ij}||_2||(B^*)^{-1}||_2||\lambda^*(p^D) - \tilde{\lambda}(\tilde{p}^D)||_2 \\ &\leq ||u^*_{ij}||_2||(B^*)^{-1}||_2(||\lambda^*(p^D) - \lambda^*(\tilde{p}^D)||_2 + ||\lambda^*(\tilde{p}^D) - \tilde{\lambda}(\tilde{p}^D)||_2) \\ &\leq ||u^*_{ij}||_2||(B^*)^{-1}||_2(\sqrt{n}C_2C_0\epsilon(L_0) + \sqrt{n}\max_{j=1,\ldots,n}\{||\lambda^*_j(.) - \tilde{\lambda}_j(.)||_\infty\}) \\ &\leq ||u^*_{ij}||_2||(B^*)^{-1}||_2(\sqrt{n}C_2C_0\epsilon(L_0) + \sqrt{n}C_1\epsilon(L_0)) = \mathcal{O}(\epsilon(L_0)). \end{aligned}$$

For the second term of (4.18), by (4.13), (4.16) and (4.17),

$$\begin{aligned} &||u^*_{ij}||_2||(B^*)^{-1}||_2||B^* - \hat{B}||_2||\hat{B}^{-1}||_2||\tilde{\lambda}(\tilde{p}^D)||_2 \\ &\leq ||u^*_{ij}||_2||(B^*)^{-1}||_2C_3\epsilon(L_0)\sqrt{2/\underline{\sigma}^*}\sqrt{n}C_2 = \mathcal{O}(\epsilon(L_0)). \end{aligned}$$

For the last term of (4.18), by (4.13), (4.15) and (4.17),

$$||u^*_{ij} - \hat{u}_{ij}||_2||\hat{B}^{-1}||_2||\tilde{\lambda}(\tilde{p}^D)||_2 \leq \epsilon(L_0)\sqrt{2nC_0^2C_2^2 + 2nC_1^2}\sqrt{2/\underline{\sigma}^*}\sqrt{n}C_2 = \mathcal{O}(\epsilon(L_0)).$$

So, $|H^*_{ij} - \hat{H}_{ij}| = \mathcal{O}(\epsilon(L_0))$ and there exists a constant $C_4 > 0$ independent of $k \geq 1$ such that

$$||H^* - \hat{H}||_2 \leq ||H^* - \hat{H}||_F = \sqrt{\textstyle\sum_{i=1}^n \sum_{j=1}^n |H^*_{ij} - \hat{H}_{ij}|^2} \leq C_4\epsilon(L_0). \tag{4.19}$$

Using similar arguments as above, by telescoping and (4.14) - (4.19), there exist constants $C_5, C_6, C_7, C_8 > 0$ independent of $k \geq 1$ such that on $\mathcal{F}$ I have

$$\begin{aligned}
||a^* - \hat{a}||_2 &\leq ||\lambda^*(p^D) - \tilde{\lambda}(\tilde{p}^D)||_2 + ||(B^*)'p^D - \hat{B}'\tilde{p}^D||_2 \\
&\leq ||\lambda^*(p^D) - \lambda^*(\tilde{p}^D)||_2 + ||\lambda^*(\tilde{p}^D) - \tilde{\lambda}(\tilde{p}^D)||_2 \\
&\qquad + ||B^*||_2||p^D - \tilde{p}^D||_2 + ||\tilde{p}^D||_2||B^* - \hat{B}||_2 \\
&\leq \sqrt{n}C_2C_0\epsilon(L_0) + \sqrt{n}C_1\epsilon(L_0) + ||B^*||_2C_0\epsilon(L_0) + (\textstyle\sum_{l=1}^n \bar{p}_l^2)^{1/2}C_3\epsilon(L_0) \\
&= C_5\epsilon(L_0) \qquad (4.20)
\end{aligned}$$

$$\begin{aligned}
||E^* - \hat{E}||_2 &\leq ||\tfrac{1}{2}(p^D - \tilde{p}^D)'H^*p^D||_2 + ||\tfrac{1}{2}(\tilde{p}^D)'H^*(p^D - \tilde{p}^D)||_2 + ||\tfrac{1}{2}(\tilde{p}^D)'(H^* - \hat{H})\tilde{p}^D||_2 \\
&\leq (\textstyle\sum_{l=1}^n \bar{p}_l^2)^{1/2}||H^*||_2C_0\epsilon(L_0) + \tfrac{1}{2}(\textstyle\sum_{l=1}^n \bar{p}_l^2)C_4\epsilon(L_0) = C_6\epsilon(L_0), \qquad (4.21)
\end{aligned}$$

$$\begin{aligned}
||F^* - \hat{F}||_2 &\leq ||a^* - \hat{a}||_2 + ||H^*||_2||p^D - \tilde{p}^D||_2 + ||H^* - \hat{H}||_2||\tilde{p}^D||_2 \\
&\leq C_5\epsilon(L_0) + ||H^*||_2C_0\epsilon(L_0) + (\textstyle\sum_{l=1}^n \bar{p}_l^2)^{1/2}C_4\epsilon(L_0) = C_7\epsilon(L_0), \qquad (4.22)
\end{aligned}$$

$$||G^* - \hat{G}||_F \leq 2||B^* - \hat{B}||_F + ||H^* - \hat{H}||_F \leq 2C_3\epsilon(L_0) + C_4\epsilon(L_0) = C_8\epsilon(L_0). \qquad (4.23)$$

Putting (4.16) and (4.20) - (4.23) together, for all $k \geq \Phi$, I obtain:

$$\begin{aligned}
||\theta^* - \hat{\theta}||_2 &\leq ||a^* - \hat{a}||_2 + ||B^* - \hat{B}||_F + ||E^* - \hat{E}||_2 + ||F^* - \hat{F}||_2 + ||G^* - \hat{G}||_F \\
&\leq C_9\epsilon(L_0) \qquad (4.24)
\end{aligned}$$

where $C_9 := C_3 + C_5 + C_6 + C_7 + C_8$. Let $M_2 = C_9 + 1$. Since $\mathcal{F}$ implies $\mathcal{E}$ (i.e., because $||\theta^* - \hat{\theta}||_2 < M_2\epsilon(L_0)$ on $\mathcal{F}$), I can bound:

$$\mathbb{P}^\pi_{\lambda^*}(\mathcal{E}^c) \leq \mathbb{P}^\pi_{\lambda^*}(\mathcal{F}^c) \leq \mathbb{P}^\pi_{\lambda^*}(\mathcal{F}_1^c) + \mathbb{P}^\pi_{\lambda^*}(\mathcal{F}_2^c). \qquad (4.25)$$

I will now bound each $\mathbb{P}^\pi_{\lambda^*}(\mathcal{F}_1^c)$ and $\mathbb{P}^\pi_{\lambda^*}(\mathcal{F}_2^c)$ separately. Note that

$$\begin{aligned}
\mathbb{P}^\pi_{\lambda^*}(\mathcal{F}_2^c) &\leq \sum_{j=1}^n \sum_{r=0}^2 \sum_{r_l \geq 0, \sum_{l=1}^n r_l = r} \mathbb{P}^\pi_{\lambda^*}\left(\left\| \frac{\partial^r(\lambda_j^*(.) - \tilde{\lambda}_j(.))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}} \right\|_\infty \geq C_1\epsilon(L_0) \right) \\
&\leq \sum_{j=1}^n \sum_{r=0}^2 \sum_{r_l \geq 0, \sum_{l=1}^n r_l = r} \mathbb{P}^\pi_{\lambda^*}\left(\left\| \frac{\partial^r(\lambda_j^*(.) - \tilde{\lambda}_j(.))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}} \right\|_\infty \geq \Psi_r \left(\frac{\log k}{\sqrt{L_0}} \right)^{\frac{s \wedge \bar{s} - r}{s + n/2}} \right) \\
&\leq \frac{n(n+1)(n+2)}{2} \frac{K}{k}, \qquad (4.26)
\end{aligned}$$

where the first inequality follows by the definition of $\mathcal{F}_2$ and the union bound, the second inequality follows by the definition of C_1 and $\epsilon(L_0)$, and the last inequality follows by Lemma 4.4.1. To compute a bound for $\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_1^c)$, first note that

$$\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_1^c) = \mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_1^c|\mathcal{F}_2)\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_2) + \mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_1^c|\mathcal{F}_2^c)\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_2^c) \leq \mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_1^c \cap \mathcal{F}_2) + \mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_2^c). \quad (4.27)$$

So, it suffices that I find a bound for $\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_1^c \cap \mathcal{F}_2)$. By Lemma 4.4.2, there exist constants $\bar{\delta}, C_{10} > 0$ independent of $k \geq 1$ such that if $||\lambda^*(.) - \tilde{\lambda}(.)||_\infty \leq \bar{\delta}$, then

$$\begin{aligned}
||p^D - \tilde{p}^D||_2 &\leq C_{10}(||\lambda^*(.) - \tilde{\lambda}(.)||_\infty + ||(\nabla\lambda^*(.) - \nabla\tilde{\lambda}(.))'||_\infty) \\
&= C_{10} \sup_{p\in\mathcal{P}} ||\lambda^*(p) - \tilde{\lambda}(p)||_\infty + C_{10} \sup_{p\in\mathcal{P}} ||(\nabla\lambda^*(p) - \nabla\tilde{\lambda}(p))'||_\infty \\
&= C_{10} \sup_{p\in\mathcal{P}} \max_{j=1,\ldots,n} |\lambda_j^*(p) - \tilde{\lambda}_j(p)| + C_{10} \sup_{p\in\mathcal{P}} \max_{i=1,\ldots,n} \sum_{j=1}^n \left| \frac{\partial\lambda_i^*(p)}{\partial p_j} - \frac{\partial\tilde{\lambda}_i(p)}{\partial p_j} \right| \\
&\leq C_{10} \max_{j=1,\ldots,n} ||\lambda_j^*(.) - \tilde{\lambda}_j(.)||_\infty + C_{10} \max_{i=1,\ldots,n} \sum_{j=1}^n \left|\left| \frac{\partial\lambda_i^*(.)}{\partial p_j} - \frac{\partial\tilde{\lambda}_i(.)}{\partial p_j} \right|\right|_\infty . \\
&\leq C_{10} \max_{j=1,\ldots,n} ||\lambda_j^*(.) - \tilde{\lambda}_j(.)||_\infty + C_{10} \sum_{j=1}^n \max_{i=1,\ldots,n} \left|\left| \frac{\partial\lambda_i^*(.)}{\partial p_j} - \frac{\partial\tilde{\lambda}_i(.)}{\partial p_j} \right|\right|_\infty . \quad (4.28)
\end{aligned}$$

Since $C_1\epsilon(L_0) \to 0$ as $k \to \infty$, there exists a constant $\Phi_2 > 0$ such that, conditioning on $\mathcal{F}_2$, for all $k \geq \Phi \geq \Phi_2$, $||\lambda^*(.) - \tilde{\lambda}(.)||_\infty = \max_{j=1,\ldots,n} ||\lambda_j^*(.) - \tilde{\lambda}_j(.)||_\infty \leq C_1\epsilon(L_0) \leq \bar{\delta}$; so, (4.28) holds. Let $C_0 = C_{10}C_1(1+n)$. Then, for all $k \geq \Phi$,

$$\begin{aligned}
&\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{F}_1^c \cap \mathcal{F}_2) = \mathbb{P}^{\pi}_{\lambda^*}(\{||p^D - \tilde{p}^D||_2 > C_0\epsilon(L_0)\} \cap \mathcal{F}_2) \\
&\leq \mathbb{P}^{\pi}_{\lambda^*}\left(C_{10} \max_{j=1,\ldots,n} \left\{ ||\lambda_j^*(.) - \tilde{\lambda}_j(.)||_\infty \right\} + C_{10} \sum_{j=1}^n \max_{i=1,\ldots,n} \left\{ \left|\left| \frac{\partial\lambda_i^*(.)}{\partial p_j} - \frac{\partial\tilde{\lambda}_i(.)}{\partial p_j} \right|\right|_\infty \right\} > C_{10}C_1(1+n)\epsilon(L_0) \right) \\
&\leq \mathbb{P}^{\pi}_{\lambda^*}\left(\max_{j=1,\ldots,n} \left\{ ||\lambda_j^*(.) - \tilde{\lambda}_j(.)||_\infty \right\} > C_1\epsilon(L_0) \right) + \sum_{j=1}^n \mathbb{P}^{\pi}_{\lambda^*}\left(\max_{i=1,\ldots,n} \left\{ \left|\left| \frac{\partial\lambda_i^*(.)}{\partial p_j} - \frac{\partial\tilde{\lambda}_i(.)}{\partial p_j} \right|\right|_\infty > C_1\epsilon(L_0) \right\} \right) \\
&\leq \sum_{j=1}^n \mathbb{P}^{\pi}_{\lambda^*}\left(||\lambda_j^*(.) - \tilde{\lambda}_j(.)||_\infty > \Psi_0 \left(\frac{\log k}{\sqrt{L_0}} \right)^{\frac{s\wedge\bar{s}}{s+n/2}} \right) + \sum_{j=1}^n \sum_{i=1}^n \mathbb{P}^{\pi}_{\lambda^*}\left(\left|\left| \frac{\partial\lambda_i^*(.)}{\partial p_j} - \frac{\partial\tilde{\lambda}_i(.)}{\partial p_j} \right|\right|_\infty > \Psi_1 \left(\frac{\log k}{\sqrt{L_0}} \right)^{\frac{s\wedge\bar{s}-1}{s+n/2}} \right) \\
&\leq n(n+1)K/k,
\end{aligned}$$

where the first inequality follows from (4.28), the third inequality follows since $\log k/\sqrt{L_0} \leq 1$ for $L_0 \geq \log^3 k$ and $k \geq \Phi \geq 3$, the last inequality follows by Lemma 4.4.1.

Let $M_3 = \max\{\Phi, n(n+1)(n+3)K\}$. Putting all the bounds together, for $k \geq \Phi$, I have

$$\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{E}^c) \leq n(n+1)K/k + n(n+1)(n+2)K/k = n(n+1)(n+3)K/k \leq M_3/k.$$

As for $k \leq \Phi$, by definition of M_3, $\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{E}^c) \leq 1 \leq \Phi/k \leq M_3/k$. This completes the proof.

4.6.3 Part 2: Proofs of Lemma 4.6.2 and Lemma 4.6.3

Let $\lambda_t := \lambda^*(p_t)$ and $\hat{\lambda}_t := \lambda(p_t; \hat{\theta}_\iota)$. I prove Lemmas 4.6.2 and 4.6.3 in turn.

Proof of Lemma 4.6.2: Let $\Omega_2 = \max\{\Omega_1, K_1, K_2, K_3\}$, where Ω_1 is as defined in the last paragraph of Part 1 in Section 4.6.1 and K_1, K_2, K_3 are positive constants to be defined later. Throughout the proof, I will implicitly assume that $\mathcal{E}$ is satisfied. I first highlight four inequalities (see (4.29) - (4.32)) that will be useful for the proof. Let $\bar{\delta}, \kappa, \omega$ be as defined in Lemma 4.4.3 and ϕ as defined in R4. First, recall that $||\hat{\theta}_\iota - \theta^*_\iota||_2 \leq ||\hat{\theta} - \theta^*||_2 \leq \bar{\delta}$ and $||\hat{\theta}_o - \theta^*_o||_2 \leq ||\hat{\theta} - \theta^*||_2 \leq \bar{\delta}$ by (4.8), and $||\delta||_2 \leq \bar{\delta}$ by (4.9). Thus, in light of part (f) of Lemma 4.4.3,

$$||p^D_{\mathbf{0}}(\theta^*) - p^D_{\delta}(\hat{\theta})||_2 \leq \phi/2. \tag{4.29}$$

Second, since $\omega\psi = \omega\sqrt{\epsilon(L_0)} \to 0$ as $k \to \infty$ (recall that $\log k/\sqrt{L_0} \leq 1$ for $k \geq 3$ since $L_0 \geq \log^3 k$), there exists a constant $K_1 > 0$ such that, for all $k \geq K_1$,

$$\omega\psi \leq \phi/4. \tag{4.30}$$

Third, note that R4 implies $\lambda^D \succ \lambda_L \mathbf{e} \succ 0$ for some $\lambda_L \in \mathbb{R}$. Hence, there exits a constant $K_2 > 0$ such that, for all $k \geq K_2$, $\psi \leq \lambda_L$ (because $\psi := \sqrt{\epsilon(L_0)} \to 0$ as $k \to \infty$) and

$$\begin{aligned} \lambda^D_\delta(\hat{\theta}) &= \lambda^D - \lambda^D_{\mathbf{0}}(\theta^*) + \lambda^D_\delta(\hat{\theta}) \succeq \lambda^D - ||\lambda^D_\delta(\hat{\theta}) - \lambda^D_{\mathbf{0}}(\theta^*)||_2\mathbf{e} \\ &\succeq \lambda^D - \kappa(||\theta^* - \hat{\theta}||_2 + ||\delta||_2)\mathbf{e} \\ &\succeq \lambda^D - \kappa M_2 \epsilon(L_0)\mathbf{e} - \kappa M_4 \epsilon(L_0)^2\mathbf{e} \succeq \lambda_L \mathbf{e}, \end{aligned} \tag{4.31}$$

where the first equality and the second inequality follow by part (e) and (f) of Lemma 4.4.3 respectively, and the last inequality follows by (4.9) and the definition of $\mathcal{E}$.

Finally, $\tilde{L_0}/k \to 0$ as $k \to \infty$. So, there exists a constant $K_3 > 0$ such that, for all

$k \geq K_3$,

$$C_{\tilde{L}_0} \;\succeq\; kC - \tilde{L}_0 A\mathbf{e} \;\succeq\; A\mathbf{e}. \tag{4.32}$$

The rest of the proof follows by induction. Fix some $k \geq \Omega_2$. If $\tau \leq \tilde{L}_0 + 1$, there is nothing to prove. Suppose that $\tau > \tilde{L}_0 + 1$. Note that $p_{\tilde{L}_0+1} = \hat{p}_{\tilde{L}_0+1}$ because $C_{\tilde{L}_0} \succeq A\mathbf{e}$ (by (4.32)) and $\hat{p}_{\tilde{L}_0+1} = p^D_\delta(\hat{\theta}) \in \mathsf{Ball}(p^D, \phi/2) \subseteq \mathcal{P}$ (by (4.29)). Hence,

$$\begin{aligned} C_{\tilde{L}_0+1} &= C_{\tilde{L}_0} - AD_{\tilde{L}_0+1} = C_{\tilde{L}_0} - A\lambda^D_\delta(\hat{\theta}) - A\tilde{\Delta}_{\tilde{L}_0+1} \\ &\succeq (k - \tilde{L}_0 - 1)A\left(\lambda^D_\delta(\hat{\theta}) - \tfrac{\tilde{\Delta}_{\tilde{L}_0+1}}{k-\tilde{L}_0-1}\right) \\ &\succeq (k - \tilde{L}_0 - 1)A\mathbf{e}\left(\lambda_L - \left\|\tfrac{\tilde{\Delta}_{\tilde{L}_0+1}}{k-\tilde{L}_0-1}\right\|_2\right) \\ &\succeq (k - \tilde{L}_0 - 1)A\mathbf{e}\left(\psi - \left\|\tfrac{\tilde{\Delta}_{\tilde{L}_0+1}}{k-\tilde{L}_0-1}\right\|_2\right) \succeq A\mathbf{e}, \end{aligned} \tag{4.33}$$

where the first inequality follows since $A\lambda^D_\delta(\hat{\theta}) \preceq kC/k - \delta = C_{\tilde{L}_0}/(k - \tilde{L}_0)$ (recall that $\lambda^D_\delta(\hat{\theta})$ is feasible to $\mathbf{QP}(\hat{\theta}; \delta)$), the second inequality follows by (4.31), the third inequality follows from the fact that $\psi \leq \lambda_L$ for $k \geq \Omega_2 \geq K_2$, and the fourth inequality follows by the definition of τ in (†). Since A is non-negative, $C_{\tilde{L}_0+1} \succeq A\mathbf{e} \succeq A_j$ for all $j \in \overline{[1, n]}$. This is my induction base case. Now, suppose that $C_s \succeq A_j$ and $\hat{p}_s \in \mathcal{P}$ for all $j = 1, \ldots, n$ for all $s \in \overline{[\tilde{L}_0 + 1, t - 1]}$ and $t - 1 < \tau$. If $t \geq \tau$, I have finished the induction. Otherwise,

$$\left\|\hat{p}_t - p^D_\delta(\hat{\theta})\right\|_2 = \left\|(\hat{B}')^{-1} \sum_{s=\tilde{L}_0+1}^{t-1} \frac{\tilde{\Delta}_s}{k - s}\right\|_2 \leq \omega \left\|\sum_{s=\tilde{L}_0+1}^{t-1} \frac{\tilde{\Delta}_s}{k - s}\right\|_2 \leq \omega\psi \leq \frac{\phi}{4}, \tag{4.34}$$

where the first equality follows because by the definition of $p(., \hat{\theta}_\iota)$ I have $\nabla_\lambda p(\lambda; \hat{\theta}_\iota)' = (\hat{B}')^{-1}$ for all λ, the first inequality follows by Lemma 4.4.3 part (a), the second inequality follows by (†), and the last inequality follows by (4.30). Combining (4.29) and (4.34), I have $\hat{p}_t \in$

$\mathsf{Ball}(p^D, 3\phi/4) \subseteq \mathcal{P}$. By the same arguments as in (4.33),

$$\begin{aligned}
C_t &= C_{\tilde{L}_0} - \textstyle\sum_{s=\tilde{L}_0+1}^{t} AD_s = C_{\tilde{L}_0} - \sum_{s=\tilde{L}_0+1}^{t} A(\hat{\lambda}_s + \tilde{\Delta}_s) \\
&= C_{\tilde{L}_0} - \sum_{s=\tilde{L}_0+1}^{t} A\left(\lambda_\delta^D(\hat{\theta}) - \sum_{v=\tilde{L}_0+1}^{s-1} \frac{\tilde{\Delta}_v}{k-v} + \tilde{\Delta}_s\right) \\
&\succeq \sum_{s=t+1}^{k} A\lambda_\delta^D(\hat{\theta}) - \sum_{s=\tilde{L}_0+1}^{t}\left(A\tilde{\Delta}_s - \sum_{v=\tilde{L}_0+1}^{s-1} \tfrac{A\tilde{\Delta}_v}{k-v}\right) \\
&= \sum_{s=t+1}^{k} A\lambda_\delta^D(\hat{\theta}) - \sum_{s=\tilde{L}_0+1}^{t} \frac{(k-t)A\tilde{\Delta}_s}{k-s} \\
&\succeq (k-t)A\mathbf{e}\left(\lambda_L - \left\|\sum_{s=\tilde{L}_0+1}^{t} \frac{\tilde{\Delta}_s}{k-s}\right\|_2\right) \;\succeq\; A\mathbf{e} \;\succeq\; A_j,
\end{aligned}$$

for all $j \in \overline{[1,n]}$. This completes the induction.

Proof of Lemma 4.6.3: Fix $k \geq \Omega_2$. Recall that I have defined at the end of Part 2 in Section 4.6.1 an auxiliary stopping time $\tilde{\tau}$ as the minimum of k and the first time $t \geq \tilde{L}_0 + 1$ such that the following condition (††) is violated:

$$\begin{aligned}
&(\dagger\dagger) \qquad \psi > \tilde{S}(t), \quad \text{where } \tilde{S}(k) := \infty \text{ and } \forall t \in \overline{[1, k-1]}, \\
&\tilde{S}(t) := \left\| \sum_{s=\tilde{L}_0+1}^{t} \frac{\Delta_s}{k-s} + \sum_{s=\tilde{L}_0+1}^{t} \frac{(\lambda_s - \hat{\lambda}_s)}{k-s} \mathbf{1}_{\{s \leq \tau\}} \right\|_2 + \frac{1}{k-t}.
\end{aligned}$$

where ψ is as defined in the definition of τ in (†). I first show that $\tau = \tilde{\tau}$ almost surely. If $\tau = t' < k$, by definition of τ, $\psi \leq S(t')$ and $\psi > S(t)$ for all $t < t'$. So,

$$\begin{aligned}
\tilde{S}(t) &= \left\| \sum_{s=\tilde{L}_0+1}^{t} \frac{\Delta_s}{k-s} + \sum_{s=\tilde{L}_0+1}^{t} \frac{(\lambda_s - \hat{\lambda}_s)}{k-s} \mathbf{1}_{\{s \leq \tau\}} \right\|_2 + \frac{1}{k-t} \\
&= \left\| \sum_{s=\tilde{L}_0+1}^{t} \frac{\Delta_s}{k-s} + \sum_{s=\tilde{L}_0+1}^{t} \frac{(\lambda_s - \hat{\lambda}_s)}{k-s} \right\|_2 + \frac{1}{k-t} = S(t).
\end{aligned}$$

for all $t \leq t'$. But, this implies $\psi \leq \tilde{S}(t')$ and $\psi > \tilde{S}(t)$ for all $t < t'$, which means that $\tilde{\tau} = t'$. If $\tau = k$, immediately I have $\tilde{S}(t) = S(t)$ for all $t < k$. Moreover, since, $\psi > S(t) = \tilde{S}(t)$ for

$t < k$, I must have $\tilde{\tau} = k = \tau$. Define the following two terms:

$$\tilde{S}_r(t) = \left|\left| \sum_{s=\tilde{L}_0+1}^{t} \frac{\Delta_s}{k-s} \right|\right|_2 \quad \text{and} \quad \tilde{S}_s(t) = \sum_{s=\tilde{L}_0+1}^{t} \frac{||\lambda_s - \hat{\lambda}_s||_2}{k-s} \mathbf{1}_{\{s \leq \tau\}}.$$

Since $\tilde{S}(t) \leq \tilde{S}_r(t) + \tilde{S}_s(t) + (k-t)^{-1}$ and $\tilde{\tau}$ is non-negative, I have:

$$\begin{aligned} \mathbf{E}^{\pi}_{\lambda^*}[k-\tau|\mathcal{E}] &= \mathbf{E}^{\pi}_{\lambda^*}[k-\tilde{\tau}|\mathcal{E}] = \sum_{t=1}^{k-1} \mathbb{P}^{\pi}_{\lambda^*}(\tilde{\tau} \leq t|\mathcal{E}) \leq \sum_{t=1}^{k-1} \mathbb{P}^{\pi}_{\lambda^*}\left(\max_{\tilde{L}_0+1 \leq s \leq t} \{\tilde{S}_r(s)\} \geq \frac{\psi}{4} \middle| \mathcal{E} \right) \\ &+ \sum_{t=1}^{k-1} \mathbb{P}^{\pi}_{\lambda^*}\left(\max_{\tilde{L}_0+1 \leq s \leq t} \{\tilde{S}_s(s)\} \geq \frac{\psi}{2} \middle| \mathcal{E} \right) + \sum_{t=1}^{k-1} \mathbb{P}^{\pi}_{\lambda^*}\left(\frac{1}{k-t} \geq \frac{\psi}{4} \middle| \mathcal{E} \right). \end{aligned} \tag{4.35}$$

Note that $\{\tilde{S}_r(t)^2\}_{t=\tilde{L}_0+1}^{k-1}$ is a submartingale with respect to $\{\mathcal{H}_t\}_{t=\tilde{L}_0+1}^{k-1}$. So, I can bound the first term after inequality in (4.35) as follows. For $t \leq k-1$,

$$\begin{aligned} \mathbb{P}^{\pi}_{\lambda^*}\left(\max_{\tilde{L}_0+1 \leq s \leq t} \{\tilde{S}_r(s)\} \geq \frac{\psi}{4} \middle| \mathcal{E} \right) &= \mathbb{P}^{\pi}_{\lambda^*}\left(\max_{\tilde{L}_0+1 \leq s \leq t} \{\tilde{S}_r(s)^2\} \geq \frac{\psi^2}{16} \middle| \mathcal{E} \right) \\ &\leq \frac{1}{\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{E})} \mathbb{P}^{\pi}_{\lambda^*}\left(\max_{\tilde{L}_0+1 \leq s \leq t} \{\tilde{S}_r(s)^2\} \geq \frac{\psi^2}{16} \right) \\ &\leq \frac{32}{\psi^2} \mathbf{E}^{\pi}_{\lambda^*}\left[\tilde{S}_r(t)^2 \right] \leq \frac{32}{\psi^2} \mathbf{E}^{\pi}_{\lambda^*}\left[\sum_{s=1}^{t} \frac{||\Delta_s||_2^2}{(k-s)^2} \right] \\ &\leq \frac{32}{\psi^2} \left[\frac{2}{(k-t)^2} + \frac{2}{k-t} \right] \leq \frac{128}{\psi^2(k-t)}, \end{aligned} \tag{4.36}$$

where the second inequality follows by Doob's submartingale inequality and (4.7), the third inequality follows beacuse $\mathbf{E}^{\pi}_{\lambda^*}[\Delta_s' \Delta_t] = 0$ if $s \neq t$, the fourth inequality follows by integral comparison and the fact that $||\Delta_t||_2^2 \leq 2$, and the last inequality follows because $k - t \geq 1$. Thus, there exists $K_5 > 0$ independent of $k \geq \Omega_2$ such that

$$\sum_{t=1}^{k-1} \mathbb{P}^{\pi}_{\lambda^*}\left(\max_{\tilde{L}_0+1 \leq s \leq t} \{\tilde{S}_r(s)\} \geq \frac{\psi}{4} \middle| \mathcal{E} \right) \leq \frac{128}{\psi^2} \sum_{t=1}^{k-1} \frac{1}{k-t} \leq K_5 \psi^{-2} \log k = K_5 \epsilon(L_0)^{-1} \log k,$$

where the equality follows by the definition of ψ in (†). I now bound the second term in (4.35). By Lemma 4.4.3 part (b) and (4.8), there exists a constant $K_4 > 0$ independent of k

such that, for all $k \geq \Omega_2 \geq \Omega_1$, I have:

$$\begin{aligned}||\lambda_s - \hat{\lambda}_s||_2 &= ||\lambda^*(p_s) - \lambda(p_s;\hat{\theta}_\iota)||_2 \\ &\leq ||\lambda^*(p_s)||_2 + ||\lambda(p_s;\theta^*_\iota)||_2 + ||\lambda(p_s;\theta^*_\iota) - \lambda(p_s;\hat{\theta}_\iota)||_2 \\ &\leq \sqrt{n}||\lambda^*(.)||_\infty + \sqrt{n}||\lambda(.;\theta^*_\iota)||_\infty + \omega\bar{\delta} \;\leq\; K_4, \end{aligned} \tag{4.37}$$

Conditioning on $\mathcal{E}$, for $s < \tau$, I can derive a sharper bound, i.e., $||\lambda_s - \hat{\lambda}_s||_2 \leq \omega_0\epsilon(L_0)$ for some constant ω_0 independent of k (see (4.44) for the derivation). Now, observe that the following holds:

$$\begin{aligned}&\mathbb{P}^\pi_{\lambda^*}\left(\max_{\tilde{L}_0+1\leq s\leq t}\{\tilde{S}_s(s)\} \geq \frac{\psi}{2}\middle|\mathcal{E}\right) = \mathbb{P}^\pi_{\lambda^*}\left(\tilde{S}_s(t) \geq \frac{\psi}{2}\middle|\mathcal{E}\right) \;\leq\; \frac{16}{\psi^4}\mathbf{E}^\pi_{\lambda^*}\left[\tilde{S}_s(t)^4\middle|\mathcal{E}\right] \\ &= \frac{16}{\psi^4}\mathbf{E}^\pi_{\lambda^*}\left[\left(\sum_{s=\tilde{L}_0+1}^{t}\frac{||\lambda_s - \hat{\lambda}_s||_2\mathbf{1}_{\{s<\tau\}}}{k-s} + \sum_{s=\tilde{L}_0+1}^{t}\frac{||\lambda_s - \hat{\lambda}_s||_2\mathbf{1}_{\{s=\tau\}}}{k-s}\right)^4\middle|\mathcal{E}\right] \\ &\leq \frac{128}{\psi^4}\mathbf{E}^\pi_{\lambda^*}\left[\left(\sum_{s=\tilde{L}_0+1}^{t}\frac{||\lambda_s - \hat{\lambda}_s||_2\mathbf{1}_{\{s<\tau\}}}{k-s}\right)^4\middle|\mathcal{E}\right] + \frac{128}{\psi^4}\mathbf{E}^\pi_{\lambda^*}\left[\left(\sum_{s=\tilde{L}_0+1}^{t}\frac{||\lambda_s - \hat{\lambda}_s||_2\mathbf{1}_{\{s=\tau\}}}{k-s}\right)^4\middle|\mathcal{E}\right] \\ &\leq \frac{128\omega_0^4\epsilon(L_0)^4}{\psi^4}\log^4\left(\frac{k}{k-t}\right) + \frac{128}{\psi^4}\left(\frac{K_4}{k-t}\right)^4 \\ &= 128\omega_0^4\epsilon(L_0)^2\log^4\left(\frac{k}{k-t}\right) + 128K_4^4\epsilon(L_0)^{-2}\left(\frac{1}{k-t}\right)^4 \end{aligned} \tag{4.38}$$

where the first equality follows by the monotonicity of $\tilde{S}_s(t)$, the first inequality follows by Markov's inequality, the second inequality follows since $(a+b)^4 \leq 8a^4 + 8b^4$, the third inequality follows by (4.37) and (4.44), and the last equality follows since $\psi = \sqrt{\epsilon(L_0)}$. Note that $\sum_{t=1}^{k-1}\log^s(\frac{k}{k-t}) \leq s!k$; so, there exists a constant $K_6 > 0$ independent of $k \geq \Omega_2$ such that:

$$\begin{aligned}\sum_{t=1}^{k-1}\mathbb{P}^\pi_{\lambda^*}\left(\max_{\tilde{L}_0+1\leq s\leq t}\{\tilde{S}_s(t)\} \geq \frac{\psi}{2}\middle|\mathcal{E}\right) &\leq \sum_{t=1}^{k-1}128\omega_0^4\epsilon(L_0)^2\log^4\left(\frac{k}{k-t}\right) + 128K_4^4\epsilon(L_0)^{-2}\sum_{t=1}^{k-1}\left(\frac{1}{k-t}\right)^4 \\ &\leq K_6\left(\epsilon(L_0)^2k + \epsilon(L_0)^{-2}\right).\end{aligned}$$

The third term of (4.35) can be bounded as follows:

$$\sum_{t=1}^{k-1}\mathbb{P}^\pi_{\lambda^*}\left(\frac{1}{k-t} \geq \frac{\psi}{4}\middle|\,\mathcal{E}\right) \;=\; \sum_{t=1}^{k-1}\mathbb{P}^\pi_{\lambda^*}\left(\frac{1}{k-t} \geq \frac{\psi}{4}\right) \;=\; 4/\psi \;=\; 4\epsilon(L_0)^{-1/2}.$$

Putting the bounds for the three terms in (4.35) together, I conclude that there exists a constant $M_5 > 0$ such that, for all $k \geq \Omega_2$, I have $\mathbf{E}^{\pi}_{\lambda^*}[k - \tau|\mathcal{E}] = \mathbf{E}^{\pi}_{\lambda^*}[k - \tilde{\tau}|\mathcal{E}] \leq K_5\epsilon(L_0)^{-1}\log k + K_6\epsilon(L_0)^2 k + K_6\epsilon(L_0)^{-2} + 4\epsilon(L_0)^{-1/2} \leq M_5(\epsilon(L_0)^2 k + \epsilon(L_0)^{-1}\log k + \epsilon(L_0)^{-2})$.

4.6.4 Part 3: Derivation of (4.11) and (4.12)

For simplicity, I suppress the dependency of $\epsilon(L_0)$ on L_0 and simply write $\epsilon(L_0)$ as ϵ throughout this section. Recall that I define $\lambda_t := \lambda^*(p_t)$ and $\hat{\lambda}_t := \lambda(p_t; \hat{\theta}_\iota)$ in Section 4.6.3. Also, by Lemma 4.6.2, for all $k \geq \Omega_2$ and all sample paths on $\mathcal{E}$,

$$\hat{\lambda}_t \;=\; \lambda(\hat{p}_t; \hat{\theta}_\iota) \;=\; \lambda^D_\delta(\hat{\theta}) - \sum_{s=\tilde{L}_0+1}^{t-1} \frac{\tilde{\Delta}_s}{k-s} \qquad \text{for all } t < \tau. \tag{4.39}$$

These will be used multiple times in the derivation of (4.11) and (4.12).

Derivation of inequality (4.11). Note that $\nabla r^*_\lambda(\lambda^D) = \nabla_\lambda q(p(\lambda^D; \theta^*_\iota); \theta^*_o)$ by (4.3) and $\lambda^D_{\mathbf{0}}(\theta^*) = \lambda^D$ by Lemma 4.4.3 part (e). So, I can write: $\nabla r^*_\lambda(\lambda^D) \cdot (\lambda^D - \lambda_t) = \nabla_\lambda q(p(\lambda^D_{\mathbf{0}}(\theta^*); \theta^*_\iota); \theta^*_o) \cdot (\lambda^D_{\mathbf{0}}(\theta^*) - \lambda_t) = \mu^D_{\mathbf{0}}(\theta^*)' A(\lambda^D_{\mathbf{0}}(\theta^*) - \lambda^D_\delta(\hat{\theta}) + \lambda^D_\delta(\hat{\theta}) - \hat{\lambda}_t + \hat{\lambda}_t - \lambda_t)$, where the last equality follows by the Karush-Kuhn-Tucker(KKT) optimality condition. Therefore, for all $k \geq \Omega$, the first term of (4.11) can be broken into two parts:

$$\begin{aligned}
&\mathbf{E}^{\pi}_{\lambda^*}\left[\textstyle\sum_{t=\tilde{L}_0+1}^{\tau-1} \nabla r^*_\lambda(\lambda^D) \cdot (\lambda^D - \lambda_t) \middle| \mathcal{E}\right] \\
&\quad = \mathbf{E}^{\pi}_{\lambda^*}\left[\textstyle\sum_{t=\tilde{L}_0+1}^{\tau-1} \mu^D_{\mathbf{0}}(\theta^*)'(A\lambda^D_{\mathbf{0}}(\theta^*) - A\lambda^D_\delta(\hat{\theta})) \middle| \mathcal{E}\right] \\
&\qquad + \mathbf{E}^{\pi}_{\lambda^*}\left[\textstyle\sum_{t=\tilde{L}_0+1}^{\tau-1} \mu^D_{\mathbf{0}}(\theta^*)' A(\lambda^D_\delta(\hat{\theta}) - \hat{\lambda}_t + \hat{\lambda}_t - \lambda_t) \middle| \mathcal{E}\right]
\end{aligned} \tag{4.40}$$

By Lemma 4.4.3 part (f), for all sample paths on $\mathcal{E}$, the set of constraints of $\mathbf{QP}(\theta^*; \mathbf{0})$ that have non-zero optimal dual variables are binding at the optimal solution $\lambda^D_\delta(\hat{\theta})$ in $\mathbf{QP}(\hat{\theta}; \delta)$. This implies that the first expectation after the equality in (4.40) is zero because, for all i, either I have $\mu^D_{\mathbf{0},i}(\theta^*) = 0$ or $(A\lambda^D_{\mathbf{0}}(\theta^*))_i - (A\lambda^D_\delta(\hat{\theta}))_i = 0$. As for the second expectation, by (4.39) and the definition of $\tilde{\Delta}_t$ (i.e., $\tilde{\Delta}_t = \Delta_t + \lambda_t - \hat{\lambda}_t$), I can write:

$$\begin{aligned}
&\mathbf{E}^{\pi}_{\lambda^*}\left[\textstyle\sum_{t=\tilde{L}_0+1}^{\tau-1} \mu^D_{\mathbf{0}}(\theta^*)' A(\lambda^D_\delta(\hat{\theta}) - \hat{\lambda}_t + \hat{\lambda}_t - \lambda_t) \middle| \mathcal{E}\right] \\
&\quad = \mathbf{E}^{\pi}_{\lambda^*}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} \mu^D_{\mathbf{0}}(\theta^*)' \left(\sum_{s=\tilde{L}_0+1}^{t-1} \frac{A\tilde{\Delta}_s}{k-s} + A\Delta_t - A\tilde{\Delta}_t\right) \middle| \mathcal{E}\right] \\
&\quad = \mathbf{E}^{\pi}_{\lambda^*}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} \mu^D_{\mathbf{0}}(\theta^*)' A\Delta_t \middle| \mathcal{E}\right] + \mathbf{E}^{\pi}_{\lambda^*}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} \left(\frac{\tau-t-1}{k-t} - 1\right) \mu^D_{\mathbf{0}}(\theta^*)' A\tilde{\Delta}_t \middle| \mathcal{E}\right].
\end{aligned} \tag{4.41}$$

Since $\{\Delta_t\}_{t=\tilde{L}_0+1}^{k-1}$ is a martingale difference sequence with respect to $\{\mathcal{H}_t\}_{t=\tilde{L}_0+1}^{k-1}$,

$$\begin{aligned}
&\mathbf{E}^{\pi}_{\lambda^*}\left[\left.\textstyle\sum_{t=\tilde{L}_0+1}^{\tau-1} \mu_{\mathbf{0}}^D(\theta^*)' A\Delta_t \right| \mathcal{E}\right] \\
&\quad = \frac{\mu_{\mathbf{0}}^D(\theta^*)'A}{\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{E})}\left\{\mathbf{E}^{\pi}_{\lambda^*}\left[\textstyle\sum_{t=\tilde{L}_0+1}^{\tau-1}\Delta_t\right] - \mathbf{E}^{\pi}_{\lambda^*}\left[\left.\textstyle\sum_{t=\tilde{L}_0+1}^{\tau-1}\Delta_t\right|\mathcal{E}^c\right]\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{E}^c)\right\} \\
&\quad \le \mu_{\mathbf{0}}^D(\theta^*)'A\mathbf{e}\,\frac{1+k\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{E}^c)}{\mathbb{P}^{\pi}_{\lambda^*}(\mathcal{E})} \;\le\; 2(1+M_3)\mu_{\mathbf{0}}^D(\theta^*)'A\mathbf{e}, \qquad (4.42)
\end{aligned}$$

where the first inequality follows from $\mathbf{E}^{\pi}_{\lambda^*}[\sum_{t=L+1}^{\tau-1}\Delta_t] = \mathbf{E}^{\pi}_{\lambda^*}[\sum_{t=L+1}^{\tau}\Delta_t] - \mathbf{E}^{\pi}_{\lambda^*}[\Delta_\tau]$ (by Optional Stopping Time Theorem) and the fact that $|\Delta_t| \prec \mathbf{e}$, and the second inequality follows by Lemma 4.6.1 and (4.7). The second term of (4.41) can be bounded as follows:

$$\begin{aligned}
&\mathbf{E}^{\pi}_{\lambda^*}\left[\left.\sum_{t=\tilde{L}_0+1}^{\tau-1}\left(\frac{\tau-t-1}{k-t}-1\right)\mu_{\mathbf{0}}^D(\theta^*)'A\tilde{\Delta}_t\right|\mathcal{E}\right] \\
&\quad \le \mathbf{E}^{\pi}_{\lambda^*}\left[(k-\tau+1)\left|\mu_{\mathbf{0}}^D(\theta^*)'\sum_{t=\tilde{L}_0+1}^{\tau-1}\frac{A\tilde{\Delta}_t}{k-t}\right|\,\middle|\,\mathcal{E}\right] \\
&\quad \le \mathbf{E}^{\pi}_{\lambda^*}\left[(k-\tau+1)\left|\left|\mu_{\mathbf{0}}^D(\theta^*)\right|\right|_2 ||A||_2\left|\left|\sum_{t=\tilde{L}_0+1}^{\tau-1}\frac{\tilde{\Delta}_t}{k-t}\right|\right|_2\,\middle|\,\mathcal{E}\right] \\
&\quad \le \psi\left|\left|\mu_{\mathbf{0}}^D(\theta^*)\right|\right|_2 ||A||_2\left(\mathbf{E}^{\pi}_{\lambda^*}\left[k-\tau|\,\mathcal{E}\right]+1\right) \\
&\quad \le ||\mu_{\mathbf{0}}^D(\theta^*)||_2||A||_2[M_5(\epsilon^2 k+\epsilon^{-1}\log k+\epsilon^{-2})+1]. \qquad (4.43)
\end{aligned}$$

where the third inequality follows by (†), and the fourth inequality follows by Lemma 4.6.3 and the fact that $\psi = \sqrt{\epsilon(L_0)} < 1$ for $k \ge \Omega \ge \Omega_3$. Putting together (4.40) - (4.43) yields

$$\mathbf{E}^{\pi}_{\lambda^*}\left[\left.\textstyle\sum_{t=\tilde{L}_0+1}^{\tau-1}\nabla r^*_{\lambda}(\lambda^D)\cdot(\lambda^D-\lambda_t)\right|\mathcal{E}\right] \;\le\; M_6(1+\epsilon^2 k+\epsilon^{-1}\log k+\epsilon^{-2})$$

where $M_6 = 2(1+M_3)\mu_{\mathbf{0}}^D(\theta^*)'A\mathbf{e} + ||\mu_{\mathbf{0}}^D(\theta^*)||_2||A||_2(1+M_5)$.

Derivation of inequality (4.12). By definition, $\lambda(p_t;\theta^*_\iota) = \lambda^*(p^D) + (\nabla\lambda^*(p^D))'(p_t - p^D)$. Since $\lambda^*(p_t) = \lambda^*(p^D) + \nabla\lambda^*(p^D)'(p_t-p^D) + (p_t-p^D)'\nabla^2\lambda^*(\xi)(p_t-p^D)$ for some $\xi \in \mathcal{P}$ and $\sup_{\hat{\xi}\in\mathcal{P}}||(p_t-p^D)'\nabla^2\lambda^*(\hat{\xi})(p_t-p^D)||_2 \le \kappa_0||p_t-p^D||_2^2$ for some $\kappa_0 > 0$ (by R1 and the compactness of $\mathcal{P}$),

$$||\lambda^*(p_t) - \lambda(p_t;\theta^*_\iota)||_2 \;\le\; \kappa_0||p_t-p^D||_2^2.$$

So, conditioning on $\mathcal{E}$, for all $t < \tau$, I have:

$$\begin{aligned}
||\lambda_t - \hat{\lambda}_t||_2 &= ||\lambda^*(p_t) - \lambda(p_t;\hat{\theta}_\iota)||_2 \\
&\leq ||\lambda^*(p_t) - \lambda(p_t;\theta_\iota^*)||_2 + ||\lambda(p_t;\theta_\iota^*) - \lambda(p_t;\hat{\theta}_\iota)||_2 \\
&\leq \kappa_0||p_t - p^D||_2^2 + \omega M_2\epsilon \\
&= \kappa_0||p_t - p_\delta^D(\hat{\theta}) + p_\delta^D(\hat{\theta}) - p_{\mathbf{0}}^D(\theta^*)||_2^2 + \omega M_2\epsilon \\
&\leq 2\kappa_0||p_t - p_\delta^D(\hat{\theta})||_2^2 + 2\kappa_0||p_\delta^D(\hat{\theta}) - p_{\mathbf{0}}^D(\theta^*)||_2^2 + \omega M_2\epsilon \\
&\leq 2\kappa_0||\hat{p}_t - p_\delta^D(\hat{\theta})||_2^2 + 2\kappa_0\kappa^2(||\delta||_2 + ||\hat{\theta} - \theta^*||_2)^2 + \omega M_2\epsilon \\
&\leq 2\kappa_0\omega^2\psi^2 + 2\kappa_0\kappa^2(M_4\epsilon^2 + M_2\epsilon)^2 + \omega M_2\epsilon \;\leq\; \omega_0\epsilon \qquad (4.44)
\end{aligned}$$

where the second inequality follows by Lemma 4.4.3 part (b) and the definition of $\mathcal{E}$, the fourth inequality follows by Lemma 4.6.2 (i.e., $p_t = \hat{p}_t$ for $t < \tau$) and Lemma 4.4.3 part (f), and the fifth inequality follows by (4.9), (4.34) and the definition of $\mathcal{E}$. Now,

$$\begin{aligned}
&\frac{1}{2}\,\mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} ||\lambda_\delta^D(\hat{\theta}) - \lambda_t||_2^2 \,|\mathcal{E}\right] \\
&\leq \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} ||\lambda_\delta^D(\hat{\theta}) - \hat{\lambda}_t||_2^2 \,|\mathcal{E}\right] + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} ||\hat{\lambda}_t - \lambda_t||_2^2 \,|\mathcal{E}\right] \\
&= \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} ||\sum_{s=\tilde{L}_0+1}^{t-1} \frac{\tilde{\Delta}_s}{k-s}||^2 \,|\mathcal{E}\right] + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} ||\hat{\lambda}_t - \lambda_t||_2^2 \,|\mathcal{E}\right] \\
&\leq 2\,\mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\left(||\sum_{s=\tilde{L}_0+1}^{t-1} \frac{\Delta_s}{k-s}||^2 + \left(\sum_{s=\tilde{L}_0+1}^{t-1} \frac{||\lambda_s - \hat{\lambda}_s||_2}{k-s}\right)^2\right)\Bigg|\,\mathcal{E}\right] \\
&\quad + \mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1} ||\hat{\lambda}_t - \lambda_t||_2^2 \,|\mathcal{E}\right] \\
&\leq \frac{2}{\mathbb{P}_{\lambda^*}^{\pi}(\mathcal{E})}\mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=1}^{\tau-1} ||\sum_{s=1}^{t-1} \frac{\Delta_s}{k-s}||^2\right] + 2\,\mathbf{E}_{\lambda^*}^{\pi}\left[\sum_{t=1}^{\tau-1}\left(\sum_{s=1}^{t-1}\frac{\omega_0\epsilon}{k-s}\right)^2\Bigg|\,\mathcal{E}\right] + \omega_0^2 k\epsilon^2 \\
&\leq 4\sum_{t=1}^{k-1}\sum_{s=1}^{t-1}\frac{\mathbf{E}_{\lambda^*}^{\pi}[||\Delta_s||_2^2]}{(k-s)^2} + 2\omega_0^2\epsilon^2\sum_{t=1}^{k-1}\left(\sum_{s=1}^{t-1}\frac{1}{k-s}\right)^2 + \omega_0^2 k\epsilon^2 \\
&\leq 8\sum_{t=1}^{k-1}\sum_{s=1}^{t-1}\frac{1}{(k-s)^2} + 2\omega_0^2\epsilon^2\sum_{t=1}^{k-1}\log^2\left(\frac{k}{k-t}\right) + \omega_0^2\epsilon^2 k \\
&\leq 8\log k + 5\omega_0^2\epsilon^2 k \;:=\; M_{10}(\log k + \epsilon^2 k), \qquad (4.45)
\end{aligned}$$

where the third inequality follows from (4.44), the fourth inequality follows by (4.7) and the fact that $\mathbf{E}^{\pi}_{\lambda^*}[\Delta'_s\Delta_t] = 0$ if $s \neq t$, and the last two inequalities follows by integral comparison and the fact that $\sum_{t=1}^{k-1}\log^s(\frac{k}{k-t}) \leq \int_1^k \log^s(\frac{k}{k-t})dt \leq s!k$.

The derivation of inequality (4.12) is completed by noting that

$$\begin{aligned}
&\frac{\bar{v}}{2}\mathbf{E}^{\pi}_{\lambda^*}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}\left|\left|\lambda^D_{\mathbf{0}}(\theta^*)-\lambda_t)\right|\right|_2^2\middle|\,\mathcal{E}\right] \\
&\leq \bar{v}\mathbf{E}^{\pi}_{\lambda^*}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}||\lambda^D_{\mathbf{0}}(\theta^*)-\lambda^D_{\delta}(\hat{\theta})||_2^2\middle|\,\mathcal{E}\right]+\bar{v}\mathbf{E}^{\pi}_{\lambda^*}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}||\lambda^D_{\delta}(\hat{\theta})-\lambda_t||_2^2\middle|\,\mathcal{E}\right] \\
&\leq \bar{v}k\mathbf{E}^{\pi}_{\lambda^*}\left[\left(\kappa||\theta^*-\hat{\theta}||_2+\kappa||\delta||_2\right)^2\middle|\,\mathcal{E}\right]+\bar{v}\mathbf{E}^{\pi}_{\lambda^*}\left[\sum_{t=\tilde{L}_0+1}^{\tau-1}||\lambda^D_{\delta}(\hat{\theta})-\lambda_t||_2^2\middle|\,\mathcal{E}\right] \\
&\leq 2\bar{v}k\left(\kappa^2M_2^2\epsilon^2+\kappa^2M_4^2\epsilon^4\right)+2\bar{v}M_{10}(\log k+\epsilon^2k) \\
&\leq 2\bar{v}M_{10}\log k+(2\bar{v}M_{10}+2\bar{v}\kappa^2M_4^2+2\bar{v}\kappa^2M_2^2)\epsilon^2k = M_7(\log k+\epsilon^2k)
\end{aligned}$$

where $M_7 := 4\bar{v}M_{10} + 2\bar{v}\kappa^2M_4^2 + 2\bar{v}\kappa^2M_2^2$. The second inequality follows by Lemma 4.4.3 part (f), the third inequality follows by the definition of $\mathcal{E}$, (4.9) and (4.45), and the fourth inequality follows by the fact that $\epsilon < 1$ for $k \geq \Omega \geq \Omega_3$.

To summarize, the proofs of Lemma 4.6.1 in Section 4.6.2, Lemma 4.6.2 and Lemma 4.6.3 in Section 4.6.3, and the derivation of (4.11) and (4.12) in Section 4.6.4 fill in the gaps in the outline in Section 4.6.1. This completes the proof of Theorem 4.5.1.

4.7 Closing remarks

I study the problem of joint learning and pricing in a general capacitated network RM problem with multiple products and multiple limited resources. I develop a heuristic called NSC that combines Spline Estimation, linear approximation of the estimated demand function, quadratic approximation of the estimated revenue function, and self-adjusting price updates. I show analytically that if the underlying demand function is sufficiently smooth, the revenue loss under NSC is $\mathcal{O}(k^{1/2+\epsilon}\log k)$ for any fixed $\epsilon > 0$. This is the tightest bound of its kind and is very close to the known theoretical lower bound of $\Omega(\sqrt{k})$. My result suggests the applicability of self-adjusting controls in dynamic pricing problems. Moreover, in proving my main result, I derive large deviation bounds for spline estimation and prove a nonparametric stability result of the optimal solution of a constrained optimization problem. These results are of independent interest and are potentially useful for other application areas.

4.8 Tables

Table 4.1: Revenue loss of Algorithm 3 in Besbes and Zeevi (2012) and NSC

	Algorithm 3 in Besbes and Zeevi (2012)			NSC		
k	Revenue loss	Stdev	Relative revenue loss(%)	Revenue loss	Stdev	Relative revenue loss(%)
500	5331	53	52.6	4681	11	46.2
1000	10490	114	51.8	8823	46	43.5
2000	20307	214	50.1	17320	85	42.7
3000	30074	300	49.5	26647	240	43.8
4000	39167	392	48.3	34421	289	42.5
5000	48922	466	48.3	40796	307	40.3
6000	57804	522	47.5	49578	477	40.8
7000	65459	594	46.1	57310	605	40.4
8000	74990	688	46.3	62319	640	38.4
9000	82891	721	45.4	70797	800	38.8
10000	89703	734	44.3	75179	814	37.1
100000	500623	2972	24.7	426665	8173	21.1
1000000	3173343	17856	15.7	1829065	40502	9.0
10000000	20342474	68912	10.0	8105010	26139	4.0

In this numerical example, I set $n = 2, m = 2, A = [1, 1; 0, 2], C = [0.1; 0.1]$. The true demand function is a logit function, and $[\lambda_1(p_1, p_2); \lambda_2(p_1, p_2)] = (1+\exp(0.4-0.015p_1)+\exp(0.8 - 0.02p_2))^{-1}\ [\exp(0.4 - 0.015p_1); \exp(0.8 - 0.02p_2)]$. For ease of performance comparison, I use $s = 4$ for both Algorithm 3 in Besbes and Zeevi (2012) and NSC. I vary k from 500 (a capacity level of 50 for each resource) to 10000000 (a capacity level of 1000000 for each resource) and run 1000 trials for each k. The fourth and the seventh columns correspond to the relative revenue loss for the corresponding heuristic π defined as $\rho^\pi(k)/J^D(k)$.

Appendix

A.1 Proofs of Results in Chapter 2

A.1.1 Proof of Lemma 2.4.1

Throughout, I use superscript j and subscript i to indicate the j^{th} column and the i^{th} row of a matrix respectively. Define $\bar{A} := A\nabla\lambda(p^D)$. By definition, a base must span the resource space which has rank m, so it must contain at least m products. Without loss of generality, suppose that $\mathcal{B} = \{1, 2, ..., m\}$. The matrix $[\bar{A}^1\, \bar{A}^2 \,...\, \bar{A}^m]$ is invertible and I can define its inverse $\bar{U} = [\bar{A}^1\, \bar{A}^2 \,...\, \bar{A}^m]^{-1}$. I now construct an n by m matrix U as follows: $U_i = \bar{U}_i$ for $i = 1, \ldots, m$ and $U_i = 0$ otherwise. Observe that $A\nabla\lambda(p^D)U = \bar{A}U = I$. Let $H = \nabla\lambda(p^D)\, U$. Since only the first m rows of $\nabla p(\lambda^D)H = U$ are non-zeros and $\mathcal{B} = \{1, 2, ..., m\}$, I conclude that H selects $\mathcal{B}$. To show the uniqueness of H, I use contradiction. Suppose not, then I have at least two n by m matrices $H \neq \tilde{H}$ that select $\mathcal{B}$. Let $U = \nabla p(\lambda^D)H, \tilde{U} = \nabla p(\lambda^D)\tilde{H}$. Since $\nabla p(\lambda^D)$ is full rank and $H - \tilde{H} \neq 0$, I conclude that $U - \tilde{U} \neq 0$. Since the last $n - m$ rows of U and $\tilde{U}$ are all zero vectors, I conclude that $U_i \neq \tilde{U}_i$ for some $1 \leq i \leq m$ which contradicts with the uniqueness of the inverse of $\bar{A}$.

A.1.2 Proof of Theorem 2.4.1

The key to the proof lies in the definition of a stopping time $\tau(\theta)$ for each of the θ^{th} problem, which can be roughly interpreted as the time when the first stock-out of any of the resources occurs. I use a martingale argument to derive an upper bound of the expectation of the remaining length of the selling season after $\tau(\theta)$, namely $\mathbf{E}[T(\theta) - \tau(\theta)]$. The main idea of the proof is to consider the revenue loss incurred before and after $\tau(\theta)$ separately. I show that both of them are in the order of $\mathbf{E}[T(\theta) - \tau(\theta)]$. Therefore, our primary task is to obtain an upper bound of $\mathbf{E}[T(\theta) - \tau(\theta)]$. The basic outline of the proof is as follows: (1) I show that, under some conditions, the resource consumption error can be explicitly written as a function of past demand errors, namely Δ_s's; (2) I introduce some technical conditions and define a stopping time $\tau(\theta)$. This can be roughly interpreted as the stock-out time of the first depleted resource. I then compute an upper bound for $\mathbf{E}[T(\theta) - \tau(\theta)]$; (3) I break down our analysis of revenue loss into two parts, before and after period τ. Not surprisingly, the latter is in the order of $\mathbf{E}[T(\theta) - \tau(\theta)]$, thanks to the bounded revenue assumption in (A2). The rest of the proof shows that the revenue loss before $\tau(\theta)$ is also in the order of $\mathbf{E}[T(\theta) - \tau(\theta)]$.

Without loss of generality, assume $T = 1$. Then $T(\theta) = \theta$. For notational clarity, I suppress the dependency on θ whenever there is no confusion. Fix a projection matrix H that selects $\mathcal{B}$. I proceed in several steps.

STEP 0

I present a well-known result in linear algebra without proof. I will use this result several times.

Lemma A.1.1 *For any real symmetric n by n matrix S, there exists an n by n orthonormal matrix $Q \in \mathbb{R}^n \times \mathbb{R}^n$ such that $Q^{-1}SQ = \Lambda$, where $\Lambda = diag(\theta_1, \ldots, \theta_n)$ is a diagonal matrix whose elements are the eigenvalues of the matrix S. In addition, for any vector $v \in \mathbb{R}^n$, I have: $v'Sv \leq \max_{1 \leq i \leq n} |\theta_i| \cdot v'v$.*

STEP 1

In this step I derive an explicit formula for resource consumption error.

Define $\delta_s := A\Delta_s$, $\tilde{\delta}_l^j := A\tilde{\Delta}_l^j$, and $\epsilon_t := \sum_{j=1}^m E^j \nabla p(\lambda^D) H \sum_{l=1}^{k_t^j} \tilde{\delta}_l^j/(\theta - t_l^j + 1)$. (I follow the convention that if the lower limit of a summation is bigger than the higher limit, then the sum is zero.) By Taylor's expansion,

$$\lambda_t = \lambda^D - \nabla\lambda(p^D)\epsilon_t + \frac{1}{2}\epsilon_t'\nabla^2\lambda(\eta_t)\epsilon_t\,, \quad \eta_t \in [p^D, p^D - \epsilon_t], \tag{46}$$

where, by a slight abuse of the notation, I use

$$\epsilon_t'\nabla^2\lambda(\eta_t)\epsilon_t := \begin{bmatrix} \epsilon_t'\nabla^2\lambda_1(\eta_t)\epsilon_t \\ \vdots \\ \epsilon_t'\nabla^2\lambda_n(\eta_t)\epsilon_t \end{bmatrix},$$

and $\nabla^2\lambda_j$ is the Hessian matrix of $\lambda_j(p_t)$. (Formula (46) holds *if* λ_t lies in the interior of Ω_λ. I will address this in STEP 2.) Since H is the projection matrix that selects $\mathcal{B}$. By definition, there exists an invertible m by m matrix M such that $\nabla p(\lambda^D)H = [\,M'\,\mathbf{0}'\,]'$ and $A\nabla\lambda(p^D) = [M^{-1}|\ldots]$, where the latter holds because $A\nabla\lambda(p^D)\,\nabla p(\lambda^D)H = AH = I$. Define M^j to be a square matrix whose j^{th} row is the same as M while the other rows are zeros. By definition, $M^j\tilde{\delta}_l^j = \tilde{\xi}_l^j\mathbf{e}_j$, where $\mathbf{e}_j$ is a column vector with a proper size whose j^{th}

element is one and the others are zeros. I can write ϵ_t as:

$$\epsilon_t = \begin{bmatrix} \sum_{j=1}^{m} \sum_{l=1}^{k_t^j} \frac{M^j \tilde{\delta}_l^j}{\theta - t_l^j + 1} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{m} \sum_{l=1}^{k_t^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1} \\ \mathbf{0} \end{bmatrix}.$$

Because $A\nabla\lambda(p^D) = [M^{-1}|...]$, I have $A\nabla\lambda(p^D)E^j\nabla p(\lambda^D)H = M^{-1}M^j$ which allows me to write the following identity as long as λ_t lies in the interior of Ω_λ:

$$\begin{aligned} A\lambda_t - A\lambda^D &= -A\nabla\lambda(p^D)\sum_{j=1}^{m} E^j \nabla p(\lambda^D) H \sum_{l=1}^{k_t^j} \frac{\tilde{\delta}_l^j}{\theta - t_l^j + 1} + \frac{1}{2} A\epsilon_t' \nabla^2 \lambda(\eta_t)\epsilon_t \qquad (47) \\ &= -M^{-1} \sum_{j=1}^{m} \sum_{l=1}^{k_t^j} \frac{M^j \tilde{\delta}_l^j}{\theta - t_l^j + 1} + \frac{1}{2} A\epsilon_t' \nabla^2 \lambda(\eta_t)\epsilon_t \\ &= -M^{-1} \sum_{j=1}^{m} \sum_{l=1}^{k_t^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1} + \frac{1}{2} A\epsilon_t' \nabla^2 \lambda(\eta_t)\epsilon_t. \end{aligned}$$

<u>STEP 2</u>

I define a stopping time τ and give an upper bound for $\mathbf{E}[\theta - \tau]$. Recall that in (A4), I assume that the absolute values of the eigenvalues of the matrices $\nabla^2\lambda_j, j = 1, \ldots, n$ are bounded above by $\bar{v}$. Let $\lambda_L = \lambda^D - \phi_L \mathbf{e}$ and $\psi = \min\left\{\psi', \psi'^2\right\}$, where

$$\psi' = \min\left\{ \frac{\min\{\phi_L, \phi_U\}}{\max\{\bar{v}, 2 \cdot ||\nabla\lambda(p^D)||_\infty\}}, \frac{\min\{A\lambda_L\}}{\max\{||A\mathbf{e}||_\infty, 2 \cdot ||M^{-1}||_\infty\}} \right\}.$$

One can directly verify that $\psi > 0$. Define a stopping time τ to be the minimum of θ and the first time when any of the following conditions is violated.

(E1) $\psi > \frac{1}{\theta - t} \left| \sum_{s=1}^{t} \left(\xi_s^j - \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j}{\theta - t_l^j + 1} \right) \right|, \forall j = 1, \ldots, m;$

(E2) $\psi > \frac{\bar{v}}{\theta - t} \sum_{s=1}^{t} \left|\left| \sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1} \right|\right|_2^2;$

(E3) $\psi > \left|\left| \sum_{j=1}^{m} \sum_{l=1}^{k_t^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1} \right|\right|_2^2.$

The three conditions listed above are somewhat technical and not easy to interpret. However, they are just stronger conditions of the two conditions below which have more obvious meaning.

(E1*) $\lambda_s \in [\lambda^D - \phi_L \mathbf{e}, \lambda^D - \phi_U \mathbf{e}] \subseteq \Omega_\lambda$, $\forall s \leq t$;

(E2*) $C_t > 0$,

where C_t denotes the remaining inventory at the end of period t. The first condition states that all the target demand rates under LPC up to period t (including t) are feasible, so are the corresponding prices. The second condition states that no stock-out happens by the end of period t. Per our discussion in STEP 1, (E1*) ensures the validity of expression (46) and (47). In addition, (E2*) ensures that all the demand requests up to period t are satisfied, so the dynamics of the resource consumption can be fully expressed by the demand error Δ_s's. Hence, under (E1*) and (E2*), I can track the inventory levels by explicitly quantifying them using past demand errors. (I emphasize that the purpose of (E1)-(E3) is simply for analytical tractability.) The following lemma reveals the connection between (E1)-(E3) and (E1*)-(E2*).

Lemma A.1.2 *I have: (E1)-(E3) $\Rightarrow$ (E1*)-(E2*). In other words, (E1*)-(E2*) hold when $t < \tau$.*

The next lemma provides an upper bound of $\mathbf{E}[\theta - \tau]$ as a function of updating schedule $\{\gamma_j\}_{j=1}^m$.

Lemma A.1.3 *Let $U_1^j(T,t)$ and $U_2^j(T,t)$ be as defined in Theorem 2.4.1. Then, there exists a constant $\bar{\Psi}$, independent of θ and the choice of the projection matrix H such that:*

$$\mathbf{E}[\theta - \tau(\theta)] \ \leq \ \bar{\Psi} \sum_{j=1}^{m} \sum_{t=1}^{\theta-1} \left(\min\left\{1, \left|\left|\nabla p(\lambda^D) H A\right|\right|_2^2 U_1^j(\theta,t)\right\} + \min\left\{1, \left|\left|\nabla p(\lambda^D) H A\right|\right|_2^2 U_2^j(\theta,t)\right\}\right).$$

Although the two lemmas above are crucial and their proofs are quite subtle, I defer the details for now and focus on the main thread of the proof.

<u>STEP 3</u>

I analyse the revenue loss incurred by LPC. Let $R_t(p_t)$ denote the revenue collected in period t under the posted price p_t. So, $R_{H,\gamma_\mathcal{B}} = \sum_{t=1}^{\theta} R_t(p_t)$. Define $\bar{\Delta}_t := R_t(p_t) - \mathbf{E}[R_t(p_t)|\mathcal{F}_t] = R_t(p_t) - r(p_t)$. Since p^D is the optimal solution to DPP, $J^{Det} = p^{D'}\lambda(p^D) = r(p^D)$. This yields

$$
\begin{aligned}
J^{Det} - \mathbf{E}\left[R_{H,\gamma_{\mathcal{B}}}\right] &= J^{Det} - \mathbf{E}\left[\sum_{t=1}^{\theta} R_t(p_t)\right] \\
&= \mathbf{E}\left[\sum_{t=1}^{\tau-1}\big(r(p^D) - R_t(p_t)\big)\right] + \mathbf{E}\left[\sum_{t=\tau}^{\theta}\big(r(p^D) - R_t(p_t)\big)\right] \\
&\leq \mathbf{E}\left[\sum_{t=1}^{\tau-1}\big(r(p^D) - R_t(p_t)\big)\right] + \mathbf{E}\left[\sum_{t=\tau}^{\theta} r(p^D)\right] \\
&\leq \mathbf{E}\left[\sum_{t=1}^{\tau-1}\big(r(p^D) - R_t(p_t)\big)\right] + \bar{r}\,\mathbf{E}[\theta - \tau + 1].
\end{aligned}
$$

For $t < \tau$, by Taylor's expansion at p^D, I have $r(p_t) = r(p^D) - \nabla r(p^D)\epsilon_t + \frac{1}{2}\epsilon_t'\nabla^2 r(\rho_t)\epsilon_t$ for some $\rho_t \in [p^D, p^D - \epsilon_t]$. So, the first term after the last inequality above can be bounded as follows:

$$
\begin{aligned}
\mathbf{E}\left[\sum_{t=1}^{\tau-1}\big(r(p^D) - R_t(p_t)\big)\right] &= \mathbf{E}\left[\sum_{t=1}^{\tau-1}\big(r(p^D) - r(p_t) - \bar{\Delta}_t\big)\right] \\
&= \mathbf{E}\left[\sum_{t=1}^{\tau-1}\left(\nabla r(p^D)\epsilon_t - \frac{1}{2}\epsilon_t'\nabla^2 r(\rho_t)\epsilon_t - \bar{\Delta}_t\right)\right] \\
&= \mathbf{E}\left[\sum_{t=1}^{\tau-1}\nabla r(\lambda^D)\nabla\lambda(p^D)\epsilon_t\right] - \frac{1}{2}\mathbf{E}\left[\sum_{t=1}^{\tau-1}\epsilon_t'\nabla^2 r_t(\rho_t)\epsilon_t\right] - \mathbf{E}\left[\sum_{t=1}^{\tau-1}\bar{\Delta}_t\right] \\
&\leq \mathbf{E}\left[\sum_{t=1}^{\tau-1}\nabla r(\lambda^D)\nabla\lambda(p^D)\epsilon_t\right] - \frac{1}{2}\mathbf{E}\left[\sum_{t=1}^{\tau-1}\epsilon_t'\nabla^2 r_t(\rho_t)\epsilon_t\right] - \mathbf{E}\left[\sum_{t=1}^{\tau}\bar{\Delta}_t\right] + \bar{r},
\end{aligned}
$$

where the third equality holds by the chain rule $\nabla r(\lambda^D)\,\nabla\lambda(p^D) = \nabla r(p^D)$ and the last inequality follows because $\mathbf{E}\left[\bar{\Delta}_\tau\right] \leq \mathbf{E}\left[R_\tau(p_\tau)\right] = \mathbf{E}\left[\mathbf{E}\left[R_\tau(p_\tau)|\tau\right]\right] \leq \bar{r}$. Note that $\{\bar{\Delta}_t\}_{t=1}^{\theta}$ is a martingale with respect to the natural filtration and τ is bounded, so $\mathbf{E}\left[\sum_{t=1}^{\tau}\bar{\Delta}_t\right] = 0$ by the optional stopping theorem. Therefore, I only need to derive upper bounds for the first two terms above, which will be the primary focus of STEP 4 and 5.

<u>STEP 4</u>

I derive an upper bound for $\mathbf{E}\left[\sum_{t=1}^{\tau-1}\nabla r(\lambda^D)\nabla\lambda(p^D)\epsilon_t\right]$. Let π and μ denote the duals associated with the inventory constraints and the constraints $\lambda_t \in \Omega_\lambda$ of DPP respectively. Note that neither depends on θ. By assumption (A5), the optimal solution of DPP is interior. As a result of Karush-Kuhn-Tucker (KKT) optimality condition, I have $\nabla r(\lambda^D) = \pi' A$ (note that $\mu = 0$ by complementary slackness). Thus, $\mathbf{E}\left[\sum_{t=1}^{\tau-1}\nabla r(\lambda^D)\nabla\lambda(p^D)\epsilon_t\right] = \mathbf{E}\left[\sum_{t=1}^{\tau-1}\pi' A\nabla\lambda(p^D)\epsilon_t\right]$. By definition of ϵ_t and $A\nabla\lambda(p^D)E^j\nabla p(\lambda^D)H = M^{-1}M^j$ (see STEP

1), I can write

$$\begin{aligned}\mathbf{E}\left[\sum_{t=1}^{\tau-1}\nabla r(\lambda^D)\nabla\lambda(p^D)\epsilon_t\right] &= \mathbf{E}\left[\sum_{t=1}^{\tau-1}\pi' M^{-1}\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{M^j\tilde{\delta}_l^j}{\theta - t_l^j + 1}\right] \\ &= \pi' M^{-1}\sum_{j=1}^{m} M^j\mathbf{E}\left[\sum_{t=1}^{\tau-1}\sum_{l=1}^{k_t^j}\frac{\tilde{\delta}_l^j}{\theta - t_l^j+1}\right] \\ &= \pi' M^{-1}\sum_{j=1}^{m} M^j\mathbf{E}\left[\sum_{l=1}^{k_{\tau-1}^j}\left(1-\frac{\theta-\tau+1}{\theta-t_l^j+1}\right)\tilde{\delta}_l^j\right]. \qquad (48)\end{aligned}$$

The last term (48) can be further broken down into two parts as follows:

$$\begin{aligned} & \pi' M^{-1}\sum_{j=1}^{m} M^j\mathbf{E}\left[\sum_{l=1}^{k_{\tau-1}^j}\left(1-\frac{\theta-\tau+1}{\theta-t_l^j+1}\right)\tilde{\delta}_l^j + (1-1)\cdot\left(\delta_{t_{k_{\tau-1}^j}}+\cdots+\delta_{\tau-1}\right)\right] \\ = \; & \pi' M^{-1}\sum_{j=1}^{m} M^j\mathbf{E}\left[\left(\delta_{t_{k_{\tau-1}^j}}+\cdots+\delta_{\tau-1}+\sum_{l=1}^{k_{\tau-1}^j}\tilde{\delta}_l^j\right)-\left(\delta_{t_{k_{\tau-1}^j}}+\cdots+\delta_{\tau-1}+\sum_{l=1}^{k_{\tau-1}^j}\frac{\theta-\tau+1}{\theta-t_l^j+1}\tilde{\delta}_l^j\right)\right] \\ \leq \; & \pi' M^{-1}\sum_{j=1}^{m} M^j\mathbf{E}\left[\delta_{t_{k_{\tau-1}^j}}+\cdots+\delta_{\tau-1}+\sum_{l=1}^{k_{\tau-1}^j}\tilde{\delta}_l^j\right] \\ & +\pi'\mathbf{E}\left[\left|M^{-1}\sum_{j=1}^{m} M^j\left(\delta_{t_{k_{\tau-1}^j}}+\cdots+\delta_{\tau-1}+\sum_{l=1}^{k_{\tau-1}^j}\frac{\theta-\tau+1}{\theta-t_l^j+1}\tilde{\delta}_l^j\right)\right|\right]. \qquad (49)\end{aligned}$$

Since $\sum_{j=1}^m M^j = M$, by definition of δ_s and $\tilde{\delta}_l^j$ (see STEP 1), I can write

$$\pi' M^{-1}\sum_{j=1}^{m} M^j\mathbf{E}\left[\delta_{t_{k_{\tau-1}^j}}+\cdots+\delta_{\tau-1}+\sum_{l=1}^{k_{\tau-1}^j}\tilde{\delta}_l^j\right] = \pi' M^{-1}\sum_{j=1}^{m} M^j\mathbf{E}\left[\sum_{s=1}^{\tau-1}\delta_s\right] = \pi'\mathbf{E}\left[\sum_{s=1}^{\tau-1}\delta_s\right].$$

Observing that $\{\sum_{s=1}^t \Delta_s\}_{t=1}^{\theta}$ is a martingale and τ is bounded, $\mathbf{E}\left[\sum_{s=1}^{\tau}\Delta_s\right] = 0$ by optional stopping theorem. Also, the elements in π and A are all nonnegative. This implies that $\pi' A\mathbf{E}[\Delta_\tau] = \pi' A\mathbf{E}[\mathbf{E}[\Delta_\tau|\tau]] \leq \pi' A\bar{\lambda}\mathbf{e}$. Thus, the first term in (49) can be bounded by

$$\pi'\mathbf{E}\left[\sum_{s=1}^{\tau-1}\delta_s\right] = \pi' A\mathbf{E}\left[\sum_{s=1}^{\tau-1}\Delta_s\right] \leq \pi' A\mathbf{E}\left[\sum_{s=1}^{\tau}\Delta_s\right] + \pi' A\bar{\lambda}\mathbf{e} = \pi' A\mathbf{e}\bar{\lambda}. \qquad (50)$$

As for the second term in (49), I have the following:

$$
\begin{aligned}
&\pi'\mathbf{E}\left[\left|M^{-1}\sum_{j=1}^{m} M^j\left(\delta_{t_{k^j_{\tau-1}}}+\cdots+\delta_{\tau-1}+\sum_{l=1}^{k^j_{\tau-1}}\frac{\theta-\tau+1}{\theta-t^j_l+1}\tilde{\delta}^j_l\right)\right|\right]\\
&=\pi'\mathbf{E}\left[\left|M^{-1}\sum_{j=1}^{m}\left(\xi^j_{t_{k^j_{\tau-1}}}+\cdots+\xi^j_{\tau-1}+\sum_{l=1}^{k^j_{\tau-1}}\frac{\theta-\tau+1}{\theta-t^j_l+1}\tilde{\xi}^j_l\right)\mathbf{e}_j\right|\right]\\
&\leq\pi'\mathbf{E}\left[||M^{-1}||_\infty\max_{j=1,\ldots,m}\left|\xi^j_{t_{k^j_{\tau-1}}}+\cdots+\xi^j_{\tau-1}+\sum_{l=1}^{k^j_{\tau-1}}\frac{(\theta-\tau+1)\tilde{\xi}^j_l}{\theta-t^j_l+1}\right|\mathbf{e}\right]\\
&=\pi'\mathbf{E}\left[||M^{-1}||_\infty\max_{j=1,\ldots,m}\left|\sum_{s=1}^{\tau-1}\left(\xi^j_s-\sum_{l=1}^{k^j_s}\frac{\tilde{\xi}^j_l}{\theta-t^j_l+1}\right)\right|\mathbf{e}\right]\\
&\leq\pi'\mathbf{E}\left[\frac{A\lambda_L}{2}(\theta-\tau+1)\right]\ \leq\ \frac{\pi' A\mathbf{e}\bar{\lambda}}{2}\mathbf{E}\left[\theta-\tau+1\right],
\end{aligned}
\tag{51}
$$

where the last equality holds because

$$
\sum_{s=1}^{\tau-1}\left(\xi^j_s-\sum_{l=1}^{k^j_s}\frac{\tilde{\xi}^j_l}{\theta-t^j_l+1}\right)\ =\ \xi^j_{t_{k^j_{\tau-1}}}+\cdots+\xi^j_{\tau-1}+\sum_{l=1}^{k^j_{\tau-1}}\frac{(\theta-\tau+1)\tilde{\xi}^j_l}{\theta-t^j_l+1}, \tag{52}
$$

and the second to the last inequality results from the definition of ψ, the condition (E1) used to define τ, and the fact that $\min\{A\lambda_L\}\mathbf{e}\leq A\lambda_L$. Combining (50) $-$ (51), I get: $\mathbf{E}\left[\sum_{t=1}^{\tau-1}\nabla r(\lambda^D)\nabla\lambda(p^D)\epsilon_t\right]\ \leq\ \pi' A\mathbf{e}\bar{\lambda}+\frac{1}{2}\pi' A\mathbf{e}\bar{\lambda}\,\mathbf{E}[\theta-\tau+1]$.

STEP 5

I now derive an upper bound for $-\frac{1}{2}\mathbf{E}\left[\sum_{t=1}^{\tau-1}\epsilon_t'\nabla^2 r_t(\rho_t)\epsilon_t\right]$ as follows:

$$
\begin{aligned}
-\frac{1}{2}\mathbf{E}\left[\sum_{t=1}^{\tau-1}\epsilon_t'\nabla^2 r_t(\rho_t)\epsilon_t\right]\ &\leq\ \mathbf{E}\left[\left|\sum_{t=1}^{\tau-1}\epsilon_t'\nabla^2 r_t(\rho_t)\epsilon_t\right|\right]\\
&\leq\ \bar{v}\,\mathbf{E}\left[\sum_{t=1}^{\tau-1}\left|\left|\sum_{j=1}^{m}\sum_{l=1}^{k^j_t}\frac{\tilde{\xi}^j_l\mathbf{e}_j}{\theta-t^j_l+1}\right|\right|_2^2\right]\leq\psi\mathbf{E}\left[\theta-\tau+1\right],
\end{aligned}
$$

where the second inequality follows from Lemma A.1.1 and assumption (A5), and the last inequality follows from condition (E2) in the definition of τ.

STEP 6

Putting together results in STEP 1 - 5 proves Theorem 2.4.1. I only need to prove Lemma A.1.2 and A.1.3 which I do below.

Proof of Lemma A.1.2. I need to show that if $t < \tau$, then (E1*) and (E2*) hold. I first show that (E1*) holds:

$$\left|\epsilon_t' \nabla^2 \lambda(\eta_t)\epsilon_t\right| \quad \leq \quad \bar{v}\,\mathbf{e}\,\epsilon_t'\epsilon_t \quad = \quad \bar{v}\,\mathbf{e}\left|\left|\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2^2 \;<\; \min\{\phi_L, \phi_U\}\,\mathbf{e}.$$

The last inequality follows from (E3) in the definition of τ. In addition, I also have

$$\begin{aligned}
\left|\left|\nabla\lambda(p^D)\epsilon_t\right|\right|_\infty &= \left|\left|\nabla\lambda(p^D)\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_\infty \;\leq\; \left|\left|\nabla\lambda(p^D)\right|\right|_\infty \cdot \left|\left|\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_\infty \\
=\; \left|\left|\nabla\lambda(p^D)\right|\right|_\infty \cdot \max_{j=1,\ldots,m}\left|\sum_{l=1}^{k_t^j}\frac{\tilde{\xi}_l^j}{\theta - t_l^j + 1}\right| &\;\leq\; \left|\left|\nabla\lambda(p^D)\right|\right|_\infty \left|\left|\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2 \\
<\; \psi'\left|\left|\nabla\lambda(p^D)\right|\right|_\infty \leq \frac{1}{2}\min\{\phi_L, \phi_U\}\,.
\end{aligned}$$

By combining the two inequalities above with (46), I get

$$|\lambda_t - \lambda^D| \leq \left|\nabla\lambda(p^D)\epsilon_t\right| + \left|\frac{1}{2}A\epsilon_t'\nabla^2\lambda(\eta_t)\epsilon_t\right| \leq \left|\left|\nabla\lambda(p^D)\epsilon_t\right|\right|_\infty \mathbf{e} + \left|\frac{1}{2}A\epsilon_t'\nabla^2\lambda(\eta_t)\epsilon_t\right| \leq \min\{\phi_L, \phi_U\}\,\mathbf{e}.$$

So, (E1*) holds. I next show that (E1)-(E3) imply (E2*). Since (E1)-(E3) imply (E1*), I know formula (46) holds for all $s \leq t$. As a result, the resource consumption error formula (47) also holds. Define $C_t^D := C - \sum_{s=1}^{t} A\lambda^D$. Then, the remaining inventory at the end of period t satisfies

$$\begin{aligned}
C_t \;&\geq\; C - \sum_{s=1}^{t} AD_s = C - \sum_{s=1}^{t} A(\Delta_s + \lambda_s + \lambda^D - \lambda^D) = C_t^D - \sum_{s=1}^{t} A(\Delta_s + \lambda_s - \lambda^D) \\
&=\; C_t^D - \sum_{s=1}^{t}\left(\delta_s - M^{-1}\sum_{j=1}^{m}\sum_{l=1}^{k_s^j}\frac{M^j\tilde{\delta}_l^j}{\theta - t_l^j + 1} + \frac{1}{2}A\epsilon_s'\nabla^2\lambda(\eta_s)\epsilon_s\right) \\
&\geq\; C_t^D - \left|M^{-1}\sum_{s=1}^{t}\sum_{j=1}^{m}\left(\xi_s^j\mathbf{e}_j - \sum_{l=1}^{k_s^j}\frac{\tilde{\xi}_l^j\mathbf{e}_j}{\theta - t_l^j + 1}\right)\right| - \left|\frac{1}{2}\sum_{s=1}^{t}A\epsilon_s'\nabla^2\lambda(\eta_s)\epsilon_s\right|. \qquad (53)
\end{aligned}$$

Because $\{\lambda^D\}$ is the optimal solution to DPP, I know that it must satisfy inventory constraint. So, $C_t^D = C - \sum_{s=1}^{t} A\lambda^D \geq \sum_{s=t+1}^{\theta} A\lambda^D$. Since I also have $\lambda^D > \lambda_L \mathbf{e}$, it must hold that $C_t^D \geq \sum_{s=t+1}^{\theta} A\lambda^D \geq A\lambda_L(\theta - t)$. As for the second term in (53), by (E1), I have

$$\begin{aligned}
& \left| M^{-1} \sum_{s=1}^{t} \sum_{j=1}^{m} \left(\xi_s^j \mathbf{e}_j - \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1} \right) \right| \\
\leq \; & \left|\left|M^{-1}\right|\right|_\infty \cdot \left|\left| \sum_{j=1}^{m} \sum_{s=1}^{t} \left(\xi_s^j \mathbf{e}_j - \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1} \right) \right|\right|_\infty \cdot \mathbf{e} \\
\leq \; & \left|\left|M^{-1}\right|\right|_\infty \cdot \max_{j=1,\ldots,m} \left| \sum_{s=1}^{t} \left(\xi_s^j - \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j}{\theta - t_l^j + 1} \right) \right| \mathbf{e} \\
< \; & \left|\left|M^{-1}\right|\right|_\infty \psi(\theta - t)\mathbf{e} \; < \; \frac{A\lambda_L}{2}(\theta - t) \; \leq \; \frac{1}{2} C_t^D.
\end{aligned}$$

For the third term in (53),the following holds by Lemma A.1.1.

$$\begin{aligned}
\left| \frac{1}{2} \sum_{s=1}^{t} A\epsilon_s' \nabla^2 \lambda(\eta_s)\epsilon_s \right| \; & \leq \; \frac{1}{2} A \sum_{s=1}^{t} \left| \epsilon_s' \nabla^2 \lambda(\eta_s) \epsilon_s \right| \leq \frac{1}{2} A\bar{v}\mathbf{e} \sum_{s=1}^{t} \epsilon_s' \epsilon_s \\
& = \; \frac{1}{2} A\bar{v}\mathbf{e} \sum_{s=1}^{t} \left|\left| \sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1} \right|\right|_2^2 \\
& < \; \frac{1}{2} A\mathbf{e}(\theta - t)\psi \; \leq \; \frac{1}{2} \left|\left|A\mathbf{e}\right|\right|_\infty \mathbf{e}(\theta - t)\psi \leq \frac{A\lambda_L}{2}(\theta - t) \; \leq \; \frac{1}{2} C_t^D.
\end{aligned}$$

Combining the two bounds above with (53), I get $C_t > 0$. So, (E2*) holds.

Proof of Lemma A.1.3. Let τ_1^j denote the minimum of θ and the first time t such that condition (E1) is violated for j^{th} resource. Also, let denote τ_i, $i = 2, 3$, denote the minimum of θ and the first time t such that condition (Ei) is violated. Note that, by definition, $\tau = \min\{(\min_j \tau_1^j), \tau_2, \tau_3\}$. Since τ is nonnegative, $\mathbf{E}[\tau] = \sum_{t=0}^{\theta-1} \Pr(\tau > t)$. So, I can write $\mathbf{E}[\theta - \tau] = \theta - \mathbf{E}[\tau] = \sum_{t=1}^{\theta-1} \Pr(\tau \leq t)$. Since $\tau \leq t$ can only happen if either τ_1^j (for some j) or τ_2 or τ_3 gets hit by time t, by sub-additivity property of probability, I can bound: $\Pr(\tau \leq t) \; \leq \; \sum_{j=1}^{m} \Pr(\tau_1^j \leq t) + \Pr(\tau_2 \leq t) + \Pr(\tau_3 \leq t)$. So, it suffices to derive a bound for each component after the inequality. I do this in turn.

STEP 1

I derive an upper bound for $\Pr(\tau_1^j \leq t)$, $j = 1, \ldots, m$. Fix t. For each $j = 1, \ldots, m$, I define a hitting time $\tilde{\tau}_1^j$ to be the minimum of t and the first time $v \le t$ such that $\psi \ \le \ |S_v^j|$, where

$$S_v^j \ = \ \begin{cases} \frac{\xi^j_{t^j_{k^j_v}} + \cdots + \xi^j_v}{\theta - t^j_{k^j_v+1} + 1} + \sum_{l=1}^{k^j_v} \frac{\tilde{\xi}^j_l}{\theta - t^j_l + 1}, & 1 \le v \le t^j_{k^j_t} - 1 \\ \\ \frac{\xi^j_{t^j_{k^j_v}} + \cdots + \xi^j_v}{\theta - v} + \sum_{l=1}^{k^j_v} \frac{\tilde{\xi}^j_l}{\theta - t^j_l + 1}, & t^j_{k^j_t} \le v \le t \end{cases} .$$

I now state a lemma which reveals the connection between τ_1^j and $\tilde{\tau}_1^j$, see STEP 4 below for proof.

Lemma A.1.4 $\Pr(\tau_1^j \le t) \ \le \ \Pr(\tilde{\tau}_1^j \le t)$.

Observe that for any given t, $\{S_v\}_{v=1}^t$ is a martingale with respect to the natural filtration $\{\mathcal{F}_v\}_{v=1}^t$. Hence, $\{|S_v|\}_{v=1}^t$ is a submartingale. By Doob's submartingale inequality and identity in (52), I have

$$\begin{aligned} \Pr(\tau_1^j \le t) \ &\le \ \Pr(\tilde{\tau}_1^j \le t) \ = \ \Pr\left(\max_{v \le t} \left| \frac{\xi^j_{t^j_{k^j_v}} + \cdots + \xi^j_v}{\theta - v} + \sum_{l=1}^{k^j_v} \frac{\tilde{\xi}^j_l}{\theta - t^j_l + 1} \right| \ge \psi \right) \\ &\le \ \min\left\{ 1, \frac{1}{\psi^2} \mathbf{E}\left[\left| \frac{\xi^j_{t^j_{k^j_t}} + \cdots + \xi^j_t}{\theta - t} + \sum_{l=1}^{k^j_t} \frac{\tilde{\xi}^j_l}{\theta - t^j_l + 1} \right|^2 \right] \right\} \\ &= \ \min\left\{ 1, \frac{1}{\psi^2} \left\{ \sum_{s=t^j_{k^j_t}}^{t} \frac{\mathbf{E}\left[(\xi^j_s)^2 \right]}{(\theta - t)^2} + \sum_{l=1}^{k^j_t} \frac{\mathbf{E}\left[\left(\tilde{\xi}^j_l\right)^2 \right]}{\left(\theta - t^j_l + 1\right)^2} \right\} \right\}, \end{aligned}$$

where the last equality holds because $\mathbf{E}[\xi^j_s \xi^j_v] = 0$ for $s \ne v$, $\mathbf{E}[\tilde{\xi}^j_l \tilde{\xi}^j_w] = 0$ for $l \ne w$, and $\mathbf{E}[\xi^j_s \tilde{\xi}^j_l] = 0$ for $s \ge t^j_{k^j_t}$ and $l \le k^j_t$. Now I want to estimate the expectations in the upper bound above. I start with the term $(\xi^j_s)^2$. By matrix norm inequality, $\left(\xi^j_s\right)^2 \le \sum_{i=1}^m \left(\xi^i_s\right)^2 = (MA\Delta_s)' (MA\Delta_s) \le ||MA||_2^2 \, \Delta_s' \Delta_s \ = \ \left|\left| \nabla p(\lambda^D) HA \right|\right|_2^2 \Delta_s' \Delta_s \le \left|\left| \nabla p(\lambda^D) \right|\right|_2^2 \, ||H||_2^2 \, ||A||_2^2 \, \Delta_s' \Delta_s$. Taking expectation on both sides and using $\mathbf{E}[\Delta_t' \, \Delta_t] = \mathbf{Var}(\Delta_t) \le 1$ (due to the assumption that at most one customer arrives in each period) yields $\mathbf{E}[(\xi^j_s)^2] \le \left|\left| \nabla p(\lambda^D) \right|\right|_2^2 \, ||H||_2^2 \, ||A||_2^2$.

By definition, $\tilde{\xi}_l^j = \sum_{s=t_{l-1}^j}^{t_l^j-1} \xi_s^j$. So I have

$$\mathbf{E}\left[\left(\tilde{\xi}_l^j\right)^2\right] = \sum_{s=t_{l-1}^j}^{t_l^j-1} \mathbf{E}\left[\left(\xi_s^j\right)^2\right] \leq \left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2 \sum_{s=t_{l-1}^j}^{t_l^j-1} \mathbf{E}\left[\Delta_s' \Delta_s\right] \leq \left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2 \left(t_l^i - t_{l-1}^i\right).$$

Putting the inequalities together, I obtain that $\sum_{j=1}^m \sum_{t=1}^{\theta-1} \Pr(\tau_1^j \leq t) \leq \sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\left\{1, \frac{\left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2}{\psi^2} U\right.$

STEP 2

I derive an upper bound for $\Pr(\tau_2 \leq t)$. Since $\left|\left|\sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2^2 \geq 0$ and $\bar{v} \geq 0$, I conclude that for all $v \leq t$, $\frac{\bar{v}}{\theta-t}\sum_{s=1}^t \left|\left|\sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2^2 \geq \frac{\bar{v}}{\theta-v}\sum_{s=1}^v \left|\left|\sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2^2$. Therefore, by Markov's inequality, the following holds:

$$\begin{aligned}
\Pr(\tau_2 \leq t) &= \Pr\left(\max_{v \leq t} \frac{\bar{v}}{\theta - v} \sum_{s=1}^v \left|\left|\sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2^2 \geq \psi\right) \\
&\leq \Pr\left(\frac{\bar{v}}{\theta - t} \sum_{s=1}^t \left|\left|\sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2^2 \geq \psi\right) \\
&\leq \min\left\{1, \frac{\bar{v}}{\psi(\theta - t)} \sum_{s=1}^t \mathbf{E}\left[\left|\left|\sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2^2\right]\right\}.
\end{aligned}$$

By similar arguments as in STEP 1, I can bound

$$\begin{aligned}
\mathbf{E}\left[\left|\left|\sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{\tilde{\xi}_l^j \mathbf{e}_j}{\theta - t_l^j + 1}\right|\right|_2^2\right] &\leq \mathbf{E}\left[\sum_{j=1}^m \sum_{l=1}^{k_s^j} \left(\frac{\tilde{\xi}_l^j}{\theta - t_l^j + 1}\right)^2\right] \leq \sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{\mathbf{E}\left[\left(\tilde{\xi}_l^j\right)^2\right]}{(\theta - t_l^j + 1)^2} \\
&\leq \left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2 \sum_{j=1}^m \sum_{l=1}^{k_s^j} \frac{t_l^j - t_{l-1}^j}{(\theta - t_l^j + 1)^2}.
\end{aligned}$$

As a result, I obtain

$$\begin{aligned}\sum_{t=1}^{\theta-1}\Pr(\tau_2\le t) &\le \sum_{j=1}^{m}\sum_{t=1}^{\theta}\min\left\{1,\frac{\bar v\left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2}{\psi(\theta-t)}\sum_{s=1}^{t}\sum_{l=1}^{k_s^j}\frac{t_l^j-t_{l-1}^j}{(\theta-t_l^j+1)^2}\right\}\\ &= \sum_{j=1}^{m}\sum_{t=1}^{\theta}\min\left\{1,\frac{\bar v\left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2}{\psi}U_2^j(\theta,t)\right\}.\end{aligned}$$

STEP 3

I derive an upper bound for $\Pr(\tau_3\le t)$. Observe that for all j, $\left\{\sum_{l=1}^{k_t^j}\frac{\tilde\xi_l^j}{\theta-t_l^j+1}\right\}_{t=1}^{\theta}$ is a martingale. Since $\left|\left|\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{\tilde\xi_l^j\mathbf{e}_j}{\theta-t_l^j+1}\right|\right|_2^2=\sum_{j=1}^{m}\left(\sum_{l=1}^{k_t^j}\frac{\tilde\xi_l^j}{\theta-t_l^j+1}\right)^2$, $\left\{\left|\left|\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{\tilde\xi_l^j\mathbf{e}_j}{\theta-t_l^j+1}\right|\right|_2^2\right\}_{t=1}^{\theta}$ is also a submartingale. So, by Doob's submartingale inequality and arguments in STEP 1,

$$\begin{aligned}\Pr(\tau_3\le t) &= \Pr\left(\max_{v\le t}\left|\left|\sum_{j=1}^{m}\sum_{l=1}^{k_v^j}\frac{\tilde\xi_l^j\mathbf{e}_j}{\theta-t_l^j+1}\right|\right|_2^2\ge\psi\right)\le\frac{1}{\psi}\mathbf{E}\left[\left|\left|\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{\tilde\xi_l^j\mathbf{e}_j}{\theta-t_l^j+1}\right|\right|_2^2\right]\\ &\le \min\left\{1,\frac{\left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2}{\psi}\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{t_l^j-t_{l-1}^j}{(\theta-t_l^j+1)^2}\right\}.\end{aligned}$$

As a result, the following inequality holds:

$$\begin{aligned}\sum_{t=1}^{\theta-1}\Pr(\tau_3\le t) &\le \sum_{j=1}^{m}\sum_{t=1}^{\theta-1}\min\left\{1,\frac{\left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2}{\psi}\sum_{l=1}^{k_t^j}\frac{t_l^j-t_{l-1}^j}{(\theta-t_l^j+1)^2}\right\}\\ &= \sum_{j=1}^{m}\sum_{t=1}^{\theta-1}\min\left\{1,\frac{\left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2}{\psi}U_1^j(\theta,t)\right\}.\end{aligned}$$

STEP 4

Putting together all the results in STEP 1-3 completes the proof of Lemma A.1.3. The last thing to do is to prove Lemma A.1.4 from STEP 1. I do this now.

Proof of Lemma A.1.4. It suffices to show that for all $v\le t$, if $\tau_1^j=v$ occurs, then $\tilde\tau_1^j\le v$ occurs as well. By definition of S_v, this is immediately true if $t_{k_t^j}^j\le v\le t$. So, I only need to check the case $1\le v\le t_{k_t^j}^j-1$. Assuming $1\le v\le t_{k_t^j}^j-1$, by definition of (E1) in STEP

1, $\tau_1^j = v$ means

$$\psi \leq \left| \frac{\xi^j_{t^j_{k^j_v}} + \cdots + \xi^j_v}{\theta - v} + \sum_{l=1}^{k^j_v} \frac{\tilde{\xi}^j_l}{\theta - t^j_l + 1} \right| \text{ and } \psi > \left| \sum_{l=1}^{k^j_v} \frac{\tilde{\xi}^j_l}{\theta - t^j_l + 1} \right|$$

$$\text{which imply that } \psi \leq \left| \frac{\xi^j_{t^j_{k^j_v}} + \cdots + \xi^j_v}{\theta - t^j_{k^j_v+1} + 1} + \sum_{l=1}^{k^j_v} \frac{\tilde{\xi}^j_l}{\theta - t^j_l + 1} \right| = |S^j_v|.$$

So, $\tilde{\tau}_1^j \leq v$ and hence $\Pr(\tau_1^j \leq t) \leq \Pr(\tilde{\tau}_1^j \leq t)$.

A.1.3 Proof of Lemma 2.4.2

I will prove a more general result of picking the best k prices. For any $v \in \mathbb{R}^n$ define $||v||_0 := |\{i : v_i \neq 0\}|$. Let $a = A'$, $x = \nabla p(\lambda^D)H$, $b = (A\nabla\lambda(p^D))'$. Since $m = 1$, a, x, b are all vectors in $\mathbb{R}^n$. The optimization problem $\min_H\{||\nabla p(\lambda^D)HA||_2 : AH = 1, ||\nabla p(\lambda^D)H||_0 \leq k\}$ is equivalent to $\min_x\{||xa'||_2^2 : b'x = 1, ||x||_0 \leq k\}$. Since $xa'ax'$ is a rank one matrix, its maximum eigenvalue is just its trace. So $||xa'||_2^2 = tr(xa'ax') = tr(x'xa'a) = ||x||_2^2||a||_2^2$. Note also that the equality constraint is equivalent to $||b||_2||x||_2\cos(b, x) = 1$, where $\cos(b, x)$ is the cosine of the angle between vectors b and x. Therefore, as long as $||x||_2 = 1/(||b||_2\cos(b, x))$, the equality constraint can be satisfied. So the problem becomes $\min_x\{||a||_2^2||b||_2^{-2}\cos^{-2}(b, x) : ||x||_0 \leq k\}$.Let $b_{(i)}$ denote the i^{th} largest element in absolute value in b, then the optimal solution x^* is parallel with a vector b^k which has the exact same elements as b in the k largest elements in absolute values but zeros in other elements. The optimal objective value is $||a||_2^2(\sum_{i=1}^k b_{(i)}^2)^{-1}$.

A.1.4 Proof of Corollary 2.4.1

I compute, part by part, the bound in Theorem 2.4.1 under periodic price update schedule. Without loss of generality, I assume that $T = 1$. For notational clarity, I suppress the dependence on θ whenever there is no confusion. I start with the summation over $U_1^j(\theta, t)$. First of all, I have

$$\sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\left\{1, \left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2 U_1(\theta, t)\right\} \leq \max\{1, \left|\left|\nabla p(\lambda^D)HA\right|\right|_2^2\} \sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\left\{1, U_1^j(\theta, t)\right\}.$$

I bound the summation after the inequality as follows: $\sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\left\{1, U_1^j(\theta, t)\right\} =$ $m\sum_{t=1}^{\theta-1} \min\left\{1, \frac{t-hk_t}{(\theta-t)^2} + \sum_{l=1}^{k_t} \frac{t_l - t_{l-1}}{(\theta-t_l+1)^2}\right\} \leq m\sum_{t=1}^{\theta-1} \min\left\{1, \frac{t-hk_t}{(\theta-t)^2} + \frac{t_{k_t} - t_{k_t-1}}{\left(\theta-t_{k_t}+1\right)^2} + \int_1^{t_{k_t}} \frac{1}{(\theta-x+1)^2}\,dx\right\} \leq$

$m\sum_{t=1}^{\theta-1}\min\left\{1, \frac{2h}{(\theta-t)^2}+\frac{1}{\theta-t}.\right\}$. The first equality follows since I update the price of the m products at the same time. The first inequality is the integral approximation and the last inequality follows from the fact that $0 \le t - k_t h \le h$. Now define $t^* = \left\lfloor \theta - \sqrt{h} \right\rfloor$. I make further approximation of the inequality above by breaking down the summation over t into two parts, before and after t^*: $\sum_{j=1}^{m}\sum_{t=1}^{\theta-1}\min\{1, U_1(\theta,t)\} \le m\left[\int_1^{t^*}\frac{2h}{(\theta-x)^2}\,dx + \int_1^{t^*}\frac{1}{\theta-x}\,dx + \theta - t^*\right] \le m\left(\frac{2h}{\theta-t^*}+\log\left(\frac{\theta-1}{\theta-t^*}\right)+\theta-t^*\right) \le m\left(1+3\sqrt{h}+\log\theta\right)$ where the first inequality follows from the integration approximation and the third inequality follows from the fact that $1 \le \sqrt{h} \le \theta - t^* \le \sqrt{h}+1$. Now I compute the summation over $U_2^j(\theta,t)$. Similarly, it suffices to bound the following: $\sum_{j=1}^{m}\sum_{t=1}^{\theta-1}\min\left\{1, U_2^j(\theta,t)\right\} \le m\sum_{t=1}^{\theta-1}\min\left\{1, \frac{1}{\theta-t}\sum_{s=1}^{t}\left(\frac{h}{(\theta-s)^2}+\frac{1}{\theta-s}\right)\right\}$. Again, I break the summation into two parts and use integral approximation:

$$
\begin{aligned}
&\sum_{j=1}^{m}\sum_{t=1}^{\theta-1}\min\{1, U_2(\theta,t)\} \\
\le\;& m\left[\sum_{t=1}^{t^*-1}\frac{1}{\theta-t}\int_1^{t+1}\left(\frac{h}{(\theta-x)^2}+\frac{1}{\theta-x}\right)dx + \theta - t^*\right] \\
\le\;& m\left[\sum_{t=1}^{t^*-1}\left(\frac{h}{(\theta-t-1)^2}+\frac{1}{\theta-t-1}\log\left(\frac{\theta-1}{\theta-t-1}\right)\right) + \theta - t^*\right] \\
\le\;& m\left[\frac{h}{(\theta-t^*)^2}+\int_1^{t^*-1}\frac{h}{(\theta-x-1)^2}\,dx + \sum_{t=1}^{t^*-1}\frac{1}{\theta-t-1}\log\left(\frac{\theta-1}{\theta-t-1}\right) + \theta - t^*\right] \\
\le\;& m\left(2+2\sqrt{h}+\log\theta+\log^2\theta\right),
\end{aligned}
$$

where the last inequality holds because $\sum_{t=1}^{t^*-1}\frac{1}{\theta-t-1}\log\left(\frac{\theta-1}{\theta-t-1}\right) \le \frac{\log\left(\frac{\theta-1}{\theta-t^*}\right)}{\theta-t^*}+\int_1^{t^*-1}\frac{\log\left(\frac{\theta-1}{\theta-t-1}\right)}{\theta-t-1}dt \le \log\theta+\log^2\theta$.

A.1.5 Proof of Corollary 2.4.2

I assume without loss of generality that $T = 1$ and suppress the dependence on θ for brevity. Note that $K(\theta)$ is well-defined since $\sum_{s=1}^{k}s^\alpha$ is strictly increasing in k and is unbounded as $k \to \infty$ for all $\alpha \ge 1$. Since $\theta > \sum_{s=1}^{K}s^\alpha \ge K^{\alpha+1}/(\alpha+1)$, I have $K \le ((\alpha+1)\theta)^{1/(\alpha+1)}$. I now analyze the performance bound. I first derive bound for the summation over $U_1^j(\theta,t)$. Similar to the proof of Corollary 2.4.1, it suffices to bound the following: $\sum_{j=1}^{m}\sum_{t=1}^{\theta-1}\min\left\{1, U_1^j(\theta,t)\right\}$. By definition, for $1 \le l \le K$, I have $\theta - t_l + 1 \ge \sum_{s=1}^{K-l+1}s^\alpha \ge \frac{(K-l+1)^{\alpha+1}}{\alpha+1} \ge \frac{(K-l+2)^{\alpha+1}}{2^{\alpha+1}(\alpha+1)}$. In addition, I also have that for $2 \le l \le K$, $t_l - t_{l-1} \le (K-l+2)^\alpha + 1 \le 2(K-l+2)^\alpha$, and for $l = 1$, $t_1 - t_0 \le \theta + 1 - \sum_{s=1}^{K}s^\alpha - 1 \le (K+1)^\alpha \le 2(K+1)^\alpha$. Then, for

$t < \theta - 1$, since $k_t < K$, I have $U_1^j(\theta, t) \leq \sum_{l=1}^{k_t+1} \frac{t_l - t_{l-1}}{(\theta - t_l + 1)^2} \leq \sum_{l=1}^{k_t+1} \frac{2(K-l+2)^\alpha (\alpha+1)^2 2^{2\alpha+2}}{(K-l+2)^{2\alpha+2}} = (\alpha + 1)^2 2^{2\alpha+3} \sum_{l=1}^{k_t+1} \frac{1}{(K-l+2)^{\alpha+2}}$. Hence, $\sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\{1, U_1^j(\theta, t)\} \leq \sum_{j=1}^m (1 + \sum_{t=1}^{\theta-2} U_1^j(\theta, t)) \leq m + m \sum_{l=1}^K 2(K - l + 2)^\alpha \sum_{s=1}^l \frac{(\alpha+1)^2 2^{2\alpha+3}}{(K-s+2)^{\alpha+2}} \leq m + m \sum_{l=1}^K 2(K - l + 2)^\alpha \int_1^{l+1} \frac{(\alpha+1)^2 2^{2\alpha+3}}{(K-s+2)^{\alpha+2}} ds \leq m + m(\alpha + 1) 2^{2\alpha+4} \sum_{l=1}^K \frac{(K-l+2)^\alpha}{(K-l+1)^{\alpha+1}} \leq m + m(\alpha + 1) 2^{3\alpha+4} \sum_{l=1}^K \frac{1}{(K-l+1)} \leq m + m(\alpha + 1) 2^{3\alpha+4} \log K$. Since $K \leq ((\alpha + 1)\theta)^{\frac{1}{\alpha+1}}$, $\sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\{1, U_1^j(\theta, t)\} \leq m(1 + 2^{3\alpha+4} \log(\alpha + 1) + 2^{3\alpha+4} \log \theta)$. As for the summation over $U_2^j(\theta, t)$, I have $\sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\{1, U_2^j(\theta, t)\} \leq m + \sum_{j=1}^m \sum_{t=1}^{\theta-2} U_2^j(\theta, t) \leq m + \sum_{j=1}^m \sum_{t=1}^{\theta-2} \frac{1}{\theta-t} \sum_{s=1}^t U_1^j(\theta, s) \leq m + \sum_{j=1}^m \sum_{t=1}^{\theta-2} \frac{1}{\theta-t} \sum_{s=1}^{\theta-2} U_1^j(\theta, s) \leq m(1 + 2^{3\alpha+4} \log(\alpha + 1) \log \theta + 2^{3\alpha+4} \log^2 \theta)$.

A.1.6 Proof of Corollary 2.4.3

I assume without loss of generality that $T = 1$ and suppress the dependence on θ for brevity. I first show that $K \leq 1 + \log_\beta \theta$. Note that since $\{t_l\}$ are strictly increasing integers, so K is well defined and by definition of t_l I have $t_{K-1} \leq \theta - 1$. By definition, I have $t_l \geq [(\beta - 1)\theta + t_{l-1}]/\beta$, so $\theta - t_l \leq (\theta - t_{l-1})/\beta \leq \theta/\beta^l$. Therefore, $\theta - 1 \geq t_{K-1} > \theta - \theta/\beta^{K-1}$ which implies that $K \leq 1 + \log_\beta \theta$. I now analyze the performance bound. By definition, I have $t_l \leq [(\beta-1)\theta + t_{l-1}]/\beta + 1$, so I have the following useful bound which will be used a couple of times later: for $l \leq K$, $(\star)$ $\frac{t_l - t_{l-1}}{\theta - t_l + 1} \leq \frac{\{[(\beta-1)\theta + t_{l-1}]/\beta + 1\} - t_{l-1}}{\theta - \{[(\beta-1)\theta + t_{l-1}]/\beta + 1\} + 1} = \frac{(\beta-1)(\theta - t_{l-1} + 1) + 1}{\theta - t_{l-1}} \leq 2\beta - 1$. I derive an upper bound for the summation over $U_1^j(\theta, t)$ first. Similar to the proof of Corollary 2.4.1, it suffices to bound the following:

$$
\begin{aligned}
\textstyle\sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\left\{1, U_1^j(\theta, t)\right\} &\leq m \sum_{t=1}^{\theta-1} \left(\frac{t - t_{k_t} + 1}{(\theta - t)^2} + \sum_{l=1}^{k_t} \frac{2\beta - 1}{\theta - t_l + 1} \right) \\
&\leq m \sum_{t=1}^{\theta-1} \left(\frac{t_{k_t+1} - 1 - t_{k_t} + 1}{(\theta - t)(\theta - t_{k_t+1} + 1)} + \sum_{l=1}^{k_t} \frac{2\beta - 1}{\theta - t_l + 1} \right) \\
&\leq m \sum_{t=1}^{\theta-1} \left(\frac{2\beta - 1}{(\theta - t)} + \sum_{l=1}^{k_t} \frac{2\beta - 1}{\theta - t_l + 1} \right) \\
&\leq m(2\beta - 1) \left(\log \theta + \sum_{t=1}^{\theta-1} \sum_{l=1}^{k_t} \frac{1}{\theta - t_l + 1} \right),
\end{aligned}
$$

where the first and the third inequalities follow from $(\star)$. Note that $\sum_{t=1}^{\theta-1} \sum_{l=1}^{k_t} \frac{1}{\theta - t_l + 1} = \sum_{j=0}^{K-1} \sum_{t=t_j}^{t_{j+1}-1} \sum_{l=1}^j \frac{1}{\theta - t_l + 1} = \sum_{j=0}^{K-1} \sum_{l=1}^j \frac{t_{j+1} - t_j}{\theta - t_l + 1} = \sum_{j=1}^{K-1} \frac{\theta - t_j}{\theta - t_j + 1} \leq K - 1 \leq \log_\beta \theta$. Hence, I have $\sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\left\{1, U_1^j(\theta, t)\right\} \leq m(2\beta - 1)\left(\log \theta + \log_\beta \theta\right)$. Now I approximate the summation over $U_2^j(\theta, t)$ as follows: $\sum_{j=1}^m \sum_{t=1}^{\theta-1} \min\left\{1, U_2^j(\theta, t)\right\} \leq m \sum_{t=1}^{\theta-1} \frac{1}{\theta-t} \sum_{s=1}^{\theta-1} \sum_{l=1}^{k_s} \frac{2\beta-1}{\theta - t_l + 1} \leq m(2\beta - 1) \log \theta \log_\beta \theta$.

A.1.7 Proof of Theorem 2.4.2

I use a slight modification of LPC with synchronous 1-Periodic Schedule as follows: follow the LPC heuristic but uses $\hat{p}_t = p^D - \nabla p(\lambda^D) H \sum_{s=1}^{t-1} \frac{A\Delta_s}{T-s}$, where H is a projection matrix. Call this heuristic π_H. Pick an H that satisfies $||\nabla p(\lambda^D)H||_0 \leq k$. Then I have $\pi_H \in \Pi_k$. Following a similar argument as Theorem 2.4.1, there exist positive constants Ψ and $\hat{\Psi}$ such that $J^{Det} - \mathbf{E}[R_{\pi_H}(\theta)] \leq \Psi + \hat{\Psi}||\nabla p(\lambda^D)HA||_2^2 \log^2 \theta$. By the proof of Lemma 2.4.2, if I minimize $||\nabla p(\lambda^D)HA||_2$ subject to $AH = 1$ and $||\nabla p(\lambda^D)H||_0 \leq k$, the optimal projection matrix H^* attains $||\nabla p(\lambda^D)H^*A||_2^2 = ||a||_2^2(\sum_{i=1}^k b_{(i)}^2)^{-1}$. Therefore, $\min_{\pi \in \Pi_k}\{J^{Det} - \mathbf{E}[R_\pi(\theta)]\} \leq J^{Det} - \mathbf{E}[R_{\pi_{H^*}}(\theta)] \leq \Psi + \bar{\Psi}(\sum_{i=1}^k b_{(i)}^2)^{-1} \log^2 \theta$, where $\bar{\Psi} = \hat{\Psi}||a||_2^2$.

A.1.8 Proof of Lemma 2.5.1

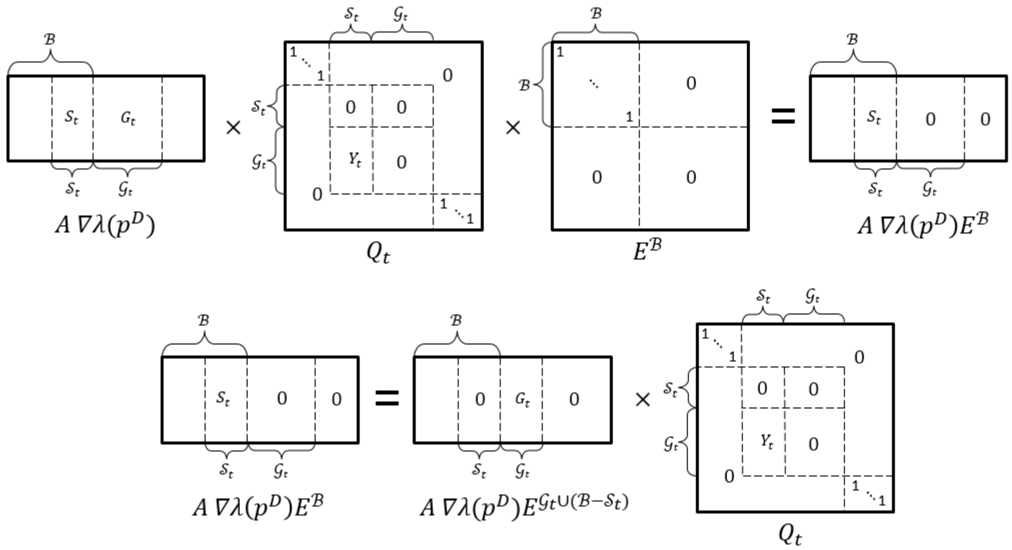
By the construction of $Q(Y_t, \mathcal{G}_t, \mathcal{S}_t)$, it is straightforward to verify that $A\nabla\lambda(p^D)Q_t E^{\mathcal{B}} = A\nabla\lambda(p^D)E^{\mathcal{B}} = A\nabla\lambda(p^D)E^{\mathcal{G}_t \cup (\mathcal{B}-\mathcal{S}_t)}Q_t$ holds. (See Figure A.1 for an illustration.) This proves (1). For (2), construct $H_t := \nabla\lambda(p^D)Q_t\nabla p(\lambda^D)H$. Note that H_t is a projection matrix since $AH_t = A\nabla\lambda(p^D)Q_t\nabla p(\lambda^D)H = A\nabla\lambda(p^D)Q_tE^{\mathcal{B}}\nabla p(\lambda^D)H = A\nabla\lambda(p^D)E^{\mathcal{B}}\nabla p(\lambda^D)H = A\nabla\lambda(p^D)\nabla p(\lambda^D)H = I$ where the second and the fourth equality follows by the fact that only the first m rows of $\nabla p(\lambda^D)H$ are nonzero. Note also that $\nabla p(\lambda^D)H_t = Q_t\nabla p(\lambda^D)H$. So, to verify that rows in $\nabla p(\lambda^D)H_t$ that correspond to products not in $\mathcal{G}_t \cup (\mathcal{B} - \mathcal{S}_t)$ are zeros, I only need to verify it for $Q_t\nabla p(\lambda^D)H$. For any $j \notin \mathcal{G}_t \cup (\mathcal{B} - \mathcal{S}_t)$, either (a) $j \in \mathcal{S}_t$ and $j \notin \mathcal{G}_t$, or (b) $j \notin \mathcal{B} \cup \mathcal{G}_t$. In case (a), the result holds since the j^{th} row of Q_t is a zero vector. In case (b), the only nonzero element in row j of Q_t is the j^{th} element, but the j^{th} row of $\nabla p(\lambda^D)H$ is a zero vector. This proves (2). Finally, since the only nonzero elements in $Q_tE^{\mathcal{S}_t}$ are in the submatrix consisting of rows in $\mathcal{G}_t$ and columns in $\mathcal{S}_t$, I conclude that the rows in $Q_tE^{\mathcal{S}_t}\nabla p(\lambda^D)H$ that correspond to products not in $\mathcal{G}_t$ are zero vectors. This completes the proof of (3).

A.1.9 Proof Sketch of Theorem 2.5.1

The proof of Theorem 2.5.1 follows the same outline of the proof for Theorem 2.4.1 with three nontrivial twists.

1) Resource Correction Equivalence. I first show that in terms of error correction, equivalent pricing is "equivalent" to LPC. In particular, let $\tilde{\epsilon}_t = \sum_{j=1}^m \sum_{l=1}^{k_t^j} Q_{t_l^j} E^j \nabla p(\lambda^D) H \frac{A\tilde{\Delta}_l^j}{T - t_l^j + 1}$. For simplicity, disregard the second order term of Taylor expansion of λ_t, then I have exactly

Figure A.1: Illustration of Lemma 2.5.1 part (1)



the same capacity error below as (47) in the proof of Theorem 2.4.1:

$$
\begin{aligned}
A\lambda_t - A\lambda^D &= -A\nabla\lambda(p^D)\tilde{\epsilon}_t = -\sum_{j=1}^{m}\sum_{l=1}^{k_t^j} A\nabla\lambda(p^D)Q_{t_l^j}E^j\nabla p(\lambda^D)H\frac{A\tilde{\Delta}_l^j}{T-t_l^j+1} \\
&= -\sum_{j=1}^{m}\sum_{l=1}^{k_t^j} A\nabla\lambda(p^D)E^j\nabla p(\lambda^D)H\frac{A\tilde{\Delta}_l^j}{T-t_l^j+1} \\
&= -M^{-1}\sum_{j=1}^{m}\sum_{l=1}^{k_t^j}\frac{\tilde{\xi}_l^j\mathbf{e}_j}{T-t_l^j+1},
\end{aligned}
$$

where the third equality follows by Lemma 2.5.1 part (1).

2) A uniform upper bound of $||Q_t||_2^2$. For any set $\mathcal{I} \subseteq \{1,\dots,n\}$ and any m by n matrix M, let $M^{\mathcal{I}}$ denote the submatrix of M that consists of columns $j \in \mathcal{I}$. Then, for any pair of $\mathcal{I}_1 \subseteq \{1,\dots,n\}, \mathcal{I}_2 \subseteq \{1,\dots,n\}$ I write $\mathcal{I}_1 \approx \mathcal{I}_2$ if $(A\nabla\lambda(p^D))^{\mathcal{I}_1}$ and $(A\nabla\lambda(p^D))^{\mathcal{I}_2}$ expands the same subspace of $\mathbb{R}^m$ and $|\mathcal{I}_1| = |\mathcal{I}_2|$. Note that $\mathcal{I}_1 \approx \mathcal{I}_2$ implies that there exists a unique $|\mathcal{I}_1|$ by $|\mathcal{I}_1|$ invertible matrix $Y(\mathcal{I}_1,\mathcal{I}_2)$ such that $M^{\mathcal{I}_1} = M^{\mathcal{I}_2}Y(\mathcal{I}_1,\mathcal{I}_2)$. Let $\bar{Q} := \sup\{||Q(Y(\mathcal{I}_1,\mathcal{I}_2),\mathcal{I}_2,\mathcal{I}_1)||_2^2 : \mathcal{I}_1 \approx \mathcal{I}_2\}$. Note that $\bar{Q}$ is bounded because there are only finite pairs of $\mathcal{I}_1,\mathcal{I}_2$ that satisfy $\mathcal{I}_1 \approx \mathcal{I}_2$. In addition, $\bar{Q}$ only depends on $A\nabla\lambda(p^D)$. I now claim that for any $\mathcal{B}, \gamma_{\mathcal{B}}, \gamma \in \mathbf{\Gamma}(\gamma_{\mathcal{B}}), t$ and $Q \in \mathcal{Q}_t(\gamma)$, $||Q||_2^2 \le \bar{Q}$. This is because, by definition, for any set of products $\mathcal{S}_t \subseteq \mathcal{B}$ being adjusted in period t under $\gamma_{\mathcal{B}}$, and any set of $\mathcal{G}_t$ being adjusted in period t under schedule $\gamma \in \mathbf{\Gamma}(\gamma_{\mathcal{B}})$, $\mathcal{G}_t$ is equivalent to $\mathcal{S}_t$ and there exists a set $\mathcal{G}_t' \subseteq \mathcal{G}_t$ such that $\mathcal{G}_t' \approx \mathcal{S}_t$. Without loss of generality, assume $\mathcal{G}_t'$ corresponds

to the first $|\mathcal{S}_t|$ elements in $\mathcal{G}_t$. Then construct a $|\mathcal{G}_t|$ by $|\mathcal{S}_t|$ matrix Y_t whose submatrix with rows in $\mathcal{G}'_t$ and columns in $\mathcal{S}_t$ equal $Y(\mathcal{S}_t, \mathcal{G}'_t)$ and remaining elements equal 0. Then, by optimality, I have that for any $Q_t \in \mathcal{Q}_t(\gamma)$, $||Q_t||_2^2 \le ||Q(Y_t, \mathcal{G}_t, \mathcal{S}_t)||_2^2 \le \bar{Q}$.

3) Bounding $\mathbf{E}[T-\tau]$. Note that in the proof of Theorem 2.4.1, I have (E2): $\psi > \frac{\bar{v}}{T-t}\sum_{s=1}^{t} \epsilon'_t \epsilon_t$ and (E3): $\psi > \epsilon'_t \epsilon_t$. Now, because the price deviation becomes $\tilde{\epsilon}_t$, I redefine (E2) and (E3) by replacing ϵ_t by ϵ'_t. Then the rest of the argument in the proof of Theorem 2.4.1 holds except that the argument and the bound in Lemma A.1.3 will be slightly different. In particular, the bounding of τ_2, τ_3 requires extra care. Let $q_{t_l^j}(j', j)$ denote the j'-th row j-th column element of the matrix $Q_{t_l^j}$. Then, the bound in STEP 2 of Lemma A.1.3 becomes:

$$
\begin{aligned}
\Pr(\tau_2 \le t) &= \Pr\left(\max_{v \le t} \frac{\bar{v}}{T-v} \sum_{s=1}^{v} \tilde{\epsilon}'_s \tilde{\epsilon}_s \ge \psi\right) \le \Pr\left(\frac{\bar{v}}{T-t} \sum_{s=1}^{t} \tilde{\epsilon}'_s \tilde{\epsilon}_s \ge \psi\right) \\
&= \Pr\left(\frac{\bar{v}}{T-t} \sum_{s=1}^{t} \left[\sum_{j'=1}^{n} \left(\sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \frac{q_{t_l^j}(j',j)\tilde{\xi}_l^j}{T-t_l^j+1}\right)^2\right] \ge \psi\right) \\
&\le \min\left\{1, \frac{\bar{v}}{\psi(T-t)} \sum_{s=1}^{t} \mathbf{E}\left[\sum_{j'=1}^{n} \left(\sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \frac{q_{t_l^j}(j',j)\tilde{\xi}_l^j}{T-t_l^j+1}\right)^2\right]\right\}
\end{aligned}
$$

Note that I have,

$$
\begin{aligned}
&\mathbf{E}\left[\sum_{j'=1}^{n} \left(\sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \frac{q_{t_l^j}(j',j)\tilde{\xi}_l^j}{T-t_l^j+1}\right)^2\right] \le \mathbf{E}\left[\sum_{j'=1}^{n} \left(\sum_{j=1}^{m} \left|\sum_{l=1}^{k_s^j} \frac{q_{t_l^j}(j',j)\tilde{\xi}_l^j}{T-t_l^j+1}\right|\right)^2\right] \\
\le\ & \mathbf{E}\left[\sum_{j'=1}^{n} m \sum_{j=1}^{m} \left(\sum_{l=1}^{k_s^j} \frac{q_{t_l^j}(j',j)\tilde{\xi}_l^j}{T-t_l^j+1}\right)^2\right] \\
=\ & m \sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \mathbf{E}\left[\sum_{j'=1}^{n} \frac{q_{t_l^j}^2(j',j)\left(\tilde{\xi}_l^j\right)^2}{(T-t_l^j+1)^2}\right] = m \sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \mathbf{E}\left[\frac{||Q_{t_l^j} E^j \nabla p(\lambda^D) H A \tilde{\Delta}_l^j||_2^2}{(T-t_l^j+1)^2}\right] \\
\le\ & m \sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \frac{||Q_{t_l^j} E^j \nabla p(\lambda^D) H A||_2^2 \mathbf{E}\left[||\tilde{\Delta}_l^j||_2^2\right]}{(T-t_l^j+1)^2} \le m \sum_{j=1}^{m} \sum_{l=1}^{k_s^j} \bar{Q} \left|\left|\nabla p(\lambda^D) H A\right|\right|_2^2 \frac{t_l^j - t_{l-1}^j}{(T-t_l^j+1)^2}
\end{aligned}
$$

where the first equality follows because $\forall l \neq l', \mathbf{E}[\tilde{\xi}_l^j \tilde{\xi}_{l'}^j] = 0$ by the martingale property. With the inequality above, I conclude that:

$$\begin{aligned}\sum_{t=1}^{T-1}\Pr(\tau_2\le t) &\le \sum_{j=1}^{m}\sum_{t=1}^{T}\min\left\{1,\frac{m\bar{v}\bar{Q}}{\psi}||\nabla p(\lambda^D)HA||_2^2U_2^j(T,t)\right\}\\ &\le \max\{1,\frac{m\bar{v}\bar{Q}}{\psi}\}\sum_{j=1}^{m}\sum_{t=1}^{T}\min\{1,||\nabla p(\lambda^D)HA||_2^2U_2^j(T,t)\}.\end{aligned}$$

I use a similar argument to modify STEP 3 in Lemma A.1.3.

$$\begin{aligned}\Pr(\tau_3\le t) &= \Pr\left(\max_{v\le t}\tilde{\epsilon}_v'\tilde{\epsilon}_v\ge\psi\right) = \Pr\left(\max_{v\le t}\left[\sum_{j'=1}^{n}\left(\sum_{j=1}^{m}\sum_{l=1}^{k_v^j}\frac{q_{t_l^j}(j',j)\tilde{\xi}_l^j}{T-t_l^j+1}\right)^2\right]\ge\psi\right)\\ &\le \Pr\left(\max_{v\le t}m\sum_{j'=1}^{n}\sum_{j=1}^{m}\left(\sum_{l=1}^{k_v^j}\frac{q_{t_l^j}(j',j)\tilde{\xi}_l^j}{T-t_l^j+1}\right)^2\ge\psi\right)\\ &\le \min\left\{1,\frac{m}{\psi}\mathbf{E}\left[\sum_{j'=1}^{n}\sum_{j=1}^{m}\left(\sum_{l=1}^{k_t^j}\frac{q_{t_l^j}(j',j)\tilde{\xi}_l^j}{T-t_l^j+1}\right)^2\right]\right\}\end{aligned}$$

The above inequality implies that:

$$\begin{aligned}\sum_{t=1}^{T-1}\Pr(\tau_3\le t) &\le \sum_{j=1}^{m}\sum_{t=1}^{T-1}\min\left\{1,\frac{m^2\bar{Q}}{\psi}||\nabla p(\lambda^D)HA||_2^2U_1^j(T,t)\right\}\\ &\le \max\{1,\frac{m^2\bar{Q}}{\psi}\}\sum_{j=1}^{m}\sum_{t=1}^{T-1}\min\left\{1,||\nabla p(\lambda^D)HA||_2^2U_1^j(T,t)\right\}.\end{aligned}$$

A.2 Simulation Parameters in Chapter 2

In all the experiments, I have 10 products and 4 resources. I use a multinomial logit demand (i.e., $\lambda_{t,i} = \exp(a_i - b_i p_{t,i})/(1 + \sum_{j=1}^{n} \exp(a_j - b_j p_{t,j}))$) with the following parameters:

$$
\begin{aligned}
a &= \begin{bmatrix} 0.5 & 0.4 & 0.3 & 0.4 & 0.5 & 0.3 & 0.2 & 0.4 & 0.6 & 0.8 \end{bmatrix}', \\
b &= \begin{bmatrix} 0.015 & 0.020 & 0.020 & 0.015 & 0.020 & 0.025 & 0.015 & 0.020 & 0.020 & 0.020 \end{bmatrix}', \\
A &= \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 0.1 \\ 0.1 \\ 0.1 \\ 0.1 \end{bmatrix}.
\end{aligned}
$$

A.3 Proof of Results in Chapter 3

In this section, I first provide some known results for the Maximum Likelihood Theory and the Newton's Method. I then prove the results in Chapter 3. All the proofs of other supporting lemmas which are used to prove Lemma 3.4.3, Lemma 3.5.1 and Theorem 3.5.1 are deferred to Section A.3.8.

A.3.1 Results for the Maximum Likelihood Theory and the Newton's Method

Theorem A.3.1 (Tail Inequality for MLE Based on IID Samples, Theorem 36.3 in Borovkov (1999)) *Let $\Theta \in \mathbb{R}^q$ be compact and convex, and let $\{\mathbb{P}^\theta : \theta \subseteq \Theta\}$ be a family of distributions on a discrete sample space $\mathcal{Y}$. Suppose Y is a random variable taking value in $\mathcal{Y}$ with distribution $\mathbb{P}^\theta$, and the following conditions hold:*

(i) $\mathbb{P}^\theta \neq \mathbb{P}^{\theta'}$ *whenever* $\theta \neq \theta'$;

(ii) For some $r > q$, $\sup_{\theta\in\Theta} \mathbf{E}_\theta[||\nabla_\theta \log \mathbb{P}^\theta(Y)||_2^r] = \gamma < \infty$;

(iii) The function $\theta \to \sqrt{\mathbb{P}^\theta(Y)}$ *is differentiable on* Θ *for any* $Y \in \mathcal{Y}$;

(iv) The Fisher information matrix, whose $(i,j)^{th}$ *entry is given by* $\mathbf{E}_\theta\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log \mathbb{P}^\theta(Y)\right]$, *is positive definite.*

If $Y_1, Y_2, ...$ *is a sequence of i.i.d. random variables taking value in* $\mathcal{Y}$ *with distribution* $\mathbb{P}^\theta$, *and* $\hat{\theta}(t) = \arg\max_{\theta\in\Theta} \prod_{l=1}^t \mathbb{P}^\theta(Y_l)$ *is the maximum likelihood estimate based on* t *i.i.d. samples, then, there exist constants* $\eta_1 > 0$ *and* $\eta_2 > 0$ *depending only on* r, q, $\mathbb{P}^\theta$ *and* Θ *such that for all* $t \geq 1$ *and all* $\delta \geq 0, \mathbb{P}^\theta(||\hat{\theta}(t) - \theta||_2 > \delta) \leq \eta_1 \exp(-t\eta_2\delta^2)$.

Theorem A.3.2 (Quadratic Convergence of Newton's Method for Convex Unconstrained Optimization Problems, Section 9.5.3 in Boyd and Vandenberghe (2004)) *Suppose* $g(z)$ *is a concave function whose unconstrained optimizer is* x^*. *Let* $\{x^{(k)}\}_{k=1}^\infty$ *be a sequence of points obtained by Newton's method. Assume there exist positive constants* m, M, L *such that*

(i) $||\nabla^2 g(z) - \nabla^2 g(y)||_2 \leq L||z-y||_2$, *and*

(ii) $-MI \preceq \nabla^2 g(z) \preceq -mI$.

Then, there exists constant $\eta = \min\{1, 3(1-2\alpha)\}m^2/L$ *where* $\alpha \in (0, 0.5)$ *such that if* $||\nabla g(x^{(k)})||_2 < \eta$*, then* $||\nabla g(x^{(k+1)})||_2 \leq \frac{L}{2m^2}||\nabla g(x^{(k)})||_2^2$.

A.3.2 Proof of Theorem 3.4.1

Throughout the proofs of this section, I fix $\pi = \text{PSC}$ and assume without loss of generality that $T = 1$. Let $L = \lceil \sqrt{k} \rceil$. For $k \geq 3$, the total expected revenue loss under PSC is:

$$\rho^\pi(k) \;\leq\; L\bar{r} + M_0\left[\epsilon(L)^2 k + \frac{\log k}{1 - \mathbb{P}^\pi(E(L) > \bar{\delta})} + L + \frac{1 + k\,\mathbb{P}^\pi\left(E(L) > \bar{\delta}\right)}{1 - \mathbb{P}^\pi(E(L) > \bar{\delta})}\right]$$

where the inequality follows because the revenue function in each period is bounded between 0 and $\bar{r}$ and also by Lemma 3.4.3. Since, by Lemma 3.4.2, $k\mathbb{P}^\pi(E(L) > \bar{\delta}) \leq k\eta_1 \exp(-\eta_2\bar{\delta}^2\lceil\sqrt{k}\rceil) \to 0$ as k tends to infinity, there exists a constant $K \geq 3$ such that for all $k > K$, I have $k\mathbb{P}^\pi(E(L) > \bar{\delta}) < \frac{1}{2}$ and $(1 - \mathbb{P}^\pi(E(L) > \bar{\delta}))^{-1} < 2$. So, for all $k > K$, I can bound

$$\begin{aligned}
\rho^\pi(k) \;&\leq\; \lceil\sqrt{k}\rceil\bar{r} + \frac{M_0\,\eta_3^2\,k}{\lceil\sqrt{k}\rceil} + 2M_0\log k + M_0\lceil\sqrt{k}\rceil + 3M_0 \\
&\leq\; 2\sqrt{k}\bar{r} + M_0\,\eta_3^2\,\sqrt{k} + 2M_0\sqrt{k} + 2M_0\sqrt{k} + 3M_0\sqrt{k} \\
&\leq\; (2\bar{r} + M_0\eta_3^2 + 7M_0)\sqrt{k},
\end{aligned}$$

where η_3 is as in Lemma 3.4.2. As for $k < K$, I have $\rho^\pi(k) \leq K\bar{r}$. The result of Theorem 3.4.1 then follows by letting $M_1 = \max\{K\bar{r}, 2\bar{r} + M_0\eta_3^2 + 7M_0\}$. This completes the proof.

A.3.3 Proof of Lemma 3.4.1

I will prove each part of the lemma in turn. Let $\bar{\delta} = \min\{\delta_1, \delta_2\}$ where δ_1 and δ_2 are strictly positive constants to be defined shortly.

Proof of part (a). This is an immediate corollary of Proposition 4.4.1. Note that, by assumption P2, I have $||\lambda(p;\theta^*) - \lambda(p;\theta)||_\infty \leq ||\lambda(p;\theta^*) - \lambda(p;\theta)||_2 \leq \omega||\theta^* - \theta||_2$ and $||(\nabla\lambda(p;\theta^*) - \nabla\lambda(p;\theta))'||_\infty = \max_{1\leq i\leq n}\sum_{j=1}^n |\frac{\partial\lambda_i}{\partial p_j}(p;\theta) - \frac{\partial\lambda_i}{\partial p_j}(p;\theta^*)| \leq n\omega||\theta^* - \theta||_2$ for all $\theta \in \Theta, p \in \mathcal{P}$. Hence, $||\lambda(.;\theta^*) - \lambda(.;\theta)||_\infty = \sup_{p\in\mathcal{P}}||\lambda(p;\theta^*) - \lambda(p;\theta)||_\infty \leq \omega||\theta^* - \theta||_2$ and $||(\nabla\lambda(.;\theta^*) - \nabla\lambda(.;\theta))'||_\infty = \sup_{p\in\mathcal{P}}||(\nabla\lambda(.;\theta^*) - \nabla\lambda(.;\theta))'||_\infty \leq n\omega||\theta^* - \theta||_2$. Therefore, by Proposition 4.4.1, $||p^D(\theta^*) - p^D(\theta)||_\infty \leq nM_6\omega||\theta^* - \theta||_2$. Let $\delta_1 = \phi(2n^{3/2}M_6\omega)^{-1}$. For all θ satisfying $||\theta - \theta^*||_2 \leq \bar{\delta} \leq \delta_1$, I have $||p^D(\theta^*) - p^D(\theta)||_2 \leq n^{1/2}||p^D(\theta^*) - p^D(\theta)||_\infty \leq n^{3/2}M_6\omega\delta_1 \leq \phi/2$. Hence, $p^D(\theta) \in \mathsf{Ball}(p^D(\theta^*), \phi/2)$. Since $\mathsf{Ball}(p^D(\theta^*), \phi) \subseteq \mathcal{P}$ by P5, I conclude that $\mathsf{Ball}(p^D(\theta), \phi/2) \subseteq \mathcal{P}$.

Since $\lambda(.;\theta^*)$ is continuously differentiable with respect to $p \in \mathcal{P}$ as implied by P1, and $\mathcal{P}$ is compact, there exists a constant $K > 0$ independent of $k > 0$ such that

$$\begin{aligned}
||\lambda^D(\theta^*) - \lambda^D(\theta)||_2 &= ||\lambda(p^D(\theta^*);\theta^*) - \lambda(p^D(\theta);\theta)||_2 \\
&\leq ||\lambda(p^D(\theta^*);\theta^*) - \lambda(p^D(\theta);\theta^*)||_2 + ||\lambda(p^D(\theta);\theta^*) - \lambda(p^D(\theta);\theta)||_2 \\
&\leq K||p^D(\theta^*) - p^D(\theta)||_2 + \omega||\theta^* - \theta||_2 \\
&\leq (\omega + n^{3/2}KM_6\omega)||\theta^* - \theta||_2,
\end{aligned}$$

where the second inequality also follows by P2. The result follows by letting $\kappa = \omega + n^{3/2}KM_6\omega$. Part (a) is proved.

Proof of part (b). Since $\mathrm{P}_\lambda(\theta)$ is a convex program for all $\theta \in \Theta$, by the Karush-Kuhn-Tucker optimality condition, $\nabla_\lambda r(\lambda^D(\theta);\theta) = A'\mu^D(\theta)$. By my assumption, A has full row rank. Thus, there exists some m by n matrix $\bar{A}$ such that $\mu^D(\theta) = \bar{A}\nabla_\lambda r(\lambda^D(\theta);\theta)$. Since the right hand side is continuous at θ^*, I conclude that $\mu^D(.)$ is continuous at θ^* as well. Part (b) is proved.

Proof of part (c). Let $\underline{\mu} = \min_{1\leq i\leq n}\{\mu_i^D(\theta^*) : \mu_i^D(\theta^*) > 0\}$. Since $\mu^D(.)$ is continuous at θ^* by part (b), there exists $\delta_2 > 0$ such that $||\mu^D(\theta) - \mu^D(\theta^*)||_2 < \underline{\mu}$ for all $\theta \in \mathsf{Ball}(\theta^*, \delta_2)$. This means for $\theta \in \mathsf{Ball}(\theta^*, \bar{\delta})$, I also have $\mu_i^D(\theta) > 0$ whenever $\mu_i^D(\theta^*) > 0$, which implies that the corresponding constraints in $\mathrm{P}(\theta)$ are binding due to Karush-Kuhn-Tucker condition. Part (c) is proved.

A.3.4 Proof of Lemma 3.4.2

The proof of Lemma 3.4.2 is a multiproduct extension of Lemma 3.7 in Broder and Rusmevichientong (2012), and is based on a well-known result in Maximum Likelihood Theory. I state this result in Theorem A.3.1 (see Section A.3.1).

To apply Theorem A.3.1 to my setting, I simply need to verify conditions (i)-(iv). First, note that Θ is a compact subset of $\mathbb{R}^q$ and $\mathcal{D}^{\tilde{q}}$ is a discrete-valued sample space. Conditions (i) and (iv), they are immediately satisfied because of S1 and S3. As for conditions (ii) and (iii), recall that

$$\left|\left|\nabla_\theta \log \mathbb{P}^{\tilde{p},\theta}(D_{1:\tilde{q}})\right|\right|_2 = \left|\left|\sum_{s=1}^{\tilde{q}} \left[\left(1-\sum_{j=1}^n D_{s,j}\right)\nabla_\theta \log\left(1-\sum_{j=1}^n \lambda_j(\tilde{p}^{(s)};\theta)\right) + \sum_{j=1}^n D_{s,j}\nabla_\theta \log \lambda_j(\tilde{p}^{(s)};\theta)\right]\right|\right|_2$$
$$\leq \sum_{s=1}^{\tilde{q}} \left(\left|\left|\nabla_\theta \log\left(1-\sum_{j=1}^n \lambda_j(\tilde{p}^{(s)};\theta)\right)\right|\right|_2 + \sum_{j=1}^n \left|\left|\nabla_\theta \log \lambda_j(\tilde{p}^{(s)};\theta)\right|\right|_2\right).$$

By Assumption P1 and S2, for all $1 \leq s \leq \tilde{q}$ and $1 \leq j \leq n$, $\lambda_j(\tilde{p}^{(s)};.) \in \mathsf{C}^1(\Theta)$ and is bounded away from zero, and $\sum_{j=1}^n \lambda_j(\tilde{p}^{(s)};.) \in \mathsf{C}^1(\Theta)$ is also bounded away from one. These imply that $||\nabla_\theta \log\left(1-\sum_{j=1}^n \lambda_j(\tilde{p}^{(s)};.)\right)||_2$ and $||\nabla_\theta \log \lambda_j(\tilde{p}^{(s)};.)||_2$, $j = 1,\ldots,n$, are both continuous functions of θ for $s = 1,\ldots,\tilde{q}$ and are, due to compactness of Θ, bounded. So, (ii) follows. As for (iii), note that $\mathbb{P}^{\tilde{p},\theta}(D_{1:\tilde{q}})$ is continuous in θ and it is also bounded away from zero. (In fact, $\mathbb{P}^{\tilde{p},\theta}(D_{1:\tilde{q}}) \geq [\lambda_{\min}^n(1-\lambda_{\max})]^{\tilde{q}}$ by S2.) So, $\theta \to \sqrt{\mathbb{P}^{\tilde{p},\theta}(D_{1:\tilde{q}})}$ is differentiable on Θ for all $D_{1:\tilde{q}} \in \mathcal{D}^{\tilde{q}}$. I have thus verified all the conditions of Theorem A.3.1.

I will now use Theorem A.3.1 to prove Lemma 3.4.2. A direct application of Theorem A.3.1 leads to $\mathbb{P}^\pi(E(L) > \delta) \leq \eta_1 \exp(-\eta_2 L\delta^2)$. Also, since $\epsilon(L)^2 = \mathbf{E}^\pi\left[E(L)^2\right] = \int_0^\infty \mathbb{P}^\pi(E(L)^2 \geq x)dx = \int_0^\infty \mathbb{P}^\pi(E(L) \geq \sqrt{x})dx \leq \int_0^\infty \eta_1 e^{-\eta_2 Lx}dx = \eta_1/(\eta_2 L)$, the result follows by taking $\eta_3 = \sqrt{\eta_1/\eta_2}$.

A.3.5 Proof of Lemma 3.4.3

Fix $\pi = \text{PSC}$. Without loss of generality, I assume that $T = 1$. Let $\mathcal{A}$ denote the event that $E(L) \leq \bar{\delta}$. I first define a stopping time and show some useful properties of this stopping time on the event of $\mathcal{A}$. Let $\lambda_L > 0$ be such that $A\lambda_L\mathbf{e} \prec C$. Define $\psi = \frac{\min\{\phi, 2\lambda_L\}}{\max\{2, 4\omega||Q||_2\}}$ and define the cumulative demand at the end of period t as $S_t := \sum_{s=1}^t D_s$. Let τ be the minimum of k and the first time $t \geq L+1$ the following condition is violated:

$$(C1) \qquad \psi > \left|\left|\sum_{s=L+1}^t \frac{\hat{\Delta}_s}{k-s}\right|\right|_2 + \left|\left|\frac{S_L - L\lambda_L\mathbf{e}}{k-t}\right|\right|_2.$$

Let C_t denote the available capacity level at the end of period t. I denote by $\hat{\lambda}_t := \lambda^D(\hat{\theta}_L) - \sum_{s=L+1}^{t-1} \frac{Q\hat{\Delta}_s}{k-s}$ the demand rate that the seller *believes* he induces in period t, and by $\lambda_t := \lambda(p(\hat{\lambda}_t;\hat{\theta}_L);\theta^*)$ the actual induced demand rate upon applying price $p(\hat{\lambda}_t;\hat{\theta}_L)$ in period t. Note that, by definition, I can also write $\hat{\lambda}_t = \lambda(p(\hat{\lambda}_t;\hat{\theta}_L);\hat{\theta}_L)$. I state two useful lemmas.

Lemma A.3.1 *For sample paths in $\mathcal{A}$, I have $C_t \succ 0$ and $\hat{\lambda}_t \in \Lambda_{\hat{\theta}_L}$ for all $L+1 \leq t < \tau$.*

Lemma A.3.2 *There exists $K_0 > 0$ independent of $k \geq 3$ such that for all $k \geq 3$*

$$\mathbf{E}^\pi[k - \tau|\mathcal{A}] \;\leq\; K_0 \left[\frac{\log k}{1 - \mathbb{P}^\pi(E(L) > \bar{\delta})} + \epsilon(L)^2 k + L\right].$$

Lemma A.3.1 essentially says that, on $\mathcal{A}$, the remaining capacity C_t is always positive and the price $p(\hat{\lambda}_t; \hat{\theta}_L)$ is always feasible before τ, and Lemma A.3.2 establishes a bound for the expected remaining time after τ. Define $r^D(\theta^*) := r(\lambda^D(\theta^*); \theta^*)$ and let R_t^π denote the revenue earned during period t under policy π. Let $\bar{\Delta}_t := R_t^\pi - r(\lambda_t; \theta^*)$. I have:

$$\begin{aligned}
&\sum_{t=L+1}^{k} r^D(\theta^*) - \mathbf{E}^\pi\left[\hat{R}^\pi(k)\right] \\
&= \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left(r^D(\theta^*) - R_t^\pi\right) + \sum_{t=\tau}^{k}\left(r^D(\theta^*) - R_t^\pi\right)\right] \\
&= \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left(r^D(\theta^*) - r(\lambda_t;\theta^*)\right) + \sum_{t=\tau}^{k}\left(r^D(\theta^*) - R_t^\pi\right)\right] - \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\bar{\Delta}_t\right] \\
&\leq \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left(r^D(\theta^*) - r(\lambda_t;\theta^*)\right) + \sum_{t=\tau}^{k}\left(r^D(\theta^*) - R_t^\pi\right)\middle|\mathcal{A}\right]\mathbb{P}^\pi(\mathcal{A}) + \bar{r}k\mathbb{P}^\pi(\mathcal{A}^c) - \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\bar{\Delta}_t\right] \\
&\leq \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left(r^D(\theta^*) - r(\lambda_t;\theta^*)\right) + \sum_{t=\tau}^{k}\left(r^D(\theta^*) - R_t^\pi\right)\middle|\mathcal{A}\right] + \bar{r}k\mathbb{P}^\pi(\mathcal{A}^c) + \bar{r} \\
&= \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left(r^D(\theta^*) - r(\lambda_t;\theta^*)\right) + \sum_{t=\tau}^{k}\left(r^D(\theta^*) - R_t^\pi\right)\middle|\mathcal{A}\right] + \bar{r} + \bar{r}k\mathbb{P}^\pi(E(L) > \bar{\delta}). \qquad (54)
\end{aligned}$$

The first inequality follows because $\bar{r}$ is the upper bound on revenue rate for each period, which is also the maximum possible revenue loss for a single period on average. As for the second inequality, note that $\{\bar{\Delta}_t\}_{t=L+1}^{k-1}$ is a martingale difference sequence with respect to $\{\mathcal{H}_t\}_{t=L+1}^{k-1}$. Thus, by the Optional Stopping Time Theorem, I have $-\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\bar{\Delta}_t\right] = -\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau}\bar{\Delta}_t\right] + \mathbf{E}^\pi\left[\bar{\Delta}_\tau\right] \leq \bar{r}$, so the second inequality holds. I now analyze the first two terms in (54). By Taylor expansion and P3,

$$\begin{aligned}
&\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left(r^D(\theta^*) - r(\lambda_t;\theta^*)\right) + \sum_{t=\tau}^{k}\left(r^D(\theta^*) - R_t^\pi\right)\middle|\mathcal{A}\right] + \bar{r} \\
&\quad \leq \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\nabla r(\lambda^D(\theta^*);\theta^*)\cdot(\lambda^D(\theta^*) - \lambda_t)\middle|\mathcal{A}\right] + \frac{\bar{v}}{2}\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}||\lambda^D(\theta^*) - \lambda_t||_2^2\middle|\mathcal{A}\right] \\
&\quad\quad + \bar{r}\left(\mathbf{E}^\pi[k-\tau|\mathcal{A}] + 2\right) \qquad (55)
\end{aligned}$$

By Lemma A.3.1, $\hat{\lambda}_t = \lambda^D(\hat{\theta}_L) - Q\sum_{s=L+1}^{t-1} \frac{\hat{\Delta}_s}{k-s} \in \Lambda_{\hat{\theta}_L}$ before τ. Also, recall that, by definition, $\hat{\Delta}_t = D_t - \hat{\lambda}_t = \Delta_t + \lambda_t - \hat{\lambda}_t$. So, I can write the first term in (55) as follows:

$$\begin{aligned}
&\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1} \nabla r(\lambda^D(\theta^*);\theta^*)\cdot(\lambda^D(\theta^*)-\lambda_t)\middle|\mathcal{A}\right] \\
&\quad = \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1} \mu^D(\theta^*)' A\left(\lambda^D(\theta^*)-\hat{\lambda}_t+\hat{\lambda}_t-\lambda_t\right)\middle|\mathcal{A}\right] \\
&\quad = \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1} \mu^D(\theta^*)'\left(A\lambda^D(\theta^*)-A\lambda^D(\hat{\theta}_L)+\sum_{s=L+1}^{t-1}\frac{A\hat{\Delta}_s}{k-s}+A\Delta_t-A\hat{\Delta}_t\right)\middle|\mathcal{A}\right] \\
&\quad = \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1} \mu^D(\theta^*)'\left(A\lambda^D(\theta^*)-A\lambda^D(\hat{\theta}_L)\right)\middle|\mathcal{A}\right] \\
&\qquad +\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1} \mu^D(\theta^*)'\left(\sum_{s=L+1}^{t-1}\frac{A\hat{\Delta}_s}{k-s}+A\Delta_t-A\hat{\Delta}_t\right)\middle|\mathcal{A}\right]. \qquad (56)
\end{aligned}$$

By Lemma 3.4.1(c), for all sample paths on $\mathcal{A}$, the set of constraints of $\mathrm{P}(\theta^*)$ that have nonzero optimal dual variables also have nonzero optimal dual variables in $\mathrm{P}(\hat{\theta}_L)$ and are thus binding at the optimal solution $\lambda^D(\hat{\theta}_L)$. This implies that the first expectation in (56) is zero because, for all i, either I have $\mu_i^D(\theta^*) = 0$ or $(A\lambda^D(\theta^*))_i - (A\lambda^D(\hat{\theta}_L))_i = 0$. As for the second term of (56), I can further write:

$$\begin{aligned}
&\mathbf{E}^\pi\left[\mu^D(\theta^*)'\sum_{t=L+1}^{\tau-1}\left(\sum_{s=L+1}^{t-1}\frac{A\hat{\Delta}_s}{k-s}+A\Delta_t-A\hat{\Delta}_t\right)\middle|\mathcal{A}\right] \\
&\quad = \mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\mu^D(\theta^*)'A\Delta_t\middle|\mathcal{A}\right]+\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left(\frac{\tau-t-1}{k-t}-1\right)\mu^D(\theta^*)'A\hat{\Delta}_t\middle|\mathcal{A}\right].
\end{aligned}$$

Since $\{\Delta_t\}_{t=L+1}^{k-1}$ is a martingale difference sequence with respect to $\{\mathcal{H}_t\}_{t=L+1}^{k-1}$, I can bound:

$$\begin{aligned}
&\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\mu^D(\theta^*)'A\Delta_t\middle|\mathcal{A}\right] \\
&\quad = \frac{\mu^D(\theta^*)'A}{\mathbb{P}^\pi(\mathcal{A})}\left\{\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\Delta_t\right]-\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\Delta_t\middle|\mathcal{A}^c\right]\mathbb{P}^\pi(\mathcal{A}^c)\right\} \\
&\quad \leq \mu^D(\theta^*)'A\mathbf{e}\,\frac{1+k\mathbb{P}^\pi(E(L)>\bar{\delta})}{1-\mathbb{P}^\pi(E(L)>\bar{\delta})},
\end{aligned}$$

where the inequality follows because $\mathbf{E}^\pi[\sum_{t=L+1}^{\tau-1}\Delta_t] = \mathbf{E}^\pi[\sum_{t=L+1}^{\tau}\Delta_t] - \mathbf{E}^\pi[\Delta_\tau] \prec \mathbf{e}$ (by Optional Stopping Time Theorem) and the fact that $|\Delta_t| \prec \mathbf{e}$. As for the second term, note that, by (C1) in the definition of τ,

$$\begin{aligned}
&\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left(\frac{\tau-t-1}{k-t}-1\right)\mu^D(\theta^*)'A\hat{\Delta}_t\,\middle|\,\mathcal{A}\right] \\
&\quad\leq\ \mathbf{E}^\pi\left[(k-\tau+1)\left|\mu^D(\theta^*)'\sum_{t=L+1}^{\tau-1}\frac{A\hat{\Delta}_t}{k-t}\right|\,\middle|\,\mathcal{A}\right] \\
&\quad\leq\ \mathbf{E}^\pi\left[(k-\tau+1)\left|\left|\mu^D(\theta^*)\right|\right|_2||A||_2\left|\left|\sum_{t=L+1}^{\tau-1}\frac{\hat{\Delta}_t}{k-t}\right|\right|_2\,\middle|\,\mathcal{A}\right] \\
&\quad\leq\ \psi\left|\left|\mu^D(\theta^*)\right|\right|_2||A||_2\left(\mathbf{E}^\pi\left[k-\tau|\,\mathcal{A}\right]+1\right).
\end{aligned}$$

Putting this together with Lemma A.3.2, for the first term in (55), I have:

$$\begin{aligned}
&\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\nabla_\lambda r(\lambda^D(\theta^*);\theta^*)\cdot(\lambda^D(\theta^*)-\lambda_t)\,\middle|\,\mathcal{A}\right] \\
&\quad\leq\ K_1\left[\frac{\log k}{1-\mathbb{P}^\pi(E(L)>\bar{\delta})}\ +\ \epsilon(L)^2k\ +\ L\ +\ \frac{1+k\mathbb{P}^\pi(E(L)>\bar{\delta})}{1-\mathbb{P}^\pi(E(L)>\bar{\delta})}\right],
\end{aligned}$$

where the constant $K_1 = \mu^D(\theta^*)'A\mathbf{e} + (1+K_0)\psi||\mu^D(\theta^*)||_2||A||_2$ is independent of $k \geq 3$.

I now bound the second term in (55). Note that

$$\begin{aligned}
&\frac{\bar{v}}{2}\,\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}||\lambda^D(\theta^*)-\lambda_t||_2^2\,\middle|\,\mathcal{A}\right] \\
&\quad\leq\ \bar{v}\,\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\hat{\lambda}_t-\lambda_t\right|\right|_2^2\,\middle|\,\mathcal{A}\right]\ +\ \bar{v}\,\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\lambda^D(\theta^*)-\hat{\lambda}_t\right|\right|_2^2\,\middle|\,\mathcal{A}\right].
\end{aligned}\tag{57}$$

Since $\lambda_t = \lambda(p(\hat{\lambda}_t;\hat{\theta}_L);\theta^*)$ and $\hat{\lambda}_t = \lambda(p(\hat{\lambda}_t;\hat{\theta}_L);\hat{\theta}_L)$, by P2, I have

$$\bar{v}\mathbf{E}^\pi[\sum_{t=L+1}^{\tau-1}||\hat{\lambda}_t-\lambda_t||_2^2|\mathcal{A}]\ \leq\ \bar{v}\omega^2k\,\mathbf{E}^\pi[||\theta^*-\hat{\theta}_L||_2^2\,|\,\mathcal{A}]\ \leq\ \bar{v}\omega^2k\,\mathbf{E}^\pi[||\theta^*-\hat{\theta}_L||_2^2]\ \leq\ \bar{v}\omega^2\epsilon(L)^2k.$$

(By definition of $\mathcal{A}$, $\mathbf{E}^\pi[||\theta^* - \hat{\theta}_L||_2^2 \,|\, \mathcal{A}] \leq \mathbf{E}^\pi[||\theta^* - \hat{\theta}_L||_2^2]$.) As for the second term in (57),

$$\begin{aligned}
&\bar{v}\,\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\lambda^D(\theta^*) - \hat{\lambda}_t\right|\right|_2^2 \,\middle|\, \mathcal{A}\right] \\
&\quad = \bar{v}\,\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\lambda^D(\theta^*) - \lambda^D(\hat{\theta}_L) + Q\sum_{s=L+1}^{t-1}\frac{\hat{\Delta}_s}{k-s}\right|\right|_2^2 \,\middle|\, \mathcal{A}\right] \\
&\quad \leq 2\bar{v}k\mathbf{E}^\pi\left[\left|\left|\lambda^D(\theta^*) - \lambda^D(\hat{\theta}_L)\right|\right|_2^2 \,\middle|\, \mathcal{A}\right] + 2\bar{v}\,||Q||_2^2\,\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\sum_{s=L+1}^{t-1}\frac{\hat{\Delta}_s}{k-s}\right|\right|_2^2 \,\middle|\, \mathcal{A}\right] \\
&\quad \leq 2\bar{v}\kappa^2\epsilon(L)^2 k + 2\bar{v}\,||Q||_2^2\,\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\sum_{s=L+1}^{t-1}\frac{\hat{\Delta}_s}{k-s}\right|\right|_2^2 \,\middle|\, \mathcal{A}\right] \qquad (58)
\end{aligned}$$

where the second inequality follows by Lemma 3.4.1(a). Using $\hat{\Delta}_t = D_t - \hat{\lambda}_t = \Delta_t + \lambda_t - \hat{\lambda}_t$, I can bound the second term in (58) as follows:

$$\begin{aligned}
&\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\sum_{s=L+1}^{t-1}\frac{\hat{\Delta}_s}{k-s}\right|\right|_2^2 \middle| \mathcal{A}\right] \\
&\quad \leq 2\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\sum_{s=L+1}^{t-1}\frac{\Delta_s}{k-s}\right|\right|_2^2 \middle| \mathcal{A}\right] + 2\mathbf{E}^\pi\left[\sum_{t=L+1}^{\tau-1}\left|\left|\sum_{s=L+1}^{t-1}\frac{\lambda_s - \hat{\lambda}_s}{k-s}\right|\right|_2^2 \middle| \mathcal{A}\right] \\
&\quad \leq \frac{2}{\mathbb{P}^\pi(\mathcal{A})}\mathbf{E}^\pi\left[\sum_{t=L+1}^{k-1}\left|\left|\sum_{s=L+1}^{t-1}\frac{\Delta_s}{k-s}\right|\right|_2^2\right] + 2\mathbf{E}^\pi\left[\sum_{t=L+1}^{k-1}\left(\sum_{s=L+1}^{t-1}\frac{\omega E(L)}{k-s}\right)^2 \middle| \mathcal{A}\right] \\
&\quad \leq \frac{2}{\mathbb{P}^\pi(\mathcal{A})}\mathbf{E}^\pi\left[\sum_{t=L+1}^{k-1}\left|\left|\sum_{s=L+1}^{t-1}\frac{\Delta_s}{k-s}\right|\right|_2^2\right] + 2\sum_{t=L+1}^{k-1}\left[\sum_{s=L+1}^{t-1}\frac{\sqrt{\mathbf{E}^\pi\left[\omega^2 E(L)^2 | \mathcal{A}\right]}}{k-s}\right]^2 \\
&\quad \leq \frac{2}{\mathbb{P}^\pi(\mathcal{A})}\mathbf{E}^\pi\left[\sum_{t=L+1}^{k-1}\sum_{s=L+1}^{t-1}\frac{||\Delta_s||_2^2}{(k-s)^2}\right] + 2\sum_{t=L+1}^{k-1}\left(\sum_{s=L+1}^{t-1}\frac{\omega\epsilon(L)}{k-s}\right)^2 \\
&\quad \leq \frac{16}{1 - \mathbb{P}^\pi(E(L) > \bar{\delta})}\log k + 6\omega^2\epsilon(L)^2 k \qquad (59)
\end{aligned}$$

where the second inequality holds by the law of total expectation and P2, the third inequality follows by first expanding the square of the sum and then applying Cauchy-Swartz inequality to the cross-terms, the fourth inequality follows by the orthogonality of martingale differences $\{\Delta_s\}$ and $\mathbf{E}^\pi[E(L)^2|\mathcal{A}] \leq \epsilon(L)^2$, and the fifth inequality holds by integral approximation. In particular, the first term after the fourth inequality can be bounded using $||\Delta_s||_2 = ||D_s -$

$\lambda_s||_2 \le ||D_s||_2 + ||\lambda_s||_2 \le 2$ and $\sum_{t=L+1}^{k-1} \sum_{s=L+1}^{t-1} \frac{1}{(k-s)^2} \le \sum_{t=L+1}^{k-1} \frac{1}{k-t} \le 1 + \log k \le 2 \log k$ (recall that $k \ge 3$) whereas the second term can be bounded using the following integral comparison:

$$\sum_{t=L+1}^{k-1} \left(\sum_{s=L+1}^{t-1} \frac{1}{k-s} \right)^2 \le \sum_{t=1}^{k-1} \left(\sum_{s=1}^{t-1} \frac{1}{k-s} \right)^2 \le \sum_{t=1}^{k-1} \left(\int_1^t \frac{1}{k-s} ds \right)^2 \le \sum_{t=1}^{k-1} \log^2 \left(\frac{k}{k-t} \right)$$
$$\le \log^2 k + \int_1^{k-1} \log^2 \left(\frac{k}{k-t} \right) dt \le \log^2 k + 2k \le 3k, \tag{60}$$

where the last inequality follows because $\log^2 k < k$ for $k \ge 1$.

Thus, for the second term in (55), I have

$$\frac{\bar{v}}{2} \mathbf{E}^\pi \left[\sum_{t=L+1}^{k-1} ||\lambda^D(\theta^*) - \lambda_t||_2^2 \middle| \mathcal{A} \right] \le K_2 \left[\frac{\log k}{1 - \mathbb{P}^\pi(E(L) > \bar{\delta})} + \epsilon(L)^2 k \right],$$

where $K_2 = \bar{v}\omega^2 + 2\bar{v}\kappa^2 + 32\bar{v}||Q||_2^2 + 12\omega^2\bar{v}||Q||_2^2$. Combining all results together, I conclude that

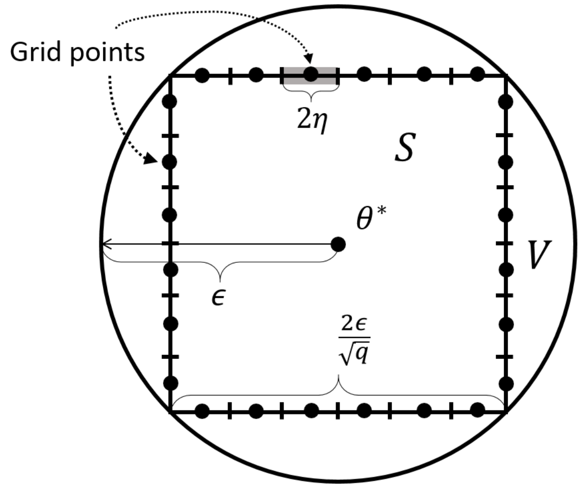
$$\sum_{t=L+1}^{k} r^D(\theta^*) - \mathbf{E}^\pi \left[\hat{R}^\pi(k) \right] \le M_0 \left[\frac{\log k}{1 - \mathbb{P}^\pi(E(L) > \bar{\delta})} + \epsilon(L)^2 k + L + \frac{1 + k\,\mathbb{P}^\pi \left(E(L) > \bar{\delta} \right)}{1 - \mathbb{P}^\pi(E(L) > \bar{\delta})} \right],$$

where $M_0 = K_1 + K_2 + \bar{r}K_0 + 3\bar{r}$. (Note that the last term in (54) can be bounded by $\bar{r}(1 + k\mathbb{P}^\pi(E(L) > \bar{\delta}))/(1 - \mathbb{P}^\pi(E(L) > \bar{\delta}))$.) This completes the proof of Lemma 3.4.3.

A.3.6 Proof of Lemma 3.5.1

I first illustrate the idea using Figure A.2. Note that $E(t) > \epsilon$ is equivalent to the event that ML estimator $\hat{\theta}_t$ is in the outside of the ball $V := \mathsf{Ball}(\theta^*, \epsilon)$. In addition, under the concavity assumption of the log-likelihood, $\hat{\theta}_t \neq \mathsf{Ball}(\theta^*, \epsilon)$ implies that at least one point on the surface of a hypercube S, which is centered at θ^* and is a subset of V, has a larger log-likelihood than the log-likelihood at θ^*. The probability of this event is a valid upper bound of $\mathbb{P}^\pi(E(t) > \epsilon)$. However, the challenge is that there are a continuum of such potential points. The idea of the proof is to consider a grid of points on the surface of that hypercube S, and the granularity of the grid is set to be fine enough so that any point on the surface of that hypercube can be closely approximated by one point on the grid. I will show that the existence of a point on the surface of S with a higher log-likelihood than the true parameter vector θ^* is extremely unlikely. I now rigorously prove this lemma.

Figure A.2: Geometric illustration of Lemma 3.5.1



Note: This illustrates the case when there are two parameters to estimate ($q = 2$). V denotes the disk (ball) centered at θ^* with radius ϵ. Note that the event of $||\theta^* - \hat{\theta}_t||_2 > \epsilon$ corresponds to the event when $\hat{\theta}_t$ lies in the exterior of V. In this example, the surface of the rectangle(hypercube) S consists of four edges.

Step 1

Fix some $0 < \tilde{\lambda}_{\min} < \tilde{\lambda}_{\max} < 1$. First, I will show that for all $D \in \mathcal{D}$, for all $p \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ and for all $\theta \in \Theta$, $\nabla_\theta \log \mathbb{P}_1^{p,\theta}(D)$ is jointly continuous in θ and p. Recall that $\nabla_\theta \log \mathbb{P}_1^{p,\theta}(D) = ((\partial/\partial\theta_1) \log \mathbb{P}_1^{p,\theta}(D); \ldots; (\partial/\partial\theta_n) \log \mathbb{P}_1^{p,\theta}(D))$ where for all $1 \leq k \leq n$,

$$\begin{aligned}\frac{\partial \log \mathbb{P}_1^{p,\theta}(D)}{\partial \theta_k} &= -\frac{(1 - \sum_{j=1}^n D_j) \log \left(1 - \sum_{j=1}^n \lambda_j(p;\theta)\right)}{1 - \sum_{j=1}^n \lambda_j(p;\theta)} \left(\sum_{j=1}^n \frac{\partial \lambda_j(p;\theta)}{\partial \theta_k}\right) \\ &\quad + \sum_{j=1}^n \frac{D_j \log(\lambda_j(p;\theta))}{\lambda_j(p;\theta)} \frac{\partial \lambda_j(p;\theta)}{\partial \theta_k}.\end{aligned}$$

Since $\lambda_j(p;.) \in \mathsf{C}^1(\Theta)$ and $\lambda(.;\theta) \in \mathsf{C}^2(\mathcal{P})$ by P1 and the denominators are strictly greater than zero, $\nabla_\theta \log \mathbb{P}_1^{p,\theta}(D)$ is jointly continuous in θ and p.

Step 2

Since Θ and $\mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ are compact, $\mathcal{D}$ is finite and $\nabla_\theta \log \mathbb{P}_1^{p,\theta}(D)$ is jointly continuous in θ and p for all $D \in \mathcal{D}$, there exists a constant $c_g > 0$ independent of θ, p, D such that for all $\theta \in \Theta$, $p \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$, and $v \in \mathbb{R}^q$ satisfying $||v||_2 = 1$, $\nabla_\theta \log \mathbb{P}_1^{p,\theta}(D) \cdot v < c_g$. Therefore, for any $v, ||v||_2 = 1$, if $p_s^\pi \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ for $1 \leq s \leq t$, then I have:

$$\nabla_\theta \log \mathbb{P}_t^{\pi,\theta}(D_{1:t}) \cdot v = \sum_{s=1}^t \nabla_\theta \log \mathbb{P}_1^{p_s^\pi,\theta}(D_s) \cdot v < c_g t. \tag{61}$$

Now, fix $\epsilon > 0$ and consider a hypercube $S \in \mathbb{R}^q$ centered at the origin with edge $2\epsilon/\sqrt{q}$. Let ∂S denote the surface of S, its area is given by $c_q(\epsilon/\sqrt{q})^{q-1}$ for a constant c_q that depends only on q. Cover ∂S with a set of identical hypercubes in $\mathbb{R}^{q-1}$ with edge 2η (see Figure A.2 for an illustration) and denote by N the number of cubes needed to cover ∂S. Then, $N = (\epsilon/(\sqrt{q}\eta))^{q-1}$. Let $v_j \in \partial S, j = 1, \dots, N$ denote the center of those $2\eta-$cubes. These points constitute a set of grid points on the surface. Then for any $x \in \partial S$, $\min_{j=1,\dots,N} ||x - v_j||_2 \leq \sqrt{q}\eta$. By W3, I have that for any $\theta' \in S + \theta^*$ and any $j = 1, \dots, N$,

$$\log \mathbb{P}_t^{\pi,\theta'}(D_{1:t}) - \log \mathbb{P}_t^{\pi,\theta^*+v_j}(D_{1:t}) \leq \nabla_\theta \log \mathbb{P}_t^{\pi,\theta^*+v_j}(D_{1:t}) \cdot (\theta' - \theta^* - v_j)$$

Let $j^*(\theta) = \arg\min_{j=1,\dots,N} ||\theta - \theta^* - v_j||_2$. I then have

$$\log \mathbb{P}_t^{\pi,\theta'}(D_{1:t}) - \log \mathbb{P}_t^{\pi,\theta^*+v_{j^*(\theta')}}(D_{1:t}) \leq c_g t ||\theta' - \theta^* - v_{j^*(\theta')}||_2 \leq c_g \sqrt{q}\eta t. \tag{62}$$

where the first inequality follows by (61). The following is the key argument for this proof:

$$\begin{aligned}
&\left\{||\hat{\theta}_t - \theta^*||_2 > \epsilon\right\} \\
&\quad\subseteq \left\{||\hat{\theta}_t - \theta^*||_\infty > \frac{\epsilon}{\sqrt{q}}\right\} \\
&\quad\subseteq \left\{\log \mathbb{P}_t^{\pi,\theta^*+v}(D_{1:t}) \geq \log \mathbb{P}_t^{\pi,\theta^*}(D_{1:t}), \text{for some } v \text{ with } ||v||_\infty = \tfrac{\epsilon}{\sqrt{q}}\right\} \\
&\quad\subseteq \left\{\log \mathbb{P}_t^{\pi,\theta^*+v_{j^*(\theta^*+v)}}(D_{1:t}) + c_g\sqrt{q}\eta t \geq \log \mathbb{P}_t^{\pi,\theta^*}(D_{1:t}), \text{for some } v \text{ with } ||v||_\infty = \tfrac{\epsilon}{\sqrt{q}}\right\} \\
&\quad\subseteq \cup_{j=1}^N \left\{\log \mathbb{P}_t^{\pi,\theta^*+v_j}(D_{1:t}) + c_g\sqrt{q}\eta t \geq \log \mathbb{P}_t^{\pi,\theta^*}(D_{1:t})\right\} \\
&\quad= \cup_{j=1}^N \left\{Z_t^\pi(v_j, D_{1:t}) \geq \exp(-c_g\sqrt{q}\eta t)\right\},
\end{aligned}$$

where $Z_t^\pi(u, D_{1:t}) := \mathbb{P}_t^{\pi,\theta^*+u}(D_{1:t})/\mathbb{P}_t^{\pi,\theta^*}(D_{1:t})$ is the likelihood ratio for any $u \in \Theta - \theta^*$. The first inclusion follows by norm inequality, the second inclusion follows by the concavity of the log-likelihood function and the definition of ML estimator, the third inclusion follows by (62), the fourth inequality follows because by definition $j^*(\theta^* + v) \in \{1, \dots, N\}$ for all v. I state a lemma below.

Lemma A.3.3 *Fix some* $0 < \tilde{\lambda}_{\min} < \tilde{\lambda}_{\max} < 1$. *Suppose that an admissible control* π *satisfies* $p_s = \pi_s(D_{1:s-1}) \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ *for all* $1 \leq s \leq t$. *Then there exists a constant* $c_h > 0$ *such that for all* π *and for all* $u \in \Theta - \theta^*$, $\mathbf{E}^\pi[\sqrt{Z_t^\pi(u, D_{1:t})}] \leq \exp(-c_h||u||_2^2 t/2)$.

By Lemma A.3.3, the following holds

$$\begin{aligned}
\mathbb{P}^\pi\left(||\hat{\theta}_t - \theta^*||_2 > \epsilon\right) &\leq \sum_{j=1}^{N} \mathbb{P}^\pi\left(Z_t^\pi(v_j, D_{1:t}) \geq \exp(-c_g\sqrt{q}\eta t)\right) \\
&\leq \sum_{j=1}^{N} \exp\left(\frac{c_g\sqrt{q}\eta t}{2}\right) \mathbf{E}^\pi\left[\sqrt{Z_t^\pi(v_j, D_{1:t})}\right] \\
&\leq \sum_{j=1}^{N} \exp\left(\frac{c_g\sqrt{q}\eta t}{2} - \frac{c_h||v_j||_2^2 t}{2}\right) \\
&\leq \left(\frac{\epsilon}{\sqrt{q}\eta}\right)^{q-1} \exp\left(-\frac{c_h\epsilon^2 t}{2q} + \frac{c_g\sqrt{q}\eta t}{2}\right),
\end{aligned}$$

where the second inequality follows by the Markov's inequality, the third inequality follows by Lemma A.3.3, and the last inequality follows because $N = (\epsilon/(\sqrt{q}\eta))^{q-1}$ and $\min_{j=1,\dots,N} ||v_j||_2 \geq \min_{j=1,\dots,N} ||v_j||_\infty \geq \epsilon/\sqrt{q}$. Now, let $\eta = \epsilon/t$, then I have

$$\mathbb{P}^\pi\left(||\hat{\theta}_t - \theta^*||_2 > \epsilon\right) \leq \min\left\{1, q^{-\frac{q-1}{2}} t^{q-1} \exp\left(-\frac{c_h\epsilon^2 t}{2q} + \frac{c_g\sqrt{q}\epsilon}{2}\right)\right\}.$$

Note that when $\epsilon \leq 1, \exp((-c_h\epsilon^2 q^{-1}t + c_g\sqrt{q}\epsilon)/2) \leq \exp(c_g\sqrt{q}/2)\exp(-c_h\epsilon^2 q^{-1}t/4)$. Note also that when $\epsilon > 1$, there exists $M > 0$ independent of ϵ such that $\exp((-c_h\epsilon^2 q^{-1}t + c_g\sqrt{q}\epsilon)/2) \leq \exp(-c_h\epsilon^2 q^{-1}t/4), \forall t > M$. With these two observations, I consider two cases below.

Case 1: $t > M$. In this case, I have $\mathbb{P}^\pi\left(||\hat{\theta}_t - \theta^*||_2 > \epsilon\right) \leq \tilde{\eta}_4 t^{q-1}\exp(-\eta_5 t\epsilon^2)$, where $\tilde{\eta}_4 = q^{-(q-1)/2}\max\{1, \exp(c_g\sqrt{q}/2)$, and $\eta_5 = c_h q^{-1}/4$.

Case 2: $t \leq M$. Let $\bar{\theta}$ be the largest distance between any two points in Θ. ($\bar{\theta} < \infty$ because Θ is bounded.) Then, I claim that for this case, $\mathbb{P}^\pi\left(||\hat{\theta}_t - \theta^*||_2 > \epsilon\right) \leq \bar{\eta}_4 t^{q-1}\exp(-\eta_5 t\epsilon^2)$ where η_5 is defined as in Case 1 and $\bar{\eta}_4 = \exp(\eta_5 M\bar{\theta}^2)$. The claim is true because: if $\epsilon > \bar{\theta}$, $\mathbb{P}^\pi\left(||\hat{\theta}_t - \theta^*||_2 > \epsilon\right) = 0$, so the bound holds; if $\epsilon \leq \bar{\theta}$, $\mathbb{P}^\pi\left(||\hat{\theta}_t - \theta^*||_2 > \epsilon\right) \leq 1 = \bar{\eta}_4\exp(-\eta_5 M\bar{\theta}^2) \leq \bar{\eta}_4 t^{q-1}\exp(-\eta_5 t\epsilon^2)$.

Combining the two cases above yields $\mathbb{P}^\pi\left(||\hat{\theta}_t - \theta^*||_2 > \epsilon\right) \leq \min\{1, \eta_4 t^{q-1}\exp(-\eta_5 t\epsilon^2)\}$

where $\eta_4 = \max\{\tilde{\eta}_4, \bar{\eta}_4\}$. Hence,

$$\begin{aligned}
\mathbf{E}^\pi\left[||\hat{\theta}_t - \theta^*||_2^2\right] &= \int_0^\infty \mathbb{P}^\pi\left(||\hat{\theta}_t - \theta^*||_2^2 \geq x\right) dx \\
&= \int_0^\infty \min\left\{1, \eta_4 t^{q-1} \exp\left(-\eta_5 t x\right)\right\} dx \\
&\leq \int_0^{\frac{2(q-1)\log t}{\eta_5 t}} dx + \int_{\frac{2(q-1)\log t}{\eta_5 t}}^\infty \left[\eta_4 t^{q-1} \exp\left(-\frac{\eta_5 t x}{2}\right)\right] \exp\left(-\frac{\eta_5 t x}{2}\right) dx \\
&\leq \frac{2(q-1)\log t}{\eta_5 t} + \eta_4 \int_{\frac{2(q-1)\log t}{\eta_5 t}}^\infty \exp\left(-\frac{\eta_5 t x}{2}\right) dx \\
&\leq \frac{2(q-1)\log t}{\eta_5 t} + \frac{2\eta_4}{\eta_5 t} \\
&\leq \frac{2\max\{1, \eta_4\}}{\eta_5} \frac{(q-1)\log t + 1}{t}
\end{aligned}$$

where the fourth inequality holds because for all $x \geq \frac{2(q-1)\log t}{\eta_5 t}$, $\eta_4 t^{q-1} \exp\left(-\frac{\eta_5 t x}{2}\right) \leq 1$. I complete the proof by letting $\eta_6 = \sqrt{2\max\{1, \eta_4\}/\eta_5}$.

A.3.7 Proof of Theorem 3.5.1

I first state an analog of Lemma 3.4.1(a) for ECP(θ) below.

Lemma A.3.4 *Suppose that $\mathcal{B}$ coincides with the set of binding constraints of $P(\theta^*)$ at the optimal solution. There exist $\tilde{\delta} > 0$ and $\tilde{\kappa} > 0$ independent of $k > 0$ such that for all $\theta \in \mathsf{Ball}(\theta^*, \tilde{\delta})$, $||x^D(\theta^*) - x^D(\theta)||_2 \leq \tilde{\kappa}||\theta^* - \theta||_2$.*

The proof of Lemma A.3.4 is similar to the proof of Lemma 3.4.1 and so is omitted. I now proceed to prove Theorem 3.5.1 in several steps.

<u>Step 1</u>

Fix π = APSC and let $k \geq 3$ throughout the proof. Throughout this section, I will assume that $T = 1$. (This is without loss of generality.) Set $L = \lceil (\log k)^{1+\epsilon} \rceil$ and $\eta = (\log k)^{-\epsilon/4}$. I first show that the set of binding constraints of P(θ^*) at the optimal solution can be *correctly* identified with a very high probability. Let $\mathcal{E}_i := \{C_i = (A\lambda^D(\theta^*))_i,\ i \notin \mathcal{B}\} \cup \{C_i > (A\lambda^D(\theta^*))_i,\ i \in \mathcal{B}\}$ denote the event that the i^{th} capacity constraint is wrongly classified. (The event $\mathcal{E}_i$ is a union of two events: either the i^{th} constraint is actually binding but not included in $\mathcal{B}$ or it is not binding but is included in $\mathcal{B}$.) By definition of η,

$$\begin{aligned}
\mathbb{P}^\pi\left(C_i = (A\lambda^D(\theta^*))_i,\, i \notin \mathcal{B}\right) &= \mathbb{P}^\pi\left(C_i = (A\lambda^D(\theta^*))_i,\, C_i - (A\lambda^D(\hat{\theta}_{t_1}))_i > \eta\right) \\
&= \mathbb{P}^\pi\left((A\lambda^D(\theta^*) - A\lambda^D(\hat{\theta}_{t_1}))_i > \eta\right) \\
&\le \mathbb{P}^\pi\left(\kappa||A||_2 E(t_1) > \eta\right) \\
&\le \eta_1 \exp\left(-\eta_2 t_1 \frac{\eta^2}{\kappa^2||A||_2^2}\right) \le \eta_1 \exp\left(-\frac{\eta_2}{\kappa^2||A||_2^2}(\log k)^{1+\frac{\epsilon}{2}}\right),
\end{aligned}$$

where the first inequality follows by Lemma 3.4.1(a), the second inequality follows by Lemma 3.4.2, and the last inequality holds by definition of t_1 and η. Define $\underline{s} := \min\{C_i - (A\lambda^D(\theta^*))_i : C_i - (A\lambda^D(\theta^*))_i > 0, i = 1, \ldots, m\}$. Since $\underline{s}$ does not scale with k, there exists a constant $\Omega_0 > 0$ such that $\eta < \underline{s}/2$ for all $k \ge \Omega_0$. So, for $k \ge \Omega_0$, by Lemmas 3.4.1(a) and 3.4.2, I can bound:

$$\begin{aligned}
\mathbb{P}^\pi\left(C_i > (A\lambda^D(\theta^*))_i,\, i \in \mathcal{B}\right) &= \mathbb{P}^\pi\left(C_i \ge (A\lambda^D(\theta^*))_i + \underline{s},\, C_i - (A\lambda^D(\hat{\theta}_{t_1}))_i \le \eta\right) \\
&\le \mathbb{P}^\pi\left((A\lambda^D(\hat{\theta}_{t_1}) - A\lambda^D(\theta^*))_i \ge \underline{s} - \eta\right) \\
&\le \mathbb{P}^\pi\left(\kappa||A||_2 E(t_1) \ge \underline{s} - \eta\right) \\
&\le \eta_1 \exp\left(-\eta_2 t_1 \frac{(\underline{s} - \eta)^2}{\kappa^2||A||_2^2}\right) \le \eta_1 \exp\left(-\frac{\eta_2\, \underline{s}^2}{4\kappa^2||A||_2^2}\log^{1+\epsilon} k\right).
\end{aligned}$$

Putting the above two bounds together, for $k \ge \Omega_0$, the probability of wrongly identifying the binding constraints can be bounded as follows:

$$\begin{aligned}
\mathbb{P}^\pi\left(\cup_{i=1}^m \mathcal{E}_i\right) &\le \sum_{i=1}^m \left[\mathbb{P}^\pi\left(C_i = (A\lambda^D(\theta^*))_i,\, i \notin \mathcal{B}\right) + \mathbb{P}^\pi\left(C_i > (A\lambda^D(\theta^*))_i,\, i \in \mathcal{B}\right)\right] \\
&\le m\,\eta_1\left[\exp\left(-\frac{\eta_2}{\kappa^2||A||_2^2}(\log k)^{1+\frac{\epsilon}{2}}\right) + \exp\left(-\frac{\eta_2\,\underline{s}^2}{4\kappa^2||A||_2^2}(\log k)^{1+\epsilon}\right)\right]. \quad (63)
\end{aligned}$$

Step 2

Let τ be the minimum of k and the first time $t \ge t_1 + 1$ such that the following condition (C1) is violated: $\psi > ||\sum_{s=t_1+1}^{t} \frac{\hat{\Delta}_s}{k-s}||_2 + ||\frac{S_L - L\lambda_L \mathbf{e}}{k-t}||_2$, where ψ is as defined in the proof of Theorem 3.4.1 and $\hat{\Delta}_s = D_s - \lambda(p_s; \hat{\theta}_{t_z})$ for $s \in (t_z, t_{z+1}]$ and $1 \le z \le Z$. Define $\mathcal{A} := \{\cap_{i=1}^m \mathcal{E}_i^c\} \cap \left\{E(t_z) \le \min\{\hat{\delta}, (\log t_z)^{-\epsilon/4}\}, \text{for all } t_z < \tau\right\}$, where $\hat{\delta} = \min\{\bar{\delta}, \tilde{\delta}, \phi/(2\omega\kappa)\}$ and $\bar{\delta}$ and $\tilde{\delta}$ are as defined in Lemma 3.4.1 and Lemma A.3.4 respectively. (Event $\mathcal{A}$ can be interpreted as the event where all binding constraints are correctly identified and the size of all subsequent estimation errors are sufficiently small.)

Note that for $t_z < \tau$, $\lambda^D(\theta^*) \in \Lambda_{\hat{\theta}_{t_z}}$ on $\mathcal{A}$. This is because $||p(\lambda^D(\theta^*); \hat{\theta}_{t_z}) - p(\lambda^D(\hat{\theta}_{t_z}); \hat{\theta}_{t_z})||_2 \leq \omega||\lambda^D(\theta^*) - \lambda^D(\hat{\theta}_{t_z})||_2 \leq \omega\kappa||\theta^* - \hat{\theta}_{t_z}||_2 \leq \phi/2$, where the first inequality follows by P1, the second inequality follows by Lemma 3.4.1(a) and the fact that $\hat{\delta} \leq \bar{\delta}$, and the last inequality follows since $\hat{\delta} \leq \phi/(2\omega\kappa)$. I then have $\lambda^D(\theta^*) \in \Lambda_{\hat{\theta}_{t_z}}$ since $p(\lambda^D(\theta^*); \hat{\theta}_{t_z}) \in \mathsf{Ball}(p^D(\hat{\theta}_{t_z}), \phi/2) \subseteq \mathcal{P}$, where the last inclusion follows by Lemma 3.4.1(a). The two important lemmas below establish the approximation error of DPUP and some important properties of the stopping time τ.

Lemma A.3.5 *There exist positive constants γ and ξ independent of $\theta \in \Theta$ such that if $||x^D(\theta) - x^{NT}_{z-1}||_2 \leq \gamma$, then $||x^D(\theta) - x^{NT}_z||_2 \leq \xi||x^D(\theta) - x^{NT}_{z-1}||_2^2$.*

Lemma A.3.6 *There exist positive constants $0 < \tilde{\lambda}_{\min} < \tilde{\lambda}_{\max} < 1$, Ω_1, and constants Γ_1 and Γ_2 independent of $k \geq \Omega_1$, such that $\tilde{\lambda}_{\min} \leq \lambda_{\min}, \tilde{\lambda}_{\max} \geq \lambda_{\max}$, and for all $k \geq \Omega_1$ and all sample paths on $\mathcal{A}$:*

(a) $||x^D(\hat{\theta}_{t_z}) - x^{NT}_z||_2^2 \leq \Gamma_1(\log t_z)^{-\epsilon/2}$ *for* $t_z < \tau$.

(b) $C_t \succ 0$, $p_t \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8) \subseteq \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ *and* $\hat{\lambda}_t \in \Lambda_{\hat{\theta}_{t_z}}$ *for all* $t \in (t_z, t_{z+1}] \cap [t_1, \tau)$.

(c) $\mathbf{E}^\pi[||x^D(\hat{\theta}_{t_z}) - x^{NT}_z||_2^2 \, \mathbf{1}_{\{t_z < \tau\}} \mid \mathcal{A}] \leq \Gamma_2/t_z$

Lemma A.3.5 essentially establishes a *uniform* locally quadratic convergence of the Newton's method for solving ECP($\hat{\theta}_{t_z}$) for all z, which is used for proving Lemma A.3.6(a) and (c). Lemma A.3.6(a) and (c) establish the approximation errors between x^{NT}_Z and the deterministic optimal solution $x^D(\hat{\theta}_{t_z})$. Note that Lemma A.3.6(b) states that $p_t \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8) \subseteq \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ for all $t_1 \leq t < \tau$. In addition, for $t \leq t_1$, $p_t \in \{\tilde{p}^{(1)}, \ldots, \tilde{p}^{(\tilde{q})}\} \subseteq \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ due to $\tilde{\lambda}_{\min} \leq \lambda_{\min}$, $\tilde{\lambda}_{\max} \geq \lambda_{\max}$ and S2. Therefore, the condition for Lemma 3.5.1 is satisfied. There exists a constant $\Omega_2 \geq \max\{\Omega_0, \Omega_1\}$ such that, for all $k \geq \Omega_2$,

$$
\begin{aligned}
k\mathbb{P}^\pi(\mathcal{A}^c) &\leq k\sum_{z=1}^{Z}\left[\mathbb{P}^\pi(E(t_z) > \hat{\delta}) + \mathbb{P}^\pi(E(t_z) > (\log t_z)^{-\frac{\epsilon}{4}}\right] + k\,\mathbb{P}^\pi\left(\cup_{i=1}^m \mathcal{E}_i\right) \\
&\leq k\sum_{z=1}^{Z}\eta_4 t_z^{q-1}\left[\exp\left(-\eta_5 t_z\hat{\delta}^2\right) + \exp\left(-\frac{\eta_5 t_z}{(\log t_z)^{\frac{\epsilon}{2}}}\right)\right] + k\,\mathbb{P}^\pi\left(\cup_{i=1}^m \mathcal{E}_i\right) \\
&\leq 2k(\log_2 k)\left[\exp\left(-\frac{\eta_5(\log k)^{1+\epsilon}\hat{\delta}^2}{2}\right) + \exp\left(-\frac{\eta_5(\log k)^{1+\epsilon}}{2(\log k)^{\frac{\epsilon}{2}}}\right)\right] + k\,\mathbb{P}^\pi\left(\cup_{i=1}^m \mathcal{E}_i\right) \\
&\leq 2k(\log_2 k)\left[\exp\left(-\frac{\eta_5(\log k)^{1+\epsilon}\hat{\delta}^2}{2}\right) + \exp\left(-\frac{\eta_5(\log k)^{1+\frac{\epsilon}{2}}}{2}\right)\right] \\
&\quad + \; m\,\eta_1 k\left[\exp\left(-\frac{\eta_2}{\kappa^2||A||_2^2}(\log k)^{1+\frac{\epsilon}{2}}\right) + \exp\left(-\frac{\eta_2\,\underline{s}^2}{4\kappa^2||A||_2^2}(\log k)^{1+\epsilon}\right)\right] \;\leq\; \frac{1}{2},
\end{aligned}
$$

where the second inequality follows by Lemma 3.5.1, the third inequality follows by a combination of $\eta_4 t_z^{q-1}\exp(-\eta_5 t_z\hat{\delta}^2/2) \to 0$ and $\eta_4 t_z^{q-1}\exp(-\eta_5 t_z(\log t_z)^{-\epsilon/2}/2) \to 0$ as $k \to \infty$, $t_z \geq t_1 \geq (\log k)^{1+\epsilon}$ for $z \geq 1$, and $Z \leq \lceil \log_2 k \rceil \leq 2\log_2 k$, the fourth inequality follows by (63), and the last inequality follows because the formula after the fourth inequality goes to zero as $k \to \infty$. Note that the above inequality also implies $\mathbb{P}^\pi(\mathcal{A}) > \frac{1}{2}$ when $k \geq \Omega_2$. Define $\Psi_\epsilon := \sum_{t=t_1+1}^{k-1}\left(\sum_{s=t_1+1}^{t-1}\frac{\bar{\epsilon}(s)}{k-s}\right)^2$ and $\Phi_\epsilon := \sum_{t=t_1+1}^{k-1}\bar{\epsilon}(s)^2$, where $\bar{\epsilon}(s) := \eta_6\sqrt{[(q-1)\log t_z + 1]/t_z}$ for all $s \in (t_z, t_{z+1}]$. By Lemma 3.5.1, $\mathbf{E}^\pi[||\hat{\theta}_t - \theta^*||_2^2\mathbf{1}_{\{t<\tau\}}|\mathcal{A}] \leq \bar{\epsilon}(t)^2$. The following result is useful to derive my bounds later.

Lemma A.3.7 *Under APSC, there exists a constant $K_3 > 0$ independent of $k \geq 1$ such that $\Psi_\epsilon < K_3(1 + (q-1)\log k)$ and $\Phi_\epsilon < K_3[1 + \log k + (q-1)(\log k)^2]$.*

Step 3

Let $K = \max\{\Omega_0, \Omega_1, \Omega_2, 3\}$. If $k < K$, the total expected revenue loss can be bounded by $K\bar{r}$. So, I will focus on the case $k \geq K$. By the same arguments as in (54) and (55), $\rho^\pi(k) \leq L\bar{r} + \sum_{t=t_1+1}^{k} r^D(\theta^*) - \mathbf{E}\left[\hat{R}^\pi(k)\right]$ and

$$
\begin{aligned}
&\sum_{t=t_1+1}^{k} r^D(\theta^*) - \mathbf{E}\left[\hat{R}^\pi(k)\right] \\
&\quad\leq \mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}\mu^D(\theta^*)'A(\lambda^D(\theta^*) - \lambda_t)\middle|\mathcal{A}\right] + \frac{\bar{v}}{2}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}||\lambda^D(\theta^*) - \lambda_t||_2^2\right] \\
&\qquad + \;\bar{r}\,\mathbf{E}^\pi[k-\tau|\mathcal{A}] \;+\; 2\,\bar{r} \;+\; \bar{r}\,k\mathbb{P}^\pi(\mathcal{A}^c) \qquad (64)
\end{aligned}
$$

Note that on $\mathcal{A}$ I have $\mu^D(\theta^*)'A\lambda^D(\theta^*) = \mu^D(\theta^*)'A\lambda_z^{NT}$ (because $B\lambda^D(\theta^*) = C_B = B\lambda_z^{NT}$ and $\mu^D(\theta^*)_i = 0$ for all $i \notin \mathcal{B}$ by KKT conditions). Therefore, similar to the proof of Lemma 3.4.3, I can bound the first term in (64) with $K_4\mathbf{E}^\pi[k-\tau+1|\mathcal{A}]$ where $K_4 := 3\mu^D(\theta^*)'A\mathbf{e} + \psi||\mu^D(\theta)||_2||A||_2$ is independent of $k \geq K$.

As for the second term in (64), recall that $\hat{\lambda}_t = \lambda_{z(t)}^{NT} - Q\sum_{s=t_1+1}^{t-1}\frac{\hat{\Delta}_s}{k-s}$ denotes the demand rate that the seller believes he is inducing during period t where $z(t)$ is the unique integer z such that $t \in (t_z, t_{z+1}]$. Note that (57) still holds. I can bound two term in (57) respectively using: $\bar{v}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}||\hat{\lambda}_t - \lambda_t||_2^2\middle|\mathcal{A}\right] = \bar{v}\sum_{t=t_1+1}^{k-1}\mathbf{E}^\pi\left[\omega^2||\hat{\theta}_t - \theta^*||_2^2\mathbf{1}_{\{t<\tau\}}\middle|\mathcal{A}\right] \leq \bar{v}\omega^2\sum_{t=t_1+1}^{k-1}\bar{\epsilon}(t)^2 \leq \bar{v}\omega^2\Phi_\epsilon$ (by P2), and

$$
\begin{aligned}
&\bar{v}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}||\lambda^D(\theta^*) - \hat{\lambda}_t||_2^2\middle|\mathcal{A}\right] \\
\leq\ & 2\bar{v}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}||\lambda^D(\theta^*) - \lambda_{z(t)}^{NT}||_2^2\middle|\mathcal{A}\right] + 2\bar{v}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}\left|\left|Q\sum_{s=t_1+1}^{t-1}\frac{\hat{\Delta}_s}{k-s}\right|\right|_2^2\middle|\mathcal{A}\right] \\
\leq\ & 2\bar{v}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}||\lambda^D(\theta^*) - \lambda_{z(t)}^{NT}||_2^2\middle|\mathcal{A}\right] \\
&+2\bar{v}||Q||_2^2\left(\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}\left|\left|\sum_{s=t_1+1}^{t-1}\frac{\Delta_s}{k-s}\right|\right|_2^2\middle|\mathcal{A}\right] + \mathbf{E}^\pi\left[\sum_{t=t_1+1}^{k-1}\left(\sum_{s=t_1+1}^{t-1}\frac{\omega E(s)\mathbf{1}_{\{s<\tau\}}}{k-s}\right)^2\middle|\mathcal{A}\right]\right) \\
\leq\ & 2\bar{v}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}||\lambda^D(\theta^*) - \lambda_{z(t)}^{NT}||_2^2\middle|\mathcal{A}\right] \\
&+2\bar{v}||Q||_2^2\left(\frac{16}{\mathbb{P}^\pi(\mathcal{A})}\log k + \sum_{t=t_1+1}^{k-1}\left[\sum_{s=t_1+1}^{t-1}\frac{\sqrt{\mathbf{E}^\pi\left[\omega^2E(s)^2\mathbf{1}_{\{s<\tau\}}\middle|\mathcal{A}\right]}}{k-s}\right]^2\right) \\
\leq\ & 2\bar{v}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}||\lambda^D(\theta^*) - \lambda_{z(t)}^{NT}||_2^2\middle|\mathcal{A}\right] + 2\bar{v}||Q||_2^2(32\log k + \omega^2\Psi_\epsilon) \\
\leq\ & K_5(\Psi_\epsilon + \log k) + 2\bar{v}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1}||\lambda^D(\theta^*) - \lambda_{z(t)}^{NT}||_2^2\middle|\mathcal{A}\right]
\end{aligned}
$$

for some constant $K_5 > 0$ independent of $k \geq K$ (the second and the third inequalities follow by the same argument as in (58) and (59) and recall that $K \geq 3$), and the fourth inequality follows since $\mathbf{E}^\pi\left[E(s)^2\mathbf{1}_{\{s<\tau\}}\middle|\mathcal{A}\right] \leq \bar{\epsilon}(s)^2$. I now analyze the last term of the above. Note that, on $\mathcal{A}$, I have for all $t < \tau$

$$\begin{aligned} \left|\left|\lambda^D(\theta^*) - \lambda^{NT}_{z(t)}\right|\right|_2 &\leq \left|\left|\lambda^D(\theta^*) - x^{NT}_{z(t)}\right|\right|_2 + \left|\left|\lambda^{NT}_{z(t)} - x^{NT}_{z(t)}\right|\right|_2 \\ &\leq 2\left|\left|\lambda^D(\theta^*) - x^{NT}_{z(t)}\right|\right|_2 \\ &= 2\left|\left|x^D(\theta^*) - x^{NT}_{z(t)}\right|\right|_2 \\ &\leq 2\left|\left|x^D(\theta^*) - x^D(\hat{\theta}_{t_{z(t)}})\right|\right|_2 + 2\left|\left|x^D(\hat{\theta}_{t_{z(t)}}) - x^{NT}_z\right|\right|_2, \end{aligned} \tag{65}$$

where the second inequality follows because $\lambda^D(\theta^*)$ lies in $\mathcal{S}_{z(t)}$ where $x^{NT}_{z(t)}$ is projected into (note that on $\mathcal{A}$, (1) $\hat{\theta}_{t_{z(t)}} \in \mathsf{Ball}(\theta^*, \phi/(2\omega\kappa))$ for $t < \tau$ which implies that, as shown previously, $\lambda^D(\theta^*) \in \Lambda_{\hat{\theta}_{t_{z(t)}}}$, and (2) the binding constraints of $\mathrm{P}(\theta^*)$ at $\lambda^D(\theta^*)$ are correctly identified which means that $B\lambda^D(\theta^*) = C_B$ and $N\lambda^D(\theta^*) \leq C_N$) and the equality follows because $\lambda^D(\theta^*) = x^D(\theta^*)$ on $\mathcal{A}$ due to the strongly concavity of the objective and the fact that $\lambda^D(\theta^*)$ is an interior solution. By Lemma A.3.4

$$\mathbf{E}^\pi\left[\sum_{s=t_1+1}^{\tau-1} ||x^D(\theta^*) - x^D(\hat{\theta}_{t_{z(s)}})||_2^2 \middle| \mathcal{A}\right] = \sum_{s=t_1+1}^{k-1} \mathbf{E}^\pi\left[\tilde{\kappa}^2||\theta^* - \hat{\theta}_{t_{z(s)}}||_2^2 \mathbf{1}_{\{s<\tau\}}\middle| \mathcal{A}\right] \leq \tilde{\kappa}^2\Phi_\epsilon$$

Furthermore, by Lemma A.3.6(a) and the fact that $t_{z+1} - t_z \leq 2t_z$ for all z, I have

$$\begin{aligned} \mathbf{E}^\pi\left[\sum_{s=t_1+1}^{\tau-1} ||x^D(\hat{\theta}_{t_{z(s)}}) - x^{NT}_{z(s)}||_2^2 \middle| \mathcal{A}\right] &= \sum_{s=t_1+1}^{k-1} \mathbf{E}^\pi\left[||x^D(\hat{\theta}_{t_{z(s)}}) - x^{NT}_{z(s)}||_2^2 \mathbf{1}_{\{s<\tau\}}\middle| \mathcal{A}\right] \\ &\leq \sum_{z=1}^{Z} (t_{z+1} - t_z)\frac{\Gamma_2}{t_z} \leq 2Z\,\Gamma_2 \leq 4\Gamma_2 \log_2 k. \end{aligned}$$

Combining the inequalities above, the second term of (64) can be bounded as follows:

$$\begin{aligned} \frac{\bar{v}}{2}\mathbf{E}^\pi\left[\sum_{t=t_1+1}^{\tau-1} ||\lambda^D(\theta^*) - \lambda_t||_2^2\right] &\leq \bar{v}\omega^2\Phi_\epsilon + K_5(\Psi_\epsilon + \log k) + 4\bar{v}\tilde{\kappa}^2\Phi_\epsilon + 16\bar{v}\Gamma_2 \log_2 \\ &\leq K_6(1 + \log k + (q-1)\log^2 k) \end{aligned}$$

for $K_6 = (\bar{v}\omega^2 + 4\bar{v}\tilde{\kappa}^2 + K_5)K_3 + K_5 + 16\bar{v}\,\Gamma_2$. To bound the third term in (64), the following lemma is useful.

Lemma A.3.8 *There exists a constant $K_7 > 0$ independent of $k \geq K$ such that for all $k \geq K$, $\mathbf{E}^\pi[k - \tau|\mathcal{A}] \leq K_7(\log k + L)$.*

Combining all the above and recalling that $L = \lceil (\log k)^{1+\epsilon} \rceil$, for all $k \geq K$, I have:

$$\begin{aligned}
\rho^\pi(k) &\leq 2\bar{r}(\log k)^{1+\epsilon} + (K_4 + \bar{r})(\mathbf{E}^\pi[k - \tau | \mathcal{A}] + 1) + K_6(1 + \log k + (q-1)\log^2 k) + \frac{5}{2}\bar{r} \\
&\leq \left(2\bar{r} + K_4 + \bar{r} + K_6 + \frac{5}{2}\bar{r}\right)\left[1 + (\log k)^{1+\epsilon} + (q-1)\log^2 k\right] \\
&\leq K_8[(\log k)^{1+\epsilon} + (q-1)\log^2 k],
\end{aligned}$$

for some constant K_8 independent of $k \geq K$. The result of Theorem 3.5.1 follows by using $M_2 = \max\{\bar{r}K, K_8\}$.

A.3.8 Proof of Supporting Lemmas

Proof of Lemma A.3.1. As in the proof of Lemma 3.4.3, I assume without loss of generality that $T = 1$. First, note that $\hat{\lambda}_t \in \Lambda_{\hat{\theta}_L}$ is equivalent to $p_t \in \mathcal{P}$. Consider sample paths on $\mathcal{A}$. If $\tau \leq L + 1$, then there is nothing to prove. Suppose that $\tau > L + 1$, I will use an induction argument to establish the result. Since $E(L) \leq \bar{\delta}$ on $\mathcal{A}$, by Lemma 3.4.1(a), $\mathsf{Ball}(p^D(\hat{\theta}_L), \frac{\phi}{2}) \subseteq \mathcal{P}$. For $t = L+1$, $||p_{L+1} - p^D(\hat{\theta}_L)||_2 = 0 < \frac{\phi}{2}$, so $p_{L+1} \in \mathcal{P}$ and hence $\hat{\lambda}_{L+1} \in \Lambda_{\hat{\theta}_L}$. In addition, I also have:

$$\begin{aligned}
C_{L+1} = C_L - AD_{L+1} &= kC - AS_L - A\left(\hat{\lambda}_{L+1} + \hat{\Delta}_{L+1}\right) \\
&= kC - LC + LC - AS_L - A\left(\lambda^D(\hat{\theta}_L) + \hat{\Delta}_{L+1}\right) \\
&\succeq (k - L - 1)C + LC - AS_L - A\hat{\Delta}_{L+1} \\
&\succeq (k - L - 1)A\lambda_L\mathbf{e} + LA\lambda_L\mathbf{e} - AS_L - A\hat{\Delta}_{L+1} \\
&= (k - L - 1)A\left(\lambda_L\mathbf{e} - \frac{S_L - L\lambda_L\mathbf{e}}{k - L - 1} - \frac{\hat{\Delta}_{L+1}}{k - L - 1}\right) \\
&\succeq (k - L - 1)A\left(\lambda_L\mathbf{e} - \left|\left|\frac{S_L - L\lambda_L\mathbf{e}}{k - L - 1}\right|\right|_2\mathbf{e} - \left|\left|\frac{\hat{\Delta}_{L+1}}{k - L - 1}\right|\right|_2\mathbf{e}\right) \\
&\succ (k - L - 1)\left(\lambda_L - \psi\right)A\mathbf{e} \\
&\succeq 0
\end{aligned}$$

(recall that $S_t = \sum_{s=1}^t D_s$) where the first inequality follows because $A\lambda^D(\hat{\theta}_L) \preceq C$, the second inequality follows because $A\lambda_L\mathbf{e} \preceq C$ by definition of λ_L, the fourth (strict) inequality follows by (C1) and $A\mathbf{e} \succ 0$, and the last inequality follows by the definition of ψ. This is my base case. Now, suppose that $C_s \succ 0, \hat{\lambda}_s \in \Lambda_{\hat{\theta}_L}$ for all $s = L+1, L+2, \ldots, t-1$, and

$t-1<\tau$. If $t \geq \tau$, I have finished the induction. If, on the other hand, $t<\tau$,

$$\left|\left|p_t - p^D(\hat{\theta}_L)\right|\right|_2 \;\leq\; \omega||Q||_2 \left|\left|\sum_{s=L+1}^{t-1} \frac{\hat{\Delta}_s}{k-s}\right|\right|_2 \;<\; \omega||Q||_2\psi \;\leq\; \frac{\phi}{4}$$

where the first inequality follows by $p_t = p(\hat{\lambda}_t;\hat{\theta}_L), p^D(\hat{\theta}_L) = p(\lambda^D(\hat{\theta}_L);\hat{\theta}_L)$ and P1, the second inequality follows by (C1) and the last inequality follows by the definition of ψ. So, by Lemma 3.4.1(a), I still have $p_t \in \mathcal{P}$ and hence $\hat{\lambda}_t \in \Lambda_{\hat{\theta}_L}$. As for the remaining capacity level C_t, by similar argument as before, I have

$$\begin{aligned}
C_t = C_L - \sum_{s=L+1}^{t} AD_s \;&=\; kC - AS_L - \sum_{s=L+1}^{t} A\left(\hat{\lambda}_s + \hat{\Delta}_s\right)\\
&=\; kC - tC + tC - AS_L - \sum_{s=L+1}^{t} A\left(\lambda^D(\hat{\theta}_L) - Q\sum_{v=L+1}^{s-1}\frac{\hat{\Delta}_v}{k-v} + \hat{\Delta}_s\right)\\
&\succeq\; (k-t)C + LC - AS_L - \sum_{s=L+1}^{t}\left(A\hat{\Delta}_s - \sum_{v=L+1}^{s-1}\frac{A\hat{\Delta}_v}{k-v}\right)\\
&\succeq\; (k-t)A\lambda_L\mathbf{e} + LA\lambda_L\mathbf{e} - AS_L - \sum_{s=L+1}^{t}\left(A\hat{\Delta}_s - \sum_{v=L+1}^{s-1}\frac{A\hat{\Delta}_v}{k-v}\right)\\
&=\; (k-t)A\left(\lambda_L\mathbf{e} - \frac{S_L - L\lambda_L\mathbf{e}}{k-t} - \sum_{s=L+1}^{t}\frac{\hat{\Delta}_s}{k-s}\right)\\
&\succeq\; (k-t)A\left(\lambda_L\mathbf{e} - \left|\left|\frac{S_L - L\lambda_L\mathbf{e}}{k-t}\right|\right|_2\mathbf{e} - \left|\left|\sum_{s=L+1}^{t}\frac{\hat{\Delta}_s}{k-s}\right|\right|_2\mathbf{e}\right)\\
&\succ\; (k-t)A(\lambda_L - \psi)\mathbf{e}\\
&\succeq\; 0.
\end{aligned}$$

This completes the induction.

Proof of Lemma A.3.2. As in the proof of Lemma 3.4.3, I assume without loss of generality that $T=1$. Because τ is non-negative, I can write $\mathbf{E}^\pi[k-\tau|\mathcal{A}] = k - \sum_{t=0}^{k-1}\mathbb{P}^\pi(\tau > t|\mathcal{A}) = \sum_{t=1}^{k-1}\mathbb{P}^\pi(\tau \leq t|\mathcal{A})$. I now bound $\mathbb{P}^\pi(\tau \leq t|\mathcal{A})$. By the union bound, I have

$$\begin{aligned}
\mathbb{P}^\pi(\tau \leq t|\mathcal{A}) \;&=\; \mathbb{P}^\pi\left(\max_{L+1\leq s\leq t}\left\{\left|\left|\frac{S_L - L\lambda_L\mathbf{e}}{k-s}\right|\right|_2 + \left|\left|\sum_{v=L+1}^{s}\frac{\hat{\Delta}_v}{k-v}\right|\right|_2\right\} \geq \psi \middle| \mathcal{A}\right)\\
&\leq\; \mathbb{P}^\pi\left(\max_{L+1\leq s\leq t}\left|\left|\sum_{v=L+1}^{s}\frac{\hat{\Delta}_v}{k-v}\right|\right|_2 \geq \frac{\psi}{2}\middle| \mathcal{A}\right) + \mathbb{P}^\pi\left(\max_{L+1\leq s\leq t}\left|\left|\frac{S_L - L\lambda_L\mathbf{e}}{k-s}\right|\right|_2 \geq \frac{\psi}{2}\middle| \mathcal{A}\right).
\end{aligned} \tag{66}$$

I first bound the first term in (66) below.

$$
\begin{aligned}
&\mathbb{P}^\pi\left(\max_{L+1\le s\le t}\left|\left|\sum_{v=L+1}^{s}\frac{\hat{\Delta}_v}{k-v}\right|\right|_2\ge\frac{\psi}{2}\middle|\mathcal{A}\right)\\
&\le \mathbb{P}^\pi\left(\max_{L+1\le s\le t}\left|\left|\sum_{v=L+1}^{s}\frac{\Delta_v}{k-v}\right|\right|_2^2\ge\frac{\psi^2}{16}\middle|\mathcal{A}\right)+\mathbb{P}^\pi\left(\max_{L+1\le s\le t}\sum_{v=L+1}^{s}\frac{||\lambda_v-\hat{\lambda}_v||_2}{k-v}\ge\frac{\psi}{4}\middle|\mathcal{A}\right)\\
&\le \frac{1}{\mathbb{P}^\pi(\mathcal{A})}\;\mathbb{P}^\pi\left(\max_{L+1\le s\le t}\left|\left|\sum_{v=L+1}^{s}\frac{\Delta_v}{k-v}\right|\right|_2^2\ge\frac{\psi^2}{16}\right)+\mathbb{P}^\pi\left(\left(\sum_{s=L+1}^{t}\frac{||\lambda_s-\hat{\lambda}_s||_2}{k-s}\right)^2\ge\frac{\psi^2}{16}\middle|\mathcal{A}\right)\\
&\le \frac{16}{\psi^2\mathbb{P}^\pi(\mathcal{A})}\;\mathbf{E}^\pi\left[\left|\left|\sum_{s=L+1}^{t}\frac{\Delta_s}{k-s}\right|\right|_2^2\right]+\frac{16}{\psi^2}\;\mathbf{E}^\pi\left[\left(\sum_{s=L+1}^{t}\frac{||\lambda_s-\hat{\lambda}_s||_2}{k-s}\right)^2\middle|\mathcal{A}\right]\\
&\le \frac{16}{\psi^2\mathbb{P}^\pi(\mathcal{A})}\;\mathbf{E}^\pi\left[\sum_{s=L+1}^{t}\frac{||\Delta_s||_2^2}{(k-s)^2}\right]+\frac{16}{\psi^2}\left(\sum_{s=L+1}^{t}\frac{\sqrt{\mathbf{E}^\pi[||\lambda_s-\hat{\lambda}_s||_2^2|\mathcal{A}]}}{k-s}\right)^2\\
&\le \frac{16}{\psi^2\mathbb{P}^\pi(\mathcal{A})}\left[\frac{4}{(k-t)^2}+\frac{4}{k-t}\right]+\frac{16}{\psi^2}\left[\frac{2\omega^2\epsilon(L)^2}{(k-t)^2}+2\omega^2\epsilon(L)^2\left(\log\left(\frac{k}{k-t}\right)\right)^2\right],
\end{aligned}
$$

where the first inequality follows by the definition of $\hat{\Delta}_v$, the triangle inequality of the norms and union bound, the second inequality follows by the law of total probability for the first term and the monotonicity of max-operator for the second term, the third inequality follows by the Doob's sub-martingale inequality for the first term and Markov's inequality for the second term, the fourth inequality follows by the orthogonality of martingale differences for the first term and Cauchy-Schwartz inequality for the second term, and the last inequality follows by $\mathbf{E}^\pi[E(L)^2|\mathcal{A}]\le\epsilon(L)^2$ and the same integral approximation bound as in (59).

As for the second term in (66), I can apply Markov's inequality and get:

$$
\begin{aligned}
\mathbb{P}^\pi\left(\max_{1\le s\le t}\left|\left|\frac{S_L-L\lambda_L\mathbf{e}}{k-s}\right|\right|_2\ge\frac{\psi}{2}\middle|\mathcal{A}\right) &\le \mathbb{P}^\pi\left(\frac{||S_L-L\lambda_L\mathbf{e}||_2^2}{(k-t)^2}\ge\frac{\psi^2}{4}\middle|\mathcal{A}\right)\\
&\le \max\left\{1,\frac{4}{\psi^2}\mathbf{E}^\pi\left[\frac{||S_L-L\lambda_L\mathbf{e}||_2^2}{(k-t)^2}\middle|\mathcal{A}\right]\right\}\\
&\le \max\left\{1,\frac{4n(1+\lambda_L)^2L^2}{\psi^2(k-t)^2}\right\},
\end{aligned}
$$

where the last inequality follows because $||S_L-L\lambda_L\mathbf{e}||_2\le||L\mathbf{e}+L\lambda_L\mathbf{e}||_2=\sqrt{n}(1+\lambda_L)L$.

Putting all the bounds together, I have for all $k \geq 3$:

$$\begin{aligned}
\mathbf{E}^{\pi}[k-\tau|\mathcal{A}] &\leq \sum_{t=1}^{k-1}\left\{\frac{16}{\psi^2\mathbb{P}^{\pi}(\mathcal{A})}\left[\frac{4}{(k-t)^2}+\frac{4}{k-t}\right]+\frac{16}{\psi^2}\left[\frac{2\omega^2\epsilon(L)^2}{(k-t)^2}+2\omega^2\epsilon(L)^2\log^2\left(\frac{k}{k-t}\right)\right]\right\} \\
&\quad +\sum_{t=1}^{k-1}\max\left\{1,\frac{4n(1+\lambda_L)^2L^2}{\psi^2(k-t)^2}\right\} \\
&\leq \frac{128}{\psi^2\mathbb{P}^{\pi}(\mathcal{A})}\sum_{t=1}^{k-1}\frac{1}{k-t}+\frac{32\omega^2\epsilon(L)^2}{\psi^2}\sum_{t=1}^{k-1}\frac{1}{(k-t)^2}+\frac{32\omega^2\epsilon(L)^2}{\psi^2}\sum_{t=1}^{k-1}\left(\log\left(\frac{k}{k-t}\right)\right)^2 \\
&\quad +\sum_{t=1}^{k-L-1}\frac{4n(1+\lambda_L)^2L^2}{\psi^2(k-t)^2}+L \\
&\leq \frac{128}{\psi^2\mathbb{P}^{\pi}(\mathcal{A})}(1+\log k)+\frac{64\omega^2}{\psi^2}\epsilon(L)^2+\frac{96\omega^2}{\psi^2}\epsilon(L)^2k+\left(\frac{4n(1+\lambda_L)^2}{\psi^2}+1\right)L \\
&\leq \frac{256}{\psi^2}\frac{\log k}{1-\mathbb{P}^{\pi}(E(L)>\bar{\delta})}+\frac{160\omega^2}{\psi^2}\epsilon(L)^2k+\left(\frac{4n(1+\lambda_L)^2}{\psi^2}+1\right)L
\end{aligned}$$

where the third inequality follows by integral approximation. The result follows by letting $K_0 = \frac{256}{\psi^2}+\frac{160\omega^2}{\psi^2}+\frac{4n(1+\lambda_L)^2}{\psi^2}+1$.

Proof of Lemma A.3.3. Recall that $\mathcal{D} = \{D \in \{0,1\}^n : \sum_{j=1}^n D_j \leq 1\}$. I define the conditional Hellinger distance as follows:

$$H_t^{\pi}(\theta_1,\theta_2,D_t|D_{1:t-1}) := \sum_{D_t\in\mathcal{D}}\left(\sqrt{\mathbb{P}_t^{\pi,\theta_1}(D_t|D_{1:t-1})}-\sqrt{\mathbb{P}_t^{\pi,\theta_2}(D_t|D_{1:t-1})}\right)^2.$$

I state a lemma and postpone its proof to the end of this subsection.

Lemma A.3.9 *Fix some* $0 < \tilde{\lambda}_{\min} < \tilde{\lambda}_{\max} < 1$. *Suppose that an admissible control* π *satisfies* $p_s = \pi_s(D_{1:s-1}) \in \mathcal{W}(\tilde{\lambda}_{\min},\tilde{\lambda}_{\max})$ *for all* $1 \leq s \leq t$. *Then there exists a positive constant* c_h *such that* $H_t^{\pi}(\theta_1,\theta_2,D_t|D_{1:t-1}) \geq c_h||\theta_1-\theta_2||_2^2$ *for all* $\theta_1,\theta_2 \in \Theta$.

For $u \in \Theta-\theta^*$, define $Z_t^{\pi}(u,D_t|D_{1:t-1}) := \mathbb{P}_t^{\pi,\theta^*+u}(D_t|D_{1:t-1})/\mathbb{P}_t^{\pi,\theta^*}(D_t|D_{1:t-1})$. Using

Lemma A.3.9, I can derive a bound for its moment below:

$$\begin{aligned}
\mathbf{E}^{\pi}\left[\sqrt{Z_t^{\pi}(u, D_t|D_{1:t-1})}\right] &= \sum_{D_t \in \mathcal{D}} \sqrt{\frac{\mathbb{P}_t^{\pi,\theta^*+u}(D_t|D_{1:t-1})}{\mathbb{P}_t^{\pi,\theta^*}(D_t|D_{1:t-1})}}\, \mathbb{P}_t^{\pi,\theta^*}(D_t|D_{1:t-1}) \\
&= \sum_{D_t \in \mathcal{D}} \sqrt{\mathbb{P}_t^{\pi,\theta^*+u}(D_t|D_{1:t-1})\mathbb{P}_t^{\pi,\theta^*}(D_t|D_{1:t-1})} \\
&= 1 - \frac{H_t^{\pi}(\theta^*, \theta^* + u, D_t|D_{1:t-1})}{2} \\
&\leq \exp\left(-\frac{H_t^{\pi}(\theta^*, \theta^* + u, D_t|D_{1:t-1})}{2}\right) \leq \exp\left(-\frac{c_h||u||_2^2}{2}\right).
\end{aligned}$$

The result of Lemma A.3.3 can now be proved by repeated conditioning: by definition,

$$\begin{aligned}
\mathbf{E}^{\pi}\left[\sqrt{Z_t^{\pi}(u, D_{1:t})}\right] &= \mathbf{E}^{\pi}\left[\mathbf{E}^{\pi}\left[\sqrt{Z_t^{\pi}(u, D_{1:t})}\,\middle|\, D_{1:t-1}\right]\right] \\
&= \mathbf{E}^{\pi}\left[\sqrt{Z_{t-1}^{\pi}(u, D_{1:t-1})}\ \mathbf{E}^{\pi}\left[\sqrt{Z_k^{\pi}(u, D_t|D_{1:t-1})}\right]\right] \\
&\leq \mathbf{E}^{\pi}\left[\sqrt{Z_{t-1}^{\pi}(u, D_{1:t-1})}\right]\exp\left(-\frac{c_h||u||_2^2}{2}\right) \\
&\leq \exp\left(-\frac{c_h||u||_2^2 t}{2}\right).
\end{aligned}$$

This completes the proof.

Proof of Lemma A.3.5. Fix $\theta \in \Theta$. Note that ECP(θ) is a convex optimization with linear equality constraints. Let m_B denote the number of columns of B, and define F to be an n by $n - m_B$ matrix whose columns are linearly independent and $BF = 0$. (In case there are multiple matrices that satisfy this condition, pick any one of them.) Then $\{x : Bx = C_B/T\} = \{x : x = Fz + \hat{x}, z \in \mathbb{R}^{n-m_B}\}$ where $\hat{x}$ satisfies $B\hat{x} = C_B/T$. Hence, ECP(θ) is equivalent to an unconstrained optimization problem $\max_{z\in\mathbb{R}^{n-m_B}} g(z;\theta) := r(Fz + \hat{x};\theta)$ in the sense that there is a one-to-one mapping between the optimizer of ECP(θ) $x^D(\theta)$ and the optimizer of the unconstrained problem $z^D(\theta)$: (1) $x^D(\theta) = Fz^D(\theta) + \hat{x}$, and (2) $z^D(\theta) = (F'F)^{-1}F'(x^D(\theta) - \hat{x})$. In addition, by Section 10.2.3 in Boyd and Vandenberghe (2004), if a feasible point of ECP(θ) $x^{(k)}$ and a feasible point of the unconstrained problem $z^{(k)}$ satisfy $x^{(k)} = Fz^{(k)} + \hat{x}$, then the Newton steps for ECP(θ) (to obtain a new feasible point $x^{(k+1)}$) and the unconstrained problem (to obtain a new feasible point $z^{(k+1)}$) coincide in the sense that $x^{(k+1)} = Fz^{(k+1)} + \hat{x}$. This relationship enables me to analyze the behavior of $x^{(k)}$ by studying $z^{(k)}$ whose convergence behavior is characterized by Theorem A.3.2 (see Section A.3.1).

Before applying Theorem A.3.2, I first show that the conditions in Theorem A.3.2 hold. Note that since Λ_θ is compact, the linear transformation of it, $\mathcal{Z}_\theta := \{z : z = (F'F)^{-1}F'(x - \hat{x}), x \in \Lambda_\theta\}$ is also compact. Also note that since $p(.;\theta) \in \mathsf{C}^2(\Lambda_\theta)$ by P1, $r(.;\theta) \in \mathsf{C}^2(\Lambda_\theta)$ and $g(.;\theta) \in \mathsf{C}^2(\mathcal{Z}_\theta)$. Hence condition (i) holds: there exists some constant L such that $||\nabla^2_{zz}g(z;\theta) - \nabla^2_{zz}g(y;\theta)||_2 \le L||z - y||_2$. Denote by $\sigma_{\min}(.), \sigma_{\max}(.)$ the smallest and the largest eigenvalues of a squared matrix. Since $\nabla^2_{zz}g(z;\theta) = F'\nabla^2_{\lambda\lambda}r(Fz+\hat{x};\theta)F$ and $-MI \preceq \nabla^2_{\lambda\lambda}r(Fz+\hat{x};\theta) \preceq -mI$ by P3, I conclude that (ii) holds: $-\bar{M}I \preceq \nabla^2_{zz}g(z;\theta) \preceq -\bar{m}I$ where $\bar{M} = M\sigma_{\max}(F'F)$ and $\bar{m} = m\sigma_{\min}(F'F)$. Then, by Theorem A.3.2, I have that there exists a constant $\eta = \min\{1, 3(1-2\alpha)\}\bar{m}^2/L$ for some $\alpha \in (0, 0.5)$ independent of θ such that if $||\nabla_z g(z^{(k)};\theta)||_2 < \eta$, then $||\nabla_z g(z^{(k+1)};\theta)||_2 < \frac{L}{2\bar{m}}||\nabla_z g(z^{(k)};\theta)||_2^2$. Note that by strong convexity of $g(.;\theta)$, $\bar{M}^{-1}||\nabla_z g(z;\theta)||_2 \le ||z - z^D(\theta)||_2 \le 2\bar{m}^{-1}||\nabla_z g(z;\theta)||_2$. Also note that for $x = Fz + \hat{x}$, $||x - x^D(\theta)||_2 \le ||F||_2||z - z^D(\theta)||_2$ and $||z - z^D(\theta)||_2 \le ||(F'F)^{-1}F'||_2||x - x^D(\theta)||_2$. Therefore,

$$
\begin{aligned}
||x^{(k+1)} - x^D(\theta)||_2 &\le ||F||_2||z^{(k+1)} - z^D(\theta)||_2 \le 2\bar{m}^{-1}||F||_2||\nabla_z g(z^{(k+1)};\theta)||_2 \\
&\le L\bar{m}^{-2}||F||_2||\nabla_z g(z^{(k)};\theta)||_2^2 \le L\bar{m}^{-2}\bar{M}||F||_2||z^{(k)} - z^D(\theta)||_2^2 \\
&\le L\bar{m}^{-2}\bar{M}||F||_2||(F'F)^{-1}F'||_2^2||x^{(k)} - x^D(\theta)||_2^2
\end{aligned}
$$

Let $\gamma = \eta$ and $\xi = L\bar{m}^{-2}\bar{M}||F||_2||(F'F)^{-1}F'||_2^2$. Note that they are both independent of θ. The result follows by letting $x^{(k+1)} = x_z^{NT}$ and $x^{(k)} = x_{z-1}^{NT}$.

Proof of Lemma A.3.6. Let $\Omega_1 = \max_{i=1,..,4}\{V_i\}$, where V_i's are positive constants to be defined later. I prove the results one by one.

(a) Let $\bar{\kappa} = \max\{\kappa, \tilde{\kappa}\}$ where κ and $\tilde{\kappa}$ are defined in Lemma 3.4.1 and Lemma A.3.1 (see Section A.3.7) respectively. Let $\Gamma_1 = \max\{1, 4\bar{\kappa}^2\}$. I proceed by induction. If $t_1 \ge \tau$, there is nothing to prove, so I consider the case when $t_1 < \tau$. Note that by DPUP algorithm, $x_1^{NT} = \lambda^D(\hat{\theta}_{t_1})$ and $x^D(\theta^*) = \lambda^D(\theta^*)$ on $\mathcal{A}$. Thus, when $t_1 < \tau$ I have

$$
\begin{aligned}
||x^D(\hat{\theta}_{t_1}) - x_1^{NT}||_2^2 &= ||x^D(\hat{\theta}_{t_1}) - \lambda^D(\hat{\theta}_{t_1})||_2^2 \\
&\le \Big(||x^D(\hat{\theta}_{t_1}) - x^D(\theta^*)||_2 + ||\lambda^D(\theta^*) - \lambda^D(\hat{\theta}_{t_1})||_2\Big)^2 \\
&\le 4\bar{\kappa}^2 E(t_1)^2 \;\le\; \Gamma_1(\log t_1)^{-\frac{\epsilon}{2}}
\end{aligned}
$$

where the last inequality follows by the definition of $\mathcal{A}$. This is my base case. I now do the inductive step. Suppose that $t_{z-1} < \tau$ and $||x^D(\hat{\theta}_{t_{z-1}}) - x_{z-1}^{NT}||_2^2 \le \Gamma_1(\log t_{z-1})^{-\epsilon/2}$. If $t_z \ge \tau$ there is nothing to prove. If $t_z < \tau$, then I need to show that $||x^D(\hat{\theta}_{t_z}) - x_z^{NT}||_2^2 \le \Gamma_1(\log t_z)^{-\epsilon/2}$ also holds. Let $V_1 > 0$ be the smallest integer satisfying $\lceil(\log V_1)^{1+\epsilon}\rceil > e^2$. Then, for

$k \geq \Omega_1 \geq V_1$, I have

$$\begin{aligned}
\left|\left|x^D(\hat{\theta}_{t_z}) - x^{NT}_{z-1}\right|\right|_2^2 &\leq 3\left|\left|x^D(\hat{\theta}_{t_z}) - x^D(\theta^*)\right|\right|_2^2 + 3\left|\left|x^D(\theta^*) - x^D(\hat{\theta}_{t_{z-1}})\right|\right|_2^2 + 3\left|\left|x^D(\hat{\theta}_{t_{z-1}}) - x^{NT}_{z-1}\right|\right|_2^2 \\
&\leq \frac{3\bar{\kappa}^2}{(\log t_z)^{\frac{\epsilon}{2}}} + \frac{3\bar{\kappa}^2}{(\log t_{z-1})^{\frac{\epsilon}{2}}} + \frac{3\Gamma_1}{(\log t_{z-1})^{\frac{\epsilon}{2}}} \\
&\leq \frac{3\bar{\kappa}^2}{(\log t_z)^{\frac{\epsilon}{2}}} + \frac{3\bar{\kappa}^2}{(\log \sqrt{t_z})^{\frac{\epsilon}{2}}} + \frac{3\Gamma_1}{(\log \sqrt{t_z})^{\frac{\epsilon}{2}}} \\
&\leq 3\left[\bar{\kappa}^2 + 2^{\frac{\epsilon}{2}}(\bar{\kappa}^2 + \Gamma_1)\right]\frac{1}{(\log t_z)^{\frac{\epsilon}{2}}},
\end{aligned}$$

where the second inequality follows by definition of $\mathcal{A}$ and induction hypothesis, the third inequality follows because $t_{z-1} \geq \frac{t_z}{2} \geq \sqrt{t_z} \geq \sqrt{t_1} = \sqrt{\lceil(\log k)^{1+\epsilon}\rceil} > e$ when $k \geq \Omega_1 \geq V_1$. Let $V_2 \geq V_1$ be such that for all $k \geq V_2$ and $z = 1, \ldots, Z$, the following holds: (1) $(\log t_z)^{\epsilon/2} \geq 3\gamma^{-2}\left[\bar{\kappa}^2 + 2^{\epsilon/2}(\bar{\kappa}^2 + \Gamma_1)\right]$ and (2) $9\xi^2\left[\bar{\kappa}^2 + 2^{\epsilon/2}(\bar{\kappa}^2 + \Gamma_1)\right]^2(\log t_z)^{-\epsilon/2} \leq 1 \leq \Gamma_1$. (Recall that γ and ξ are the constants for the locally quadratic convergence of Newton's method defined in Lemma A.3.5.) Inequality (1) ensures that $||x^D(\hat{\theta}_{t_z}) - x^{NT}_{z-1}||_2 \leq \gamma$ for all $k \geq \Omega_1 \geq V_2$ and inequality (2) ensures, by the locally quadratic convergence of the Newton's method, that $||x^D(\hat{\theta}_{t_z}) - x^{NT}_z||_2^2 \leq \xi^2||x^D(\hat{\theta}_{t_z}) - x^{NT}_{z-1}||_2^4 \leq \Gamma_1(\log t_z)^{-\epsilon/2}$. This completes the induction.

(b) First, I claim that there exist $0 < \tilde{\lambda}_{\min} < \tilde{\lambda}_{\max} < 1$ such that (1) $\tilde{\lambda}_{\min} \leq \lambda_{\min}$ and $\tilde{\lambda}_{\max} \geq \lambda_{\max}$, and (2) if $p_t \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8)$ for $t \in [t_1+1, \tau)$, then $p_t \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ for all $1 \leq t < \tau$, which will be used to prove Lemma A.3.6(c). If this is true, then Lemma 3.5.1 can be used to bound $E(t_z)$ as long as $t_z < \tau$. I now find such $\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max}$ below.

I first consider $p \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8)$. Define $V_p := \mathsf{Ball}(p^D(\theta^*), 7\phi/8)$ (note that by my notation, V_p is a *closed* ball) and $V_\lambda(\theta) := \{x \in \Lambda_\theta : x \in \lambda(p;\theta), p \in V_p\}$. Also, define $O_p := \{p \in \mathcal{P} : ||p - p^D(\theta^*)||_2 < \phi\}$ (note that this is an *open* ball) and $O_\lambda(\theta) := \{x \in \Lambda_\theta : x \in \lambda(p;\theta), p \in O_p\}$. Note that $V_p \subseteq O_p \subseteq \mathcal{P}$ by P5. This implies that $V_\lambda(\theta) \subseteq O_\lambda(\theta) \subseteq \Lambda_\theta$. In addition, since $p(.;\theta)$ is continuous in λ by P1 and O_p is an open set, $O_\lambda(\theta)$ is an open set. Therefore, $O_\lambda(\theta)$ lies in the interior of Λ_θ, and hence, $V_\lambda(\theta) \subseteq O_\lambda(\theta)$ also lies in the interior of Λ_θ. This implies that for any $\theta \in \Theta$, $\lambda_{\min}(\theta) := \inf_{p \in V_p} \min_{1 \leq j \leq n} \lambda_j(p;\theta) > 0$ and $\lambda_{\max}(\theta) := \sup_{p \in V_p} \sum_{j=1}^n \lambda_j(p;\theta) < 1$. Since Θ is compact and $\lambda_{\min}(\theta)$ and $\lambda_{\max}(\theta)$ are continuous functions, there exists some $\theta', \theta'' \in \Theta$ such that $\sup_{\theta \in \Theta} \lambda_{\max}(\theta) = \lambda_{\max}(\theta') < 1$ and $\inf_{\theta \in \Theta} \lambda_{\min}(\theta) = \lambda_{\min}(\theta'') > 0$. Hence, for all $p \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8) = V_p$ and for all θ, $1 - \sum_{j=1}^n \lambda_j(p;\theta) \geq 1 - \sup_{\theta \in \Theta} \sup_{p \in V_p} \sum_{j=1}^n \lambda_j(p;\theta) = 1 - \sup_{\theta \in \Theta} \lambda_{\max}(\theta) = 1 - \lambda_{\max}(\theta') > 0$ and $\lambda_j(p;\theta) \geq \inf_{\theta \in \Theta} \inf_{p \in V_p} \min_{1 \leq j \leq n} \lambda_j(p;\theta) \geq \inf_{\theta \in \Theta} \lambda_{\min}(\theta) \geq \lambda_{\min}(\theta'') > 0$ for all $1 \leq j \leq n$. Set $\tilde{\lambda}_{\max} = \max\{\lambda_{\max}, \lambda_{\max}(\theta')\}, \tilde{\lambda}_{\min} = \min\{\lambda_{\min}, \lambda_{\min}(\theta'')\}$ where $\lambda_{\max}$ and $\lambda_{\min}$ are as defined in S2. Note that by S2, for $p \in \{\tilde{p}^{(1)}, \ldots, \tilde{p}^{(\tilde{q})}\}$, $1 - \sum_{j=1}^n \lambda_j(p;\theta) \geq 1 - \lambda_{\max} \geq 1 - \tilde{\lambda}_{\max}$

and $\lambda_j(p;\theta) \geq \lambda_{\min} \geq \tilde{\lambda}_{\min}$ for all $1 \leq j \leq n$ and for all $\theta \in \Theta$. This completes the proof of the claim: if $t \leq t_1$, then $p_t \in \{\tilde{p}^{(1)}, \ldots, \tilde{p}^{(n)}\} \subseteq \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$; if $t_1 < t < \tau$, then $p_t \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8) \subseteq \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$.

Note that $\hat{\lambda}_t \in \Lambda_{\hat{\theta}_t}$ is equivalent to $p_t \in \mathcal{P}$ which is immediately satisfied if $p_t \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8) \subseteq \mathsf{Ball}(p^D(\theta^*), \phi) \subseteq \mathcal{P}$ (the last inequality follows by P5). This means that I only need to show $C_t \succ 0$ and $p_t \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8)$ for $t_1 \leq t < \tau$. Let $V_3 \geq V_2$ be such that for all $k \geq V_3$ and $z = 1, \ldots, Z$, $\left(2\sqrt{\Gamma_1} + 3\kappa\right)(\log t_z)^{-\epsilon/4} < \phi/(8\omega)$. I now prove the result by induction. If $\tau \leq t_1 + 1$, then there is nothing to prove. Suppose that $\tau > t_1 + 1$. Since $E(t_1) \leq \bar{\delta}$ on $\mathcal{A}$, by Lemma 3.4.1(a), $p^D(\hat{\theta}_1) \in \mathsf{Ball}(p^D(\theta^*), \phi/2)$. For $t = t_1 + 1$, I then have $||p_{t_1+1} - p^D(\theta^*)||_2 = ||p^D(\hat{\theta}_1) - p^D(\theta^*)||_2 \leq \phi/2$, so $p_{t_1+1} \in \mathcal{P}$. In addition, similar to Lemma A.3.1, I also have $C_{t_1+1} = kC - LC + LC - AS_L - A(\lambda_1^{NT} + \hat{\Delta}_{t_1+1}) \succeq (k - L - 1)C + LC - AS_L - A\hat{\Delta}_{t_1+1} \succ 0$ where the first inequality follows by the fact that $A\lambda_1^{NT} \preceq C$, and the second inequality follows by the same argument as in Lemma A.3.1. This is the base case. Now suppose $C_s \succ 0, p_s \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ for all $s \leq t-1$ for some $t - 1 < \tau$ with $t - 1 \in [t_z, t_{z+1})$. If $t \geq \tau$, there is nothing to prove. So I only need to show that $C_t \succ 0, p_t \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ when $t < \tau$. Note that when $t < \tau$, I have $t_z \leq t < \tau$. Hence, by definition of $\mathcal{A}$, I have

$$
\begin{aligned}
||p_t - p^D(\theta^*)||_2 &\leq ||p_t - p(\lambda_z^{NT}; \hat{\theta}_{t_z})||_2 + ||p(\lambda_z^{NT}; \hat{\theta}_{t_z}) - p^D(\hat{\theta}_{t_z})||_2 + ||p^D(\hat{\theta}_{t_z}) - p^D(\theta^*)||_2 \\
&\leq w||Q||_2 \left|\left| \sum_{s=t_1+1}^{t-1} \frac{\hat{\Delta}_s}{k-s} \right|\right|_2 + ||p(\lambda_z^{NT}; \hat{\theta}_{t_z}) - p(\lambda^D(\hat{\theta}_{t_z}); \hat{\theta}_{t_z})||_2 + \frac{\phi}{2} \\
&\leq \frac{\phi}{4} + \omega||\lambda_z^{NT} - \lambda^D(\hat{\theta}_{t_z})||_2 + \frac{\phi}{2} \leq \frac{\phi}{4} + \frac{\phi}{8} + \frac{\phi}{2} = \frac{7\phi}{8}
\end{aligned}
$$

where the second inequality follows by Lemma 3.4.1(a) and the fact that $E(t_z) < \bar{\delta}$ on $\mathcal{A}$, the last inequality results from the following inequality

$$
\begin{aligned}
\left|\left|\lambda_z^{NT} - \lambda^D(\hat{\theta}_{t_z})\right|\right|_2 &\leq \left|\left|\lambda_z^{NT} - \lambda^D(\theta^*)\right|\right|_2 + \left|\left|\lambda^D(\theta^*) - \lambda^D(\hat{\theta}_{t_z})\right|\right|_2 \\
&\leq 2\left|\left|x_z^{NT} - x^D(\hat{\theta}_{t_z})\right|\right|_2 + 2\left|\left|x^D(\hat{\theta}_{t_z}) - x^D(\theta^*)\right|\right|_2 + \left|\left|\lambda^D(\theta^*) - \lambda^D(\hat{\theta}_{t_z})\right|\right|_2 \\
&\leq 2\sqrt{\Gamma_1}(\log t_z)^{-\frac{\epsilon}{4}} + 3\bar{\kappa}E(t_z) \\
&\leq \left(2\sqrt{\Gamma_1} + 3\bar{\kappa}\right)(\log t_z)^{-\frac{\epsilon}{4}} < \frac{\phi}{8\omega},
\end{aligned}
$$

where the second inequality follows by (65) and the fourth inequality follows by the definition of $\mathcal{A}$. Hence, $p_t \in \mathsf{Ball}(p^D(\theta^*), 7\phi/8)$. For C_t, by a similar argument to Lemma A.3.1, I have $C_t = kC - tC + tC - AS_L - \sum_{s=t_1+1}^{t} A(\lambda_{z(s)}^{NT} - Q\sum_{v=t_1+1}^{s-1} \frac{\hat{\Delta}_v}{k-v} + \hat{\Delta}_s) \succeq (k-t)C + LC - AS_L - \sum_{s=t_1+1}^{t}(A\hat{\Delta}_s - \sum_{v=t_1+1}^{s-1} \frac{A\hat{\Delta}_v}{k-v}) \succ 0$. This completes the induction.

(c) Let $V_4 \geq V_3$ be such that $27\xi^2 \left(5\bar{\kappa}^4 \left[8\eta_4 + 4(q-1)^2(\log t_z)^2\right]/(\eta_5^2 t_z) + 2\Gamma_1\Gamma_2/(\log t_{z-1})^{\frac{\epsilon}{2}}\right) < 1$ for all $k \geq V_4$ and $z = 1, \ldots, Z$, where $\Gamma_2 = \max\{1, 4\bar{\kappa}^2\eta_3^2\}$, η_4 and η_5 are as in Lemma 3.5.1. Again, I show by induction. For $z = 1$, I have:

$$\begin{aligned}
& \mathbf{E}^\pi[||x^D(\hat{\theta}_{t_1}) - x_1^{NT}||_2^2 \mathbf{1}_{\{t_1<\tau\}}|\mathcal{A}] = \mathbf{E}^\pi[||x^D(\hat{\theta}_{t_1}) - \lambda^D(\hat{\theta}_{t_1})||_2^2 \mathbf{1}_{\{t_1<\tau\}}|\mathcal{A}] \\
\leq\ & 2\,\mathbf{E}^\pi[||x^D(\hat{\theta}_{t_1}) - x^D(\theta^*)||_2^2 \mathbf{1}_{\{t_1<\tau\}}|\mathcal{A}] + 2\,\mathbf{E}^\pi[||\lambda^D(\hat{\theta}_{t_1}) - \lambda^D(\theta^*)||_2^2 \mathbf{1}_{\{t_1<\tau\}}|\mathcal{A}] \\
\leq\ & 4\bar{\kappa}^2 \frac{\eta_3^2}{t_1} \ \leq\ \frac{\Gamma_2}{t_1},
\end{aligned}$$

where the second to the last inequality follows by Lemma 3.4.2. This is my base case. I now do the inductive step. Suppose that $\mathbf{E}^\pi[||x^D(\hat{\theta}_{t_s}) - x_s^{NT}||_2^2 \mathbf{1}_{\{t_s<\tau\}}|\mathcal{A}] \leq \Gamma_2 t_s^{-1}$ holds for $s = z-1$, I need to show that same thing holds for $s = z$. Then, for $k \geq \Omega_1 \geq V_4$, I have:

$$\begin{aligned}
& \mathbf{E}^\pi\left[\left|\left|x^D(\hat{\theta}_{t_z}) - x_z^{NT}\right|\right|_2^2 \mathbf{1}_{\{t_z<\tau\}}\middle|\,\mathcal{A}\right] \leq \xi^2 \mathbf{E}^\pi\left[\left|\left|x^D(\hat{\theta}_{t_z}) - x_{z-1}^{NT}\right|\right|_2^4 \mathbf{1}_{\{t_z<\tau\}}\middle|\,\mathcal{A}\right] \\
\leq\ & 27\xi^2 \left\{\mathbf{E}^\pi\left[\left|\left|x^D(\hat{\theta}_{t_z}) - x^D(\theta^*)\right|\right|_2^4 \mathbf{1}_{\{t_z<\tau\}}\middle|\,\mathcal{A}\right] + \mathbf{E}^\pi\left[\left|\left|x^D(\theta^*) - x^D(\hat{\theta}_{t_{z-1}})\right|\right|_2^4 \mathbf{1}_{\{t_z<\tau\}}\middle|\,\mathcal{A}\right]\right. \\
& \left. + \mathbf{E}^\pi\left[\left|\left|x^D(\hat{\theta}_{t_{z-1}}) - x_{z-1}^{NT}\right|\right|_2^4 \mathbf{1}_{\{t_z<\tau\}}\middle|\,\mathcal{A}\right]\right\} \\
\leq\ & 27\xi^2 \left\{\bar{\kappa}^4 \mathbf{E}^\pi[E(t_z)^4 \mathbf{1}_{\{t_z<\tau\}}|\mathcal{A}] + \bar{\kappa}^4 \mathbf{E}^\pi_{\theta^*}[E(t_{z-1})^4 \mathbf{1}_{\{t_z<\tau\}}|\mathcal{A}] + \frac{\Gamma_1}{(\log t_{z-1})^{\frac{\epsilon}{2}}} \frac{\Gamma_2}{t_{z-1}}\right\} \\
\leq\ & 27\xi^2 \left\{\frac{8\eta_4 + 4(q-1)^2(\log t_z)^2}{\eta_5^2 t_z^2}\bar{\kappa}^4 + \frac{8\eta_4 + 4(q-1)^2(\log t_{z-1})^2}{\eta_5^2 t_{z-1}^2}\bar{\kappa}^4 + \frac{\Gamma_1}{(\log t_{z-1})^{\frac{\epsilon}{2}}} \frac{2\Gamma_2}{t_z}\right\} \\
\leq\ & 27\xi^2 \left\{\frac{5\bar{\kappa}^4 \left[8\eta_4 + 4(q-1)^2(\log t_z)^2\right]}{\eta_5^2 t_z} + \frac{2\Gamma_1\Gamma_2}{(\log t_{z-1})^{\frac{\epsilon}{2}}}\right\} \frac{1}{t_z} \\
\leq\ & \frac{1}{t_z} \ \leq\ \frac{\Gamma_2}{t_z},
\end{aligned}$$

where the first inequality follows by Lemma A.3.6(a), the third inequality follows by Lemma A.3.4, Lemma A.3.6(a) and the induction hypothesis, and the fourth inequality holds because Lemma A.3.6(b) shows that $p_s \in \mathcal{W}(\tilde{\lambda}_{\min}, \tilde{\lambda}_{\max})$ for $s < \tau$ which means that the condition for Lemma 3.5.1 is satisfied, so

$$
\begin{aligned}
&\mathbf{E}^\pi\left[E(t)^4\mathbf{1}_{\{t<\tau\}}\middle|\mathcal{A}\right]\\
&\quad\le\int_0^\infty \mathbb{P}^\pi\left(||\hat{\theta}_t-\theta^*||_2^4\ge x\right)dx\\
&\quad\le\int_0^\infty \min\left\{1,\eta_4 t^{q-1}\exp\left(-\eta_5 t\sqrt{x}\right)\right\}dx\\
&\quad\le\int_0^{\left(\frac{2(q-1)\log t}{\eta_5 t}\right)^2}dx+\int_{\left(\frac{2(q-1)\log t}{\eta_5 t}\right)^2}^\infty\left[\eta_4 t^{q-1}\exp\left(-\frac{\eta_5 t\sqrt{x}}{2}\right)\right]\exp\left(-\frac{\eta_5 t\sqrt{x}}{2}\right)dx\\
&\quad\le\frac{4(q-1)^2(\log t)^2}{\eta_5^2 t^2}+\eta_4\int_{\left(\frac{2(q-1)\log t}{\eta_5 t}\right)^2}^\infty\exp\left(-\frac{\eta_5 t\sqrt{x}}{2}\right)dx\\
&\quad\le\frac{4(q-1)^2(\log t)^2}{\eta_5^2 t^2}+\eta_4\int_0^\infty\exp\left(-\frac{\eta_5 t\sqrt{x}}{2}\right)dx\\
&\quad\le\frac{8\eta_4+4(q-1)^2(\log t)^2}{\eta_5^2 t^2}.
\end{aligned}
$$

This completes the induction.

Proof of Lemma A.3.7. I first derive a bound for Φ_ϵ. By definition $t_z=\lceil(t_{z+1}-L)/2\rceil+L$ for $z>1$, so $t_z-L\ge(t_{z+1}-L)/2$. This implies that $t_{z+1}-t_z\le t_z$ for all $z>1$. For $z=1$, I also have $t_2-t_1=1\le L=t_1$. Recall that $Z\le\lceil\log_2 k\rceil\le 2\log_2 k$. Thus, I can bound Φ_ϵ as follows:

$$
\begin{aligned}
\Phi_\epsilon=\sum_{s=t_1+1}^{k-1}\bar{\epsilon}(s)^2 \;=\; \sum_{z=1}^{Z}(t_{z+1}-t_z)\bar{\epsilon}(t_z)^2 \;&\le\; \sum_{z=1}^{Z}(t_{z+1}-t_z)\,\eta_6^2\,\frac{(q-1)\log t_z+1}{t_z}\\
&\le\; \eta_6^2 Z[(q-1)\log k+1]\\
&\le\; K_\Phi[1+\log k+(q-1)\log^2 k]
\end{aligned}
$$

for some positive constant K_Φ independent of $k\ge1$.

I now derive a bound for Ψ_ϵ. To do that, I first show that there exists a constant $K>3$ such that for all $k\ge K$, I have (1)$(\log k)^{1+\epsilon}/k<1/19$, (2)$Z\ge3$ and (3)$t_{Z-2}\le k/3$. Note that as $k\to\infty$, I have $(\log k)^{1+\epsilon}/k\to0$, $Z\to\infty$ and $t_{z+1}-L\to\infty$ for $z=Z-2,Z-1,Z$. This implies that $t_z-L=\lceil(t_{z+1}-L)/2\rceil\le2(t_{z+1}-L)/3$ for $z=Z-2,Z-1,Z$ when k is large. Therefore, there exists a constant $K>3$ such that for all $k\ge K$, I have $(\log k)^{1+\epsilon}/k<1/19$, $Z\ge3$ and $t_{Z-2}\le\frac{8}{27}(t_{Z+1}-L)+L=\frac{8}{27}k+\frac{19}{27}(\log k)^{1+\epsilon}<\frac{k}{3}$.

Since $\bar{\epsilon}(t_z)=\eta_6\sqrt{[(q-1)\log t_z+1]/t_z}\le\eta_6\sqrt{q}$, I conclude that for $k<K$, $\Psi_\epsilon\le$

$k(k\eta_6\sqrt{q})^2 \leq K^3\eta_6^2 q$. I now focus on the case when $k \geq K$. Note that,

$$\Psi_\epsilon = \sum_{t=t_1+1}^{k-1}\left(\sum_{s=t_1+1}^{t-1}\frac{\bar\epsilon(s)}{k-s}\right)^2 \leq 2\sum_{t=t_1+1}^{k-1}\left(\sum_{s=t_1+1}^{t_{Z-2}}\frac{\bar\epsilon(s)}{k-s}\right)^2 + 2\sum_{t=t_{Z-2}+1}^{k-1}\left(\sum_{s=t_{Z-2}+1}^{t-1}\frac{\bar\epsilon(s)}{k-s}\right)^2. \tag{67}$$

Since $t_{Z-2} > k/4$ (recall that $t_{z+1} \leq 2t_z$ and $t_{Z+1} = k$), I have $\bar\epsilon(s) < \eta_6\sqrt{4[(q-1)\log k + 1]/k}$ for all $s > t_{Z-2}$. So, for all $k \geq K$, the second term in (67) can be bounded by

$$\begin{aligned} &\frac{8\eta_6^2[1+(q-1)\log k]}{k}\sum_{t=t_{Z-2}+1}^{k-1}\left(\sum_{s=t_{Z-2}+1}^{t-1}\frac{1}{k-s}\right)^2 \\ \leq\ &\frac{8\eta_6^2[1+(q-1)\log k]}{k}3k \leq K_{\Psi,2}[1+(q-1)\log k] \end{aligned}$$

for some positive constant $K_{\Psi,2} = 24\eta_6^2$ independent of $k \geq K$, where the first inequality follows by a similar argument as in (60) and $k \geq K > 3$. As for the first term in (67), for all $k \geq K$, I have

$$\begin{aligned} 2\sum_{t=t_1+1}^{k-1}\left(\sum_{s=t_1+1}^{t_{Z-2}}\frac{\bar\epsilon(s)}{k-s}\right)^2 &\leq 2k\left(\sum_{s=t_1+1}^{t_{Z-2}}\frac{\bar\epsilon(s)}{k-s}\right)^2 \\ &\leq 2k\left(\sum_{z=1}^{Z-3}\frac{t_{z+1}-t_z}{k-t_{z+1}}\eta_6\sqrt{\frac{1+(q-1)\log t_z}{t_z}}\right)^2 \\ &\leq 4k\eta_6^2\left(\sum_{z=1}^{Z-3}\frac{t_{z+1}-t_z}{k-t_{z+1}}\sqrt{\frac{1+(q-1)\log k}{t_{z+1}}}\right)^2 \\ &\leq 4k\eta_6^2[1+(q-1)\log k]\left(\int_1^{t_{Z-2}}\frac{1}{k-x}\sqrt{\frac{1}{x}}dx\right)^2 \\ &\leq 4k\eta_6^2[1+(q-1)\log k]\left(\frac{2\log(\frac{\sqrt2}{\sqrt2-1})}{\sqrt k}\right)^2 \leq K_{\Psi,1}[1+(q-1)\log k] \end{aligned}$$

where $K_{\Psi,1} = 16\eta_6^2\log^2(\frac{\sqrt2}{\sqrt2-1})$. The second inequality follows by Lemma 3.5.1. The third inequality follows because $t_{z+1} \leq 2t_z$. Note that the function $f(x) = \frac{1}{(k-x)\sqrt x}$ is decreasing when $x < \frac{k}{3}$. Since $t_{Z-2} < \frac{k}{3}$, the fourth inequality holds by integral approximation. The fifth inequality follows by

$$\begin{aligned}\int_1^{t_{Z-2}} \frac{1}{k-x}\sqrt{\frac{1}{x}}dx &= \frac{1}{\sqrt{k}}\int_1^{t_{Z-2}}\left(\frac{1}{\sqrt{k}-\sqrt{x}}+\frac{1}{\sqrt{k}+\sqrt{x}}\right)d\sqrt{x} \\ &\leq \frac{2}{\sqrt{k}}\log\left(\frac{\sqrt{k}-1}{\sqrt{k}-\sqrt{t_{Z-2}}}\right) \leq \frac{2\log(\frac{\sqrt{2}}{\sqrt{2}-1})}{\sqrt{k}}.\end{aligned}$$

Thus, I conclude that there exists some positive constant K_Ψ independent of $k \geq 1$ such that $\Psi_\epsilon \leq \max\{(K_{\Psi,1}+K_{\Psi,2})[1+(q-1)\log k], K^3\eta_6^2 q\} \leq K_\Psi[1+(q-1)\log k]$. I complete the proof by letting $K_3 = \max\{K_\Phi, K_\Psi\}$.

Proof of Lemma A.3.8. The proof of Lemma A.3.8 is very similar to that of Lemma A.3.2, with some nontrivial twists. Per the proof of Lemma A.3.2, I only need to bound $\mathbb{P}^\pi(\tau \leq t|\mathcal{A})$. Note that I have

$$\begin{aligned}&\mathbb{P}^\pi(\tau \leq t|\mathcal{A}) \\ \leq\ & \mathbb{P}^\pi\left(\max_{L+1\leq s\leq t}\left\{\left|\left|\frac{S_L - L\lambda_L\mathbf{e}}{k-s}\right|\right|_2 + \left|\left|\sum_{v=L+1}^{s}\frac{\Delta_v}{k-v}\right|\right|_2 + \sum_{v=L+1}^{s}\frac{||\lambda_v-\hat{\lambda}_v||_2\mathbf{1}_{\{v\leq\tau\}}}{k-v}\right\} \geq \psi\middle|\mathcal{A}\right) \\ \leq\ & \mathbb{P}^\pi\left(\max_{L+1\leq s\leq t}\left|\left|\frac{S_L-L\lambda_L\mathbf{e}}{k-s}\right|\right|_2 \geq \frac{\psi}{2}\middle|\mathcal{A}\right) + \mathbb{P}^\pi\left(\max_{L+1\leq s\leq t}\left|\left|\sum_{v=L+1}^{s}\frac{\Delta_v}{k-v}\right|\right|_2 \geq \frac{\psi}{4}\middle|\mathcal{A}\right) \\ &+\mathbb{P}^\pi\left(\max_{L+1\leq s\leq t}\sum_{v=L+1}^{s}\frac{||\lambda_v-\hat{\lambda}_v||_2\mathbf{1}_{\{v\leq\tau\}}}{k-v} \geq \frac{\psi}{4}\middle|\mathcal{A}\right) \\ \leq\ & \max\left\{1, \frac{4n(1+\lambda_L)^2L^2}{\psi^2(k-t)^2}\right\} + \frac{16}{\psi^2\mathbb{P}^\pi(\mathcal{A})}\left[\frac{4}{(k-t)^2}+\frac{4}{k-t}\right] \\ &+\mathbb{P}^\pi\left(\max_{L+1\leq s\leq t}\sum_{v=L+1}^{s}\frac{||\lambda_v-\hat{\lambda}_v||_2\mathbf{1}_{\{v\leq\tau\}}}{k-v} \geq \frac{\psi}{4}\middle|\mathcal{A}\right) \qquad (68)\end{aligned}$$

where the last inequality follows by the same argument in Lemma A.3.2. I now bound the last term in (68):

$$\mathbb{P}^\pi\left(\max_{L+1\le s\le t}\sum_{v=L+1}^{s}\frac{||\lambda_v-\hat{\lambda}_v||_2\mathbf{1}_{\{v\le\tau\}}}{k-v}\ge\frac{\psi}{4}\middle|\mathcal{A}\right)\le\frac{16}{\psi^2}\left(\sum_{s=L+1}^{t}\frac{\sqrt{\mathbf{E}^\pi[||\lambda_s-\hat{\lambda}_s||_2^2\mathbf{1}_{\{\tau\le s\}}|\mathcal{A}]}}{k-s}\right)^2$$

$$\le \frac{16}{\psi^2}\left(\sum_{s=L+1}^{t}\frac{\sqrt{\mathbf{E}^\pi[||\lambda_s-\hat{\lambda}_s||_2^2\mathbf{1}_{\{\tau<s\}}|\mathcal{A}]}}{k-s}+\frac{\sqrt{\mathbf{E}^\pi[||\lambda_s-\hat{\lambda}_s||_2^2\mathbf{1}_{\{\tau=s\}}|\mathcal{A}]}}{k-s}\right)^2$$

$$\le \frac{32}{\psi^2}\left(\sum_{s=L+1}^{t}\frac{\sqrt{\mathbf{E}^\pi[||\lambda_s-\hat{\lambda}_s||_2^2\mathbf{1}_{\{\tau<s\}}|\mathcal{A}]}}{k-s}\right)^2+\frac{32}{\psi^2}\left(\sum_{s=L+1}^{t}\frac{\sqrt{\mathbf{E}^\pi[||\lambda_s-\hat{\lambda}_s||_2^2\mathbf{1}_{\{\tau=s\}}|\mathcal{A}]}}{k-s}\right)^2$$

$$\le \frac{32\omega^2}{\psi^2}\left(\sum_{s=L+1}^{t}\frac{\bar{\epsilon}(s)}{k-s}\right)^2+\frac{32}{\psi^2}\left(\sum_{s=L+1}^{t}\frac{\sqrt{2}\sqrt{\mathbf{E}^\pi[\mathbf{1}_{\{\tau=s\}}|\mathcal{A}]}}{k-s}\right)^2$$

$$\le \frac{32\omega^2}{\psi^2}\left(\sum_{s=L+1}^{t}\frac{\bar{\epsilon}(s)}{k-s}\right)^2+\frac{128}{\psi^2}\left(\frac{1}{k-t}\right)$$

where the first inequality follows the same argument as in the proof of Lemma A.3.2, the fourth inequality follows by Lemma 3.5.1 and the fact that for any two points $x_1, x_2 \in \Delta^{n-1}$ I have $||x_1-x_2||_2^2 \le 2$, and the last inequality follows because by Cauchy-Schwartz inequality,

$$\left(\sum_{s=L+1}^{t}\frac{\sqrt{\mathbf{E}^\pi[\mathbf{1}_{\{\tau=s\}}|\mathcal{A}]}}{k-s}\right)^2\le\left(\sum_{s=L+1}^{t}\frac{1}{(k-s)^2}\right)\left(\sum_{s=L+1}^{t}\mathbf{E}^\pi[\mathbf{1}_{\{\tau=s\}}|\mathcal{A}]\right)\le\frac{1}{(k-t)^2}+\frac{1}{k-t}\le\frac{2}{k-t}$$

Finally, I have for all $k \ge K \ge \Omega_2 \ge 3$,

$$\begin{aligned}
\mathbf{E}^\pi[k-\tau|\mathcal{A}] &= \sum_{t=1}^{k-1}\mathbb{P}^\pi(\tau\le t|\mathcal{A})\le\frac{256}{\psi^2}\frac{\log k}{1-\mathbb{P}^\pi(E(L)>\bar{\delta})}+\left(\frac{4n(1+\lambda_L)^2}{\psi^2}+1\right)L\\
&\quad+\frac{32\omega^2}{\psi^2}\sum_{t=1}^{k-1}\left(\sum_{s=L+1}^{t}\frac{\bar{\epsilon}(s)}{k-s}\right)^2+\frac{128}{\psi^2}\sum_{t=1}^{k-1}\left(\frac{1}{k-t}\right)\\
&\le \frac{256}{\psi^2}\frac{\log k}{1-\mathbb{P}^\pi(E(L)>\bar{\delta})}+\left(\frac{4n(1+\lambda_L)^2}{\psi^2}+1\right)L\\
&\quad+\frac{64\omega^2}{\psi^2}\sum_{t=1}^{k-1}\left(\sum_{s=L+1}^{t-1}\frac{\bar{\epsilon}(s)}{k-s}\right)^2+\frac{64\omega^2}{\psi^2}\sum_{t=1}^{k-1}\frac{\bar{\epsilon}(t)^2}{(k-t)^2}+\frac{128}{\psi^2}\sum_{t=1}^{k-1}\left(\frac{1}{k-t}\right)\\
&\le \frac{512}{\psi^2}\log k+\left(\frac{4n(1+\lambda_L)^2}{\psi^2}+1\right)L+\frac{64K_3\omega^2 q}{\psi^2}\log k+\frac{128\omega^2\eta_6^2 q}{\psi^2}+\frac{128}{\psi^2}\\
&\le K_7(\log k+L)
\end{aligned}$$

where $K_7 = 640/\psi^2 + 64K_3\omega^2 q/\psi^2 + 128\omega^2\eta_6^2 q/\psi^2 + (4n(1+\lambda_L)^2/\psi^2 + 1)$, the first inequality follows by a similar argument as in Lemma A.3.2, and the third inequality follows by Lemma A.3.7 and the fact that $\bar{\epsilon}(t) \leq \eta_6\sqrt{q}$.

Proof of Lemma A.3.9. Note that, for any $\theta_1, \theta_2 \in \Theta, \theta_1 \neq \theta_2$, by Fatou's lemma, I have

$$
\begin{aligned}
\liminf_{\theta'\to\theta_1,\theta''\to\theta_2} \frac{H_t^\pi(\theta',\theta'',D_t|D_{1:t-1})}{||\theta'-\theta''||_2^2} &= \liminf_{\theta'\to\theta_1,\theta''\to\theta_2} \sum_{D_t\in\mathcal{D}} \frac{\left(\sqrt{\mathbb{P}_t^{\pi,\theta'}(D_t|D_{1:t-1})} - \sqrt{\mathbb{P}_t^{\pi,\theta''}(D_t|D_{1:t-1})}\right)^2}{||\theta'-\theta''||_2^2} \\
&\geq \sum_{D_t\in\mathcal{D}} \liminf_{\theta'\to\theta_1,\theta''\to\theta_2} \frac{\left(\sqrt{\mathbb{P}_t^{\pi,\theta'}(D_t|D_{1:t-1})} - \sqrt{\mathbb{P}_t^{\pi,\theta''}(D_t|D_{1:t-1})}\right)^2}{||\theta'-\theta''||_2^2} \\
&= \frac{H_t^\pi(\theta_1,\theta_2,D_t|D_{1:t-1})}{||\theta_1-\theta_2||_2^2} \;>\; 0, \qquad (69)
\end{aligned}
$$

where the last inequality follows by W1. Let $\underline{\sigma}(.)$ denote the smallest eigenvalues of a real symmetric matrix. If I now set $\theta_1 = \theta_2 = \theta$, since $\sqrt{\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})}$ is continuously differentiable in θ, there exists $\tilde{\theta}$ on the line segment connecting θ' and θ'' such that

$$
\begin{aligned}
&\liminf_{\theta'\to\theta,\theta''\to\theta} \frac{H_t^\pi(\theta',\theta'',D_t|D_{1:t-1})}{||\theta'-\theta''||_2^2} \\
&\geq \sum_{D_t\in\mathcal{D}} \liminf_{\theta'\to\theta,\theta''\to\theta} \left[\left(\frac{\partial}{\partial\theta}\sqrt{\mathbb{P}_t^{\pi,\tilde{\theta}}(D_t|D_{1:t-1})}\right)' \frac{\theta'-\theta''}{||\theta'-\theta''||_2}\right]^2 \\
&= \sum_{D_t\in\mathcal{D}} \liminf_{\theta'\to\theta,\theta''\to\theta} \frac{(\theta'-\theta'')'}{||\theta'-\theta''||_2} \left(\frac{\partial}{\partial\theta}\sqrt{\mathbb{P}_t^{\pi,\tilde{\theta}}(D_t|D_{1:t-1})}\right)\left(\frac{\partial}{\partial\theta}\sqrt{\mathbb{P}_t^{\pi,\tilde{\theta}}(D_t|D_{1:t-1})}\right)' \frac{\theta'-\theta''}{||\theta'-\theta''||_2} \\
&\geq \sum_{D_t\in\mathcal{D}} \liminf_{\theta'\to\theta,\theta''\to\theta} \underline{\sigma}\left(\left(\frac{\partial}{\partial\theta}\sqrt{\mathbb{P}_t^{\pi,\tilde{\theta}}(D_t|D_{1:t-1})}\right)\left(\frac{\partial}{\partial\theta}\sqrt{\mathbb{P}_t^{\pi,\tilde{\theta}}(D_t|D_{1:t-1})}\right)'\right) \\
&= \sum_{D_t\in\mathcal{D}} \underline{\sigma}\left(\left(\frac{\partial}{\partial\theta}\sqrt{\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})}\right)\left(\frac{\partial}{\partial\theta}\sqrt{\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})}\right)'\right) \\
&= \sum_{D_t\in\mathcal{D}} \frac{\underline{\sigma}\left((\frac{\partial}{\partial\theta}\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1}))(\frac{\partial}{\partial\theta}\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1}))'\right)}{4\,\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})} \\
&= \frac{1}{4}\sum_{D_t\in\mathcal{D}} \underline{\sigma}\left(\left(\frac{\partial}{\partial\theta}\log\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})\right)\left(\frac{\partial}{\partial\theta}\log\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})\right)'\right)\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1}) \\
&\geq \frac{c_f}{4} \;>\; 0 \qquad (70)
\end{aligned}
$$

where the first inequality follows by Fatou's Lemma as in (69) and the Mean Value Theorem, and the third equality follows because

$$\frac{\partial}{\partial\theta}\sqrt{\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})} = \frac{\frac{\partial}{\partial\theta}\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})}{2\sqrt{\mathbb{P}_t^{\pi,\theta}(D_t|D_{1:t-1})}}$$

(by chain rule) and the last two inequalities follow by the definition of Fisher information and W2. To prove Lemma A.3.9, it suffices to show that, for any $\theta_1, \theta_2 \in \Theta$, $H_t^{\pi}(\theta_1, \theta_2, D_t|D_{1:t-1})/||\theta_1 - \theta_2||_2^2 \geq c_h$ for some $c_h > 0$ independent of θ_1, θ_2. (If $\theta_1 = \theta_2$, the ratio is to be understood as its limit.) Suppose not, since the ratio is always non-negative, there exists two sequences $\theta_1^n \to \theta_1, \theta_2^n \to \theta_2$ such that $\liminf_{n\to\infty} H_t^{\pi}(\theta_1^n, \theta_2^n, D_t|D_{1:t-1})/||\theta_1^n - \theta_2^n||_2^2 = 0$. But, this contradicts with (69) when $\theta_1 \neq \theta_2$ and with (70) when $\theta_1 = \theta_2$. This completes the proof.

A.4 Proofs of Results in Chapter 4

A.4.1 Proof of Lemma 4.4.1

.

Let $\bar{\delta}_l := (\bar{p}_l - \underline{p}_l)/(d+1)$. The proof of Lemma 4.4.1 depends on two important lemmas, which I first state and prove later:

Lemma A.4.1 *Define $\mathcal{X} := \otimes_{l=1}^n [0, \bar{x}_l]$ where $0 < \bar{x}_l \le 1$ for all $l \in \overline{[1,n]}$. Let $f : \mathcal{X} \to \mathbb{R}$ be a function that satisfies N1-N2. Let s be a positive integer and $\bar{s}$ be as defined in N1. There exists $g \in \otimes_{l=1}^n \mathsf{P}^{(s\wedge\bar{s})-1}([0,\bar{x}_l])$ such that for any $r \in \overline{[0, s\wedge\bar{s}]}$, and any $r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}$ satisfying $\sum_{l=1}^n r_l = r$, the following holds*

$$\left|\left|\frac{\partial^r (f-g)(.)}{\partial x_1^{r_1}\dots\partial x_n^{r_n}}\right|\right|_\infty \le C_{n,r} W \left[\max_{l=1,\dots,n}\{\bar{x}_l\}\right]^{s\wedge\bar{s}-r},$$

where $C_{n,r} > 0$ only depends on n, r and W is as defined in N2.

Lemma A.4.2 *Suppose $s \ge 2$. Let $\mathcal{L}$, $\{y_{l,i}\}_{l=1,i=1}^{n,2s+d}$, $\{\beta_{l,i,j}\}_{l=1,i=1,j=1}^{n,s+d,s}$ and $\{N_{i_1,\dots,i_n}(.)\}_{i_1=1,\dots,i_n=1}^{s+d,s+d}$ be as defined in the Technical Details for Spline Approximations in* Section 4.4.1*. The following properties hold:*

a. $\mathcal{L}f = f, \forall f \in \otimes_{l=1}^n \mathsf{P}^{s-1}([\underline{p}_l, \bar{p}_l])$.

b. For all $l \in \overline{[1,n]}, i \in \overline{[1,s+d]}, j \in \overline{[1,s]}$, *I have* $|\beta_{l,i,j}| \le (y_{l,i+s} - y_{l,i})^{j-1} \le (s\bar{\delta}_l)^{j-1}$.

c. For all $i_l \in \overline{[1,s+d]}, l \in \overline{[1,n]}$, *any* $r \in \overline{[0,s-2]}$ *and any* $r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}$ *satisfying* $\sum_{l=1}^n r_l = r$, $N_{i_1,\dots,i_n}(.)$ *is nonnegative and* $\partial^r N_{i_1,\dots,i_n}(p)/(\partial p_1^{r_1}\dots\partial p_n^{r_n}) = 0$ *for all* $p \notin \otimes_{l=1}^n (y_{l,i_l}, y_{l,i_l+s})$.

d. $\sum_{i_1=1}^{s+d}\cdots\sum_{i_n=1}^{s+d} |N_{i_1,\dots,i_n}(p)| = 1$ *for all* $p \in \mathcal{P}$.

e. Fix any $r \in \overline{[0,s-2]}$ *and any* $r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}$ *satisfying* $\sum_{l=1}^n r_l = r$,

$$\left|\left|\frac{\partial^r N_{i_1,\dots,i_n}(.)}{\partial p_1^{r_1}\dots\partial p_n^{r_n}}\right|\right|_\infty \le C_{r,s}\left[\min_{l=1,\dots,n}\{\bar{\delta}_l\}\right]^{-r}$$

where $C_{r,s} > 0$ is a constant that only depends on r and s.

I first discuss the meaning of the two lemmas above. Since a spline function is essentially a sequence of local polynomial functions attached together, to understand its approximation accuracy, I need to first answer the following question: Suppose that I use a polynomial function g to approximate a deterministic function f on a small region, how does the approximation error depend on the smoothness index of f, the degree of g, and the size of the region? Lemma A.4.1 derives a bound for approximation error as a function of these factors. Lemma A.4.2 summarizes some useful properties of the spline function constructed using the B-Spline approach (see Section 4.4.1 for more details). I now proceed to prove Lemma 4.4.1.

Let $K' = \max\{K_1, K_2, K_3\}$ where the constants K_1 is defined below and K_2, K_3 are defined later in Step 3 (below (82)). Let $K = \exp(\log^2 K')$. Since $L_0 \geq \log^3 k$, $d \to \infty$ as $k \to \infty$. So there exists a constant $K_1 \geq 3$ such that for all $k \geq K_1$ and for all $l \in \overline{[1, n]}$, $2s\bar{\delta}_l \leq 1$. This observation allows me to invoke Lemma A.4.1 later. Note that for $k \leq K'$, the desired result holds because for any $x > 0$,

$$\mathbb{P}\left(\left\|\frac{\partial^r(\lambda_j^*(p) - \tilde{\lambda}_j(p))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}}\right\|_\infty \geq x\right) \leq 1 = K\exp(-\log^2 K').$$

Hence, in the remaining of the proof, I will focus only on the case when $k > K'$. I proceed in several steps:

<u>Step 1</u>

My objective in this step is to compute an upper bound for

$$\left\|\frac{\partial^r(\lambda_j^*(.) - \tilde{\lambda}_j(.))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}}\right\|_\infty.$$

Fix some $\tilde{i}_l \in \overline{[s, s+d]}$ for all $l \in \overline{[1, n]}$, $j \in \overline{[1, n]}$ and $r \in \overline{[0, (s-2) \wedge \bar{s}]}$. Define two hypercubes $H_{\tilde{i}_1,\dots,\tilde{i}_n} := \otimes_{l=1}^n [y_{l,\tilde{i}_l}, y_{l,\tilde{i}_l+1}]$ and $\tilde{H}_{\tilde{i}_1,\dots,\tilde{i}_n} := \otimes_{l=1}^n [y_{l,\tilde{i}_l-s+1}, y_{l,\tilde{i}_l+s}]$. For any $p \in H_{\tilde{i}_1,\dots,\tilde{i}_n}$, and any $r_l \in \mathbb{Z}_+, l \in \overline{[1, n]}$ satisfying $\sum_{l=1}^n r_l = r$, I have:

$$\left|\frac{\partial^r(\lambda_j^*(p) - \tilde{\lambda}_j(p))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}}\right| \leq \left|\frac{\partial^r(\lambda_j^*(p) - \mathcal{L}\lambda_j^*(p))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}}\right| + \left|\frac{\partial^r(\mathcal{L}\lambda_j^*(p) - \tilde{\lambda}_j(p))}{\partial p_1^{r_1} \dots \partial p_n^{r_n}}\right|. \tag{71}$$

I now bound the terms after the inequality separately.

Bounding the first term in (71). Let $\mathcal{X} = \tilde{H}_{\tilde{i}_1,\dots\tilde{i}_n}$. Since $2s\bar{\delta}_l \leq 1$ for $k \geq K' \geq K_1$, by Lemma A.4.1, there exists $g \in \otimes_{l=1}^n \mathsf{P}^{(s\wedge\bar{s})-1}([\underline{p}_l, \bar{p}_l])$ such that for all $p \in \tilde{H}_{\tilde{i}_1,\dots,\tilde{i}_n}$ and

$r \in \overline{[0, (s-2)\wedge\bar{s}]}$,

$$\begin{aligned}
\left|\frac{\partial^r(\lambda_j^*(p) - g(p))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right| &\leq C_{n,r} W\left[\max_{l=1,\ldots,n}\{2s\bar{\delta}_l\}\right]^{s\wedge\bar{s}-r} \\
&\leq C_{n,r} W(2s)^{s\wedge\bar{s}-r}\left[\max_{l=1,\ldots,n}\left\{\frac{\bar{p}_l - \underline{p}_l}{d}\right\}\right]^{s\wedge\bar{s}-r}, \qquad (72)
\end{aligned}$$

where $C_{n,r}$ is a positive constant that only depends on n and r. Note that for all $i_l \in \overline{[\tilde{i}_l - s + 1, \tilde{i}_l]}$ and for all $r_l \in \overline{[1, s]}$, $l \in \overline{[1, n]}$, I have $(\tau_{1,i_1,r_1}, \ldots, \tau_{n,i_n,r_n}) \in \tilde{H}_{\tilde{i}_1,\ldots,\tilde{i}_n}$. Thus, there exists a constant C_0 independent of k such that for any $i_l \in \overline{[\tilde{i}_l - s + 1, \tilde{i}_l]}$, $l \in \overline{[1, n]}$, the following holds:

$$\begin{aligned}
&\left|\gamma_{i_1,\ldots,i_n}\lambda_j^* - \gamma_{i_1,\ldots,i_n} g\right| \\
&\leq \sum_{j_1=1}^{s}\sum_{r_1=1}^{j_1}\cdots\sum_{j_n=1}^{s}\sum_{r_n=1}^{j_n}\frac{\left(\prod_{l=1}^n \beta_{l,i_l,j_l}\right)\left|\lambda_j^*(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n}) - g(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})\right|}{\left|\prod_{l=1}^n\prod_{s_l=1,s_l\neq r_l}^{j_l}(\tau_{l,i_l,r_l} - \tau_{l,i_l,s_l})\right|} \\
&\leq \sum_{j_1=1}^{s}\sum_{r_1=1}^{j_1}\cdots\sum_{j_n=1}^{s}\sum_{r_n=1}^{j_n}\frac{\prod_{l=1}^n(\bar{\delta}_l s)^{j_l-1}}{\prod_{l=1}^n(\bar{\delta}_l/s)^{j_l-1}}\left|\lambda_j^*(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n}) - g(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})\right| \\
&= \sum_{j_1=1}^{s}\sum_{r_1=1}^{j_1}\cdots\sum_{j_n=1}^{s}\sum_{r_n=1}^{j_n} s^{2(\sum_{l=1}^n j_l - n)}\left|\lambda_j^*(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n}) - g(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})\right| \\
&\leq \left(\frac{s+s^2}{2}\right)^n s^{2(ns-n)} C_{n,0} W(2s)^{s\wedge\bar{s}}\left[\max_{l=1,\ldots,n}\left\{\frac{\bar{p}_l - \underline{p}_l}{d}\right\}\right]^{s\wedge\bar{s}} \leq \frac{C_0}{d^{s\wedge\bar{s}}}, \qquad (73)
\end{aligned}$$

where the first inequality follows by the definition of $\gamma_{i_1,\ldots,i_n}$, the second inequality follows by Lemma A.4.2 part (b), and the third inequality follows by (72). Then, for any $p \in H_{\tilde{i}_1,\ldots,\tilde{i}_n}$, I have:

$$\begin{aligned}
\left|\frac{\partial^r(\mathcal{L}\lambda_j^*(p) - \mathcal{L}g(p))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right| &\leq \sum_{i_1=1}^{s+d}\cdots\sum_{i_n=1}^{s+d}\left|\gamma_{i_1,\ldots,i_n}\lambda_j^* - \gamma_{i_1,\ldots,i_n} g\right|\left|\frac{\partial^r N_{i_1,\ldots,i_n}(p)}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right| \\
&= \sum_{i_1=\tilde{i}_1-s+1}^{\tilde{i}_1}\cdots\sum_{i_n=\tilde{i}_n-s+1}^{\tilde{i}_n}\left|\gamma_{i_1,\ldots,i_n}\lambda_j^* - \gamma_{i_1,\ldots,i_n} g\right|\left|\frac{\partial^r N_{i_1,\ldots,i_n}(p)}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right| \\
&\leq s^n\frac{C_0}{d^{s\wedge\bar{s}}}C_{r,s}\left[\min_{l=1,\ldots,n}\{\bar{\delta}_l\}\right]^{-r} \leq \frac{s^n C_0 C_{r,s}}{d^{s\wedge\bar{s}}}\left[\min_{l=1,\ldots,n}\left\{\frac{\bar{p}_l - \underline{p}_l}{2d}\right\}\right]^{-r} \\
&\leq 2^r s^n C_0 C_{r,s}\left[\min_{l=1,\ldots,n}\left\{\bar{p}_l - \underline{p}_l\right\}\right]^{-r} d^{r-s\wedge\bar{s}}, \qquad (74)
\end{aligned}$$

where the equality follows because, by Lemma A.4.2 part (c), $\partial^r N_{i_1,\ldots,i_n}(p)/(\partial p_1^{r_1}\ldots\partial p_n^{r_n})=0$ for $p\notin\otimes_{l=1}^n(y_{l,i_l},y_{l,i_l+s})$, the second inequality follows by (73) and Lemma A.4.2 part (e), and the third inequality follows since $d+1\le 2d$.

Putting things together, by Lemma A.4.2 part (a) (note that $s\wedge\bar{s}\le s$), (72) and (74), I have the following inequality for all $p\in H_{\tilde{i}_1,\ldots,\tilde{i}_n}$:

$$\begin{aligned}
\left|\frac{\partial^r(\lambda_j^*(p)-\mathcal{L}\lambda_j^*(p))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right| &\le \left|\frac{\partial^r(\lambda_j^*(p)-g(p))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right|+\left|\frac{\partial^r(g(p)-\mathcal{L}g(p))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right|+\left|\frac{\partial^r(\mathcal{L}g(p)-\mathcal{L}\lambda_j^*(p))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right| \\
&\le \left[C_{n,r}W\left[2s\max_{1\le l\le n}\{\bar{p}_l-\underline{p}_l\}\right]^{s\wedge\bar{s}-r}+2^r s^n C_0 C_{r,s}\left[\min_{l=1,\ldots,n}\left\{\bar{p}_l-\underline{p}_l\right\}\right]^{-r}\right]\frac{1}{d^{s\wedge\bar{s}-r}} \\
&\le \frac{C_1}{d^{s\wedge\bar{s}-r}}, \qquad (75)
\end{aligned}$$

for some C_1 independent of k. Since the right hand side of (75) does not depend on $\tilde{i}_1,\ldots,\tilde{i}_n$, the inequality holds uniformly for all $p\in\mathcal{P}$.

Bounding the second term in (71). Define $\xi^j_{i_1,\ldots,i_n}:=\max_{1\le r_1,\ldots,r_n\le s}\{|\lambda_j^*(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})-\tilde{\lambda}_j(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})|\}$ and $\xi^j:=\max_{1\le i_1,\ldots,i_n\le s+d}\{\xi^j_{i_1,\ldots,i_n}\}$. For any $i_l\in\overline{[\tilde{i}_l-s+1,\tilde{i}_l]}$, $l\in\overline{[1,n]}$, by similar argument as in (73), I have:

$$\begin{aligned}
&\left|\gamma_{i_1,\ldots,i_n}\lambda_j^*-c^j_{i_1,\ldots,i_n}\right| \\
&\quad\le \sum_{j_1=1}^{s}\sum_{r_1=1}^{j_1}\cdots\sum_{j_n=1}^{s}\sum_{r_n=1}^{j_n}\frac{\left(\prod_{l=1}^n\beta_{l,i_l,j_l}\right)|\lambda_j^*(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})-\tilde{\lambda}_j(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})|}{\left|\prod_{l=1}^n\prod_{s_l=1,s_l\ne r_l}^{j_l}(\tau_{l,i_l,r_l}-\tau_{l,i_l,s_l})\right|} \\
&\quad\le \sum_{j_1=1}^{s}\sum_{r_1=1}^{j_1}\cdots\sum_{j_n=1}^{s}\sum_{r_n=1}^{j_n} s^{2(\sum_{l=1}^n j_l-n)}|\lambda_j^*(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})-\tilde{\lambda}_j(\tau_{1,i_1,r_1},\ldots,\tau_{n,i_n,r_n})| \\
&\quad\le \left(\frac{s+s^2}{2}\right)^n s^{2(ns-n)}\,\xi^j_{i_1,\ldots,i_n}.
\end{aligned}$$

So, for all $p\in H_{\tilde{i}_1,\ldots,\tilde{i}_n}$, there exists constant $C_2>0$ independent of k and $\tilde{i}_1,\ldots,\tilde{i}_n$ such that

$$\begin{aligned}
\left|\frac{\partial^r(\mathcal{L}\lambda_j^*(p)-\tilde{\lambda}_j(p))}{\partial p_1^{r_1}\dots\partial p_n^{r_n}}\right| &\leq \sum_{i_1=1}^{s+d}\cdots\sum_{i_n=1}^{s+d}\left|\gamma_{i_1,\dots,i_n}\lambda_j^*-c^j_{i_1,\dots,i_n}\right|\;\left|\frac{\partial^r N_{i_1,\dots,i_n}(p)}{\partial p_1^{r_1}\dots\partial p_n^{r_n}}\right| \\
&= \sum_{i_1=\tilde{i}_1-s+1}^{\tilde{i}_1}\cdots\sum_{i_n=\tilde{i}_n-s+1}^{\tilde{i}_n}\left|\gamma_{i_1,\dots,i_n}\lambda_j^*-c^j_{i_1,\dots,i_n}\right|\;\left|\frac{\partial^r N_{i_1,\dots,i_n}(p)}{\partial p_1^{r_1}\dots\partial p_n^{r_n}}\right| \\
&\leq s^n\left(\frac{s+s^2}{2}\right)^n s^{2(ns-n)}\xi^j C_{r,s}\left[\min_{l=1,\dots,n}\{\bar{\delta}_l\}\right]^{-r} \\
&\leq \left(\frac{1+s}{2}\right)^n s^{2ns}\xi^j C_{r,s}\left[\min_{l=1,\dots,n}\left\{\frac{\bar{p}_l-\underline{p}_l}{2d}\right\}\right]^{-r} \\
&= 2^{r-n}C_{r,s}\,(1+s)^n\,s^{2ns}\left[\min_{l=1,\dots,n}\left\{\bar{p}_l-\underline{p}_l\right\}\right]^{-r}\xi^j d^r = C_2\xi^j d^r, \quad (76)
\end{aligned}$$

where the second inequality follows by Lemma A.4.2 part (e) and $C_{r,s}$ only depends on r and s, the third inequality follows since $d+1\leq 2d$.

Note that the right hand side of (76) does not depend on $\tilde{i}_1,\dots,\tilde{i}_n$, (76) holds uniformly for all $p\in\mathcal{P}$. So, by (71), (75) and (76), I conclude that there exists C_3 independent of k such that

$$\begin{aligned}
&\left\|\frac{\partial^r(\lambda_j^*(.)-\tilde{\lambda}_j(.))}{\partial p_1^{r_1}\dots\partial p_n^{r_n}}\right\|_\infty = \sup_{p\in\mathcal{P}}\left|\frac{\partial^r(\lambda_j^*(p)-\tilde{\lambda}_j(p))}{\partial p_1^{r_1}\dots\partial p_n^{r_n}}\right| \\
&\quad\leq \max_{s\leq\tilde{i}_1,\dots,\tilde{i}_n\leq s+d}\left\{\sup_{p\in H_{\tilde{i}_1,\dots,\tilde{i}_n}}\left\{\left|\frac{\partial^r(\lambda_j^*(p)-\mathcal{L}\lambda_j^*(p))}{\partial p_1^{r_1}\dots\partial p_n^{r_n}}\right|+\left|\frac{\partial^r(\mathcal{L}\lambda_j^*(p)-\tilde{\lambda}_j(p))}{\partial p_1^{r_1}\dots\partial p_n^{r_n}}\right|\right\}\right\} \\
&\quad\leq \left(\frac{C_1}{d^{s\wedge\bar{s}-r}}+C_2\xi^j d^r\right)\;\leq\; C_3\left(\frac{1}{d^{s\wedge\bar{s}-r}}+\xi^j d^r\right). \quad (77)
\end{aligned}$$

<u>Step 2</u>

I now analyze the term ξ^j. Note that $\xi^j=\max_{p\in\tilde{\mathcal{G}}}|\lambda_j^*(p)-\tilde{\lambda}_j(p)|$ where $\tilde{\mathcal{G}}:=\{(\tau_{1,i_1,j_1};\dots;\tau_{n,i_n,j_n}): i_l\in\overline{[1,s+d]}, j_l\in\overline{[1,s]},\forall l\in\overline{[1,n]}\}$ is as defined in Section 4.4.1. So, for all $x\geq 0$, I can bound

$$\mathbb{P}\left(\max_{p\in\tilde{\mathcal{G}}}|\lambda_j^*(p)-\tilde{\lambda}_j(p)|\geq x\right)\leq\mathbb{P}\left(\max_{p\in\tilde{\mathcal{G}}}\{\tilde{\lambda}_j(p)-\lambda_j^*(p)\}\geq x\right)+\mathbb{P}\left(\max_{p\in\tilde{\mathcal{G}}}\{\lambda_j^*(p)-\tilde{\lambda}_j(p)\}\geq x\right). \quad (78)$$

I now bound the two terms after the inequality separately. For $x\geq 0$ and $t>0$, since

$|\tilde{\mathcal{G}}| = s^n(s+d)^n$, the following holds:

$$\begin{aligned}
\mathbb{P}\left(\max_{p\in\tilde{\mathcal{G}}}\{\tilde{\lambda}_j(p) - \lambda_j^*(p)\} \ge x\right) &= \mathbb{P}\left(t\max_{p\in\tilde{\mathcal{G}}}\{\tilde{\lambda}_j(p) - \lambda_j^*(p)\} \ge tx\right) \\
&\le \exp(-tx)\mathbf{E}\left[\exp\left(t\max_{p\in\tilde{\mathcal{G}}}\{\tilde{\lambda}_j(p) - \lambda_j^*(p)\}\right)\right] \\
&\le \exp(-tx)\sum_{p\in\tilde{\mathcal{G}}}\mathbf{E}\left[\exp\left(t(\tilde{\lambda}_j(p) - \lambda_j^*(p))\right)\right] \\
&\le \exp(-tx)s^n(s+d)^n\max_{p\in\tilde{\mathcal{G}}}\left\{\mathbf{E}\left[\exp\left(t(\tilde{\lambda}_j(p) - \lambda_j^*(p))\right)\right]\right\}. \quad (79)
\end{aligned}$$

Note that there exists a $p^* \in \tilde{\mathcal{G}}$ such that the expectation $\mathbf{E}\left[\exp\left(t(\tilde{\lambda}_j(p) - \lambda_j^*(p))\right)\right]$ in (79) attains its maximum. So, for all $0 < t \le L_0$,

$$\begin{aligned}
\max_{p\in\tilde{\mathcal{G}}}\left\{\mathbf{E}\left[\exp\left(t(\tilde{\lambda}_j(p) - \lambda_j^*(p))\right)\right]\right\} &= \mathbf{E}\left[\exp\left(t(\tilde{\lambda}_j(p^*) - \lambda_j^*(p^*))\right)\right] \\
&= \exp(-t\lambda_j^*(p^*))\left\{\mathbf{E}\left[\exp\left(\frac{t}{L_0}\mathsf{Bernoulli}(\lambda_j^*(p^*))\right)\right]\right\}^{L_0} \\
&= \exp(-t\lambda_j^*(p^*))\left\{1 - \lambda_j^*(p^*) + \lambda_j^*(p^*)\exp\left(\frac{t}{L_0}\right)\right\}^{L_0} \\
&\le \exp(-t\lambda_j^*(p^*))\left\{\exp\left(\lambda_j^*(p^*)\left[\exp\left(\frac{t}{L_0}\right) - 1\right]\right)\right\}^{L_0} \\
&= \exp(-t\lambda_j^*(p^*))\exp\left(\lambda_j^*(p^*)L_0\sum_{j=1}^{\infty}\frac{1}{j!}\left(\frac{t}{L_0}\right)^j\right) \\
&= \exp\left(\lambda_j^*(p^*)L_0\sum_{j=2}^{\infty}\frac{1}{j!}\left(\frac{t}{L_0}\right)^j\right) \\
&\le \exp\left(\lambda_j^*(p^*)t^2/L_0\right) \le \exp(t^2/L_0), \quad (80)
\end{aligned}$$

where the second equality follows because $\tilde{\lambda}_j(p^*)$ is the average of L_0 independent Bernoulli random variables with success probability $\lambda_j^*(p^*)$ and the last inequality follows from the fact that $\sum_{j=2}^{\infty}(j!)^{-1}(t/L_0)^j \le (t/L_0)^2\sum_{j=2}^{\infty}[j(j-1)]^{-1} \le (t/L_0)^2$. Hence, by (79) and (80), for all $0 < t \le L_0$,

$$\mathbb{P}\left(\max_{p\in\tilde{\mathcal{G}}}\{\tilde{\lambda}_j(p) - \lambda_j^*(p)\} \ge x\right) \le \exp(t^2/L_0 - tx + \log(s^n(s+d)^n)). \quad (81)$$

Following similar arguments, for all $0 < t \le L_0$, there exists some $q^* \in \tilde{\mathcal{G}}$ such that

$$\begin{aligned}
\mathbb{P}\left(\max_{p\in\tilde{\mathcal{G}}}\{\lambda_j^*(p) - \tilde{\lambda}_j(p)\} \ge x\right) &\le \exp(-tx)\left[\max_{p\in\tilde{\mathcal{G}}}\left\{\mathbf{E}\left[\exp\left(t(\lambda_j^*(p) - \tilde{\lambda}_j(p))\right)\right]\right\}\right] s^n(s+d)^n \\
&\le \exp(-tx)\left[\exp(t\lambda_j^*(q^*))\exp\left(\lambda_j^*(q^*)L_0\sum_{j=1}^{\infty}\frac{(-1)^j}{j!}\left(\frac{t}{L_0}\right)^j\right)\right] s^n(s+d)^n \\
&\le \exp(\lambda_j^*(q^*)t^2/L_0 - tx)s^n(s+d)^n \\
&\le \exp(t^2/L_0 - tx + \log(s^n(s+d)^n)). \qquad (82)
\end{aligned}$$

Pick $x = 4L_0^{-1/2}(s+d)^{n/2}s^{n/2}\log k$ and $t = L_0x/2$. I now show that under this choice of x and t, the inequalities (81) and (82) hold for large k, i.e., $t \le L_0$ when k is large. Recall that I have set $d = \lceil (L_0^{1/2}/\log k)^{1/(s+n/2)}\rceil$. Since $L_0 \ge \log^3 k$, for $k \ge 3$, I have $L_0^{1/2}/\log k \ge 1$. This implies that

$$(L_0^{1/2}/\log k)^{1/(s+n/2)} \le d \le 2(L_0^{1/2}/\log k)^{1/(s+n/2)}. \qquad (83)$$

I then have that for all $k \ge 3$, the following holds

$$x = \frac{4\log k}{\sqrt{L_0}}(s+d)^{\frac{n}{2}}s^{\frac{n}{2}} \le \frac{4\log k}{\sqrt{L_0}}(s+1)^{\frac{n}{2}}s^{\frac{n}{2}}d^{\frac{n}{2}} \le 4(s+1)^{\frac{n}{2}}s^{\frac{n}{2}}2^{\frac{n}{2}}\left(\frac{\log k}{\sqrt{L_0}}\right)^{\frac{s}{s+n/2}} \to 0, \text{ as } k \to \infty.$$

Hence, there exists a constant $K_2 \ge 3$ such that for all $k \ge K_2 \ge 3$, I have $x \le 2$ and hence $t = L_0x/2 \le L_0$. The following inequality holds for $k \ge K_2$:

$$\begin{aligned}
&\mathbb{P}\left(\xi^j \ge 4L_0^{-1/2}d^{\frac{n}{2}}(s+s^2)^{\frac{n}{2}}\log k\right) \le \mathbb{P}\left(\xi^j \ge 4L_0^{-1/2}s^{\frac{n}{2}}(s+d)^{\frac{n}{2}}\log k\right) \\
&= \mathbb{P}\left(\max_{p\in\tilde{\mathcal{G}}}|\tilde{\lambda}_j(p) - \lambda_j^*(p)| \ge 4L_0^{-1/2}s^{\frac{n}{2}}(s+d)^{\frac{n}{2}}\log k\right) \\
&\le 2\exp\left(-\frac{x^2L_0}{4}\right)s^n(s+d)^n = 2s^n(s+d)^n\exp\left(-4s^n(s+d)^n\log^2 k\right) \\
&\le 2s^n(s+d)^n\exp\left(-2s^n(s+d)^n\right)\exp\left(-\log^2 k\right) \le K_3\exp\left(-\log^2 k\right), \qquad (84)
\end{aligned}$$

where $K_3 = \sup_{x\ge 0}\{x\exp(-x)\}$ is a positive constant, the first inequality follows since $s + d \le (s+1)d$ for $d \ge 1$, the second inequality follows by (78), (81) and (82). Let

$\Psi_r = C_3(1 + 2^{n/2+r+2}(s+s^2)^{n/2})$ which is independent of k (C_3 is defined in (77)), then

$$\begin{aligned}\Psi_r\left(\frac{\log k}{\sqrt{L_0}}\right)^{\frac{s\wedge\bar{s}-r}{s+n/2}} &\geq C_3\left[\left(\frac{\log k}{\sqrt{L_0}}\right)^{\frac{s\wedge\bar{s}-r}{s+n/2}} + 2^{\frac{n}{2}+r+2}(s+s^2)^{\frac{n}{2}}\left(\frac{\log k}{\sqrt{L_0}}\right)^{\frac{s-r}{s+n/2}}\right] \\ &\geq C_3\left(\frac{1}{d^{s\wedge\bar{s}-r}} + \frac{4\log k}{\sqrt{L_0}}(s+s^2)^{\frac{n}{2}}d^{\frac{n}{2}+r}\right), \end{aligned} \tag{85}$$

where the first inequality follows since $\log k/L_0^{1/2} \leq 1$ and the second inequality follows by (83). So,

$$\begin{aligned}&\mathbb{P}\left(\left\|\frac{\partial^r(\lambda_j^*(p)-\tilde{\lambda}_j(p))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right\|_\infty \geq \Psi_r\left(\frac{\log k}{\sqrt{L_0}}\right)^{\frac{s\wedge\bar{s}-r}{s+n/2}}\right) \\ &\leq \mathbb{P}\left(\left\|\frac{\partial^r(\lambda_j^*(p)-\tilde{\lambda}_j(p))}{\partial p_1^{r_1}\ldots\partial p_n^{r_n}}\right\|_\infty \geq C_3\left(\frac{1}{d^{s\wedge\bar{s}-r}} + \frac{4\log k}{\sqrt{L_0}}(s+s^2)^{\frac{n}{2}}d^{\frac{n}{2}+r}\right)\right) \\ &\leq \mathbb{P}\left(C_3\left(\frac{1}{d^{s\wedge\bar{s}-r}} + \xi^j d^r\right) \geq C_3\left(\frac{1}{d^{s\wedge\bar{s}-r}} + \frac{4\log k}{\sqrt{L_0}}(s+s^2)^{\frac{n}{2}}d^{\frac{n}{2}+r}\right)\right) \\ &\leq \mathbb{P}\left(\xi^j \geq \frac{4\log k}{\sqrt{L_0}}d^{\frac{n}{2}}(s+s^2)^{\frac{n}{2}}\right) \leq K_3\exp(-\log^2 k) \leq K\exp(-\log^2 k),\end{aligned}$$

where the first inequality follows by (85), the second inequality follows by (77) and the fourth inequality follows by (84). This completes the proof.

Proof of Lemma A.4.1: For $k \in \mathbb{Z}_{++}$, define $\mathcal{I}^k := \{a = (a_1; \ldots; a_n) : a_l \in \overline{[0,k]}, \text{for all } l \in \overline{[1,n]}, \text{ and } \sum_{l=1}^n a_l = k\}$. Define $g(x) = \sum_{k=0}^{(s\wedge\bar{s})-1}\sum_{a\in\mathcal{I}^k} h_f(x,\mathbf{0},a)$ for all $x \in \mathcal{X}$ where

$$h_f(x,y,a) := \frac{\partial^{a_1+\cdots+a_n}f(y)}{\partial x_1^{a_1}\ldots\partial x_n^{a_n}}\prod_{l=1}^n\frac{(x_l-y_l)^{a_l}}{a_l!}, \;\; \forall y \in \mathcal{X}.$$

It is straightforward to verify that $g(.) \in \otimes_{l=1}^n \mathsf{P}^{s\wedge\bar{s}-1}([0,\bar{x}_l])$. For some $k \in \overline{[0, s\wedge\bar{s}]}$, consider any $a = (a_1; \ldots; a_n) \in \mathcal{I}^k$, any $x, y \in \mathcal{X}$, any $r \in \overline{[0,k]}$, and any $r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}$ satisfying $\sum_{l=1}^n r_l = r$. If $a_l < r_l$ for some $l \in \overline{[1,n]}$, $\partial^r h_f(x,y,a)/(\partial x_1^{r_1}\ldots\partial x_n^{r_n}) = 0$. Hence, the following hold:

$$\left|\frac{\partial^r h_f(x,y,a)}{\partial x_1^{r_1}\ldots\partial x_n^{r_n}}\right| = \begin{cases} 0, & \text{if } a_l < r_l, \forall l \in \overline{[1,n]}; \\ \left|\frac{\partial^{a_1+\cdots+a_n}f(y)}{\partial x_1^{a_1}\ldots\partial x_n^{a_n}}\prod_{l=1}^n\frac{(x_l-y_l)^{a_l-r_l}}{(a_l-r_l)!}\right| \leq \frac{W}{\prod_{l=1}^n(a_l-r_l)!}\left[\max_{l=1,\ldots,n}\{\bar{x}_l\}\right]^{k-r}, & \text{otherwise,}\end{cases} \tag{86}$$

where the inequality follows by N2, $\bar{x}_l \leq 1$ and $a_l - r_l \geq 0$ for all $l \in \overline{[1,n]}$, and $\sum_{l=1}^n(a_l - r_l) =$

$k-r$. So, for any $r \in \overline{[0, s \wedge \bar{s}]}$ and for any $r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}$ satisfying $\sum_{l=1}^n r_l = r$,

$$\begin{aligned}
\left\|\frac{\partial^r (f-g)(.)}{\partial x_1^{r_1}\dots\partial x_n^{r_n}}\right\|_\infty &= \sup_{x\in\mathcal{X}}\left|\frac{\partial^r (f-g)(x)}{\partial x_1^{r_1}\dots\partial x_n^{r_n}}\right| \le \sum_{a\in\mathcal{I}^{s\wedge\bar{s}}}\sup_{x,y\in\mathcal{X}}\left|\frac{\partial^r h_f(x,y,a)}{\partial x_1^{r_1}\dots\partial x_n^{r_n}}\right| \\
&\le W\left[\max_{l=1,\dots,n}\{\bar{x}_l\}\right]^{s\wedge\bar{s}-r}\sum_{\substack{\sum_{l=1}^n a_l = s\wedge\bar{s}\\ a_l\ge r_l,\forall l\in\overline{[1,n]}}}\frac{1}{\prod_{l=1}^n (a_l-r_l)!} \\
&= W\left[\max_{l=1,\dots,n}\{\bar{x}_l\}\right]^{s\wedge\bar{s}-r}\frac{1}{(s\wedge\bar{s}-r)!}\sum_{\substack{\sum_{l=1}^n w_l = s\wedge\bar{s}-r\\ w_l\in\mathbb{Z}_+,\forall l\in\overline{[1,n]}}}\frac{(s\wedge\bar{s}-r)!}{\prod_{l=1}^n w_l!} \\
&= W\left[\max_{l=1,\dots,n}\{\bar{x}_l\}\right]^{s\wedge\bar{s}-r}\frac{n^{s\wedge\bar{s}-r}}{(s\wedge\bar{s}-r)!}
\end{aligned}$$

where the first inequality follows by the Lagrangian remainder formula, the second inequality follows by (86), and the last inequality follows by the multinomial theorem. The result follows by letting $C_{n,r} = \sup_{k\ge r} n^{k-r}/(k-r)! \le n^n/n! < \infty$.

Proof of Lemma A.4.2: Recall that $N^s_{l,i_l}(.)$ are defined in Section 4.4.1 as the building blocks of the tensor-product B-Spline basis functions. Let $D^\sigma, D^\sigma_+, D^\sigma_-$ respectively denote the σ^{th} order derivative, right derivative, left derivative of a single variate real function. I first state some known results of spline functions that will be used to prove Lemma A.4.2.

Theorem A.4.1 (Theorem 6.18 in Schumaker (2007)) *For any $l \in \overline{[1,n]}$, $\mathcal{L}_l f = f$ for all $f \in \mathsf{P}^{s-1}([\underline{p}_l, \bar{p}_l])$.*

Theorem A.4.2 (Lemma 6.19 in Schumaker (2007)) *For all $l \in \overline{[1,n]}, i \in \overline{[1, s+d]}, j \in \overline{[1,s]}$, $|\beta_{l,i,j}| \le (y_{l,i+s} - y_{l,i})^{j-1} \le (s\,\bar{\delta}_l)^{j-1}$.*

Theorem A.4.3 (Theorem 4.17 in Schumaker (2007)) *Let $s > 1$. Fix $l \in \overline{[1,n]}$ and $i_l \in \overline{[1, s+d]}$. Suppose $y_{l,i_l} < y_{l,i_l+s}$. Then $N^s_{l,i_l}(p_l) > 0$ when $p_l \in (y_{l,i_l}, y_{l,i_l+s})$, and $N^s_{l,i_l}(p_l) = 0$ when $p_l \notin [y_{l,i_l}, y_{l,i_l+s}]$. At the end points of (y_{l,i_l}, y_{l,i_l+s}),*

$$\begin{aligned}
(-1)^{k+s-\mu_{l,i_l}} D^k_+ N^s_{l,i_l}(y_{l,i_l}) &= 0 \qquad k = 0,1,\dots,s-1-\mu_{l,i_l} \\
(-1)^{s-\nu_{l,i_l+s}} D^k_- N^s_{l,i_l}(y_{l,i_l+s}) &= 0 \qquad k = 0,1,\dots,s-1-\nu_{l,i_l+s}
\end{aligned}$$

where $\mu_{l,i_l} = \max\{j : y_{l,i_l} =, \dots, = y_{l,i_l+j-1}\}$ and $\nu_{l,i_l+s} = \max\{j : y_{l,i_l+s} =, \dots, = y_{l,i_l+s-j+1}\}$.

Theorem A.4.4 (Theorem 4.20 in Schumaker (2007)) *Fix* $l \in \overline{[1,n]}$ *and* $i_l \in \overline{[s,s+d]}$. *For all* $p_l \in [y_{l,i_l}, y_{l,i_l+1})$, $\sum_{v_l=i_l+1-s}^{i_l} N^s_{l,v_l}(p_l) = 1$.

Theorem A.4.5 (Theorem 4.22 in Schumaker (2007)) *Fix* $l \in \overline{[1,n]}$. *Suppose that* k *and* p_l *are such that* $y_{l,k} \leq p_l < y_{l,k+1}$, *and define* $\delta_{l,i_l,k,j} = \min\{(y_{l,v+j} - y_{l,v}) : y_{l,i_l} \leq y_{l,v} \leq y_{l,k} < y_{l,k+1} \leq y_{l,v+j} \leq y_{l,i_l+s}\}$, *for* $j \in \overline{[1,s]}$. *Suppose* $\sigma > 0$ *and* $\delta_{l,i_l,k,s-\sigma+1} > 0$. *Then* $|D^\sigma_+ N^s_{l,i_l}(p_l)| \leq \Gamma_{s,\sigma}/(\prod_{q=1}^{\sigma} \delta_{l,i_l,k,s-q})$ *where* $\Gamma_{s,\sigma} = \frac{(s-1)!}{(s-\sigma-1)!}\begin{pmatrix} \sigma \\ \lfloor \sigma/2 \rfloor \end{pmatrix} \leq 2^\sigma \frac{(s-1)!}{(s-\sigma-1)!}$.

I now proceed to prove each part in Lemma A.4.2 one by one.

Proof of part (a)

Note that $\mathcal{L} = \mathcal{L}_1 \circ \mathcal{L}_2 \circ \cdots \circ \mathcal{L}_n$. For any $f \in \otimes_{l=1}^n \mathsf{P}^{s-1}[\underline{p}_l, \bar{p}_l]$, I can apply Theorem A.4.1 iteratively n times to obtain $\mathcal{L}f = \mathcal{L}_1 \circ \cdots \circ \mathcal{L}_n f = \mathcal{L}_1 \circ \cdots \circ \mathcal{L}_{n-1} f = \cdots = f$, where $\mathcal{L}_l f$ is understood as applying $\mathcal{L}_l$ to f which is viewed a single variate function of p_l.

Proof of part (b)

This follows directly from Theorem A.4.2.

Proof of part (c)

By my definition of $\{y_{l,i}\}_{l=1,i=1}^{n,2s+d}$, $\mu_{l,i_l} = \nu_{l,i_l+s} = 1$ for all $l \in \overline{[1,n]}$ and $i_l \in \overline{[1,s+d]}$. Hence, by Theorem A.4.3, for any $l \in \overline{[1,n]}$, $i_l \in \overline{[1,s+d]}$ and $r_l \in \overline{[0,s-2]}$, $N^s_{l,i_l}(.)$ is nonnegative and $D^{r_l} N^s_{l,i_l}(p_l) = 0$ for all $p_l \notin (y_{l,i_l}, y_{l,i_l+s})$. Hence $N_{i_1,\ldots,i_n}$ is nonnegative and, for any $r \in \overline{[0,s-2]}$ and any $r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}$ satisfying $\sum_{l=1}^n r_l = r$,

$$\frac{\partial^r N_{i_1,\ldots,i_n}(p)}{\partial p_1^{r_1} \ldots \partial p_n^{r_n}} = \prod_{l=1}^n D^{r_l} N^s_{l,i_l}(p_l) = 0$$

for all $p \notin \otimes_{l=1}^n (y_{l,i_l}, y_{l,i_l+s})$, where the second equality follows since $r_l \leq r \leq s-2$.

Proof of part (d)

Let $H_{i_1,\ldots,i_n} = \otimes_{l=1}^n [y_{l,i_l}, y_{l,i_l+1}]$ for any $i_l \in \overline{[s,s+d]}, l \in \overline{[1,n]}$. By Theorem A.4.4 and the fact that $\{N_{i_1,\ldots,i_n}(.)\}_{i_1=1,\ldots,i_n=1}^{s+d,\ldots,s+d}$ are all continuous functions (because $s \geq 2$), I have $\sum_{v_1=i_1+1-s}^{i_1} \cdots \sum_{v_n=i_n+1-s}^{i_n} N_{v_1,\ldots,v_n}(p) = 1$ for $p \in H_{i_1,\ldots,i_n}$. Moreover, by Lemma A.4.2 part (c), $N_{i_1,\ldots,i_n}(.)$ is nonnegative and $N_{i_1,\ldots,i_n}(p) = 0$ for $p = (p_1; \ldots; p_n) \notin \otimes_{l=1}^n (y_{l,i_l}, y_{l,i_l+s})$. Fix some $\tilde{i}_l \in \overline{[s,s+d]}$, $l \in \overline{[1,n]}$. For all $p \in H_{\tilde{i}_1,\ldots,\tilde{i}_n}$, $\sum_{i_1=1}^{s+d} \cdots \sum_{i_n=1}^{s+d} |N_{i_1,\ldots,i_n}(p)| = \sum_{i_1=1}^{s+d} \cdots \sum_{i_n=1}^{s+d} N_{i_1,\ldots,i_n}(p) = \sum_{v_1=\tilde{i}_1+1-s}^{\tilde{i}_1} \cdots \sum_{v_n=\tilde{i}_n+1-s}^{\tilde{i}_n} N_{v_1,\ldots,v_n}(p) = 1$. The result follows since the equality holds for all $\tilde{i}_l \in \overline{[s,s+d]}, l \in \overline{[1,n]}$, and $\mathcal{P} = \cup_{\tilde{i}_1=s,\ldots,\tilde{i}_n=s}^{s+d,\ldots,s+d} H_{\tilde{i}_1,\ldots,\tilde{i}_n}$.

Proof of part (e)

Fix $r \in \overline{[0, s-2]}$, and consider any $r_l \in \mathbb{Z}_+, l \in \overline{[1,n]}$ satisfying $\sum_{l=1}^{n} r_l = r$. Since $N^s_{l,i_l}(.) \in \mathsf{C}^{s-2}([\underline{p}_l, \bar{p}_l])$ and $r_l \leq r \leq s-2$, $D^{r_l}_{+} N^s_{l,i_l}(p_l) = D^{r_l}_{-} N^s_{l,i_l}(p_l) = D^{r_l} N^s_{l,i_l}(p_l)$ for all $p_l \in [\underline{p}_l, \bar{p}_l]$. Fix some $i_l \in \overline{[s, s+d]}$ for all $l \in \overline{[1,n]}$. Suppose that $p_l \in [y_{l,i_l}, y_{l,i_l+1})$. Then, if $r_l = 0$, $|D^{r_l} N^s_{l,i_l}(p_l)| = |N^s_{l,i_l}(p_l)| \leq 1 = 2^0 \frac{(s-1)!}{(s-0-1)!} \bar{\delta}^0_l$ where the inequality follows by Lemma A.4.2 part (d). Otherwise, $r_l \geq 1$, and $s - r_l \geq s - r \geq 2 > 1$. So $\delta_{l,i_l,k,s-q} \geq \bar{\delta}_l > 0$ for all $q = 1, \ldots, r_l$ (recall that $\delta_{l,i_l,k,j}$ is as defined in Theorem A.4.5). Then, by Theorem A.4.5, $|D^{r_l} N^s_{l,i_l}(p_l)| \leq 2^{r_l} \frac{(s-1)!}{(s-r_l-1)!} \bar{\delta}^{-r_l}_l$.

Now, for any $p = (p_1; \ldots; p_n) \in \mathcal{P}$, if $p_l \in (y_{l,i_l}, y_{l,i_l+s})$ for all $l \in \overline{[1,n]}$, the following holds,

$$\left| \frac{\partial^r N_{i_1,\ldots,i_n}(p)}{\partial p_1^{r_1} \ldots \partial p_n^{r_n}} \right| = \prod_{l=1}^{n} \left| D^{r_l} N^s_{l,i_l}(p_l) \right| \leq \prod_{l=1}^{n} 2^{r_l} \frac{(s-1)!}{(s-r_l-1)!} \bar{\delta}^{-r_l}_l \leq 2^r \left[\frac{(s-1)!}{(s-r-1)!} \right]^n \left[\min_{l=1,\ldots,n} \{\bar{\delta}_l\} \right]^{-r}.$$

Otherwise, there exists some l_0 such that $p_{l_0} \notin (y_{l_0,i_{l_0}}, y_{l_0,i_{l_0}+s})$. By Lemma A.4.2 part (c),

$$\left| \frac{\partial^r N_{i_1,\ldots,i_n}(p)}{\partial p_1^{r_1} \ldots \partial p_n^{r_n}} \right| = 0 \leq 2^r \left[\frac{(s-1)!}{(s-r-1)!} \right]^n \left[\min_{l=1,\ldots,n} \{\bar{\delta}_l\} \right]^{-r}.$$

So the result follows by letting $C_{r,s} = 2^r[(s-1)!/(s-r-1)!]^n$.

A.4.2 Proof of Proposition 4.4.1

I first show the feasibility of $\tilde{\mathbf{NP}}(\delta)$. Define $\tilde{h}(.) = \tilde{f}(g^{-1}(.)) : \mathcal{Y} \to \mathbb{R}^n$ and $h(.) = f(g^{-1}(.)) : \mathcal{Y} \to \mathbb{R}^n$. By condition (ii), $h(.)$ is strongly concave. Also, define $\tilde{\delta}(y) := \tilde{g}(g^{-1}(y)) - y$. Consider the following two optimization problems:

$$(\mathbf{NP}_y) \quad \max_{y \in \mathcal{Y}} \ \{h(y): \ Uy \preceq V\} \quad \text{and} \quad (\tilde{\mathbf{NP}}_y(\delta)) \quad \max_{y \in \mathcal{Y}} \ \left\{\tilde{h}(y): \ Uy + U\tilde{\delta}(y) \preceq V - \delta\right\}.$$

Note that $\mathbf{NP}_y$ is equivalent to $\mathbf{NP}$ and $\tilde{\mathbf{NP}}_y(\delta)$ is equivalent to $\tilde{\mathbf{NP}}(\delta)$. Thus, $y^* := g(x^*)$ is the optimal solution to $\mathbf{NP}_y$ and $Uy^* \preceq V$. Since $g^{-1}(.)$ is continuous by condition (i) and x^* is in the interior of $\mathcal{X}$ by condition (iv), y^* is in the interior of $\mathcal{Y}$ and there exists a constant $\bar{\phi} > 0$ such that $y^* - \bar{\phi}\mathbf{e} \subseteq \mathcal{Y}$. Let $\bar{\delta} = \min_i\{\bar{\phi}(U\mathbf{e})_i\}$ (note that since U does not have zero rows and all its components are non-negative, $\bar{\delta} > 0$). I claim that if $||Ug(.) - U\tilde{g}(.) + \delta||_\infty \leq \bar{\delta}$, then $y^* - \bar{\phi}\mathbf{e}$ is a feasible solution of $\tilde{\mathbf{NP}}_y(\delta)$. To see this, simply note that

$$\begin{aligned} ||U\tilde{\delta}(.) + \delta||_\infty &= \sup_{y \in \mathcal{Y}} ||U\tilde{g}(g^{-1}(y)) - Uy + \delta||_\infty = \sup_{y \in \mathcal{Y}} ||U\tilde{g}(g^{-1}(y)) - Ug(g^{-1}(y)) + \delta||_\infty \\ &= \sup_{x \in \mathcal{X}} ||U\tilde{g}(x) - Ug(x) + \delta||_\infty = ||Ug(.) - U\tilde{g}(.) + \delta||_\infty. \end{aligned}$$

So, $U(y^* - \bar{\phi}\mathbf{e}) + U\tilde{\delta}(y^* - \bar{\phi}\mathbf{e}) + \delta \preceq Uy^* - \bar{\phi}U\mathbf{e} + ||U\tilde{\delta}(.) + \delta||_\infty \mathbf{e} \preceq Uy^* + \bar{\delta}\mathbf{e} - \bar{\phi}U\mathbf{e} \preceq V$, where the last inequality follows by the definition of $\bar{\delta}$ and the fact that y^* is feasible to $\mathbf{NP}_y$. This proves that $\tilde{\mathbf{NP}}_y(\delta)$ is feasible. Thus, $\tilde{\mathbf{NP}}(\delta)$ is feasible. Since the feasible region of $\tilde{\mathbf{NP}}_y(\delta)$ is compact and $\tilde{h}(.)$ is continuous, $\tilde{\mathbf{NP}}_y(\delta)$ has an optimal solution. Let $\tilde{y}_\delta$ denote an optimal solution of $\tilde{\mathbf{NP}}_y(\delta)$ (note that $\tilde{y}_\delta$ may not be unique).

I now proceed to derive a bound of $||y^* - \tilde{y}_\delta||_2$, which will be used later to obtain the desired bound for $||x^* - \tilde{x}_\delta||_2$. To bound $||y^* - \tilde{y}_\delta||_2$, I will use the optimal solution of an auxiliary optimization problem below:

$$(\tilde{\mathbf{NP}}_y^{ax}(\delta)) \qquad \max_{y \in \mathcal{Y}} \left\{ h(y) : \ Uy + U\tilde{\delta}(y) \preceq V - \delta \right\}.$$

The above problem has the same feasible region as $\tilde{\mathbf{NP}}_y(\delta)$, so it is feasible. Let y_δ^{ax} denote an optimal solution of $\tilde{\mathbf{NP}}_y^{ax}(\delta)$. Since $||y^* - \tilde{y}_\delta||_2 \leq ||y^* - y_\delta^{ax}||_2 + ||y_\delta^{ax} - \tilde{y}_\delta||_2$, to bound $||y^* - \tilde{y}_\delta||_2$, I only need to bound $||y^* - y_\delta^{ax}||_2$ and $||y_\delta^{ax} - \tilde{y}_\delta||_2$. To derive an upper bound of $||y^* - y_\delta^{ax}||_2$, I need to use the following lemma (the proof is given later).

Lemma A.4.3 *Consider the family of perturbed optimization problems below:*

$$(\mathbf{NP}_y(\epsilon)) \qquad \max_{y \in \mathcal{Y}} \ \{ h(y) : \ Uy \preceq V + \epsilon \}.$$

Suppose that $h(.)$ is strongly concave and twice continuously differentiable, $\mathcal{Y}$ is a convex compact set, U is a non-negative matrix and has full row rank, and the optimal solution of $\mathbf{NP}_y(\mathbf{0})$ lies in the interior of $\mathcal{Y}$. If $y^*(\epsilon)$ is an optimal solution for $\mathbf{NP}_y(\epsilon)$, then $||y^*(\mathbf{0}) - y^*(\epsilon)||_2 \leq K||\epsilon||_\infty$ for some $K > 0$ independent of ϵ.

Note that the assumptions of Lemma A.4.3 hold (i.e., $h(.) = f(g^{-1}(.))$ is twice continuously differentiable because $f(.)$ and $g^{-1}(.)$ are both twice continuously differentiable. Also, as shown earlier, the optimal solution of $\mathbf{NP}_y(\mathbf{0})$, $y^*(\mathbf{0}) = y^*$, is in the interior of $\mathcal{Y}$). By the strong concavity of $h(.)$, y_δ^{ax} is the *unique* optimal solution of $\mathbf{NP}_y(-\delta - U\tilde{\delta}(y_\delta^{ax}))$. Thus, by Lemma A.4.3, there exists a constant $K_1 > 0$ independent of $\tilde{f}, \tilde{g}, \delta$ such that

$$||y^* - y_\delta^{ax}||_2 \ \leq \ K_1||U\tilde{\delta}(y_\delta^{ax}) + \delta||_\infty \ \leq \ K_1(||U||_\infty ||g(.) - \tilde{g}(.)||_\infty + ||\delta||_\infty). \tag{87}$$

I now derive a bound for $||\tilde{y}_\delta - y_\delta^{ax}||_2$. Since $\tilde{\mathbf{NP}}_y^{ax}(\delta)$ and $\tilde{\mathbf{NP}}_y(\delta)$ have the same constraints, $\tilde{y}_\delta$ is feasible for $\tilde{\mathbf{NP}}_y^{ax}(\delta)$ and y_δ^{ax} is feasible for $\tilde{\mathbf{NP}}_y(\delta)$. By the strong concavity of $h(.)$, there exists a constant $v > 0$ depending only on $h(.)$ such that

$$h(\tilde{y}_\delta) \;\leq\; h(y_\delta^{ax}) + \nabla h(y_\delta^{ax}) \cdot (\tilde{y}_\delta - y_\delta^{ax}) - \frac{v}{2}||\tilde{y}_\delta - y_\delta^{ax}||_2^2 \;\leq\; h(y_\delta^{ax}) - \frac{v}{2}||\tilde{y}_\delta - y_\delta^{ax}||_2^2, \quad (88)$$

where the last inequality follows because $\nabla h(y_\delta^{ax}) \cdot (\tilde{y}_\delta - y_\delta^{ax}) \leq 0$ (otherwise, y_δ^{ax} cannot be the optimal solution of $\tilde{\mathbf{NP}}_y^{ax}(\delta)$). Note also that $\tilde{h}(y_\delta^{ax}) \leq \tilde{h}(\tilde{y}_\delta)$. Combining this with (88), by Mean Value Theorem, I have

$$\begin{aligned} \frac{v}{2}||\tilde{y}_\delta - y_\delta^{ax}||_2^2 \;&\leq\; [h(y_\delta^{ax}) - \tilde{h}(y_\delta^{ax})] - [h(\tilde{y}_\delta) - \tilde{h}(\tilde{y}_\delta)] \;\leq\; (\nabla h(\xi) - \nabla \tilde{h}(\xi))'(y_\delta^{ax} - \tilde{y}_\delta) \\ &\leq\; ||(\nabla h(.) - \nabla \tilde{h}(.))'||_\infty ||\tilde{y}_\delta - y_\delta^{ax}||_\infty \;\leq\; ||(\nabla h(.) - \nabla \tilde{h}(.))'||_\infty ||\tilde{y}_\delta - y_\delta^{ax}||_2, \end{aligned}$$

for some $\xi \in \mathcal{Y}$. This means that $||\tilde{y}_\delta - y_\delta^{ax}||_2 \leq \frac{2}{v}||(\nabla h(.) - \nabla \tilde{h}(.))'||_\infty$. Combining this with (87),

$$\begin{aligned} ||y^* - \tilde{y}_\delta||_2 \;&\leq\; ||y^* - y_\delta^{ax}||_2 + ||y_\delta^{ax} - \tilde{y}_\delta||_2 \\ &\leq\; K_1||U||_\infty ||g(.) - \tilde{g}(.)||_\infty + K_1||\delta||_\infty + \frac{2}{v}||(\nabla h(.) - \nabla \tilde{h}(.))'||_\infty \\ &\leq\; K_1||U||_\infty ||g(.) - \tilde{g}(.)||_\infty + K_1||\delta||_\infty + \frac{2}{v}||(\nabla g^{-1}(.))'||_\infty ||(\nabla f(.) - \nabla \tilde{f}(.))'||_\infty, \end{aligned} \quad (89)$$

where the last inequality holds because $\nabla h(y) - \nabla \tilde{h}(y) = \nabla g^{-1}(y)[\nabla f(g^{-1}(y)) - \nabla \tilde{f}(g^{-1}(y))]$ for all $y \in \mathcal{Y}$. This means that the following inequality also holds:

$$\begin{aligned} &||x^* - \tilde{x}_\delta||_2 \\ &\leq \sqrt{n}||g^{-1}(y^*) - g^{-1}(\tilde{y}_\delta)||_\infty \leq \sqrt{n}||(\nabla g^{-1}(.))'||_\infty ||y^* - \tilde{y}_\delta||_\infty \leq \sqrt{n}||(\nabla g^{-1}(.))'||_\infty ||y^* - \tilde{y}_\delta||_2 \\ &\leq \sqrt{n}||(\nabla g^{-1}(.))'||_\infty \left(K_1||U||_\infty ||g(.) - \tilde{g}(.)||_\infty + K_1||\delta||_\infty + \frac{2}{v}||(\nabla g^{-1}(.))'||_\infty ||(\nabla f(.) - \nabla \tilde{f}(.))'||_\infty\right) \\ &\leq K(||(\nabla f(.) - \nabla \tilde{f}(.))'||_\infty + ||g(.) - \tilde{g}(.)||_\infty + ||\delta||_\infty), \end{aligned}$$

where $K = \sqrt{n}||(\nabla g^{-1}(.))'||_\infty (K_1||U||_\infty + K_1 + \frac{2}{v}||(\nabla g^{-1}(.))'||_\infty)$. This completes the proof.

Proof of Lemma A.4.3. I claim that there exists $\bar{\epsilon} := \min\{\bar{\epsilon}_1, \bar{\epsilon}_2\} > 0$, where $\bar{\epsilon}_1, \bar{\epsilon}_2$ are strictly positive constants to be defined later, such that, for all $||\epsilon||_\infty \leq \bar{\epsilon}$, $||y^*(\mathbf{0}) - y^*(\epsilon)||_2 \leq K_1||\epsilon||_\infty$ for some $K_1 > 0$ independent of ϵ. Note that, if this claim is true, Lemma A.4.3 can be proven as follows. Define $l := \sup_{y_1, y_2 \in \mathcal{Y}} ||y_1 - y_2||_2$ and let $K_2 = l/\bar{\epsilon}$. Then, for all ϵ with $||\epsilon||_\infty > \bar{\epsilon}$, $||y^*(\epsilon) - y^*(\mathbf{0})||_2 \leq l = K_2\bar{\epsilon} \leq K_2||\epsilon||_\infty$. So, Lemma A.4.3 follows by letting $K = \max\{K_1, K_2\}$. I now prove my claim.

I first introduce two optimization problems whose optimal solutions are closely related to $y^*(\mathbf{0})$ and $y^*(\epsilon)$. The first optimization problem is almost identical to $\mathbf{NP}_y(\mathbf{0})$ except that the domain is $\mathbb{R}^n$ instead of $\mathcal{Y}$:

$$(\bar{\mathbf{NP}}_y) \qquad \max_{y \in \mathbb{R}^n} \{h(y) : \; Uy \preceq V\}.$$

To define my second optimization problem, first note that, since U has full row rank, there exists an n by m matrix H such that $UH = I$. For any $\epsilon \in \mathbb{R}^m$, by a change of variables $y = z + H\epsilon$, I can transform $\mathbf{NP}_y(\epsilon)$ into an equivalent optimization problem below:

$$(\mathbf{NP}_z(\epsilon)) \qquad \max_{z \in \mathcal{Y} - H\epsilon} \{h_\epsilon(z) : \; Uz \preceq V\},$$

where $h_\epsilon(z) := h(z + H\epsilon)$. Two important observations are in order. The first observation relates $y^*(\mathbf{0})$ to the first optimization problem $\bar{\mathbf{NP}}_y$ whereas the second observation relates $y^*(\epsilon)$ to the second optimization problem $\mathbf{NP}_z(\epsilon)$.

Observation 1: $y^*(\mathbf{0})$ is the unique optimal solution to $\bar{\boldsymbol{NP}}_y$.

Suppose that this is not true. Then, there exists $\tilde{y} \neq y^*(\mathbf{0})$ satisfying $U\tilde{y} \preceq V$ such that $h(\tilde{y}) \geq h(y^*(\mathbf{0}))$. Let $d = (\tilde{y} - y^*(\mathbf{0}))/||\tilde{y} - y^*(\mathbf{0})||_2$. Since $y^*(\mathbf{0})$ is in the interior of $\mathcal{Y}$, $Uy^*(\mathbf{0}) \preceq V$ and $Uy^*(\mathbf{0}) + Ud||\tilde{y} - y^*(\mathbf{0})||_2 = U\tilde{y} \preceq V$; so, for sufficiently small $t > 0$, $\bar{y}_d(t) := y^*(\mathbf{0}) + td$ is a feasible solution for $\mathbf{NP}_y(\mathbf{0})$. Note that, by the strong concavity of $h(.)$, $h(y^*(\mathbf{0})) + \nabla h(y^*(\mathbf{0})) \cdot (\tilde{y} - y^*(\mathbf{0})) > h(\tilde{y}) \geq h(y^*(\mathbf{0}))$. This means that $\nabla h(y^*(\mathbf{0})) \cdot d > 0$. Hence, for sufficiently small $t > 0$, $h(\bar{y}_d(t)) = h(y^*(\mathbf{0})) + t\nabla h(y^*(\mathbf{0})) \cdot d + \mathcal{O}(t^2) > h(y^*(\mathbf{0}))$. However, by the strong concavity of $h(.)$, $y^*(\mathbf{0})$ is the *unique* optimal solution of $\mathbf{NP}_y(\mathbf{0})$. A contradiction is found. Hence, Observation 1 holds.

Observation 2: There exists $\bar{\epsilon}_1 > 0$ such that for all ϵ with $||\epsilon||_\infty \leq \bar{\epsilon}_1$, $\boldsymbol{NP}_y(\epsilon)$ has a unique optimal solution $y^*(\epsilon)$ and $z^*(\epsilon) := y^*(\epsilon) - H\epsilon$ is the unique optimal solution of $\boldsymbol{NP}_z(\epsilon)$.

I now prove Observation 2. Since $y^*(\mathbf{0})$ lies in the interior of $\mathcal{Y}$, there exists a constant $\bar{\phi} > 0$ such that $\{x : ||x - y^*(\mathbf{0})||_\infty \leq \bar{\phi}\} \subseteq \mathcal{Y}$. Let $\bar{\epsilon}_1 := \min_i\{\bar{\phi}(U\mathbf{e})_i\}$. Note that $\bar{\epsilon}_1 > 0$ since U is non-negative and has full row rank. Moreover, for all ϵ with $||\epsilon||_\infty \leq \bar{\epsilon} \leq \bar{\epsilon}_1$, $y^*(\mathbf{0}) - \bar{\phi}\mathbf{e} \in \mathcal{Y}$ is a feasible solution of $\mathbf{NP}_y(\epsilon)$ because $U(y^*(\mathbf{0}) - \bar{\phi}\mathbf{e}) \preceq V - \bar{\phi}U\mathbf{e} \preceq V - \bar{\epsilon}_1\mathbf{e} \preceq V + \epsilon$. This means that $\mathbf{NP}_y(\epsilon)$ has a unique optimal solution $y^*(\epsilon)$ (because its feasible region is convex, compact, not empty, and its objective function $h(.)$ is strongly concave). Hence, by definition of $\mathbf{NP}_z(\epsilon)$, $z^*(\epsilon)$ is its unique optimal solution. So, Observation 2 holds.

To bound $||y^*(\mathbf{0}) - y^*(\epsilon)||_2$, I first derive a bound for $||y^*(\mathbf{0}) - z^*(\epsilon)||_2$. Let $\bar{\epsilon}_2 := \bar{\phi}/||H||_\infty$. Then, for all ϵ with $||\epsilon||_\infty \leq \bar{\epsilon} \leq \bar{\epsilon}_2$, since $||y^*(\mathbf{0}) + H\epsilon - y^*(\mathbf{0})||_\infty = ||H\epsilon||_\infty \leq ||H||_\infty||\epsilon||_\infty \leq$

$\bar{\phi}$, $y^*(\mathbf{0}) + H\epsilon \in \mathcal{Y}$. This means that $y^*(\mathbf{0})$ is feasible for $\mathbf{NP}_z(\epsilon)$ and $h_\epsilon(y^*(\mathbf{0})) \leq h_\epsilon(z^*(\epsilon))$. Note that $z^*(\epsilon) \in \mathbb{R}^n$ is also feasible for $\bar{\mathbf{NP}}_y$, so

$$\begin{aligned} h(z^*(\epsilon)) &\leq h(y^*(\mathbf{0})) + \nabla h(y^*(\mathbf{0})) \cdot (z^*(\epsilon) - y^*(\mathbf{0})) - \frac{v}{2}||z^*(\epsilon) - y^*(\mathbf{0})||_2^2 \\ &\leq h(y^*(\mathbf{0})) - \frac{v}{2}||z^*(\epsilon) - y^*(\mathbf{0})||_2^2 \end{aligned} \quad (90)$$

for some $v > 0$ that only depends on $h(.)$. The first inequality follows by the strong concavity of $h(.)$ and the second inequality follows because $\nabla h(y^*(\mathbf{0})) \cdot (z^*(\epsilon) - y^*(\mathbf{0})) \leq 0$ (otherwise, $y^*(\mathbf{0})$ cannot be the optimal solution of $\bar{\mathbf{NP}}_y$). Note also that, for all ϵ with $||\epsilon||_\infty \leq \bar{\epsilon}$, the following holds

$$\begin{aligned} \frac{v}{2}||z^*(\epsilon) - y^*(\mathbf{0})||_2^2 &\leq h(y^*(\mathbf{0})) - h(z^*(\epsilon)) - h_\epsilon(y^*(\mathbf{0})) + h_\epsilon(z^*(\epsilon)) \\ &= (\nabla h(\xi_1) - \nabla h_\epsilon(\xi_1)) \cdot (y^*(\mathbf{0}) - z^*(\epsilon)) \\ &\leq ||\nabla h(\xi_1) - \nabla h_\epsilon(\xi_1)||_\infty ||y^*(\mathbf{0}) - z^*(\epsilon)||_\infty \\ &\leq ||\nabla h(\xi_1) - \nabla h(\xi_1 + H\epsilon)||_\infty ||y^*(\mathbf{0}) - z^*(\epsilon)||_2 \\ &= ||\nabla^2 h(\xi_2) H\epsilon||_\infty ||y^*(\mathbf{0}) - z^*(\epsilon)||_2 \\ &\leq K_0 ||H||_\infty ||\epsilon||_\infty ||y^*(\mathbf{0}) - z^*(\epsilon)||_2 \end{aligned} \quad (91)$$

for some $\xi_1 \in \mathcal{Z}, \xi_2 \in \bar{\mathcal{Z}}$ where $\mathcal{Z} := \{z : ||z - y||_\infty \leq \bar{\phi} \text{ for some } y \in \mathcal{Y}\}$ and $\bar{\mathcal{Z}} := \{z : ||z - x||_\infty \leq \bar{\phi} \text{ for some } x \in \mathcal{Z}\}$ are both compact and convex, and $K_0 := \sup_{z \in \bar{\mathcal{Z}}} ||\nabla^2 h(z)||_\infty$ only depends on $h(.)$. The first inequality of (91) follows by (90) and $h_\epsilon(y^*(\mathbf{0})) \leq h_\epsilon(z^*(\epsilon))$. The first equality of (91) follows by the Mean Value Theorem and the fact that $y^*(\mathbf{0}) \in \mathcal{Z}$ and $z^*(\epsilon) \in \mathcal{Z}$ (the latter inclusion holds because, since $z^*(\epsilon) \in \mathcal{Y} - H\epsilon$, there exists a $y \in \mathcal{Y}$ such that $||z^*(\epsilon) - y||_\infty = ||H\epsilon||_\infty \leq ||H||_\infty \bar{\epsilon} \leq ||H||_\infty \bar{\epsilon}_2 = \bar{\phi}$). Similarly, the second equality also follows by the Mean Value Theorem and the fact that $\xi_1 \in \mathcal{Z} \subseteq \bar{\mathcal{Z}}$ and $\xi_1 + H\epsilon \in \bar{\mathcal{Z}}$ (the latter inclusion holds since $\xi_1 \in \mathcal{Z}$ and $||\xi_1 + H\epsilon - \xi_1||_\infty \leq ||H||_\infty \bar{\epsilon} \leq ||H||_\infty \bar{\epsilon}_2 = \bar{\phi}$). Note that (91) is equivalent to $||z^*(\epsilon) - y^*(\mathbf{0})||_2 \leq 2v^{-1} K_0 ||H||_\infty ||\epsilon||_\infty$. Let $K_1 = 2v^{-1} K_0 ||H||_\infty + ||H||_2 \sqrt{n}$. Then, for all ϵ with $||\epsilon||_\infty \leq \bar{\epsilon}$, I can bound:

$$\begin{aligned} ||y^*(\epsilon) - y^*(\mathbf{0})||_2 &= ||z^*(\epsilon) + H\epsilon - y^*(\mathbf{0})||_2 \leq ||z^*(\epsilon) - y^*(\mathbf{0})||_2 + ||H||_2 ||\epsilon||_2 \\ &\leq 2v^{-1} K_0 ||H||_\infty ||\epsilon||_\infty + ||H||_2 ||\epsilon||_2 \leq K_1 ||\epsilon||_\infty. \end{aligned}$$

This proves the claim I stated at the beginning and completes the proof of Lemma A.4.3.

A.4.3 Proof of Lemma 4.4.2

I now prove Lemma 4.4.2 using Proposition 4.4.1. Let $g(.) = \lambda^*(.), \tilde{g}(.) = \tilde{\lambda}(.), f(.) = r^*(.), \tilde{f}(.) = \tilde{r}(.), U = A, V = C/T, \delta = \mathbf{0}, \mathcal{X} = \mathcal{P}, \mathcal{Y} = \Lambda_{\lambda^*}$. Note that $r^*(.), \lambda^*(.)$ are twice continuously differentiable by P1 and $\tilde{r}(.), \tilde{\lambda}(.)$ are continuously differentiable since $\tilde{\lambda}(.) \in \mathsf{C}^{s-2}(\mathcal{P})$ and $s-2 \geq 1$. Also $\mathcal{P}$ is convex and A is nonnegative with full row rank. I first verify the conditions (i) - (iv) in Proposition 4.4.1. By P1, $\lambda^*(.)$ has a twice continuously differentiable inverse function $p^*(.)$, and Λ_{λ^*} is assumed to be convex, so (i) holds. By R3, $r^*_\lambda(.) := r^*(p^*(.))$ is strongly concave, so (ii) holds. By R4, $\mathbf{P}$ is feasible and its optimal solution p^D lies in the interior of $\mathcal{P}$, so both (iii) and (iv) hold. For any $p \in \mathcal{P}$, I have

$$\begin{aligned} ||(\nabla r^*(p) - \nabla \tilde{r}(p))'||_\infty &= ||(\lambda^*(p) + \nabla \lambda^*(p)p - \tilde{\lambda}(p) - \nabla \tilde{\lambda}(p)p)'||_\infty \\ &\leq ||(\lambda^*(p) - \tilde{\lambda}(p))'||_\infty + ||p'(\nabla \lambda^*(p) - \nabla \tilde{\lambda}(p))'||_\infty \\ &\leq n||\lambda^*(p) - \tilde{\lambda}(p)||_\infty + ||p'||_\infty ||(\nabla \lambda^*(p) - \nabla \tilde{\lambda}(p))'||_\infty \\ &\leq n||\lambda^*(.) - \tilde{\lambda}(.)||_\infty + (\textstyle\sum_{l=1}^n \bar{p}_l)||(\nabla \lambda^*(.) - \nabla \tilde{\lambda}(.))'||_\infty. \end{aligned}$$

Therefore, by Proposition 4.4.1, there exists $\bar{\delta}_1 > 0$ and $K_1 > 0$ such that for all $\tilde{\lambda}(.)$ satisfying $||A\lambda(.) - A\tilde{\lambda}(.)||_\infty \leq \bar{\delta}_1$, $\tilde{\mathbf{P}}$ is feasible and

$$\begin{aligned} ||p^D - \tilde{p}^D||_2 &\leq K_1(||A\lambda^*(.) - A\tilde{\lambda}(.)||_\infty + ||(\nabla r^*(.) - \nabla \tilde{r}(.))'||_\infty) \\ &\leq K_1[(n + ||A||_\infty)||\lambda^*(.) - \tilde{\lambda}(.)||_\infty + (\textstyle\sum_{l=1}^n \bar{p}_l)||(\nabla \lambda^*(.) - \nabla \tilde{\lambda}(.))'||_\infty] \\ &\leq K(||\lambda^*(.) - \tilde{\lambda}(.)||_\infty + ||(\nabla \lambda^*(.) - \nabla \tilde{\lambda}(.))'||_\infty) \end{aligned}$$

where $K = K_1(n + ||A||_\infty + \sum_{l=1}^n \bar{p}_l)$ is independent of $\tilde{\lambda}(.)$. Let $\bar{\delta} := \bar{\delta}_1/||A||_\infty$. Since $||\lambda(.) - \tilde{\lambda}(.)||_\infty \leq \bar{\delta}$ means that $||A\lambda(.) - A\tilde{\lambda}(.)||_\infty \leq ||A||_\infty \bar{\delta} = \bar{\delta}_1$, the result follows.

A.4.4 Derivation of the equality (4.2)

Recall that $u^*_{ij} := [\frac{\partial^2 \lambda^*_1(p^D)}{\partial p_i \partial p_j}; \ldots; \frac{\partial^2 \lambda^*_n(p^D)}{\partial p_i \partial p_j}]$. Note that the following identity holds:

$$\begin{aligned} H^*_{ij} &= \left[B^* \nabla^2 r^*_\lambda(\lambda^D)(B^*)'\right]_{ij} - [B^* + (B^*)']_{ij} \\ &= \sum_{l=1}^n \sum_{k=1}^n \frac{\partial \lambda^*_k(p^D)}{\partial p_i} \frac{\partial^2 r^*_\lambda(\lambda^D)}{\partial \lambda_k \partial \lambda_l} \frac{\partial \lambda^*_l(p^D)}{\partial p_j} - \left[\frac{\partial \lambda^*_i(p^D)}{\partial p_j} + \frac{\partial \lambda^*_j(p^D)}{\partial p_i}\right]. \end{aligned} \quad (92)$$

Note also that $r^*(p) = p'\lambda^*(p) = r^*_\lambda(\lambda^*(p))$. Taking its second order derivative, I have

$$\frac{\partial^2 r^*(p^D)}{\partial p_i \partial p_j} = \sum_{l=1}^n \sum_{k=1}^n \frac{\partial \lambda^*_k(p^D)}{\partial p_i} \frac{\partial^2 r^*_\lambda(\lambda^D)}{\partial \lambda_k \partial \lambda_l} \frac{\partial \lambda^*_l(p^D)}{\partial p_j} + \sum_{l=1}^n \frac{\partial r^*_\lambda(\lambda^D)}{\partial \lambda_l} \frac{\partial^2 \lambda^*_l(p^D)}{\partial p_i \partial p_j} \quad (93)$$

$$\frac{\partial^2 r^*(p^D)}{\partial p_i \partial p_j} = \frac{\partial \lambda^*_i(p^D)}{\partial p_j} + \frac{\partial \lambda^*_j(p^D)}{\partial p_i} + \sum_{l=1}^n p^D_l \frac{\partial^2 \lambda^*_l(p^D)}{\partial p_i \partial p_j}. \quad (94)$$

Hence, combining (93) and (94) with (92), I have $H^*_{ij} = \sum_{l=1}^n (p^D_l - \frac{\partial r^*_\lambda(\lambda^D)}{\partial \lambda_l}) \frac{\partial^2 \lambda^*_l(p^D)}{\partial p_i \partial p_j} = (u^*_{ij})'(p^D - \nabla r^*_\lambda(\lambda^D))$. Note that $\lambda^*(p^D) + \nabla\lambda^*(p^D)p^D = \nabla r^*(p^D) = \nabla\lambda^*(p^D)\nabla r^*_\lambda(\lambda^D)$, so $p^D - \nabla r^*_\lambda(\lambda^D) = -\nabla\lambda^*(p^D)^{-1}\lambda^*(p^D) = -(B^*)^{-1}\lambda^D$. Hence, $H^*_{ij} = -(u^*_{ij})'(B^*)^{-1}\lambda^D$.

A.4.5 Proof of Lemma 4.4.3

I will prove each part of the lemma in turn. Let $\bar{\delta} = \min\{\bar{\delta}_1, \bar{\delta}_2, \bar{\delta}_3, \bar{\delta}_4, \bar{\delta}_5\}$, $\kappa = \max\{\kappa_1, \kappa_2\}$, where $\bar{\delta}_1, \ldots, \bar{\delta}_5$ and κ_1, κ_2 are strictly positive constants to be defined shortly.

Proof of Parts (a)-(c)

Let $\sigma_{\max}(X)$ and $\sigma_{\min}(X)$ denote the maximum and minimum eigenvalues of a symmetric real matrix X, respectively. Since $B^* = \nabla\lambda^*(p^D)$ is invertible, $B^*(B^*)'$ is positive definite; so, $\bar{\sigma}^* := \sigma_{\max}(B^*(B^*)') > 0$ and $\underline{\sigma}^* := \sigma_{\min}(B^*(B^*)') > 0$. Define $\bar{\delta}_1 = \underline{\sigma}^*/(4\sqrt{\bar{\sigma}^*}) > 0$. Note that, for all $\theta_\iota \in \mathsf{Ball}(\theta^*_\iota, \bar{\delta})$, $||B - B^*||_2 \le ||B - B^*||_F \le ||\theta_\iota - \theta^*_\iota||_2 \le \bar{\delta} \le \bar{\delta}_1$. Therefore, for all $v \in \mathbb{R}^n$ such that $||v||_2 = 1$, I have:

$$\begin{aligned} v'B'Bv &= v'(B - B^* + B^*)'(B - B^* + B^*)v \\ &= v'(B^*)'B^*v + v'(B^*)'(B - B^*)v + v'(B - B^*)'B^*v + v'(B - B^*)'(B - B^*)v \\ &\ge \underline{\sigma}^* - 2||v||_2^2||(B^*)'||_2||B - B^*||_2 \ \ge\ \underline{\sigma}^* - 2\sqrt{\bar{\sigma}^*}\,\underline{\sigma}^*/(4\sqrt{\bar{\sigma}^*}) \ =\ \underline{\sigma}^*/2. \end{aligned}$$

This means that $\sigma_{\min}(B'B) \ge \underline{\sigma}^*/2 > 0$ and B is invertible. Since $(B'B)^{-1} = B^{-1}(B^{-1})'$,

$$||(B')^{-1}||_2 = \sqrt{\sigma_{\max}(B^{-1}(B^{-1})')} = \sqrt{\sigma_{\min}(B'B)^{-1}} \le \sqrt{2/\underline{\sigma}^*}.$$

By definition, $\lambda(p; \theta_\iota) = a + B'p$, $\lambda(p; \theta^*_\iota) = a^* + (B^*)'p$, $\frac{\partial \lambda_j}{\partial p_i}(p; \theta_\iota) = B_{ij}$, $\frac{\partial \lambda_j}{\partial p_i}(p; \theta^*_\iota) = B^*_{ij}$ for all $i, j \in \overline{[1, n]}$, and $p(\lambda; \theta_\iota) = -(B')^{-1}a + (B')^{-1}\lambda$ (recall that B is invertible). So, for all

$p \in \mathcal{P}$ and $\lambda, \lambda' \in \lambda(\mathcal{P}; \theta_\iota)$,

$$\begin{aligned}
||\lambda(p;\theta_\iota) - \lambda(p;\theta_\iota^*)||_2 &\leq ||a - a^*||_2 + ||(B - B^*)'||_2 ||p||_2 \\
&\leq ||a - a^*||_2 + ||(B - B^*)'||_F \left(\sum_{l=1}^n \bar{p}_l^2 \right)^{\frac{1}{2}} \\
&\leq \left[1 + \left(\sum_{l=1}^n \bar{p}_l^2 \right)^{\frac{1}{2}} \right] ||\theta_\iota - \theta_\iota^*||_2, \\
\left| \frac{\partial \lambda_j}{\partial p_i}(p;\theta_\iota) - \frac{\partial \lambda_j}{\partial p_i}(p;\theta_\iota^*) \right| &= |B_{ij} - B_{ij}^*| \leq ||\theta_\iota - \theta_\iota^*||_2, \text{ and}
\end{aligned}$$

$$||p(\lambda;\theta_\iota) - p(\lambda';\theta_\iota)||_2 = ||(B')^{-1}(\lambda - \lambda')||_2 \leq ||(B')^{-1}||_2 ||\lambda - \lambda'||_2 \leq \sqrt{2/\underline{\sigma}^*} ||\lambda - \lambda'||_2.$$

The results follow by letting $\omega = \max\{1 + (\sum_{l=1}^n \bar{p}_l^2)^{\frac{1}{2}}, \sqrt{2/\underline{\sigma}^*}\}$.

Proof of part (d)

To prove that $q(p(.;\theta_\iota);\theta_o)$ is strongly concave, I need to show that its Hessian, $B^{-1}G(B')^{-1}$, is negative definite. Let $\bar{\delta}_2 = ||G^*||_2$. For all $\theta_o \in \mathsf{Ball}(\theta_o^*, \bar{\delta})$, $||\theta_o - \theta_o^*||_2 \leq \bar{\delta} \leq \bar{\delta}_2$, so

$$||B^* - B||_2 \leq ||B^* - B||_F \leq \bar{\delta} \tag{95}$$

$$||G^* - G||_2 \leq ||G^* - G||_F \leq \bar{\delta} \tag{96}$$

$$||G||_2 \leq ||G^*||_2 + ||G - G^*||_2 \leq ||G^*||_2 + \bar{\delta}_2 = 2||G^*||_2. \tag{97}$$

Recall that, by Lemma 4.4.3 part (a), B is invertible and $||B^{-1}||_2 = ||(B')^{-1}||_2 \leq \omega$. So,

$$\begin{aligned}
&||(B^*)^{-1}G^*((B^*)')^{-1} - B^{-1}G(B')^{-1}||_2 \\
&= ||((B^*)^{-1} - B^{-1})G^*((B^*)')^{-1} + B^{-1}(G^* - G)((B^*)')^{-1} + B^{-1}G(((B^*)')^{-1} - (B')^{-1})||_2 \\
&\leq ||(B^*)^{-1} - B^{-1}||_2 ||G^*||_2 ||((B^*)')^{-1}||_2 + ||B^{-1}||_2 ||G^* - G||_2 ||((B^*)')^{-1}||_2 \\
&\qquad + ||B^{-1}||_2 ||G||_2 ||((B^*)')^{-1} - (B')^{-1}||_2 \\
&\leq ||(B^*)^{-1}||_2 ||B^* - B||_2 ||B^{-1}||_2 ||G^*||_2 ||((B^*)')^{-1}||_2 + ||B^{-1}||_2 ||G^* - G||_2 ||((B^*)')^{-1}||_2 \\
&\qquad + ||B^{-1}||_2 ||G||_2 ||(B^*)^{-1}||_2 ||B^* - B||_2 ||B^{-1}||_2 \\
&\leq ||B^{-1}||_2 ||(B^*)^{-1}||_2 (||(B^*)^{-1}||_2 ||G^*||_2 + 1 + 2||G^*||_2 ||B^{-1}||_2)\bar{\delta} \\
&\leq \omega ||(B^*)^{-1}||_2 (||(B^*)^{-1}||_2 ||G^*||_2 + 1 + 2||G^*||_2 \omega)\bar{\delta} \\
&\leq C\bar{\delta},
\end{aligned}$$

for some $C > 0$ that only depends on θ^* and ω. The second inequality above holds because

$(B^*)^{-1} - B^{-1} = (B^*)^{-1}(B - B^*)B^{-1}$ and the third inequality follows from (95)-(97).

By (4.3), $(B^*)^{-1}G^*((B^*)')^{-1} = \nabla^2 r^*_\lambda(\lambda^D)$. So, $\sigma_{\max}((B^*)^{-1}G^*((B^*)')^{-1}) \leq -\underline{v}$ by R3. Let $\bar{\delta}_3 = \underline{v}/(2C)$. Then, for all v such that $||v||_2 = 1$,

$$\begin{aligned} v'(B^{-1}G(B')^{-1})v &= v'((B^*)^{-1}G^*((B^*)')^{-1})v + v'(B^{-1}G(B')^{-1} - (B^*)^{-1}G^*((B^*)')^{-1})v \\ &\leq \sigma_{\max}((B^*)^{-1}G^*((B^*)')^{-1}) + ||B^{-1}G(B')^{-1} - (B^*)^{-1}G^*((B^*)')^{-1}||_2 \\ &\leq -\underline{v} + C\bar{\delta} \ \leq \ -\underline{v} + C\bar{\delta}_3 \ = \ -\underline{v}/2. \end{aligned}$$

This means that $B^{-1}G(B')^{-1}$ is negative definite and, thus, $q(p(.;\theta_\iota);\theta_o)$ is strongly concave.

Proof of part (e)

Consider the following optimization problem:

$$(\mathbf{QP}_\lambda(\theta;\delta)) \qquad \max_{\lambda \in \lambda(\mathcal{P};\theta_\iota)} \left\{ q(p(\lambda;\theta_\iota);\theta_o) : \quad A\lambda \preceq \frac{C}{T} - \delta \right\}.$$

This problem is equivalent to $\mathbf{QP}(\theta;\delta)$, except that it optimizes over λ instead of p. Recall that, by Lemma 4.4.3 part (d), $q(p(.;\theta_\iota);\theta_o)$ is strongly concave; so, $\mathbf{QP}_\lambda(\theta;\delta)$ is a convex program and, if it is feasible, it has a unique optimal solution that is characterized by the KKT condition. Since $\mathbf{P}_\lambda$ is a convex program and has a unique optimal solution $\lambda^D \in \Lambda_{\lambda^*}$, by the KKT conditions, I have:

$$\begin{cases} \nabla r^*_\lambda(\lambda^D) = A'\mu^D \\ (A\lambda^D - \frac{C}{T})'\mu^D = 0 \\ \mu^D \succeq \mathbf{0}, A\lambda^D \preceq \frac{C}{T} \end{cases} \quad \text{which, by (4.3), implies} \quad \begin{cases} \nabla_\lambda q(p(\lambda^D;\theta^*_\iota);\theta^*_o) = A'\mu^D \\ (A\lambda^D - \frac{C}{T})'\mu^D = 0 \\ \mu^D \succeq \mathbf{0}, A\lambda^D \preceq \frac{C}{T} \end{cases}$$

Since $\lambda^D = \lambda(p^D;\theta^*_\iota) \in \lambda(\mathcal{P};\theta^*_\iota)$ is feasible to $\mathbf{QP}_\lambda(\theta^*;\mathbf{0})$, by the sufficiency of KKT conditions for optimality in a strongly convex program, λ^D and μ^D are also the *unique* optimal primal and dual solution of $\mathbf{QP}_\lambda(\theta^*;\mathbf{0})$. Hence, $p^D = p(\lambda^D;\theta^*_\iota)$ is also the unique optimal solution of $\mathbf{QP}(\theta^*;\mathbf{0})$. This proves that $\lambda^D = \lambda^D_{\mathbf{0}}(\theta^*), \mu^D = \mu^D_{\mathbf{0}}(\theta^*)$ and $p^D = p^D_{\mathbf{0}}(\theta^*)$.

Proof of part (f)

The proof relies on Proposition 4.4.1. Let $g(.) = \lambda(.;\theta^*_\iota), \tilde{g}(.) = \lambda(.;\theta_\iota), f(.) = q(.;\theta^*_o), \tilde{f}(.) = q(.;\theta_o), U = A, V = C/T, \mathcal{X} = \mathcal{P}, \mathcal{Y} = \lambda(\mathcal{P};\theta^*_\iota)$. I first verify conditions (i)-(iv) of Proposition 4.4.1. Since $\lambda(.;\theta^*_\iota)$ is linear and $B^* = \nabla\lambda^*(p^D)$ is invertible, it has an inverse function $p(.;\theta^*_\iota)$ that is linear and, hence, twice continuously differentiable. Moreover, the set $\lambda(\mathcal{P};\theta^*_\iota)$ is convex because $\mathcal{P}$ is convex and convexity is preserved under affine transformation. So, (i) holds. By (4.3) and the fact that $r^*_\lambda(.)$ is strongly concave, $q(p(.;\theta^*_\iota);\theta^*_o)$ is strongly concave, so (ii) holds. As shown earlier, p^D is the optimal solution of $\mathbf{QP}(\theta^*;\mathbf{0})$ and it is in the interior of $\mathcal{P}$, so (iii) and (iv) hold. Therefore, by Proposition 4.4.1, there exists some constant $\tilde{\delta} > 0$ such that if $||A\lambda(.;\theta_\iota^*) - A\lambda(.;\theta_\iota) + \delta||_\infty \leq \tilde{\delta}$, then $\mathbf{QP}(\theta;\delta)$ is feasible and there exists some constant K independent of θ such that the unique optimal solution of $\mathbf{QP}(\theta;\delta)$ (i.e., $p_\delta^D(\theta)$) satisfies the following:

$$
\begin{aligned}
&||p_{\mathbf{0}}^D(\theta^*) - p_\delta^D(\theta)||_2 \\
&\quad \leq K\left(||(\nabla q(.;\theta_o^*) - \nabla q(.;\theta_o))'||_\infty + ||\lambda(.;\theta_\iota^*) - \lambda(.;\theta_\iota)||_\infty + ||\delta||_\infty\right) \\
&\quad = K\left(\sup_{p\in\mathcal{P}}\{||(F^* + p'G^* - F - p'G)'||_\infty\} + \sup_{p\in\mathcal{P}}\{||\lambda(p;\theta_\iota^*) - \lambda(p;\theta_\iota)||_\infty\} + ||\delta||_\infty\right) \\
&\quad \leq K\left(\sup_{p\in\mathcal{P}}\{||(F^* + p'G^* - F - p'G)'||_2\} + \sup_{p\in\mathcal{P}}\{||\lambda(p;\theta_\iota^*) - \lambda(p;\theta_\iota)||_2\} + ||\delta||_\infty\right) \\
&\quad \leq K\left(||F^* - F||_2 + \sup_{p\in\mathcal{P}}\{||p'||_2\}||G^* - G||_2 + \omega||\theta_\iota^* - \theta_\iota||_2 + ||\delta||_2\right) \\
&\quad \leq K\left(||F^* - F||_2 + (\textstyle\sum_{l=1}^n \bar{p}_l^2)^{1/2}||G^* - G||_F + \omega||\theta_\iota^* - \theta_\iota||_2 + ||\delta||_2\right) \\
&\quad \leq K\left(1 + (\textstyle\sum_{l=1}^n \bar{p}_l^2)^{1/2}\right)||\theta_o^* - \theta_o||_2 + K\omega||\theta_\iota^* - \theta_\iota||_2 + K||\delta||_2 \\
&\quad \leq \kappa_1\left(||\theta^* - \theta||_2 + ||\delta||_2\right) \\
&\quad \leq \kappa(||\theta^* - \theta||_2 + ||\delta||_2),
\end{aligned}
$$

where $\kappa_1 = K(2 + (\sum_{l=1}^n \bar{p}_l^2)^{1/2} + \omega)$. Let $\bar{\delta}_4 = \min\{\tilde{\delta}/2, \tilde{\delta}/(2\omega||A||_2), \phi/(6\kappa)\}$. Then, for $\theta_\iota \in \mathsf{Ball}(\theta_\iota^*, \bar{\delta})$ and $\delta \in \mathsf{Ball}(\mathbf{0}, \bar{\delta})$, I have

$$
\begin{aligned}
||A\lambda(.;\theta_\iota^*) - A\lambda(.;\theta_\iota) + \delta||_\infty &\leq \sup_{p\in\mathcal{P}} ||A\lambda(p;\theta_\iota^*) - A\lambda(p;\theta_\iota)||_\infty + ||\delta||_\infty \\
&\leq \sup_{p\in\mathcal{P}} ||A\lambda(p;\theta_\iota^*) - A\lambda(p;\theta_\iota)||_2 + ||\delta||_2 \ \leq \ \omega||A||_2||\theta_\iota^* - \theta_\iota||_2 + ||\delta||_2 \\
&\leq \omega||A||_2\bar{\delta} + \bar{\delta} \ \leq \ \omega||A||_2\bar{\delta}_4 + \bar{\delta}_4 \ \leq \ \omega||A||_2\frac{\tilde{\delta}}{2\omega||A||_2} + \frac{\tilde{\delta}}{2} = \tilde{\delta}.
\end{aligned}
$$

Moreover, $p_\delta^D(\theta) \in \mathsf{Ball}(p_{\mathbf{0}}^D(\theta^*), \phi/2)$ because

$$
\begin{aligned}
||p_{\mathbf{0}}^D(\theta^*) - p_\delta^D(\theta)||_2 &\leq \kappa(||\theta^* - \theta||_2 + ||\delta||_2) \ \leq \ \kappa(||\theta_\iota^* - \theta_\iota||_2 + ||\theta_o^* - \theta_o||_2 + ||\delta||_2) \\
&\leq 3\kappa\bar{\delta} \ \leq \ 3\kappa\bar{\delta}_4 \ \leq \ 2\kappa\phi/(6\kappa) = \ \phi/3 < \ \phi/2.
\end{aligned}
$$

Since $\mathsf{Ball}(p_{\mathbf{0}}^D(\theta^*), \phi) \subseteq \mathcal{P}$ (by R4 and $p_{\mathbf{0}}^D(\theta^*) = p^D$), $\mathsf{Ball}(p_\delta^D(\theta), \phi/2) \subseteq \mathcal{P}$. Let $\kappa_2 =$

$2\kappa_1||B^*||_2 + \omega$. I have:

$$\begin{aligned}
||\lambda^D_{\mathbf{0}}(\theta^*) - \lambda^D_\delta(\theta)||_2 &= ||\lambda(p^D_{\mathbf{0}}(\theta^*);\theta^*_\iota) - \lambda(p^D_\delta(\theta);\theta_\iota)||_2 \\
&\leq ||\lambda(p^D_{\mathbf{0}}(\theta^*);\theta^*_\iota) - \lambda(p^D_\delta(\theta);\theta^*_\iota)||_2 + ||\lambda(p^D_\delta(\theta);\theta^*_\iota) - \lambda(p^D_\delta(\theta);\theta_\iota)||_2 \\
&\leq ||B^*||_2||p^D_{\mathbf{0}}(\theta^*) - p^D_\delta(\theta)||_2 + \omega||\theta^* - \theta||_2 \\
&\leq (\kappa_1||B^*||_2 + \omega)||\theta^* - \theta||_2 + \kappa_1||B^*||_2||\delta||_2 \\
&= \kappa_2(||\theta^* - \theta||_2 + ||\delta||_2) \\
&\leq \kappa(||\theta^* - \theta||_2 + ||\delta||_2).
\end{aligned}$$

I will now show that the constraints of $\mathbf{QP}(\theta;\delta)$ that correspond to rows $\{i : \mu^D_{\mathbf{0},i}(\theta^*) > 0\}$ are binding. Since $\mathbf{QP}_\lambda(\theta;\delta)$ is a feasible convex program, by KKT condition, $\nabla_\lambda q(p(\lambda^D_\delta(\theta);\theta_\iota);\theta_o) = A'\mu^D_\delta(\theta)$. By my assumption, A has full row rank. So, there exists some m by n matrix $\bar{A}$ such that $\mu^D_\delta(\theta) = \bar{A}\nabla_\lambda q(p(\lambda^D_\delta(\theta);\theta_\iota);\theta_o)$. Since the right hand side is jointly continuous in $(\theta;\delta)$ at $(\theta^*;\mathbf{0})$, $\mu^D_\delta(\theta)$ must also be continuous in $(\theta;\delta)$ at $(\theta^*;\mathbf{0})$. Let $\underline{\mu} := \min_{1\leq i\leq n}\{\mu^D_{\mathbf{0},i}(\theta^*) : \mu^D_{\mathbf{0},i}(\theta^*) > 0\}$. By continuity, there exists $\bar{\delta}_5 > 0$ such that $||\mu^D_\delta(\theta) - \mu^D_{\mathbf{0}}(\theta^*)||_2 < \underline{\mu}$ for all $\theta = (\theta_o;\theta_\iota)$ and δ satisfying $||\theta_\iota - \theta^*_\iota||_2 \leq \bar{\delta}_5$, $||\theta_o - \theta^*_o||_2 \leq \bar{\delta}_5$ and $||\delta||_2 \leq \bar{\delta}_5$. Since, by definition, $\bar{\delta} \leq \bar{\delta}_5$, for all $\theta = (\theta_o;\theta_\iota)$ and δ satisfying $||\theta_\iota - \theta^*_\iota||_2 \leq \bar{\delta}$, $||\theta_o - \theta^*_o||_2 \leq \bar{\delta}$ and $||\delta||_2 \leq \bar{\delta}$, I have $\mu^D_{\delta,i}(\theta) > 0$ whenever $\mu^D_{\mathbf{0},i}(\theta^*) > 0$; so, the corresponding constraints in $\mathbf{QP}_\lambda(\theta;\delta)$ are binding due to the KKT condition that $(A\lambda^D_\delta(\theta) - \frac{C}{T})'\mu^D_\delta(\theta) = 0$.

Bibliography

Atar, R., M. Reiman. 2012. Aymptotically optimal dynamic pricing for network revenue management. *Stochastic Systems* **2** 232 – 276.

Auer, P., N. Cesa-Bianchi, P. Fischer. 2002a. Finite-time analysis of the multiarmed bandit problem. *Machine Learning* **47** 235–256.

Auer, P., N. Cesa-Bianchi, Y. Freund, R. Schapire. 2002b. The nonstochastic multiarmed bandit problem. *SIAM J. COMPUT.* **32** 48–77.

Badanidiyuru, A., R. Kleinberg, A. Slivkins. 2013. Bandits with knapsacks. *FOCS '13 Proceedings of the 2013 IEEE 54th Annual Symposium on Foundations of Computer Science*.

Badanidiyuru, A., R. Kleinberg, A. Slivkins. 2015. Bandits with knapsacks. *Working paper* .

Ben-Tal, A., L. E. Ghaoui, A. Nemirovski. 2009. *Robust Optimization*. Princeton University Press.

Bertsimas, D., D. A. Iancu, P. A. Parrilo. 2010. Optimality of affine policies in multistage robust optimization. *Math. Oper. Res.* **35** 363–394.

Besbes, O., Y. Gur, A. Zeevi. 2015. Non-stationary stochastic optimization. *Oper. Res.* **63**(5) 1227–1244.

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

Besbes, O., A. Zeevi. 2012. Blind network revenue management. *Oper. Res.* **60** 1537–1550.

Besbes, O., A. Zeevi. 2015. On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Sci.* **61** 723–739.

Bitran, G., R. Caldentey. 2003. An overview of pricing models for revenue management. *Manufacturing Service Oper. Management* **5** 203–229.

Bonnans, J., A. Shapiro. 2000. *Perturbation Analysis of Optimization Problems*. Springer.

Borovkov, A. A. 1999. *Mathematical Statistics*. CRC Press.

Boyd, S., L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press.

Broder, J., P. Rusmevichientong. 2012. Dynamic pricing under a general parametric choice model. *Oper. Res.* **60** 965–980.

Calafiore, G. C. 2009. An affine control method for optimal dynamic asset allocation with transaction costs. *SIAM J. Control Optim.* **48** 2254–2274.

Chen, L., T. Homem de Mello. 2010. Re-solving stochastic programming models for airline revenue management. *Annals of Operations Research* **177** 91–114.

Chen, Q., S. Jasin, I. Duenyas. 2016. Realtime pricing with minimal and flexible price adjustment. *Management Sci.* **62** 2437–2455.

Chen, Y., V. Farias. 2013. Simple policies for dynamic pricing with imperfect forecasts. *Oper. Res.* **61** 612–624.

Ciocan, D. F., V. Farias. 2012. Model predictive control for dynamic resource allocation. *Math. Oper. Res.* **37** 501–525.

Combes, R., C. Jiang, R. Srikant. 2015. Bandits with budgets: Regret lower bounds and optimal algorithms. *Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*.

den Boer, A. 2014. Dynamic pricing with multiple products and partially specified demand distribution. *Math. Oper. Res.* **39** 863–888.

den Boer, A., B. Zwart. 2014. Simultaneously learning and optimizing using controlled variance pricing. *Management Sci.* **60** 770–783.

Easley, D., N.M. Kiefer. 1988. Controlling a stochastic process with unknown parameters. *Econometrica: Journal of the Econometric Society* 1045–1064.

Elmaghraby, W., P. Keskinocak. 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Sci.* **49** 1287–1309.

Erdelyi, A., H. Topaloglu. 2011. Using decomposition methods to solve pricing problems in network revenue management. *Journal of Pricing and Revenue Management* **10** 325–343.

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

Ferreira, K. J., D. Simchi-Levi, H. Wang. 2016. Online network revenue management using thompson sampling. *Under review* .

Flajolet, A., P. Jaillet. 2015. Logarithmic regret bounds for bandits with knapsacks. *Working paper* .

Gallego, G., G. van Ryzin. 1994. Optimal dynamic pricing of inventory with stochastic demand over finite horizons. *Management Sci.* **40** 999–1020.

Gallego, G., G. van Ryzin. 1997. A multiproduct dynamic pricing problem and its applications to network yield management. *Oper. Res.* **45** 24–41.

Golrezaei, N., H. Nazerzadeh, P. Rusmevichientong. 2014. Real-time optimization of personalized assortments. *Management Sci.* **60** 1532–1551.

Gyorfi, L., M. Kohler, A Krzyzak, H Walk. 2002. *A Distribution-Free Theory of Nonparametric Regression*. Springer.

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

Haws, K. L., W. O. Bearden. 2006. Dynamic pricing and consumer fairness perceptions. *Journal of Consumer Research* **33** 304 – 311.

Hazan, E., et al. 2016. Introduction to online convex optimization. *Foundations and Trends in Optimization* **2**(3-4) 157–325.

Irvine, D. 2014. How airlines make less than $ 6 per passenger. posted in www.cnn.com on tuesday, june 3, 2014. .

Jasin, S. 2014. Re-optimization and self-adjusting price control for network revenue management. *Oper. Res.* **62** 1168–1178.

Jasin, S., S. Kumar. 2012. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Math. Oper. Res.* **37** 313–345.

Jasin, S., S. Kumar. 2013. Analysis of deterministic lp-based heuristics for revenue management. *Oper. Res.* **61** 1312–1320.

Keskin, N.B., A. Zeevi. 2014. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Oper. Res.* **62** 1142–1167.

Keskin, N.B., A. Zeevi. 2016. Chasing demand: Learning and earning in a changing environment. *Math. Oper. Res. forthcoming* .

Koushik, D., J. A. Higbie, C. Eister. 2012. Retail price optimization at intercontinental hotels group. *Interface* **42** 45 – 57.

Kunnumkal, S., H. Topaloglu. 2010. A new dynamic programming decomposition method for the network revenue management problem with customer choice behavior. *Production and Oper. Management* **19** 575–590.

Lai, T., H. Robbins. 1985. Asymptotic efficient adaptive allocation rules. *Advances in Applied Mathematics* **6** 4–22.

Lei, Y., S. Jasin, A. Sinha. 2014. Near-optimal bisection search for nonparametric dynamic pricing with inventory constraint. *Under review* .

Maglaras, C., J. Meissner. 2006. Dynamic pricing strategies for multiproduct revenue management problems. *Manufacturing Service Oper. Management* **8** 136–148.

Marn, M. V., E. V. Roegner, C. C. Zawada. 2003. The power of pricing. *The McKinsey Quarterly* **1** 26–39.

McLennan, A. 1984. Price dispersion and incomplete learning in the long run. *Journal of Economic dynamics and control* **7**(3) 331–347.

Pekgun, P., R. P. Menich, S. Acharya, P. G. Finch, F. Deschamps, K. Mallery, J. van Sistine, K. Christianson, J. Fuller. 2013. Carlson rezidor hotel group maximizes revenue through improved demand management and price optimization. *Interface* **43** 21 – 36.

Pzman, A. 2013. *Design of experiments in nonlinear models : asymptotic normality, optimality criteria and small-sample properties*. Springer.

Ramasastry, A. 2005. Web sites change prices based on customers' habits. posted in www.cnn.com on friday, june 24, 2005. .

Reiman, M. I., Q. Wang. 2008. An asymptotically optimal policy for a quantity-based network revenue management problem. *Math. Oper. Res.* **33** 257–282.

Rigby, D., K. Miller, J. Chernoff, S. Tager. 2012. Are consumers waiting for better deals? *Bain Retail Holiday Newsletter* **4** 1–20.

Rothschild, M. 1974. A two-armed bandit theory of market pricing. *Journal of Economic Theory* **9**(2) 185–202.

Sahay, A. 2007. How to reap higher profits with dynamic pricing. *MIT Sloan Management Review* **48** 53–60.

Schumaker, L. 2007. *Spline Functions: Basic Theory (Third Edition)*. Cambridge University Press.

Secomandi, N. 2008. An analysis of the control-algorithm re-solving issue in inventory and revenue management. *Manufacturing Service Oper. Management* **10** 468–483.

Sen, A. 2013. A comparison of fixed and dynamic pricing policies in revenue management. *Omega* **41** 586–597.

Talluri, K., G. van Ryzin. 2005. *The theory and Practice of Revenue Management*. Springer.

Wang, Z., S. Deng, Y. Ye. 2014. Closing the gaps: A learning-while-doing algorithm for single-product revenue management problems. *Oper. Res.* **62** 318–331.