Online Sequential Optimization with Biased Gradients: Theory and Applications to Censored Demand

Woonghee Tim Huh* Paat Rusmevichientong†
University of British Columbia University of Southern California

May 23, 2010
Revised: December 30, 2011

Abstract

We study a class of stochastic optimization problems where though the objective functions may not be convex, they satisfy a generalization of convexity, called the sequentially convex property. We focus on a setting where the distribution of the underlying uncertainty is unknown and the manager must make a decision in real-time based on historical data. Since sequentially convex functions are not necessarily convex, they pose difficulties in applying standard adaptive methods for convex optimization. We propose a nonparametric algorithm based on a gradient descent method, and show that the $T$-season average expected cost differs from the minimum cost by at most $O\left(\frac{1}{\sqrt{T}}\right)$. Our analysis is based on a careful quantification of the bias that is inherent in gradient estimation due to the adaptive nature of the problem. We demonstrate the usefulness of the concept of sequential convexity by applying it to three canonical problems in inventory control, capacity allocation, and the lifetime buy decision, under the assumption that the manager does not know the demand distributions and has access only to historical sales (censored demand) data.

1 Introduction

Many classical problems in operations research involve sequential optimization under uncertainty, where the decision maker makes a decision sequentially in each period to optimize a certain objective function. A traditional approach to these problems is to model the uncertainty through a probability distribution, and to maximize the expected profit by solving a dynamic programming equation.

In this paper, we are interested in sequential decision-making problems with the following three

* Sauder School of Business, University of British Columbia, Vancouver BC Canada V5T 1Z2. E-mail: tim.huh@sauder.ubc.ca.
† Marshall School of Business, University of Southern California, Los Angeles CA 90089, USA. E-mail: rusmevic@marshall.usc.edu
characteristics: i) the distribution of the underlying uncertainty is unknown; ii) the manager must make the decision in real-time based on historical data; and iii) the decision the manager makes in each period may influence the quality of observed data in that period, which in turn affects future decisions. Examples of problems with such characteristics include periodic-review inventory control with cyclic demand and lost sales, allocation of fixed capacity to multiple demand classes, and lifetime buy decisions towards the end of a lifecycle for a product with multiple obsolete parts. When the distribution of the underlying demand uncertainty is known in advance, these problems have been extensively studied in the literature and can be solved efficiently using dynamic programming, and we can characterize their optimal policies. However, when the distribution information is unavailable, relatively little is known regarding how to manage such a system.

While each of these problems has been studied in the domain of inventory management, revenue management or lifecycle management, there has not been any systematic effort to investigate the commonality of these problems. In this paper, we introduce the notion of sequentially convex stochastic optimization, a framework that encompasses the above three problems. We consider a setting where the manager makes a decision repeatedly over time in an adaptive manner. The main characteristics of our formulation are the two distinct types of temporal dependence: (i) The decision in each period is constrained and influenced by past decisions and past realizations of the underlying uncertainty. For example, the inventory level in a period is bounded below by the carried-over inventory. (ii) The decision in each period also affects future rewards. For instance, the amount of capacity reserved in the current period affects the total revenue that we can obtain in the future. In this sense, the model that we study here is inherently dynamic because the information is gathered over time and the decision in each period influences the rewards in the future.

The methodological approach that we develop here extends the current research on online convex optimization by incorporating non-convexity and biased gradients, allowing for sequentially convex objective functions, and enabling us to address a much larger class of applications. To our knowledge, this is the first application of online convex optimization for adaptively minimizing non-convex functions.

In addition, our framework leads to new results for adaptive methods in each domain area. For the inventory control problem, when the demands in each period do not have identical distribution, our algorithm represents the first adaptive method with the provable convergence rate. For the capacity allocation problem, the convergence rate of our algorithm is faster than the best
known algorithm. Furthermore, we are the first to consider the adaptive decision making for the lifetime problem. We will elaborate on our contributions and provide relevant literature review in later sections.

The paper is organized as follows. In Section 2, we extend the classical result in online convex optimization to allow for biased gradient estimators. In Section 3, we introduce the notion of sequentially convex functions, and develop a general theory for minimizing such functions. We apply the theory to problems in inventory control with seasonal demands, capacity allocation and lifetime buy decision in Sections 4, 5 and 6, respectively. We present computational results in Section 7, and conclude in Section 8.

2 Preliminaries: Stochastic Online Convex Optimization with Biased Gradients

In online convex optimization, the optimizer does not know any information about the objective function except its convexity. At each iteration, the optimizer chooses a feasible solution and obtains some information about the objective function at that feasible solution. If this information is the exact gradient of the function, Zinkevich (2003) presented an adaptive algorithm whose average regret over \( T \) periods – the difference between the average cost over \( T \) periods and the minimum cost – diminishes at the rate of \( O\left(1/\sqrt{T}\right)\). Flaxman et al. (2005) extended this work to stochastic online convex optimization, where the optimizer has access only to an unbiased estimate of the gradient in each iteration, and showed the average expected regret of \( O\left(1/\sqrt{T}\right)\). Later, Hazan et al. (2007) improved the convergence rate to \( O\left(\log T/T\right)\) by imposing additional requirements on the objective function.

All of these results require the availability of an unbiased estimate of the gradient, which is not the case in our setting. Thus, we extend the results of stochastic online convex optimization to allow for biased gradient estimates. We show that even with biased gradients, the descent algorithm still converges to the optimal solution at the same asymptotic rate, provided that the bias is sufficiently “small”. To our knowledge, this is the first attempt in the online convex optimization literature to address biases in the gradient estimator. The analysis that we provide shows the expected regret of \( O\left(1/\sqrt{T}\right)\), but it can be improved to \( O\left(\log T/T\right)\) by imposing similar technical conditions as in Hazan et al. (2007). The main result of this section is stated in Theorem 2.1, which will be used in subsequent analyses. The proof of the theorem appears in Appendix A. While the original context of online convex optimization is in an adversarial setting, we position our result in a stochastic setting so that its connection to later applications becomes more evident. For any \( z \in \mathbb{R}^n \), let \( \|z\| \)
denote the Euclidean norm of \( z \), and for any set \( S \subseteq \mathbb{R}^n \), let \( P_S(\cdot) \) denote the projection operator on \( S \) with respect to \( \|\cdot\| \) and let \( \text{diam}(S) \) denote the diameter of \( S \).

**Theorem 2.1 (Optimization with Biased Gradients).** Let \( \Phi : S \to \mathbb{R} \) be a convex function on a compact convex set \( S \subseteq \mathbb{R}^n \). For any \( z \in S \), let \( G(z) \) be an \( n \)-dimensional random vector and let \( \text{Bias}(z) = \mathbb{E}[G(z)] - \nabla \Phi(z) \). Suppose that there exists \( B > 0 \) such that \( \|G(z)\| \leq B \) holds with probability one for all \( z \in S \). Let a sequence \( (Z^t : t \geq 1) \) of random vectors be defined by

\[
Z^{t+1} = P_S(Z^t - \epsilon^t \cdot G(Z^t)), \quad \text{where} \quad \epsilon^t = \frac{\text{diam}(S)}{B} \cdot \frac{1}{\sqrt{t}},
\]

and \( Z^1 \) is any point in \( S \). Then, for all \( T \geq 1 \),

\[
\frac{1}{T} \sum_{t=1}^{T} \mathbb{E} \left[ \Phi(Z^t) - \min_{z \in S} \Phi(z) \right] \leq \frac{2B \text{diam}(S)}{\sqrt{T}} + \frac{\text{diam}(S)}{T} \sum_{t=1}^{T} \mathbb{E} \| \text{Bias}(Z^t) \|
\]

Note that \( \text{Bias}(Z^t) \) is a vector representing the expected difference between the true gradient \( \nabla \Phi(Z^t) \) and our estimate \( G(Z^t) \). Theorem 2.1 shows that, as \( T \) becomes large, the average expected regret will converge to zero as long as the average expected bias over time also converges to zero.

### 3 A General Theory for Sequentially Convex Functions

We consider the problem of minimizing a function \( f : \mathcal{K} \to \mathbb{R} \), where \( \mathcal{K} = \mathcal{K}_n \times \mathcal{K}_{n-1} \times \cdots \times \mathcal{K}_1 \), and \( \mathcal{K}_i \) is a compact interval in \( \mathbb{R} \) for each \( i \). Let \( \overline{\mathcal{K}} = \max_{i=1,...,n} \{ \text{diam}(\mathcal{K}_i) \} \). For each \( z = (z_n, \ldots, z_1) \in \mathcal{K} \), let \( z_{[i,k]} = (z_i, z_{i-1}, \ldots, z_k) \) denote a vector of length \( i - k + 1 \) if \( n \geq i \geq k \geq 1 \); otherwise, \( z_{[i,k]} \) is an empty vector. We assume that the function \( f \) is continuous in \( \mathcal{K} \) and attains its minimum value at \( z^* = (z^*_n, \ldots, z^*_1) \in \mathcal{K} \). We now introduce the class of functions that will be considered in this paper.

**Definition 3.1 (Sequential Convexity).** The function \( f \) is **sequentially convex** if there exists a sequence of functions \( \{f_i : \mathcal{K}_i \times \mathcal{K}_{i-1} \times \cdots \times \mathcal{K}_1 \to \mathbb{R} \mid i = 1, \ldots, n \} \) such that

(a) For all \( z \in \mathcal{K} \), \( f(z) = \sum_{i=1}^{n} f_i(z_i | z_{[i-1,1]}) \),

(b) For all \( i \), \( f_i(\cdot | z_{[i-1,1]}^*) \) is a convex function that attains its minimum value at \( z^*_i \), and

(c) There exists \( M \geq 1 \) such that for all \( z = (z_n, \ldots, z_1) \in \mathcal{K} \) and \( n \geq i \geq k \geq 1 \),

\[
\bigg| f_i \left( z_i \mid z_{[i-1,k+1]}, z_k, z_{[k-1,1]}^* \right) - f_i \left( z_i \mid z_{[i-1,k+1]}, z_k^*, z_{[k-1,1]}^* \right) \bigg| \\
\leq M \left\{ f_k \left( z_k \mid z_{[k-1,1]}^* \right) - f_k \left( z_k^* \mid z_{[k-1,1]}^* \right) \right\},
\]
\[
\left| f'_i \left( z_i \mid z_{[i-1,k+1]}, z_k, z^*_{[k-1,1]} \right) - f'_i \left( z_i \mid z_{[i-1,k+1]}, z_k^*, z^*_{[k-1,1]} \right) \right| \\
\leq M \left\{ f_k \left( z_k \mid z^*_{[k-1,1]} \right) - f_k \left( z_k^* \mid z^*_{[k-1,1]} \right) \right\},
\]

where \( f'_i \left( z_i \mid z^*_{[i-1,1]} \right) \) denotes a subgradient of \( f_i \left( \cdot \mid z^*_{[i-1,1]} \right) \) at \( z_i \).

We refer to the sequence of functions \( \{f_i\} \) as a **sequential decomposition** of \( f \).

We note that a sequentially convex function is **not** necessarily convex. However, it exhibits a property that its optimal solution can be obtained by sequentially optimizing one component at a time, and that each single-dimensional problem is convex given that all the previous decisions are made optimally. Condition (c) in the definition of sequential convexity means that when we change the value of \( z_k^* \) to \( z_k \), then for any \( i \geq k \), the change in \( f_i \left( z_i \mid z_{[n,k]}, z^*_{[k-1,1]} \right) \) and its subgradient \( f'_i \left( z_i \mid z_{[n,k]}, z^*_{[k-1,1]} \right) \) is uniformly bounded by the change in \( f_k \left( z_k \mid z^*_{[k-1,1]} \right) \).

Since \( f \) is sequentially convex, to find its minimum value, it is desirable to obtain an unbiased estimator of \( f'_i \left( z_i \mid z^*_{[i-1,1]} \right) \). However, such information requires the knowledge of \( z^*_{[i-1,1]} \), which is not feasible in our adaptive setting. As shown in the following theorem, it turns out that for sequentially convex functions, it suffices to have just an estimate of the subgradient \( f'_i \left( z_i \mid z^*_{[i-1,1]} \right) \) at \( z^* \) in each iteration. Our proposed algorithm – which we refer to as the **Adaptive Sequentially Convex Function Minimization** algorithm – generates a sequence \( \{Z^t \in K : t \geq 1\} \) of vectors where each \( Z^{t+1} \in K \) depends only on \( Z^t \) and an estimate \( G^t (Z^t) \) of the subgradient at \( Z^t \). The algorithm is given in Figure 1, and its convergence result is stated in Theorem 3.1.
Adaptive Sequentially Convex Function Minimization

**Input:** A sequentially convex function $f$ with a decomposition $\{f_i\}$.

**Initialization:** Let $Z^1$ be any vector in $K$.

**Description:** For any $t = 1, 2, \ldots$

- Let $G^t(Z^t) = (G^t_n(Z^t), \ldots, G^t_1(Z^t))$ denote an estimate of the gradient such that for all $i$ and $z$, $E[G^t_i(Z^t) \mid Z^t = z] = f'_i(z^t_i | z^t_{i-1,1})$ and $\|E[G^t_i(Z^t) \mid Z^t = z]\| \leq B$ for some $B > 0$.
- Update the vector $Z^{t+1} = (Z^{t+1}_n, \ldots, Z^{t+1}_1)$ as follows: for all $i$,

$$Z^{t+1}_i = P_{K_i} (Z^t_i - \epsilon^t_i G^t_i(Z^t)),$$

where $P_{K_i}$ denotes a projection operator onto the set $K_i$ and the $\epsilon^t = \overline{K}/(B \sqrt{T})$ represents the step size in period $t$.

**Output:** A sequence of vectors $\{Z^t : t \geq 1\}$.

**Theorem 3.1** (Adaptive Algorithm for Sequentially Convex Functions). For any sequentially convex function $f$ with a sequential decomposition $\{f_i\}$, let $\{Z^t : t \geq 1\}$ denote a sequence of vectors generated by the Adaptive Sequentially Convex Function Minimization algorithm. Then, for any $T \geq 1$,

$$0 \leq \frac{1}{T} \sum_{t=1}^{T} \mathbb{E} [f(Z^t) - f(z^*)] \leq \frac{2nMB\overline{K}(1 + \overline{K}M)^{n-1}}{\sqrt{T}}.$$ 

Before we proceed to the proof, we emphasize that the application of Theorem 3.1 requires only a sequentially convex objective function and an estimate of the gradient at $Z^t$. As shown in Sections 4, 5, and 6, this requirement is satisfied for a broad class of applications. Moreover, verifying the definition of sequential convexity for each application is straightforward, and in all of our applications, it is easy to obtain a gradient estimate based on sales data. Once these two conditions are satisfied, we can immediately apply the above theorem.

**Dependence on Problem Parameters:** In all of our applications, the constant $M$ corresponds to an upper bound on the density function associated with the demand random variable in each period, and thus, it is independent of the dimension $n$. However, the diameter $\overline{K}$ of the domain and the upper bound $B$ on the gradient estimate usually scale linearly with $n$. Thus, the constant appearing in the denominator of the above error bound can be $O(n^{n+2})$. Nonetheless, as a function of $T$, the number of iterations, we note that the running average expected cost converges to the
minimum cost at the rate of $O(T^{-1/2})$. In contrast, the best known adaptive algorithm for the capacity rationing problem in Section 5 has a convergence rate of $O(T^{-\beta/2^n})$, which deteriorates with $n$. As shown in Section 7, for the capacity allocation problem, our algorithm will converge more quickly as the dimension $n$ increases.

Here is the proof of Theorem 3.1.

**Proof.** It follows from the definition of sequential convexity and a telescoping sum that for any $z$,

$$f(z) - f(z^*) = \sum_{k=1}^{n} f_k(z_k | z_{[k-1,1]}) - f_k(z_k^* | z_{[k-1,1]}^*)$$

$$= \sum_{k=1}^{n} \sum_{i=1}^{k} f_i(z_i | z_{[i-1,1]}^*, z_{i,1}^*, z_{i-1,1}^*) - f_k(z_k | z_{[k-1,i+1]}^*, z_{k,i}^*, z_{[i-1,1]}^*)$$

$$\leq M \sum_{k=1}^{n} \sum_{i=1}^{k} f_i(z_i | z_{[i-1,1]}^*) - f_i(z_i^* | z_{[i-1,1]}^*)$$

$$\leq nM \sum_{i=1}^{n} f_i(z_i | z_{[i-1,1]}^*) - f_i(z_i^* | z_{[i-1,1]}^*)$$

which implies that

$$\frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[f(Z_t^i) - f(z^*)] \leq nM \sum_{i=1}^{n} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[f_i(Z_t^i | z_{[i-1,1]}^*) - f_i(z_i^* | z_{[i-1,1]}^*)] = nM \sum_{i=1}^{n} \Delta_i^T,$$

where for any $i$, we define $\Delta_i^T = \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[f_i(Z_t^i | z_{[i-1,1]}^*) - f_i(z_i^* | z_{[i-1,1]}^*)]$. Since the function $f_i(\cdot | z_{[i-1,1]}^*)$ is convex, we apply Theorem 2.1 with $S = \mathcal{K}_i \subset \mathbb{R}$ to obtain

$$\Delta_i^T \leq \frac{2B\mathcal{K}}{\sqrt{T}} + \frac{\mathcal{K}}{T} \sum_{t=1}^{T} \mathbb{E}[\text{Bias}(Z_t^i)] \leq \frac{2B\mathcal{K}}{\sqrt{T}} + \frac{\mathcal{K}}{T} \sum_{t=1}^{T} \mathbb{E}[f_i'(Z_t^i | z_{[i-1,1]}^*)] - f_i'(z_i^* | z_{[i-1,1]}^*)].$$

By using a telescoping sum and property of sequentially convex functions, we note that for any $z$,

$$f_i'(z_i | z_{[i-1,1]+1}) - f_i'(z_i | z_{[i-1,1]}^*) = \sum_{k=1}^{i-1} f_i'(z_i | z_{[i-1,k+1]}^*, z_k^*, z_{[i-1,k]}^*) - f_i'(z_i | z_{[i-1,k+1]}^*, z_k^*, z_{[i-1,k]}^*)$$

$$\leq \sum_{k=1}^{i-1} f_i'(z_i | z_{[i-1,k+1]}^*, z_k^*, z_{[i-1,k]}^*) - f_i'(z_i | z_{[i-1,k+1]}^*, z_k^*, z_{[i-1,k]}^*)$$

$$\leq M \sum_{k=1}^{i-1} f_k(z_k | z_{[k-1,1]}^*) - f_k(z_k^* | z_{[k-1,1]}^*)$$

$$\leq \frac{2B\mathcal{K}}{\sqrt{T}} + \mathcal{K} \sum_{t=1}^{T} \mathbb{E}\left[\sum_{k=1}^{i-1} f_k(z_k | z_{[k-1,1]}^*) - f_k(z_k^* | z_{[k-1,1]}^*)\right].$$

This implies that

$$\Delta_i^T \leq \frac{2B\mathcal{K}}{\sqrt{T}} + \mathcal{K} \sum_{t=1}^{T} \mathbb{E}\left[\sum_{k=1}^{i-1} f_k(z_k | z_{[k-1,1]}^*) - f_k(z_k^* | z_{[k-1,1]}^*)\right].$$
Since \(i\) is arbitrary, we get the following recursion: for any \(i\), 
\[
\Delta T_i \leq \left( \frac{2B\bar{K}}{\sqrt{T}} \right) + \bar{K}M \sum_{k=1}^{i-1} \Delta T_k.
\]
By expanding the recursion, it is easy to verify 
\[
\sum_{i=1}^{n} \Delta T_i \leq \frac{2B\bar{K}(1 + \bar{K}M)^{n-1}}{\sqrt{T}},
\]
which gives the desired result.

4 Application I: Inventory Control with Seasonal and Censored Demand

Most papers on inventory theory employ either deterministic or stochastic demand models, and the underlying assumption is that the inventory manager knows either the exact demands in the future or the demand distribution. However, it is quite common in practice that the manager does not know the demand distribution, but can only infer it from historical data. In the case where unsatisfied demand is lost, the information available to the manager may be further limited because he only observes the sales data (which is the minimum of the realized demand and the stocking quantity), in which case we say the demand observation is censored. In this case, the manager’s stocking decision in each period affects both the cost in that period and the quality of information that he can use for his future decisions.

When the information on the demand distribution is not available, the most common method is the use of Bayesian methods, where the distribution family is known while the exact parameters are not known; for example, Scarf (1959, 1960), Karlin (1960), Iglehart (1964), Chen and Plambeck (2008), and Liyanage and Shanthikumar (2005). When information about the distribution family to which the demand belongs to is not available, there are several non-parametric approaches: robust min-max approach (Scarf, 1958; Jagannathan, 1977; Gallego and Moon, 1993; Perakis and Roels, 2008), stochastic approximation (Burnetas and Smith, 2000; Kunnumkal and Topaloglu, 2008), successive piece-wise linear approximation (Godfrey and Powell, 2001; Powell et al., 2004), multi-armed bandit (Chang et al., 2005), and the shadow dynamic programming approach (Levi et al., 2007). Huh and Rusmevichientong (2009) use a stochastic online convex optimization method of Flaxman et al. (2005), and Huh et al. (2011) apply Kaplan-Meier estimator.

While there has been much interest in adaptively optimizing inventory decisions, most of the papers mentioned above address the setting where the demands over time are independent and identically distributed; the only exceptions, to our knowledge, are Levi et al. (2007) who consider uncensored demand setting and Kunnumkal and Topaloglu (2008) who show convergence without
providing a convergence rate. The inventory problem with independent and identically distributed distributions, also known as the repeated newsvendor problem, is a much simpler problem – it can be solved by finding an appropriate fractile of demand distribution without the sequential impact of decisions. In this paper, we consider a multi-period problem where the optimal decision in each period cannot be made myopically. In particular, we consider seasonal demand where the demand distributions are not identical over time, but repeated in a cyclic manner. For this problem, if the manager has full access to the demand distributions, it is well known that base-stock policies are optimal (Veinott, 1965), and the optimal policy can be computed by solving a dynamic programming equation (Iglehart and Karlin, 1962; Song and Zipkin, 1993). In the setting where the manager has no prior information regarding demand distribution setting, we develop an adaptive policy, whose $T$-season average expected regret is $O\left(1/\sqrt{T}\right)$. This result represents the first convergence rate for this class of problems.

4.1 Model and Main Result

We assume instantaneous replenishment, and consider a non-stationary inventory system where demand distributions follow a cyclic pattern of length $n$, where $n$ is the number of periods in each selling season. We index each season forward by $t = 1, 2, \ldots$, and index each period within a season backward by $i = n, n-1, \ldots, 1$. We use $(t, i)$ to represent period $i$ of season $t$. Let $D^t_i$ denote the nonnegative random variable representing demand in period $(t, i)$. The manager does not know the demand distributions, but knows that demands are independent and cyclic; that is, for each $i$, the demands in period $i$ of each season $D^1_i, D^2_i, \ldots$ are independent and identically distributed random variables. We let $D_i$ denote the generic demand random variable having the same distribution as $D^t_i$.

We also assume that the demand in each period of each selling season \{\(D^t_i \mid t \geq 1, i = n, \ldots, 1\}\} are independent random variables. We assume that for any $n \geq i \geq 1$, the random variable $D_i$ has a continuous density function that is bounded above by $M \geq 1$. We also assume that $D_i$ has a bounded support; thus, there exists an upper bound $K$ such that $\sum_{i=1}^{n} D_i \leq K$ with probability one.

The following sequence of events take place in period $(t, i)$: (1) At the beginning of each period $(t, i)$, the manager observes the initial inventory level $x^t_i$. We assume $x^t_i = 0$ if $i = n$. (2) The manager orders sufficient inventory to raise it up to $y^t_i$ units, where $y^t_i \geq x^t_i$. We assume instantaneous replenishment. (3) Then, demand $D^t_i$ is realized. We denote the realized demand by $d^t_i$. The manager, however, does not observe the realized demand $d^t_i$, but only observes the sales quantity $\min\{d^t_i, y^t_i\}$. (4) An appropriate overage and underage cost is realized depending on the
quantity of excess inventory or excess demand at the rate of $h$ or $b$ per unit.\footnote{All of the results in this section easily extend to the case of period-dependent costs.} Its expected cost, as a function of the inventory level after ordering, is given by $Q_i(y_i)$ where

$$Q_i(y) = h \cdot E[y - D_i^+] + b \cdot E[D_i - y]^+.$$  

(5) Excess demand is lost and excess inventory is carried over to the next period, except in the last period of each season. Thus, the beginning inventory level of the next period satisfies $x_{t-1}^i = [y_t^i - d_t^i]^+$ for $n \geq i \geq 2$, except that the first period of each season has no carry-over inventory, that is, $x_n^i = 0$ for each season $t$.

When the distributions of the demands $D_n, D_{n-1}, \ldots, D_1$ are known in advance, it is a well-known result that order-up-to policies are optimal (see, for example, Zipkin (2000)). An order-up-to policy is characterized by a sequence of order-up-to levels $z = (z_n, \ldots, z_1)$, where for any $i$, $z_i$ denotes the order-up-to level in period $i$. Under the policy $z = (z_n, \ldots, z_1)$, the after-ordering inventory level in each period $(t, i)$ is set to $y_t^i = \max\{x_t^i, z_i\}$, where $x_t^i$ denotes the current on-hand inventory level. Since the demand in each of $n$ periods is bounded above by $K$ with probability one, it suffices to restrict our attention to order-up-to policies in $K = \prod_{i=1}^n [0, K]$.

For any $z \in K$, let $f(z)$ denote the total expected overage and underage cost in a season under the order-up-to policy $z$, and let $z^* = (z_n^*, \ldots, z_1^*)$ denote the optimal order-up-to levels, that is, $f(z^*) = \min_{z \in K} f(z)$. The main result of this section is stated in the following theorem, which shows that the function $f$ is sequentially convex with a decomposition $\{f_i\}$, and describes how we can compute the gradient $f_i'(z_i | z_{i-1})$ based on censored demand observations. Thus, we can apply the Adaptive Sequentially Convex Function Minimization algorithm from Section 3.

Before we proceed to the statement of the theorem, let us introduce the following notation.

For any $n \geq i \geq 1$ and $z \in K$, let $g_i(y | z_{[i-1,1]})$ denote the total expected cost from period $i$ to period 1 provided that the after-ordering inventory level in period $i$ is $y$, and that we will use order-up-to levels $z_{i-1}, \ldots, z_1$ in the remaining $i-1$ periods. Then, we have the following dynamic programming recursion:

$$g_i(y | z_{[i-1,1]}) = Q_i(y) + E_{D_i} \left[ g_{i-1} \left( \max\{y - D_i, z_i\} | z_{[i-2,1]} \right) \right],$$

where $g_0(\cdot) = 0$. Thus, we have $f(z) = g_n(z_n | z_{[n-1,1]})$.

**Theorem 4.1** (Sequential Convexity in Inventory Control). In the seasonal demand inventory control model, the cost function $f$ is sequentially convex with a sequential decomposition $\{f_i\}$ defined
by: for any $z \in \mathcal{K}$ and each $i$,

$$f_i(z_i \mid z_{[i-1,1]}) = g_i(z_i \mid z_{[i-1,1]}) - g_{i-1}(z_{i-1} \mid z_{[i-2,1]})\,.$$  

Moreover, we can compute an unbiased estimate of the gradient estimate of $f'_i(z_i \mid z_{[i-1,1]})$ using the sales information, that is,

$$f'_i(z_i \mid z_{[i-1,1]}) = E\left[G_i(Z) \mid Z = z\right],$$

where

$$G_i(z) = h \cdot (i - \Lambda_i(z) + 1) - (h + b) \cdot 1\left\{z_i = \sum_{\ell = \Lambda_i(z)}^{i} S_{\ell}\right\},$$

and for any $\ell$, $S_{\ell} = \min\{D_{\ell}, z_{\ell}\}$ denotes the sales (censored demand) in period $\ell$, and

$$\Lambda_i(z) = \min \left\{ k \mid i \geq k \geq 1 \text{ and } z_i - \sum_{\ell = j+1}^{i} S_{\ell} > z_j \text{ for all } i \geq j \geq k \right\}.$$

Note that $\Lambda_i(z)+1$ denotes the first time after period $i$ when we place an order because, by definition, the starting inventory exceeds the base-stock level in each of periods $\{i - 1, i - 2, \ldots, \Lambda_i(z)\}$.

### 4.2 Proof of Theorem 4.1

The proof of Theorem 4.1 consists of verifying that the function $f$ and $\{f_i\}$ satisfies the definition of sequential convexity, and showing that we can estimate the gradient of $f'_i(z_i \mid z_{[i-1,1]})$ from sales data. Note that by our construction it follows that $f$ is a continuous function that achieves a minimum at $z^*$. Moreover, it is a well-known result that $g_i(\cdot \mid z^*_{[i-1,1]})$ is a convex function whose minimum occurs at $z^*_i$, which implies that $f_i(\cdot \mid z^*_{[i-1,1]})$ is also convex with a minimum at $z^*_i$.

The following lemma shows that $\{f_i\}$ has the desired property. The proof of this result appears in Appendix B.1.

**Lemma 4.1** (Bounds on the Difference). For any $z = (z_n, \ldots, z_1) \in \mathcal{K}$ and $n \geq i > k \geq 1$,

$$\left|f_i(z_i \mid z_{[i-1,k+1]}, z_k, z^*_{[k-1,1]}) - f_i(z_i \mid z_{[i-1,k+1]}, z_k, z^*_{[k-1,1]})\right| \leq f_k(z_k \mid z^*_{[k-1,1]}) - f_k(z_k \mid z^*_{[k-1,1]}).$$

The next lemma gives an explicit expression for the derivative of $f'_i(z_i \mid z_{[i-1,1]})$ for any $z \in \mathcal{K}$, and shows that it satisfies the definition of sequential convexity. The proof of this result appears in Appendix B.2. To facilitate our exposition, let us introduce the following notation. For any $n \geq i \geq j \geq 1$, let $D[i, j] = D_i + D_{i-1} + \ldots + D_j$ denote the total demand in periods $i, i - 1, \ldots, j$, respectively.
and set \( D[i,j] = 0 \) if \( i < j \). Also, for any \( z \in \mathcal{K} \), let the random variable \( \Lambda_i(z_i | z_{[i-1,1]}^j) \) represent the last time (after period \( i \)) in which the on-hand inventory in all of the preceding periods is greater than the order-up-to level, given the after-ordering inventory in period \( i \) is \( z_i \), that is,

\[
\Lambda_i(z_i | z_{[i-1,1]}^j) = \min \{ k | i \geq k \geq 1 \text{ and } z_i - D[i,j+1] > z_j \text{ for all } i \geq j \geq k \}.
\]

The definition of \( \Lambda_i(z_i | z_{[i-1,1]}^j) \) is similar to the definition of \( \Lambda_i(z) \) in the statement of Theorem 4.1. Here, \( \Lambda_i(z_i | z_{[i-1,1]}^j) \) depends on demands \( D_i, D_{i-1}, \ldots, D_1 \) instead of sales quantities. Note that by the definition of order-up-to policies, no order is placed during periods \( i, i-1, \ldots, \Lambda_i(z_i | z_{[i-1,1]}^j) \).

**Lemma 4.2** (Bound on the Difference in Derivatives). For any \( z \in \mathcal{K} \) and \( n \geq i \geq 1 \),

\[
f'_i(z_i | z_{[i-1,1]}^j) = \mathbb{E} \left[ h \cdot (i - \Lambda_i(z_i | z_{[i-1,1]}^j) + 1) - (h + b) \cdot 1 \{ z_i < D[i, \Lambda_i(z_i | z_{[i-1,1]}^j)] \} \right],
\]

and

\[
\left| f'_i(z_i | z_{[i-1,k+1]}^j, z_k, z_{[k-1,1]}^j) - f'_i(z_i | z_{[i-1,k+1]}^j, z_k^*, z_{[k-1,1]}^j) \right| \leq \bar{M} \left( f_k(z_k | z_{[k-1,1]}^j) - f_k(z_k^* | z_{[k-1,1]}^j) \right).
\]

Lemmas 4.1 and 4.2 show that \( \{f_i\} \) is a sequential decomposition of \( f \). To apply Theorem 3.1, we need to show that we can compute the derivative of \( f_i \) using sales data. This result is stated in the following lemma whose proof appears in Appendix B.3.

**Lemma 4.3** (Estimating Derivatives from Sales). For any \( z \in \mathcal{K} \) and \( n \geq i \geq 1 \),

\[
\mathbb{E} [G_i(Z) | Z = z] = f'_i(z_i | z_{[i-1,1]}^j).
\]

The proof of Theorem 4.1 follows immediately from the above three lemmas. Thus, we can apply Theorem 3.1 with the parameters \( \mathcal{K}, \bar{M}, \) and in this case, \( B = n \max\{b, h\} \) because \( |f'_i(z_i | z_{[i-1,1]}^j)| \leq \max\{b, h\} \) for all \( z \) and \( i \).

5 Application II: Capacity Rationing with Censored Demand

Capacity allocation of a single resource among several demand classes is one of the classical problems in revenue management (see, for example, Talluri and van Ryzin, 2004; Phillips, 2005). When the overall capacity is fixed and the demand classes arrive sequentially, the revenue manager must balance the immediate revenue from accepting current demand and the potential future revenue from higher fare classes. Examples include seats on a single-leg flight being sold to customers in multiple fare classes, and hotel rooms for a given night being sold to customers at different
rates. For this problem, Brumelle and McGill (1993) and Robinson (1995) characterize the optimal solution when the explicit knowledge of the demand distributions is available.

In many applications, however, the revenue manager does not know the demand distributions, and the available data is often limited to past sales, representing censored demand observations. We study the capacity allocation problem with monotonic fare classes, which is repeated over multiple selling seasons. A fixed capacity of units becomes available at the beginning of each season, and then the demand for each class arrives sequentially. Any demand that is not satisfied immediately upon arrival is lost and unobserved. At the end of a season, any remaining capacity perishes. The manager, who does not know the demand distributions, observes only the historical sales (censored demand) data as well as the protection levels used in the past. In this setting, we develop an adaptive non-anticipatory policy.

To our knowledge, the only other adaptive algorithm in the literature for the capacity allocation problem with censored demand that precedes this paper is by van Ryzin and McGill (2000). They consider a setting similar to ours. Their algorithm make use of the optimality condition of Brumelle and McGill (1993). Using the results from the stochastic approximation theory (Robbins and Monro, 1951), they show that, after $T$ selling seasons, the sequence of protection levels generated by their algorithm converges to the optimal protection levels at the rate of $O\left( T^{\beta/2n} \right)$ for some $\beta \in (0, 1)$, where $n+1$ denotes the number of demand classes. Note that the asymptotic convergence rate depends on $n$, the number of demand classes.

Instead of relying on the optimality condition of Brumelle and McGill (1993), our proposed algorithm is based directly on the sequential convexity of the revenue function. We show that the $T$-season average expected regret is $O\left( 1/\sqrt{T} \right)$. Note that the asymptotic convergence rate does not deteriorate as the number of demand classes increases, in contrast to the convergence rate of van Ryzin and McGill (2000).

We mention that there has been a limited number of papers on non-parametric algorithms in the revenue management literature. For the capacity allocation problem, Ball and Queyranne (2009) and Karaesmen et al. (2008) apply the competitive ratio approach, and Eren and Maglaras (2006) take the maximum entropy approach. Other applications where pricing is a decision variable include Rusmevichientong et al. (2006), Besbes and Zeevi (2006) and Kachani et al. (2006).

As before, we index each selling season forward by $t = 1, 2, \ldots$, and index each period within a season backward by $i = n + 1, n, n - 1, \ldots, 1$. We use $(t, i)$ to denote the $i^{th}$ period of the $t^{th}$
selling season. Let $D_t^i$ denote the demand in period $i$ of season $t$, and we assume that for each $i$, the random variables $D_t^1, D_t^2, \ldots$ are independent and identically distributed, having a continuous density function, and we let $D_i$ denote a random variable having the same distribution as $D_t^i$. We assume that there exists a positive constant $\mathcal{M} \geq 1$ such that for all $i$, the continuous density function of $D_i$ is bounded above by $\mathcal{M}$ for all $i$. Also, assume that $D_i \leq K$ with probability one for all $i$.

At the beginning of each period $i$, the manager observes the amount of remaining capacity, and must decide how much of this capacity to use in the current period, and how much to reserve for future periods. Given that we have $x$ units of remaining capacity at the beginning period $i$, if the manager reserves $y$ units for future periods (with $y \leq x$), then the capacity available for the current period will be $x - y$. For simplicity, assume that each unit of capacity can be used to satisfy one unit of demand. Let $p_i$ denote the exogenous selling price for each unit of demand in period $i$. In this case, the expected revenue in period $i$ will be $p_i \mathbb{E}[\min\{D_i, x - y\}]$ and the remaining capacity for the next time period is $\max\{x - D_i, y\}$. We assume that prices increase within a season, that is, $p_{n+1} \leq p_n \leq \cdots \leq p_1$. In this formulation, we allow the reserved capacity $y$ to exceed $x$. When $y > x$, it means that the manager will purchase an additional $y - x$ units of capacity for future periods at a cost of $p_i$ per unit. It can be shown that the optimal quantity to reserve decreases within a season, and thus the purchase option that we allow in our model is never exercised provided that the initial capacity is sufficiently high. Therefore, the optimal policy in our setting corresponds to the optimal policy in a setting where the purchase option is now allowed. Note that when $y > x$, $\min\{D_i, x - y\} = x - y$.

For any $i$, let $J_i(x)$ denote the value function, corresponding to the maximum expected profit that can be obtained in periods $i, i - 1, \ldots, 1$, given that we have $x$ units of capacity remaining. It is easy to verify that $J_i(\cdot)$ satisfies the following dynamic programming equation:

$$J_i(x) = \max_{y \geq 0} \{p_i \mathbb{E}[\min\{D_i, x - y\}] + J_{i-1}(\max\{x - D_i, y\})\},$$

where $x - y$ denotes the amount of capacity that will be used in the current period if $y \leq x$; or the amount of capacity that is purchased for future periods if $y > x$. We assume that any left over capacity at the end of period 1 has no value, so $J_0(\cdot) = 0$. It is easy to verify that the the value functions $J_i(\cdot)$ satisfy the following properties (see, for example, Brumelle and McGill, 1993).

**Lemma 5.1 (Value Function Property).** For each $i$, $J_i(x)$ is a continuously differentiable, increasing, and concave function in $x$. The optimal policy is characterized by a sequence $y^* = (y^*_n, y^*_{n-1}, \ldots, y^*_2, y^*_1)$.
of protection levels with \( y^*_n \geq y^*_{n-1} \geq \cdots \geq y_1^* \), where \( y^*_{i-1} \) denotes the amount of capacity that will be reserved for periods \( i-1, \ldots, 1 \) and \( y^*_{i-1} = \max \{ y : J^*_i(y) \geq p_i \} \) for all \( i \).

We are interested in the setting where the demand distributions are unknown and we only observe the sales in each period. In this setting, we have the following sequence of events in period \((t,i)\): (1) The manager observes the remaining capacity \( x^i_t \). (2) The manager determines the amount of capacity \( y^*_{i-1} \) that will be reserved for periods \( i-1, i-2, \ldots, 1 \). This decision can depend only on the historical sales and protection levels used in the past. For convenience, we define \( y^i_0 = 0 \). (3) Demand \( D^i_t \) for class \( i \) is realized, and we denote the realized demand by \( d^i_t \). However, the manager only observes the sales quantity \( s^i_t \) given by \( s^i_t = \min\{d^i_t, x^i_t - y^*_{i-1} \} \). Note that if \( y^*_{i-1} > x^i_t \), then we interpret \( s^i_t \) as the amount of capacity that will be purchased for future periods. (4) The profit \( p_i \cdot s^i_t \) is collected and recorded. If \( i > 1 \), the remaining capacity of \( x^i_{t-1} = x^i_t - s^i_t \) is available for period \( i-1 \). If \( i = 1 \), any unused capacity has no value.

For each \( i \in \{ n+1, n, \ldots, 1 \} \), let \( \mathcal{R}_i (y_i \mid y_{[i-1,1]}) \) denote the total expected profits in periods \( i, i-1, \ldots, 1 \) given that we have \( y_i \) units of capacity available at the beginning of period \( i \) and we use protection levels \( y_{[i-1,1]} = (y_{i-1}, \ldots, y_1) \) for the remaining periods. It follows from Lemma 5.1 that \( \mathcal{R}_i (y_i \mid y_{[i-1,1]}) \) satisfies the following recursion:

\[
\mathcal{R}_i (y_i \mid y_{[i-1,1]}) = E[p_i \cdot \min\{D_i, y_i - y_{i-1} \}] + E[\mathcal{R}_{i-1} (\max\{y_i - D_i, y_{i-1} \} \mid y_{[i-2,1]})],
\]

and we have that \( \mathcal{R}_i (\cdot \mid y^*_{[i-1,1]} ) = J_i (\cdot) \) for all \( i \). Let \( C > 0 \) denote the initial capacity at the beginning of period \( n+1 \) in each selling season. We are interested in determining \( \mathcal{R}_{n+1} (C \mid y^*_{[n,1]} ) \).

To do this, we introduce a series of functions that will be used for sequential convex decomposition. Let \( f_i \) be defined by:

\[
f_i (y_i \mid y_{[i-1,1]}) = \mathcal{R}_i (y_i \mid y_{[i-1,1]}) - p_{i+1} y_i ,
\]

and let \( f(y) = \sum_{i=1}^{n} f_i (y_{i} \mid y_{[i-1,1]}) \). The following lemma shows that \( f \) achieves its maximum at \( y^* = (y^*_n, \ldots, y^*_1) \) and it provides an upper bound on the revenue function. The proof of this result is given in Appendix C.1.

**Lemma 5.2** (Maximum Value and Revenue Upper Bound). Under the capacity control problem, \( \max_y f(y) = f(y^*) \) and for any protection levels \( y = (y_n, y_{n-1}, \ldots, y_1) \),

\[
\mathcal{R}_{n+1} (C \mid y^*_{[n,1]}) - \mathcal{R}_{n+1} (C \mid y_{[n,1]}) \leq f(y^*) - f(y) .
\]
The main result of this section is stated in the following theorem. Let $p_0 = 0$. Also, for each period $\ell$, let $S_\ell$ denote the sales in period $\ell$.

**Theorem 5.1** (Sequential Convexity in Capacity Allocation). In the capacity allocation model, $-f$ is a sequentially convex function with a sequential decomposition $\{-f_i\}$. Moreover, we can compute an unbiased estimate of the gradient estimate of $f'_i(y_i | y_{[i-1,1]})$ using the sales information, that is,

$$f'_i(y_i | y_{[i-1,1]}) = E[G_i(Y) | Y = y] ,$$

where for any $y$,

$$G_i(y) = p\Lambda_i(y) - p_{i+1}$$

and we define $\Lambda_i(y)$ by

$$\Lambda_i(y) = \max \left\{ k : i \geq k \geq 1 \text{ and } \sum_{\ell=k}^i S_\ell \geq y_k - y_{k-1} \right\} ,$$

if there exists $k$ such that $\sum_{\ell=k}^i S_\ell \geq y_k - y_{k-1}$ holds; otherwise, we set $\Lambda_i(y) = 0$.

### 5.1 Proof of Theorem 5.1

It follows from the definition that for any $i$,

$$f'_i(y_i | y^*_i) = R'_i(y_i | y^*_i) - p_{i+1} ,$$

and from Lemma 5.1, we have that $y^*_i = \max\{ y : R'_i(y | y^*_i) \geq p_{i+1} \}$. Thus, $f_i(\cdot | y^*_i)$ is a concave function with a maximum at $y^*_i$. This verifies parts (a) and (b) of sequential convexity. Part (c) of the sequential convexity definition follows immediately from the following lemma that establishes an upper bound on the difference in the function $f_i$ and the difference in its derivative. The proof of this lemma uses a similar technique to the proofs of Lemmas 4.1 and 4.2, and the details are given in Appendix C.2.

**Lemma 5.3** (Bounds on the Differences). For any $y = (y_n, \ldots, y_1) \in K$ and $n \geq i > k \geq 1$,

$$|f_i(y_i | y^*_{[i-1,k+1]}, y_k^*, y^*_{[k-1,1]}) - f_i(y_i | y_{[i-1,k+1]}, y_k, y^*_{[k-1,1]})| \leq f_k(y^*_k | y^*_{[k-1,1]}) - f_k(y_k | y^*_{[k-1,1]}) ,$$

and

$$|f'_i(y_i | y_{[i-1,k+1]}, y_k, y^*_{[k-1,1]}) - f'_i(y_i | y_{[i-1,k+1]}, y^*_k, y^*_{[k-1,1]})| \leq M \left[ f_k(y^*_k | y^*_{[k-1,1]}) - f_k(y_k | y^*_{[k-1,1]}) \right] .$$
The above lemma proves that \( \{-f_i\} \) is a sequential decomposition of \(-f\). Then, Theorem 5.1 follows immediately from the following lemma, whose proof appears in Appendix C.3.

**Lemma 5.4 (Estimating Derivatives from Sales).** For any \( z \in K \) and \( n \geq i \geq 1 \),

\[
    f'_i(z_i | z_{i-1,1}) = E[G_i(Z) \mid Z = z].
\]

Note that we can apply Theorem 3.1 to the capacity allocation problem with the parameters \( \mathcal{K}, \mathcal{M} \), and in this case, \( B = np_1 \) because \( |f'_i(y_i | y_{i-1,1})| \leq p_1 \) for all \( y \) and \( i \).

6 Application III: Lifetime Buy

The lifetime buy decision refers to the problem of deciding how many parts to buy at the end of their lifecycles if the lifecycle of the product that uses these components extends beyond the lifecycles of these parts. This problem has become important as the components parts become obsolete quickly due to technological advances, resulting in shorter and more frequent product lifecycles. This problem has been studied by Pierskalla (1969), Brown et al. (1964), David et al. (1997), and Cattani and Souza (2003) in the case of a single part becoming obsolete. More recently, Bradley and Guerrero (2009) considered the case of multiple parts becoming obsolete, and they characterized the optimal solution as a threshold-based policy.

All of the above-mentioned papers assume that the demand distributions are known a priori. However, it is sometimes difficult to obtain information, and our adaptive approach here relies on historical data, applicable when products undergo ongoing upgrades based on newer versions of component parts. We will show how our theory of sequential convexity can be applied to this example. For simplicity, we will focus on the case where the product consists of component 1 and component 2, where the lifecycle of component 2 ends before that of component 1. The extension to the general \( n \) components is straightforward. Let \( D_2 \) be the random product demand after component 2 becomes obsolete until the end of component 1’s lifecycle, and let \( D_1 \) be the random product demand between the end of component 1’s lifecycle and the end of the product’s lifecycle. We assume that the density of \( D_2 \) is bounded above by \( \mathcal{M} \). Let \( c_1 \) and \( c_2 \) denote the per unit ordering cost for components 1 and 2, respectively, and let \( \pi \) denote the profit margin (excluding the ordering cost).

The decision variables are the total amount of “lifetime buy” for components 1 and 2, respectively, which will be denoted by \( y_1 \) and \( y_2 \). Given the decisions \( y_1 \) and \( y_2 \), here is the sequence
of events: (1) At the end of component 2’s lifecycle, the manager makes a lifetime purchase of $y_2$ units of component 2 to stock up to meet the future demand of the product. As in Bradley and Guerrero (2009), since component 1 continues production, we assume that we can procure it instantaneously using a regular ordering channel. (2) Demand $D_2$ for the product is realized, but we only observe the sales given by $\min\{D_2, y_2\}$. An appropriate profit is earned based on the resulting sales and ordering decision. Let $x_2 = [y_2 - D_2]^+$ denote the remaining inventory of component 2. Note that $x_2$ acts as an upper bound on the amount of the lifetime purchase for component 1. (3) At the end of component 1’s lifecycle, the manager makes a lifetime purchase of component 1 in the amount of $\min\{y_1, x_2\}$ to stock up to meet the future demand of the product, where $y_1$ is the decision that represents the target production quantity. (4) Demand $D_1$ for the product is realized, but we only observed the sales $\min\{D_1, \min\{y_1, x_2\}\}$. An appropriate profit is earned based on the resulting sales and ordering decision.

Let $\Pi_1(y_1)$ denote the total expected profit from the end of the lifecycle of component 1 until the product becomes obsolete given that we have a lifetime purchase target decision of $y_1$, and let $\Pi_2(y_2|y_1)$ be the total expected profit from the end of the lifecycle of component 2 until the product becomes obsolete given that we have decided to make a lifetime purchase of $y_1$ and $y_2$. Note that the functions $\Pi_1(\cdot)$ and $\Pi_2(\cdot|\cdot)$ satisfy the following dynamic programming recursion:

$$\Pi_1(y_1) = \pi \mathbb{E} \left[ \min\{y_1, D_1\} \right] - c_1 y_1$$

$$= (\pi - c_1) \mathbb{E}[D_1] - \mathbb{E} \left\{ (\pi - c_1)(D_1 - y_1)^+ + c_1 (y_1 - D_1)^+ \right\}$$

$$\Pi_2(y_2|y_1) = (\pi - c_1) \mathbb{E} \left[ \min\{y_2, D_2\} \right] - c_2 y_2 + \mathbb{E} \left[ \Pi_1 \left( \min\{y_1, [y_2 - D_2]^+\} \right) \right]$$

$$= (\pi - c_1 - c_2) \mathbb{E}[D_2] - \mathbb{E} \left\{ (\pi - c_1 - c_2)(D_2 - y_2)^+ + c_2 (y_2 - D_2)^+ \right\} + \mathbb{E} \left[ \Pi_1 \left( \min\{y_1, x_2\} \right) \right],$$

where $x_2 = [y_2 - D_2]^+$ denotes the remaining inventory of component 2. The corresponding optimization problem is given by:

$$\max_{y_1, y_2} \Pi_2(y_2|y_1),$$

and let $y_1^*$ and $y_2^*$ denote the optimal “planned” lifetime purchases for components 1 and 2, respectively. It is a standard result for this model that $y_1^* = \arg \max \Pi_1(y_1)$ and $y_2^* = \arg \max \Pi_2(y_2|y_1^*)$, and that both $\Pi_1(\cdot)$ and $\Pi_2(\cdot|y_1^*)$ are concave functions.

**Connection to Inventory Example:** The lifetime buy example shares some similarities with the inventory example in Section 4. The one-period reward function is essentially the same as the overage plus underage cost in the classical inventory example. The main difference is the way the
“state variable” is updated. In the inventory example, the inventory after ordering in period 1 is
given by \( \max\{y_1, y_2 - D_2\} = \max\{y_1, x_2\} \), while in the lifetime buy example, the amount of lifetime
purchase for component 1 is \( \min\{y_1, x_2\} \).

The sequential decomposition for our lifetime buy example is given by:

\[
f_1(y_1) = \Pi_1(y_1) \quad \text{and} \quad f_2(y_2|y_1) = \Pi_2(y_2|y_1) - \Pi_1(y_1) ,
\]
and let \( f(y_2, y_1) = f_2(y_2|y_1) + f_1(y_1) \). The following theorem shows that \( \{f_1, f_2\} \) is a sequential
convex decomposition.

**Theorem 6.1** (Sequential Convexity in Lifetime Buy). In the lifetime buy model, \(-f\) is a sequentially convex function with a sequential decomposition \( \{-f_i\} \). Furthermore, the derivative of each
\( f_i \) can be estimated from sales data.

**Proof.** By definition, \( f_1(\cdot) \) and \( f_2(\cdot|y_1^*\) achieve its maximum at \( y_1^* \) and \( y_2^* \), respectively. Thus,
\( \{-f_1, -f_2\} \) satisfies parts (a) and (b) in the definition of a sequential decomposition. Now, we will
show that

\[
| f_2(y_2|y_1^*) - f_2(y_2|y_1) | \leq f_1(y_1^*) - f(y_1) .
\]

To prove this, note that the two systems have the same planned lifetime buy of \( y_2 \) for component 2,
but they differ in the amount of lifetime buy for component 1. Thus, the inventory \( X_2 = [y_2 - D_2]^+ \)
of component 2 will be the same in both systems, and we have

\[
f_2(y_2|y_1^*) - f_2(y_2|y_1) = \mathbb{E}_{X_2} [ f_1(\min\{X_2, y_1^*\}) - f_1(\min\{X_2, y_1\}) ] \leq f_1(y_1^*) - f_1(y_1) ,
\]
where the last inequality follows from the fact that \( f_1(\cdot) \) is a concave function that achieves a
maximum at \( y_1^* \).

We now consider the derivative condition. The expression for the derivatives of \( f_1 \) and \( f_2 \) follows
immediately from the definition. To complete the proof, note that

\[
f_2'(y_2|y_1^*) - f_2'(y_2|y_1) = \mathbb{E}_{D_2} [ f_1'(y_2 - D_2) \cdot (1\{0 < y_2 - D_2 < y_1^*\} - 1\{0 < y_2 - D_2 < y_1\}) ] .
\]

There are two cases to consider: \( y_1 < y_1^* \) and \( y_1 > y_1^* \). Suppose that \( y_1 < y_1^* \). Then, we have that

\[
1\{0 < y_2 - D_2 < y_1\} \leq 1\{0 < y_2 - D_2 < y_1^*\}
\]

with probability one, which implies that

\[
| f_2'(y_2|y_1^*) - f_2'(y_2|y_1) | \leq \mathbb{E}_{D_2} [ | f_1'(y_2 - D_2) | \cdot 1\{y_1 < y_2 - D_2 < y_1^*\} ] = \int_{y_1}^{y_1^*} f_1'(w) \psi(w) dw ,
\]
where $\psi(\cdot)$ denotes the density function of the random variable $W = y_2 - D_2$. Using the fact that $f_1'(w) \geq 0$ for all $y_1 < w < y_1^*$ and that the density of $D_2$ is bounded above by $\mathcal{M}$, it is easy to verify that $\psi$ is also bounded above by $\mathcal{M}$, and we show that
\[
|f_2'(y_2|y_1^*) - f_2'(y_2|y_1)| \leq \mathcal{M} (f_1(y_1^*) - f_1(y_1)) .
\]
Similarly, we can argue for the case $y_1 > y_1^*$.

Finally, it can be shown that
\[
\begin{align*}
f_1'(y_1) &= E_{D_1} [\pi 1[D_1 \geq y_1] - c_1] \\
f_2'(y_2|y_1) &= E_{D_1,D_2} [\{\pi - c_1\} 1[D_2 \geq y_2] - c_2 + \{\pi 1[D_1 \geq y_2 - D_2] - c_1\} 1[y_2 - D_2 \leq y_1]] .
\end{align*}
\]

It can be verified that the expression inside the expectation can be computed based on sales, and thus sample path derivatives can also be computed from sales data.

\section{Experiments}

In this section, we focus on the capacity allocation problem with censored demand from Section 5, and evaluate the performance of our adaptive algorithm on several demand distributions and parameter settings. We choose the revenue management application as a testbed for our experiment because we wish to compare our proposed method with the adaptive algorithm of van Ryzin and McGill (2000) – which we refer to as the VM Algorithm – that is based on stochastic approximation. For our experiments, we modify the adaptive algorithm in Figure 1 slightly by setting the step size $\epsilon_t = O(1/t)$, which is similar to the step size used in van Ryzin and McGill (2000). We also maintain a sequence of monotone protection levels in every season. We refer to our modified algorithm as the Adaptive Revenue Management (ARM) Algorithm.

In Section 7.1, we consider four demand classes, replicating the setup considered in the paper by van Ryzin and McGill (2000). Then, in Section 7.2, we consider a larger number of demand classes. When the number of demand classes is small, our numerical results show that the revenues generated by both the VM and ARM algorithms are comparable. However, as the number of demand classes increases, our ARM algorithm converges to the optimal revenue more quickly.

\subsection{Four Demand Classes}

In this section, we consider the setting involving four demand classes, replicating the original experiments conducted in van Ryzin and McGill (2000). The demand for each class follows a
Gaussian distribution with given parameters. Table 1 shows the mean and the standard deviation of the demand distributions, along with the corresponding fare and the optimal protection levels.

<table>
<thead>
<tr>
<th>Class</th>
<th>Fare</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Optimal Protection Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1,050</td>
<td>17.3</td>
<td>5.8</td>
<td>16.7</td>
</tr>
<tr>
<td>2</td>
<td>$567</td>
<td>45.1</td>
<td>15.0</td>
<td>44.6</td>
</tr>
<tr>
<td>3</td>
<td>$527</td>
<td>73.6</td>
<td>17.4</td>
<td>134.0</td>
</tr>
<tr>
<td>4</td>
<td>$350</td>
<td>19.8</td>
<td>6.6</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 1: Fares, demand distributions, and the optimal protection levels for each demand class.

The stepsize of the VM Algorithm is given by $\varepsilon_{VM,t} = \frac{200}{t+10}$ as in the original experiment in van Ryzin and McGill (2000), and the stepsize of our ARM Algorithm is set to $\varepsilon^t = \frac{C}{p_1^t}$ where $C$ and $p_1$ denote the capacity and the most expensive fare, respectively. As in the original experiments, we consider four different parameter settings, corresponding to different capacity levels and initial protection levels. The four settings are given in Table 2. Note that in Case II and IV, we set the initial protection level for class 3 to the value of the capacity.

<table>
<thead>
<tr>
<th>Capacity</th>
<th>Initial Protection Levels for Classes 1, 2 &amp; 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>124</td>
<td>Low: (0, 15, 65)</td>
</tr>
<tr>
<td>164</td>
<td>Case I</td>
</tr>
<tr>
<td></td>
<td>Case II</td>
</tr>
<tr>
<td></td>
<td>Case III</td>
</tr>
<tr>
<td></td>
<td>Case IV</td>
</tr>
</tbody>
</table>

Table 2: Capacities and initial protection levels for each of the four cases considered in the experiment involving four demand classes.

Figure 2 shows the comparison between the running average revenue under the optimal protection levels, the VM Algorithm, and our ARM Algorithm over 1000 problem instances. We use the revenue model that is used in van Ryzin and McGill (2000), i.e., without the purchase option in each period. For each problem instance, we consider 1000 time periods and plot the running average revenue over time. The dash lines above and below the solid lines represent the 95% confidence interval. As seen from the figures, in all four cases, the revenue generated by both the VM Algorithm and our ARM Algorithm are comparable, converging to the same value.
Running Average Revenue Under VM and ARM Algorithms

(1000 problem instances, 4 class, 124 capacity, low initial values)

NOTE: Dash lines correspond to 95% confidence intervals

$70,500$ $70,600$ $70,700$ $70,800$ $70,900$ $71,000$ $71,100$ $71,200$ $71,300$ $71,400$ $71,500$ $71,600$

10 110 210 310 410 510 610 710 810 910

Time (t)

Running Avg. Rev. Up To Time t

VM Algorithm

ARM Algorithm

Optimal

Figure 2: Running average revenue under the VM and ARM Algorithms with four demand classes under four different parameter settings.

### 7.2 Larger Numbers of Demand Classes

In this section, we compare the performance of the VM and ARM Algorithms when the number of demand classes is larger. We consider 8 and 12 demand classes. The demand distribution for each class is normally distributed. We generate the mean and standard deviation, along with the fare, of each class as follows. From the 4-class setting in Section 7.1, let $C_0 = 124$ and let $I_0 = \{(f_i, \mu_i, \sigma_i) : 1 \leq i \leq 4\}$ denote the collection of fare, mean, and standard deviation for each of the four demand classes considered in Section 7.1. Let

\[
C_1 = 1.1 \times C_0 \quad \text{and} \quad I_1 = \{1.1 \times (f_i, \mu_i, \sigma_i) : 1 \leq i \leq 4\}
\]

\[
C_2 = 1.2 \times C_0 \quad \text{and} \quad I_2 = \{1.2 \times (f_i, \mu_i, \sigma_i) : 1 \leq i \leq 4\}.
\]
We use the following parameters for the 8-class and 12-class settings.

<table>
<thead>
<tr>
<th>Settings</th>
<th>Capacity</th>
<th>Fares, Means, and Stdev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>8-class</td>
<td>( C_0 + C_1 )</td>
<td>( I_0 \cup I_1 )</td>
</tr>
<tr>
<td>12-class</td>
<td>( C_0 + C_1 + C_2 )</td>
<td>( I_0 \cup I_1 \cup I_2 )</td>
</tr>
</tbody>
</table>

Table 3: Parameters for 8-class and 12-class settings, respectively.

The initial protection levels for each class \( i \) is set to the total expected demand from class 1 through \( i \). We use the same step sizes for both algorithms given by \( \varepsilon^t = C/(p_1 \cdot t) \).

Figure 3 shows the running average revenue under both algorithms for 8-class and 12-class settings, respectively, averaged over 1000 problem instances. For each problem instance, we run both algorithms for 10,000 time periods, and plot the running average revenue. From the figure, the running average revenue under our ARM ALGORITHM is higher than the revenue under the VM ALGORITHM, and the difference appears to be statistically significant. This result is consistent with the observation that the convergence rate of the VM ALGORITHM deteriorates as the number of demand classes increases.

8 Conclusion

Motivated by applications in supply chain, revenue management, and lifetime buy problems, we present a general theory for minimizing sequentially convex functions. To this effect, we have extended existing results in online stochastic optimization to allow for biased gradient estimates.
believe our work offers venues into many directions for future research, including finding a broader
class of objective functions and identifying additional applications. It would also be interesting to
explore ways to reduce biases in our gradient estimates.

Acknowledgement
We are grateful to the associate editor and the referees for their thoughtful and detailed com-
ments on the earlier version of the paper. Their suggestions greatly improve the quality and the
presentation of our work.

References
Ball, M. O., and M. Queyranne. 2009. Toward Robust Revenue Management: Competitive Analysis of Online


Bradley, J. R., and H. H. Guerrero. 2009. Lifetime Buy Decisions with Multiple Obsolete Parts. *Production and


(1): 127–137.


Research* 146 (1): 216–228.


Chen, L., and E. Plambeck. 2008. Dynamic Inventory Management with Learning about Demand Distribution and

David, I., E. Greenshtein, and A. Mehrez. 1997. A Dynamic-Programming Approach to Continuous-Review Obso-


Gradient Descent Without a Gradient. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete


Huh, W. T., and P. Rusmevichientong. 2009. A Non-Parametric Approach to Stochastic Inventory Planning with


Appendices for Online Supplement

A Proof of Theorem 2.1

Let \( z^* = \arg\min_{z \in S} \Phi(z) \). Also, let \( \langle \cdot, \cdot \rangle \) denote a dot product operator. We first claim that, for any \( T \geq 1 \),

\[
\sum_{t=1}^{T} \mathbb{E} \left[ \langle \nabla \Phi(z'), z' - z^* \rangle \right] \leq \sum_{t=1}^{T} \left\{ \frac{\mathbb{E} \| z' - z^* \|^2}{2\epsilon^t} - \frac{\mathbb{E} \| z'^{t+1} - z^* \|^2}{2\epsilon^t} \right\} + B^2 \sum_{t=1}^{T} \frac{\epsilon^t}{2} + \frac{\text{diam}(S) \sum_{t=1}^{T} \mathbb{E} \| \text{Bias}(z') \|}{2} .
\]

(1)

Since \( z'^{t+1} \) is recursively defined in terms of \( z'^t \) and the projection operator is non-expansive, that is, it does not increase the distance between two points,

\[
\mathbb{E} \| z'^{t+1} - z^* \|^2 \leq \mathbb{E} \| z'^t - \epsilon^t \cdot G(z'^t) - z^* \|^2 \\
= \mathbb{E} \| (z'^t - z^*) - \epsilon^t \cdot G(z'^t) \|^2 \\
= \mathbb{E} \| z'^t - z^* \|^2 + (\epsilon^t)^2 \mathbb{E} \| G(z'^t) \|^2 - 2\epsilon^t \cdot \mathbb{E} \left[ \langle G(z'^t), z'^t - z^* \rangle \right],
\]

for any \( t \geq 1 \). Thus, by arranging the above inequality,

\[
\mathbb{E} \left[ \langle G(z'^t), z'^t - z^* \rangle \right] \leq \frac{\mathbb{E} \| z'^t - z^* \|^2}{2\epsilon^t} - \frac{\mathbb{E} \| z'^{t+1} - z^* \|^2}{2\epsilon^t} + \frac{\epsilon^t}{2} \mathbb{E} \| G(z'^t) \|^2 .
\]

(2)

Also, it follows from the definition of \( G(z'^t) \) and \( \text{Bias}(z'^t) \) that

\[
\mathbb{E} \left[ \langle G(z'^t), z'^t - z^* \rangle \right] = \mathbb{E} \left[ \mathbb{E} \left[ \langle G(z'^t), z'^t - z^* \rangle \mid z'^t \right] \right] \\
= \mathbb{E} \left[ \langle \mathbb{E} \left[ G(z'^t) \mid z'^t \right], z'^t - z^* \rangle \right] \\
= \mathbb{E} \left[ \langle \nabla \Phi(z'^t) + \text{Bias}(z'^t), z'^t - z^* \rangle \right] \\
= \mathbb{E} \left[ \langle \nabla \Phi(z'^t), z'^t - z^* \rangle \right] + \mathbb{E} \left[ \langle \text{Bias}(z'^t), z'^t - z^* \rangle \right].
\]

(3)

Combining (2) and (3),

\[
\mathbb{E} \left[ \langle \nabla \Phi(z'^t), z'^t - z^* \rangle \right] = \mathbb{E} \left[ \langle G(z'^t), z'^t - z^* \rangle \right] - \mathbb{E} \left[ \langle \text{Bias}(z'^t), z'^t - z^* \rangle \right] \\
\leq \frac{\mathbb{E} \| z'^t - z^* \|^2}{2\epsilon^t} - \frac{\mathbb{E} \| z'^{t+1} - z^* \|^2}{2\epsilon^t} + \frac{\epsilon^t}{2} \mathbb{E} \| G(z'^t) \|^2 - \mathbb{E} \left[ \langle \text{Bias}(z'^t), z'^t - z^* \rangle \right] .
\]

Thus, it follows

\[
\sum_{t=1}^{T} \mathbb{E} \left[ \langle \nabla \Phi(z'^t), z'^t - z^* \rangle \right] \leq \sum_{t=1}^{T} \left\{ \frac{\mathbb{E} \| z'^t - z^* \|^2}{2\epsilon^t} - \frac{\mathbb{E} \| z'^{t+1} - z^* \|^2}{2\epsilon^t} \right\} + B^2 \sum_{t=1}^{T} \frac{\epsilon^t}{2} + \frac{\text{diam}(S) \sum_{t=1}^{T} \mathbb{E} \| \text{Bias}(z'^t) \|}{2} ,
\]

26
where the last inequality follows from Cauchy-Schwarz's inequality. This establishes the inequality (1) in the claim.

We will now establish an upper bound on the first two terms in the statement of the claim in (1).

\[
\sum_{t=1}^{T} \left\{ \frac{E \| Z_t - z^* \|^2}{2 \epsilon_t} - \frac{E \| Z_{t+1} - z^* \|^2}{2 \epsilon_t} \right\} + B^2 \sum_{t=1}^{T} \frac{\epsilon_t}{2} \\
\leq \frac{E \| Z_1 - z^* \|^2}{2 \epsilon_1} + \frac{1}{2} \sum_{t=1}^{T} \left\{ \frac{1}{\epsilon_t} - \frac{1}{\epsilon_{t+1}} \right\} E \| Z_t - z^* \|^2 + B^2 \sum_{t=1}^{T} \epsilon_t \\
\leq \frac{\text{diam}(S)^2}{2} \left\{ \frac{1}{\epsilon_1} + \sum_{t=1}^{T} \left\{ \frac{1}{\epsilon_{t+1}} - \frac{1}{\epsilon_t} \right\} \right\} + B^2 \sum_{t=1}^{T} \epsilon_t \\
= \frac{\text{diam}(S)^2}{2 \epsilon_1} + B^2 \sum_{t=1}^{T} \epsilon_t \\
= \frac{\text{diam}(S) B \sqrt{T+1}}{2} + \frac{\text{diam}(S) B}{2} \sum_{t=1}^{T} \frac{1}{\sqrt{t}} \leq 2 \text{diam}(S) B \sqrt{T},
\]

where the last equality follows from the definition of \( \epsilon_t \), for \( t \in \{1, \ldots, T+1\} \), and the final inequality follows from the fact that \( \sqrt{T+1} \leq 2 \sqrt{T} \) and the fact that \( \sum_{t=1}^{T} \leq 2 \sqrt{T} \).

The result of Theorem 2.1 follows from the convexity of \( \Phi \), which implies that

\[
\frac{1}{T} \sum_{t=1}^{T} E \left[ \Phi(Z_t') - \Phi(z^*) \right] \leq \frac{1}{T} \sum_{t=1}^{T} \langle \nabla \Phi(Z_t'), Z_t' - z^* \rangle \leq \frac{2 B \text{diam}(S)}{\sqrt{T}} + \frac{\text{diam}(S)}{T} \sum_{t=1}^{T} E \| \text{Bias}(Z_t') \|.
\]

B Proof of Lemmas in Section 4

B.1 Proof of Lemma 4.1

We compare two sets of order-up-to policies from period \( i \) to 1 defined by the following order-up-to levels: \( (z_{i,k+1}, z_k, z_{k-1,1}) \) and \( (z_{i,k+1}, z^*_k, z^*_{k-1,1}) \). Since the order-up-to levels are the same for periods from \( i \) to \( k+1 \), the costs incurred by both policies during these periods are the same. Let \( X_k \) be the random variable representing the inventory level at the beginning of period \( k \). Thus, the inventory level after ordering in period \( k \) will be \( \max\{X_k, z_k\} \) for the first system, and \( \max\{X_k, z^*_k\} \) for the second system. Thus,

\[
f_i(z_i|z_{i-1,k+1}, z_k, z_{k-1,1}) - f_i(z_i|z_{i-1,k+1}, z^*_k, z^*_{k-1,1}) \\
= E_{X_k} \left[ f_k(\max\{X_k, z_k\}|z^*_{k-1,1}) - f_k(\max\{X_k, z^*_k\}|z^*_{k-1,1}) \right] .
\]

Consider the following claim: for any real-valued convex function \( \phi \) defined on a closed interval, any pair of \( w \in \mathbb{R} \) and \( w' \in \mathbb{R} \) satisfies \( 0 \leq \phi(\max\{w', w\}) - \phi(\max\{w', w^*\}) \leq \phi(w) - \phi(w^*) \), where \( w^* \) is a minimizer of \( \phi \). The proof of this claim follows by considering all possible orderings among \( w^*, w \) and \( w' \), and using the fact that \( \phi(w) \) is weakly decreasing if \( w < w^* \), and weakly increasing if \( w > w^* \). Since \( f_k(z^*_{k-1,1}) \) is a convex function achieving its minimum at \( z^*_k \), the above claim implies, for any realized value \( x_k \) of \( X_k \),

\[
0 \leq f_k(\max\{x_k, z_k\}|z^*_{k-1,1}) - f_k(\max\{x_k, z^*_k\}|z^*_{k-1,1}) \leq f_k(z_k|z^*_{k-1,1}) - f_k(z^*_k|z^*_{k-1,1}) .
\]

By taking an expectation with respect to \( X_k \), we obtain the required result.
B.2 Proof of Lemma 4.2

From the definition of \( f_i \) and \( g_i \),

\[
f'_i \left( z_i \mid z_{[i-1,i]} \right) = g'_i \left( z_i \mid z_{[i-1,i]} \right)
= Q'_i(z_i) + \frac{d}{dz_i} \mathbb{E}_{D_i} \left[ g_{i-1} \left( \max \{ z_i - D_i, z_{i-1} \} \mid z_{[i-2,1]} \right) \right]
= Q'_i(z_i) + \mathbb{E}_{D_i} \left[ g'_{i-1} \left( z_i - D_i \mid z_{[i-2,1]} \right) \cdot 1 \{ z_i - D_i > z_{i-1} \} \right]
= Q'_i(z_i) + \mathbb{E}_{D_i} \left[ f'_{i-1} \left( z_i - D_i \mid z_{[i-2,1]} \right) \cdot 1 \{ z_i - D_i > z_{i-1} \} \right].
\]

By applying the above argument iteratively, we obtain

\[
f'_i \left( z_i \mid z_{[i-1,i]} \right) = \mathbb{E} \left[ \sum_{k=1}^{i} Q'_k(z_i - D[i, k+1]) \cdot 1 \{ k \geq \Lambda_i \left( z_i \mid z_{[i-1,i]} \right) \} \right]
\]

where the expectation is taken over \((D_i, D_{i-1}, \ldots, D_1)\). Recall \( Q'_k(z_i) = h - (h + b) \cdot \mathbb{E} \{ z_i < D_k \} \). For any \( k \in \{ i, i-1, \ldots, 1 + \Lambda_i \left( z_i \mid z_{[i-1,i]} \right) \} \), \( z_i - D[i, k] > z_{i-1} \geq 0 \) holds by the definition of \( \Lambda_i \left( z_i \mid z_{[i-1,i]} \right) \), implying \( z_i - D[i, k+1] > D_k \). If \( k = \Lambda_i \left( z_i \mid z_{[i-1,i]} \right) \), then we have \( z_i - D[i, k+1] < D_k \), which is equivalent to \( z_i < D[i, k] \). Thus,

\[
\sum_{k = \Lambda_i \left( z_i \mid z_{[i-1,i]} \right)}^{i} Q'_k(z_i - D[i, k+1])
= \mathbb{E} \left[ h \cdot (i - \Lambda_i \left( z_i \mid z_{[i-1,i]} \right)) \right] + \mathbb{E} \left[ h - (h + b) \cdot 1 \{ z_i < D \left[ i, \Lambda_i \left( z_i \mid z_{[i-1,i]} \right) \right] \} \right],
\]

and we obtain the first desired result.

For the second result, observe that the telescoping sum and the triangle inequality imply

\[
\left| f'_i \left( z_i \mid z_{[i-1,i]} \right) - f'_i \left( z_i \mid z_{[i-1,i]} \right) \right| \leq \sum_{k=1}^{i-1} \left| f'_i \left( z_i \mid z_{[k-1,k]}, z_{[k,k+1]} \right) \right| - f'_i \left( z_i \mid z_{[k-1,k]}, z_{[k',k']} \right) \right|.
\]

Thus, it suffices to show the following result that, for any \( z = (z_n, \ldots, z_1) \in \mathcal{K} \) and \( n \geq i > k \geq 1 \),

\[
\left| f'_i \left( z_i \mid z_{[i-1+1]}, z_{[i,k]}, z_{[k+1]}' \right) - f'_i \left( z_i \mid z_{[i-1,i]}, z_{[k+1]}, z_{[k,k+1]} \right) \right| \leq \mathbb{M} \cdot \left| f_k \left( z_i \mid z_{[k-1,k+1]} \right) - f_k \left( z_i \mid z_{[k-1,k+1]} \right) \right|.
\]

From the above derivative expression \( f'_i \left( z_i \mid z_{[i-1,i]} \right) \), for any \( z \in \mathcal{K} \),

\[
f'_i \left( z_i \mid z_{[i-1,i]} \right) = \mathbb{E} \left[ \sum_{j=k+1}^{i} Q'_j (z_{[i, j+1]} \mid z_{[i-1,i]}) \right] + \mathbb{E} \left[ f'_k \left( z_i \mid z_{[k,k+1]} \mid z_{[k,k+1]} \right) \right] \cdot 1 \{ k \geq \Lambda_i \left( z_i \mid z_{[i-1,i]} \right) \}.
\]

Since the event \( j \geq \Lambda_i \left( z_i \mid z_{[i-1,i]} \right) \) depends on \( z_{[i, j]} \) only through \( z_{[i,j]} = (z_i, z_{i-1}, \ldots, z_j) \),

\[
f'_i \left( z_i \mid z_{[i-1,i]}, z_{[i-1,i+1]}, z_{[k,k+1]} \right) - f'_i \left( z_i \mid z_{[i-1,i]}, z_{[k,k+1]}, z_{[k,k+1]} \right)
= \mathbb{E} \left[ f'_k \left( W \mid z_{[k,k+1]} \right) \right] \cdot 1 \{ k \geq \Lambda_i \left( z_i \mid z_{[i-1,i]}, z_{[k,k+1]} \right) \}
- \mathbb{E} \left[ f'_k \left( W \mid z_{[k,k+1]} \right) \right] \cdot 1 \{ k \geq \Lambda_i \left( z_i \mid z_{[i-1,i]}, z_{[k,k+1]} \right) \}.
\]

28
where $W = z_i - D[i, k + 1]$.

We consider the case where $z_k^* \leq z_k$. (The case of $z_k^* > z_k$ can be argued similarly.) Then,

$$
\Lambda_i \left( z_i \mid z_{[i-1,k+1]}, z_k^*, z_{[k-1,1]} \right) \leq \Lambda_i \left( z_i \mid z_{[i-1,k+1]}, z_k, z_{[k-1,1]} \right),
$$

which implies that

$$
f_i' \left( z_i \mid z_{[i-1,k+1]}, z_k, z_{[k-1,1]} \right) - f_i' \left( z_i \mid z_{[i-1,k+1]}, z_k^*, z_{[k-1,1]} \right) = \mathbb{E} \left[ f_i'(W \mid z_{[k-1,1]}) \cdot 1 \left\{ \Lambda_i \left( z_i \mid z_{[i-1,k+1]}, z_k^*, z_{[k-1,1]} \right) \leq k < \Lambda_i \left( z_i \mid z_{[i-1,k+1]}, z_k, z_{[k-1,1]} \right) \right\} \right].
$$

However, by definition of $\Lambda_i$ and the definition of $W = z_i - D[i, k + 1]$, the event in the indicator of the above equation is a subset of the event that $W \in [z_k^*, z_k]$. Then,

$$
\left| f_i' \left( z_i \mid z_{[i-1,k+1]}, z_k, z_{[k-1,1]} \right) - f_i' \left( z_i \mid z_{[i-1,k+1]}, z_k^*, z_{[k-1,1]} \right) \right| \leq \mathbb{E} \left[ f_i'(W \mid z_{[k-1,1]}) \cdot 1 \left[ W \in [z_k^*, z_k] \right] \right] = \int_{z_k^*}^{z_k} f_i'(w \mid z_{[k-1,1]}) \psi(w) dw,
$$

where $\psi$ denotes the density function of $W = z_i - D[i, k + 1]$.

We claim that the density function of $W$ is bounded above by $\overline{M}$. To see this claim, consider $D_1$ and $D_2$ whose density functions $\psi_1$ and $\psi_2$ are bounded above by $\overline{M}_1$ and $\overline{M}_2$, respectively. Let $\psi_{12}$ be the density function of $D_1 + D_2$. Then, for any $w \geq 0$,

$$
\psi_{12}(w) = \int_0^w \phi_1(v) \phi_2(w - v) dv \leq \overline{M}_1 \int_0^w \phi_1(v) dv \leq \overline{M}_2.
$$

By applying a similar logic iteratively, we complete the proof of the claim.

To complete the proof of the lemma, note that the convex function $f_k \left( \cdot \mid z_{[k-1,1]}^* \right)$ achieves its minimum value at $z_k^*$, and thus, $f_k'(w \mid z_{[k-1,1]}^*) \geq 0$ for all $z_k^* \leq w \leq z_k$. Therefore,

$$
\int_{z_k^*}^{z_k} f_k'(w \mid z_{[k-1,1]}^*) \psi(w) dw \leq \overline{M} \int_{z_k^*}^{z_k} f_k'(w \mid z_{[k-1,1]}^*) dw = \overline{M} \cdot \left[ f_k(z_k \mid z_{[k-1,1]}^*) - f_k(z_k^* \mid z_{[k-1,1]}^*) \right].
$$

### B.3 Proof of Lemma 4.3

For $n \geq i \geq 1$, recall the definitions of $\Lambda_i(z)$ and $\Lambda_i(z_i \mid z_{[i-1,1]})$. To prove the result, it suffices to show that, for each $i$, (i) $\Lambda_i(z) = \Lambda_i(z_i \mid z_{[i-1,1]})$ and (ii) $1 \left\{ z_i \leq \sum_{j=\Lambda_i(z_i \mid z_{[i-1,1]})}^{j=i} D_j \right\} = 1 \left\{ z_i = \sum_{j=\Lambda_i(z)}^{j=S_j} S_j \right\}$ for each sample path of demand realizations. From the definition of $\Lambda_i(z_i \mid z_{[i-1,1]})$, it can be shown that in periods between $i$ and $\Lambda_i(z_i \mid z_{[i-1,1]}) + 1$ inclusive, the sales quantity is the same as the demand. Thus, it can be shown that (i) holds. Furthermore, since no inventory is ordered in these periods, it can be shown that the two events in (ii) are the same.

### C Proof of Lemmas in Section 5

#### C.1 Proof of Lemma 5.2

We first prove the following claim by induction on $i$:

$$
\max_{(y_{k}, y_{k-1}, \ldots, y_{1})} \sum_{k=1}^{i} f_k \left( y_k \mid y_{[k-1,1]} \right) = \sum_{k=1}^{i} f_k \left( y_k^* \mid y_{k-1,1}^* \right).
$$
For $i = 1$, we have that $f_1(y_i) = \mathcal{R}_1(y_i) - p_2 y_1$. It follows from Lemma 5.1 that $y^*_i = \max\{y : \mathcal{R}^*_i(y) \geq p_i\}$. Thus, $f_1(\cdot)$ is a concave function with a maximum at $y^*_1$. Suppose the result holds for $k = 1, 2, \ldots, i - 1$. Then, 

$$
\max_{y_i, y_{i-1}, \ldots, y_1} \sum_{k=1}^{i} f_k(y_k \mid y_{k-1,1}) = \max_{y_i, y_{i-1}, \ldots, y_1} f_i(y_i \mid y_{i-1,1}) + \sum_{k=1}^{i-1} f_k(y_k \mid y_{k-1,1}) = \max_{y_i} - p_{i+1} y_i + \max_{y_{i-1}, \ldots, y_1} \mathcal{R}_i(y_i \mid y_{i-1,1}) + \sum_{k=1}^{i-1} f_k(y_k \mid y_{k-1,1}) = \max_{y_i} - p_{i+1} y_i + \mathcal{R}_i\left(y_i \mid y^*_i, y^*_{i-1,1}\right) + \sum_{k=1}^{i-1} f_k(y_k \mid y_{k-1,1}) \, ,
$$

where the last equality follows from Lemma 5.1 which shows that $\max_{y_{i-1}, \ldots, y_1} \mathcal{R}_i(y_i \mid y_{i-1,1}) = J_i(y_i) = \mathcal{R}_i\left(y_i \mid y^*_i, y^*_{i-1,1}\right)$, and the induction hypothesis that $\max_{y_{i-1}, \ldots, y_1} \sum_{k=1}^{i-1} f_k(y_k \mid y_{k-1,1}) = \sum_{k=1}^{i-1} f_k(y_k^* \mid y^*_{k-1,1})$.

Finally, it follows from Lemma 5.1 that $y^*_i = \max\{y : J'_i(y) \geq p_{i+1}\} = \max\{y : \mathcal{R}^*_i(y_i \mid y^*_{i-1,1}) \geq p_{i+1}\}$, and thus, the maximum objective value is achieved at $y^*_i$, which completes the induction and proves the claim.

Now, we will now show that $\mathcal{R}_{n+1}\left(C \mid y^*_{n+1}\right) - \mathcal{R}_{n+1}\left(C \mid y_{n+1}\right) \leq f(y^*) - f(y)$ . By definition and telescoping sums,

$$
\mathcal{R}_{n+1}\left(C \mid y^*_{n+1}\right) - \mathcal{R}_{n+1}\left(C \mid y_{n+1}\right) = \sum_{i=1}^{n} \left[\mathcal{R}_{n+1}\left(C \mid y_{n+1}, y_i^*, y^*_{i-1,1}\right) - \mathcal{R}_{n+1}\left(C \mid y_{n+1}, y_i, y^*_{i-1,1}\right)\right] .
$$

Let $Y_{i+1}$ be a random variable that denotes the remaining capacity at the beginning of period $i + 1$ given that we have used protection levels $y_n, \ldots, y_{i+1}$. Note that $Y_{i+1}$ depends on $D_{n+1}, \ldots, D_{i+2}$. Then, it follows from the definition of $\mathcal{R}(\cdot)$ that

$$
\mathcal{R}_{n+1}\left(C \mid y_{n+1}, y_i^*, y^*_{i-1,1}\right) - \mathcal{R}_{n+1}\left(C \mid y_{n+1}, y_i, y^*_{i-1,1}\right) = \mathbb{E}\left[\mathcal{R}_{i+1}\left(Y_{i+1} \mid y_i^*, y^*_{i-1,1}\right) - \mathcal{R}_{i+1}\left(Y_{i+1} \mid y_i, y^*_{i-1,1}\right)\right] = \mathbb{E}\left[p_{i+1} \left(\min\{Y_{i+1} - y_i^*, D_{i+1}\} - \min\{Y_{i+1} - y_i, D_{i+1}\}\right) + \mathcal{R}_i\left(W_i^* \mid y^*_{i-1,1}\right) - \mathcal{R}_i\left(W_i \mid y^*_{i-1,1}\right)\right] = \mathbb{E}\left[p_{i+1} \left\{W_i - W_i^*\right\} + \mathcal{R}_i\left(W_i^* \mid y^*_{i-1,1}\right) - \mathcal{R}_i\left(W_i \mid y^*_{i-1,1}\right)\right] = \mathbb{E}\left[f_i\left(W_i \mid y^*_{i-1,1}\right) - f_i\left(W_i^* \mid y^*_{i-1,1}\right)\right] \leq f_i\left(y_i^* \mid y^*_{i-1,1}\right) - f_i\left(y_i \mid y^*_{i-1,1}\right) ,
$$

where $W_i = \max\{Y_{i+1} - D_{i+1}, y_i\}$ and $W_i^* = \max\{Y_{i+1} - D_{i+1}, y_i^*\}$. The final inequality follows from the fact that the function $f_i\left(\cdot \mid y^*_{i-1,1}\right)$ is concave and achieves its maximum at $y_i^*$. By telescoping sum, we
have that
\[
\begin{align*}
& f_i \left( y^*_i \mid y_{[i-1, 1]} \right) - f_i \left( y_i \mid y_{[i-1, 1]} \right) \\
= & \ f_i \left( y^*_i \mid y_{[i-1, 1]} \right) - f_i \left( y_i \mid y_{[i-1, 1]} \right) + \sum_{k=1}^{i-1} f_i \left( y_i \mid y_{[i-1, k+1]}, y^*_k, y^*_k \mid k-1, 1 \right) - f_i \left( y_i \mid y_{[i-1, k+1]}, y_k, y^*_k \mid k-1, 1 \right) \\
= & \ f_i \left( y^*_i \mid y_{[i-1, 1]} \right) - f_i \left( y_i \mid y_{[i-1, 1]} \right) + \sum_{k=1}^{i-1} \mathcal{R}_i \left( y_i \mid y_{[i-1, k+1]}, y^*_k, y^*_k \mid k-1, 1 \right) - \mathcal{R}_i \left( y_i \mid y_{[i-1, k+1]}, y_k, y^*_k \mid k-1, 1 \right) \\
= & \ f_i \left( y^*_i \mid y_{[i-1, 1]} \right) - f_i \left( y_i \mid y_{[i-1, 1]} \right) + \sum_{k=1}^{i-1} \mathbb{E} \left[ f_k \left( W_k \mid y^*_{[k-1, 1]} \right) - f_k \left( W_k \mid y_{[k-1, 1]} \right) \right] \\
\geq & \ f_i \left( y^*_i \mid y_{[i-1, 1]} \right) - f_i \left( y_i \mid y_{[i-1, 1]} \right),
\end{align*}
\]

where for any \( k \), \( W_k = \max \{ Y_{k+1} - D_{k+1}, y_k \} \) and \( W_k^* = \max \{ Y_{k+1} - D_{k+1}, y_k^* \} \). The final inequality follows from the fact that the function \( f_k \left( \cdot \mid y^*_{[k-1, 1]} \right) \) is concave and achieves its maximum at \( y_k^* \), which implies that for all \( x \in \mathbb{R} \),
\[
f_k \left( \max \{ x, y_k^* \} \mid y^*_{[k-1, 1]} \right) - f_k \left( \max \{ x, y_k \} \mid y^*_{[k-1, 1]} \right) \geq 0.
\]

Putting everything together, we have that
\[
\mathcal{R}_{n+1} \left( C \mid y^*_{[n, 1]} \right) - \mathcal{R}_{n+1} \left( C \mid y_{[n, 1]} \right) \\
= \sum_{i=1}^{n} \left[ \mathcal{R}_{n+1} \left( C \mid y_{[i-1, i]}, y_i^*, y^*_{[i-1, 1]} \right) - \mathcal{R}_{n+1} \left( C \mid y_{[i+1, i]}, y_i, y^*_{[i-1, 1]} \right) \right] \\
\leq \sum_{i=1}^{n} f_i \left( y_i^* \mid y^*_{[i-1, 1]} \right) - f_i \left( y_i \mid y^*_{[i-1, 1]} \right) \leq \sum_{i=1}^{n} f_i \left( y_i^* \mid y^*_{[i-1, 1]} \right) - f_i \left( y_i \mid y_{[i-1, 1]} \right) \leq f(y^*) - f(y).
\]

C.2 Proof of Lemma 5.3

We will first establish a bound on the difference between \( f_i \). As in Section 4, we compare the following two sets of protection levels: \( \left( y_{[i, k+1]}, y_k, y^*_k \mid [k-1, 1] \right) \) and \( \left( y_{[i, k+1]}, y_k^*, y^*_{[k-1, 1]} \right) \). These protection levels are the same for periods \( i \) to \( k+1 \), which implies that the profits during these periods will be the same under both policies. Let \( X_{k+1} \) denote the available capacity at the beginning of period \( k+1 \). Then, we have that
\[
\begin{align*}
& f_i \left( y_i \mid y_{[i-1, k+1]}, y_k, y^*_k \mid [k-1, 1] \right) - f_i \left( y_i \mid y_{[i-1, k+1]} + y_k, y^*_k \mid [k-1, 1] \right) \\
= & \mathcal{R}_i \left( y_i \mid y_{[i-1, k+1]}, y_k^*, y^*_{[k-1, 1]} \right) - \mathcal{R}_i \left( y_i \mid y_{[i-1, k+1]}, y_k, y^*_k \mid [k-1, 1] \right) \\
= & p_{k+1} \mathbb{E} \left[ \min \{ D_{k+1} + X_{k+1} - y_k^* \} \right] - p_{k+1} \mathbb{E} \left[ \min \{ D_{k+1} + X_{k+1} - y_k \} \right] \\
+ & \mathbb{E} \left[ \mathcal{R}_k \left( \max \{ X_{k+1} - D_{k+1}, y_k^* \} \mid y^*_{[k-1, 1]} \right) - \mathcal{R}_k \left( \max \{ X_{k+1} - D_{k+1}, y_k \} \mid y^*_{[k-1, 1]} \right) \right] \\
= & -p_{k+1} \mathbb{E} \left[ \max \{ X_{k+1} - D_{k+1}, y_k^* \} \right] + p_{k+1} \mathbb{E} \left[ \max \{ X_{k+1} - D_{k+1}, y_k \} \right] \\
+ & \mathbb{E} \left[ \mathcal{R}_k \left( \max \{ X_{k+1} - D_{k+1}, y_k^* \} \mid y^*_{[k-1, 1]} \right) - \mathcal{R}_k \left( \max \{ X_{k+1} - D_{k+1}, y_k \} \mid y^*_{[k-1, 1]} \right) \right] \\
= & \mathbb{E} \left[ f_k \left( \max \{ X_{k+1} - D_{k+1}, y_k^* \} \mid y^*_{[k-1, 1]} \right) - f_k \left( \max \{ X_{k+1} - D_{k+1}, y_k \} \mid y^*_{[k-1, 1]} \right) \right]
\end{align*}
\]

where the third equality follows from the fact that for any triplets of real numbers \( a, b, \) and \( c, \min \{ c, a - b \} = a - \max \{ a - c, b \} \). Since \( f_k \left( \cdot \mid y^*_{[k-1, 1]} \right) \) is a concave function with a maximum at \( y_k^* \), it is easy to verify that for any \( x \in \mathbb{R} \),
\[
0 \leq f_k \left( \max \{ x, y_k^* \} \mid y^*_{[k-1, 1]} \right) - f_k \left( \max \{ x, y_k \} \mid y^*_{[k-1, 1]} \right) \leq f_k \left( y_k^* \mid y^*_{[k-1, 1]} \right) - f_k \left( y_k \mid y^*_{[k-1, 1]} \right).
\]
We will now derive the expression for the difference in the derivative. It follows from the definition that
\[ f'_i (y_i \mid y_{i-1,1}) = -p_{i+1} + R'_i (y_i \mid y_{i-1,1}), \]
and it follows from the definition of \( R_i \) that
\[
\begin{align*}
  f'_i (y_i \mid y_{i-1,1}) & = -p_{i+1} + p_i E \left[ 1 \left\{ y_i - y_{i-1} \leq D_i \right\} \right] + E \left[ 1 \left\{ D_i < y_i - y_{i-1} \right\} \cdot R'_{i-1} \left( y_{i-1} - D_i \mid y_{i-1,1} \right) \right] \\
  & = -p_{i+1} + p_i + E \left[ 1 \left\{ D_i < y_i - y_{i-1} \right\} \cdot (-p_i + R'_{i-1} \left( y_i - D_i \mid y_{i-1,1} \right)) \right] \\
  & = -p_{i+1} + p_i + E \left[ 1 \left\{ i - 1 \geq \Lambda_i (y_i \mid y_{i-1,1}) \right\} \cdot (f'_{i-1} (y_i - D_i \mid y_{i-1,1})) \right] \\
  & = -p_{i+1} + p_i + E \left[ 1 \left\{ i - 1 \geq \Lambda_i (y_i \mid y_{i-1,1}) \right\} \cdot (-p_i + R_{i-1}) \right] \\
  & \quad + E \left[ 1 \left\{ i - 2 \geq \Lambda_i (y_i \mid y_{i-1,1}) \right\} \cdot (f'_{i-2} (y_i - D_i - D_i - 1 \mid y_{i-1,1})) \right],
\end{align*}
\]
where \( \Lambda_i (y_i \mid y_{i-1,1}) = \max \left\{ k \mid i \geq k \geq 1 \text{ and } D[i, k] \geq y_i - y_{k-1} \right\} \) if the right-hand side expression is well-defined; otherwise, 0. Repeated application of the above argument shows that for any \( y \) and \( i \geq k, \)
\[
  f'_i (y_i \mid y_{i-1,1}) = E \left[ \sum_{j=k+1}^i (-p_{j+1} + p_j) \cdot 1 \left\{ j \geq \Lambda_i (y_i \mid y_{i-1,1}) \right\} \right] \\
  \quad + E \left[ f'_k (y_i - D[i, k + 1] \mid y_{[k-1,1]}) \cdot 1 \left\{ k \geq \Lambda_i (y_i \mid y_{i-1,1}) \right\} \right].
\]

Then, using virtually the same argument as in Lemma 4.2, we can show
\[
  \left| f'_i (y_i \mid y_{[i-1,k+1]}, y_k, y_{[i-1,k+1]}) - f'_i (y_i \mid y_{[i-1,k+1]}, y_k, y_{[i-1,k+1]}) \right| \leq \mathcal{M} \left[ f_k (y_k \mid y_{[k-1,1]}) - f_k (y_k \mid y_{[k-1,1]}) \right].
\]

### C.3 Proof of Lemma 5.4

Comparing the definitions of \( \Lambda_i (y) \) and \( \Lambda_i (y_i \mid y_{i-1,1}) \), the main difference is whether sales quantity or demand is used. It can be argued that these values are identical for each sample path since the sales quantity is the same as demand in periods between \( i \) and \( \Lambda_i (y) - 1 \) or \( \Lambda_i (y_i \mid y_{i-1,1}) - 1 \), inclusive. This explains how the sales data can be used in computing the sample-path derivative that is originally given in terms of demand data.