Joint Assortment and Inventory Planning for Heavy Tailed Demand

Omar El Housni[1], Vineet Goyal[2], Salal Humair[3], Omar Mouchtaki[4], Ali Sadighian[3], and Jingchen Wu[3]

[1] *School of Operations Research and Information Engineering, Cornell Tech*
[2] *Department of Industrial Engineering and Operations Research, Columbia University*
[3] *Amazon.com*
[4] *Graduate School of Business, Columbia University*

April 22, 2021

Abstract

We study a joint assortment and inventory optimization problem faced by an online retailer who needs to decide on both the assortment along with the inventories of a set of N substitutable products before the start of the selling season to maximize the expected profit. The problem raises both algorithmic and modeling challenges. One of the main challenges is to tractably model dynamic stock-out based substitution where a customer may substitute to the most preferred product that is available if their first choice is not offered or stocked-out. We first consider the joint assortment and inventory optimization problem for a Markov Chain choice model and present a near-optimal algorithm for the problem. Our results significantly improve over the results in Gallego and Kim (2020) where the regret can be linear in T (where T is the number of customers) in the worst case. We build upon their approach and give an algorithm with regret $\tilde{\mathcal{O}}(\sqrt{NT})$ with respect to an LP upper bound. Our algorithm achieves a good balance between expected revenue and inventory costs by identifying a subset of products that can pool demand from the universe of substitutable products without significantly cannibalizing the revenue in the presence of dynamic substitution behavior of customers. We also present a multi-step choice model that captures the complex choice process in an online retail setting characterized by a large universe of products and a heavy-tailed distribution of mean demands. Our model captures different steps of the choice process including search, formation of a consideration set and eventual purchase. We conduct computational experiments that show that our algorithm empirically outperforms previous approaches both on synthetic and realistic instances.

Keywords: Inventory planning, stock-out based substitution, assortment optimization, heavy-tailed demand, sample average approximation, Markov Chain choice model.

1 Introduction

We study a joint assortment and inventory optimization problem where we need to decide an assortment of substitutable products along with their inventories in advance of the realization of

[1]oe46@cornell.edu, [2]vg2277@columbia.edu, [3]{salal, alisadi, jcwu}@amazon.com, [4]om2316@columbia.edu

the demand. Demand for substitutable products is highly correlated and depends on the set of available products. This is because of dynamic stock-out based substitutions (SOBS) where a customer substitutes to a less preferred product if their most preferred product is not offered or stocked-out. Such dynamic substitution and correlation in demand makes the joint assortment and inventory decision quite challenging even for simple models where a static optimal assortment can be computed efficiently.

In this paper, we focus on the assortment and inventory problem in the context of online retailing. While some of the challenges are similar to brick-and-mortar settings, aspects particular to online retailing add complexity to the problem. First, the universe of products is significantly larger than brick-and-mortar stores. Second, the interaction between the customer and an online retailer is complex and involves multiple layers including search, recommendation, consideration phase and possibly an eventual purchase. This complex interaction adds significant modeling challenges to capture the choice process that cannot be captured by a simple choice model amenable to easy optimization. For example, the distribution of customer search queries is usually heavy tailed (Szpektor et al. (2011); Jansen and Spink (2006)), the subset of products displayed to customers and their ranking on a page is governed by search engines that have their own algorithms for surfacing results, customers browse/click probabilities are known to decay sharply as a function of product ranking in the display list (Silverstein et al. (1999)), and customers usually browse a small set of products before making purchases or walking away (Agarwal et al. (2011); Ursu (2018); Honka et al. (2019)). The confluence of these human and system behaviors means a small number of products tend to get a large fraction of the total views. As a consequence, the distribution of mean demand for products is usually heavy-tailed (Brynjolfsson et al. (2006); Goel et al. (2010)). The heavy tailed nature of the mean demands adds to the challenge and standard approaches from literature for optimizing assortment and inventory, such as fluid approximations of substitutable demand, may not work well.

Revenue management under parametric choice models has a rich literature. Parametric models enable encoding a relatively complex choice model by a small-number of features. Several revenue management problems have been solved under widely used parametric choice models. For instance, under the Multinomial Logit model (MNL), assortment optimization can be solved optimally. Aouad and Segev (2019) give a PTAS (Polynomial Time Approximation Scheme) for the inventory problem under the MNL choice model. Other works developed approaches to solve revenue management problems on Markov chain choice model (Blanchet et al. (2016); Feldman and Topaloglu (2017); Désir et al. (2020)). Howard and Sheth (1969) present a consider-then-choose model for consideration sets in customer's behavior. It has been empirically validated by Jeuland (1979) and Crompton and Ankomah (1993) among others. Wang and Sahin (2018) present a model that incorporates search cost in the customer choice process and study the assortment selection and the pricing problem under this model. More recently, a line of work has studied multiple variants of the assortment selection problem under consider-then-choose models (Li et al. (2018); Aouad et al. (2019, 2020)). In particular, Aouad et al. (2019) introduce the click-based MNL choice model to capture customer choice process in online retail settings. Clicks of customers represent the consideration set among which they make the final purchase decision. We extend this model by considering the journey of a customer starting from the query and ending at the purchase decision.

While the algorithms developed in these works depend on the structure of the underlying choice model, our goal in this paper is to abstract the complex choice process in online retail in a manner that preserves the heavy tailed nature of the demand; and design a near optimal algorithm for jointly computing the assortment and inventory that does not depend on the complexity of the choice process.

Our Contributions. We develop a modular approach to decompose the challenge of *both* modeling a complex choice process and solving the resulting joint assortment and inventory optimization problem. In particular, we approximate the underlying complex choice process as a Markov Chain (MC) choice model (Blanchet et al. (2016)) and formulate the joint assortment and inventory optimization using the MC model. Our main contributions are the following.

SAA based near-optimal regret algorithm. We present a sample average approximation (SAA) based algorithm for the joint assortment and inventory optimization problem that achieves a worst-case regret of $\tilde{\mathcal{O}}(\sqrt{NT})$[1]. Here, N is the number of substitutable products in the universe and T is the number of customer visits.

We would like to note that Gallego and Kim (2020) consider the inventory optimization problem under the Markov Chain model and present an SAA based heuristic. However, we show that their approach can lead to a solution with $\Omega(T)$-regret in the worst-case as compared to the optimal solution[2]. The linear regret arises from the fact that their SAA based LP is not able to accurately capture dynamic substitution where every customer must select the most preferred product among the available ones. In fact, their LP allows assigning less preferred products to customers if that is more profitable for the seller even if more preferred ones have available inventory. We present families of instances where this drawback leads to solutions with linear regret.

Our algorithm builds upon the SAA based approach and overcomes the above drawback by identifying upfront the set of products that can potentially cannibalize the revenue from products in the optimal assortment and limiting the inventory of such products. In general, it can be challenging to identify the set of cannibalizing products. However, for the Markov Chain model, using ideas of reduced price introduced by Désir et al. (2020), one can efficiently identify the set of cannibalizing products. Furthermore, we also consider an alternate sampling approach to more

[1]The $\tilde{\mathcal{O}}$ notation hides logarithmic factors.

[2]The asymptotic optimality shown in Gallego and Kim (2020) holds under certain restrictive assumptions where the number of products and number of sampled demand scenarios are constant while the number of customers goes to infinity.

effectively capture the demand stochasticity for a large universe of products. We show that our algorithm has a regret of $\tilde{\mathcal{O}}(\sqrt{NT})$ that is near-optimal.

Parsimonious choice model from query to purchase. We present a parsimonious choice model that captures the complexities of customer choice and system interaction and produces a heavy tailed demand distribution across products. In particular, we consider a setting where a customer journey starts from a query. The customer is then shown a ranked set of products that are relevant to the received query. The ranking is built by sampling without replacement products according to a multinomial distribution. The customer then builds a small consideration set of products by browsing and clicking on products following the ranking of the displayed assortment. The browsing phase stops once the customer gets impatient and switches to the purchase phase. The purchase decision is finally made among products in the consideration set according to an MNL model. We show that we can efficiently estimate from choice data the parameters of the model. Specifically, we show that the log-likelihood is jointly concave in the parameters. We also present a computational study to show that our algorithm significantly outperforms prior approaches.

1.1 Related literature

Assortment and inventory optimization are important problems that have been studied extensively in the literature. A first version of the problem has been studied by van Ryzin and Mahajan (1999) who analyze the trade-offs between assortment breadth and inventory costs for an MNL choice model. However, they consider a static substitution model for the customers that depends statically on the initial assortment. Variants of this seminal work have been studied by Cachon et al. (2005) when customers have a search cost and in Topaloglu (2013) where the assumption of identical price and cost across products is relaxed. However, these works do not consider dynamic stock-out based substitution which is important to model especially in our setting of heavy-tailed demand.

Dynamic substitution models that consider stock-out events has also been studied extensively in the literature. Parlar and Goyal (1984) study a two-stage substitution framework in a two product probabilistic substitution model. Smith and Agrawal (2000) design a heuristic to solve this problem in a more general setting with an arbitrary number of products. Kök and Fisher (2007) couple inventory decision with a procedure to estimate substitution parameters and benchmark it on real data. Netessine and Rudi (2003) and Nagarajan and Rajagopalan (2008) develop other two-stage substitution methods and provide heuristics to solve them. However, algorithms developed by these previous approaches rely heavily on the structure of stock-out substitution. In practice SOBS can be complex and it may be intractable to design algorithms based on the structural properties in general.

In this paper, we consider a sample-path-based inventory model where dynamic substitution is completely represented in the demand function. In this more general setting, Mahajan and van Ryzin (2001) show that profit function is not necessarily quasi-concave and provide a gradient-based algorithm which converges to stationary points. Gaur and Honhon (2006) propose a heuristic that yields an upper-bound on the profit for the one-dimensional location model. Honhon et al. (2010) and Honhon and Seshadri (2013) give an algorithm in the heterogeneous customer preference list model and give an upper-bound on profit in the case of random proportions of customer types. Goyal et al. (2016) consider the capacity constrained problem where there is a bound on the total number of units of inventory. They develop a polynomial time approximation scheme (PTAS) for the nested preference list model. Following the estimation based procedure of Kök and Fisher (2007), Müller et al. (2020) propose a data-driven approach to solve the problem. They compare a model based setting and a model-free heuristic where inventory decision is the output of a neural-network.

Farahat and Lee (2018) and Gallego and Kim (2020) are the most closely related to our work. Farahat and Lee (2018) propose a methodology referred to as *approximate similarity transformation* from which they derive a heuristic with provable upper-bound on the generated profit. Gallego and Kim (2020) present a heuristic for the inventory optimization problem under a Markov Chain choice model and show that it is asymptotically optimal under certain settings. Numerically, Gallego and Kim (2020) show that their heuristic performs significantly better than the independent Newsvendor Problem and matches Farahat and Lee (2018) performances.

2 Notations, preliminaries and problem definition

We consider a set of products $\mathcal{N} := \{1, \ldots, N\}$ and the no-purchase option 0 with the convention that $\mathcal{N}_+ = \mathcal{N} \cup \{0\}$. Let $(p_i)_{i\in\mathcal{N}}$ and $(c_i)_{i\in\mathcal{N}}$ denote respectively the prices and the costs of products. The seller faces T i.i.d. customers with choice model given by π that specifies for any $S \subset \mathcal{N}$ and $i \in S$, $\pi(i, S)$, the probability that a random customer selects product i if the offered set is S. When customers arrive on the online store, they observe the set of offered products and make a choice according to the model π. The goal of the seller is to decide on the assortment and inventory $\mathbf{q} \in \mathbb{N}^{\mathcal{N}}$ for all products in order to maximize the expected profit from T customers arriving sequentially. This optimization problem can be formulated as

$$\max_{\mathbf{q}\in\mathbb{N}^{\mathcal{N}}} \mathcal{P}_T(\mathbf{q}) := R_T(\mathbf{q}) - \mathbf{c}^\top \mathbf{q}, \tag{2.1}$$

where for $t \in \{1, \ldots, T\}$, we use $R_t(\mathbf{q})$ to denote the expected revenue generated from t customers with inventory $\mathbf{q}$. For any inventory level $\mathbf{q}$, $R_t(\mathbf{q})$ is defined recursively as follows: $R_0(\mathbf{q}) = 0$ and for $t \in \{1, \ldots, T\}$,

$$R_t(\mathbf{q}) = \sum_{i \in S_{\mathbf{q}}} \pi(i, S_{\mathbf{q}}) \cdot (p_i + R_{t-1}(\mathbf{q} - \mathbf{e}_i)) + \pi(0, S_{\mathbf{q}}) \cdot R_{t-1}(\mathbf{q}), \tag{2.2}$$

where $\pi(0, S_{\mathbf{q}})$ is the no-purchase probability and $S_{\mathbf{q}} := \{i \in \mathcal{N} : q_i > 0\}$ denotes the assortment of products induced by $\mathbf{q}$. Here, $\mathbf{e_i}$ is the i^{th} vector of the canonical basis. Note that $\mathbf{c}^\top \mathbf{q} = \sum_{i \in \mathcal{N}} c_i q_i$ is the cost of inventory $\mathbf{q}$, $R_T(\mathbf{q})$ is the expected revenue from all T customers and $\mathcal{P}_T(\mathbf{q})$ refers to the expected profit from T customers.

Our work focuses on the inventory planning problem under general choice models. Specifically, we consider a general choice model that satisfies a natural assumption of substitutability where the choice probability of a product does not increase if we add more products to the assortment. Formally, we assume that for any $A \subseteq B \subseteq \mathcal{N}$, and any $i \in A$, $\pi(i, A) \geq \pi(i, B)$. Solving this problem is challenging in general. In fact, it has been shown in Mahajan and van Ryzin (2001) that the revenue function is not quasi-concave in this setting. Our objective is to find an algorithm with near optimal regret. Given a demand distribution π the regret incurred by an algorithm ϕ is defined as

$$\mathcal{R}_\pi^\phi(T) = \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{\phi(\pi)})\right], \tag{2.3}$$

where $\mathbf{q}^*$ is the optimal solution of problem (2.1), $\mathbf{q}^{\phi(\pi)}$ is the solution returned by the algorithm ϕ and the expectation takes into account any randomness in the algorithm ϕ. Evaluating the solution of the dynamic program (2.2) is a hard task. For this reason, we consider the fluid relaxation of problem (2.1) in which we ignore demand randomness.

$$\max_{S \subseteq \mathcal{N}, \mathbf{q}} \quad T \cdot \sum_{i \in S} p_i \pi(i, S) - \mathbf{c}^\top \mathbf{q} \tag{2.4a}$$

$$\text{subject to} \quad T \cdot \pi(i, S) \leq q_i, \qquad \forall i \in \mathcal{N} \tag{2.4b}$$

$$q_i \geq 0, \qquad \forall i \in \mathcal{N}. \tag{2.4c}$$

Problem (2.4) is closely related to the static unconstrained assortment selection problem where the goal is to determine a subset of products from the universe $\mathcal{N}$ and offer it to a single customer in order to maximize the expected profit under the choice model π. This problem is formally defined as

$$\max_{S \subset \mathcal{N}} \sum_{i \in S} (p_i - c_i) \pi(i, S). \tag{2.5}$$

Let S^* be the optimal solution of problem (2.5). We refer to S^* as the optimal unconstrained assortment. The solution of problem (2.4) is given by S^* and $q_i := T \cdot \pi(i, S^*)$ for all $i \in \mathcal{N}$. Moreover, the optimal objective value of problem (2.4) is equal to T times the optimal objective value of problem (2.5). The total expected demand captured by this inventory decision is

$$D^*(T) := T \cdot \sum_{i \in S^*} \pi(i, S^*).$$

Let z_{fluid} denote the optimal objective value of problem (2.4), in particular,

$$z_{\mathsf{fluid}} := T \cdot \sum_{i \in S^*} (p_i - c_i)\pi(i, S^*).$$

Note that z_{fluid} is an upper bound on (2.1). Formally, the following holds.

Proposition 2.1. $z_{\mathsf{fluid}} \geq \mathcal{P}_T(\mathbf{q}^*)$.

The proof of Proposition 2.1 is presented in Appendix A. Note that for arbitrary choice models, solving the static assortment problem (2.5) can be challenging. We consider approximations of the choice model for which this problem is tractable. In particular, we approximate customer preferences by a Markov Chain choice model. More specifically, starting from a general choice process π, we define the Markov Chain parameters as follows

$$\lambda_i = \pi(i, \mathcal{N}), \qquad \forall i \in \mathcal{N},$$
$$\rho_{ij} = \frac{\pi(j, \mathcal{N} \setminus \{i\}) - \pi(j, \mathcal{N})}{\pi(i, \mathcal{N})}, \qquad \forall i, j \in \mathcal{N}.$$

Here, λ_i denotes the probability that the first choice product for a random customer is product i, and ρ_{ij} denotes the probability of substituting from product i to product j if product i is not available. For any pair-wise disjoint subsets $U, V, W \subset \mathcal{N}_+$, we denote by $\mathbb{P}_i(U \prec V \prec W)$ the probability that starting from $i \in \mathcal{N}$, the customer visits a product in U before any product in $V \cup W$, and visits a product in V before any product in W.

3 SAA algorithm of Gallego and Kim (2020) and its limitations

In this section, we discuss the details and limitations of the SAA based algorithm by Gallego and Kim (2020) for inventory optimization problems. In particular, we show that the worst-case regret of the algorithm scales as $\Omega(T)$.

The SAA-based algorithm presented in Gallego and Kim (2020) solves the inventory optimization problem (2.1) with a choice model approximated by the above surrogate MC choice model. They first consider L samples of first choice demand $(\mathbf{D}^\ell)_{\ell \in [L]}$ where for each $\ell \in [L]$ and $i \in \mathcal{N}$, D_i^ℓ is defined as the number of customers in scenario ℓ whose first choice is product i over all $\mathcal{N}$. Note that $[L] := \{1, \ldots, L\}$ and $\mathbf{D}^\ell \in \mathbb{N}^{\mathcal{N}}, \forall \ell \in [L]$. For any first-stage inventory $\mathbf{q}$, the sales of product $i \in \mathcal{N}$ in scenario $\ell \in [L]$ are approximated with the following set of constraints,

$$x_i^\ell + y_i^\ell - \sum_{k \in \mathcal{N}, k \neq i} \rho_{ki} y_k^\ell = D_i^\ell \tag{3.1}$$

$$x_i^\ell \leq q_i. \tag{3.2}$$

Here, x_i^ℓ denotes the sales of product i in scenario ℓ, y_i^ℓ is the total number of customers substituting out of i in scenario ℓ and constraints (3.1) are the flow-balance equations based on the MC model. The constraint (3.2) ensures that the sales of product i in scenario ℓ are less than the inventory q_i. The SAA problem is formulated as follows,

$$\max_{\mathbf{x}^\ell, \mathbf{y}^\ell, \mathbf{q}} \quad \frac{1}{L} \sum_{\ell=1}^{L} \mathbf{p}^\top \mathbf{x}^\ell - \mathbf{c}^\top \mathbf{q} \tag{3.3a}$$

$$\text{subject to} \quad x_i^\ell \leq q_i, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L], \tag{3.3b}$$

$$x_i^\ell + y_i^\ell - \sum_{k \in \mathcal{N}, k \neq i} \rho_{ki} y_k^\ell = D_i^\ell, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L], \tag{3.3c}$$

$$x_i^\ell \geq 0, \; y_i^\ell \geq 0, \; q_i \geq 0, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L]. \tag{3.3d}$$

Limitation of the formulation. To model SOBS exactly, for each product $i \in \mathcal{N}$, the solution of the linear program (3.3) must satisfy that for every scenario ℓ,

$$x_i^\ell < q_i \;\Rightarrow y_i^\ell = 0.$$

However, (3.3) does not include such a constraint. Consequently, in some scenarios, the formulation allows a more profitable product to be assigned to a customer even if a less profitable but more preferred product is available. This can lead to a significant over-estimation of the revenue and a highly sub-optimal inventory decision. In particular, we show that this limitation induces a linear regret for the algorithm in the worst case.

3.1 $\Omega(T)$-regret for SAA approach of Gallego and Kim (2020)

We present a family of instances where the SAA based algorithm in Gallego and Kim (2020) has a linear, $\Omega(T)$-regret. The number of scenarios in our family of instances depends on the number of customers unlike the setting considered in Gallego and Kim (2020). The approximation bounds for SAA based algorithms typically improve as we increase the number of sampled scenarios (Kleywegt et al. (2002)). However, counter-intuitively, in this case, as the number of sampled scenarios increases, the algorithm leads to a solution with a significantly worse regret with high probability. In particular, we have the following theorem.

Theorem 3.1. *There exists a family of instances of the Markov Chain based choice model π, such that for L and T sufficiently large, the regret of the SAA algorithm is linear in T, i.e.,*

$$R_{\pi}^{SAA}(T) = \Omega(T).$$

We would like to note again that Gallego and Kim (2020) show that SAA is asymptotically optimal for the setting of constant number of products N and constant number of scenarios L that do not depend on the number of customers T which goes to infinity. Note that for T customers, the variance of the demand depends on T. Therefore, for the SAA objective to be a good approximation of the true objective, the number of scenarios should also depend on T (see Kleywegt et al. (2002)). This is not the case in the setting of asymptotic optimality considered in Gallego and Kim (2020) and therefore, the objective in their formulation can be significantly different from the true objective for their solution.

Proof of Theorem 3.1. Fix N. We consider the following instance with $(N+1)$ products. The arrival probabilities $(\lambda_i)_{i\in\mathcal{N}_+}$ are given by: $\lambda_1 = \frac{1}{4}$, $\lambda_0 = \frac{3}{4}$ and $\lambda_i = 0$ for $i \in \{2, \dots, N+1\}$. If product 1 is not available, the customer substitutes to any product among $\{2, \dots, N+1\}$ with probability $\frac{1}{N}$, i.e., $\rho_{1j} = \frac{1}{N}$ for all $j \in \{2, \dots N+1\}$. Moreover, $\rho_{j0} = 1$ for all $j \in \{2, \dots, N+1\}$. For $i \in \mathcal{N}$, the prices and costs are given by,

$$p_i = \begin{cases} \frac{1}{4} & \text{if } i = 1, \\ 1 & \text{if } i \in \{2, \dots, N+1\}. \end{cases}$$

$$c_i = \begin{cases} 0 & \text{if } i = 1, \\ \frac{1}{2} & \text{if } i \in \{2, \dots, N+1\}. \end{cases}$$

The instance is presented in Figure 1. We refer to this instance as the cannibalizing instance.

In this instance, the optimal unconstrained assortment for a single customer is $\{2, \dots, N+1\}$. We first characterize the SAA solution when T and L are sufficiently large. In particular, we show that the SAA solution carries $\frac{T}{4}$ units for the highly sub-optimal product 1 with high probability. More formally, let $\mathbf{q}^{\text{SAA}}$ denote the optimal solution of problem (3.3) and let $\mathcal{E}$ be the event defined as follows:

$$\mathcal{E} = \left\{ q_1^{\text{SAA}} \geq \frac{T}{4} - 2\sqrt{\frac{T^2 \log(2T)}{L}} \right\}. \tag{3.4}$$

Lemma 3.2. *For $T \geq 2$ and $L \geq 4e^{2T} \log(2T)$ we have, $\mathbb{P}(\mathcal{E}) \geq 1 - \frac{1}{T}$.*

The proof of Lemma 3.2 is deferred to Appendix B. The above lemma is a consequence of the approximation of SOBS in the SAA formulation (3.3). The formulation does not guarantee that

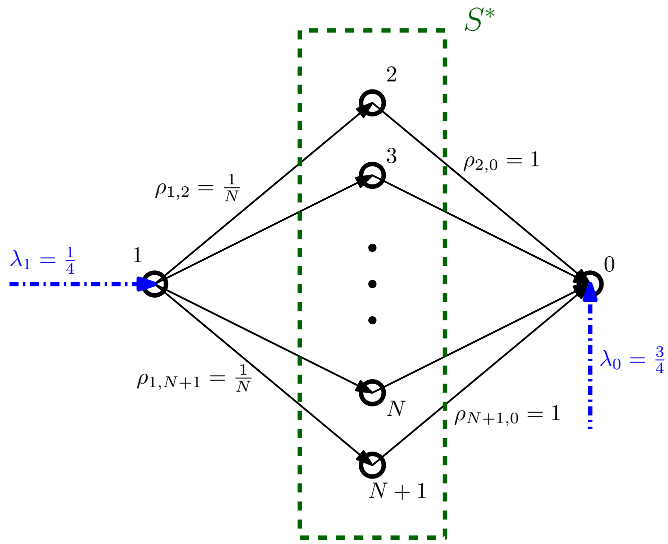

Figure 1: Cannibalizing instance

in each scenario, the most preferred product will be assigned to all customers. This can lead to adding inventory in products that have low cost but can cannibalize the revenue of more profitable ones without hurting the SAA objective value.

Specifically, in the instance of Figure 1, we show that with high probability (over random L scenarios), the SAA algorithm stocks $\frac{T}{4}$ units of product 1. The SAA formulation does not constrain product 1 to be be fully exhausted before any demand substitutes to $\{2, \ldots, N+1\}$. It rather allows substitution even when full inventory of product 1 is available. However, the true objective value decreases significantly which leads to $\Omega(T)$ regret. Formally, we show the following.

Lemma 3.3. *For L and T sufficiently large and $L \geq T$, we have*

$$\mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{SAA}) \mid \mathcal{E}\right] \geq \frac{T}{16} - \frac{7}{2}\sqrt{N^3 T \log(2T)} - 2.$$

The proof of Lemma 3.3 is presented in Appendix C. Finally, we decompose the regret by conditioning on the event $\mathcal{E}$ and its complement $\mathcal{E}^c$. We obtain that,

$$\begin{aligned}
\mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{\text{SAA}})\right] &= \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{\text{SAA}}) \mid \mathcal{E}\right]\mathbb{P}(\mathcal{E}) + \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{\text{SAA}}) \mid \mathcal{E}^c\right]\mathbb{P}(\mathcal{E}^c) \\
&\overset{(a)}{\geq} \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{\text{SAA}}) \mid \mathcal{E}\right]\mathbb{P}(\mathcal{E}) - \frac{T}{2} \cdot \mathbb{P}(\mathcal{E}^c) \\
&\overset{(b)}{\geq} \left[\frac{T}{16} - \frac{7}{2}\sqrt{N^3 T \log(2T)} - 2\right]\mathbb{P}(\mathcal{E}) - \frac{T}{2} \cdot \mathbb{P}(\mathcal{E}^c) \\
&\overset{(c)}{\geq} \left[\frac{T}{16} - \frac{7}{2}\sqrt{N^3 T \log(2T)} - 2\right]\left(1 - \frac{1}{T}\right) - \frac{1}{2},
\end{aligned}$$

where (a) holds since $\mathcal{P}_T(\mathbf{q}^*)$ is not a random variable, so $\mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) \mid \mathcal{E}^c\right] = \mathcal{P}_T(\mathbf{q}^*)$ and by optimality of $\mathbf{q}^*$, $\mathcal{P}_T(\mathbf{q}^*) \geq \mathcal{P}_T(\mathbf{0}) = 0$. Furthermore, $\mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^{\mathrm{SAA}}) \mid \mathcal{E}^c\right] \leq \frac{T}{2}$ as the maximum profit achievable on any selling season is $\frac{T}{2}$, which corresponds to the profit generated by selling the most profitable product, which yields profit $\frac{1}{2}$, to all customers. Moreover, (b) is a consequence of Lemma 3.3. Lastly, (c) is derived from Lemma 3.2. We obtain that for T sufficiently large,

$$\mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{\mathrm{SAA}})\right] \geq \frac{T}{16} + o(T). \qquad \square$$

3.2 SAA algorithm and large number of products

In this section, we illustrate the limitation of the SAA algorithm in Gallego and Kim (2020) when there is a large number of products in the universe. Note that the SAA formulation (3.3) only considers stochasticity in first choice demand in the sampled scenarios. Any subsequent substitution is modeled by a fluid approximation in the formulation. As a result, there are instances where the SAA solution reduces to the inventory solution given by problem (2.4). Recall that this problem does not address the stochasticity of the demand. It only carries an inventory proportional to the expected demand on each product in the optimal unconstrained assortment. This can lead to significantly suboptimal profit, especially when the number of products is large.

Consider the instance described in Figure 2. We refer to this example as the pooling example. The arrival probabilities are $\lambda_1 = 1$ and $\lambda_i = 0$ for $i \in \{2, \ldots, N+1\} \cup \{0\}$. The transition probabilities satisfy, $\rho_{1j} = \frac{1}{N}$ for all $j \in \{2, \ldots, N+1\}$, $\rho_{j,N+2} = 1$ for all $j \in \{2, \ldots, N+1\}$ and $\rho_{N+2,0} = 1$. We define prices and costs as follows:

$$p_i = \begin{cases} \frac{1}{2} & \text{if } i \in \{1, \ldots, N+1\}, \\ \frac{1}{2} - \epsilon & \text{if } i = N+2. \end{cases}$$

$$c_i = \begin{cases} \frac{1}{2} & \text{if } i = 1, \\ \frac{3}{8} & \text{if } i \in \{2, \ldots, N+2\}. \end{cases}$$

In this instance, there is no stochasticity in the first choice demand. Therefore the SAA formulation (3.3) reduces to the fluid problem (2.4). In particular, this implies the following lemma.

Proposition 3.4. *Consider the instance described in Figure 2. An optimal solution of problem* (3.3) *for this instance is given by,*

$$q_i^{SAA} = \begin{cases} 0 & \text{if } i \in \{1, N+2\} \\ \frac{T}{N} & \text{if } i \in \{2, \ldots, N+1\}. \end{cases}$$

The instance presented in Figure 2 materializes the trade-off between revenue generated and risk of left-over. In this example, products among $\{2, \ldots, N+1\}$ generate the maximum revenue.

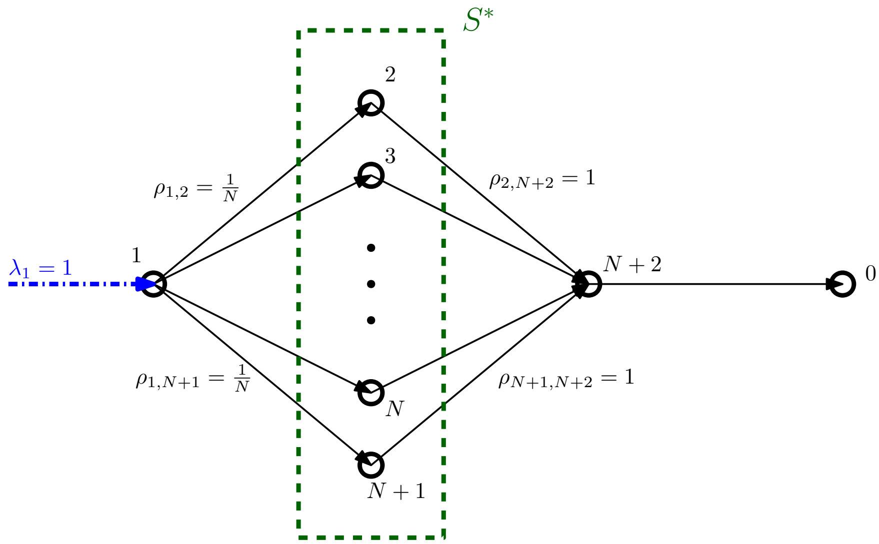

Figure 2: Pooling example

However, as N grows large, the risk of left-over units in these products grows. Product $N + 2$ is slightly suboptimal in comparison. However, it can satisfy any demand that substitutes out of product among $\{2, \ldots, N + 1\}$. Therefore, product $N + 2$ is a good candidate to pool risk related to stock-out and capture demand that was not satisfied by the optimal assortment $\{2, \ldots, N + 1\}$. For ϵ sufficiently small and N sufficiently large, we show that the optimal inventory solution is to carry order T units of product $N + 2$ (see Proposition 4.5). Such risk-pooling is not captured by the SAA problem (3.3) as it does not account for stochasticity of demand for products among $\{2, \ldots, N + 1\}$. We present a detailed proof of Proposition 3.4 in Appendix E.

4 Our approach: Lower Confidence Bound SAA

In this section, we present our algorithm to solve the joint assortment and inventory optimization problem (2.1). We first discuss the intuition for our algorithm and present the key components that address the limitations in Gallego and Kim (2020).

Approximating Stock-out based Substitution (SOBS). Recall that SAA allows arbitrary assignments of products to customers even if a more preferred product is available. In other words, it is possible that in the optimal solution of (3.3) there exists some scenario $\ell \in [L]$ and some product $i \in \mathcal{N}$ such that $y_i^\ell > 0$ even when $x_i^\ell < q_i$. This can lead to significant errors as we may include inventory for some product $i \in \mathcal{N}$ that cannibalizes the revenue for more profitable products as we illustrate in Theorem 3.1. Ideally, we could model exactly SOBS by including the constraint,

$$(x_i^\ell - q_i) \cdot y_i^\ell = 0, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L].$$

However, this constraint is non-convex and therefore, intractable. We consider the following approach to approximate SOBS: we identify a subset of products that could potentially decrease the expected profit significantly from the set of optimal products and explicitly constrain the inventory for such products to be zero. We refer to such products as *cannibalizing products*.

Intuitively, a product is considered to be a cannibalizing product if including it in the assortment could significantly decrease the expected revenue of the assortment. Clearly, the classification of whether the product is cannibalizing or not depends on the rest of the assortment. As an approximation, we consider the set of products that are cannibalizing the profit of the unconstrained optimal assortment by more than an appropriately chosen threshold. In particular, we use the notion of reduced prices defined in Désir et al. (2020) to identify the set of cannibalizing products under the Markov Chain choice model and constrain the inventory of all these products to be 0.

Capturing Demand Stochasticity. Recall SAA approach in Gallego and Kim (2020) can lead to a solution that does not capture stochasticity in demand. It ends up carrying in some cases inventory corresponding to the expected demand over T customers of products in an unconstrained optimal assortment. When the variance of the demand in such products is high, such a solution can be highly suboptimal (as we show in Proposition 4.5). To address this, we consider the following approach to sample scenarios. Let S^* be the unconstrained optimal assortment. For each scenario $\ell \in [L]$, we sample the demand D_i^ℓ for all $i \in S^*$ assuming S^* is offered and available to T customers.

We consider the SAA formulation (3.3) with these $(\mathbf{D}^\ell)_{\ell \in [L]}$ as the initial demands for the products in the unconstrained optimal assortment. Clearly, these samples are approximations of what SOBS will actually lead to as we do not take into consideration stock out events that may occur during the selling season. However, since we have explicitly removed products that are significantly cannibalizing, we can show that the SAA objective provides a good approximation of the true objective value with a near-optimal regret guarantee.

4.1 Our Algorithm

We now describe formally the details of our algorithm. We refer to it as the Lower Confidence Bound SAA (LCB-SAA).

Step 1 (Computing optimal unconstrained assortment). We compute S^*, the solution of the unconstrained assortment optimization problem (2.5) under the Markov Chain choice model. This problem can be efficiently solved using a linear program (see Blanchet et al. (2016)).

Step 2 (Identifying cannibalizing products). For a given assortment S, product i is cannibalizing if the expected revenue from a random customer with first choice as product i decreases by at least γ when we include i in the assortment. This definition is closely related to the notion of reduced price introduced in Désir et al. (2020).

Definition 4.1. (Reduced Price (Désir et al. (2020))) The reduced price of product i given an assortment S is defined as

$$p_i^S = p_i - \sum_{j \in S} \mathbb{P}_i \left(\{j\} \prec S \cup \{0\} \setminus \{j\} \right) \cdot p_j.$$

We define the set of cannibalizing products (denoted by $\mathcal{F}^c$) as

$$\mathcal{F}^c := \{ i \in \mathcal{N} \, : \, p_i^{S^*} \leq -\gamma \},$$

where $\gamma > 0$ is an appropriately chosen threshold. We constrain the inventory of all cannibalizing products $\mathcal{F}^c$ to be equal to 0. The set of feasible products, $\mathcal{F}$ is therefore, given by

$$\mathcal{F} := \{ i \in \mathcal{N} \, : \, p_i^{S^*} > -\gamma \}.$$

Step 3 (Deciding inventory of S^*). Note that for any $i \in S^*$, it is more profitable to sell product i if there is demand as opposed to letting the customer substitute to some other product. Therefore, we would like to satisfy as much demand as possible by products in S^*. The challenge arises from the stochasticity of demand for products in S^* and the risk of unsold inventory. Therefore, we stock a level of inventory which we refer to as the lower confidence bound (LCB) inventory level, that can all be sold with high probability. Specifically, we consider,

$$q_i^{\text{LCB}} = \pi(i, S^*)T - 2\sqrt{\pi(i, S^*)T \log(NT)}, \qquad \forall i \in S^*.$$

This particular choice of inventory, q_i^{LCB} for all $i \in S^*$ ensures that all inventory of S^* is consumed with high probability.

Step 4 (Deciding safety-stock inventory). We solve a second joint assortment and inventory problem which aims at deciding how to allocate safety stock to capture the demand not satisfied by q_i^{LCB} for all $i \in S^*$. More precisely, we consider a surrogate problem approximating SOBS. Let

$$T' := 2 \sum_{i \in S^*} \sqrt{\pi(i, S^*)T \log(NT)}.$$

Note that T' is roughly the expected lost sales if we only stock q_i^{LCB} for all $i \in S^*$. The inventory for this demand is not necessarily allocated to products in S^* due to the risk of high leftover which may hurt the profit. Therefore, we consider a joint assortment and inventory problem for this residual demand that appropriately captures the stochasticity in demand for all $i \in S^*$ and identifies the demand pooling products that better balance the tradeoff between expected revenue and expected leftover inventory cost. We refer to these products as risk pooling products or safety stock.

To decide the allocation of safety stock, we consider a problem with T' i.i.d. customers who choose according to the Markov Chain choice model $(\rho_{ij})_{i,j\in\mathcal{N}}$ as before. However, to appropriately model the demand stochasticity for all $i \in S^*$, we consider a different arrival probability using $\pi(\cdot, S^*)$. In particular, we assume that the first choice for any customer is $i \in S^*$ with probability $\pi(i, S^*)$. Subsequent substitutions occur as before according to the transition matrix ρ.

The variant of SAA to determine the safety stock can now be described as follows. Let $\mathcal{F}^c$ denote the set of cannibalizing products. We constrain the inventory of cannibalizing products to 0, i.e.,

$$q_i = 0, \qquad \forall i \in \mathcal{F}^c.$$

We sample demand scenarios $(D^\ell)_{\ell\in[L]}$ with first choice demand D_i^ℓ for all $i \in S^*$ and for all scenarios $\ell \in [L]$ in this modified Markov Chain. Note that by construction, $D_i^\ell = 0$ for $i \in \mathcal{N} \setminus S^*$. We now consider the following LP to allocate safety-stock on the remaining products in $\mathcal{F}$.

$$\max_{\mathbf{x}^\ell,\mathbf{y}^\ell,\mathbf{q}} \quad \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{p}^\top \mathbf{x}^\ell - \mathbf{c}^\top \mathbf{q} \tag{4.1a}$$

$$\text{subject to} \quad x_i^\ell \le q_i, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L], \tag{4.1b}$$

$$x_i^\ell + y_i^\ell - \sum_{k\in\mathcal{N},k\neq i} \rho_{ki} y_k^\ell = D_i^\ell, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L], \tag{4.1c}$$

$$q_i = 0, \qquad \forall i \in \mathcal{F}^c, \tag{4.1d}$$

$$x_i^\ell \ge 0,\ y_i^\ell \ge 0,\ q_i \ge 0, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L]. \tag{4.1e}$$

Algorithm 1 presents formally the procedure. The parameters of this algorithm are γ and L.

Algorithm 1 LCB-SAA

Compute the unconstrained optimal assortment S^*.
Compute reduced prices $p_i^{S^*}$ for each product $i \in \mathcal{N}$.
for $i \in S^*$ **do**
  $q_i^{\mathrm{LCB}} \leftarrow \pi(i, S^*)T - 2\sqrt{\pi(i, S^*)T \log(NT)}$.
end for
$\mathcal{F} \leftarrow \{i \in \mathcal{N} \mid p_i^{S^*} > -\gamma\}$.
$T' \leftarrow 2\sum_{i\in S^*} \sqrt{\pi(i, S^*)T \log(NT)}$.
Sample $(\mathbf{D}^\ell)_{\ell\in[L]}$ from a multinomial with probability $\pi(\cdot, S^*)$ and T' trials.
Solve LP (4.1) and denote by $\mathbf{q}^{\mathrm{pool}}$ its solution.
return $\mathbf{q}^{\mathrm{LCB}} + \mathbf{q}^{\mathrm{pool}}$.

4.2 Regret bound

In the following theorem, we show that Algorithm 1 incurs a regret of $\tilde{\mathcal{O}}\left(\sqrt{NT}\right)$.

Theorem 4.2. *Let π be a Markov Chain choice model. For $0 \leq \gamma \leq \sqrt{\frac{D^*(T)}{NT^2}}$, define $\mathbf{q}^{LCB\text{-}SAA}$ as the solution returned by Algorithm 1. Then,*

$$z_{\mathsf{fluid}} - \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^{LCB\text{-}SAA})\right] = \mathcal{O}\left(\sqrt{ND^*(T)\log(NT)}\right),$$

where $D^*(T) = T\sum_{i\in S^*}\pi(i, S^*)$.

Note that Theorem 4.2 implies that the regret of our solution is $\tilde{\mathcal{O}}\left(\sqrt{NT}\right)$. This holds because $z_{\mathsf{fluid}} \geq \mathcal{P}_T(\mathbf{q}^*)$ (Proposition 2.1) and the expected demand $D^*(T) \leq T$. We would like to note that in online retail, $D^*(T)$ can be significantly smaller than T because of small conversion rate. Therefore, the bound in Theorem 4.2 can be significantly stronger than $\tilde{\mathcal{O}}\left(\sqrt{NT}\right)$.

We would like to note that the above regret bound is achievable also by a fluid solution that carries inventory equal to the expected demand for all products: $q_i^{\text{fluid}} = \pi(i, S^*)T$ for all $i \in S^*$. However, the fluid solution does not address either of the two main challenges in this problem including *i*) approximating stock-out based substitution, and *ii*) capturing stochasticity of demand to balance the trade-offs between expected revenue and inventory cost. Furthermore, there are family of instances where the expected profit of a fluid solution is significantly worse than the expected profit of the solution given by our algorithm. We show in Proposition 4.5 that the expected profit of the fluid solution is $\Omega(T)$ worse than both the optimal solution and the solution of our algorithm in the instance described by Figure 2.

In contrast, Algorithm 1 explicitly addresses both the fundamental challenges of approximating SOBS and identifying demand pooling products. In particular, Step 4 of deciding the safety stock inventory in our algorithm explicitly identifies products that can pool highly stochastic demand if needed. Adding inventory of products outside of the optimal assortment, S^* runs the risk of cannibalizing the revenue from the optimal products. Therefore, our algorithm explicitly identifies such revenue cannibalizing products and forces inventory of such products to be zero (Step 2). This allows us to design a good tractable approximation of SOBS via a Markov chain choice model. Therefore, even though the regret bound of our solution and the fluid solution are similar, Steps 2 and 4 in Algorithm 1 attempt to address the fundamental complexities in this problem and result in solutions with many desirable properties.

Proof of Theorem 4.2. For the sake of simplicity of notation, denote by $\mathbf{q}$ the inventory level of LCB-SAA. Let $(\mathbf{q}_t)_{t\in\{0,\dots,T\}}$ be a sequence of random variables such that for all $t \in \{1, \dots, T\}$, $\mathbf{q}_t$ represents the inventory level after the purchase decision of customer t and $\mathbf{q}_0 = \mathbf{q}$. The proof consists in lower bounding the profit generated by our policy.

Given a Markov Chain choice model, we associate each customer to a random walk that starts at a certain product according to the arrival probabilities and transition from one product to another according to the transition probabilities until hitting the no-purchase option 0. We say that a customer visits product $i \in \mathcal{N}$ before product $j \in \mathcal{N}$ if the random walk associated to that customer hits product i before hitting product j. For a customer $t \in \{1, \dots, T\}$, we define the following relation $\prec_t$: for $U, V \subset \mathcal{N}_+$, we say that $U \prec_t V$ if and only if there exists a product $i \in U$ such that customer t visits product i before any product in V.

Recall $\mathcal{F}$ is the set of non-cannibalizing products and let $\mathcal{F}^{\text{pool}} := \mathcal{F} \backslash S^*$. We define the sequence of random variables $(X_i^t)_{i \in \mathcal{F}, t \in \{1,\dots,T\}}$ as follows: for any $i \in \mathcal{F}$ and $t \in \{1, \dots, T\}$,

$$X_i^t = \mathbb{1}\{\{i\} \prec_t S^* \cup \{0\} \setminus \{i\}\},$$

and $X_i = \sum_{t=1}^T X_i^t$. Remark that, by definition of $\mathcal{F}^{\text{pool}}$, we have $X_i^t = \mathbb{1}\{\{i\} \prec_t S^* \cup \{0\}\}$ for every $i \in \mathcal{F}^{\text{pool}}$. Consider the sequence of random variables $(Z_i^t)_{i \in \mathcal{F}, t \in \{1,\dots,T\}}$ such that for all $i \in \mathcal{F}$ and $t \in \{1, \dots, T\}$,

$$Z_i^t = \mathbb{1}\{\{i\} \prec_t S_{\mathbf{q}_t} \cup \{0\} \setminus \{i\}\} \cdot \mathbb{1}\{i \in S_{\mathbf{q}_t}\}.$$

Note that $Z_i^t = 1$ if and only if customer t purchases product i. Furthermore, for each $i \in \mathcal{F}$, let $\tilde{S}_i$ denote the total number of purchases of product i. Our first lemma states the following.

Lemma 4.3. *For any realization of customer preferences and for any $i \in S^*$, we have,*

$$\tilde{S}_i \geq \min(X_i, q_i) - \sum_{t=1}^T \sum_{j \in \mathcal{F}^{pool}} Z_j^t \cdot X_i^t \cdot X_j^t.$$

The proof of Lemma 4.3 is presented in Appendix D. We denote by $\tilde{\mathcal{P}}_T$ the profit of LCB-SAA during a selling season. Recall $\mathbf{q}^{\text{LCB}}$ and $\mathbf{q}^{\text{pool}}$ defined in Algorithm 1. Lemma 4.3 implies that the profit satisfies almost surely the following inequality,

$$\begin{aligned}\tilde{\mathcal{P}}_T(\mathbf{q}) &\geq \sum_{i \in S^*} p_i \left(\min(X_i, q_i) - \sum_{t=1}^T \sum_{j \in \mathcal{F}^{\text{pool}}} Z_j^t \cdot X_i^t \cdot X_j^t \right) + \sum_{i \in \mathcal{F}^{\text{pool}}} p_i \tilde{S}_i - \sum_{i \in \mathcal{N}} c_i q_i \\ &\overset{(a)}{=} \sum_{i \in S^*} \left(p_i \cdot \min(X_i, q_i) - c_i q_i^{\text{LCB}} \right) + \sum_{j \in \mathcal{F}^{\text{pool}}} \sum_{t=1}^T Z_j^t \left(p_j - \sum_{i \in S^*} p_i \cdot X_i^t \cdot X_j^t \right) - \sum_{i \in \mathcal{N}} c_i q_i^{\text{pool}},\end{aligned} \tag{4.2}$$

where (a) holds because for every $i \in \mathcal{F}$, $\tilde{S}_i = \sum_{t=1}^T Z_i^t$. Next, we will derive a lower bound for the expectation of each of the three terms in (4.2).

Step 1: Recall for $i \in S^*$, $q_i^{\text{LCB}} = \pi(i, S^*)T - 2\sqrt{\pi(i, S^*)T \log(NT)}$. Let us define the event $\mathcal{B}_i := \{X_i < q_i^{\text{LCB}}\}$, and let $\mathcal{B} := \cup_{i \in S^*} \mathcal{B}_i$. Remark that for $i \in S^*$, X_i is the sum of T i.i.d. Bernoulli random variables with mean $\pi(i, S^*)$. By the multiplicative Chernoff's inequality (see

Theorem 1.10.5 in Doerr (2020)), we get

$$\mathbb{P}(\mathcal{B}_i) = \mathbb{P}\left(X_i \leq \mathbb{E}[X_i]\left(1 - 2\sqrt{\frac{\log(NT)}{\pi(i,S^*)T}}\right)\right) \leq \exp\left(-\frac{2\log(NT)\mathbb{E}[X_i]}{\pi(i,S^*)T}\right) = \frac{1}{N^2T^2},$$

and by a union bound we obtain that,

$$\mathbb{P}(\mathcal{B}) \leq \sum_{i\in S^*} \mathbb{P}(\mathcal{B}_i) \leq \frac{|S^*|}{N^2T^2} \leq \frac{1}{NT^2}. \tag{4.3}$$

Let $\bar{c} := \max_{i\in\mathcal{N}} c_i$ and $\bar{p} := \max_{i\in\mathcal{N}} p_i$, note that,

$$\sum_{i\in S^*} \mathbb{E}\left[p_i \min(X_i, q_i) - c_i q_i^{\mathrm{LCB}} \,\middle|\, \mathcal{B}\right] \geq -\bar{c} \sum_{i\in S^*} q_i^{\mathrm{LCB}} \geq -\bar{c}\cdot T.$$

Moreover, conditionally on $\mathcal{B}^c$ (the complementary event of $\mathcal{B}$), we have

$$\begin{aligned}\sum_{i\in S^*} \mathbb{E}\left[p_i \min(X_i, q_i) - c_i q_i^{\mathrm{LCB}} \,\middle|\, \mathcal{B}^c\right] &\overset{(a)}{\geq} \sum_{i\in S^*} (p_i - c_i) q_i^{\mathrm{LCB}} \\ &= T\cdot \sum_{i\in S^*}(p_i - c_i)\pi(i,S^*) - 2\sum_{i\in S^*}(p_i - c_i)\sqrt{\pi(i,S^*)T\log(NT)} \\ &\geq z_{\mathsf{fluid}} - \bar{p}\cdot T',\end{aligned}$$

where (a) holds because for $i \in S^*$, $q_i \geq q_i^{\mathrm{LCB}}$ and conditionally on $\mathcal{B}^c$, we have $X_i \geq q_i^{\mathrm{LCB}}$. Therefore,

$$\begin{aligned}\sum_{i\in S^*} \mathbb{E}\left[p_i \min(X_i, q_i) - c_i q_i^{\mathrm{LCB}}\right] &\geq (z_{\mathsf{fluid}} - \bar{p}\cdot T')\,\mathbb{P}(\mathcal{B}^c) - \bar{c}\cdot T\cdot \mathbb{P}(\mathcal{B}) \\ &\overset{(a)}{\geq} (z_{\mathsf{fluid}} - \bar{p}\cdot T') - \frac{z_{\mathsf{fluid}}}{NT^2} - \frac{\bar{c}}{NT} \\ &\overset{(b)}{\geq} (z_{\mathsf{fluid}} - \bar{p}\cdot T') - \frac{\bar{p}+\bar{c}}{NT},\end{aligned} \tag{4.4}$$

where (a) follows from (4.3) and (b) holds because $z_{\mathsf{fluid}} \leq \bar{p}\cdot T$.

Step 2: For every $i \in S^*$, $j \in \mathcal{F}^{\mathrm{pool}}$ and $t \in \{1,\ldots,T\}$, we have that X_i^t, X_j^t and Z_j^t are binary random variables, therefore,

$$\mathbb{E}\left[Z_j^t\left(p_j - \sum_{i\in S^*} p_i\cdot X_i^t\cdot X_j^t\right)\right] = p_j\cdot \mathbb{P}\left(Z_j^t = 1\right) - \sum_{i\in S^*} p_i\cdot \mathbb{P}\left(X_i^t = 1\,,\, X_j^t = 1\,,\, Z_j^t = 1\right).$$

Moreover, for every $t \in \{1,\ldots,T\}$, let $(W_n^t)_{n\in\mathbb{N}}$ be the random walk associated to customer t on the Markov Chain. For every $j \in \mathcal{F}^{\mathrm{pool}}$, consider the stopping time (in which we omit the dependence in customer t) $\tilde{T}_1^j := \min\{n \in \mathbb{N} \mid W_n^t \in S_{\mathbf{q}_t} \cup S^* \cup \{j, 0\}\}$ and for every $i \in S^*$ define the stopping time

$\tilde{T}_2^i := \min\{n \in \mathbb{N} \mid W_n^t \in S^* \cup \{0\}\}$. Remark that, $\{X_j^t = 1\} \cap \{Z_j^t = 1\} = \{W_{\tilde{T}_1^j}^t = j\} \cap \{j \in S_{\mathbf{q}_t}\}$ for all $j \in \mathcal{F}^{\text{pool}}$. Similarly, $\{X_i^t = 1\} = \{W_{\tilde{T}_2^i}^t = i\}$ for all $i \in S^*$. Thus, for all $i \in S^*$ and $j \in \mathcal{F}^{\text{pool}}$,

$$\begin{aligned}
\mathbb{P}\left(X_i^t = 1\,,\, X_j^t = 1\,,\, Z_j^t = 1\right) &= \mathbb{P}\left(X_i^t = 1 \mid X_j^t = 1\,,\, Z_j^t = 1\right) \cdot \mathbb{P}\left(X_j^t = 1\,,\, Z_j^t = 1\right) \\
&\overset{(a)}{=} \mathbb{P}\left(W_{\tilde{T}_2^i}^t = i \mid W_{\tilde{T}_1^j}^t = j\right) \cdot \mathbb{P}\left(X_j^t = 1\,,\, Z_j^t = 1\right) \\
&\overset{(b)}{=} \mathbb{P}\left(W_{\tilde{T}_2^i}^t = i \mid W_0^t = j\right) \cdot \mathbb{P}\left(X_j^t = 1\,,\, Z_j^t = 1\right) \\
&= \mathbb{P}_j\left(\{i\} \prec_t S^* \cup \{0\} \setminus \{i\}\right) \cdot \mathbb{P}\left(X_j^t = 1\,,\, Z_j^t = 1\right) \\
&\leq \mathbb{P}_j\left(\{i\} \prec_t S^* \cup \{0\} \setminus \{i\}\right) \cdot \mathbb{P}\left(Z_j^t = 1\right),
\end{aligned}$$

where (a) follows from the independence of the random walk $(W_n^t)_{n\in\mathbb{N}}$ and the event $\{j \in S_{\mathbf{q}_t}\}$. The equality (b) holds because conditional on $\{W_{\tilde{T}_1^j}^t = j\}$ and since $j \notin S^*$, we get that $W_n^t \notin S^* \cup \{0\}$ for any $n \leq \tilde{T}_1^j$, i.e., $\tilde{T}_1^j \leq \tilde{T}_2^i$ for any $i \in S^*$ which implies (b) by the strong Markov property. Using the definition of reduced price we obtain that,

$$\begin{aligned}
\mathbb{E}\left[Z_j^t\left(p_j - \sum_{i\in S^*} p_i \cdot X_i^t \cdot X_j^t\right)\right] &\geq \left(p_j - \sum_{i\in S^*} \mathbb{P}_j\left(\{i\} \prec_t S^* \cup \{0\} \setminus \{i\}\right) \cdot p_i\right) \cdot \mathbb{P}\left(Z_j^t = 1\right) \\
&= p_j^{S^*} \cdot \mathbb{P}\left(Z_j^t = 1\right) \overset{(a)}{\geq} -\gamma \cdot \mathbb{P}\left(Z_j^t = 1\right) \geq -\gamma,
\end{aligned}$$

where (a) follows from the definition of feasible products $\mathcal{F}$. Therefore, we obtain that,

$$\sum_{t=1}^{T} \sum_{j\in\mathcal{F}^{\text{pool}}} \mathbb{E}\left[Z_j^t\left(p_j - \sum_{i\in S^*} p_i \cdot X_i^t \cdot X_j^t\right)\right] \geq -\gamma \cdot N \cdot T \geq -\sqrt{ND^*(T)}, \tag{4.5}$$

where the last inequality follows from $0 \leq \gamma \leq \sqrt{\frac{D^*(T)}{NT^2}}$.

Step 3: We bound the cost of inventory by using the following lemma.

Lemma 4.4. *Let $\mathbf{q}^{pool}$ be an optimal solution of* (4.1) *where the demand scenarios are sampled with T' customers, then $\sum_{i\in\mathcal{N}} c_i q_i^{pool} \leq \bar{p} \cdot T'$ almost surely.*

The proof of Lemma 4.4 is presented in Appendix D. Finally, we conclude that,

$$\begin{aligned}
\mathbb{E}\left[\tilde{\mathcal{P}}_T(\mathbf{q})\right] &\overset{(a)}{\geq} z_{\mathsf{fluid}} - \bar{p} \cdot T' - \frac{\bar{p} + \bar{c}}{NT} - \sqrt{ND^*(T)} - \bar{p} \cdot T' \\
&\overset{(b)}{\geq} z_{\mathsf{fluid}} - \frac{\bar{p} + \bar{c}}{NT} - \sqrt{ND^*(T)} - 4\bar{p}\sqrt{ND^*(T)\log(NT)},
\end{aligned}$$

where (a) is a consequence of the decomposition derived in (4.2) and of the bounds computed in (4.4), (4.5) and Lemma 4.4. Moreover, (b) follows from the Cauchy-Schwarz inequality, i.e.,

$T' = 2\sum_{i\in S^*}\sqrt{\pi(i,S^*)T\log(NT)} \leq 2\sqrt{N\log(NT)T\sum_{i\in S^*}\pi(i,S^*)}$. □

4.3 Practical implementation for large number of products

Algorithm 1 is appealing because of the interpretability of the inventory solution that also simplifies our analysis of the regret bound. While the value we set for LCB in Algorithm 1 ensures a provable sub-linear regret for our approach; we can practically decide on inventory levels by using SAA. Algorithm 2 presents a practical implementation of LCB-SAA. It differs from Algorithm 1 in that we do not explicitly fix the inventory of the products in the unconstrained optimal assortment to the lower confidence bound of their respective demand. Rather we only specify the unconstrained optimal assortment and the set of cannibalizing products whose inventory is constrained to 0. The LP formulation (4.1) jointly optimizes the inventory of the products in the optimal assortment and allocates safety stock for products that are not cannibalizing based on the sampled scenarios. Recall that L and γ are parameters of the algorithm.

Algorithm 2 Practical LCB-SAA

Compute the unconstrained optimal assortment S^*.
Compute reduced prices $p_i^{S^*}$ for each product $i \in \mathcal{N}$.
$\mathcal{F} \leftarrow \{i \in \mathcal{N} \mid p_i^{S^*} > -\gamma\}$.
Sample $(\mathbf{D}^\ell)_{\ell\in[L]}$ from a multinomial with probability $\pi(\cdot, S^*)$ and T trials.
Solve LP (4.1) and denote by $\mathbf{q}^{\text{final}}$ its solution.
return $\mathbf{q}^{\text{final}}$.

We show that in the setting described in Section 3.2, the performance of our approach is significantly better than the performance of the solution of problem (2.4). More formally, we show the following result.

Proposition 4.5. *Consider the Markov Chain instance, π described in Figure 2. We consider the policy which prescribes the inventory level that solves* (2.4)*, and refer to it by fluid policy. We refer to the policy given by Algorithm 2 by LCB-SAA. For $\epsilon = \frac{1}{\sqrt{T}}$, we have*

$$\begin{aligned} R_\pi^{LCB\text{-}SAA}(T) &= \tilde{\mathcal{O}}\left(\sqrt{T}\right), \\ R_\pi^{fluid}(T) &= \Omega\left(T\right). \end{aligned}$$

The proof is deferred to Appendix F. We note that, as shown in Section 3.2, SAA has the same solution as the fluid policy of Proposition 4.5. This result formally shows that accounting for stochasticity in the optimal unconstrained assortment leads to more accurate solutions which outperform approaches that do not account for stochasticity. In particular, we see that pooling safety-stock by allocating inventory to product $N+2$ is crucial when the number of products is large. This aspect of the demand cannot be captured by a deterministic model or by a model that only considers stochasticity in first choice.

5 Customer choice process from query to purchase

In this section, we present a model for the customer journey on an online store. In this context, customer interactions with the retailer are complex and occur in multiple steps. Our goal is to present a feature based model that captures the various elements of these complex interactions resulting in customer choices that produce heavy-tailed demand. The resulting choice model is a complex parametric formulation and even the assortment optimization problem (without inventory decisions) is intractable. Therefore, we consider a Markov Chain based approximation of this parametric choice model for the joint assortment and inventory optimization.

We consider the setting where each customer interaction with the online retailer begins with a search query. Let $\mathcal{Q}$ denote the universe of search queries and Ψ denote a mapping from $\mathcal{Q}$ to a subset of products i.e., for any $q \in \mathcal{Q}$, $\Psi(q)$ is a subset of candidate products associated to the query q and from which the online retailer selects a subset of product to display. The customer is shown a random ranking of a subset of products in $\Psi(q)$ where a product with high estimated utility has a higher probability to appear at top in the ranked results. We refer to this as the *ranked impression set* or simply the impression set. The customer observes the impression set and selects a *consideration set* of products. We assume that the customers only observe partial attributes of products in the impression page and select a subset of these products (referred to as the consideration set) to observe detailed features and make the eventual selection.

For each product $i \in \mathcal{N}$, we denote by $\mathbf{x_i}$ (resp. $\mathbf{y_i}$) the vector of features observed during the initial impression (resp. product details page). For instance feature $\mathbf{x_i}$ may represent price range or rating, whereas, $\mathbf{y_i}$ includes the real price and other detailed features. This models a realistic setting where the impression page only displays partial attributes. Below we present more details of the choice process for a customer with query $q \in \mathcal{Q}$ and subset of products $\Psi(q)$ mapped to query q.

We assume i.i.d. customers arrive on the online store and have two unknown utility vectors, $\boldsymbol{\beta}$ for click utility and $\boldsymbol{\theta}$ for purchase utility.

Construction of ranked impression set. For any query $q \in \mathcal{Q}$, we sample a random ranking of $\Psi(q)$ as follows. The ranking is such that higher utility products (based on historical utilities) have a higher probability of appearing at the top. Let $\hat{\boldsymbol{\beta}}$ be a utility vector used by the decision-maker as a proxy for the true preference of customers. For instance, $\hat{\boldsymbol{\beta}}$ may be computed based on previous historical data. We construct a sequence of random variables $(\tilde{s}_j)_{j\in\{1,\ldots,K\}}$, where K is the maximum number of product displayed in the impression set. The sequence is constructed sequentially such that for each product $i \in \Psi(q)$,

$$\mathbb{P}(\tilde{s}_1 = i) = \frac{e^{\hat{\boldsymbol{\beta}}^\top \mathbf{y_i}}}{\sum_{k \in \Psi(q)} e^{\hat{\boldsymbol{\beta}}^\top \mathbf{y_k}}},$$

and for each rank $j \in \{2, \ldots, K\}$ product $i \in \Psi(q)$ is sampled without replacement with probability

$$\mathbb{P}(\tilde{s}_j = i) = \frac{e^{\hat{\boldsymbol{\beta}}^\top \mathbf{y_i}}}{\sum_{k \in \Psi(q) \setminus \{\tilde{s}_1, \ldots, \tilde{s}_{j-1}\}} e^{\hat{\boldsymbol{\beta}}^\top \mathbf{y_k}}}.$$

Consideration set. We next describe the process by which the customer selects a consideration set C. We assume that the customer examines the ranked impression set in the order of the ranking. Product $i \in \mathcal{N}$ is added to the consideration set with a probability that depends only on the partial attributes $\mathbf{x_i}$. Formally, we have that

$$\mathbb{P}(i \in C \mid \text{customer examines } i) = \frac{e^{\boldsymbol{\theta}^\top \mathbf{x_i}}}{1 + e^{\boldsymbol{\theta}^\top \mathbf{x_i}}}.$$

We also assume that the customer is impatient and after each click stops browsing more products with probability

$$\frac{1}{1 + e^{\mu}},$$

where μ is the impatience parameter. In our setting we assume that μ is fixed but one may consider settings for which μ depends on the current consideration set either through cardinality or aggregated utility. Therefore, the number of products in the consideration set is geometrically distributed.

Purchase decision within the consideration set. The customer observes all attributes of the products in the consideration set and makes the eventual selection. We model this choice according to a MNL choice model. In particular, customer selects product $i \in C$ with probability,

$$\mathbb{P}(i \text{ is selected} \mid C \text{ is offered}) = \frac{e^{\boldsymbol{\beta}^\top \mathbf{y_i}}}{1 + \sum_{k \in C} e^{\boldsymbol{\beta}^\top \mathbf{y_k}}}.$$

5.1 Parameter estimation

We show that the parameters $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ and μ of the above choice model can be computed efficiently from historical observations. We assume that we can infer the consideration set of the customer from his browsing history. Specifically, we observe clicks to detail pages of the product which is a good proxy for the product being in the consideration set. More specifically, we observe for all customers t, $(\Psi(q_t), \sigma_t, S_t, C_t, i_t, e_t)$, where q_t is the query sent to the website, σ_t is a mapping from S_t to $\{1, \ldots, |S_t|\}$ that defines the ranking in the impression set S_t where, 1 corresponds to the highest ranked product whereas $|S_t|$ is the ranking of the last product in the impression set. We use i_t to denote the selected product by the customer among the consideration set C_t. Finally, e_t is a binary variable which is equal to 1 if the customer starts the purchase phase because he reached the end of the impression set or 0 if he starts the purchase phase because of impatience.

The goal is to estimate the vectors of parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ along with the impatience parameter μ. Remark, that $\hat{\boldsymbol{\beta}}$ is not an unknown parameter of the model. It represents a parameter generated from historical data in order to construct a ranked impression set.

The estimation problem requires to maximize the log-likelihood under these parameters. In particular, let $\mathcal{H}_T = \{(\Psi(q_t), \sigma_t, S_t, C_t, i_t, e_t), \forall 1 \leq t \leq T\}$ be the set of previously observed data, then the log-likelihood is defined as

$$\mathcal{L}(\mu, \boldsymbol{\theta}, \boldsymbol{\beta}) := \log \left(\mathbb{P} \left(\mathcal{H}_T | \mu, \boldsymbol{\theta}, \boldsymbol{\beta} \right) \right).$$

Proposition 5.1. *When customers are independent and identically distributed, the log-likelihood $\mathcal{L}$ is jointly concave in the unknown parameters $\mu, \boldsymbol{\theta}$ and $\boldsymbol{\beta}$.*

The proof of Proposition 5.1 is deferred to Appendix G. Proposition 5.1 implies that the maximum likelihood estimator can be computed efficiently by solving a convex optimization problem.

5.2 Parsimonious heavy-tail distribution

We present empirical evidences to show that our choice model preserves some important characteristics of the demand observed in practice. In particular, we focus on the heavy-tailed nature of demand which arises in practice, especially in e-commerce settings where retailers carry a broad selection of products and a large number of low demand products account for a significant proportion of total sales (Brynjolfsson et al. (2006); Goel et al. (2010)).

Besides its interpretability and expressiveness, our choice process is also appealing because it allows a parsimonious generation of a heavy-tail distribution for clicks and purchases. When query distribution is heavy-tailed, our choice process preserves the heavy-tail property when Ψ maps different queries to disjoint sets. More precisely, assume that the query space $\mathcal{Q}$ is finite and that for any $q, q' \in \mathcal{Q}$, such that $q \neq q'$ we have, that $|\Psi(q) \cap \Psi(q')|$ is relatively small. Then, if the distribution of queries is heavy-tailed, we observe that distributions of click and purchase are also heavy-tailed. We represent in Figure 3 purchase distribution and click distribution for an instance of our model with 50 products and 5 query types. The query is generated according to a power-law and there is no intersection between candidates product. We observe that, in that case, click and purchase distributions have a heavy-tailed nature.

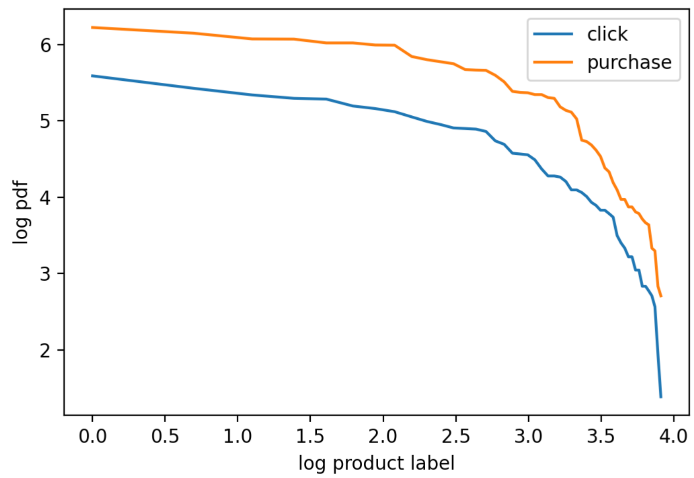

Figure 3: **Heavy-tailed nature of click and purchase:** log-log plot of the purchase and click probability density function.

6 Numerical experiments

We compare the empirical performance of our heuristic with previous approaches in the literature on adversarial instances, synthetic instances as well as instances of the multi-step query to purchase model using parameters that closely capture realistic settings.

6.1 Empirical performance on worst case instances (Theorem 3.1)

We first present numerical results illustrating the performance of SAA on the class of worst-case instances from Theorem 3.1 as described in Figure 1. Our random instances follow the structure of Markov Chain as in Figure 1. However, we explore other settings of the Markov Chain parameters as follows: we consider $\lambda_0 = \lambda_1 = \frac{1}{2}$. The transition probability vector from node 1 to other nodes is sampled uniformly in the unit simplex (Dirichlet distribution with parameters 1). The prices of products 2 to $N + 1$ are sampled from a triangular distribution supported on $[10, 50]$ with mode equal to 40. The price of product 1 is 1. All costs are computed using a margin of 33% except c_1 which is set to 0.

We compare the performance of LCB-SAA and SAA with $L = 500$ scenarios, product universe of size in $\{10, 20\}$ and time horizons of length in $\{100, 200, 500\}$. For each policy, the performance, also referred to as profit gap, is measured by the ratio between the profit of the policy and the value of the LP relaxation defined in problem (2.4).

The average profit gap is computed numerically by running these algorithms on 400 randomly generated instances. For each instance, the average profit of a policy is computed by simulating 100 selling-seasons. We observe that LCB-SAA significantly outperforms SAA on this class of instance. This is even more visible on instances where the ratio between the number of customers and the number of products is large. Table 1 presents the $25^{th}, 50^{th}$ and 75^{th} percentiles of the distribution of profit gap (in percentage).

		$T = 100$			$T = 200$			$T = 500$		
Products	ALG	25th	50th	75th	25th	50th	75th	25th	50th	75th
N = 10	LCB-SAA	54.8	57.3	59.2	68.3	69.6	71.1	80.4	81.0	82.0
	SAA	6.8	9.9	12.4	25.4	29.3	32.8	52.0	54.0	55.6
N = 20	LCB-SAA	36.6	39.0	40.7	53.4	55.3	57.0	70.8	72.2	72.8
	SAA	3.8	5.2	6.5	17.2	20.3	23.0	44.2	47.2	48.9

Table 1: Percentiles of the profit gap distribution for LCB-SAA and SAA for adversarial instances

6.2 Synthetic instances

The setting described in Section 6.1 compares both approaches on a family of instances that is adversarial for SAA. We now evaluate these policies on a family of Markov Chain choice models for which arrival probabilities are heavy-tailed and prices are negatively correlated with demand. These instances are characterized by arrival probabilities generated by a discretized log-normal distribution.

Products are classified between two price segments, "high price products" and "low price products". Prices are generated from a triangular distribution with the same support, but we choose the mode of high price products to be larger than the one for low price products. In these synthetic instances, we choose the support to be $[10, 50]$ for both categories and the mode for low (resp. high) price products to be 20 (resp. 40). To ensure negative correlation between arrival demand and price, we put product $i \in \mathcal{N}$ in the low price category according to a Bernoulli distribution with parameter λ_i. Costs of products are generated by considering a hypothetical fixed margin of 33%.

Finally, the transition matrix is computed by associating each product $i \in \mathcal{N}$ to a utility vector $\mathbf{v_i} := \frac{\mathbf{u_i}}{\|\mathbf{u_i}\|} \in \mathbb{R}^2$ where each $\mathbf{u_i}$ is sampled from a standard Gaussian. The transition from product i to product $j \in \mathcal{N}$ is defined as

$$\rho_{ij} = \frac{e^{\mathbf{v_i}^\top \mathbf{v_j}}}{1 + \sum_{k \in \mathcal{N}} e^{\mathbf{v_i}^\top \mathbf{v_k}}}.$$

We report in Table 2 percentiles of the distribution of profit gap (in percentage) for each policy. As expected, we observe that the gap between LCB-SAA and SAA is tighter for this family of instance. However, we can still observe that LCB-SAA improves by approximately 1% upon SAA. This gap tends to vanish as T grows which is a natural consequence of the asymptotic optimality of SAA when fixing N and L.

Products	ALG	$T = 100$			$T = 200$			$T = 500$		
		25th	50th	75th	25th	50th	75th	25th	50th	75th
N = 10	LCB-SAA	85.4	87.1	89.7	91.1	92.0	93.5	95.0	95.6	96.8
	SAA	83.4	85.3	88.8	89.9	90.8	92.8	94.3	95.0	96.8
N = 20	LCB-SAA	82.9	84.3	85.8	89.9	90.7	91.6	94.7	95.2	95.8
	SAA	80.8	82.4	84.1	88.3	89.4	90.6	94.0	94.4	95.3

Table 2: Percentiles of the profit gap distribution for LCB-SAA and SAA for synthetic instances

6.3 Model misspecification

We compare both algorithms on a dataset generated from our query-to-purchase choice process. In this sequence of experiments, the underlying model is not a Markov Chain decision model therefore both algorithms suffer from model *misspesification* when deciding inventory levels. In order to compare both approaches, we first construct an approximate Markov choice model by computing estimators of the following quantities

$$\lambda_i = \pi(i, \mathcal{N}), \qquad \forall i \in \mathcal{N}$$
$$\rho_{ij} = \frac{\pi(j, \mathcal{N} \setminus \{i\}) - \pi(j, \mathcal{N})}{\pi(i, \mathcal{N})}, \qquad \forall i, j \in \mathcal{N}.$$

This is achieved by simulating 5000 customer arrivals with assortment being $\mathcal{N}$, and repeating this process for each product $i \in \mathcal{N}$, with assortment $\mathcal{N} \setminus \{i\}$.

Similarly to previous experiments we define high and low price ranges which are now defined on different supports. The low price range is supported on $[20, 50]$ with a mode equal to 35 whereas the high price range is supported on $[50, 300]$ with a mode equal to 75. We obtain costs by sampling margins for each products. Margins are uniformly sampled from the interval $[0.4, 0.5]$ which corresponds to a markup ranging from 66% to 100%. These values corresponds to typical margins for fashion retail as reported by (Ghemawat et al. (2003); Şen (2008)).

We present in Figure 4 results obtained by our algorithm for 50 products on the query-to-purchase choice model with a time horizon of 500 customers. Both algorithms decide inventory on a single instance by using $20,000$ scenarios. Their performance is then evaluated on 400 selling seasons. We can see from this figure that LCB-SAA is stochastically dominating SAA. It suggests that the inventory decision of LCB-SAA is able to capture selling seasons with high demand while still being conservative enough to avoid losses in selling seasons with less demand.

7 Conclusion

Inventory planning under heavy-tail demand is a challenging practical problem faced by online retailers. The challenges arise from many considerations including *i*) a complex choice process

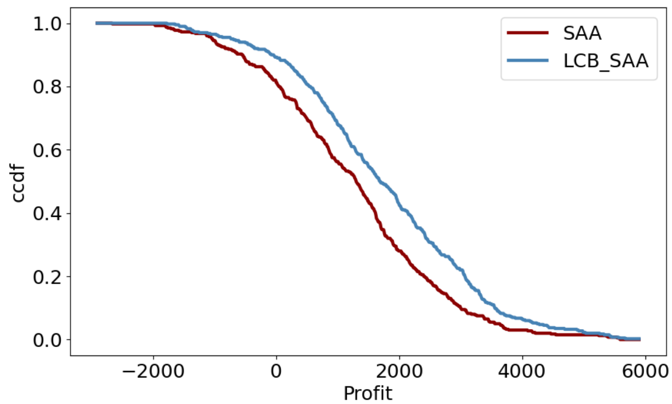

Figure 4: Empirical complementary cumulative distribution (ccdf) for SAA and LCB-SAA.

of customers with the online retailer, and *ii*) the computational difficulty in modeling stock-out based substitutions. The main contributions in this paper are two-fold. First, we approximate the complex interaction of the customers with the online seller and the resulting choice process using a Markov Chain choice model and formulate the inventory optimization problem under a Markov Chain choice model. We present a near-optimal algorithm that achieves a regret of $\tilde{\mathcal{O}}(\sqrt{NT})$ with respect to an LP upper bound. This is near-optimal and improves significantly over prior approaches where we show that the regret can be linear in T.

Furthermore, we also present a model for the complex interactions of customers with the online seller that captures the customer journey from query-to-purchase. We show that parameters of this model can be learned efficiently from data observed by the online retailers such as clicks and purchase decisions. We also study the empirical performance of our algorithm as compared to the LP upper bound as well as prior approaches. We observe that our algorithm performs significantly better than prior approaches both on synthetic instances as well as instances that capture realistic settings.

References

Agarwal, A., Hosanagar, K., and Smith, M. D. (2011). Location, location, location: An analysis of profitability of position in online advertising markets. *Journal of Marketing Research*, 48(6):1057–1073.

Aouad, A., Farias, V., and Levi, R. (2020). Assortment optimization under consider-then-choose choice models. *Management Science*.

Aouad, A., Feldman, J., Segev, D., and Zhang, D. (2019). Click-based mnl: Algorithmic frameworks for modeling click data in assortment optimization. *Available at SSRN 3340620*.

Aouad, A. and Segev, D. (2019). The stability of mnl-based demand under dynamic customer substitution and its algorithmic implications. *Available at SSRN 3498325*.

Blanchet, J., Gallego, G., and Goyal, V. (2016). A markov chain approximation to choice modeling. *Operations Research*, 64(4):886–905.

Brynjolfsson, E., Hu, Y. J., and Smith, M. D. (Summer 2006). From niches to riches: Anatomy of the long tail. *Sloan Management Review*, 47(4):67–71.

Cachon, G. P., Terwiesch, C., and Xu, Y. (2005). Retail assortment planning in the presence of consumer search. *Manufacturing & Service Operations Management*, 7(4):330–346.

Crompton, J. L. and Ankomah, P. K. (1993). Choice set propositions in destination decisions. *Annals of Tourism Research*, 20(3):461–476.

Désir, A., Goyal, V., Segev, D., and Ye, C. (2020). Constrained assortment optimization under the markov chain–based choice model. *Management Science*, 66(2):698–721.

Doerr, B. (2020). *Probabilistic Tools for the Analysis of Randomized Optimization Heuristics*, pages 1–87. Springer International Publishing, Cham.

Farahat, A. and Lee, J. (2018). The multiproduct newsvendor problem with customer choice. *Operations research*, 66(1):123–136.

Feldman, J. B. and Topaloglu, H. (2017). Revenue management under the markov chain choice model. *Operations Research*, 65(5):1322–1342.

Gallego, G. and Kim, S. (2020). Joint pricing and inventory decisions for substitutable products. *Available at SSRN 3513274*.

Gaur, V. and Honhon, D. (2006). Assortment planning and inventory decisions under a locational choice model. *Management Science*, 52(10):1528–1543.

Ghemawat, P., Nueno, J. L., and Dailey, M. (2003). *ZARA: Fast fashion*, volume 1. Harvard Business School Boston, MA.

Goel, S., Broder, A., Gabrilovich, E., and Pang, B. (2010). Anatomy of the long tail: Ordinary people with extraordinary tastes. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, WSDM '10, page 201–210, New York, NY, USA. Association for Computing Machinery.

Goyal, V., Levi, R., and Segev, D. (2016). Near-optimal algorithms for the assortment planning problem under dynamic substitution and stochastic demand. *Operations Research*, 64(1):219–235.

Hoeffding, W. (1994). Probability inequalities for sums of bounded random variables. In *The Collected Works of Wassily Hoeffding*, pages 409–426. Springer.

Honhon, D., Gaur, V., and Seshadri, S. (2010). Assortment planning and inventory decisions under stockout-based substitution. *Operations research*, 58(5):1364–1379.

Honhon, D. and Seshadri, S. (2013). Fixed vs. random proportions demand models for the assortment planning problem under stockout-based substitution. *Manufacturing & Service Operations Management*, 15(3):378–386.

Honka, E., Hortaçsu, A., and Wildenbeest, M. (2019). Chapter 4 - empirical search and consideration sets. In Dubé, J.-P. and Rossi, P. E., editors, *Handbook of the Economics of Marketing, Volume 1*, volume 1 of *Handbook of the Economics of Marketing*, pages 193 – 257. North-Holland.

Howard, J. Y. S. and Sheth, J. (1969). Jn (1969)" the theory of buyer behavior. *New York*.

Jansen, B. J. and Spink, A. (2006). How are we searching the world wide web? a comparison of nine search engine transaction logs. *Information Processing & Management*, 42(1):248 – 263. Formal Methods for Information Retrieval.

Jeuland, A. P. (1979). Brand choice inertia as one aspect of the notion of brand loyalty. *Management Science*, 25(7):671–682.

Kleywegt, A. J., Shapiro, A., and Homem-de Mello, T. (2002). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502.

Kök, A. G. and Fisher, M. L. (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research*, 55(6):1001–1021.

Li, M. M., Liu, X., Huang, Y., and Shi, C. (2018). Integrating empirical estimation and assortment personalization for e-commerce: A consider-then-choose model. *Xiang and Huang, Yan and Shi, Cong, Integrating Empirical Estimation and Assortment Personalization for E-Commerce: A Consider-Then-Choose Model (September 10, 2018)*.

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

Mousavi, N. (2010). How tight is chernoff bound.

Müller, S., Huber, J., Fleischmann, M., and Stuckenschmidt, H. (2020). Data-driven inventory management under customer substitution. *Available at SSRN 3624026*.

Nagarajan, M. and Rajagopalan, S. (2008). Inventory models for substitutable products: optimal policies and heuristics. *Management Science*, 54(8):1453–1466.

Netessine, S. and Rudi, N. (2003). Centralized and competitive inventory models with demand substitution. *Operations research*, 51(2):329–335.

Parlar, M. and Goyal, S. (1984). Optimal ordering decisions for two substitutable products with stochastic demands. *Opsearch*, 21(1):1–15.

Şen, A. (2008). The us fashion industry: A supply chain review. *International Journal of Production Economics*, 114(2):571–593.

Silverstein, C., Marais, H., Henzinger, M., and Moricz, M. (1999). Analysis of a very large web search engine query log. *SIGIR Forum*, 33(1):6–12.

Smith, S. A. and Agrawal, N. (2000). Management of multi-item retail inventory systems with demand substitution. *Operations Research*, 48(1):50–64.

Szpektor, I., Gionis, A., and Maarek, Y. (2011). Improving recommendation for long-tail queries via templates. In *Proceedings of the 20th International Conference on World Wide Web*, WWW '11, page 47–56, New York, NY, USA. Association for Computing Machinery.

Topaloglu, H. (2013). Joint stocking and product offer decisions under the multinomial logit model. *Production and Operations Management*, 22(5):1182–1199.

Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science*, 37(4):530–552.

van Ryzin, G. and Mahajan, S. (1999). On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*, 45(11):1496–1509.

Wang, R. and Sahin, O. (2018). The impact of consumer search cost on assortment planning and pricing. *Management Science*, 64(8):3649–3666.

Weissman, T., Ordentlich, E., Seroussi, G., Verdu, S., and Weinberger, M. J. (2003). Inequalities for the l1 deviation of the empirical distribution. *Hewlett-Packard Labs, Tech. Rep.*

A Proof of Proposition 2.1

Let $\mathbf{q}$ be an inventory decision and denote by $(X_i^t)_{i\in\mathcal{N},t\in\{1,\ldots,T\}}$ the sequence of random variables such that for every $t \in \{1,\ldots,T\}$ and $i \in \mathcal{N}$ the variable $X_i^t = 1$ if the customer t purchases product i, conditionally on $\mathbf{q}$ being the initial inventory level, and $X_i^t = 0$ otherwise. The profit $\tilde{\mathcal{P}}_T(\mathbf{q})$ generated on a selling season after prescribing inventory level $\mathbf{q}$ is given by,

$$\begin{aligned}\tilde{\mathcal{P}}_T(\mathbf{q}) &= \sum_{i\in\mathcal{N}}\sum_{t=1}^{T} p_i X_i^t - \sum_{i\in\mathcal{N}} c_i q_i \\ &\le \sum_{i\in\mathcal{N}}\sum_{t=1}^{T} (p_i - c_i)\, X_i^t,\end{aligned}$$

where the inequality holds because $\sum_{t=1}^{T} X_i^t \le q_i$ for all $i \in \mathcal{N}$. By taking expectation, we get

$$\begin{aligned}\mathcal{P}_T(\mathbf{q}) &= \mathbb{E}\left[\tilde{\mathcal{P}}_T(\mathbf{q})\right] \\ &\le \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i\in\mathcal{N}} (p_i - c_i)\, X_i^t\right] \\ &\overset{(a)}{=} \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i\in\mathcal{N}} (p_i - c_i)\, \mathbb{E}\left[X_i^t \mid S_1,\ldots,S_T\right]\right] \\ &= \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i\in\mathcal{N}} (p_i - c_i)\, \pi(i,S_t)\right] \\ &\le \mathbb{E}\left[\sum_{t=1}^{T}\max_{S\subset\mathcal{N}}\sum_{i\in\mathcal{N}} (p_i - c_i)\, \pi(i,S)\right] \\ &= T\cdot\max_{S\subset\mathcal{N}}\sum_{i\in\mathcal{N}} (p_i - c_i)\, \pi(i,S) = z_{\mathsf{fluid}},\end{aligned}$$

where the random variables $(S_t)_{t\in\{1,\ldots,T\}}$ in (a) are defined such that for every $t \in \{1,\ldots,T\}$, S_t is the set of products offered to customer t, i.e., the set of products for which there is some remaining inventory at time t. Note that the equality in (a) holds by the law of total expectation.

B Proof of Lemma 3.2.

In this section, we present the proof of Lemma 3.2. Our proof relies on two key lemmas. We first present the structure of our proof and the statements of the two key lemmas and then give their proofs. Let $(\mathbf{x}^{\mathrm{SAA},\ell})_{\ell\in[L]}$,$(\mathbf{y}^{\mathrm{SAA},\ell})_{\ell\in[L]}$ and $\mathbf{q}^{\mathrm{SAA}}$ denote the solution of the SAA algorithm. For any

$T \geq 1$, define the following events,

$$\begin{aligned}
\mathcal{E}_1' &:= \left\{ \exists \ell \in [L] \text{ s.t. } D_1^\ell > \frac{3T}{4} \right\} \\
\mathcal{E}_2' &:= \left\{ \sum_{i=2}^{N+1} q_i^{\text{SAA}} \leq \frac{T}{2} + 2\sqrt{\frac{T^2 \log(2T)}{L}} \right\} \\
\mathcal{E}' &:= \mathcal{E}_1' \cap \mathcal{E}_2'.
\end{aligned}$$

First, we show that the events $\mathcal{E}_1'$ and $\mathcal{E}_2'$ happen conjointly with high probability. Formally, we have the following lemma.

Lemma B.1. *For any $L \geq 4e^{2T} \log(2T)$, we have $\mathbb{P}(\mathcal{E}') \geq 1 - \frac{1}{T}$.*

Second, we show that conditionally on the event $\mathcal{E}'$, the solution of problem (3.3) carries too many units of product 1. In fact, in order to handle the scenario $\ell \in [L]$ satisfying $D_1^\ell > \frac{3T}{4}$, the SAA algorithm can either increase the inventory for products $\{2, \ldots, N+1\}$, or carry some inventory in product 1. We show that it necessarily chooses the second option and carries a constant fraction of T inventory from the suboptimal product 1. Formally, we show that the event $\mathcal{E}'$ implies the event $\mathcal{E}$ defined in (3.4).

Lemma B.2. *For any $T \geq 2$ and $L \geq 4e^{2T} \log(2T)$, we have $\mathcal{E}' \subset \mathcal{E}$.*

From Lemma B.1 and Lemma B.2, we get $\mathbb{P}(\mathcal{E}) \geq \mathbb{P}(\mathcal{E}') \geq 1 - \frac{1}{T}$, which is the desired result for Lemma 3.2. To complete our proof, let us give the proofs of Lemma B.1 and Lemma B.2.

Proof of Lemma B.1.
Step 1: We first bound the probability of event $\mathcal{E}_1'$ by using the following probabilistic bound.

Proposition B.3 (Theorem 2 (Mousavi (2010))). *Let $X_1, \ldots, X_m$ be i.i.d. Bernoulli random variables such that $\mathbb{P}(X_1 = 1) = p$. If $p \leq \frac{1}{4}$ then for any $t \geq 0$,*

$$\mathbb{P}\left(\sum_{i=1}^m X_i - mp > t \right) \geq \frac{1}{4} \exp\left(-\frac{2t^2}{mp} \right).$$

Recall that for every $\ell \in [L]$, D_1^ℓ is the sum of T i.i.d. Bernoulli random variables with mean $\frac{1}{4}$. By applying Proposition B.3 we obtain that,

$$\mathbb{P}\left(D_1^\ell > \frac{3T}{4} \right) = \mathbb{P}\left(D_1^\ell > \mathbb{E}\left[D_1^\ell \right] + \frac{T}{2} \right) \geq \frac{1}{4} e^{-2T}. \tag{B.1}$$

Let $c_T := \frac{1}{4} e^{-2T}$. One has that,

$$\mathbb{P}\left(\forall \ell \leq L,\, D_1^\ell \leq \frac{3T}{4} \right) \overset{(a)}{=} \left(1 - \mathbb{P}\left(D_1^1 > \frac{3T}{4} \right) \right)^L \overset{(b)}{\leq} (1 - c_T)^L \overset{(c)}{\leq} (1 - c_T)^{\frac{\log(2T)}{c_T}} \overset{(d)}{\leq} \frac{1}{2T},$$

where (a) holds because $(D_1^\ell)_{\ell\in[L]}$ are i.i.d., (b) is a consequence of (B.1), (c) holds because $L \geq \frac{\log(2T)}{c_T}$ and (d) comes from the inequality $(1-\frac{1}{t})^t \leq e^{-1}$, $\forall t > 0$ applied to $t = \frac{1}{c_T}$. Therefore, $\mathbb{P}(\mathcal{E}_1') \geq 1 - \frac{1}{2T}$.

Step 2: We derive a bound on the probability of the event $\mathcal{E}_2'$. First, we state in (B.2) the LP solved by SAA for the problem instance defined in Figure 1.

$$\max_{\mathbf{x}^\ell,\mathbf{y}^\ell,\mathbf{q}} \quad \frac{1}{L}\sum_{\ell=1}^{L}\mathbf{p}^\top\mathbf{x}^\ell - \mathbf{c}^\top\mathbf{q} \tag{B.2a}$$

$$\text{subject to} \quad x_i^\ell \leq q_i, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L], \tag{B.2b}$$

$$x_i^\ell + y_i^\ell = \frac{1}{N}y_1^\ell, \qquad \forall i \in \{2,\dots,N+1\}, \forall \ell \in [L], \tag{B.2c}$$

$$x_1^\ell + y_1^\ell = D_1^\ell, \qquad \forall \ell \in [L], \tag{B.2d}$$

$$x_i^\ell \geq 0, y_i^\ell \geq 0, q_i \geq 0, \qquad \forall i \in \mathcal{N}, \forall \ell \in [L]. \tag{B.2e}$$

Define the random events,

$$\mathcal{A}_1' = \left\{ \frac{1}{L}\sum_{\ell=1}^{L} D_1^\ell \leq \frac{T}{4} + \sqrt{\frac{T^2\log(2T)}{L}} \right\}$$

$$\mathcal{A}_2' = \left\{ \forall i \in \{2,\dots,N+1\},\ \frac{1}{L}\sum_{\ell=1}^{L} x_i^{\mathrm{SAA},\ell} \leq \frac{T}{4N} + \sqrt{\frac{T^2\log(2T)}{N^2L}} \right\}.$$

We show that the following sequence of inclusions holds,

$$\mathcal{A}_1' \subset \mathcal{A}_2' \subset \mathcal{E}_2'. \tag{B.3}$$

From (B.2d), we remark that for each $\ell \in [L]$, $y_1^{\mathrm{SAA},\ell} \leq D_1^\ell$ and from (B.2c) we obtain that for each $i \in \{2,\dots,N+1\}$ and $\ell \in [L]$, $x_i^{\mathrm{SAA},\ell} \leq \frac{y_1^{\mathrm{SAA},\ell}}{N} \leq \frac{D_1^\ell}{N}$. Hence we conclude that, $\mathcal{A}_1' \subset \mathcal{A}_2'$.

Next, we use the following notation. We denote by $Obj\left((\mathbf{x}^\ell)_{\ell\in[L]},\mathbf{q}\right)$ the objective value of a feasible solution $(\mathbf{x}^\ell)_{\ell\in[L]},\mathbf{q}$ for problem (B.2), i.e.,

$$Obj\left((\mathbf{x}^\ell)_{\ell\in[L]},\mathbf{q}\right) := \frac{1}{L}\sum_{\ell=1}^{L}\mathbf{p}^\top\mathbf{x}^\ell - \mathbf{c}^\top\mathbf{q}.$$

We will show that if $\mathcal{A}_2'$ holds, then $\mathcal{E}_2'$ holds. Suppose for the sake of contradiction that $\mathcal{A}_2'$ holds and $\sum_{i=2}^{N+1} q_i^{\mathrm{SAA}} > \frac{T}{2} + 2\sqrt{\frac{T^2\log(2T)}{L}}$. This implies that

$$\sum_{i=2}^{N+1} q_i^{\mathrm{SAA}} > 2\sum_{i=2}^{N+1}\frac{1}{L}\sum_{\ell=1}^{L} x_i^{\mathrm{SAA},\ell}. \tag{B.4}$$

We construct the following feasible solution where for all $\ell \in [L]$,

$$\tilde{x}_i^\ell = \begin{cases} x_1^{\mathrm{SAA},\ell} & \text{if } i = 1 \\ 0 & \text{if } i \in \{2, \dots, N+1\}. \end{cases}$$

$$\tilde{y}_i^\ell = \begin{cases} D_1^\ell - x_1^{\mathrm{SAA},\ell} & \text{if } i = 1 \\ \frac{1}{N}\left(D_1^\ell - x_1^{\mathrm{SAA},\ell}\right) & \text{if } i \in \{2, \dots, N+1\}. \end{cases}$$

$$\tilde{q}_i = \begin{cases} q_1^{\mathrm{SAA}} & \text{if } i = 1 \\ 0 & \text{if } i \in \{2, \dots, N+1\}. \end{cases}$$

For each $\ell \in [L]$, $\tilde{\mathbf{x}}^\ell$, $\tilde{\mathbf{y}}^\ell$ and $\tilde{\mathbf{q}}$ clearly satisfy the constraints of problem (B.2). Moreover,

$$\begin{aligned} Obj\left((\mathbf{x}^{\mathrm{SAA},\ell})_{\ell\in[L]}, \mathbf{q}^{\mathrm{SAA}}\right) - Obj\left((\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}, \tilde{\mathbf{q}}\right) &= \frac{1}{4}\frac{1}{L}\sum_{\ell=1}^{L} x_1^{\mathrm{SAA},\ell} + \sum_{i=2}^{N+1}\left(\frac{1}{L}\sum_{\ell=1}^{L} x_i^{\mathrm{SAA},\ell} - \frac{q_i^{\mathrm{SAA},\ell}}{2}\right) \\ &\quad - \left(\frac{1}{4}\frac{1}{L}\sum_{\ell=1}^{L}\tilde{x}_1^\ell + \sum_{i=2}^{N+1}\left(\frac{1}{L}\sum_{\ell=1}^{L}\tilde{x}_i^\ell - \frac{\tilde{q}_i}{2}\right)\right) \\ &= \sum_{i=2}^{N+1}\left(\frac{1}{L}\sum_{\ell=1}^{L} x_i^{\mathrm{SAA},\ell} - \frac{1}{2}q_i^{\mathrm{SAA}}\right) < 0, \end{aligned}$$

where the last inequality is implied by (B.4). Hence, we have a feasible solution that has a strictly better objective value than the optimal SAA solution, which is a contradiction. Therefore, $\mathcal{A}_2' \subset \mathcal{E}_2'$. Thus, we conclude that (B.3) holds. Hence,

$$\begin{aligned} \mathbb{P}\left(\mathcal{E}_2'\right) \geq \mathbb{P}\left(\mathcal{A}_1'\right) &= 1 - \mathbb{P}\left(\frac{1}{L}\sum_{\ell=1}^{L} D_1^\ell \geq \frac{T}{4} + \sqrt{\frac{T^2\log(2T)}{L}}\right) \\ &\overset{(a)}{\geq} 1 - \exp\left(-2\frac{L^2\frac{T^2}{L}\log(2T)}{LT^2}\right) = 1 - \frac{1}{4T^2} \geq 1 - \frac{1}{2T}, \end{aligned}$$

where (a) follows from Hoeffding inequality for bounded random variables (Theorem 2 in Hoeffding (1994)) applied to $(D_1^\ell)_{\ell\in[L]}$ which are bounded random variable in $[0, T]$ with mean $\frac{T}{4}$. We finally conclude by a union bound that,

$$\mathbb{P}(\mathcal{E}) = \mathbb{P}(\mathcal{E}_1' \cap \mathcal{E}_2') \geq 1 - \frac{1}{2T} - \frac{1}{2T} = 1 - \frac{1}{T}.$$

□

Proof of Lemma B.2. We will show that if $\mathcal{E}'$ holds, then $\mathcal{E}$ holds. Assume for the sake of contradiction that $\mathcal{E}'$ holds and $q_1^{\mathrm{SAA}} < \frac{T}{4} - 2\sqrt{\frac{T^2\log(2T)}{L}}$. Since $\mathcal{E}_1 = \mathcal{E}_1' \cap \mathcal{E}_2'$, then $\mathcal{E}_1'$ holds and therefore there exists $\ell \in [L]$ such that, $D_1^\ell > \frac{3T}{4}$. Without loss of generality assume that $D_1^1 > \frac{3T}{4}$.

We will construct a feasible solution $(\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}, (\tilde{\mathbf{y}}^\ell)_{\ell\in[L]}, \tilde{\mathbf{q}}$ for problem (B.2) that contradicts the optimality of $(\mathbf{x}^{\mathrm{SAA},\ell})_{\ell\in[L]}, (\mathbf{y}^{\mathrm{SAA},\ell})_{\ell\in[L]}, \mathbf{q}^{\mathrm{SAA}}$. Define $\tilde{\mathbf{q}}$ such that,

$$\tilde{q}_i = \begin{cases} \frac{T}{4} - 2\sqrt{\frac{T^2\log(2T)}{L}} & \text{if } i = 1 \\ \frac{1}{N}\sum_{j=2}^{N+1} q_j^{\mathrm{SAA}} & \text{if } i \in \{2,\dots,N+1\}. \end{cases}$$

For all $\ell \in \{2,\dots,L\}$, define $\tilde{\mathbf{x}}^\ell$ and $\tilde{\mathbf{y}}^\ell$ as,

$$\tilde{x}_i^\ell = \begin{cases} x_1^{\mathrm{SAA},\ell} & \text{if } i = 1 \\ \frac{1}{N}\sum_{j=2}^{N+1} x_j^{\mathrm{SAA},\ell} & \text{if } i \in \{2,\dots,N+1\}. \end{cases}$$

$$\tilde{y}_i^\ell = \begin{cases} y_1^{\mathrm{SAA},\ell} & \text{if } i = 1 \\ \frac{1}{N}\sum_{j=2}^{N+1} y_j^{\mathrm{SAA},\ell} & \text{if } i \in \{2,\dots,N+1\}. \end{cases}$$

For $\ell = 1$, define

$$\begin{aligned} \tilde{x}_i^1 &= \begin{cases} \frac{T}{4} - 2\sqrt{\frac{T^2\log(2T)}{L}} & \text{if } i = 1 \\ \frac{1}{N}\sum_{j=2}^{N+1} x_j^{\mathrm{SAA},1} & \text{if } i \in \{2,\dots,N+1\}. \end{cases} \\ \tilde{y}_i^1 &= \begin{cases} D_1^1 - \frac{T}{4} + 2\sqrt{\frac{T^2\log(2T)}{L}} & \text{if } i = 1 \\ \frac{\tilde{y}_1^1}{N} - \tilde{x}_i^1 & \text{if } i \in \{2,\dots,N+1\}. \end{cases} \end{aligned}$$

Feasibility. Remark that, for every $\ell \in [L]$ and $i \in \{2,\dots,N+1\}$,

$$\tilde{x}_i^\ell = \frac{1}{N}\sum_{j=2}^{N+1} x_j^{\mathrm{SAA},\ell} \le \frac{1}{N}\sum_{j=2}^{N+1} q_j^{\mathrm{SAA}} = \tilde{q}_i.$$

For $i = 1$, it is clear that $\tilde{x}_1^1 \le \tilde{q}_1$. For any $\ell \in \{2,\dots,L\}$,

$$\tilde{x}_1^\ell = x_1^{\mathrm{SAA},\ell} \le q_1^{\mathrm{SAA}} \overset{(a)}{\le} \frac{T}{4} - 2\sqrt{\frac{T^2\log(2T)}{L}} = \tilde{q}_1,$$

where (a) follows from the assumption on q_1^{SAA}. Therefore, the constraint (B.2b) is satisfied. Furthermore, for $\ell \in \{2,\dots,L\}$ and $i \in \{2,\dots,N+1\}$,

$$\tilde{x}_i^\ell + \tilde{y}_i^\ell = \frac{1}{N}\sum_{j=2}^{N+1} x_j^{\mathrm{SAA},\ell} + y_j^{\mathrm{SAA},\ell} \overset{(a)}{=} \frac{1}{N}\sum_{j=2}^{N+1} \frac{y_1^\ell}{N} = \frac{y_1^\ell}{N},$$

where (a) holds because the SAA solution satisfies (B.2c). When $\ell = 1$, (B.2c) holds trivially. Note that $(\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}$ and $(\tilde{\mathbf{y}}^\ell)_{\ell\in[L]}$ clearly satisfy (B.2d). Lastly, let us check the non-negativity of our variables. Remark that $\tilde{q}_1 \geq 0$ for $T \geq 2$. In fact, $\tilde{q}_1 \geq 0$ is equivalent to $\frac{\log(2T)}{L} \leq \frac{1}{64}$ which is true since from $L \geq 4e^{2T}\log(2T)$, we obtain for $T \geq 2$, $\frac{\log(2T)}{L} \leq \frac{1}{4}\exp(-2T) < \frac{1}{64}$. This also shows that $\tilde{x}_1^1 \geq 0$. The only non-trivial remaining case is to show that $\tilde{\mathbf{y}}^1 \geq 0$. Since $D_1^1 > \frac{3}{4}T$, we get

$$\tilde{y}_1^1 = D_1^1 - \frac{T}{4} + 2\sqrt{\frac{T^2\log(2T)}{L}} \geq \frac{T}{2} + 2\sqrt{\frac{T^2\log(2T)}{L}} \geq 0.$$

This also implies that for $i \in \{2, \ldots, N+1\}$,

$$\begin{aligned}
\frac{\tilde{y}_1^1}{N} - \tilde{x}_i^1 &\geq \frac{T}{2N} + 2\sqrt{\frac{T^2\log(2T)}{N^2L}} - \tilde{x}_i^1 \\
&= \frac{T}{2N} + 2\sqrt{\frac{T^2\log(2T)}{N^2L}} - \frac{1}{N}\sum_{j=2}^{N+1} x_j^{\mathrm{SAA},\ell} \\
&\geq \frac{T}{2N} + 2\sqrt{\frac{T^2\log(2T)}{N^2L}} - \frac{1}{N}\sum_{j=2}^{N+1} q_j^{\mathrm{SAA}} \geq 0,
\end{aligned}$$

where the last inequality comes from the fact that $\mathcal{E}_2'$ holds. Therefore $\tilde{\mathbf{y}}^1 \geq 0$.

Objective value. We have

$$\begin{aligned}
Obj\left((\mathbf{x}^{\mathrm{SAA},\ell})_{\ell\in[L]}, \mathbf{q}^{\mathrm{SAA}}\right) - Obj\left((\tilde{\mathbf{x}}^{\ell})_{\ell\in[L]}, \tilde{\mathbf{q}}\right) &= \frac{1}{L}\sum_{\ell=1}^{L}\sum_{i=1}^{N+1} p_i\left(x_i^{\mathrm{SAA},\ell} - \tilde{x}_i^{\ell}\right) + \sum_{i=1}^{N+1} c_i\left(\tilde{q}_i - q_i^{\mathrm{SAA}}\right) \\
&\overset{(a)}{=} \frac{1}{L}p_1(x_1^{\mathrm{SAA},1} - \tilde{x}_1^1) + c_1(\tilde{q}_1 - q_1^{\mathrm{SAA}}) \\
&\overset{(b)}{=} \frac{1}{4L}\left(x_1^{\mathrm{SAA},1} - \frac{T}{4} + 2\sqrt{\frac{T^2\log(2T)}{L}}\right) \\
&\leq \frac{1}{4L}\left(q_1^{\mathrm{SAA},1} - \frac{T}{4} + 2\sqrt{\frac{T^2\log(2T)}{L}}\right) \overset{(c)}{<} 0,
\end{aligned}$$

where (a) holds because for each $\ell \in [L]$,

$$\sum_{i=2}^{N+1} p_i\left(x_i^{\mathrm{SAA},\ell} - \tilde{x}_i^{\ell}\right) = \sum_{i=2}^{N+1}\left(x_i^{\mathrm{SAA},\ell} - \frac{1}{N}\sum_{j=2}^{N+1} x_j^{\mathrm{SAA},\ell}\right) = \sum_{i=2}^{N+1} x_i^{\mathrm{SAA},\ell} - \frac{1}{N}\sum_{j=2}^{N+1}\sum_{i=2}^{N+1} x_j^{\mathrm{SAA},\ell} = 0,$$

and a similar argument holds to show that $\sum_{i=2}^{N+1} c_i\left(\tilde{q}_i - q_i^{\mathrm{SAA}}\right) = 0$. Moreover, (b) holds simply because $p_1 = 1/4$ and $c_1 = 0$. Finally, (c) is a consequence of the assumption over q_1^{SAA}. Hence, we contradict the optimality of the SAA solution. Therefore, we conclude $\mathcal{E}' \subset \mathcal{E}$.

□

C Proof of Lemma 3.3

Proof. We define $\mathbf{q}^{\text{fluid}}$ as,

$$q_i^{\text{fluid}} = \begin{cases} 0 & \text{if } i = 1 \\ \frac{T}{4N} & \text{if } i \in \{2, \dots, N+1\}. \end{cases}$$

Given a Markov Chain choice model, recall that we say that a customer visits product $i \in \mathcal{N}$ before product $j \in \mathcal{N}$ if the random walk associated to this customer hits product i before hitting product j. For a customer $t \in \{1, \dots, T\}$, the relation $\prec_t$ is defined as $U \prec_t V$ if and only if there exists a product $i \in U$ such that customer t visits product i before any product in V. For $i \in \{2, \dots, N+1\}$ define $\tilde{D}_i$ as follows:

$$\tilde{D}_i = \sum_{t=1}^{T} \mathbb{1}\left\{\{1\} \prec_t \{i\} \prec_t \{0\}\right\}$$

$$\tilde{D}_1 = \sum_{t=1}^{T} \mathbb{1}\left\{\{0\} \prec_t \{1\}\right\}.$$

Note that, $\tilde{\mathbf{D}} := (\tilde{D}_i)_{i \in \{1,\dots,N+1\}}$ follows a multinomial distribution with T trials and a vector of probability $\mathbf{d}$ such that, $d_1 = \frac{3}{4}$ and for $i \in \{2, \dots, N+1\}$, $d_i = \frac{1}{4} \times \frac{1}{N} = \frac{1}{4N}$. Let I be the interval,

$$I = \left[\frac{T}{4N} - 2\sqrt{NT\log(2T)}, \frac{T}{4N} + 2\sqrt{NT\log(2T)}\right],$$

and define the event,

$$\mathcal{G} = \left\{\tilde{D}_i \in I, \text{ for all } i \in \{2, \dots, N+1\}\right\}.$$

The profit $\tilde{\mathcal{P}}_T(\mathbf{q}^{\text{fluid}})$ over a selling season is expressed as a function of $\tilde{\mathbf{D}}$ as follows:

$$\tilde{\mathcal{P}}_T(\mathbf{q}^{\text{fluid}}) = \sum_{i=2}^{N+1} \min\left(q_i^{\text{fluid}}, \tilde{D}_i\right) - \frac{1}{2}\sum_{i=2}^{N+1} q_i^{\text{fluid}} = \sum_{i=2}^{N+1} \min\left(\frac{T}{4N}, \tilde{D}_i\right) - \frac{T}{8}.$$

Hence,

$$\mathbb{E}\left[\tilde{\mathcal{P}}_T(\mathbf{q}^{\text{fluid}}) \;\middle|\; \mathcal{G}\right] = \mathbb{E}\left[\sum_{i=2}^{N+1} \min\left(\frac{T}{4N}, \tilde{D}_i\right) \;\middle|\; \mathcal{G}\right] - \frac{T}{8} \overset{(a)}{\geq} \frac{T}{8} - 2\sqrt{N^3 T \log(2T)}, \tag{C.1}$$

where (a) holds because, conditionally on $\mathcal{G}$, $\tilde{D}_i \geq \frac{T}{4N} - 2\sqrt{NT\log(2T)}$ for all $i \in \{2, \dots, N+1\}$.

We now bound the profit generated by the SAA algorithm. We define the sequence of random variables $(\tilde{Q}_\tau)_{\tau \in \{1,\dots,T\}}$ such that for every $\tau \in \{1,\dots,T\}$,

$$\tilde{Q}_\tau := \sum_{t=1}^{\tau} \mathbb{1}\left\{\{1\} \prec_t \{0\}\right\}.$$

We define $\tilde{T}_1$ as the following stopping time

$$\tilde{T}_1 = \min\{\tau \in \{1,\dots,T\} \mid \tilde{Q}_\tau \geq q_1^{\mathrm{SAA}}\},$$

with the convention that the minimum of the empty set is ∞. Note that, $\tilde{T}_1$ corresponds to the time at which product 1 runs out of stock. For $i \in \{2,\dots,N+1\}$, define

$$\tilde{R}_i := \sum_{t=\tilde{T}_1+1}^{T} \mathbb{1}\left\{\{1\} \prec_t \{i\} \prec_t \{0\}\right\}.$$

Conditionally on $\{\tilde{T}_1 < \infty\}$, we have, almost surely,

$$\begin{aligned}\sum_{i=2}^{N+1} \tilde{R}_i &= \sum_{i=2}^{N+1} \tilde{D}_i - \sum_{t=1}^{\tilde{T}} \sum_{i=2}^{N+1} \mathbb{1}\left\{\{1\} \prec_t \{i\} \prec_t \{0\}\right\} = \sum_{i=2}^{N+1} \tilde{D}_i - \sum_{t=1}^{\tilde{T}} \mathbb{1}\left\{\{1\} \prec_t \{0\}\right\} \\ &= \sum_{i=2}^{N+1} \tilde{D}_i - \tilde{Q}_{\tilde{T}_1} = \sum_{i=2}^{N+1} \tilde{D}_i - q_1^{\mathrm{SAA}}. \end{aligned} \tag{C.2}$$

Therefore conditionally on $\{\tilde{T}_1 < \infty\}$, the profit of the SAA inventory decision on one selling season satisfies almost surely,

$$\begin{aligned}\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{SAA}}) &= \frac{1}{4} q_1^{\mathrm{SAA}} + \sum_{i=2}^{N+1} \min\left(\tilde{R}_i, q_i^{\mathrm{SAA}}\right) - \frac{1}{2} \sum_{i=2}^{N+1} q_i^{\mathrm{SAA}} \\ &\overset{(a)}{\leq} \frac{1}{4} q_1^{\mathrm{SAA}} + \frac{1}{2} \sum_{i=2}^{N+1} \min\left(\tilde{R}_i, q_i^{\mathrm{SAA}}\right) \\ &\leq \frac{1}{4} q_1^{\mathrm{SAA}} + \frac{1}{2} \sum_{i=2}^{N+1} \tilde{R}_i \\ &\overset{(b)}{=} \frac{1}{2} \sum_{i=2}^{N+1} \tilde{D}_i - \frac{1}{4} q_1^{\mathrm{SAA}}, \end{aligned}$$

where (a) holds because for $i \in \{2,\dots,N+1\}$, $q_i^{\mathrm{SAA}} \geq \min\left(\tilde{R}_i, q_i^{\mathrm{SAA}}\right)$ and (b) follows from (C.2).

Conditionally on $\{\tilde{T}_1 < \infty\}$, product 1 runs out of stock whereas for $\{\tilde{T}_1 = \infty\}$, we only sell a fraction of the inventory of product 1 and we do not sell any other product. Therefore, the profit generated over a selling season under the event $\{\tilde{T}_1 = \infty\}$ is lower than the one generated conditionally on $\{\tilde{T}_1 < \infty\}$. Hence, we always have $\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{SAA}}) \leq \frac{1}{2} \sum_{i=2}^{N+1} \tilde{D}_i - \frac{1}{4} q_1^{\mathrm{SAA}}$.

Recall $\mathcal{E}$ the event defined in (3.4). We have

$$\begin{aligned}
\mathbb{E}\left[\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{SAA}}) \;\middle|\; \mathcal{E},\mathcal{G}\right] &\leq \mathbb{E}\left[\frac{1}{2}\sum_{i=2}^{N+1}\tilde{D}_i - \frac{1}{4}q_1^{\mathrm{SAA}} \middle| \mathcal{E},\mathcal{G}\right] \\
&\overset{(a)}{\leq} \frac{1}{2}\left(\frac{T}{4} + 2\sqrt{N^3T\log(2T)}\right) - \frac{T}{16} + \frac{1}{2}\sqrt{\frac{T^2\log(2T)}{L}} \\
&\overset{(b)}{\leq} \frac{T}{16} + \frac{3}{2}\sqrt{N^3T\log(2T)}, \qquad \text{(C.3)}
\end{aligned}$$

where (a) holds because conditionally on $\mathcal{E}$, the inventory level satisfies $q_1^{\mathrm{SAA}} \geq \frac{T}{4} - 2\sqrt{\frac{T^2\log(2T)}{L}}$ and conditionally on $\mathcal{G}$, $\tilde{D}_i \leq \frac{T}{4N} + 2\sqrt{NT\log(2T)}$, for all $i \in \{2,\ldots,N+1\}$. Lastly, (b) follows from $L \geq T$ and $N \geq 1$. We conclude by decomposing the regret as follows.

$$\begin{aligned}
\mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{\mathrm{SAA}}) \mid \mathcal{E}\right] &\geq \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^{\mathrm{fluid}}) - \mathcal{P}_T(\mathbf{q}^{\mathrm{SAA}}) \mid \mathcal{E}\right] \\
&= \mathbb{E}\left[\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{fluid}}) - \tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{SAA}}) \mid \mathcal{E},\mathcal{G}\right]\mathbb{P}(\mathcal{G}|\mathcal{E}) \\
&\qquad + \left[\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{fluid}}) - \tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{SAA}}) \mid \mathcal{E},\mathcal{G}^c\right]\mathbb{P}(\mathcal{G}^c|\mathcal{E}) \\
&\overset{(a)}{\geq} \left(\mathbb{E}\left[\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{fluid}}) \mid \mathcal{G}\right] - \mathbb{E}\left[\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{SAA}}) \mid \mathcal{E},\mathcal{G}\right]\right)\mathbb{P}(\mathcal{G}) - \frac{5}{8}T\cdot\mathbb{P}(\mathcal{G}^c) \\
&\overset{(b)}{\geq} \left[\frac{T}{16} - \frac{7}{2}\sqrt{N^3T\log(2T)}\right]\mathbb{P}(\mathcal{G}) - T\cdot\mathbb{P}(\mathcal{G}^c),
\end{aligned}$$

where (a) holds because $\mathcal{G}$ and $\mathcal{E}$ are independent, and it uses $\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{fluid}}) \geq -\sum_{i\in\mathcal{N}} c_i q_i^{\mathrm{fluid}} = -\frac{T}{8}$ and $\tilde{\mathcal{P}}_T(\mathbf{q}^{\mathrm{SAA}}) \leq \frac{T}{2}$ since the profit generated by the most profitable product is $\frac{1}{2}$. Note that (b) is a consequence of (C.1) and (C.3). To conclude the proof, we bound the probability of $\mathcal{G}$ using the following concentration result.

Proposition C.1 (Weissman et al. (2003))**.** *Let $\mathbf{X}$ be a multinomial random variable with T trials and a vector of probabilities $\mathbf{d}$ of dimension k. Then for any $k \geq 2$ and $\delta \in (0,1)$, we have*

$$\mathbb{P}\left(\|\mathbf{X} - \mathbb{E}[\mathbf{X}]\|_1 \geq \sqrt{2kT\log(2/\delta)}\right) \leq \delta.$$

Recall that $\tilde{\mathbf{D}} := (\tilde{D}_i)_{i\in\{1,\ldots,N+1\}}$ follows a multinomial distribution with T trials and a vector of probability $\mathbf{d}$ (of dimension $N+1$) such that, $d_1 = \frac{3}{4}$ and for $i \in \{2,\ldots,N+1\}$, $d_i = \frac{1}{4N}$.

Therefore,

$$\begin{aligned}
\mathbb{P}(\mathcal{G}) &= \mathbb{P}\left(\max_{i\in\{2,\dots,N+1\}} |\tilde{D}_i - \mathbb{E}\left[\tilde{D}_i\right]| \le \sqrt{4NT\log(2T)}\right)\\
&\ge \mathbb{P}\left(\max_{i\in\{1,\dots,N+1\}} |\tilde{D}_i - \mathbb{E}\left[\tilde{D}_i\right]| \le \sqrt{4NT\log(2T)}\right)\\
&\overset{(a)}{\ge} \mathbb{P}\left(\|\tilde{\mathbf{D}} - \mathbb{E}\left[\tilde{\mathbf{D}}\right]\|_1 \le \sqrt{4NT\log(2T)}\right)\\
&\overset{(b)}{\ge} \mathbb{P}\left(\|\tilde{\mathbf{D}} - \mathbb{E}\left[\tilde{\mathbf{D}}\right]\|_1 \le \sqrt{2(N+1)T\log(2T)}\right)\\
&\overset{(c)}{\ge} 1 - \frac{1}{T},
\end{aligned}$$

where (a) holds because the norm one is larger than the infinite norm, (b) is induced by $2N \ge N+1$ for $N \ge 1$ and (c) is a consequence of proposition C.1. This enables us to conclude that,

$$\mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T(\mathbf{q}^{\mathrm{SAA}}) \mid \mathcal{E}\right] \ge \frac{T}{16} - \frac{7}{2}\sqrt{N^3T\log(2T)} - 2.$$

□

D Proof of lemmas used in Theorem 4.2

Proof of Lemma 4.3. For any $i \in S^*$, we have by definition of Z_i^t that, $\tilde{S}_i = \sum_{t=1}^T Z_i^t$. Therefore,

$$\begin{aligned}
\tilde{S}_i &= \sum_{t=1}^T \mathbb{1}\{\{i\} \prec_t S_{\mathbf{q}_t} \cup \{0\} \setminus \{i\}\} \cdot \mathbb{1}\{i \in S_{\mathbf{q}_t}\}\\
&= \min\left(\sum_{t=1}^T \mathbb{1}\{\{i\} \prec_t S_{\mathbf{q}_t} \cup \{0\} \setminus \{i\}\}, q_i\right)\\
&\overset{(a)}{\ge} \min\left(\sum_{t=1}^T \mathbb{1}\{\{i\} \prec_t \left(\mathcal{F}^{\mathrm{pool}} \cap S_{\mathbf{q}_t}\right) \cup S^* \cup \{0\} \setminus \{i\}\}, q_i\right)\\
&= \min\left(\sum_{t=1}^T X_i^t \cdot \mathbb{1}\{\{i\} \prec_t (\mathcal{F}^{\mathrm{pool}} \cap S_{\mathbf{q}_t}) \cup \{0\} \setminus \{i\}\}, q_i\right)\\
&\overset{(b)}{=} \min\left(\sum_{t=1}^T X_i^t \left(1 - \mathbb{1}\left\{\exists\, j \in \mathcal{F}^{\mathrm{pool}} \cap S_{\mathbf{q}_t} \text{ s.t. } \{j\} \prec_t \{i\}\right\}\right), q_i\right)\\
&\ge \min(X_i, q_i) - \sum_{t=1}^T X_i^t \cdot \mathbb{1}\left\{\exists\, j \in \mathcal{F}^{\mathrm{pool}} \cap S_{\mathbf{q}_t} \text{ s.t. } \{j\} \prec_t \{i\}\right\},
\end{aligned}$$

where (a) follows from the inclusion $S_{\mathbf{q}_t} = \left(\mathcal{F}^{\text{pool}} \cap S_{\mathbf{q}_t}\right) \cup (S^* \cap S_{\mathbf{q}_t}) \subset \left(\mathcal{F}^{\text{pool}} \cap S_{\mathbf{q}_t}\right) \cup S^*$ and (b) holds because, when product $\{i\} \prec_t \{0\}$, we have,

$$\left\{\{i\} \prec_t (\mathcal{F}^{\text{pool}} \cap S_{\mathbf{q}_t}) \cup \{0\} \setminus \{i\}\right\} = \left\{\exists\, j \in \mathcal{F}^{\text{pool}} \cap S_{\mathbf{q}_t} \text{ s.t. } \{j\} \prec_t \{i\}\right\}^c.$$

Next, remark that,

$$\mathbb{1}\left\{\exists\, j \in \mathcal{F}^{\text{pool}} \cap S_{\mathbf{q}_t} \text{ s.t. } \{j\} \prec_t \{i\}\right\} = \sum_{j \in \mathcal{F}^{\text{pool}}} \mathbb{1}\{\{j\} \prec_t (\mathcal{F}^{\text{pool}} \cap S_{\mathbf{q}_t}) \cup \{0\} \setminus \{j\}\} \cdot \mathbb{1}\{\{j\} \prec_t \{i\}\} \cdot \mathbb{1}\{j \in S_{\mathbf{q}_t}\},$$

and that $X_i^t \cdot \mathbb{1}\{\{j\} \prec_t \{i\}\} = X_i^t \cdot X_j^t$ for $i \in S^*$ and $j \in \mathcal{F}^{\text{pool}}$. We also note that,

$$\begin{aligned}
\mathbb{1}\{\{j\} \prec_t (\mathcal{F}^{\text{pool}} \cap S_{\mathbf{q}_t}) \cup \{0\} \setminus \{j\}\} \cdot X_j^t &= \mathbb{1}\{\{j\} \prec_t (\mathcal{F}^{\text{pool}} \cap S_{\mathbf{q}_t}) \cup S^* \cup \{0\} \setminus \{j\}\} \cdot X_j^t \\
&\leq \mathbb{1}\{\{j\} \prec_t S_{\mathbf{q}_t} \cup \{0\} \setminus \{j\}\} \cdot X_j^t,
\end{aligned}$$

where the last inequality holds because of the same argument as in (a). Therefore, we obtain that,

$$\begin{aligned}
\tilde{S}_i &\geq \min(X_i, q_i) - \sum_{t=1}^{T} \sum_{j \in \mathcal{F}^{\text{pool}}} X_i^t \cdot X_j^t \cdot \mathbb{1}\{\{j\} \prec_t S_{\mathbf{q}_t} \cup \{0\} \setminus \{j\}\} \cdot \mathbb{1}\{j \in S_{\mathbf{q}_t}\} \\
&= \min(X_i, q_i) - \sum_{t=1}^{T} \sum_{j \in \mathcal{F}^{\text{pool}}} X_i^t \cdot X_j^t \cdot Z_j^t.
\end{aligned}$$

□

Proof of Lemma 4.4. Let $(\mathbf{x}^\ell)_{\ell \in [L]}$,$(\mathbf{y}^\ell)_{\ell \in [L]}$, $\mathbf{q}^{\text{pool}}$ be an optimal solution of the LP (4.1). For every $\ell \in [L]$, the constraint (4.1c) implies that,

$$\sum_{i=1}^{N} x_i^\ell + \sum_{i=1}^{N} y_i^\ell - \sum_{i=1}^{N} \sum_{k=1, k \neq i}^{N} \rho_{ki} y_k^\ell = \sum_{i=1}^{N} D_i^\ell. \tag{D.1}$$

Moreover,

$$\sum_{i=1}^{N} \sum_{k=1, k \neq i}^{N} \rho_{ki} y_k^\ell = \sum_{k=1}^{N} y_k^\ell \sum_{i=1, i \neq k}^{N} \rho_{ki} = \sum_{k=1}^{N} (1 - \rho_{k0}) y_k^\ell.$$

Therefore, (D.1) implies that,

$$\sum_{i=1}^{N} x_i^\ell = \sum_{i=1}^{N} D_i^\ell - \sum_{i=1}^{N} \rho_{i0} y_i^\ell \leq \sum_{i=1}^{N} D_i^\ell \leq T'.$$

Suppose for the sake of contradiction that $\sum_{i \in \mathcal{N}} c_i q_i^{\text{pool}} > \bar{p} \cdot T'$, hence the optimal objective value of (4.1) verifies

$$Obj((\mathbf{x}^\ell)_{\ell\in[L]}, \mathbf{q}^{\text{pool}}) < \frac{\bar{p}}{L}\sum_{\ell=1}^{L}\sum_{i=1}^{N} x_i^\ell - \bar{p}\cdot T' \leq 0,$$

which is a contradiction since $\mathbf{x}^\ell = \mathbf{y}^\ell = 0$ for all $\ell \in [L]$ and $\mathbf{q} = 0$ is a feasible solution with an objective value equal to 0. We conclude that $\sum_{i\in\mathcal{N}} c_i q_i^{\text{pool}} \leq \bar{p}\cdot T'$. □

E Proof of Proposition 3.4

There is no stochasticity in demand arrival in the instance described by Figure 2. Hence for any $\ell \in [L]$, the sampled demand $\mathbf{D}^\ell$ satisfies

$$\mathbf{D}^\ell = T\mathbf{e_1},$$

where $\mathbf{e_1}$ is the basis vector with the first coefficient being equal to one and the others being equal to 0. In this case, problem (3.3) is equivalent to the following LP.

$$\max_{\mathbf{x}^\ell,\mathbf{y}^\ell,\mathbf{q}} \quad \frac{1}{L}\sum_{\ell=1}^{L}\mathbf{p}^\top\mathbf{x}^\ell - \mathbf{c}^\top\mathbf{q} \tag{E.1a}$$

$$\text{subject to} \quad x_i^\ell \leq q_i, \qquad \forall i \in \{1,\dots,N+2\}, \forall \ell \in [L], \tag{E.1b}$$

$$x_1^\ell + y_1^\ell = T, \qquad \forall \ell \in [L], \tag{E.1c}$$

$$x_i^\ell + y_i^\ell = \frac{y_1^\ell}{N}, \qquad \forall\, i \in \{2,\dots,N+1\}, \forall \ell \in [L], \tag{E.1d}$$

$$x_{N+2}^\ell + y_{N+2}^\ell = \sum_{j=2}^{N+1} y_j^\ell, \qquad \forall \ell \in [L], \tag{E.1e}$$

$$x_i^\ell \geq 0, y_i^\ell \geq 0, q_i \geq 0, \qquad \forall i \in \{1,\dots,N+2\}, \forall \ell \in [L]. \tag{E.1f}$$

For each $\ell \in [L]$, consider the solution defined as follows.

$$\tilde{x}_i^\ell = \begin{cases} 0 & \text{if } i \in \{1, N+2\} \\ \frac{T}{N} & \text{if } i \in \{2,\dots,N+1\}. \end{cases}$$

$$\tilde{y}_i^\ell = \begin{cases} T & \text{if } i = 1 \\ 0 & \text{if } i \in \{2,\dots,N+2\}. \end{cases}$$

$$\tilde{q}_i = \begin{cases} 0 & \text{if } i \in \{1, N+2\} \\ \frac{T}{N} & \text{if } i \in \{2,\dots,N+1\}. \end{cases}$$

It is clear that $(\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}, (\tilde{\mathbf{y}}^\ell)_{\ell\in[L]}, \tilde{\mathbf{q}}$ is feasible for (E.1). Moreover, its objective value is given by,

$$Obj((\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}, \tilde{\mathbf{q}}) = \frac{T}{8}.$$

Finally, for any feasible solution $(\mathbf{x}^\ell)_{\ell\in[L]}, (\mathbf{y}^\ell)_{\ell\in[L]}, \mathbf{q}$ of problem (E.1), the total sales are less than T, i.e., $\sum_{i=1}^{N+2} x_i^\ell \leq T$ for every $\ell \in [L]$. Furthermore, $\sum_{i=2}^{N+2} x_i^\ell \leq \sum_{i=2}^{N+2} q_i$. Therefore,

$$Obj((\mathbf{x}^\ell)_{\ell\in[L]}, \mathbf{q}) \leq \frac{1}{2}\min\left(\sum_{i=2}^{N+2} q_i, T\right) - \frac{3}{8}\sum_{i=2}^{N+2} q_i \leq \frac{T}{8}.$$

Hence, $(\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}, (\tilde{\mathbf{y}}^\ell)_{\ell\in[L]}, \tilde{\mathbf{q}}$ is optimal.

F Proof of Proposition 4.5

Let $\mathbf{q}^{\text{fluid}}$ denote the inventory level of the fluid policy. Note that, by definition, $(S^*,\mathbf{q}^{\text{fluid}})$ is an optimal solution for problem (2.4). It is clear that for the instance described in Figure 2, the optimal solution of (2.4) satisfies $S^* = \{2, \dots, N+1\}$ and

$$q_i^{\text{fluid}} = \begin{cases} 0 & \text{if } i \in \{1, N+2\} \\ \frac{T}{N} & \text{if } i \in \{2, \dots, N+1\}. \end{cases}$$

In the next lemma, we show that the regret of the fluid policy is linear in T for the instance defined in Figure 2.

Lemma F.1. *Let $N = T$ and $\epsilon = \frac{1}{\sqrt{T}}$. For $T \geq 4$, the regret of the fluid policy on the Markov Chain instance π defined in Figure 2 satisfies,*

$$R_\pi^{\textit{fluid}}(T) \geq \frac{1}{8}T - \sqrt{T}.$$

Proof. For a customer $t \in \{1, \dots, T\}$, recall the notation $\prec_t$ defined in the proof of Theorem 4.2. Consider the sequence of random variables $\left(\tilde{D}_i\right)_{i\in\{2,\dots,N+1\}}$ such that for each $i \in \{2, \dots, N+1\}$,

$$\tilde{D}_i = \sum_{t=1}^{T} \mathbb{1}\{\{i\} \prec_t \{0\}\}.$$

When $T = N$, we have $q^{\text{fluid}} = 1$ for all $i \in \{2, \dots, N+1\}$. Thus, for every $i \in \{2, \dots, N+1\}$, product i is purchased if and only if $\{\tilde{D}_i \geq 1\}$. This implies that the profit $\tilde{\mathcal{P}}_T\left(\mathbf{q}^{\text{fluid}}\right)$ during a selling season satisfies,

$$\tilde{\mathcal{P}}_T\left(\mathbf{q}^{\text{fluid}}\right) = \frac{1}{2}\sum_{i=2}^{N+1}\mathbb{1}\{\tilde{D}_i \geq 1\} - \frac{3}{8}\sum_{i=2}^{N+1} q_i^{\text{fluid}} \stackrel{(a)}{=} \frac{T}{8} - \frac{1}{2}\sum_{i=2}^{N+1}\mathbb{1}\{\tilde{D}_i = 0\},$$

where (a) holds because,$\mathbb{1}\{\tilde{D}_i \geq 1\} = 1 - \mathbb{1}\{\tilde{D}_i = 0\}$. For each $i \in \{2, \dots, N+1\}$, $\tilde{D}_i$ follows a binomial distribution with T trials and with success probability equal to $\frac{1}{T}$. Therefore,

$$\mathbb{E}\left[\tilde{\mathcal{P}}_T\left(\mathbf{q}^{\text{fluid}}\right)\right] = \frac{T}{8} - \frac{1}{2}\sum_{i=2}^{N+1}\mathbb{P}\left(\tilde{D}_i = 0\right) = \frac{T}{8} - \frac{1}{2}\left(1 - \frac{1}{T}\right)^T \cdot T \stackrel{(a)}{\leq} \frac{T}{8} - \frac{1}{2e}\left(1 - \frac{1}{T}\right) \cdot T \stackrel{(b)}{\leq} 0,$$

where (a) follows from $e^{-1}\left(1 - \frac{1}{T}\right) \leq \left(1 - \frac{1}{T}\right)^T$ for $T \geq 2$ and (b) holds because, when $T \geq 4$, we have $\frac{1}{2e}\left(1 - \frac{1}{T}\right) \geq \frac{3}{8e} \geq \frac{1}{8}$. We consider the solution $\mathbf{q}^{risk-pool}$ defined by,

$$q_i^{risk-pool} = \begin{cases} 0 & \text{if } i \in \{1, \dots, N+1\} \\ T & \text{if } i = N+2. \end{cases}$$

It is clear that for $\epsilon = \frac{1}{\sqrt{T}}$,

$$\mathcal{P}_T(\mathbf{q}^{risk-pool}) = \left(\frac{1}{8} - \epsilon\right)T = \frac{1}{8}T - \sqrt{T}.$$

Therefore the regret of the fluid policy satisfies,

$$R_\pi^{\text{fluid}}(T) = \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^*) - \mathcal{P}_T\left(\mathbf{q}^{\text{fluid}}\right)\right] \geq \mathcal{P}_T(\mathbf{q}^{risk-pool}) - \mathcal{P}_T\left(\mathbf{q}^{\text{fluid}}\right) \geq \frac{1}{8}T - \sqrt{T}.$$

□

Let $\mathbf{q}^{\text{LCB-SAA}}$ denote the inventory solution prescribed by Algorithm 2 on the Markov Chain instance π defined in Figure 2. In the next lemma, we show that $z_{\textsf{fluid}} - \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^{\text{LCB-SAA}})\right] = O(\sqrt{T})$, and since $z_{\textsf{fluid}} \geq \mathcal{P}_T(\mathbf{q}^*)$, by Proposition 2.1, we conclude that the regret of LCB-SAA is bounded by $\mathcal{O}\left(\sqrt{T}\right)$.

Lemma F.2. *Let $\epsilon = \frac{1}{\sqrt{T}}$. For $\gamma > 0$, $T \geq 64$, and $L \geq 2500\log(T)$, we have that,*

$$z_{\textsf{fluid}} - \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^{LCB\text{-}SAA})\right] \leq \sqrt{T} + \frac{5}{8}.$$

Proof. Recall the optimal unconstrained assortment S^* is $\{2, \dots, N+1\}$ for the instance described in Figure 2 and $z_{\textsf{fluid}} = \frac{T}{8}$. The solution given by Algorithm 2 is the solution of problem (4.1) where, for any $\ell \in [L]$, the demand scenarios $\mathbf{D}^\ell$ are sampled from a multinomial with probability $\pi(\cdot, S^*)$ and T trials. Moreover, the reduced prices are: $p_i^{S^*} = 0$ for $i \in S^*$, $p_1^{S^*} = 0$ and $p_{N+2}^{S^*} = \frac{1}{2} - \epsilon$. So for any threshold $\gamma > 0$, we get $\mathcal{F} = \mathcal{N}$. The LP (4.1) for this instance is given by,

$$
\begin{aligned}
\max_{(\mathbf{x}^\ell)_{\ell\in[L]},(\mathbf{y}^\ell)_{\ell\in[L]},\mathbf{q}} \quad & \frac{1}{L}\sum_{\ell=1}^{L} \mathbf{p}^\top \mathbf{x}^\ell - \mathbf{c}^\top \mathbf{q} && \text{(F.1a)}\\
\text{subject to} \quad & x_i^\ell \le q_i, \qquad \forall i \in \{1,\dots,N+2\}, \forall \ell \in [L], && \text{(F.1b)}\\
& x_1^\ell + y_1^\ell = 0, \qquad \forall \ell \in [L], && \text{(F.1c)}\\
& x_i^\ell + y_i^\ell = D_i^\ell + \frac{1}{N} y_1^\ell, \qquad \forall\, i \in \{2,\dots,N+1\}, \forall \ell \in [L], && \text{(F.1d)}\\
& x_{N+2}^\ell + y_{N+2}^\ell = \sum_{j=2}^{N+1} y_j^\ell, \qquad \forall \ell \in [L], && \text{(F.1e)}\\
& x_i^\ell \ge 0, y_i^\ell \ge 0, q_i \ge 0, \qquad \forall i \in \{1,\dots,N+2\}, \forall \ell \in [L]. && \text{(F.1f)}
\end{aligned}
$$

First, remark that the constraint (F.1c) implies that for every $\ell \in [L]$, $x_1^\ell = 0$. Therefore, for any optimal solution, we must have $q_1 = 0$. Moreover, note that for every $i \in S^*$, $\pi(i, S^*) = \frac{1}{T}$. As a consequence, for every $\ell \in [L]$, $T \ge 4$ and $i \in \{2,\dots,N+1\}$,

$$
\mathbb{P}\left(D_i^\ell = 0\right) = \left(1 - \frac{1}{T}\right)^T \overset{(a)}{\ge} e^{-1}\left(1 - \frac{1}{T}\right) \overset{(b)}{\ge} \frac{3}{4e}, \tag{F.2}
$$

where (a) follows from $e^{-1}\left(1 - \frac{1}{T}\right) \le \left(1 - \frac{1}{T}\right)^T$ for $T \ge 2$, and (b) holds because $T \ge 4$. For any $i \in \{2,\dots,N+1\}$, denote by

$$
N_i = \sum_{\ell\in[L]} \mathbb{1}\{D_i^\ell = 0\}.
$$

and define the event $\mathcal{B}_i := \left\{N_i < \frac{3L}{4e} - \sqrt{L\log(T)}\right\}$. Let $\mathcal{B} := \bigcup_{i=2}^{N+1} \mathcal{B}_i$. Let $(\mathbf{x}^\ell)_{\ell\in[L]}, (\mathbf{y}^\ell)_{\ell\in[L]}, \mathbf{q}$ be an optimal solution for problem (F.1). We now show that, when $\mathcal{B}^c$ holds, we get $q_i = 0$ for every $i \in \{2,\dots,N+1\}$. First, conditional on $\mathcal{B}^c$, we have for all $i \in \{2,\dots,N+1\}$,

$$
\begin{aligned}
\frac{1}{2L}\sum_{\ell\in[L]} x_i^\ell - \frac{3q_i}{8} &\le \frac{1}{2L}(L - N_i)\cdot \max_{\ell\in[L]} x_i^\ell - \frac{3q_i}{8}\\
&\overset{(a)}{\le} \left(\frac{1}{2} - \frac{3}{8e} + \frac{1}{2}\sqrt{\frac{\log(T)}{L}}\right)\cdot \max_{\ell\in[L]} x_i^\ell - \frac{3q_i}{8}\\
&\overset{(b)}{\le} \left(\frac{1}{8} - \frac{3}{8e} + \frac{1}{2}\sqrt{\frac{\log(T)}{L}}\right) q_i,
\end{aligned} \tag{F.3}
$$

where (a) holds because $\mathcal{B}^c$ holds and (b) follows from $x_i^\ell \le q_i$ for every $\ell \in [L]$. For the sake of simplicity of notation, let $C_L := \frac{1}{8} - \frac{3}{8e} + \frac{1}{2}\sqrt{\frac{\log(T)}{L}}$ and remark that $\frac{1}{8} - \frac{3}{8e} < -\frac{1}{100}$, thus for every $L \ge 2500\log(T)$, we have $C_L < 0$.

For the sake of contradiction, assume that $\sum_{i=2}^{N+2} q_i > 0$. Then, conditionally on $\mathcal{B}^c$, the objective value of problem (F.1) satisfies,

$$
\begin{aligned}
Obj\left((\mathbf{x}^\ell)_{\ell\in[L]},\mathbf{q}\right) &= \sum_{i=2}^{N+1}\left(\frac{1}{2L}\sum_{\ell=1}^{L}x_i^\ell - \frac{3q_i}{8}\right) + \frac{1}{L}\sum_{\ell=1}^{L}\left(\frac{1}{2}-\frac{1}{\sqrt{T}}\right)x_{N+2}^\ell - \frac{3q_{N+2}}{8} \\
&\overset{(a)}{\le} C_L\sum_{i=2}^{N+1}q_i + \frac{1}{L}\sum_{\ell=1}^{L}\left(\frac{1}{2}-\frac{1}{\sqrt{T}}\right)x_{N+2}^\ell - \frac{3q_{N+2}}{8} \\
&\overset{(b)}{<} \left(\frac{1}{2}-\frac{1}{\sqrt{T}}\right)\min(q_{N+2},T) - \frac{3q_{N+2}}{8} \le \frac{T}{8}-\sqrt{T},
\end{aligned}
$$

where (a) is follows from (F.3) and (b) holds because $C_L < 0$ and, for each $\ell \in [L]$, we have $x_{N+2}^\ell \le T$ and $x_{N+2}^\ell \le q_{N+2}$. Consider the solution $(\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}, (\tilde{\mathbf{y}}^\ell)_{\ell\in[L]}, \tilde{\mathbf{q}}$ such that, for $\ell \in [L]$,

$$
\begin{aligned}
\tilde{x}_i^\ell &= \begin{cases} 0 & \text{if } i \in \{1,\dots,N+1\} \\ T & \text{if } i = N+2, \end{cases} \\
\tilde{y}_i^\ell &= \begin{cases} 0 & \text{if } i = 1 \\ D_i^\ell & \text{if } i \in \{2,\dots,N+1\} \\ 0 & \text{if } i = N+2. \end{cases} \\
\tilde{q}_i &= \begin{cases} 0 & \text{if } i \in \{1,\dots,N+1\} \\ T & \text{if } i = N+2, \end{cases}
\end{aligned}
$$

It is clear that $(\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}, (\tilde{\mathbf{y}}^\ell)_{\ell\in[L]}, \tilde{\mathbf{q}}$ is feasible for (F.1). Moreover, its objective value satisfies,

$$
Obj\left((\tilde{\mathbf{x}}^\ell)_{\ell\in[L]},\tilde{\mathbf{q}}\right) = \left(\frac{1}{2}-\sqrt{T}\right)T - \frac{3}{8}T = \left(\frac{1}{8}-\frac{1}{\sqrt{T}}\right)T = \frac{T}{8}-\sqrt{T}.
$$

This contradicts the optimality of $(\mathbf{x}^\ell)_{\ell\in[L]}, (\mathbf{y}^\ell)_{\ell\in[L]}, \mathbf{q}$. Thus, when $\mathcal{B}^c$ holds, we get $\sum_{i=2}^{N+1} q_i = 0$, i.e., $q_i = 0$ for all $i \in \{1,\dots,N+1\}$, and therefore, for $T \ge 64$,

$$
Obj\left((\mathbf{x}^\ell)_{\ell\in[L]},\mathbf{q}\right) = \left(\frac{1}{2}-\frac{1}{\sqrt{T}}\right)\frac{1}{L}\sum_{\ell=1}^{L}x_{N+2}^\ell - \frac{3q_{N+2}}{8} \le \left(\frac{1}{2}-\frac{1}{\sqrt{T}}\right)\min\left(T,q_{N+2}\right) - \frac{3q_{N+2}}{8} \le \frac{T}{8}-\sqrt{T}.
$$

Hence, when $\mathcal{B}^c$ holds, the solution $(\tilde{\mathbf{x}}^\ell)_{\ell\in[L]}, (\tilde{\mathbf{y}}^\ell)_{\ell\in[L]}, \tilde{\mathbf{q}}$ is optimal for (F.1) and,

$$
\mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^{\text{LCB-SAA}})\right] \overset{(a)}{\ge} \mathbf{E}\left[\mathcal{P}_T(\tilde{q}) \mid \mathcal{B}^c\right]\mathbb{P}\left(\mathcal{B}^c\right) - \frac{T}{2}\mathbb{P}\left(\mathcal{B}\right) = \left(\frac{T}{8}-\sqrt{T}\right)\mathbb{P}\left(\mathcal{B}^c\right) - \frac{T}{2}\mathbb{P}\left(\mathcal{B}\right),
$$

where (a) follows from Lemma 4.4 and $\max_{i\in\mathcal{N}} p_i = \frac{1}{2}$.

On the other hand, note that, for $i \in \{2,\dots,N+1\}$, N_i is the sum of L bounded i.i.d. random variables in $[0,1]$, therefore,

$$\mathbb{P}\left(\mathcal{B}_i\right) \overset{(a)}{\leq} \mathbb{P}\left(N_i < \mathbb{E}\left[N_i\right] - \sqrt{L\log\left(T\right)}\right) \overset{(b)}{\leq} e^{-2\frac{L\log(T)}{L}} = \frac{1}{T^2},$$

where (a) is a consequence of (F.2) and (b) comes from the Hoeffding inequality for bounded random variables (Theorem 2 in Hoeffding (1994)). By a union bound, we have that,

$$\mathcal{P}\left(\mathcal{B}\right) \leq \sum_{i=2}^{N+1} \mathcal{P}\left(\mathcal{B}_i\right) \leq \frac{N}{T^2} = \frac{1}{T},$$

Recall that $z_{\mathsf{fluid}} = \frac{T}{8}$. We conclude that,

$$z_{\mathsf{fluid}} - \mathbb{E}\left[\mathcal{P}_T(\mathbf{q}^{\text{LCB-SAA}})\right] \leq \frac{T}{8} - \left(\left(\frac{T}{8} - \sqrt{T}\right)\mathbb{P}\left(\mathcal{B}^c\right) - \frac{T}{2}\mathbb{P}\left(\mathcal{B}\right)\right) \leq \sqrt{T} + \frac{5}{8}.$$

where the last inequality follows from $\mathcal{P}\left(\mathcal{B}\right) \leq \frac{1}{T}$.

□

G Proof of Proposition 5.1

Let $\mathcal{H}_T$ be the set of historical data. Recall that $\hat{\boldsymbol{\beta}}$ is known to the decision-maker. By independence we have,

$$\mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(\mathcal{H}_T\right) = \prod_{t=1}^{T} \mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(i_t \mid C_t \text{ is considered, } e_t\right) \cdot \mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(C_t \text{ is considered, } e_t \mid \sigma_t,\, S_t,\, \Psi(q_t)\right) \cdot P_t,$$

where for any event $\mathcal{A}$, we shorten notation as $\mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(\mathcal{A}\right) = \mathbb{P}\left(\mathcal{A} \mid \mu, \boldsymbol{\theta}, \boldsymbol{\beta}\right)$. Moreover, we note that $P_t := \mathbb{P}\left(\sigma_t,\, S_t,\, \Psi(q_t) \mid \mu, \boldsymbol{\theta}, \boldsymbol{\beta}\right)$ is a constant that does not depend on $\mu, \boldsymbol{\theta}, \boldsymbol{\beta}$ but only on $\hat{\boldsymbol{\beta}}$ which is known. Recall that,

$$\mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(i_t \mid C_t \text{ is considered, } e_t\right) = \frac{e^{\boldsymbol{\beta}^\top \mathbf{y_{i_t}}}}{1 + \sum_{k\in C_t} e^{\boldsymbol{\beta}^\top \mathbf{y_k}}}.$$

Recall σ_t is the ranking of S_t and let $i_t^* := \max_{i\in C_t} \sigma_t(i)$. For the sake of simplicity of notation, for every $e \in \{0,1\}$, let

$$p_t(e) := \mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(C_t \text{ is considered, } e_t = e \mid \sigma_t,\, S_t,\, \Psi(q_t)\right).$$

Then,

$$\begin{aligned}
p_t(0) &= \frac{1}{1+e^{\mu}}\left(\frac{e^{\mu}}{1+e^{\mu}}\right)^{|C_t|-1}\prod_{i\in C_t}\mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(i\in C_t\right)\prod_{i\notin C_t,\sigma_t(i)\le i_t^*}\mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(i\notin C_t\right)\\
&= \frac{1}{1+e^{\mu}}\left(\frac{e^{\mu}}{1+e^{\mu}}\right)^{|C_t|-1}\prod_{i\in C_t}\frac{e^{\boldsymbol{\theta}^\top\mathbf{x_i}}}{1+e^{\boldsymbol{\theta}^\top\mathbf{x_i}}}\prod_{i\notin C_t,\sigma_t(i)\le i_t^*}\frac{1}{1+e^{\boldsymbol{\theta}^\top\mathbf{x_i}}}\\
&= \frac{1}{1+e^{\mu}}\left(\frac{e^{\mu}}{1+e^{\mu}}\right)^{|C_t|-1}\prod_{i\in C_t}e^{\boldsymbol{\theta}^\top\mathbf{x_i}}\prod_{i\in S_t,\sigma_t(i)\le i_t^*}\frac{1}{1+e^{\boldsymbol{\theta}^\top\mathbf{x_i}}}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
p_t(1) &= \left(\frac{e^{\mu}}{1+e^{\mu}}\right)^{|C_t|}\prod_{i\in C_t}\mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(i\in C_t\right)\prod_{i\notin C_t}\mathbb{P}_{\mu,\boldsymbol{\theta},\boldsymbol{\beta}}\left(i\notin C_t\right)\\
&= \left(\frac{e^{\mu}}{1+e^{\mu}}\right)^{|C_t|}\prod_{i\in C_t}\frac{e^{\boldsymbol{\theta}^\top\mathbf{x_i}}}{1+e^{\boldsymbol{\theta}^\top\mathbf{x_i}}}\prod_{i\notin C_t}\frac{1}{1+e^{\boldsymbol{\theta}^\top\mathbf{x_i}}}\\
&= \left(\frac{e^{\mu}}{1+e^{\mu}}\right)^{|C_t|}\prod_{i\in C_t}e^{\boldsymbol{\theta}^\top\mathbf{x_i}}\prod_{i\in S_t}\frac{1}{1+e^{\boldsymbol{\theta}^\top\mathbf{x_i}}}.
\end{aligned}$$

Therefore, the log-likelihood is given by,

$$\begin{aligned}
\mathcal{L}(\mu,\theta,\beta) = \sum_{t=1}^{T}\Bigg\{&|C_t|\left(\mu-\log\left(1+e^{\mu}\right)\right)+\sum_{i\in C_t}\boldsymbol{\theta}^\top\mathbf{x_i}-\sum_{i\in S_t,\sigma_t(i)\le i_t^*}\log(1+e^{\boldsymbol{\theta}^\top\mathbf{x_i}})\\
&-e_t\sum_{i\in S_t,\sigma_t(i)>i_t^*}\log(1+e^{\boldsymbol{\theta}^\top\mathbf{x_i}})-(1-e_t)\cdot\mu\\
&+\boldsymbol{\beta}^\top\mathbf{y_{i_t}}-\log(1+\sum_{k\in C_t}e^{\boldsymbol{\beta}^\top\mathbf{y_k}})+\log\left(P_t\right)\Bigg\}.
\end{aligned}$$

Moreover, for any $n\ge 1$, $(x_1,\ldots,x_n)\mapsto\log(1+\sum_{i=1}^n e^{x_i})$ is a convex function. Therefore, $\mathcal{L}$ is a sum of concave functions so it is concave.