Dynamic Assortment Optimization with a Multinomial Logit Choice Model and Capacity Constraint

Paat Rusmevichientong* — Cornell University

Zuo-Jun Max Shen† — UC Berkeley

David B. Shmoys‡ — Cornell University

September 22, 2009

Abstract

We consider an assortment optimization problem where a retailer chooses an assortment of products that maximizes the profit subject to a capacity constraint. The demand is represented by a multinomial logit choice model. We consider both the static and dynamic optimization problems. In the static problem, we assume that the parameters of the logit model are known in advance; we then develop a simple algorithm for computing a profit-maximizing assortment based on the geometry of lines in the plane, and derive structural properties of the optimal assortment. For the dynamic problem, the parameters of the logit model are unknown and must be estimated from data. By exploiting the structural properties found for the static problem, we develop an adaptive policy that learns the unknown parameters from past data, and at the same time, optimizes the profit. Numerical experiments based on sales data from an online retailer indicate that our policy performs well.

1. Introduction

The problem of learning customer preferences and offering a profit-maximizing assortment of products, subject to a capacity constraint, has applications in retail, online advertising, and revenue management. For instance, given a limited shelf capacity, a retailer must determine the assortment of products that maximizes the profit (see, for example, Mahajan and van Ryzin, 1998, 2001). The retailer might not know the demand a priori, and a customer's product selection often depends on the assortment offered. In this case, the retailer can learn the demand distribution by offering

*School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA. E-mail: paatrus@cornell.edu

†Department of Industrial Engineering and Operations Research, University of California–Berkeley, 4129 Etcheverry Hall, Berkeley, CA 94720, USA. E-mail: shen@ieor.berkeley.edu

‡School of Operations Research and Information Engineering and Department of Computer Science, Cornell University, Ithaca, NY 14853, USA. E-mail: shmoys@cs.cornell.edu

different assortments, observing purchases, and estimating the demand model from past sales and assortment decisions (Caro and Gallien, 2007).

In online advertising, the capacity constraint may represent the limited number of locations on the web page where the ads can appear. The demand for each product corresponds to the number of customers who click on the ad. The probability that a customer will click on a particular ad will likely depend on the assortment of ads shown. Given the uncertainty in the demand for each ad and the limited number of locations where the ads can be shown, we must decide on the assortment of ads that will generate the most profit, adjusting our assortment decisions, and refining our demand estimates over time as new data arrive.

Modeling customer choice behavior and estimating demand distributions are active areas of research in revenue management. When the choice model is known, the focus is to determine the assortment of itinerary and fare combinations that maximizes the total revenue (see, for example, Talluri and van Ryzin, 2004; Liu and van Ryzin, 2008; Kunnumkal and Topaloglu, 2008; and Zhang and Adelman, 2008). When the parameters of the demand distribution are unknown, researchers have developed techniques for estimating the choice model from sales data (Ratliff et al., 2007; Vulcano et al., 2008).

1.1 The Model

Motivated by the above applications, we formulate a stylized dynamic assortment optimization model that captures some of the issues commonly present in these problems, namely the capacity constraint, the uncertainty in the demand distribution, and the dependence of the purchase or selection probability on the assortment offered. Assume that we have N products indexed by $1, 2, \ldots, N$. Let $\boldsymbol{w} = (w_1, \ldots, w_N) \in \mathbb{R}_+^N$ denote a vector of marginal profits, where for each i, $w_i > 0$ denotes the marginal profit of product i. The option of no purchase is denoted by 0 with $w_0 = 0$. Through appropriate scaling, we will assume without loss of generality that $w_i \leq 1$ for all i. Due to a capacity constraint, we can offer at most C products to the customers, where $C \geq 2$. The goal is to determine a profit-maximizing assortment of at most C products.

We represent the demand using the multinomial logit (MNL) choice model, which is one of the most commonly used choice models in economics, marketing, and operations management (see, Ben-Akiva and Lerman (1985), Anderson et al. (1992), Mahajan and van Ryzin (1998), and the references therein). Under the MNL model, each customer chooses the product that maximizes her utility, where the utility U_i of product i is given by: $U_i = \mu_i + \zeta_i$, where $\mu_i \in \mathbb{R}$ denotes the mean utility that the customer assigns to product i. We assume that $\zeta_0, \ldots, \zeta_N$ are independent and identically distributed random variables having a Gumbel distribution with location parameter 0 and scale parameter 1. Without loss of generality, we set $\mu_0 = 0$, and let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N) \in \mathbb{R}^N$ denote the vector of mean utilities.

Following the terminology in Vulcano et al. (2008), we define a "customer preference vector" $\boldsymbol{v} = (v_1, \ldots, v_N) \in \mathbb{R}_+^N$, where $v_i = e^{\mu_i}$ for all i, and set $v_0 = 1$. Given an assortment $S \subseteq \{1, 2, \ldots, N\}$, the probability $\theta_i(S)$ that a customer chooses product i is given by:

$$\theta_i(S) = \begin{cases} v_i / \left(1 + \sum_{k \in S} v_k\right), & \text{if} \quad i \in S \cup \{0\}, \\ 0, & \text{otherwise} , \end{cases} \tag{1}$$

and the expected profit $f(S)$ associated with the assortment S is defined by:

$$f(S) = \sum_{i \in S} w_i \theta_i(S) = \frac{\sum_{i \in S} w_i v_i}{1 + \sum_{i \in S} v_i} \ . \tag{2}$$

We will consider two problems: *static* and *dynamic* optimizations. In static optimization, we assume that $\boldsymbol{v}$ is known in advance, and we wish to find the assortment with at most C products that gives the maximum expected profit, corresponding to the following combinatorial optimization problem:

$$(\textbf{Capacitated MNL}) \qquad Z^* = \max \left\{ f(S) \ : \ S \subseteq \{1, \ldots, N\} \text{ and } |S| \leq C \right\} \ . \tag{3}$$

In dynamic optimization, on the other hand, the vector $\boldsymbol{v}$ is *unknown*, and we have to infer its value by offering different assortments over time and observing the customer selections. For simplicity, we will assume that we can offer an assortment to a single customer in each time period[1]. For each assortment $S \subseteq \{1, \ldots, N\}$ and $t \geq 1$, let the random variables $X_t(S)$ and $Y_t(S)$ denote the selection and the reward, respectively, associated with offering the assortment S in period t. The random variable $X_t(S)$ takes values in $\{0\} \cup S$ and has the following probability distribution: for each $i \in \{0\} \cup S$,

$$\Pr\{X_t(S) = i\} = \theta_i(S) = \frac{v_i}{1 + \sum_{k \in S} v_k} \ .$$

In addition, the random variable $Y_t(S)$ is given by $Y_t(S) = w_{X_t(S)}$, and we have that $\mathsf{E}[Y_t(S)] = \sum_{i \in S} w_i \Pr\{X_t(S) = i\} = f(S)$. For each t, let $\mathcal{H}_t$ denote the set of possible histories until the end of period t. A policy $\psi = (\psi_1, \psi_2, \ldots)$ is a sequence of functions, where $\psi_t : \mathcal{H}_{t-1} \to \{S \subseteq \{1, \ldots, N\} : |S| \leq C\}$ selects an assortment of size C or less in period t based on the history until the end of period $t - 1$. The T-period cumulative regret under the policy ψ is defined by:

$$\text{Regret}(T, \psi) = \sum_{t=1}^{T} \mathsf{E}\left[Z^* - Y_t(S_t)\right] = \sum_{t=1}^{T} \mathsf{E}\left[Z^* - f(S_t)\right] \ ,$$

[1]This assumption is introduced primarily to simplify our exposition. Our analysis extends to the setting where a single assortment is offered to multiple customers in each period.

where $S_1, S_2, \ldots,$ denote the sequence of assortments offered under the policy ψ. Note that S_t is a random variable that depends on the selections $X_1(S_1), X_2(S_2), \ldots, X_{t-1}(S_{t-1})$ of customers in the preceding $t-1$ periods. We are interested in finding a policy that minimizes the regret, which is equivalent to maximizing the total expected reward $\sum_{t=1}^{T} \mathsf{E}\left[f(S_t)\right]$.

1.2 Contributions and Organization

Our work illuminates the structure of the capacitated assortment optimization problem under the MNL model, both in static and dynamic settings. For static optimization, Example 2.1 shows how the optimal assortment, under a capacity constraint, exhibits different structural properties from the optimal solution in the uncapacitated setting. This example demonstrates that we must be careful in applying our intuition from the uncapacitated problem. Megiddo (1979) presents a recursive algorithm for optimizing a rational objective function, which can be applied to the **Capacitated MNL** problem. Because the algorithm recursively invokes subroutines and traverses a computational tree, we do not know how changes in the customer preference vector $\boldsymbol{v}$ affect the optimal assortment and profit. The lack of a transparent relationship between $\boldsymbol{v}$ and Z^* makes it very difficult to apply and analyze this algorithm in a dynamic setting, where $\boldsymbol{v}$ is unknown and must be estimated from data.

So, in Section 2.1, we describe an alternative algorithm – which we refer to as StaticMNL – that is non-recursive and is based on a simple geometry of lines in the two-dimensional plane. Although the StaticMNL algorithm builds upon the ideas introduced in Megiddo (1979), our algorithm demonstrates a simple and transparent relationship between the preference vector $\boldsymbol{v}$ and the optimal assortment, enabling us to extend it to the dynamic optimization setting. To our knowledge, this is the first result that characterizes the sensitivity of the optimal assortment to changes in the customer preferences (see Theorem 2.4).

The StaticMNL algorithm generates a sequence $\mathcal{A} = \langle A_0, A_1, \ldots, A_K \rangle$ of assortments, with $K = O(N^2)$, that is guaranteed to contain the optimal solution (Theorem 2.2). By exploiting the properties of the MNL model, we show in Theorem 2.5 that the number of distinct assortments in the sequence $\mathcal{A}$ is of order $O(NC)$. We also prove that the sequence of profits $\langle f(A_0), f(A_1), \ldots, f(A_K) \rangle$ is unimodal, and establish a lower bound on the difference between the profit of any two consecutive assortments in the sequence $\mathcal{A}$ (Theorem 2.6). We then show how we can exploit these structural properties to derive an efficient search algorithm based on golden ratio search (Lemma 2.7). To our knowledge, these structural properties associated with the MNL model are new, and they can potentially be extended to more complex choice models such as the nested logit (see, for example, Rusmevichientong et al., 2009).

We exploit the geometric insights and the wealth of structural properties to extend the StaticMNL algorithm to the dynamic optimization setting, where $\boldsymbol{v}$ is unknown. In Section 3, we describe a policy that adaptively learns the customer preference over time, and establish an $O(\log^2 T)$ upper bound on the regret. Saure and Zeevi (2008) has improved the regret bound to $O(\log T)$ and prove that this is the minimal possible regret. Our analysis of the regret also establishes a connection between our estimates of the customer preference and maximum likelihood estimation, enabling us to generalize our results to the linear-in-parameters utility model. The results of the numerical experiments in Section 4 based on sales data from an online retailer show that our policy performs well.

Although our results build upon the existing work in the literature, our refinements enable us to discover previously unknown relationships between the customer preference and the optimal assortment in a capacitated setting. The newly discovered insights help us to develop a policy for joint parameter estimation and assortment optimization. The synthesis of results from diverse communities to address an important practical problem represents one of the main contributions of our work.

1.3 Literature Review

This paper contributes to the literature in both static and dynamic assortment planning. The static assortment planning (where the underlying demand distribution is assumed be known in advance) has an extensive literature, and we refer the reader to Kok et al. (2008) for an excellent review of the current state of the art. We will focus on a few papers that are closely related to our work.

Our work is part of a growing literature on modeling customer choice behavior in revenue management. Talluri and van Ryzin (2004) consider the multi-period single-resource revenue management problem under a general discrete choice model, where the objective is to determine the assortment of fare products over time that maximizes the total revenue. They characterize the optimal assortments in terms of nondominated sets. This pioneering work has been extended to the general network revenue management setting (see, for example, Shen and Su, 2007; Kunnumkal and Topaloglu, 2008; Zhang and Adelman, 2008, and the references therein).

The **Capacitated MNL** model can be viewed as a single-period problem, and is an extension of the unconstrained optimization problem studied by Gallego et al. (2004) and Liu and van Ryzin (2008), who describe a beautiful algorithm for finding the optimal assortment based on sorting the products in a descending order of marginal profits. They use this subroutine as part of the column-generation method for solving the choice-based linear programming model for network revenue management. As shown in Example 2.1, when there is a capacity constraint, sorting the products based on marginal profits alone can lead to suboptimal solutions. It turns out that the assortments generated by the StaticMNL algorithm are similar to the nondominated sets introduced by Talluri and van Ryzin (2004) (more on this in Section 2.1).

Our dynamic optimization formulation can be viewed as an instance of the multiarmed bandit problem (see, Lai and Robbins, 1985 and Auer et al., 2002). We can view each product as an arm whose reward profile is unknown. In each period, we can choose up to C arms (equivalent to offering up to C products) with the goal of minimizing regret. Many researchers have studied this problem (see, for example, Anantharam et al., 1987a,b), but most of the literature assumes that the reward of each arm is independent of the assortment of arms chosen, which is not applicable to our setting where there are substitution effects.

To our knowledge, the first paper that consider the dynamic optimization where the reward is contingent on the assortment of arms chosen is the pioneering work of Caro and Gallien (2007), who consider a stylized multiarmed bandit model assuming independent demand at first, and develop an effective dynamic index policy for determining the product assortment. They then extend their model to account for substitution among products, but their substitution model is very different from the MNL model considered here.

Although this is not our focus, our work in dynamic optimization is also related to the research on estimating the customer choice behavior from data. Much of the research in this area focuses on developing techniques for inferring the underlying demand from censored observations. Ratliff et al. (2007) describe a heuristic for unconstraining the demand across multiple flights and fare classes. Vulcano et al. (2008) consider an approach for estimating substitutes and lost sales based on first-choice demand. In our formulation, we ignore the inventory consideration and assume that every product in the assortment is available to the customers. We focus primarily on designing an adaptive assortment policy with minimal regret.

2. Static Optimization

In this section, we assume that the preference vector $\boldsymbol{v}$ is known and develop an algorithm for solving the **Capacitated MNL** problem. As shown in the following example, in contrast to the unconstrained problem (Gallego et al., 2004; Liu and van Ryzin, 2008), a greedy policy that sequentially adds one product to an assortment (until the profit decreases) may lead to a suboptimal solution when we have a capacity constraint.

Example 2.1. Consider 4 products with $w_1 = 9.5$, $w_2 = 9.0$, $w_3 = 7.0$, $w_4 = 4.5$, and $v_1 = 0.2$,

$v_2 = 0.6$, $v_3 = 0.3$, $v_4 = 5.2$. The expected profit for each assortment is given in the table below, and the maximum of each column is highlighted in bold.

S	$f(S)$	S	$f(S)$	S	$f(S)$	S	$f(S)$
$\{1\}$	1.583	$\{1,2\}$	4.056	$\{1,2,3\}$	**4.476**	$\{1,2,3,4\}$	**4.493**
$\{2\}$	3.375	$\{1,3\}$	2.667	$\{1,2,4\}$	4.386		
$\{3\}$	1.615	$\{1,4\}$	3.953	$\{1,3,4\}$	4.090		
$\{4\}$	**3.774**	$\{2,3\}$	3.947	$\{2,3,4\}$	4.352		
		$\{2,4\}$	**4.253**				
		$\{3,4\}$	3.923				

The optimal assortment for each value of C is given by:

C	1	2	3	4
Optimal Assortment	$\{4\}$	$\{2,4\}$	$\{1,2,3\}$	$\{1,2,3,4\}$

When $C = 1$ and $C = 2$, the optimal assortments include product 4, which has the lowest marginal profit, but this product is not included in the optimal assortment when $C = 3$. Yet, it re-appears when there is no capacity constraint ($C = 4$). Moreover, for $C = 3$, under the greedy policy that sequentially adds a product until the profit decreases, we would get the assortment $\{1, 2, 4\}$, which is suboptimal.

Although we can apply the algorithm of Megiddo (1979) to solve the **Capacitated MNL** problem, the algorithm is recursive and difficult to visualize. Moreover, the algorithm does not provide a simple and transparent relationship between the preference vector $\boldsymbol{v}$ and the optimal assortment, making it difficult to analyze its performance in a dynamic setting. In the next section, we present a geometric algorithm that can be easily visualized (see Example 2.3), and the simple geometry associated with our method illuminates how the optimal assortment changes with the parameter $\boldsymbol{v}$, providing the first sensitivity analysis for this class of problems (Theorem 2.4). We also derive novel structural properties (Theorems 2.5 and 2.6), which can be used to develop an efficient search procedure for the optimal assortment (Theorem 2.7).

2.1 A Geometric Non-Recursive Polynomial-time Algorithm

The key idea underlying our algorithm is the observation that we can express the optimal profit Z^* in Equation (3) as follows:

$$\begin{aligned} Z^* &= \max \{\lambda \in \mathbb{R} : \exists X \subseteq \{1, \ldots, N\}, |X| \leq C, \text{ and } f(X) \geq \lambda\} \\ &= \max \left\{ \lambda \in \mathbb{R} : \exists X \subseteq \{1, \ldots, N\}, |X| \leq C, \text{ and } \sum_{i \in X} v_i (w_i - \lambda) \geq \lambda \right\} \\ &= \max \left\{ \lambda \in \mathbb{R} : \max_{X : |X| \leq C} \sum_{i \in X} v_i (w_i - \lambda) \geq \lambda \right\} , \end{aligned}$$

where the second equality follows from the definition of the profit function $f(\cdot)$ given in Equation (2). Let the functions $A : \mathbb{R} \to \{X \subseteq \{1, \ldots, N\} : |X| \leq C\}$ and $g : \mathbb{R} \to \mathbb{R}$ be defined by: for each $\lambda \in \mathbb{R}$,

$$A(\lambda) = \arg \max_{X:|X|\leq C} \sum_{i \in X} v_i \left(w_i - \lambda\right) \qquad \text{and} \qquad g(\lambda) = \sum_{i \in A(\lambda)} v_i \left(w_i - \lambda\right) \ , \tag{4}$$

where we break ties arbitrarily. Therefore, $Z^* = \max\{f\left(A(\lambda)\right) : \lambda \in \mathbb{R}\}$, and to find the optimal assortment, it suffices to enumerate $A(\lambda)$ for all values of $\lambda \in \mathbb{R}$. We will show that the collection of assortments $\{A(\lambda) : \lambda \in \mathbb{R}\}$ has at most $O(N^2)$ sets. For each $\lambda \in \mathbb{R}$, it follows from Proposition 1 in Talluri and van Ryzin (2004) that $A(\lambda)$ can be interpreted as a nondominated set among all subsets of size C or less[2].

Before we describe the algorithm, let us provide some geometric intuition. For $i = 0, 1, \ldots, N$, let the linear function $h_i : \mathbb{R} \to \mathbb{R}$ be defined by: for each $\lambda \in \mathbb{R}$,

$$h_0(\lambda) = 0 \quad \text{and} \quad h_i(\lambda) = v_i \left(w_i - \lambda\right), \quad \text{for} \ \ i = 1, \ldots, N. \tag{5}$$

The number of intersection points among the $N+1$ lines $h_0(\cdot), \ldots, h_N(\cdot)$ is at most $\binom{N+1}{2} = O(N^2)$. Suppose we sort these intersection points based on their x-coordinates. It follows from Equation (4) that, for each $\lambda \in \mathbb{R}$, $A(\lambda)$ corresponds to the top C lines among $h_0(\cdot), h_1(\cdot), \ldots, h_N(\cdot)$ whose values at λ are nonnegative. Then, for an arbitrary λ strictly between two consecutive intersection points, the ordering of the values $h_j(\lambda) = v_j\left(w_j - \lambda\right)$ remain constant and their values do not change sign. Therefore, for any λ strictly between the two consecutive intersection points, $A(\lambda)$ remains the same. So, to enumerate $A(\lambda)$ for all $\lambda \in \mathbb{R}$, it suffices to enumerate all of the intersection points among the $N+1$ lines. This observation forms the basis of the StaticMNL algorithm described below.

For all $0 \leq i < j \leq N$ with $v_i \neq v_j$, let $\mathcal{I}(i,j)$ denote the x-coordinate of the intersection point between the lines $h_i(\cdot)$ and $h_j(\cdot)$, that is,

$$h_i\left(\mathcal{I}(i,j)\right) = h_j\left(\mathcal{I}\left(i,j\right)\right) \quad \Leftrightarrow \quad \mathcal{I}(i,j) = \frac{v_i w_i - v_j w_j}{v_i - v_j} \ .$$

Let $\boldsymbol{\tau} = \left(\left(i_1, j_1\right), \ldots, \left(i_K, j_K\right)\right)$ denote the ordering of the intersection points, that is, $i_\ell < j_\ell$ for each $\ell = 1, \ldots, K$ and

$$-\infty \equiv \mathcal{I}(i_0, j_0) < \mathcal{I}(i_1, j_1) \leq \mathcal{I}(i_2, j_2) \leq \cdots \leq \mathcal{I}(i_K, j_K) < \mathcal{I}(i_{K+1}, j_{K+1}) \equiv +\infty, \tag{6}$$

[2]To apply Proposition 1 in Talluri and van Ryzin (2004), for any assortment S of size C or less, we can define $R(S) = \sum_{i \in S} v_i w_i$ and $Q(S) = \sum_{i \in S} v_i$.

where we have added the two end points $\mathcal{I}(i_0, j_0)$ and $\mathcal{I}(i_{K+1}, j_{K+1})$ to facilitate our exposition. Also, let $\boldsymbol{\sigma}^0 = \left(\sigma_1^0, \ldots, \sigma_N^0\right)$ denote the ordering of the customer preference weights from highest to lowest, that is,

$$v_{\sigma_1^0} \geq \cdots \geq v_{\sigma_N^0} \ . \tag{7}$$

The ordering $\boldsymbol{\sigma}^0$ is the ordering of the lines $h_1(\cdot), h_2(\cdot), \ldots, h_N(\cdot)$ at $\lambda = -\infty$ from the highest to the lowest values. The StaticMNL algorithm maintains the following *four* pieces of information associated with the interval $\left(\mathcal{I}\left(i_\ell, j_\ell\right), \mathcal{I}\left(i_{\ell+1}, j_{\ell+1}\right)\right)$:

1. The ordering $\boldsymbol{\sigma}^\ell = \left(\sigma_1^\ell, \ldots, \sigma_N^\ell\right)$ of the lines $h_1(\cdot), \ldots, h_N(\cdot)$ from the highest to the smallest values, that is, for all $\lambda \in \Big(\mathcal{I}\left(i_\ell, j_\ell\right)\ ,\ \mathcal{I}\left(i_{t+1}, j_{t+1}\right)\Big)$,

$$h_{\sigma_1^\ell}(\lambda) \geq h_{\sigma_2^\ell}(\lambda) \geq \cdots \geq h_{\sigma_N^\ell}(\lambda) \ .$$

2. The set G^ℓ corresponding to the first C elements according to the ordering $\boldsymbol{\sigma}^\ell$, that is

$$G^\ell = \left\{\sigma_1^\ell, \ldots, \sigma_C^\ell\right\} \ .$$

3. The set B^ℓ of lines whose values have become negative, that is,

$$B^\ell = \left\{i : h_i(\lambda) < 0 \text{ for } \lambda \in \Big(\mathcal{I}\left(i_\ell, j_\ell\right)\ ,\ \mathcal{I}\left(i_{t+1}, j_{t+1}\right)\Big)\right\} \ .$$

   Since the lines $h_1(\cdot), \ldots, h_N(\cdot)$ are strictly decreasing, $B^\ell \subseteq B^{\ell+1}$ for all ℓ.
4. The assortment $A^\ell = G^\ell \setminus B^\ell$.

The formal description of the StaticMNL policy is given as follows.

<u>StaticMNL</u>

Inputs: The number of intersection points K, the ordering $\boldsymbol{\tau} = \left(\left(i_1, j_1\right), \ldots, \left(i_K, j_K\right)\right)$ of the intersection points, and the ordering $\boldsymbol{\sigma}^0$ of the preference vector $\boldsymbol{v}$. Let $A^0 = G^0 = \left\{\sigma_1^0, \ldots, \sigma_C^0\right\}$ and $B^0 = \emptyset$.

Description: For $\ell = 1, 2, \ldots, K$,

- If $i_\ell \neq 0$, let the permutation $\boldsymbol{\sigma}^\ell$ be obtained from $\boldsymbol{\sigma}^{\ell-1}$ by transposing i_ℓ and j_ℓ and set $B^\ell = B^{\ell-1}$.
- If $i_\ell = 0$, let $\boldsymbol{\sigma}^\ell = \boldsymbol{\sigma}^{\ell-1}$ and $B^\ell = B^{\ell-1} \cup \{j_\ell\}$.
- Let $G^\ell = \left\{\sigma_1^\ell, \ldots \sigma_C^\ell\right\}$ and $A^\ell = G^\ell \setminus B^\ell$.

Output: The sequence of assortments $\mathcal{A} = \left\langle A^\ell : \ell = 0, 1, \ldots, K\right\rangle$.

ℓ	$\mathcal{I}(i_\ell, j_\ell)$	$\boldsymbol{\sigma}^\ell$	G^ℓ	B^ℓ	A^ℓ
0	$-\infty$	$(3,2,1)$	$\{3,2\}$	$\emptyset$	$\{3,2\}$
1	$\mathcal{I}(2,3)$	$(2,3,1)$	$\{2,3\}$	$\emptyset$	$\{2,3\}$
2	$\mathcal{I}(1,3)$	$(2,1,3)$	$\{2,1\}$	$\emptyset$	$\{2,1\}$
3	$\mathcal{I}(0,3)$	$(2,1,3)$	$\{2,1\}$	$\{3\}$	$\{2,1\}$
4	$\mathcal{I}(1,2)$	$(1,2,3)$	$\{1,2\}$	$\{3\}$	$\{1,2\}$
5	$\mathcal{I}(0,2)$	$(1,2,3)$	$\{1,2\}$	$\{3,2\}$	$\{1\}$
6	$\mathcal{I}(0,1)$	$(1,2,3)$	$\{1,2\}$	$\{3,2,1\}$	$\emptyset$

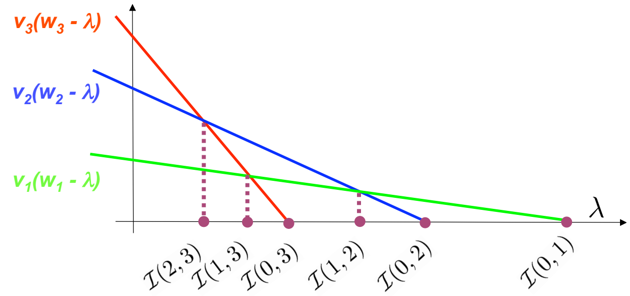

Figure 1: An illustration of the StaticMNL algorithm with three products for Example 2.3.

The main result of this section is stated in the following theorem, which establishes the running time and the correctness of the StaticMNL algorithm. The proof of this result follows immediately from the discussion in the beginning of the section, and the observation that A^ℓ correspond to $A(\lambda)$ (defined in Equation (4)) for $\lambda \in \left[\mathcal{I}\left(i_\ell, j_\ell\right), \mathcal{I}\left(i_{\ell+1}, j_{\ell+1}\right)\right)$.

Theorem 2.2 (Solution to **Capacitated MNL**)**.** *The running time of the* StaticMNL *algorithm is of order* $O(N^2)$*, and* $Z^* = \max\left\{f\left(A(\lambda)\right) : \lambda \in \mathbb{R}\right\} = \max\left\{f\left(A^\ell\right) : \ell = 0, 1, \ldots, K\right\}$.

The following example illustrates an application of the StaticMNL algorithm to a problem with three products.

Example 2.3 (Application of StaticMNL)**.** Consider an example with 3 products whose corresponding lines $h_1(\cdot)$, $h_2(\cdot)$, and $h_3(\cdot)$ are shown in Figure 1. In this example, the products are ordered so that $v_1 < v_2 < v_3$ and we have six intersection points with

$$\mathcal{I}(2,3) < \mathcal{I}(1,3) < \mathcal{I}(0,3) < \mathcal{I}(1,2) < \mathcal{I}(0,2) < \mathcal{I}(0,1)$$

Suppose that we have a capacity constraint $C = 2$. The output of the StaticMNL algorithm is given in the table in Figure 1. We note that, for each ℓ, the assortment A^ℓ corresponds to the top two lines whose values are nonnegative within the interval $\left(\mathcal{I}\left(i_\ell, j_\ell\right), \mathcal{I}\left(i_{\ell+1}, j_{\ell+1}\right)\right)$.

The following result follows immediately from the description of the StaticMNL algorithm.

Theorem 2.4 (Sensitivity of the Optimal Assortment)**.** *The output of the* StaticMNL *algorithm depends only on the ordering* $\boldsymbol{\sigma}^0$ *of the customer preference weights* v_i *and the ordering* $\boldsymbol{\tau}$ *of the intersection points* $\mathcal{I}(i,j)$.

Theorem 2.4 has the following important implication for the dynamic optimization problem that we will consider in Section 3, where the customer preference vector $\boldsymbol{v}$ is not known. Suppose that we have an approximation $\widehat{\boldsymbol{v}} = (\widehat{v}_1, \ldots, \widehat{v}_N) \in \mathbb{R}_+^N$ of the customer preference weights. Given

$\widehat{\boldsymbol{v}}$, we can estimate the intersection points $\widehat{\mathcal{I}}(i_k, j_k)$. We can then use the ordering of $\widehat{\boldsymbol{v}}$ and the ordering of the estimated intersection points as inputs to the StaticMNL algorithm. The above theorem tells us that as long as the estimated orderings coincide with the true orderings, then the output of the StaticMNL will be *exactly the same* as if we know the true value of $\boldsymbol{v}$.

2.2 Properties of $\mathcal{A}$

In the next two sections, we exploit the geometry associated with the MNL model to derive structural properties of the sequence of assortments $\mathcal{A} = \langle A^0, A^1, \ldots, A^K \rangle$ generated by the StaticMNL algorithm. Before we proceed, let us introduce the following assumption that will be used throughout the rest of the paper. We emphasize that Assumption 2.1 is introduced primarily to simplify our exposition and to facilitate the discussion of the key ideas without having to worry about degenerate cases.

Assumption 2.1 (Distinct Customer Preferences and Intersection Points)**.**

(a) The products are indexed so that $0 < v_1 < v_2 < \cdots < v_N$.

(b) The intersection points are distinct, that is, for any $(i,j) \neq (s,t)$, $\mathcal{I}(i,j) \neq \mathcal{I}(s,t)$.

Assumption 2.1(a) requires the values of v_i to be distinct, while Assumption 2.1(b) requires that the marginal profit w_i are distinct, and no three lines among $h_1(\cdot), \ldots, h_N(\cdot)$ can intersect at the same point. As a consequence of Assumption 2.1, we observe that every pair of lines $h_i(\cdot)$ and $h_j(\cdot)$ will intersect each other, and thus, the number of intersection points K is exactly $\binom{N+1}{2}$. In addition, we also have a strict ordering of the intersection points, that is,

$$-\infty \equiv \mathcal{I}(i_0, j_0) < \mathcal{I}(i_1, j_1) < \mathcal{I}(i_2, j_2) < \cdots < \mathcal{I}(i_K, j_K) < \mathcal{I}(i_{K+1}, j_{K+1}) \equiv +\infty \ .$$

The main result of this section is stated in the following theorem. The proof of this result is given in Appendix A.

Theorem 2.5 (Properties of $\mathcal{A}$)**.** *Under Assumption 2.1,*

(a) For each $\ell = 1, \ldots, K$, *if* $A^\ell \neq A^{\ell-1}$, *then*

$$A^\ell = \begin{cases} \left(A^{\ell-1} \setminus \{j_\ell\}\right) \cup \{i_\ell\}, & \text{if } \left|A^\ell\right| = C \ , \\ \left(A^{\ell-1} \setminus \{j_\ell\}\right), & \text{if } \left|A^\ell\right| < C \ , \end{cases}$$

(b) For any $s < C$, *there is exactly one distinct assortment of size* s *in the sequence* $\mathcal{A}$.

(c) There are at most $C(N - C + 1)$ *distinct non-empty assortments in the sequence* $\mathcal{A}$.

Let us briefly describe the intuition behind the proof the above result. Since the slope of the lines $h_i(\cdot)$ are negative, we can establish a partial ordering among the intersection points of any three lines (Lemma A.1). This relationship enables us to show that the ordering $\boldsymbol{\sigma}^\ell$ is obtained from $\boldsymbol{\sigma}^{\ell-1}$ by a transposition of two adjacent products (to be defined precisely in Lemma A.2). These two results then allow us to establish Theorem 2.5.

The above theorem shows that each assortment in the sequence $\mathcal{A}$ is obtained by either interchanging a pair of products or removing a product. Moreover, if the capacity C is fixed, finding the optimal assortment requires us to search only through $O(N)$ assortments, which is on the same order as the uncapacitated optimization problem (see Gallego et al., 2004; Liu and van Ryzin, 2008). Given the result of Theorem 2.5, throughout the rest of the paper, we will assume that the assortments in the sequence $\mathcal{A}$ are distinct.

2.3 Unimodality of $\mathcal{A}$ and Application to Sampling-based Golden Ratio Search

In this section, we will show that the sequence of profits $\left\langle f(A^i) : i = 0, 1, \ldots, K \right\rangle$ is unimodal. To facilitate our discussion, let $\beta \in (0,1)$ be defined by

$$\beta = \frac{\min\left\{\min_i v_i \ , \ \min_{i \neq j} |v_i - v_j| \ , \ \min_{(i,j) \neq (s,t)} |\mathcal{I}(i,j) - \mathcal{I}(s,t)|\right\}}{(1 + C \max_i v_i)} \ . \qquad (8)$$

For any assortment S of size C or less and $\{i, j\} \subseteq S$, we have that

$$|\theta_i(S) - \theta_j(S)| = \frac{|v_i - v_j|}{1 + \sum_{k \in S} v_k} \ \geq \ \beta \ ,$$

and thus, the parameter β is a lower bound on the difference between the selection probabilities of any two products. Under Assumption 2.1, the parameter β is always positive.

Theorem 2.6 (Unimodality of the Profit Function Over $\mathcal{A}$)**.** *Under Assumptions 2.1, if the assortments in the sequence $\mathcal{A}$ are distinct, then there exists $q \in \{0, 1, \ldots, K\}$ such that*

$$f\left(A^0\right) < f\left(A^1\right) < \cdots < f\left(A^{q-2}\right) < f\left(A^{q-1}\right) \leq f(A^q) \quad \textit{and} \quad f\left(A^q\right) > f\left(A^{q+1}\right) > \cdots > f\left(A^K\right) \ ,$$

and for each $\ell \notin \{q, q+1\}$, $\left|f(A^\ell) - f(A^{\ell-1})\right| \geq \beta^2$.

Proof. Consider an arbitrary ℓ. There are two cases to consider: $\left|A^\ell\right| = C$ and $\left|A^\ell\right| < C$. Suppose that $\left|A^\ell\right| = C$. It follows from Theorem 2.5(a) that $A^\ell = \left(A^{\ell-1} \setminus \{j_\ell\}\right) \cup \{i_\ell\}$ with $1 \leq i_\ell < j_\ell$. Let

$X = A^{\ell-1} \setminus \{i_\ell, j_\ell\}$. Then, we have

$$\begin{aligned}
f\left(A^{\ell}\right) - f\left(A^{\ell-1}\right) &= \frac{\sum_{k\in X} w_k v_k + w_{i_\ell} v_{i_\ell}}{1 + \sum_{k\in X} v_k + v_{i_\ell}} - \frac{\sum_{k\in X} w_k v_k + w_{j_\ell} v_{j_\ell}}{1 + \sum_{k\in X} v_k + v_{j_\ell}} \\
&= \frac{\left(1 + \sum_{k\in X} v_k\right)\left(w_{i_\ell} v_{i_\ell} - w_{j_\ell} v_{j_\ell}\right) - \left(\sum_{k\in X} w_k v_k\right)\left(v_{i_\ell} - v_{j_\ell}\right) + \left(w_{i_\ell} - w_{j_\ell}\right) v_{i_\ell} v_{j_\ell}}{\left(1 + \sum_{k\in X} v_k + v_{i_\ell}\right)\left(1 + \sum_{k\in X} v_k + v_{j_\ell}\right)} \\
&= \frac{\left(v_{j_\ell} - v_{i_\ell}\right) \cdot \left\{ -\left(1 + \sum_{k\in X} v_k\right) \frac{\left(w_{i_\ell} v_{i_\ell} - w_{j_\ell} v_{j_\ell}\right)}{v_{i_\ell} - v_{j_\ell}} + \sum_{k\in X} w_k v_k - \frac{\left(w_{i_\ell} - w_{j_\ell}\right) v_{i_\ell} v_{j_\ell}}{v_{i_\ell} - v_{j_\ell}} \right\}}{\left(1 + \sum_{k\in X} v_k + v_{i_\ell}\right)\left(1 + \sum_{k\in X} v_k + v_{j_\ell}\right)} \\
&= \frac{\left(v_{j_\ell} - v_{i_\ell}\right) \cdot \left\{ -\left(1 + \sum_{k\in X} v_k\right) \mathcal{I}\left(i_\ell, j_\ell\right) + \left(\sum_{k\in X} w_k v_k\right) + h_{i_\ell}\left(\mathcal{I}\left(i_\ell, j_\ell\right)\right) \right\}}{\left(1 + \sum_{k\in X} v_k + v_{i_\ell}\right)\left(1 + \sum_{k\in X} v_k + v_{j_\ell}\right)} ,
\end{aligned}$$

where the last equality follows from the fact that

$$\mathcal{I}\left(i_\ell, j_\ell\right) = \frac{w_{i_\ell} v_{i_\ell} - w_{j_\ell} v_{j_\ell}}{v_{i_\ell} - v_{j_\ell}} \quad \text{and} \quad h_{i_\ell}\left(\mathcal{I}\left(i_\ell, j_\ell\right)\right) = \frac{-\left(w_{i_\ell} - w_{j_\ell}\right) v_{i_\ell} v_{j_\ell}}{v_{i_\ell} - v_{j_\ell}} .$$

Since $h_{i_\ell}\left(\mathcal{I}\left(i_\ell, j_\ell\right)\right) = v_{i_\ell}\left(w_{i_\ell} - \mathcal{I}\left(i_\ell, j_\ell\right)\right)$ and $A^\ell = X \cup \{i_\ell\}$, we have that

$$\begin{aligned}
f\left(A^{\ell}\right) - f\left(A^{\ell-1}\right) &= \frac{\left(v_{j_\ell} - v_{i_\ell}\right) \cdot \left\{ -\mathcal{I}(i_\ell, j_\ell) + \sum_{k\in A^\ell} v_k \left(w_k - \mathcal{I}\left(i_\ell, j_\ell\right)\right) \right\}}{\left(1 + \sum_{k\in X} v_k + v_{i_\ell}\right)\left(1 + \sum_{k\in X} v_k + v_{j_\ell}\right)} \\
&= \frac{\left(v_{j_\ell} - v_{i_\ell}\right) \cdot \left\{ g\left(\mathcal{I}(i_\ell, j_\ell)\right) - \mathcal{I}(i_\ell, j_\ell) \right\}}{\left(1 + \sum_{k\in X} v_k + v_{i_\ell}\right)\left(1 + \sum_{k\in X} v_k + v_{j_\ell}\right)} ,
\end{aligned}$$

where the last equality follows from the definition of $g(\cdot)$ defined in Equation (4), which shows that $g(\lambda) = \sum_{k\in A^\ell} v_k(w_k - \lambda)$ for $\mathcal{I}(i_\ell, j_\ell) \le \lambda < \mathcal{I}(i_{\ell+1}, j_{\ell+1})$. In the case where $\left|A^\ell\right| < C$, we have that $A^\ell = A^{\ell-1} \setminus \{j_\ell\}$ by Theorem 2.5(a). Using the same argument as above, we can show that

$$f\left(A^{\ell}\right) - f\left(A^{\ell-1}\right) = \frac{v_{j_\ell} \cdot \left\{ g\left(\mathcal{I}(i_\ell, j_\ell)\right) - \mathcal{I}(i_\ell, j_\ell) \right\}}{\left(1 + \sum_{\ell\in X} v_\ell\right)\left(1 + \sum_{\ell\in X} v_\ell + v_{j_\ell}\right)} .$$

In both cases, the denominator in the expression $f\left(A^\ell\right) - f\left(A^{\ell-1}\right)$ is always positive. Also, since $i_\ell < j_\ell$, it follows from Assumption 2.1 that $v_{j_\ell} - v_{i_\ell}$ is always positive. From the definition, $g(\cdot)$ is a continuous, piecewise linear, non-increasing, convex function, and $g(Z^*) = Z^*$. Thus, $g(\lambda) - \lambda > 0$ for all $\lambda < Z^*$ and $g(\lambda) - \lambda > 0$ for all $\lambda > Z^*$. Let $q \in \{0, 1, \ldots, K\}$ denote the largest index such that $\mathcal{I}\left(i_q, j_q\right) \le Z^*$. It follows from Assumption 2.1(b) that

$$\mathcal{I}\left(i_1, j_1\right) < \cdots < \mathcal{I}\left(i_{q-1}, j_{q-1}\right) < \mathcal{I}\left(i_q, j_q\right) \le Z^* < \mathcal{I}\left(i_{q+1}, j_{q+1}\right) < \cdots < \mathcal{I}\left(i_K, j_K\right) .$$

Then, it follows that $f\left(A^0\right) < f\left(A^1\right) < \cdots < f\left(A^{q-1}\right) \le f(A^q)$ and $f\left(A^q\right) > f\left(A^{q+1}\right) > \cdots > f\left(A^K\right)$, which is the desired result.

The function $g(\lambda) - \lambda$ is a piecewise linear, strictly decreasing, and convex function which is zero at Z^*, and the absolute value of its subgradient is bounded below by one. Thus, it is easy to verify that for all $\lambda \in \mathbb{R}$, $|g(\lambda) - \lambda| \geq |\lambda - Z^*|$. Thus, if $\ell \notin \{q, q+1\}$, we have that $|g(\mathcal{I}(i_\ell, j_\ell)) - \mathcal{I}(i_\ell, j_\ell)| \geq \min_{(k,m)\neq(s,t)} |\mathcal{I}(k,m) - \mathcal{I}(s,t)|$. Therefore, using the expression for $f(A^\ell) - f(A^{\ell-1})$, we conclude that $\left|f\left(A^\ell\right) - f\left(A^{\ell-1}\right)\right| \geq \beta^2$, which completes the proof. □

Given the sequence of assortments $\mathcal{A}$, if we can evaluate the profit function $f(\cdot)$, it follows from Theorem 2.6 that we can apply the standard Golden Ratio Search (see, for example, Press et al., 1999) to find the optimal assortment in $O(\log(NC))$ iterations. It turns out that the unimodality structure of the sequence of profits can also be exploited to yield an efficient search algorithm in the dynamic optimization problem, where the preference vector $\boldsymbol{v}$ is unknown. In this case, for each assortment A^ℓ in the sequence $\mathcal{A}$, we can offer it to a sample of customers and compute the average profit from the resulting sales. We know from Theorem 2.6 that there is a gap of β^2 in the difference between the expected profit of two consecutive assortments. This suggests that, if the number of customers is sufficiently large, we can use the average profit as a proxy for $f(A^\ell)$, and apply the standard golden ratio search procedure. This idea is the basis of the sampling-based golden ratio search described below.

Sampling-Based Golden Ratio Search (Sampling GRS)

Input: A sequence $\mathcal{A}$ of assortments and a time horizon T.

Algorithm Description: We perform the standard Golden Ratio search. Whenever we need to compare the values of two assortments A^{ℓ_1} and A^{ℓ_2} in the sequence $\mathcal{A}$, we check to see if each assortment has been offered to at least $\lceil 2(\log T)/\beta^4 \rceil$ independent customers. If not, then offer each of these assortments to the customers until we have at least $\lceil 2(\log T)/\beta^4 \rceil$ observations for each assortment. If we have enough data, compare the two assortments based on the average profits obtained and proceed as in the classical Golden Ratio Search algorithm. At the end of the Golden Ratio Search, we are left with a single assortment, offer that assortment until the end of the horizon[3].

The idea of using the sample average as an approximation of the true expectation has been applied in many applications (see, for example, Shapiro, 2003; Swamy and Shmoys, 2005; Levi et al., 2007). We present the algorithm and its analysis primarily for the sake of completeness. We will use this algorithm in the next section when we present an adaptive policy for generating a sequence of assortments. The following lemma establishes a performance bound associated with our sampling-based golden ratio search; the proof appears in Appendix B.

[3]Instead of waiting until we are left with a single assortment, we can terminate the search procedure when we are left with a few assortments, say four or five. Then, we can apply the standard multiarmed bandit algorithm (Lai and Robbins, 1985; Auer et al., 2002). The analysis is essentially the same.

Lemma 2.7 (Regret for Sampling-based GRS)**.** *Suppose that the sequence $\mathcal{A}$ is given, but the customer preference vector $\boldsymbol{v}$ is unknown. Then, there exists a positive constant a_1 that depends only on C, $\boldsymbol{v}$, and $\boldsymbol{w}$ such that, for any $T \geq 1$, the regret under the* Sampling-based Golden Ratio Search *is bounded above by*

$$\text{Regret}\left(T, \text{Sampling GRS}\right) \leq a_1 (\log N) \log T \ .$$

We note that by exploiting the unimodality of the sequence of profits and the gap between the expected profit of any two consecutive assortments, we obtain a regret bound that scales with $\log N$, instead of N under the traditional bandit algorithms that tries every assortment in the sequence $\mathcal{A}$. The Sampling-based GRS algorithm, however, requires a prior knowledge of the gap β. Since we have an explicit formula for β in Equation (8), we can potentially estimate its value from our estimates of $\boldsymbol{v}$ (more on this in the next section).

3. Dynamic Optimization

In this section, we address the dynamic optimization problem, where the preference vector $\boldsymbol{v} \in \mathbb{R}_+^N$ is unknown and must be estimated from past sales and assortment decisions. It follows from Theorem 2.4 that the ordering of $\boldsymbol{v}$ and the ordering of the intersection points among the lines $h_0(\cdot), \ldots, h_N(\cdot)$ *completely determine* the outputs of the StaticMNL algorithm. So, instead of estimating the actual values of $\boldsymbol{v}$, it suffices to estimate the orderings. There is a simple relationship between the selection probabilities and the orderings of the intersection points and the customer preference weights. It follows from the definition of the intersection point $\mathcal{I}(i,j)$ that for any assortment S containing both i and j,

$$\mathcal{I}(i,j) = \frac{w_i v_i - w_j v_j}{v_i - v_j} = \frac{w_i \theta_i\left(S\right) - w_j \theta_j\left(S\right)}{\theta_i\left(S\right) - \theta_j\left(S\right)} \ , \tag{9}$$

and $v_i \leq v_j$ if and only if $\theta_i\left(S\right) \leq \theta_j\left(S\right)$.

We can estimate the selection probabilities by counting the number of customers who select a particular product from an assortment. This idea forms the basis of our proposed policy for the dynamic optimization. To facilitate our exposition, let $\mathcal{E}$ denote the collection of subsets of size C that "cover" all pairs of products, that is, for all i and j, there exists an assortment $S \in \mathcal{E}$ with $|S| = C$ and $\{i,j\} \subseteq S$. It is easy to verify that $|\mathcal{E}| \leq 5(N/C)^2$. Here is an example of $\mathcal{E}$.

Example 3.1. Suppose that $N = 6$ and $C = 3$. We can define $\mathcal{E}$ as follows:

$$\mathcal{E} = \{\{1,2,3\},\{4,5,6\},\{1,2,4\},\{1,2,5\},\{1,2,6\},\{1,3,4\},\{3,5,6\}\} \, .$$

It can be observed that, for all i and j, there exists $S \in \mathcal{E}$ such that $\{i,j\} \subseteq S$.

Our policy, which we refer to as Adaptive Assortment (AA), operates in cycles. Each cycle $m \geq 1$ consists of an exploration phase, followed by an exploitation phase. In the exploration phase of cycle m, we offer each assortment $S \in \mathcal{E}$ to a single customer and observe her selection. At the end of the exploration phase, for each assortment $S \in \mathcal{E}$ and $i \in S$, we estimate the selection probability $\widehat{\Theta}_i(m, S)$ based on the fraction of customers who selected product i during the past m cycles. We can then estimate of the ordering of $\boldsymbol{v}$ based on the ordering of $\widehat{\Theta}_i(m, S)$. Also, for any $\{i, j\} \subseteq S$, we can estimate the intersection point between the lines $h_i(\cdot)$ and $h_j(\cdot)$ based on $\widehat{\Theta}_i(m, S)$ and $\widehat{\Theta}_j(m, S)$. This gives us an estimated ordering of the intersection points. Using these *estimated* orderings as inputs to the StaticMNL algorithm, we obtain as an output a sequence of assortments $\widehat{\mathcal{A}}(m)$. We will show that, with a high probability, the sequence $\widehat{\mathcal{A}}(m)$ *coincides* with the sequence of assortments $\mathcal{A}$ that we would have obtained had we known $\boldsymbol{v}$ a priori. Since the sequence of assortments $\mathcal{A}$ is unimodal under the profit function $f(\cdot)$ by Theorem 2.6, in the exploitation phase of cycle m, we apply the Sampling-based Golden Ratio Search from Section 2.3 for V_m periods, where V_m is the parameter to be determined. This concludes cycle m.

Before we proceed to the formal description of the AA policy, let us highlight the main results of this section. In Theorem 3.2, we establish a large deviation inequality for the estimated selection probabilities and the estimated intersection points, which is used to show that our estimated sequence of assortment $\widehat{\mathcal{A}}(m)$ is correct with high probability (Theorem 3.3). The regret bound is then given in Theorem 3.4, and we conclude this section by pointing out the connection to maximum likelihood estimation and extensions to linear-in-parameters utility model. A formal description of the Adaptive Assortment policy is given below.

Adaptive Assortment (AA)

Parameter: The number of periods V_m associated with the exploitation phase of cycle m.

Description: For each cycle $m \geq 1$, complete the following two phases:

1. **Exploration Phase ($|\mathcal{E}|$ periods):**

   (a) Offer each assortment $S \in \mathcal{E}$ to a single customer. For any $i \in S$, let $\widehat{\Theta}_i(m, S)$ denote the estimated selection probability of product i based on the customers who have been offered assortment S during the exploration phases in the past m cycles, that is,

   $$\widehat{\Theta}_i(m, S) = \frac{1}{m} \sum_{q=1}^{m} \mathbb{1}\left[X(q, S) = i\right] ,$$

   where for any $q \leq m$, $X(q, S)$ denote the selection of the customer in the exploration phase of the q^{th} cycle when she is offered the assortment S.

(b) For each $S \in \mathcal{E}$ and for each $\{i,j\} \subset S$, let $\widehat{\mathcal{I}}_{ij}(m,S)$ denote an estimated intersection point between lines $h_i(\cdot)$ and $h_j(\cdot)$ based on the estimated probabilities $\widehat{\Theta}_i(m,S)$ and $\widehat{\Theta}_j(m,S)$, that is,

$$\widehat{\mathcal{I}}_{ij}(m,S) = \frac{w_i \widehat{\Theta}_i(m,S) - w_j \widehat{\Theta}_j(m,S)}{\widehat{\Theta}_i(m,S) - \widehat{\Theta}_j(m,S)} \ ,$$

provided that $\widehat{\Theta}_i(m,S) \neq \widehat{\Theta}_j(m,S)$; otherwise, set $\widehat{\mathcal{I}}_{ij}(m,S)$ to some arbitrary number.

(c) For each $i \neq j$, find a set $S_{ij} \in \mathcal{E}$ that contains both i and j, and estimate the pairwise ordering between v_i and v_j using $\widehat{\Theta}_i(m,S_{ij})$ and $\widehat{\Theta}_j(m,S_{ij})$. Let $\widehat{\boldsymbol{\sigma}}(m)$ denote an estimated ordering of the customer preference weights based on the estimated pairwise orderings. Let $\widehat{\boldsymbol{\tau}}(m)$ denote the ordering of the estimated intersection points $\left\{\widehat{\mathcal{I}}_{ij}(m,S_{ij}) : i \neq j\right\}$ from the lowest to the highest value.

(d) Apply the StaticMNL algorithm *using the estimated orderings* $\widehat{\boldsymbol{\sigma}}(m)$ *and* $\widehat{\boldsymbol{\tau}}(m)$ *as inputs.* Let $\widehat{\mathcal{A}}(m)$ denote the sequence of assortments produced by the StaticMNL algorithm.

2. **Exploitation Phase (V_m periods):** Using the sequence $\widehat{\mathcal{A}}(m)$ of assortments as input, apply the Sampling-based Golden Ratio Search for V_m periods.

The following theorem establishes a large deviation inequality associated with our estimated selection probabilities $\Theta_i(m,S)$ and the estimated intersection points $\widehat{\mathcal{I}}_{ij}(m,S)$. The proof of this result is given in Appendix C.

Theorem 3.2 (Large Deviation Inequalities)**.** *Under Assumption 2.1, for each* $0 < \epsilon < 1$ *and* $m \geq 1$*,*

$$\Pr\left\{\max_{S\in\mathcal{E}}\left\{\max_{\{i,j\}\subseteq S: i\neq j}\left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right|, \max_{i\in S}\left|\theta_i(S) - \widehat{\Theta}_i(m,S)\right|\right\} > \epsilon\right\} \leq \frac{10N^2}{C} e^{-m\,\epsilon^2\,\beta^4/72} \ ,$$

where β is defined in Equation (8).

It follows from the definition of β in Equation (8) that for any $(i,j) \neq (s,t)$ and for any $S \supseteq \{i,j\}$,

$$|\mathcal{I}(i,j) - \mathcal{I}(s,t)| \geq \beta \quad \text{and} \quad |\theta_i(S) - \theta_j(S)| \geq \beta \ .$$

Consider the event that

$$\max_{S\in\mathcal{E}}\left\{\max_{\{i,j\}\subseteq S: i\neq j}\left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right|, \max_{i\in S}\left|\theta_i(S) - \widehat{\Theta}_i(m,S)\right|\right\} \leq \beta/2 \ .$$

According to Theorem 3.2, this event happens with probability of at least $1 - O(e^{-m\beta^6})$. When this event happens, it follows that the ordering $\widehat{\boldsymbol{\tau}}(m)$ of the intersection points based on $\widehat{\mathcal{I}}_{ij}(m,S)$ and the ordering $\widehat{\boldsymbol{\sigma}}(m)$ of the preference weights based on the estimated selection probabilities $\Theta_i(m,S)$ will coincide with the true orderings $\boldsymbol{\tau}$ and $\boldsymbol{\sigma}^0$ defined in Equations (6) and (7), respectively. From Thereom 2.4, we know that when this happens, we are guaranteed that the outputs of the StaticMNL algorithm – using the estimated orderings as inputs – will be *exactly the same* as if we had known $\boldsymbol{\tau}$ and $\boldsymbol{\sigma}^0$. This result is summarized in the following theorem.

Theorem 3.3 (Accuracy of Estimated Assortments)**.** *For each $m \geq 1$,*

$$\Pr\left\{\widehat{\mathcal{A}}(m) = \mathcal{A}\right\} \geq \Pr\left\{\widehat{\boldsymbol{\sigma}}(m) = \boldsymbol{\sigma}^0 \ \text{ and } \ \widehat{\boldsymbol{\tau}}(m) = \boldsymbol{\tau}\right\} \geq 1 - \frac{10N^2}{C} e^{-m\,\beta^6/288} \ .$$

The main result of this section is stated in the following theorem that gives a bound on the cumulative regret under the Adaptive Assortment policy. The result follows directly from Theorems 3.2 and 3.3, and the observation that when $\widehat{\mathcal{A}}(m) = \mathcal{A}$, the regret under Sampling-based Golden Ratio Search increases logarithmically over time by Lemma 2.7. The detail is given in Appendix D.

Theorem 3.4 (Regret Bound for Dynamic Optimization)**.** *For any $\alpha < \beta^6/288$, if $V_m = \lfloor e^{\alpha m} \rfloor$ for all m, then there exists a positive constant a_2 that depends only on C, $\boldsymbol{v}$, $\boldsymbol{w}$, and α such that for any $T \geq 1$,*

$$\text{Regret}(T, \text{AA}) \leq a_2 N^2 \log^2 T \ .$$

Connection to Maximum Likelihood Estimate and Extension to Linear-in-Parameters Utilities: We conclude this section by discussing the connection between $\widehat{\Theta}_i(m,S)$ and the maximum likelihood estimate. Consider an assortment S that was offered to the customers in the exploration phases of the past m cycles, and for each $i \in \{0\} \cup S$, let $N_i(m,S)$ denote the number of customers who have selected product i. Note that $\widehat{\Theta}_i(m,S) = N_i(m,S)/m$. Then, the maximum likelihood estimate $\widehat{\boldsymbol{\mu}}(m,S) = (\widehat{\mu}_i(m,S) : i \in S)$ is given by

$$\begin{aligned}
\widehat{\boldsymbol{\mu}}(m,S) &= \arg \max_{(u_i : i \in S)} \sum_{\ell \in S \cup \{0\}} N_\ell(m,S) \log\left(\frac{e^{u_\ell}}{1 + \sum_{k \in S} e^{u_k}}\right) \\
&= \arg \max_{(u_i : i \in S)} - \sum_{\ell \in S \cup \{0\}} \frac{N_\ell(m,S)}{m} \log\left(\frac{N_\ell(m,S)/m}{e^{u_\ell} \Big/ \left(1 + \sum_{k \in S} e^{u_k}\right)}\right) \\
&= \arg \min_{(u_i : i \in S)} \mathcal{KL}\left(\left(\widehat{\Theta}_\ell(m,S) : \ell \in \{0\} \cup S\right) \ \Big|\Big| \ \left(\frac{e^{u_\ell}}{1 + \sum_{k \in S} e^{u_k}} : \ell \in \{0\} \cup S\right)\right) ,
\end{aligned}$$

where for any two probability distributions $(p_1, \ldots, p_k) \in \mathbb{R}_+^k$ and $(q_1, \ldots, q_k) \in \mathbb{R}_+^k$ with $\sum_{\ell=1}^k p_\ell = \sum_{\ell=1}^k q_\ell = 1$, $\mathcal{KL}\left((p_1, \ldots, p_k) \ \big|\big| \ (q_1, \ldots, q_k)\right) = \sum_{\ell=1}^k p_\ell \log \frac{p_\ell}{q_\ell}$ denotes the KL-divergence between the two distributions. Using the standard property of the KL-divergence (Cover and Thomas, 2006), for all $\ell \in \{0\} \cup S$, we have that[4]

$$\widehat{\Theta}_\ell(m,S) = \frac{e^{\widehat{\mu}_\ell(m,S)}}{1 + \sum_{k \in S} e^{\widehat{\mu}_k(m,S)}} \ .$$

The above relationship between the estimated selection probabilities and the maximum likelihood estimate allows us to extend our model to the setting where the mean utilities $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N)$ are a linear combination of the features associated with product i, that is, for $i = 1, 2, \ldots, N$,

$$\mu_i = \sum_{\ell=1}^{F} \alpha_\ell \phi_{i,\ell},$$

where for each $i = 1, \ldots, N$, $\boldsymbol{\phi}_i = (\phi_{i,1}, \ldots, \phi_{i,F}) \in \mathbb{R}^F$ is an F-dimensional vector that represents the features of the i^{th} product. Examples of product features might include its price, customer reviews, or brands. We assume the feature vectors $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_N$ are known in advance, and we only need to estimate the coefficients $\alpha_1, \ldots, \alpha_F$. When $F \leq C$, instead of estimating the selection probabilities, we can estimate the coefficients $\alpha_1, \ldots, \alpha_F$ directly, and use them to compute the maximum likelihood estimate $\widehat{\boldsymbol{\mu}}(m,S)$, which will give rise to the estimated selection probabilities.

4. Numerical Experiments

In this section, we report the results of our numerical experiments. In the next section, we describe our motivation, the dataset, and our model of the mean utilities. We then consider the static optimization problem, and compare the optimal assortment under the StaticMNL algorithm with the assortment computed under other policies. Then, in Section 4.2, we assume the mean utilities are unknown and apply the Adaptive Assortment algorithm from the previous section.

4.1 Dataset, Model, and Static Optimization

Before we can evaluate the performance of both the StaticMNL and Adaptive Assortment algorithms, we need to identify a set of products and specify their mean utilities. To help us understand the range of utility values that we might encounter in actual applications, we estimate the utilities using data on DVD sales at a large online retailer. We consider DVDs that are sold during a three-month period from July 1, 2005 through September 30, 2005. During this period, the retailer sold over 4.3 million DVDs, spanning across 51,764 DVD titles.

To simplify our analysis, we restrict our attention to customers who have purchased DVDs that account for the top 33% of the total sales, and assume that each customer purchases at most one

[4]An alternative proof of this result is given in Theorem 1 in Vulcano et al. (2008).

DVD. This gives us total of $1,409,261$ customers in our dataset. The products correspond to the 200 best-selling DVDs that account for about 65% of the total sales among our customers. We assume that all 200 DVDs are available for purchase, and when customers do not purchase these DVDs, we assign them to the no-purchase alternative. We observe that the best-selling selling DVD in our dataset was purchased by only about 2.6% of the customers. In fact, among the top 10 best-selling DVDs, each one was sold to only around 1.1% - 2.6% of the customers. Thus, only a small fraction of the customers purchased each DVD.

We assume a linear-in-parameters utility model described in Section 3. The attributes of each DVD that we consider are the selling price (averaged over 3 months of data), customer reviews, total votes received by the reviews, running time, and the number of discs in the DVD collection. We obtain data on customer reviews and the number of discs of each DVD from Amazon.com web site through a publicly available interface via Amazon.com E-Commerce Services (`http://aws.amazon.com`). Each visitor at the Amazon.com web site can provide a review and a rating for each DVD. The rating is on a scale of 1 to 5, with 5 representing the most favorable review. Each review can be voted by other visitors as either "helpful" or "not helpful". For each DVD, we consider all reviews up until June 30, 2005, and compute features such as the average rating, the proportion of reviews that give a 5 rating, and the average number of helpful votes received by each review, and so on.

Under the linear-in-parameters utility model, for $i \in \{1, 2, \ldots, 200\}$, the mean utility μ_i of DVD i is given by $\mu_i = \alpha_0 + \sum_{k=1}^{F} \alpha_k \phi_{i,k}$, where $(\phi_{i,1}, \ldots, \phi_{i,F})$ denotes the features of DVD i, and $\mu_0 = 0$. The estimated coefficients $\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_F$ are obtained by maximizing the logarithm of the likelihood function, that is,

$$\begin{aligned}(\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_F) \quad = \quad \arg \max_{(u_0, u_1, \ldots, u_F) \in \mathbb{R}^F} & \left\{ N_0 \log \left(\frac{1}{1 + \sum_{\ell=1}^{200} e^{u_0 + \sum_{k=1}^{F} u_k \phi_{\ell,k}}} \right) \right. \\ & \left. + \sum_{i=1}^{200} N_i \log \left(\frac{e^{u_0 + \sum_{k=1}^{F} u_k \phi_{i,k}}}{1 + \sum_{\ell=1}^{200} e^{u_0 + \sum_{k=1}^{F} u_k \phi_{\ell,k}}} \right) \right\},\end{aligned}$$

where N_i denotes the number of customers who purchased DVD i, and N_0 denotes the number of customers who did not purchase any of the 200 DVDs in our dataset.

We use the software BIOGEME developed by Bierlaire (2003) to determine the most relevant DVD features and estimate the corresponding coefficients. It turns out that the two most relevant attributes are the total number of votes received by the reviews of each DVD and the *price per disc* (computed as the selling price divided by the number of discs in the DVD collection). We estimate that for each DVD $i = 1, \ldots, 200$,

$$\mu_i = -4.31 + \left(3.50 \times 10^{-5} \times \phi_{i,1}\right) - \left(0.038 \times \phi_{i,2}\right) \ , \tag{10}$$

where

$$\begin{aligned} \phi_{i,1} &= \text{Total Number of Votes Received by All Reviews of DVD } i, \\ \phi_{i,2} &= \text{Price Per Disc Associated with DVD } i \end{aligned}$$

All estimated coefficients (-4.31, 3.50×10^{-5}, and -0.038) are statistically significant with p-values of 0.00, 0.04, and 0.06, respectively. We also checked for any correlation between the product features, and found them to have statistically no correlation. Figure 2 shows the histograms of the price per disc associated with each DVD. We observe that for over half of the DVDs, the price per disc is between \$6 and \$12.

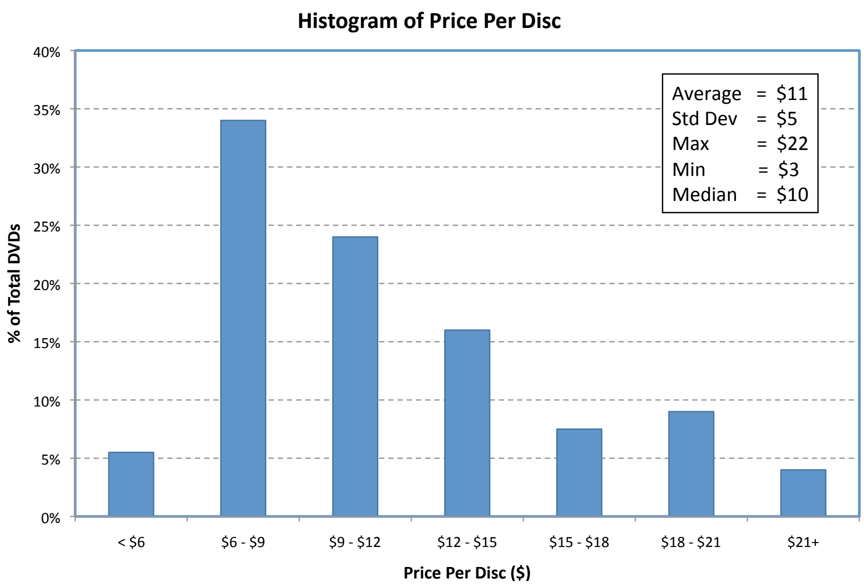

Figure 2: Histogram of the price per disc associated with the 200 best-selling DVDs.

In our model selection, we have also considered the selling price of each DVD as a candidate feature, but the selling price turns out *not* to be a statistically significant feature in explaining customer purchase patterns in the data. We believe this happens because many DVDs in our dataset are multi-disc collections. For example, the most popular DVD in the dataset is the 7-disc collection of *Lost - The Complete First Season*. Although the selling price of this DVD is about \$39 (which is considerably more expensive than other DVDs), the *price per disc* of this DVD collection is less than \$6, which is significantly smaller than the price per disc of other DVDs in the dataset. We hypothesize that the price per disc is a more appropriate measure of the "true price signal" perceived by the customers. Developing a rigorous justification of this hypothesis remains an open research question.

Figure 3 shows a histogram of the estimated utility across the 200 best-selling DVDs, along with descriptive statistics. The average utility is approximately -4.67, with a maximum of -3.59 and a minimum of -5.13. We note that since the utility of the option of no purchase is set to 0 and the fraction of customers who purchase each DVD is at most 2.6%, the mean utility of each DVD will be negative. We emphasize that our model of the mean utility is quite simplistic and it is unlikely to capture details and factors that affect each customer's purchasing decision. Rather, our goal is to obtain a rough estimate of the range of utilities values that one might encounter in actual applications. Developing a more sophisticated choice model is a subject of ongoing research and is beyond the scope of this paper.

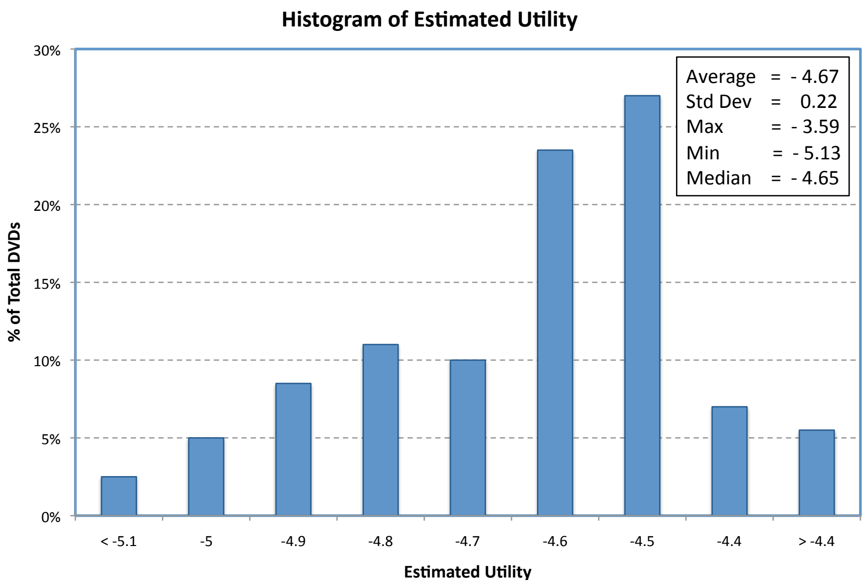

Figure 3: Histogram of the estimated utility based on Equation (10) for the 200 best-selling DVDs.

Static Optimization: Using the estimated utility of each DVD from Equation (10), we compare the expected profit from the optimal assortment under the StaticMNL algorithm and the greedy assortment that only considers the top C most expensive DVDs. In this experiment, we set the marginal profit of each DVD to be its selling price (averaged over the three-month period that we collect the data). The comparison between the assortments is shown in Table 1 when the capacity $C = 10$. We choose this value as an approximation to the number of DVDs that the retailer can display on a web page. From the table, we observe that the optimal assortment yields about 10% = (\$7.35 - \$6.67)/\$6.67 increase in the expected profit. We note that if we consider a set consisting of the top 10 most popular DVDs, the expected profit of this assortment is approximately \$2.68.

We note that the optimal and the greedy assortments contain 8 DVDs in common. The other 2 DVDs in the optimal assortment correspond to *Star Wars Trilogy* and *Firefly (The Complete Series)*. Although their selling prices are less than the 10 most expensive DVDs in the greedy assortment, they are quite popular and have high estimated utilities. The *Star Wars Trilogy* DVD

Optimal Assortment: **Expected Profit = $7.35**

Title	Number of Discs	Estimated Utility	Selling Price ($)
24 - Seasons 1-3	20	-4.51	115.49
The Boston Red Sox 2004 World Series Collector's Edition	12	-4.60	92.03
Star Trek Enterprise - The Complete Second Season	7	-4.79	91.67
Band of Brothers	6	-4.51	79.35
The Lord of the Rings - The Motion Picture Trilogy	12	-4.31	77.94
Six Feet Under - The Complete Fourth Season	5	-4.84	70.12
The Sopranos - The Complete Fifth Season	4	-4.89	64.97
The O.C. - The Complete Second Season	7	-4.55	48.97
Star Wars Trilogy	4	-3.59	45.45
Firefly - The Complete Series	4	-4.01	32.29

Greedy Assortment (Top 10 Most Expensive DVDs): **Expected Profit = $6.67**

Title	Number of Discs	Estimated Utility	Selling Price ($)
24 - Seasons 1-3	20	-4.51	115.49
The Boston Red Sox 2004 World Series Collector's Edition	12	-4.60	92.03
Star Trek Enterprise - The Complete Second Season	7	-4.79	91.67
Band of Brothers	6	-4.51	79.35
The Lord of the Rings - The Motion Picture Trilogy	12	-4.31	77.94
Six Feet Under - The Complete Fourth Season	5	-4.84	70.12
The Sopranos - The Complete Fifth Season	4	-4.89	64.97
Shelley Duvall's Faerie Tale Theatre -The Complete Collection Gift Set	4	-4.76	49.95
The O.C. - The Complete Second Season	7	-4.55	48.97
Thundercats - Season One, Volume One	6	-4.59	46.12

Table 1: The greedy and the optimal assortments when $C = 10$, along with their expected profits.

had over 7,300 reviews and these reviews received over 32,000 votes. This is the highest number of votes received by any DVD in our dataset. Similarly, the *Firefly (Complete Series)* received over 17,000 votes, which is the third highest number of votes among all DVDs. As shown in Table 1, these two DVDs thus have two of the highest estimated utilities (-3.59 and -4.01), contributing to their selection as part of the optimal assortment. Interestingly, the DVD with the second highest number of votes is *What the Bleep Do We Know!?* with about 18,000 votes. However, this DVD has one of the highest price per disc (approximately $18.92), so its estimated utility is quite small, and thus, it does not appear in the optimal assortment.

Table 2 compares the expected profits of the greedy and the optimal assortments as the capacity C increases from 1 to 20. For each value of C, the greedy assortment corresponds to the C most expensive DVDs. For small values of C, we observe that the increases in expected profit can be quite significant. As the capacity C increases, the optimal and the greedy assortments can be significantly different. For $C = 20$, there are 6 DVDs in the optimal assortment which do not appear in the top 20 most expensive DVDs.

Capacity (C)	Expected Profit of the Greedy Assortment ($)	Expected Profit of the Optimal Assortment ($)	Improvement in Profit Relative to the Greedy Assortment (%)	# of DVDs in the Optimal Assortment that are NOT in the Greedy Assortment
1	1.25	1.25	0%	0
2	2.15	2.43	13%	1
3	2.87	3.39	18%	2
4	3.67	4.23	15%	2
5	4.62	5.00	8%	1
6	5.11	5.66	11%	1
7	5.53	6.13	11%	1
8	5.88	6.56	11%	2
9	6.30	6.96	10%	2
10	6.67	7.35	10%	2
11	7.04	7.70	9%	2
12	7.97	8.04	1%	1
13	8.26	8.37	1%	2
14	8.54	8.68	2%	3
15	8.81	8.99	2%	4
16	9.12	9.28	2%	3
17	9.39	9.57	2%	4
18	9.68	9.86	2%	5
19	9.94	10.13	2%	6
20	10.22	10.40	2%	6

Table 2: Expected profits under the greedy and the optimal assortments for different values of the capacity.

4.2 Dynamic Optimization

In this section, we assume that the coefficients in underlying utility model in Equation (10) are *not* known in advance, and we adaptively estimate these coefficients based on past sales and assortment decisions using the Adaptive Assortment (AA) policy described in Section 3 with $C = 10$. It follows from Theorem 3.4 that the T-period regret is bounded above by

$$\text{Regret}(T, \text{AA}) = \sum_{t=1}^{T} \mathsf{E}\left[Z^* - f(S_t)\right] \leq a_2 N^2 \log^2 T \quad \Leftrightarrow \quad \frac{1}{T}\sum_{t=1}^{T} \mathsf{E}\left[f(S_t)\right] \geq Z^* - a_2 N^2 \left(\frac{\log^2 T}{T}\right)$$

where S_t denote the assortment offered in period t under the AA policy. The above lower bound shows the running average expected profit converges to the optimal profit Z^* (which in this case is equal to \$7.35 from Table 1).

The goal of this section is to understand the rate of convergence of the AA policy when applied to our dataset. We apply the AA policy for 100,000 periods and track the running average profit over time. We then repeat this experiment 1,000 times, generating 1,000 independent profit trajectories. For $g = 1, \ldots, 1000$, let Y_t^g denote observed profit in period t of the g^{th} experiment. In Figure 4, we plot the function

$$t \;\rightarrow\; \frac{\frac{1}{1000}\sum_{g=1}^{1000} \frac{1}{t}\sum_{s=1}^{t} Y_s^g}{Z^*}$$

as t ranges from 1 to $100,000$. This represents the running average profit over time (averaged over

1000 experiments) as a fraction of the optimal expected profit Z^*. The dash lines around the solid line represent the 95% confidence interval, which reflects the variability of the running average profit across 1000 experiments. The horizon line corresponds to the expected profit of the greedy assortment relative to Z^*, which is equal to 90.7% = \$6.67/\$7.35 from Table 1. We observe that the running average profit under the Adaptive Assortment policy matches the greedy profit after about 15,000 customers (corresponding to 15,000 periods), representing approximately one day of total sales (since we have $1,409,261$ customers over a three-month period).

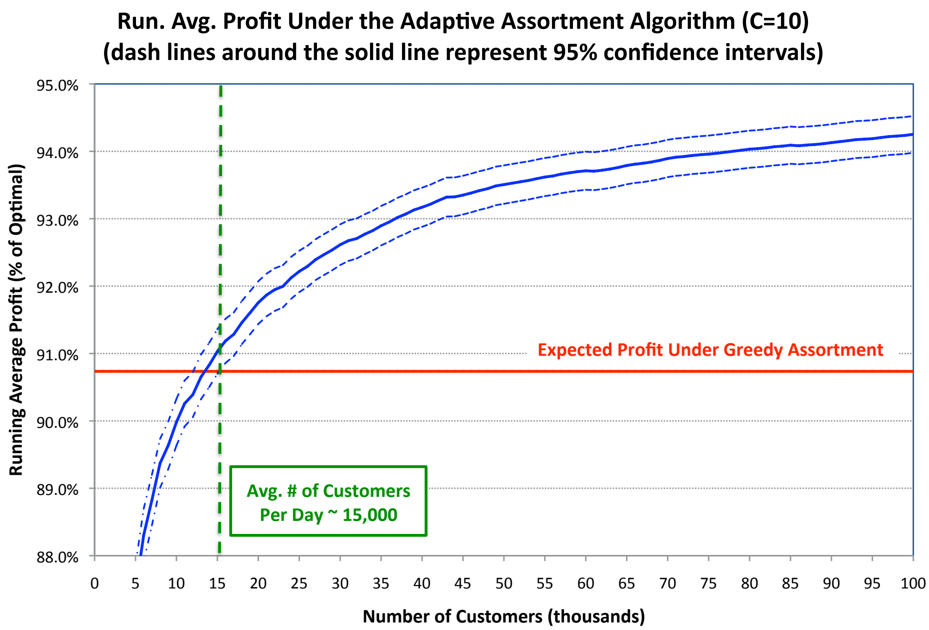

Figure 4: Running average profit under the Adaptive Assortment algorithm.

5. Extensions and Future Work

We describe algorithms for solving the capacitated assortment optimization problem with a multinomial logit choice model in static and dynamic settings. Our work opens up a number of interesting research directions. Our formulation assumes that each customer independently subscribes to the same choice model. This assumption is appropriate when we have a single market and the customers within the market are relatively homogenous. In many cases, the customers are heterogenous and their mean utilities vary across segments. We can model this using the latent-class logit model, and use the maximum likelihood estimation technique to estimate the parameters (see, for example, Bodapati and Gupta, 2004). It would be interesting to establish a large deviation inequality for the maximum likelihood estimate and to develop a regret bound for this class of model.

Although the MNL choice model has been used successfully in many applications, it exhibits the Independence of Irrelevant Alternatives (IIA), where the ratio of the selection probabilities between two alternatives is independent of the assortment containing them. In certain applications, this property may not be consistent with the actual customer choice behavior. Rusmevichientong et al. (2009) has partially extended the solution technique developed in this paper to the nested logit choice model, which is one of the most popular extensions of the MNL model that alleviates the IIA property (McFadden, 1981). Under this model, the products are partitioned into groups. Although products within the same group still exhibit the IIA property, the likelihood of choosing products from two different groups depends on the assortment containing them, providing a potentially more realistic model of customer choice behavior. In this case, the assortment $A(\lambda)$ defined in Equation (4) corresponds to the optimal solution of a sum-of-ratio optimization problem, which is NP-hard and we must resort to approximation methods to compute $A(\lambda)$. We believe that the structural properties developed in this paper can be extended to this model as well. For our DVD dataset, we might apply the nested logit model by grouping the DVDs based on their genre. However, estimating the parameters in the nested logit can be complicated because the likelihood function is not concave and approximation methods are required (see, for example, Silberhorn et al., 2008).

Our model also assumes that the cost of changing an assortment from one customer to the next is negligible. This assumption is reasonable in the online setting where the cost of changing the ads or product recommendations on the web page is minimal. However, in settings where there are significant costs associated with switching product assortments, our model might not be appropriate. In addition, we implicitly assume that we have enough supply of each product to ignore all inventory considerations. Incorporating inventory constraints is an exciting direction for future research.

Acknowledgement

The authors would like to thank Candi Yano for her helpful comments and suggestions on the earlier draft of our manuscript, which inspired us to consider the multinomial logit choice model. We are grateful for all of her help. We would like to thank Mike Todd for insightful comments on linear algebra and for pointers to implementations of the Golden Ratio Search and Newton-Raphson algorithms, Yinyu Ye for helpful discussions on the problem, and Miao Song for all of her help with the BIOGEME software. We also want to thank Costis Maglaras for helpful discussions about the applications of the MNL model during the first author's visit to the Columbia Graduate School of Business. Finally, we are grateful for Garrett van Ryzin, the associated editor, and the three referees for their thoughtful and detailed comments on the earlier version of the manuscripts; their suggestions greatly improve the quality and the presentation of our work. We also want to thank the second referee for encouraging us to improve our estimated logit choice model in Section 4, enabling us to discover the price per disc variable that turns out to be important in explaining the customer purchase patterns. We are grateful for all of his/her suggestions. This research is supported in part by the National Science Foundation through grants DMS-0732196, CMMI-0746844, CMMI-0621433, CMMI-0727640, CCF-0635121, and CCF-0832782.

References

Anantharam, V., P. Varaiya, and J. Walrand. 1987a. Asymptotically efficient adaptive allocation rules for the multi-armed bandit problem with multiple plays, Part I: I.I.D. rewards. *IEEE Transactions on Automatic Control* AC-32 (11): 968–976.

Anantharam, V., P. Varaiya, and J. Walrand. 1987b. Asymptotically efficient adaptive allocation rules for the multi-armed bandit problem with multiple plays, Part II : Markovian rewards. *IEEE Transactions on Automatic Control* AC-32 (11): 977–982.

Anderson, S., A. de Palma, and J. F. Thisse. 1992. *Discrete choice theory of product differentiation*. Cambridge, MA: MIT Press.

Auer, P., N. Cesa-Bianchi, and F. P. 2002. Finite-time analysis of the multiarmed bandit problem. *Machine Learning* 47:235–256.

Ben-Akiva, M., and S. Lerman. 1985. *Discrete choice analysis: Theory and application to travel demand.* Cambridge, MA: MIT Press.

Bierlaire, M. 2003. BIOGEME: A Free Package For the Estimation of Discrete Choice Models. In *Proceedings of the 3rd Swiss Transportation Research Conference*. Ascona, Switzerland.

Bodapati, A. V., and S. Gupta. 2004. The recoverability of segmentation structure from store-level aggregation data. *Journal of Marketing Research* 41:351–364.

Caro, F., and J. Gallien. 2007. Dynamic assortment with demand learning for seasonal consumer goods. *Management Science* 53 (2): 276–292.

Cover, T. M., and J. A. Thomas. 2006. *Elements of information theory.* Wiley.

Gallego, G., G. Iyengar, R. Phillips, and A. Dubey. 2004. Managing flexible products on a network. Working Paper, Columbia University.

Kok, A. G., M. Fisher, and R. Vaidyanathan. 2008. Assortment planning: Review of literature and industry practice. In *Retail Supply Chain Management*: Springer.

Kunnumkal, S., and H. Topaloglu. 2008. A refined deterministic linear program for the network revenue management problem with customer choice behavior. *Naval Research Logistics* 55 (6): 563–580.

Lai, T. L., and H. Robbins. 1985. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics* 6 (1): 4–22.

Levi, R., R. O. Roundy, and D. B. Shmoys. 2007. Provably near-optimal sampling-based algorithms for stochastic inventory control models. *Mathematics of Operations Research* 32 (4): 821–838.

Liu, Q., and G. J. van Ryzin. 2008. On the choice-based linear programming model for network revenue management. *Manufacturing and Service Operations Management* 10 (2): 288–310.

Mahajan, S., and G. J. van Ryzin. 1998. Retail inventories and consumer choice. In *Quantitative Models for Supply Chain Management*, ed. S. Tayur, R. Ganeshan, and M. Magazine, 491–551: Kluwer Academic Publishers.

Mahajan, S., and G. J. van Ryzin. 2001. Stocking retail assortments under dynamic consumer substitution. *Operations Research* 49 (3): 334–351.

McFadden, D. 1981. Econometric models of probabilistic choice. In *Structural Analysis of Discrete Data*: MIT Press.

Megiddo, N. 1979. Combinatorial optimization with rational objective functions. *Mathematics of Operations Research* 4 (4): 414–424.

Press, W. H., S. A. Teukolsky, and W. T. V. et al.. 1999. *Numerical recipes in C, the art of scientific computing*. Cambridge University Press.

Ratliff, R., B. V. Rao, C. P. Narayan, and K. Yellepeddi. 2007. A multi-flight recapture heuristic for estimating unconstrained demand from airline bookings. *Journal of Revenue and Pricing Management* 7 (2): 153–171.

Rusmevichientong, P., Z.-J. M. Shen, and D. B. Shmoys. 2009. A PTAS for capacitated sum-of-ratios optimization. *Operations Research Letters* 37 (4): 230–238.

Saure, D., and A. Zeevi. 2008. Optimal dynamic assortment planning. Working Paper, Columbia Graduate School of Business.

Shapiro, A. 2003. Stochastic programming. In *Handbook in Operations Research and Management Science*. Elsevier.

Shen, Z.-J. M., and X. Su. 2007. Customer behavior modeling in revenue management and auctions: A review and new research opportunities. *Production and Operations Management* 16 (6): 713–728.

Silberhorn, N., Y. Boztug, and L. Hildebrandt. 2008, 653. Estimation with the nested logit model: specifications and software particularities. *OR Spectrum* 30 (4): 635.

Swamy, C., and D. B. Shmoys. 2005. Sampling-based approximation algorithms for multi-stage stochastic optimization. In *Proceedings of the 46th Annual IEEE Symposium on the Foundations of Computer Science*.

Talluri, K., and G. J. van Ryzin. 2004. Revenue management under a general discrete choice model of consumer behavior. *Management Science* 50 (1): 15–33.

Vulcano, G., G. J. van Ryzin, and R. Ratliff. 2008. Estimating primary demand for substitutable products from sales transaction data. Working paper, Stern Business School.

Zhang, D., and D. Adelman. 2008. An approximate dynamic programming approach to network revenue management with customer choice. To appear in *Transportation Science*.

A. Proof of Theorem 2.5

The proof of Theorem 2.5 makes use of the following lemmas. Recall that, under Assumption 2.1, the number of intersection points K is exactly $\binom{N+1}{2}$, and the StaticMNL algorithm maintains the following four pieces of information associated with the interval $\left(\mathcal{I}\left(i_\ell, j_\ell\right), \mathcal{I}\left(i_{\ell+1}, j_{\ell+1}\right)\right)$: 1) the ordering $\boldsymbol{\sigma}^\ell = \left(\sigma_1^\ell, \ldots, \sigma_N^\ell\right)$ of the lines $h_1(\cdot), \ldots, h_N(\cdot)$ from the highest to the smallest values, 2) the set G^ℓ corresponding to the first C elements according to the ordering $\boldsymbol{\sigma}^\ell$, 3) the set B^ℓ of lines whose values have become negative, and 4) the assortment $A^\ell = G^\ell \setminus B^\ell$. The first lemma (whose proof appears in Appendix A.1) provides a simple characterization of the ordering among intersection points among any three lines.

Lemma A.1. *Under Assumption 2.1, for all $0 \leq i < j < k \leq N$, one of the following conditions must hold: either $\mathcal{I}(i,j) < \mathcal{I}(i,k) < \mathcal{I}(j,k)$ or $\mathcal{I}(j,k) < \mathcal{I}(i,k) < \mathcal{I}(i,j)$.*

For any ordering $\boldsymbol{\gamma} : \{1, \ldots, N\} \to \{1, \ldots, N\}$, x is *adjacent* to y under $\boldsymbol{\gamma}$ if $\left|\boldsymbol{\gamma}^{-1}(x) - \boldsymbol{\gamma}^{-1}(y)\right| = 1$. The next lemma characterizes the orderings $\boldsymbol{\sigma}^\ell$ generated under the StaticMNL algorithm. The proof of this lemma appears in Appendix A.2.

Lemma A.2. *Under Assumption 2.1, for each $\ell = 1, 2, \ldots, K$, either $\boldsymbol{\sigma}^\ell = \boldsymbol{\sigma}^{\ell-1}$ or $\boldsymbol{\sigma}^\ell$ is obtained by transposing two adjacent items under $\boldsymbol{\sigma}^{\ell-1}$.*

The next lemma shows that if a set G^ℓ (corresponding to the first C elements in the permutation $\boldsymbol{\sigma}^\ell$) is distinct from all the previous sets, then one of its elements must be strictly smaller. The proof of this result appears in Appendix A.3.

Lemma A.3. *Under Assumption 2.1, for each $\ell = 1, 2, \ldots, K$, if $G^\ell \neq G^{\ell-1}$, then G^ℓ is obtained from $G^{\ell-1}$ by transposition between $j_\ell = \sigma_C^{\ell-1}$ and $i_\ell = \sigma_{C+1}^{\ell-1}$ with $0 < i_\ell < j_\ell$, that is, $G^\ell = \left(G^{\ell-1} \setminus \{j_\ell\}\right) \cup \{i_\ell\}$ and $\sum_{\ell \in G^\ell} \ell < \sum_{\ell \in G^{\ell-1}} \ell$.*

The next lemma shows that the size of the assortments generated by the StaticMNL algorithm is non-increasing. The proof appears in Appendix A.4.

Lemma A.4. *Let θ denote the smallest index such that $\left|A^\theta\right| < C$. Then, $A^\theta = A^{\theta-1} \setminus \{j_\theta\}$, and for all $\ell > \theta$, either $A^\ell = A^{\ell-1}$ or $A^\ell = A^{\ell-1} \setminus \{j_\ell\}$.*

We are now ready to give a proof of Theorem 2.5.

Proof. To prove part (a) of Theorem 2.5, suppose that $A^\ell \neq A^{\ell-1}$. If $\left|A^\ell\right| = C$, it follows from Lemma A.4 that $\ell < \theta$. This implies that $A^\ell = G^\ell$ and $A^{\ell-1} = G^{\ell-1}$. It follows from Lemma A.3 that $A^\ell = \left(A^{\ell-1} \setminus \{j_\ell\}\right) \cup \{i_\ell\}$. On the other hand, if $\left|A^\ell\right| < C$, the desired result follows from Lemma A.4.

Lemma A.4 also shows that there are exactly $C-1$ distinct non-empty assortments of size $C-1$ or less, which establishes part (b) of Theorem 2.5. To complete the proof, it suffices to count the number of distinct assortments of size C. Note that if A^ℓ is an assortment of size C, it must be the case that $A^\ell = G^\ell$. Therefore, the number of distinct assortments of size C is bounded above by the number of distinct subsets among the sets G^ℓ.

Under Assumption 2.1, we know that, among $h_1(\cdot), \ldots, h_N(\cdot)$, each of the $\binom{N}{2}$ pairs of lines will intersect each other. Moreover, by Assumption 2.1, we have that $\sigma^0 = (N, N-1, \ldots, 2, 1)$, corresponding to the order of $h_i(\lambda)$'s at $\lambda = -\infty$, and $\sigma^K = (1, 2, \ldots, N-1, N)$ for $\lambda = +\infty$. Note that

$$\sum_{\ell \in G^0} \ell - \sum_{\ell \in G^K} \ell = \{N + (N-1) + \cdots + (N-C+1)\} - \{1 + 2 + \cdots + C\}$$
$$= NC - 2(1 + 2 + \cdots + C - 1) - C = NC - C(C-1) - C = C(N-C).$$

By Lemma A.3, whenever G^t is distinct from G^{t-1}, the total value $\sum_{\ell \in G^t} \ell$ is strictly less than $\sum_{\ell \in G^{t-1}} \ell$. Thus, in addition to G^0, there can be at most $C(N-C)$ distinct subsets of G^ℓ's. Therefore, the number of distinct subsets among $G^0, G^1, \ldots, G^K$ is at most $C(N-C)+1$. Thus, the maximum number of distinct non-empty assortments is at most $C(N-C)+1+(C-1) = C(N-C+1)$, which is the desired result. ☐

A.1 Proof of Lemma A.1

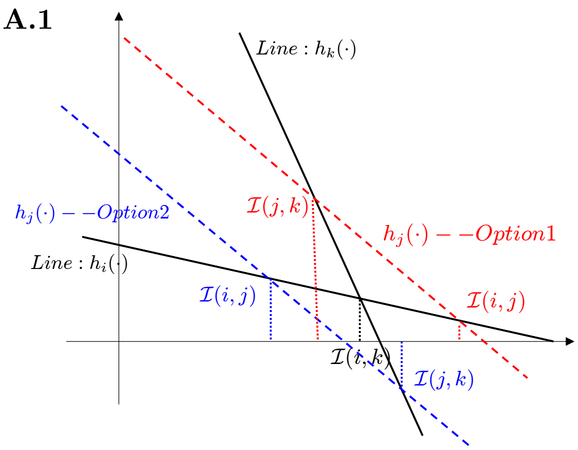

Figure 5: A geometric proof of Lemma A.1.

Proof. As shown in Figure 5, the proof of this lemma follows from a simple geometric intuition. By Assumption 2.1, we have that $v_i < v_j < v_k$, which implies that the line $h_i(\cdot)$ will intersect with $h_k(\cdot)$. Since v_j is between v_i and v_k, the are two possible options for line $h_j(\cdot)$ as shown in the two dash lines in Figure 5. In both cases, we observe that $\mathcal{I}(i,k)$ is always between $\mathcal{I}(i,j)$ and $\mathcal{I}(j,k)$, giving the desired results. A more algebraic proof is also straightforward. We omit the details due to space constraints. ☐

A.2 Proof of Lemma A.2

Proof. We will prove the lemma by induction. By Assumption 2.1, we know that $\boldsymbol{\sigma}^0 = (N, N-1, \ldots, 2, 1)$. We want to show that either $\boldsymbol{\sigma}^1 = \boldsymbol{\sigma}^0$ or $\boldsymbol{\sigma}^1$ is obtained by transposing two adjacent products. Suppose that $\boldsymbol{\sigma}^1 \neq \boldsymbol{\sigma}^0$. It then follows from the definition that $\boldsymbol{\sigma}^1$ is obtained from $\boldsymbol{\sigma}^0$ by a transposition of i_1 and j_1, where $\mathcal{I}(i_1, j_1)$ corresponds to the smallest intersection point. Since this is the first intersection point, it follows from Lemma A.1 that we must have $j_1 = i_1 + 1$, which is the desired result. Now, suppose that for each $s = 1, \ldots, \ell-1$, if $\boldsymbol{\sigma}^s \neq \boldsymbol{\sigma}^{s-1}$, then $\boldsymbol{\sigma}^s$ is obtained from $\boldsymbol{\sigma}^{s-1}$ by transposing two adjacent products. To complete the induction, we will now establish the result for $\boldsymbol{\sigma}^\ell$ via proof by contradiction.

Suppose on the contrary that $\boldsymbol{\sigma}^\ell \neq \boldsymbol{\sigma}^{\ell-1}$ and $\boldsymbol{\sigma}^\ell$ is obtained by $\boldsymbol{\sigma}^{\ell-1}$ by transposing i_ℓ and j_ℓ, which are **NOT** adjacent under $\boldsymbol{\sigma}^{\ell-1}$. This implies there is another product m that is between i_ℓ and j_ℓ under $\boldsymbol{\sigma}^{\ell-1}$, that is, $i_\ell = \sigma_u^{\ell-1}$, $m = \sigma_v^{\ell-1}$, and $j_\ell = \sigma_w^{\ell-1}$ where either $1 \le u < v < w \le N$ or $1 \le w < v < u \le N$.

We first claim that we cannot have $1 \le u < v < w \le N$. This follows because we know from Assumption 2.1(a) that $\sigma^0 = (N, N-1, \ldots, 2, 1)$. If $1 \le u < v < w \le N$ holds, this implies that we have $\boldsymbol{\sigma}^{\ell-1} =$

$(\ldots, i_\ell, \ldots, m, \ldots, j_\ell, \ldots)$. Since $i_\ell < j_\ell$ and the transpositions in all of the previous iterations involve adjacent items by induction, it must be the case that i_ℓ and j_ℓ have switched places once before, that is, we must have already encountered the intersection point $\mathcal{I}(i_\ell, j_\ell)$ in the earlier iterations. Contradiction!

So, we have $1 \leq w < v < u \leq N$, which implies that $\boldsymbol{\sigma}^{\ell-1} = (\ldots, j_\ell, \ldots, m, \ldots, i_\ell, \ldots)$. By construction, we know that $i_\ell < j_\ell$. Thus, there are three cases to consider for the value of m.

Case 1: $m < i_\ell < j_\ell$. In this case, m is smaller than i_ℓ but appears earlier in the ordering $\boldsymbol{\sigma}^{\ell-1}$. Since $\sigma^0 = (N, N-1, \ldots, 2, 1)$ and all previous transpositions involve adjacent items by induction, it must be the case that i_ℓ and m interchange their positions in the earlier iteration, implying that $\mathcal{I}(m, i_\ell) < \mathcal{I}(i_\ell, j_\ell)$. It thus follows from Lemma A.1 that

$$\mathcal{I}(m, i_\ell) < \mathcal{I}(m, j_\ell) < \mathcal{I}(i_\ell, j_\ell),$$

implying that we must have already encountered the intersection point $\mathcal{I}(m, j_\ell)$ before this iteration. By induction, this means that m and j_ℓ should have interchanged their places and we must have m *before* j_ℓ (note that $m \geq 1$). This contradicts our definition of $\boldsymbol{\sigma}^{\ell-1}$!

Case 2: $i_\ell < m < j_\ell$. In this case, it follows from Lemma A.1 that either

$$\mathcal{I}(i_\ell, m) < \mathcal{I}(i_\ell, j_\ell) < \mathcal{I}(m, j_\ell) \quad \text{OR} \quad \mathcal{I}(m, j_\ell) < \mathcal{I}(i_\ell, j_\ell) < \mathcal{I}(i_\ell, m).$$

We will show that either one of the above condition will lead to a contradiction. Suppose that $\mathcal{I}(i_\ell, m) < \mathcal{I}(i_\ell, j_\ell) < \mathcal{I}(m, j_\ell)$. This implies that we must have encountered the intersection point $\mathcal{I}(i_\ell, m)$ before the current iteration. Therefore, by induction, m and i_ℓ should have switched places and we must have m *after* i_ℓ in $\boldsymbol{\sigma}^{\ell-1}$. This again contradicts our definition of $\boldsymbol{\sigma}^{\ell-1}$! Now, suppose that $\mathcal{I}(m, j_\ell) < \mathcal{I}(i_\ell, j_\ell) < \mathcal{I}(i_\ell, m)$. Then, we must have encountered the intersection point $\mathcal{I}(m, j_\ell)$ before the current iteration. Therefore, by induction, m and j_ℓ should have switched places and we must have m *before* j_ℓ in $\boldsymbol{\sigma}^{\ell-1}$. This again contradicts our definition of $\boldsymbol{\sigma}^{\ell-1}$!

Case 3: $i_\ell < j_\ell < m$. The proof for this case is similar to Case 1 and we omit the details.

Thus, all three cases lead to contradictions. Therefore, it must be the case that $\boldsymbol{\sigma}^{\ell}$ is obtained from $\boldsymbol{\sigma}^{\ell-1}$ by transposing two adjacent products. This completes the induction. $\square$

A.3 Proof of Lemma A.3

Proof. By definition, if $G^\ell \neq G^{\ell-1}$, it must be the case that we encounter an intersection point $\mathcal{I}(i_\ell, j_\ell)$ with $0 < i_\ell < j_\ell$. In this case, $\boldsymbol{\sigma}^\ell$ is obtained from $\boldsymbol{\sigma}^{\ell-1}$ by transposition of i_ℓ and j_ℓ. Moreover, we know from Lemma A.2 that i_ℓ is adjacent to j_ℓ under $\boldsymbol{\sigma}^{\ell-1}$. Note that G^ℓ and $G^{\ell-1}$ correspond to the first C elements for the ordering $\boldsymbol{\sigma}^\ell$ and $\boldsymbol{\sigma}^{\ell-1}$, respectively. Thus, in order for G^ℓ to be different from $G^{\ell-1}$, it must be the case that either 1) $\sigma_C^{\ell-1} = j_\ell$ and $\sigma_{C+1}^{\ell-1} = i_\ell$, or 2) $\sigma_C^{\ell-1} = i_\ell$ and $\sigma_{C+1}^{\ell-1} = j_\ell$.

We will first show that option 2) is not feasible. Recall that under Assumption 2.1, we have $\boldsymbol{\sigma}^0 = (N, N-1, \ldots, 2, 1)$. Since $i_\ell < j_\ell$ and every transposition only occur between adjacent items (by Lemma A.2), if $\sigma_C^{\ell-1} = i_\ell$ and $\sigma_{C+1}^{\ell-1} = j_\ell$, then it must be the case that i_ℓ and j_ℓ have interchanged places before; that is, we have already encountered the intersection point $\mathcal{I}(i_\ell, j_\ell)$ in the earlier iterations. This is a contradiction! Therefore, we can only have $\sigma_C^{\ell-1} = j_\ell$ and $\sigma_{C+1}^{\ell-1} = i_\ell$. In this case, we have $G^\ell = \left(G^{\ell-1} \setminus \{j_\ell\}\right) \cup \{i_\ell\}$, which implies that $\sum_{\ell \in G^\ell} \ell - \sum_{\ell \in G^{\ell-1}} \ell = i_\ell - j_\ell < 0$, which is the desired result. $\square$

A.4 Proof of Lemma A.4

Proof. Since $\left|A^0\right| = C$, $\theta \geq 1$. We will first show that $A^\theta = A^{\theta-1} \setminus \{j_\theta\}$. By definition, the set A^θ is created when we encounter the θ^{th} intersection point $\mathcal{I}(i_\theta, j_\theta)$. We claim that we must have $i_\theta = 0$. Suppose on the contrary that we have $0 < i_\theta < j_\theta$. In this case, $B^\theta = B^{\theta-1}$. Thus, for $\left|A^\theta\right| < C = \left|A^{\theta-1}\right|$, it must be the case that $G^\theta \neq G^{\theta-1}$. It follows from Lemma A.3 that G^θ is obtained from $G^{\theta-1}$ by transposition of $j_\theta = \sigma_C^{\theta-1}$ and $i_\theta = \sigma_{C+1}^{\theta-1}$ with $0 < i_\theta < j_\theta$. Since $\left|A^\theta\right| < C$, it follows that $B^\theta \ni i_\theta$, which implies that we must have already encountered the intersection $\mathcal{I}(0, i_\theta)$ earlier. By Lemma A.1,

$$\mathcal{I}(0, i_\theta) < \mathcal{I}(0, j_\theta) < \mathcal{I}(i_\theta, j_\theta),$$

which implies that the intersection point $\mathcal{I}(0, j_\theta)$ appeared before the current iteration, and thus, $B^{\theta-1} \ni j_\theta$, which implies that $\left|A^{\theta-1}\right| < C$. Contradiction! Therefore, we must have $i_\theta = 0$. Then, we know that $\boldsymbol{\sigma}^\theta = \boldsymbol{\sigma}^{\theta-1}$, $G^\theta = G^{\theta-1}$, and $B^\theta = B^{\theta-1} \cup \{j_\theta\}$. Since $\left|A^{\theta-1}\right| = C > \left|A^\theta\right|$, $G^{\theta-1} \cap B^{\theta-1} = \emptyset$ and $G^\theta \cap B^\theta \neq \emptyset$. Since only j_θ is added to the set $B^{\theta-1}$, we have $G^\theta \cap B^\theta = \{j_\theta\}$, and therefore, $A^\theta = G^\theta \setminus B^\theta = G^\theta \setminus \{j_\theta\} = G^{\theta-1} \setminus \{j_\theta\} = A^{\theta-1} \setminus \{j_\theta\}$, which is the desired result.

To complete the proof of Lemma A.4, consider any $\ell > \theta$. Let $r = \left|A^{\ell-1}\right|$. Since $\ell > \theta$, we have that $r < C$. By the definition, $\sigma^{\ell-1} = \left(\sigma_1^{\ell-1}, \ldots, \sigma_N^{\ell-1}\right)$ represent the ordering of the lines $h_1(\cdot), \ldots, h_N(\cdot)$ during the interval $\Big(\mathcal{I}(i_{\ell-1}, j_{j-1}), \mathcal{I}(i_\ell, j_\ell)\Big)$ with

$$h_{\sigma_1^{\ell-1}}(\mathcal{I}(i_\ell, j_\ell)) \geq h_{\sigma_2^{\ell-1}}(\mathcal{I}(i_\ell, j_\ell)) \geq \cdots \geq h_{\sigma_N^{\ell-1}}(\mathcal{I}(i_\ell, j_\ell)) \tag{11}$$

Since $G^{\ell-1} = \left\{\sigma_1^{\ell-1}, \ldots, \sigma_C^{\ell-1}\right\}$, $A^{\ell-1} = G^{\ell-1} \setminus B^{\ell-1}$, and $\left|A^{\ell-1}\right| = r < C$, it follows from the above ordering that

$$h_{\sigma_r^{\ell-1}}(\mathcal{I}(i_\ell, j_\ell)) \geq 0 > h_{\sigma_{r+1}^{\ell-1}}(\mathcal{I}(i_\ell, j_\ell)) \geq \cdots \geq h_{\sigma_N^{\ell-1}}(\mathcal{I}(i_\ell, j_\ell)) \ ,$$

and $B^{\ell-1} = \left\{\sigma_{r+1}^{\ell-1}, \sigma_{r+2}^{\ell-1}, \ldots, \sigma_N^{\ell-1}\right\}$.

We will show that either $A^\ell = A^{\ell-1}$ or $A^\ell = A^{\ell-1} \setminus \{j_\ell\}$. Consider the ℓ^{th} intersection point $\mathcal{I}(i_\ell, j_\ell)$. There are two cases to consider: $i_\ell = 0$ and $i_\ell \geq 1$. Suppose that $i_\ell = 0$. Then, we claim that $j_\ell = \sigma_r^{\ell-1}$. Since $B^{\ell-1} = \left\{\sigma_{r+1}^{\ell-1}, \sigma_{r+2}^{\ell-1}, \ldots, \sigma_N^{\ell-1}\right\}$, we know that $j_\ell \in \left\{\sigma_1^{\ell-1}, \ldots, \sigma_r^{\ell-1}\right\}$. If, on the contrary, $j_\ell = \sigma_k^{\ell-1}$ where $k < r$, this means that $\mathcal{I}\left(0, \sigma_k^{\ell-1}\right) < \mathcal{I}\left(0, \sigma_r^{\ell-1}\right)$; that is, the value associated with the line $h_{\sigma_r^{\ell-1}}$ remains non-negative. However, by the ordering in (11), it must be the case that $0 > h_{\sigma_k^{\ell-1}}(\mathcal{I}(i_\ell, j_\ell)) \geq h_{\sigma_r^{\ell-1}}(\mathcal{I}(i_\ell, j_\ell))$. Contradiction! Therefore, $j_\ell = \sigma_r^{\ell-1}$, which implies that $A^\ell = A^{\ell-1} \setminus \left\{\sigma_r^{\ell-1}\right\}$, which is the desired result.

On the other hand, suppose that $i_\ell \geq 1$. In this case, we have $B^\ell = B^{\ell-1}$. We claim that we must either have $\{i_\ell, j_\ell\} \subset B^\ell$ or $\{i_\ell, j_\ell\} \subset \left(B^\ell\right)^c$; that is, either both elements are in B^ℓ or both are not. This will show that $A^\ell = A^{\ell-1}$, which is the desired result. To prove this, note that since $0 < i_\ell < j_\ell$, it follows from Lemma A.1 that either $\mathcal{I}(0, i_\ell) < \mathcal{I}(0, j_\ell) < \mathcal{I}(i_\ell, j_\ell)$ or $\mathcal{I}(i_\ell, j_\ell) < \mathcal{I}(0, j_\ell) < \mathcal{I}(0, i_\ell)$. If $\mathcal{I}(0, i_\ell) < \mathcal{I}(0, j_\ell) < \mathcal{I}(i_\ell, j_\ell)$ holds, then the intersection points $\mathcal{I}(0, i_\ell)$ and $\mathcal{I}(0, j_\ell)$ appear in the earlier iterations, implying that $\{i_\ell, j_\ell\} \subset B^\ell$. On the other hand, if $\mathcal{I}(i_\ell, j_\ell) < \mathcal{I}(0, j_\ell) < \mathcal{I}(0, i_\ell)$, then $i_\ell \notin B^\ell$ and $j_\ell \notin B^\ell$. □

B. Proof of Lemma 2.7

Let $r = \left(\sqrt{5} - 1\right)/2$ denote the golden ratio and let $W_{GRS} = \left\lceil 2(\log T)/\beta^4 \right\rceil$. Since the classical Golden Ratio search reduces the size of the target set by factor of r in each iteration, the maximum number of iterations is at most $\lceil \log|\mathcal{A}| / \log(1/r) \rceil$, where $|\mathcal{A}|$ denotes the number of assortments in the sequence $\mathcal{A}$.

For $k = 1, \ldots, \lceil \log|\mathcal{A}| / \log(1/r) \rceil$, let $1 \leq a_k < b_k < c_k < d_k \leq |\mathcal{A}|$ denote the 4 indices associated with the k^{th} iteration of the golden search algorithm. In the k^{th} iteration, we consider the assortments $A^{a_k}, A^{a_k+1}, \ldots, A^{d_k}$, and we compare the assortment A^{b_k} and A^{c_k} based on their average profits. If the average profit of A^{b_k} is larger than that of A^{c_k}, the resulting target set for the next iteration is $A^{a_k}, \ldots, A^{c_k}$. Otherwise, the new target set will be $A^{b_k}, \ldots, A^{d_k}$. Let $\mathcal{B}_k$ denote the event that resulting target set after the k^{th} iteration does *not* contain the optimal assortment, that is, we made an error in the k^{th} iteration. We wish to show that $\Pr\{\mathcal{B}_k\} \leq 2ke^{-\beta^4 W_{GRS}/2}$

Let $q \in \{1, \ldots, |\mathcal{A}|\}$ denote the index of the optimal assortment. Consider the first iteration where we compare the average profit of A^{b_1} and A^{c_1}. Let $Y_1^{b_1}, \ldots, Y_{W_{GRS}}^{b_1}$ denote the observed profit associated with the selections of the W_{GRS} customers who were offered the assortment A^{b_1}. Similarly, let $Y_1^{c_1}, \ldots, Y_{W_{GRS}}^{c_1}$ denote the observed profits associated with A^{c_1}. Note that if $b_1 \leq q \leq c_1$, then $\Pr\{\mathcal{B}_1\} = 0$ because the resulting target set will always contain the optimal assortment, regardless of the outcome of the comparison. So, suppose that $q < b_1$. Then, we know from Theorem 2.6 that $f(A^{b_1}) - f(A^{c_1}) > \beta^2$. Using the fact that $\mathsf{E}\left[Y_i^{b_1}\right] = f(A^{b_1})$ and $\mathsf{E}\left[Y_i^{c_1}\right] = f(A^{c_1})$ for all $1 \leq i \leq W_{GRS}$, we can bound the probability of error in the first iteration as follows

$$\begin{aligned}
\Pr\{\mathcal{B}_1\} &= \Pr\left\{\frac{1}{W_{GRS}}\sum_{i=1}^{W_{GRS}} Y_i^{b_1} < \frac{1}{W_{GRS}}\sum_{i=1}^{W_{GRS}} Y_i^{c_1}\right\} \\
&\leq \Pr\left\{\frac{1}{W_{GRS}}\sum_{i=1}^{W_{GRS}} \left(Y_i^{c_1} - f(A^{c_1})\right) + \frac{1}{W_{GRS}}\sum_{i=1}^{W_{GRS}} \left(f(A^{b_1}) - Y_i^{b_1}\right) > f(A^{b_1}) - f(A^{c_1})\right\} \\
&\leq \Pr\left\{\frac{1}{W_{GRS}}\sum_{i=1}^{W_{GRS}} \left(Y_i^{c_1} - f(A^{c_1})\right) + \frac{1}{W_{GRS}}\sum_{i=1}^{W_{GRS}} \left(f(A^{b_1}) - Y_i^{b_1}\right) > \beta^2\right\} \\
&\leq \Pr\left\{\frac{1}{W_{GRS}}\sum_{i=1}^{W_{GRS}} \left(Y_i^{c_1} - f(A^{c_1})\right) > \frac{\beta^2}{2}\right\} + \Pr\left\{\frac{1}{W_{GRS}}\sum_{i=1}^{W_{GRS}} \left(f(A^{b_1}) - Y_i^{b_1}\right) > \frac{\beta^2}{2}\right\} \\
&\leq 2e^{-\beta^4 W_{GRS}/2},
\end{aligned}$$

where the last inequality follows from the classical Chernoff-Hoeffding Inequality and the fact that $w_\ell \leq 1$ for all ℓ. The argument for the case where $q > c_1$ is exactly the same. To get a bound on $\Pr\{\mathcal{B}_2\}$, note that $\Pr\{\mathcal{B}_2\} = \Pr\{\mathcal{B}_2|\mathcal{B}_1\}\Pr\{\mathcal{B}_1\} + \Pr\{\mathcal{B}_2|\mathcal{B}_1^c\}\Pr\{\mathcal{B}_1^c\} \leq \Pr\{\mathcal{B}_1\} + \Pr\{\mathcal{B}_2|\mathcal{B}_1^c\}$. Using exactly the same argument as above, we can show that $\Pr\{\mathcal{B}_2|\mathcal{B}_1^c\} \leq 2e^{-\beta^4 W_{GRS}/2}$. Therefore, $\Pr\{\mathcal{B}_2\} \leq 2\left(2e^{-\beta^4 W_{GRS}/2}\right)$. Repeated applications show that $\Pr\{\mathcal{B}_k\} \leq 2ke^{-\beta^4 W_{GRS}/2}$.

To determine an upper bound on the regret after T periods, we note that the total number of samples that we used is at most

$$W_{GRS} \cdot \left(\left\lceil \frac{\log|\mathcal{A}|}{\log(1/r)} \right\rceil + 3\right) \leq 4W_{GRS} \cdot \left\lceil \frac{\log|\mathcal{A}|}{\log(1/r)} \right\rceil,$$

where the factor of 3 reflects the fact that the standard Golden Ratio search starts with 4 points in the first iteration and adds one additional search point in subsequent iterations. The maximum expected regret incurred from these samples is at most $4 \cdot W_{GRS} \cdot \left\lceil \frac{\log|\mathcal{A}|}{\log(1/r)} \right\rceil$. because $\max_\ell w_\ell \leq 1$. Moreover, if we correctly identify the optimal assortment at the end of the search, the regret will be zero thereafter. Thus, the maximum expected regret after we conclude the search algorithm is at most

$$T \cdot \Pr\left\{\mathcal{B}_{\lceil \log|\mathcal{A}| / \log(1/r) \rceil}\right\} \leq 2T \left\lceil \frac{\log|\mathcal{A}|}{\log(1/r)} \right\rceil e^{-\beta^4 W_{GRS}/2},$$

where the right hand side follows from the upper bound on the error probability. Using the definition of

W_{GRS}, we have the following upper bound on the expected regret after T periods:

$$\begin{aligned}
&4 \cdot W_{GRS} \cdot \left\lceil \frac{\log |\mathcal{A}|}{\log(1/r)} \right\rceil + 2T \left\lceil \frac{\log |\mathcal{A}|}{\log(1/r)} \right\rceil e^{-\beta^4 W_{GRS}/2} \\
\leq\ & 4\left(\frac{2\log T}{\beta^4} + 1\right)\left(\frac{\log |\mathcal{A}|}{\log(1/r)} + 1\right) + 2\left(\frac{\log |\mathcal{A}|}{\log(1/r)} + 1\right) \\
=\ & 4\left(\frac{2\log T}{\beta^4} + 1\right)\left(\frac{\log(|\mathcal{A}|/r)}{\log(1/r)}\right) + 2\left(\frac{\log(|\mathcal{A}|/r)}{\log(1/r)}\right) \\
=\ & \frac{8\log(|\mathcal{A}|/r)}{\beta^4 \log(1/r)} \log T + 6\frac{\log(|\mathcal{A}|/r)}{\log(1/r)},
\end{aligned}$$

and the desired result follows from the fact that $1/r \leq 2$ and $1/\log(1/r) \leq 2.079$, and from Theorem 2.5 which shows that $|\mathcal{A}| \leq C(N - C + 1) \leq N^2$.

C. Proof of Theorem 3.2

Proof. By the union bound and the fact that $\epsilon \geq \min\{\epsilon, \beta/3\}$, we have that

$$\begin{aligned}
&\Pr\left\{\max_{S\in\mathcal{E}}\left\{\max_{\{i,j\}\subseteq S: i\neq j}\left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right|, \max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(m,S\right)\right|\right\} > \epsilon\right\} \\
\leq\ & \sum_{S\in\mathcal{E}} \Pr\left\{\max_{\{i,j\}\subseteq S: i\neq j}\left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right| > \epsilon \ \text{ OR } \ \max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(m,S\right)\right| > \min\{\epsilon, \beta/3\}\right\} \\
\leq\ & \sum_{S\in\mathcal{E}} \Pr\left\{\max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(m,S\right)\right| > \epsilon\beta^2/12\right\} ,
\end{aligned}$$

where the last inequality follows from the fact that the event

$$\max_{\{i,j\}\subseteq S: i\neq j}\left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right| > \epsilon \ \text{ OR } \ \max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(m,S\right)\right| > \min\{\epsilon, \beta/3\}$$

is a subset of the event

$$\max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(m,S\right)\right| > \epsilon\beta^2/12 \ .$$

To see this, note that $\min\{\epsilon, \beta/3\} \geq \epsilon\beta^2/12$ because $0 < \beta < 1$. If $\max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(m,S\right)\right| > \min\{\epsilon, \beta/3\}$, then the result is trivially true.

On the other hand, suppose that

$$\max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(m,S\right)\right| \leq \min\{\epsilon, \beta/3\} \quad \text{and} \quad \max_{\{i,j\}\subseteq S: i\neq j}\left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right| > \epsilon \ .$$

It follows from the definition of β in Equation (8) that for any $(i,j) \neq (s,t)$, $|\mathcal{I}(i,j) - \mathcal{I}(s,t)| \geq \beta$, and for any $i \neq j$ and $S \in \mathcal{E}$, $|\theta_i(S) - \theta_j(S)| \geq \beta$. This implies that for each $\{i,j\} \subseteq S$ and $i \neq j$, $\left|\widehat{\Theta}_i\left(m,S\right) - \widehat{\Theta}_j\left(m,S\right)\right| \geq \beta/3$. Therefore,

$$\begin{aligned}
\left|\frac{\theta_i(S)}{\theta_i(S) - \theta_j(S)} - \frac{\widehat{\Theta}_i(m,S)}{\widehat{\Theta}_i(m,S) - \widehat{\Theta}_j(m,S)}\right| &\leq \frac{2\max_{\ell\in S}\left|\theta_\ell(S) - \widehat{\Theta}_\ell(m,S)\right|}{\left(\theta_i(S) - \theta_j(S)\right)\left(\widehat{\Theta}_i(m,S) - \widehat{\Theta}_j(m,S)\right)} \\
&\leq \frac{6}{\beta^2}\max_{\ell\in S}\left|\theta_\ell(S) - \widehat{\Theta}_\ell(m,S)\right| ,
\end{aligned}$$

and it follows from the definition of $\mathcal{I}(i,j)$ and $\mathcal{I}_{ij}(m,S)$ and the fact that $\max_\ell w_\ell \leq 1$ that

$$\left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right| \leq \frac{12}{\beta^2}\max_{\ell\in S}\left|\theta_\ell(S) - \widehat{\Theta}_\ell(m,S)\right| \ .$$

Since $\{i,j\} \subseteq S$ is arbitrary, we have that

$$\epsilon < \max_{\{i,j\}\subseteq S: i\neq j} \left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right| \leq \frac{12}{\beta^2} \max_{\ell \in S} \left|\theta_\ell(S) - \widehat{\Theta}_\ell(m,S)\right| ,$$

which once again implies that $\max_{i\in S} \left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(c,S\right)\right| > \epsilon\beta^2/12$.

Note that for each $i \in S$, $\widehat{\Theta}_i(c,S)$ is simply the average of c independent Bernoulli random variables with parameter $\theta_i(S)$. Therefore, it follows from the standard Chernoff-Hoeffding Inequality that

$$\Pr\left\{\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(c,S\right)\right| > \epsilon\beta^2/12\right\} \leq 2\,e^{-2c(\epsilon\beta^2/12)^2} = 2\,e^{-c\,\epsilon^2\,\beta^4/72} ,$$

and thus, $\Pr\left\{\max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(c,S\right)\right| > \epsilon\beta^2/12\right\} \leq 2Ce^{-c\,\epsilon^2\,\beta^4/72}$. Putting everything together and using the fact that $|\mathcal{E}| \leq 5(N/C)^2$, we have that

$$\Pr\left\{\max_{S\in\mathcal{E}}\left\{\max_{\{i,j\}\subseteq S: i\neq j}\left|\mathcal{I}(i,j) - \widehat{\mathcal{I}}_{ij}(m,S)\right|, \max_{i\in S}\left|\theta_i\left(S\right) - \widehat{\Theta}_i\left(c,S\right)\right|\right\} > \epsilon\right\} \leq \frac{10N^2}{C} e^{-c\,\epsilon^2\,\beta^4/72} ,$$

which is the desired result. □

D. Proof of Theorem 3.4

Since $w_\ell \leq 1$ for all ℓ, the regret incurred in each period is bounded above by 1. Consider an arbitrary cycle m. We incurred a total regret of $|\mathcal{E}|$ during the exploration phase of cycle m. At the end of the exploration phase, if $\widehat{\mathcal{A}}(m) = \mathcal{A}$, then the regret incurred from the SAMPLING-BASED GRS algorithm during the exploitation phase is bounded above by $a_1(\log N)\log V_m$ by Lemma 2.7. On the other hand, if $\widehat{\mathcal{A}}(m) \neq \mathcal{A}$, we incurred a total regret of at most V_m. Thus, the total regret incurred during the m^{th} cycle is bounded above by

$$|\mathcal{E}| + V_m \Pr\left\{\widehat{\mathcal{A}}(m) \neq \mathcal{A}\right\} + a_1(\log N)\log V_m \leq |\mathcal{E}| + \frac{10N^2}{C} e^{-(\alpha-\beta^6/288)m} + a_1\alpha(\log N)m ,$$

where the inequality follows from Theorem 3.3 and the fact that $V_m = \lfloor e^{\alpha m}\rfloor \leq e^{\alpha m}$. Therefore, the total cumulative regret after M cycles (corresponding to $|\mathcal{E}|\,M + \sum_{m=1}^M V_m$ periods) is bounded above by

$$\begin{aligned}
\mathrm{Regret}\left(|\mathcal{E}|\,M + \sum_{m=1}^{M} V_m, \mathrm{AA}\right) &\leq |\mathcal{E}|\,M + a_1\alpha\log N\sum_{m=1}^{M} m + \frac{10N^2}{C}\sum_{m=1}^{M} e^{-(\alpha-\beta^6/288)m} \\
&\leq |\mathcal{E}|\,M + a_1\alpha(\log N)M^2 + \frac{10N^2/C}{1 - e^{-(\alpha-\beta^6/288)}}
\end{aligned}$$

Consider an arbitrary time period T. Let $M_0 = \left\lceil \frac{\log T}{\alpha}\right\rceil$. Note that the total time periods after M_0 cycles is at least T because

$$M_0\,|\mathcal{E}| + \sum_{m=1}^{M_0} V_m = M_0\,|\mathcal{E}| + \sum_{m=1}^{M_0}\lfloor e^{\alpha m}\rfloor = M_0(|\mathcal{E}| - 1) + \sum_{m=1}^{M_0} e^{\alpha m} \geq e^{\alpha M_0} \geq T .$$

Since the cumulative regret is nondecreasing, it follows that

$$\begin{aligned}
\mathrm{Regret}\left(T, \mathrm{AA}\right) &\leq \mathrm{Regret}\left(|\mathcal{E}|\,M_0 + \sum_{m=1}^{M_0} V_m, \mathrm{AA}\right) \\
&\leq |\mathcal{E}|\,M_0 + a_1\alpha(\log N)M_0^2 + \frac{10N^2/C}{1 - e^{-(\alpha-\beta^6/288)}} \;\leq\; a_2N^2\log^2 T ,
\end{aligned}$$

for some constant a_2 that depends on C, $\boldsymbol{v}$, and $\boldsymbol{w}$. The last inequality above follows from the fact that $|\mathcal{E}| = O(N^2)$.