Submitted to *Operations Research*
manuscript (Please, provide the manuscript number!)

Near-Optimal Algorithms for Capacity Constrained Assortment Optimization

Antoine Désir
Department of Industrial Engineering and Operations Research, Columbia University ad2918@columbia.edu

Vineet Goyal
Department of Industrial Engineering and Operations Research, Columbia University vg2277@columbia.edu

Assortment optimization is an important problem that arises in many practical applications such as retailing and online advertising. In an assortment optimization problem, the goal is to select a subset of items that maximizes the expected revenue in the presence of the substitution behavior of consumers specified by a *choice model*. In this paper, we consider the capacity constrained version of the assortment optimization problem under several choice models including Multinomial logit (MNL), Nested Logit (NL) and the mixture of Multinomial logit (MMNL) models. The goal is to select a revenue maximizing subset of items with total weight or capacity at most a given bound. We present a fully polynomial time approximation scheme (FPTAS) for these models when the number of mixtures or nests is constant. Our FPTAS uses ideas similar to the FPTAS for the knapsack problem.

The running time of our algorithm depends exponentially on the number of mixtures in the MMNL model. We show that surprisingly the exponential dependence on the number of mixtures is necessary for any near-optimal algorithm for the MMNL choice model. In particular, we show that there is no algorithm with running time polynomial in the number of items, n and mixtures, K that obtains an approximation better than $O(1/K^{1-\delta})$ for any $\delta > 0$ for even the unconstrained assortment optimization over a general MMNL model. Our reduction provides a procedure to construct a natural family of hard benchmark instances for the assortment optimization problem over MMNL that may be of independent interest. These instances are quite analogous to the consideration set based models (Jagabathula and Rusmevichientong (2014)) where the consideration set arises from a graphical model. We also present some special cases of MMNL and NL models where we can obtain an FPTAS with a polynomial dependence on the number of mixtures.

Key words: Assortment Optimization, FPTAS

1. Introduction

Assortment optimization problems arise widely in many practical applications such as retailing and online advertising. One of the key operational decision faced by a retailer is to select a subset of items to offer from a universe of n substitutable items, that maximizes the expected revenue. The demand of any item depends on the set of offered items due to substitution behavior of consumers. For a given substitution behavior of consumers, the goal in the assortment optimization problem is to find a subset of items that maximizes the total expected revenue. In this paper, we study a capacity constrained assortment optimization problem for a large family of choice models where there is a constraint on the total weight or capacity of the items selected in the offer set. The capacity constraint on the assortment is a natural constraint that arises in many applications from budget or space constraints.

The substitution behavior of consumers is captured by a choice model that for every $S \subseteq [n], j \in S$, specifies the probability that a random consumer selects item j in offer set, S. Many parametric choice models have been extensively studied in the literature in several areas including marketing, transportation, economics and operations management (see McFadden (1980), Ben-Akiva and Lerman (1985) and Wierenga (2008)). The Multinomial logit (MNL) model is by far the most popular model in practice. It was introduced independently by Luce (1959) and Plackett (1975) and was referred to as the Plackett-Luce model. It came to be known as the Multinomial logit model after the work of McFadden (1978). The popularity of the MNL model arises from the tractability of estimation and the corresponding assortment optimization problem for this model (see Talluri and Van Ryzin (2004)); even though some of the model justifications (for instance, Independence of Irrelevant Alternatives property (IIA) property (see Ben-Akiva and Lerman (1985)) of the MNL model) are not reasonable for many applications. A more complex choice model can capture a richer substitution behavior but leads to increased complexity of the assortment optimization problem. McFadden and Train (2000) show that any choice model arising from the random utility model can be approximated as closely as required by a mixture of a finite (but unknown) number of MNL models. In general, the problem of selecting the right model that explains the substitution behavior of consumers is a challenging one.

1.1. Our Contributions

In this paper, we assume that the choice model is given and focus on the assortment optimization problem under capacity constraint for a fairly general class of choice models. Rusmevichientong et al. (2014) show that the unconstrained assortment optimization is NP-hard for the mixture of MNL model (MMNL) even for the case of mixture of two MNL models. They give a PTAS for the assortment optimization for the MMNL model with a constant number of MNL models in the mixture where the running time depends exponentially on $1/\epsilon$ for accuracy level $\epsilon > 0$. Our main contributions are the following.

FPTAS for Capacitated Assortment. We study the capacity constrained assortment optimization problem for MMNL choice model, Cap-MMNL-Assort, with a constant number of mixtures and present a fully polynomial time approximation scheme (FPTAS) for the problem. In other words, for any $\epsilon > 0$, our algorithm computes a $(1-\epsilon)$-approximation of the optimal assortment in time polynomial in the input size and $1/\epsilon$. This is the best possible approximation for a NP-hard problem. Furthermore, we show that even for the case of MNL model, the capacity constrained assortment optimization is NP-hard. Therefore, our algorithm gives the best possible approximation for the capacity constrained assortment optimization problem even for the MNL model.

Our FPTAS also gives a $(1-\epsilon)$-approximation for the following more general capacitated sum of ratio optimization problem, Cap-Sum-Ratios

$$\max_{\boldsymbol{x}\in\{0,1\}^n} \left\{ \sum_{k=1}^{K} \frac{\boldsymbol{c}_k^T \boldsymbol{x}}{1+\boldsymbol{d}_k^T \boldsymbol{x}} \;\middle|\; \sum_{j=1}^{n} w_j x_j \le W \right\},$$

where K is a constant. Rusmevichientong et al. (2009) give a PTAS for the Cap-Sum-Ratios problem for constant K based on a linear programming formulation. Mittal and Schulz (2013) give an FPTAS for the Cap-Sum-Ratios problem. However, they use a black-box construction of an approximate Pareto-optimal frontier introduced by Papadimitriou and Yannakakis (2000).

We present an explicit FPTAS for Cap-MMNL-Assort based on ideas from the FPTAS for the knapsack problem. In particular, we reduce the Cap-MMNL-Assort to a variant of the multi-dimensional knapsack problem. We would like to note that the multi-dimensional knapsack problem does not admit a FPTAS (see Frieze and Clarke (1984)). However, in our reformulation to the multi-dimensional knapsack problem, we can violate some of the knapsack constraints which allows us to obtain an FPTAS. Our approach is quite general and can be adapted for more general constrained assortment optimization including joint cardinality and capacity constraints.

We would like to note that the running time of our algorithm is polynomial in the input size and $1/\epsilon$, but is exponential in K (number of mixtures in the MMNL model). Therefore, we obtain an FPTAS only when the MMNL model is a mixture of a constant number of MNL models. We show that this exponential (or super-polynomial) dependence in the number of mixtures is necessary for any near-optimal algorithm for the MMNL-Assort problem.

Hardness of approximation for the MMNL model. We show that the unconstrained assortment optimization for the MMNL model, MMNL-Assort, is hard to approximate within any reasonable factor when the number of mixtures is not constant. More specifically, there is no polynomial time algorithm (polynomial in number of items and mixtures: n, K and the input size) with an approximation factor better than $O(1/K^{1-\delta})$ for any constant $\delta > 0$ for the unconstrained MMNL-Assort unless $NP \subseteq BPP$. This implies that if we require a near-optimal algorithm for the assortment optimization over the MMNL model, a super-polynomial dependence on the number of mixtures is necessary.

We prove the hardness by an approximation preserving reduction from the maximum independent set problem. In our reduction, we consider an MMNL instance where there is a product as well as a MNL segment corresponding to each vertex in the instance for the independent set problem. The MNL segment for any vertex only contains products corresponding to neighbors of that vertex. This is quite analogous to the consideration set model considered in Jagabathula and Rusmevichientong (2014) where we can think of the consideration set arising from a natural graphical model. Our reduction provides a natural family of hard benchmark instances for MMNL-Assort that may be of independent interest.

Polynomial dependence in K: Special Cases. We present special cases of the Nested logit and MMNL models where we can get an FPTAS with running time polynomial in the number of nests or mixtures. Davis et al. (2011) consider the unconstrained NL-Assort problem and give a polynomial time algorithm for a special case of parameters, namely, the dissimilarity parameter for each nest is smaller than one and the utility of the no purchase option is 0 for each nest. We give an FPTAS for the Cap-NL-Assort for this special case with running time polynomial in the number of nests.

We also consider a special case of the MMNL model considered in Rusmevichientong et al. (2014) where the different MNL segments differ only in the utility of the no-purchase option. We show that we can adapt our framework to obtain an FPTAS for the capacity constrained version with running time polynomial in the number of mixtures.

Outline. The rest of the paper is organized as follows. In Section 2, we present an LP based optimal algorithm for the cardinality constrained assortment optimization for the MNL model and show that the capacity constrained version is NP-hard even for this model. In Section 3 we present the FPTAS for Cap-MMNL-Assort and Cap-NL-Assort problems. Next, in Section 4, we present the hardness of approximation result for MMNL-Assort. Finally, we discuss the special cases of NL and MMNL in Section 5.

2. Capacitated Assortment Optimization for MNL model

In this section, we consider the assortment optimization problem under capacity constraint for the MNL choice model. The MNL model is given by $(n+1)$ parameters $u_0, \ldots, u_n$ which represent the preference weights of each item as well as the preference weight of the no purchase option. For any $S \subseteq [n]$, $j \in S_+ = S \cup \{0\}$, the choice probability of product j is given by

$$\pi(j,S) = \frac{u_j}{\sum_{i \in S_+} u_i}.$$

Each item $i \in [n]$ is also assigned a revenue r_j and a weight w_i. We denote by W the total available capacity. The capacity constrained assortment optimization, Cap-MNL-Assort, can be formulated as follows.

$$\max_{\boldsymbol{x} \in \{0,1\}^n} \left\{ \sum_{j=1}^{n} r_j \frac{u_j x_j}{u_0 + \sum_{i=1}^{n} u_i x_i} \;\middle|\; \sum_{j=1}^{n} w_j x_j \leq W \right\}.$$

2.1. Special case of Cardinality Constraints

We first consider the special case of cardinality constrained assortment problem, Cardinality-MNL-Assort, where there is an upper bound on the number of products in the assortment. We present an LP based optimal algorithm for this case. We would like to note that Davis et al. (2013) and Rusmevichientong et al. (2010) give an optimal algorithm for cardinality constrained assortment problem for the MNL and a special case of NL choice model by using structural properties of the optimal solution. However, our proof of optimality for the LP based algorithm is based on the properties of an optimal extreme point solution. In particular, we prove the following theorem.

Theorem 1. Cardinality-MNL-Assort *is equivalent to the following linear program*

$$z_{LP} = \max \left\{ \sum_{j=1}^{n} r_j p_j \,\middle|\, u_0 p_0 + \sum_{j=1}^{n} p_j = 1 \,,\, \sum_{j=1}^{n} \frac{p_j}{u_j} \leq k p_0 \,,\, 0 \leq p_j \leq u_j p_0 \right\}, \tag{1}$$

where k is the upper bound on the number of items in the assortment. Furthermore, if $\boldsymbol{p}^*$ is an optimal solution, then $S^* = \{j \,|\, p_j^* = u_j p_0^*\}$ is an optimal assortment.

Proof. We first show that the above LP is a relaxation of Cardinality-MNL-Assort. For any feasible solution $S \subseteq [n]$ for Cardinality-MNL-Assort, we have the following feasible solution to the LP

$$p_0 = \frac{1}{\sum_{i \in S_+} u_i} \quad \text{and} \quad p_j = \begin{cases} \frac{u_j}{\sum_{i \in S_+} u_i} & \text{if } j \in S \\ 0 & \text{otherwise} \end{cases} \quad \forall j \geq 1.$$

Moreover, the two solutions give the same objective value which implies that $z_{LP} \geq z^*$.

We now show that any basic solution $\boldsymbol{p}^*$ to (1) satisfies $p_j^* \in \{0, u_j p_0^*\}$ for all $j \in [n]$. We have $n+1$ variables in (1) and only one equality constraint. Therefore, in a basic optimal solution, at least n inequalities are tight among

$$\sum_{j=1}^{n} \frac{p_j}{u_j} \leq k p_0 \quad \text{and} \quad 0 \leq p_j \leq u_j p_0 \quad \forall j \geq 1.$$

Consequently, $p_j \in \{0, u_j p_0\}$ for at least $(n-1)$ variables. Suppose exactly $(n-1)$ variables satisfy $p_j^* \in \{0, u_j p_0^*\}$ and one of the variable, say p_1^*, satisfies $0 < p_1^* < u_1 p_0^*$. Therefore, the inequality $\sum_{j=1}^{n} \frac{p_j}{u_j} \leq k p_0$ must be tight and

$$k p_0^* = \sum_{j=1}^{n} \frac{p_j^*}{u_j} = \frac{p_1^*}{u_1} + \sum_{j=2}^{n} \frac{p_j^*}{u_j} = \rho p_0 + k' p_0$$

where k' is an integer and $0 < \rho < 1$. This yields a contradiction. Therefore, any basic solution leads to an integral solution of the original problem which means that $z_{LP} \leq z^*$. □

2.2. General Capacity Constraints: Hardness

We show that the general capacity constrained assortment optimization, Cap-MNL-Assort, is NP-hard even for the MNL choice model. We prove this by a reduction from the knapsack problem.

THEOREM 2. Cap-MNL-Assort *is NP-hard.*

Proof We give a reduction from the knapsack. In an instance of the knapsack problem on n items, we are given weights $c_1, \ldots, c_n$ and profits $p_1, \ldots, p_n$ and a knapsack capacity C. The goal is to find the most profitable assortment of items.

Consider the following instance for Cap-MNL-Assort:

$$u_0 = 1, \quad W = C \quad \text{and} \quad \forall j \geq 1, \quad u_j = p_j, \quad r_j = 1, \quad w_j = c_j.$$

For this instance, the problem becomes

$$\max_{\boldsymbol{x} \in \{0,1\}^n} \left\{ \frac{\sum_{i=1}^{n} p_i x_i}{1 + \sum_{i=1}^{n} p_i x_i} \;\middle|\; \sum_{i=1}^{n} c_i x_i \leq C \right\}.$$

Note that the function $f(x) = \dfrac{x}{1+x}$ is increasing in x. Therefore, maximizing $f(\boldsymbol{p}^{\boldsymbol{T}}\boldsymbol{x})$ is equivalent to maximizing $\boldsymbol{p}^{\boldsymbol{T}}\boldsymbol{x}$, hence the reduction to the knapsack problem. □

In view of the above hardness, a natural goal is to design near-optimal algorithms for the capacity constrained assortment optimization problem. We do that in the following section.

3. FPTAS for Capacitated Assortment under MMNL model

In this section, we present an FPTAS for the capacity constrained assortment optimization problem for a mixture of MNL model (Cap-MMNL-Assort). The MMNL model is a generalization of the MNL model and is given by a distribution over K different MNL models. For all $k \in [K]$ and $j \in [n]$, let $u_{j,k}$ denote the MNL parameters for segment k and θ_k denote the probability of segment k. For any $S \subseteq [n]$, $j \in S_+ = S \cup \{0\}$, the choice probability of product j is given by

$$\pi(j, S) = \sum_{k=1}^{K} \theta_k \frac{u_{j,k}}{\sum_{i \in S_+} u_{i,k}}.$$

Each item $i \in [n]$ has revenue r_j and weight w_i. Let W denote the total available capacity. The Cap-MMNL-Assort can be formulated as follows.

$$\max_{\boldsymbol{x} \in \{0,1\}^n} \left\{ \sum_{k=1}^{K} \theta_k \frac{\sum_{j=1}^{n} r_j u_{j,k} x_j}{u_{0,k} + \sum_{j=1}^{n} u_{j,k} x_j} \;\middle|\; \sum_{j=1}^{n} w_j x_j \leq W \right\}$$

As mentioned earlier, even the unconstrained problem is NP-hard even for the mixture of two MNL choice models (see Rusmevichientong et al. (2014) where the authors give a PTAS for the unconstrained problem for a constant number of mixtures). Here, we present an FPTAS for the Cap-MMNL-Assort problem when the number of mixtures is constant. Our algorithm utilizes the rational structure of the objective function and is based on solving a polynomial number of dynamic programs. Since the objective function is a sum of ratios, we guess the value of each numerator ($\sum_{j\in S^*} r_j u_{j,k}$) and each denominator ($\sum_{j\in S^*} u_{j,k}$), for an optimal solution, S^* within a factor of $(1+\epsilon)$. We then try to find a feasible assortment (satisfying the capacity constraint) with the numerator and denominator values approximately equal to the guesses using a dynamic program that is similar in spirit to the FPTAS for the knapsack problem (see for example Lawler (1979)).

Let r (resp. R) be the minimum (resp. maximum) revenue and u (resp. U) be the minimum (resp. maximum) value of the utility parameters over all segments. We assume wlog. that $u_{j,k} > 0$ for all j,k. Otherwise, we can replace $u_{j,k}$ by $\hat{u}_{j,k} = \epsilon ur/(nR)$ for all j, k such that $u_{j,k} = 0$ where $u = \min\{u_{i,k} \mid u_{i,k} > 0\}$. This only changes the objective function by a factor of $(1+\epsilon)$ (see Appendix A). For a given $\epsilon > 0$, we use the following set of guesses.

$$\Gamma_{\epsilon,K} = (\Gamma_\epsilon)^K \quad \text{and} \quad \Delta_{\epsilon,K} = (\Delta_\epsilon)^K.$$

where

$$\Gamma_\epsilon = \{ru(1+\epsilon)^\ell, \ell = 0,\dots,L_1\} \text{ and } \Delta_\epsilon = \{u(1+\epsilon)^\ell, \ell = 0,\dots,L_2\}, \tag{2}$$

and $L_1 = O(\log(nRU/r)/\epsilon)$ and $L_2 = O(\log((n+1)U/r)/\epsilon)$. Note that for constant K, the number of guesses is polynomial in the input size and $1/\epsilon$. For a given guess $(\boldsymbol{h},\boldsymbol{g}) \in \Gamma_{\epsilon,K} \times \Delta_{\epsilon,K}$, we discretize the coefficients as follows,

$$\tilde{r}_{i,k} = \left\lfloor \frac{r_i u_{i,k}}{\epsilon h_k/n} \right\rfloor \quad \text{and} \quad \tilde{u}_{i,k} = \left\lceil \frac{u_{i,k}}{\epsilon g_k/(n+1)} \right\rceil. \tag{3}$$

We use a dynamic program to find a feasible assortment S such that for all $k \in [K]$

$$\sum_{j\in S} r_j u_{j,k} \geq h_k \quad \text{and} \quad \sum_{j\in S} u_{j,k} \leq g_k. \tag{4}$$

Let us now present the dynamic program. For each $(\boldsymbol{i},\boldsymbol{j},p) \in [I]^K \times [J]^K \times [n]$, let $F(\boldsymbol{i},\boldsymbol{j},p)$ be the minimum weight of any subset $S \subseteq \{1,\dots,p\}$ such that for all $k \in [K]$,

$$\sum_{s\in S} \tilde{r}_{s,k} \geq i_k \quad \text{and} \quad \sum_{s\in S_+} \tilde{u}_{s,k} \leq j_k.$$

We can compute $F(\boldsymbol{i},\boldsymbol{j},p)$ for $(\boldsymbol{i},\boldsymbol{j},p) \in [I]^K \times [J]^K \times [n]$ using the following recursion.

$$\begin{aligned} F(\boldsymbol{i},\boldsymbol{j},1) &= \begin{cases} w_1 & \text{if } \mathbf{0} \leq \boldsymbol{i} \leq \tilde{\boldsymbol{r}}_1 \text{ and } \boldsymbol{j} \geq \tilde{\boldsymbol{u}}_0 + \tilde{\boldsymbol{u}}_1 \\ 0 & \text{if } \boldsymbol{i} \leq \mathbf{0} \text{ and } \boldsymbol{j} \geq \tilde{\boldsymbol{u}}_0 \\ \infty & \text{otherwise} \end{cases} \\ F(\boldsymbol{i},\boldsymbol{j},p+1) &= \min\{F(\boldsymbol{i},\boldsymbol{j},p), w_{p+1} + F(\boldsymbol{i} - \tilde{\boldsymbol{r}}_{p+1}, \boldsymbol{j} - \tilde{\boldsymbol{u}}_{p+1}, p)\} \end{aligned} \tag{5}$$

Let $\boldsymbol{I}$ (resp. $\boldsymbol{J}$) be the vector with all components being I (resp. J). In order to show that (5) correctly finds a subset satisfying (4), we have the following lemma.

LEMMA 1. *For any guess $\boldsymbol{h}$, $\boldsymbol{g}$, if there exists a feasible S such that (4) is satisfied, then $F(\boldsymbol{I},\boldsymbol{J},n) \le W$. Moreover, if $F(\boldsymbol{I},\boldsymbol{J},n) \le W$, then the DP finds a subset $\tilde{S}$ such that for all $k \in [K]$,*

$$\sum_{j\in S} r_{j,k}u_{j,k} \ge h_k(1-2\epsilon) \quad \textit{and} \quad \sum_{j\in S} u_{j,k} \le g_k(1+2\epsilon).$$

Proof Consider S satisfying (4) for given guesses $\boldsymbol{h},\boldsymbol{g}$. Scaling the inequalities yields for all $k\in[K]$

$$\sum_{j\in S}\frac{r_j u_{j,k}}{\epsilon h_k/n} \ge \frac{h_k}{\epsilon h_k/n} \text{ and } \sum_{j\in S}\frac{u_{j,k}}{\epsilon g_k/(n+1)} \le \frac{g_k}{\epsilon g_k/(n+1)}.$$

Rounding down and up the previous inequalities gives for all k

$$\sum_{j\in S}\tilde{r}_{j,k} \ge \lfloor n/\epsilon\rfloor - n = I \text{ and } \sum_{j\in S}\tilde{u}_{j,k} \le \left\lceil \frac{(n+1)}{\epsilon}\right\rceil + (n+1) = J,$$

which implies that $F(\boldsymbol{I},\boldsymbol{J},n)\le W$.

Conversely, suppose $F(\boldsymbol{I},\boldsymbol{J},n)\le W$ and let $\tilde{S}$ be the corresponding subset. We have

$$\sum_{j\in\tilde{S}} r_j u_{j,k} \ge I\frac{\epsilon h_k}{n} \ge h_k(1-2\epsilon) \quad \text{and} \quad \sum_{j\in\tilde{S}}\tilde{u}_{j,k} \le J\frac{\epsilon g_k}{n+1} \le g_k(1+2\epsilon).$$

□

We can now present the FPTAS for Cap-MMNL-Assort.

Algorithm 1 FPTAS for Cap-MMNL-Assort

```
procedure FPTAS(ε)
    for (h, g) ∈ Γ_{ε,K} × Δ_{ε,K} do
        Compute discretization of coefficient r̃_{i,k} and ũ_{i,k} using (3)
        Compute F(i, j, p) for all (i, j, p) ∈ [I]^K × [J]^K × [n] using (5)
        If F(I, J, n) ≤ W, then let S̃_{h,g} be a the corresponding subset
    end for
    return S that maximizes the expected revenue over {S̃_{h,g}, (h, g) ∈ Γ_{ε,K} × Δ_{ε,K}}
end procedure
```

THEOREM 3. *Algorithm 1 returns an $(1-\epsilon)$-optimal solution to* Cap-MMNL-Assort*. Moreover, the running time is $O\left(\log(nRU)^K \log(nU)^K n^{2K+1}/\epsilon^{4K}\right)$.*

Proof. Let S^* be the optimal solution to Cap-MMNL-Assort and $(\hat{\boldsymbol{\ell}}_1,\hat{\boldsymbol{\ell}}_2)$ such that for all $k\in[K]$

$$ru\,(1+\epsilon)^{\hat{\ell}_{1,k}} \le \sum_{i\in S^*} r_i u_{i,k} \le ru\,(1+\epsilon)^{\hat{\ell}_{1,k}+1} \text{ and } u\,(1+\epsilon)^{\hat{\ell}_{2,k}} \le \sum_{i\in S^*_+} u_{i,k} \le u\,(1+\epsilon)^{\hat{\ell}_{2,k}+1}.$$

From Lemma 1, we know that for $(\boldsymbol{h},\boldsymbol{g}) = (ru\,(1+\epsilon)^{\hat{\boldsymbol{\ell}}_1}, u\,(1+\epsilon)^{\hat{\boldsymbol{\ell}}_2})$, Algorithm 1 returns $\tilde{S}$ such that for all $k \in [K]$

$$\sum_{i\in\tilde{S}} r_i u_{i,k} \geq ru\,(1+\epsilon)^{\hat{\ell}_{1,k}}\,(1-2\epsilon) \;\text{ and }\; \sum_{i\in\tilde{S}_+} u_{i,k} \leq u\,(1+\epsilon)^{\hat{\ell}_{2,k}}\,(1+2\epsilon).$$

Consequently,

$$f(\tilde{S}) = \sum_{k=1}^{K} \theta_k \frac{\sum_{i\in\tilde{S}} r_i u_{i,k}}{\sum_{i\in\tilde{S}_+} u_{i,k}} \geq \frac{1-2\epsilon}{1+2\epsilon} f(S^*) \geq (1-4\epsilon) f(S^*).$$

Running Time. We try $L_1^K \cdot L_2^K$ guesses for the numerators and denominators values, $(\boldsymbol{h},\boldsymbol{g})$, of the optimal solution. For each guess, we formulate a dynamic program with $O\,(n^{2K+1}/\epsilon^{2K})$ states. Therefore, the running time of Algorithm 1 is $O\,(L_1^K L_2^K n^3/\epsilon^2) = O\,(\log(nRU)\log(nU) n^{2K+1}/\epsilon^{4K})$ which is polynomial in input size and $1/\epsilon$. □

Algorithm 1 is therefore an FPTAS for Cap-MMNL-Assort when the number of mixtures K is constant and we get the following result.

COROLLARY 1. *There is a fully polynomial time approximation scheme (FPTAS) for* Cap-MMNL-Assort *when the number of mixtures, K, is constant.*

Note that since MNL is a special case of MMNL, this algorithm gives an FPTAS for Cap-MNL-Assort.

We now consider the capacitated assortment optimization problem, Cap-NL-Assort for the Nested logit model. In a Nested logit model, the set of items is partitioned into nests (or subsets) and the choice probability for any item j is decomposed in the probability of selecting the nest containing j and the probability of selecting j in that nest. Suppose there are K nests $N_1,\ldots,N_K$ and each nest N_k contains n items with revenue $r_{i,k}$ and utility parameter $u_{i,k}$. Each nest N_k has a *dissimilarity parameter*, $\gamma_k \in [0,1]$ that models the influence of nest k over others. For a set of assortments $(S_1,\ldots,S_K)$, the probability that nest k is selected is given by

$$Q_k(S_1,\ldots,S_K) = \frac{U_k(S_k)^{\gamma_k}}{U_0 + \sum_{j=1}^{K} U_j(S_j)^{\gamma_j}}.$$

Let R_k denote the expected revenue of nest k conditional on nest k being selected. Then

$$R_k(S_k) = \sum_{i\in S_k} r_{i,k} \frac{u_{i,k}}{u_{0,k} + \sum_{j\in S_k} u_{j,k}} = \frac{\sum_{i\in S_k} r_{i,k} u_{i,k}}{U_k(S_k)}.$$

Additionally, each item is assigned a weight $w_{i,k}$. Let W denote the total available capacity and W_k be the available capacity for nest k for $k \in [K]$. We introduce the following general capacitated assortment optimization for the NL model, Cap-NL-Assort

$$\begin{aligned} \max_{(S_1,\ldots,S_K)\subseteq[n]^K} \; & \sum_{k=1}^{K} Q_k(S_1,\ldots,S_K)R_k(S_k) \\ & \sum_{i\in S_k} w_{i,k} \leq W_k, \; \forall k \in [K] \\ & \sum_{k\in[K]} \sum_{i\in S_k} w_{i,k} \leq W. \end{aligned} \tag{6}$$

Note that this general problem contains both capacity constraints for each nest as well as a capacity constraint across the nests. The objective function is a weighted sum of ratios and we can adapt Algorithm 1 to solve this problem as well.

Corollary 2. *There is a fully polynomial time approximation scheme (FPTAS) for* Cap-NL-Assort *when the number of nests, K, is constant.*

The adapted algorithm is described in Appendix B. Note that Mittal and Schulz (2013) give an FPTAS when there is only a single capacity constraint across nests. Also, Gallego and Topalogulu (2012) give a 2-approximation for Cap-NL-Assort.

For both Cap-MMNL-Assort and Cap-NL-Assort, we obtain an FPTAS when K (the number of mixtures in MMNL or the number of nests in NL) is constant. The running time is however exponential in K. We show in the next section that a polynomial dependence in K is not possible.

4. Hardness of approximation for Assortment Optimization for MMNL model

In this section, we show that MMNL-Assort is hard to approximate within any reasonable factor when the number of MNL segments, K is not constant. In particular, we show that there is no polynomial time algorithm (polynomial in n, K and the input size) with an approximation factor better than $O(1/K^{1-\delta})$ for any constant $\delta > 0$ for the MMNL assortment optimization problem unless $NP \subseteq BPP$. Aouad et al. (2014) show that the assortment optimization problem is hard to approximate within a factor of $O(1/K^{1-\delta})$ for any $\delta > 0$ when the choice model is given by a distribution over K permutations by a approximation preserving reduction from the independent set problem. We adapt the reduction in Aouad et al. (2014) to show a hardness of approximation for the assortment optimization under MMNL choice problem.

Theorem 4. *There is no polynomial time algorithm (polynomial in n, K and the input size) that approximates the unconstrained assortment optimization problem for the MMNL model within a factor $O(1/K^{1-\delta})$ for any constant $\delta > 0$ unless $NP \subseteq BPP$.*

Proof. We prove this by a reduction from the independent set problem. In a maximum independent set problem, we are given an undirected graph $G = (V, E)$ where $V = \{v_1, \ldots, v_n\}$. The goal is to find a maximum cardinality subset of vertices that are independent.

We construct an instance of MMNL-Assort as follows. We have one product and one MNL segment corresponding to each vertex in G. Therefore, $n = K = |V|$ in the MMNL model. For any MNL segment k corresponding to $v_k \in V$, we only consider a subset of products corresponding to a subset of neighbors of v_k in G. In particular, we consider the following utility parameters.

$$\begin{aligned} u_{j,k} &= \begin{cases} 1 & \text{if } j = k \text{ or } j = 0 \\ n^2 & \text{if } (v_j, v_k) \in E \text{ and } j < k \\ 0 & \text{otherwise} \end{cases} \\ r_i &= n^{3(i-1)},\ i \in [n] \\ \theta_k &= \frac{\theta}{n^{3(k-1)}},\ k \in [n] \end{aligned} \tag{7}$$

where $\theta \in [1/2, 1]$ is an appropriate normalizing constant. Note that the utility of any product $j \in [n]$ for segment $k \in [n]$, $u_{j,k} > 0$ only if $(v_j, v_k) \in E$ and $j < k$.

We first show that if there is an independent set, $\mathcal{I} \subseteq V$ where $|\mathcal{I}| = t$, we can find an assortment with revenue $\theta t/2$. Consider the set of products, S corresponding to vertices in $\mathcal{I}$, i.e.,

$$S = \{j \mid v_j \in \mathcal{I}\}.$$

Then, it is easy to observe that the revenue of S is exactly $\theta \cdot t/2$.

Next, we show that if there is an assortment S with expected revenue $R(S)$, then there exists an independent set of size at least $\lfloor 2 \cdot R(S)/\theta \rfloor$. For any segment $k \in [K]$, let R_k denote the contribution of segment k to the expected revenue of assortment S, i.e.,

$$R_k = \theta_k \cdot \frac{\sum_{j \in S} r_j u_{j,k}}{u_{0,k} + \sum_{j \in S} u_{j,k}}, \text{ and } R(S) = \sum_{k=1}^{K} R_k.$$

We show $R_k \geq \theta/2$ or $R_k \leq (2\theta)/n^2$. Let

$$N(k) = \{j \mid (v_j, v_k) \in E,\ j < k\}.$$

Case 1 ($N(k) = \emptyset$): If $k \notin S$, then $R_k = 0$. On the other hand, if $k \in S$, then

$$R_k = \theta_k \cdot \frac{r_k u_{k,k}}{u_{0,k} + u_{k,k}} = \frac{\theta}{n^{3(k-1)}} \cdot \frac{n^{3(k-1)}}{2} = \frac{\theta}{2}. \tag{8}$$

Case 2 ($N(k) \neq \emptyset$): In this case, $|N(k)| \geq 1$. Therefore,

$$R_k = \frac{\theta}{n^{3(k-1)}} \cdot \frac{n^{3(k-1)} + n^2 \cdot \sum_{j \in N(k)} n^{3(j-1)}}{2 + |N(k)| \cdot n^2} \leq \frac{2 \cdot \theta}{n^2}.$$

Therefore,

$$\left(|\{k \in S \mid N(k) = \emptyset\}| \cdot \frac{\theta}{2}\right) \leq R(S) \leq \left(|\{k \in S \mid N(k) = \emptyset\}| \cdot \frac{\theta}{2}\right) + \frac{2 \cdot \theta}{n}. \tag{9}$$

We can now construct an independent set, $\mathcal{I}$ as follows:

$$\mathcal{I} = \{v_k \in V \mid k \in S, N(k) = \emptyset\}.$$

We claim that $\mathcal{I}$ is an independent set. For the sake of contradiction, suppose there exist $v_i, v_j \in \mathcal{I}$ $(i < j)$ such that $(v_i, v_j) \in E$. Since $v_i, v_j \in \mathcal{I}$, $i, j \in S$ and $N(i) = N(j) = \emptyset$. Moreover, since $i < j$ and $(v_i, v_j) \in E$, $i \in N(j)$ which implies $N(j) \neq \emptyset$; a contradiction. Therefore, $\mathcal{I}$ is an independent set. Also,

$$|\mathcal{I}| = |\{k \in S \mid N(k) = \emptyset\}| = \left\lfloor \frac{2 \cdot R(S)}{\theta} \right\rfloor,$$

where the second equality follows from (9). Therefore, if $\mathcal{I}^*$ is the optimal independent set and R^* is the optimal expected revenue of the corresponding MMNL-Assort instance (7), then

$$\left\lfloor \frac{2 \cdot R^*}{\theta} \right\rfloor \leq |\mathcal{I}^*| \leq \frac{2 \cdot R^*}{\theta}.$$

Consequently, an α-approximation for MMNL-Assort implies an $O(\alpha)$-approximation for the maximum independent set problem. Since the maximum independent set is hard to approximation within a factor better than $O(1/n^{1-\delta})$ (where $|V| = n = K$) for any constant $\delta > 0$ (see Feige et al. (1996)), the above reduction implies the same hardness of approximation for MMNL-Assort. □

The above theorem shows that the unconstrained assortment optimization for MMNL is hard to approximate. Consequently, the capacity constrained version of the problem is also hard. The approximation preserving reduction from the independent set problem gives several interesting insights. First, note that each MNL segment in the reduction only contains a subset of products corresponding to a subset of vertices in the neighborhood of the corresponding vertex. This is quite analogous to the consideration set model considered in Jagabathula and Rusmevichientong (2014) where a local neighborhood defines the consideration set. Such graphical model based consideration set instances are quite natural and our reduction shows that Cap-MMNL-Assort is hard even for these naturally occurring instances. Therefore, our reduction gives a procedure to construct naturally arising hard benchmark instances of Cap-MMNL-Assort that may be of independent interest.

We can extend the hardness of approximation even for the continuous relaxation of MMNL-Assort.

THEOREM 5. *Consider the following continuous relaxation of the* MMNL-Assort *problem.*

$$\max_{\boldsymbol{x} \in [0,1]^n} \left\{ \sum_{k=1}^{K} \theta_k \frac{\sum_{j=1}^{n} r_j u_{j,k} x_j}{u_{0,k} + \sum_{j=1}^{n} u_{j,k} x_j} \right\} \tag{10}$$

There is no approximation algorithm (with running time polynomial in K) that has an approximation factor better than $O(1/K^{1-\delta})$ for any constant $\delta > 0$ unless $NP \subseteq BPP$.

We present the proof in Appendix C.

5. FPTAS for Models with Arbitrary Number of Mixtures or Nests

In this section, we consider special cases of choice models where we can give near-optimal algorithms with a polynomial dependence on the number of mixtures or nests.

5.1. Special Case of Nested logit

We consider the special case of Cap-NL-Assort considered in Gallego and Topalogulu (2012). More precisely, we assume that the dissimilarity parameter for each nest, $\gamma \leq 1$ and the utility of no-purchase in each nest k, $u_{0,k} = 0$ for all $k \in [K]$. Moreover, we assume that there is a separate capacity constraint on products in each nest but not across nests. We can formulate the capacitated assortment optimization problem over this special case of MMNL model as follows.

$$\begin{aligned} \max_{(S_1,\ldots,S_K)\subseteq[n]^K} \; & \sum_{k=1}^{K} Q_k(S_1,\ldots,S_K)R_k(S_k) \\ & \sum_{i\in S_k} w_{i,k} \leq W_k, \; \forall k \in [K]. \end{aligned} \tag{11}$$

Note that Gallego and Topalogulu (2012) give a 2-approximation for this problem. They also give an equivalent LP formulation of the problem that we build on to develop our FPTAS.

THEOREM 6 **(Gallego and Topalogulu (2012))**. Cap-NL-Assort *is equivalent to the following linear program*

$$\begin{aligned} \min_{z,\boldsymbol{y}} \quad & z \\ & U_0 x \geq \sum_{k=1}^{K} y_k, \\ & y_k \geq U_k(S_k)^{\gamma_k}(R_k(S_k) - z), \forall S_k \in F_k, \forall k \in [K]. \end{aligned} \tag{12}$$

where for all $k \in [K]$,

$$F_k = \left\{ S_k \subseteq N_k \,\middle|\, \sum_{i\in S_k} w_{i,k} \leq W_k \right\}.$$

Note that this LP has exponentially many constraints. However, we show that we can approximately solve the separation problem in polynomial time.

Separation Problem: For a given $(z, y_1, \ldots, y_K)$, the separation problem is the following. For each nest $k \in [K]$, decide whether

$$y_k \geq \max_{S_k\subseteq[n]} \left\{ U_k(S_k)^{\gamma_k}(R_k(S_k) - z) \,\middle|\, \sum_{i\in S_k} w_{i,k} \leq W_k \right\}, \tag{13}$$

or give a separating hyperplane given by the maximizer $S_k \subseteq [n]$. Therefore, to solve the separation, we need to solve the above maximization problem for each nest. For notational convenience, we ignore the subscript k and focus on the maximization problem for one particular nest. Since the utility of no-purchase, $u_0 = 0$, we can rewrite the maximization problem as follows.

$$\max_{x\in\{0,1\}^n} \left\{ \left(\sum_{i=1}^n u_i x_i\right)^{\gamma-1} \left(\sum_{j=1}^n (r_j u_j - z) x_j\right) \middle| \sum_{i=1}^n w_i x_i \le W \right\}.$$

Notice that since $\gamma \le 1$, adding an item such that $(r_j u_j - z) < 0$ will only decrease the objective function. Therefore, we can equivalently consider the following problem:

$$\max_{x\in\{0,1\}^n} \left\{ \left(\sum_{i=1}^n u_i x_i\right)^{\gamma-1} \left(\sum_{j=1}^n (r_j u_j - z)_+ x_j\right) \middle| \sum_{i=1}^n w_i x_i \le W \right\}. \tag{14}$$

The above problem is quite similar to the Cap-MNL-Assort problem since we need to guess only two quantities:

$$\left(\sum_{i=1}^n u_i x_i\right), \text{ and } \left(\sum_{j=1}^n (r_j u_j - z)_+ x_j\right).$$

Therefore, an appropriate variant of Algorithm 1 returns a $(1-\epsilon)$-approximation for (14). Since we can solve the (14) for each nest independently, we can approximately solve the separation problem (13). The running time is $O\left(K \log(nRU)\log(nU)n^3/\epsilon^4\right)$ which is polynomial in K. Furthermore, we show that a $(1-\epsilon)$-approximation for the separation problem (13) implies a $(1-\epsilon)$-approximation for the assortment optimization problem (12). In particular, we have the following theorem.

THEOREM 7. *Suppose there is a $(1-\epsilon)$-approximation for the separation problem* (13) *with running time polynomial in the input size and $1/\epsilon$. Then we can get a $(1-\epsilon)$-approximation for the optimization problem* (12) *in running time polynomial in the input size and $1/\epsilon$.*

We present the proof in Appendix D. Therefore, we have the following corollary.

COROLLARY 3. *There is a FPTAS for* Cap-NL-Assort *with running time polynomial in n, K and the input size, if the dissimilarity parameter $\gamma \le 1$ and the utility of no-purchase in each nest k, $u_{0,k} = 0$ for all $k \in [K]$.*

5.2. Special Case of MMNL

We consider the special case of MMNL considered in Rusmevichientong et al. (2014) where the various MNL segments differ only in the utility of the no-purchase option. In particular, the utility of each item j is equal to u_j for all segments $k = 1, \ldots, K$. Also, let v_k be the utility of the no-purchase option for segment $k \in [K]$. We can formulate the capacitated assortment optimization problem over this special case of MMNL model as follows.

$$\max_{\boldsymbol{x}\in\{0,1\}^n} \left\{ \sum_{k=1}^K \theta_k \frac{\sum_{j=1}^n r_j u_j x_j}{v_k + \sum_{j=1}^n u_j x_j} \middle| \sum_{j=1}^n w_j x_j \le W \right\}.$$

Note that in this case, since the utilities are the same across mixtures, we do not need to guess each numerator and denominator. Instead, we only need to guess

$$\left(\sum_{i\in S^*} r_j u_j\right), \text{ and } \left(\sum_{i\in S^*} u_j\right).$$

Therefore, the number of guesses does not depend on the number of mixtures. As in the special case of Cap-NL-Assort, this special case of Cap-MMNL-Assort is quite similar to Cap-MNL-Assort where we need only two guesses. Note that the running time of our FPTAS for the general MMNL model depends exponentially on the number of guesses (which is equal to twice the number of MNL segments). Therefore, we have the following result.

Theorem 8. *There is a FPTAS for the capacitated assortment optimization over the MMNL model with an arbitrary number of mixtures if the various MNL segments differ only in the utility of the no-purchase option.*

6. Conclusion

In this paper, we study the capacity constrained version of the assortment optimization under different choice models including MMNL and NL choice models and give an FPTAS for these problems when the number of mixtures or nests, K is constant. Our FPTAS has an exponential dependence on K. Therefore, we require the number of mixtures or nests to be constant. However, we show that such a dependence is necessary for any near-optimal algorithm for the capacitated assortment optimization problem. In particular, we show that there is no algorithm for the unconstrained MMNL-Assort with a polynomial dependance on K with an approximation factor better than $O(1/K^{1-\delta})$ for any $\delta > 0$. Therefore, there is no reasonable approximation algorithm for Cap-MMNL-Assort in general whose running time is polynomial in K. Furthermore, the proof of this hardness provides a procedure to construct a natural family of hard benchmark instances for the assortment optimization problem over MMNL that may be of independent interest. Finally, we present special cases of Nested logit model and the MMNL model, where we get an FPTAS for the capacitated assortment optimization problem for an arbitrary number of nests or mixtures.

Acknowledgments

The authors gratefully acknowledge Danny Segev for valuable discussions.

References

Aouad, Ali, Vivek Farias, Retsef Levi, Danny Segev. 2014. personal communication .

Ben-Akiva, M., S. Lerman. 1985. *Discrete choice analysis: theory and application to travel demand*, vol. 9. MIT press.

Davis, James, Guillermo Gallego, Huseyin Topaloglu. 2013. Assortment planning under the multinomial logit model with totally unimodular constraint structures. *Technical Report* .

Davis, J.M., G. Gallego, H. Topaloglu. 2011. Assortment optimization under variants of the nested logit model. Tech. rep., Technical report, Cornell University.

Feige, Uriel, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, Mario Szegedy. 1996. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM (JACM)* **43**(2) 268–292.

Frieze, Alan M, MRB Clarke. 1984. Approximation algorithms for the m-dimensional 0–1 knapsack problem: Worst-case and probabilistic analyses. *European Journal of Operational Research* **15**(1) 100–109.

Gallego, G., H. Topalogulu. 2012. Constrained assortment optimization for the nested logit model. Tech. rep., Working Paper, Cornell University.

Grötschel, Martin, László Lovász, Alexander Schrijver. 1981. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica* **1**(2) 169–197.

Jagabathula, Srikanth, Paat Rusmevichientong. 2014. A two-stage model of consideration set and choice: Learning, revenue prediction, and applications. *Submitted* .

Lawler, Eugene L. 1979. Fast approximation algorithms for knapsack problems. *Mathematics of Operations Research* **4**(4) 339–356.

Luce, R.D. 1959. *Individual choice behavior: A theoretical analysis*. Wiley.

McFadden, D. 1978. Modeling the choice of residential location. *Transportation Research Record* .

McFadden, D. 1980. Econometric models for probabilistic choice among products. *Journal of Business* 13–29.

McFadden, D., K. Train. 2000. Mixed mnl models for discrete response. *Journal of applied Econometrics* **15**(5) 447–470.

Mittal, Shashi, Andreas S Schulz. 2013. A general framework for designing approximation schemes for combinatorial optimization problems with many objectives combined into one. *Operations Research* **61**(2) 386–397.

Papadimitriou, Christos H, Mihalis Yannakakis. 2000. On the approximability of trade-offs and optimal access of web sources. *In Proceedings of 41st Annual IEEE Symposium on Foundations of Computer Science*. 86–92.

Plackett, R.L. 1975. The analysis of permutations. *Applied Statistics* 193–202.

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

Rusmevichientong, Paat, Zuo-Jun Max Shen, David B Shmoys. 2010. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations research* **58**(6) 1666–1680.

Rusmevichientong, Paat, David Shmoys, Chaoxu Tong, Huseyin Topaloglu. 2014. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management* .

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

Wierenga, B. 2008. Handbook of marketing decision models, vol. 121 of international series in operations research & management science.

Appendix A: Assumption of $u_{i,k} > 0$ in Cap-MMNL-Assort

We show that wlog. we can assume $u_{i,k} > 0$ for all $i \in [n], k \in [K]$ in the Cap-MMNL-Assort problem. Let $u = \min \{u_{i,k} \mid u_{i,k} > 0\}$. Suppose $u_{j,k} = 0$ for some j, k. Then, consider the following modified utility parameters for all j, k.

$$\hat{u}_{j,k} = \begin{cases} \epsilon ur/(nR) & \text{if } u_{j,k} = 0 \\ u_{j,k} & \text{otherwise} \end{cases}$$

We show that replacing $u_{j,k}$ by $\hat{u}_{j,k}$ in Cap-MMNL-Assort changes the expected revenue of any subset by a factor of $[1-\epsilon, 1+\epsilon]$. In particular, for any $x \in \{0,1\}^n$, for all $k \in [K]$,

$$\sum_{j=1}^{n} r_j u_{j,k} x_j \le \left(\sum_{j=1}^{n} r_j \hat{u}_{j,k} x_j \right) \le \sum_{j=1}^{n} r_j u_{j,k} x_j + \frac{r_{j,k}}{R} \cdot \epsilon r u \le (1+\epsilon) \cdot \sum_{j=1}^{n} r_j u_{j,k} x_j .$$

Similarly for all $k \in [K]$,

$$u_{0,k} + \sum_{j=1}^{n} u_{j,k} x_j \le \left(\hat{u}_{0,k} + \sum_{j=1}^{n} \hat{u}_{j,k} x_j \right) \le (1+\epsilon) \cdot \left(u_{0,k} + \sum_{j=1}^{n} u_{j,k} x_j \right) .$$

Therefore, for each rational terms in the expression for the expected revenue, both the numerator and denominator increase by a factor of at most $(1+\epsilon)$. Let z^* be the optimal value of Cap-MMNL-Assort and $\hat{z}$ be the optimal value of the modified problem with parameters, $\hat{u}_{j,k}$. Using the previous set of inequalities, we have $(1-\epsilon)\hat{z} \le z^* \le (1+\epsilon)\hat{z}$ and we can equivalently approximate the modified problem.

Appendix B: Proof of Corollary 2

Let r (resp. R) be the minimum (resp. maximum) revenue, u (resp. U) be the minimum (resp. maximum) utility parameter and γ (resp. Γ) be the minimum (resp. maximum) dissimilarity parameter. As earlier, we can assume wlog. that $u > 0$.

For a given guess $(\boldsymbol{h}, \boldsymbol{g}) \in \Gamma_{\epsilon,K} \times \Delta_{\epsilon,K}$, we use a similar scaling as for Algorithm 1. More precisely, the discretized coefficients are

$$\tilde{r}_{i,k} = \left\lfloor \frac{r_{i,k} u_{i,k}}{\epsilon h_k / n} \right\rfloor \quad \text{and} \quad \tilde{u}_{i,k} = \left\lceil \frac{u_{i,k}}{\epsilon g_k/(n+1)} \right\rceil . \tag{15}$$

We need to solve K different dynamic programs for each guess $(\boldsymbol{h},\boldsymbol{g})$ to find feasible assortments $S_1,\ldots,S_K$ such that for all $k\in[K]$,

$$\sum_{j\in S_k} r_{j,k}u_{j,k}\geq h_k \quad \text{and} \quad \sum_{j\in S_k} u_{j,k}\leq g_k. \tag{16}$$

For $k=1,\ldots,K$, for each $(i,j,p)\in[I]\times[J]\times[n]$, let $F_k(i,j,k)$ be the minimum weight of any subset $S_k\subseteq\{1,\ldots,p\}$ such that

$$\sum_{s\in S_k}\tilde{r}_{s,k}\geq i \ \text{ and } \sum_{s\in S_{k+}}\tilde{u}_{s,k}\leq j.$$

To compute $F_k(i,j,p)$ for $(i,j,p)\in[I]\times[J]\times[n]$, we use the following recursion

$$\begin{aligned} F_k(i,j,1) &= \begin{cases} w_{1,k} & \text{if } 0\leq i\leq \tilde{r}_{1,k} \text{ and } j\geq \tilde{u}_{0,k}+\tilde{u}_{1,k} \\ 0 & \text{if } i\leq 0 \text{ and } j\geq \tilde{u}_{0,k} \\ \infty & \text{otherwise} \end{cases} \\ F_k(i,j,p+1) &= \min\{F_k(i,j,p), w_{p+1,k}+F_k(i-\tilde{r}_{p+1},j-\tilde{u}_{p+1},p)\} \end{aligned} \tag{17}$$

We can now present the complete algorithm. To prove the approximation bound, let $S^*=$

Algorithm 2 FPTAS for Cap-NL-Assort

procedure FPTAS(ϵ)
  for $(\boldsymbol{h},\boldsymbol{g})\in\Gamma_{\epsilon,K}\times\Delta_{\epsilon,K}$ **do**
    Compute discretization of coefficient $\tilde{r}_{i,k}$ and $\tilde{u}_{i,k}$ using (15)
    Compute $F_k(i,j,p)$ for all $(i,j,p)\in[I]\times[J]\times[n]$ and all $k\in[K]$ using (5)
    If $F_k(I,J,n)\leq W_k$ for all $k=1,\ldots,K$ and $\sum_{k=1}^{n}F_k(I,J,n)\leq W$, let $\tilde{S}_{\boldsymbol{h},\boldsymbol{g}}=(S_1,\ldots,S_K)$ be the corresponding feasible subsets
  end for
  return S that maximizes the expected revenue over $\{\tilde{S}_{\boldsymbol{h},\boldsymbol{g}},(\boldsymbol{h},\boldsymbol{g})\in\Gamma_{\epsilon,K}\times\Delta_{\epsilon,K}\}$
end procedure

$(S_1^*,\ldots,S_K^*)$ be the optimal solution of Cap-NL-Assort. Let $\hat{\boldsymbol{\ell}}_{\mathbf{1}}$ and $\hat{\boldsymbol{\ell}}_{\mathbf{2}}$ such that for all $k=1,\ldots,K$

$$ru\left(1+\epsilon\right)^{\hat{\ell}_{1,k}}\leq\sum_{i\in S_k^*}r_{i,k}\leq ru\left(1+\epsilon\right)^{\hat{\ell}_{1,k}+1}$$
$$u\left(1+\epsilon\right)^{\hat{\ell}_{2,k}}\leq\sum_{i\in S_{k+}^*}u_{i,k}\leq u\left(1+\epsilon\right)^{\hat{\ell}_{2,k}+1}.$$

Using Lemma 1, we know that Algorithm 2 returns sets $\tilde{S}=(\tilde{S}_1,\ldots,\tilde{S}_K)$ such that for all $k=1,\ldots,K$,

$$\sum_{i\in S_k}r_{i,k}u_{i,k}\geq ru\left(1+\epsilon\right)^{\hat{\ell}_{1,k}}(1-2\epsilon) \ \text{ and } \sum_{i\in S_{k+}}u_{i,k}\leq u\left(1+\epsilon\right)^{\hat{\ell}_{2,k}}(1+2\epsilon).$$

Therefore, for all $k = 1, \ldots, K$

$$U_k(S_k)^{\gamma_k} \leq U_k(S_k^*)^{\gamma_k}(1+\epsilon)^{\gamma_k} \leq U_k(S_k^*)^{\gamma_k}(1+\epsilon)^{\Gamma}$$
$$U_k(S_k)^{\gamma_k} \geq \frac{U_k(S_k^*)^{\gamma_k}}{(1+\epsilon)^{\gamma_k}} \geq \frac{U_k(S_k^*)^{\gamma_k}}{(1+\epsilon)^{\gamma}}.$$

Consequently,

$$\begin{aligned}\sum_{k=1}^{K} Q_k(S_1, \ldots, S_K) R(S_k) &\geq \frac{(1-4\epsilon)}{(1+\epsilon)^{\Gamma+\gamma}} f(\mathsf{OPT}) \\ &\geq (1-4\epsilon)(1-(\Gamma+\gamma)\epsilon) f(\mathsf{OPT}) \\ &\geq (1-(4+\Gamma+\gamma)\epsilon) f(\mathsf{OPT}).\end{aligned}$$

which concludes the proof.

Running Time. Using a similar analysis as in the proof of Theorem 3, the running time is $O\left(\log(nRU)^K \log(nU)^K n^{2K+1}/\epsilon^{4K}\right)$.

Appendix C: Proof of Theorem 5

As in Theorem 4, we prove this by a reduction from the independent set problem where we are given an undirected graph $G = (V, E)$ and the goal is to find a maximum cardinality subset of vertices that are independent. Let $V = \{v_1, \ldots, v_n\}$.

We construct an instance of MMNL-Assort similar to the proof of Theorem 4. We have one product and one MNL segment corresponding to each vertex in G. Therefore, $n = K = |V|$ and we consider the following utility parameters:

$$\begin{aligned} u_{j,k} &= \begin{cases} 1 & \text{if } j = k \text{ or } j = 0 \\ n^3 & \text{if } (v_j, v_k) \in E \text{ and } j < k \\ 0 & \text{otherwise} \end{cases} \\ r_i &= n^{3(i-1)},\ i \in [n] \\ \theta_k &= \frac{\theta}{n^{3(k-1)}},\ k \in [n] \end{aligned} \tag{18}$$

where $\theta \in [1/2, 1]$ is an appropriate normalizing constant.

Consider an optimal independent set, $\mathcal{I}^*$ of size t^*. Consider the following assortment

$$S = \{j \mid v_j \in \mathcal{I}^*\}.$$

It is easy to observe that the expected revenue of S is exactly $\theta t^*/2$.

Conversely, consider an optimal fractional assortment $\boldsymbol{x}^* \in [0,1]^n$ with revenue z^*. Then we show that there exists an independent set of size $\lfloor 2z^*/\theta \rfloor$. Let $\epsilon = 1/4n$. Consider a modified solution $\tilde{\boldsymbol{x}}$ defined as follows. For all $k \in [K]$,

$$\tilde{x}_k = \begin{cases} 0 & \text{if } x_k^* \leq \epsilon \\ x_k^* & \text{otherwise.} \end{cases}$$

Also, let $\tilde{z}$ be the revenue associated with solution $\tilde{\boldsymbol{x}}$. It is easy to observe that the revenue of each nest only decreases by at most $\theta\epsilon$. Consequently,

$$\tilde{z} \geq z^* - n\theta\epsilon \geq z^* - \frac{\theta}{4} \geq \frac{z^*}{2},$$

where the last inequality follows as $z^* \geq \theta/2$. For any $k \in [K]$, let

$$\tilde{z}_k = \theta_k \cdot \frac{\sum_{j=1}^n r_j u_{j,k} \tilde{x}_k}{u_{0,k} + \sum_{j=1}^n u_{j,k} \tilde{x}_j}, \text{ and } \tilde{z} = \sum_{k=1}^K \tilde{z}_k.$$

We show that for any $k \in [K]$, $\tilde{z}_k \geq \theta/(5n)$ or $\tilde{z}_k \leq \theta/n^2$. Let

$$N(k) = \{j \mid j < k, \, (v_j, v_k) \in E, \, \tilde{x}_j \geq \epsilon\}.$$

Case 1 $(N(k) = \emptyset)$: In this case

$$\tilde{z}_k = \frac{\theta \tilde{x}_k}{1 + \tilde{x}_k} \leq \frac{\theta}{2}.$$

Therefore, if $\tilde{x}_k < \epsilon$, it implies $\tilde{x}_k = 0$ (by construction) and $\tilde{z}_k = 0$.

Case 2 $(N(k) \neq \emptyset)$: In this case,

$$\begin{aligned}
\tilde{z}_k &\leq \frac{\theta}{n^{3(k-1)}(1 + n^3 \sum_{j \in N(k)} \tilde{x}_j)} \left(n^{3(k-1)} + n^3 \sum_{j \in N(k)} n^{3(j-1)} \tilde{x}_j \right) \\
&\leq \frac{\theta}{n^{3(k-1)}(2 + n^3\epsilon)} \left(n^{3(k-1)} + n^3 \sum_{j=1}^{k-1} n^{3(j-1)} \right) \\
&\leq \frac{2\theta}{n^2},
\end{aligned}$$

where the second inequality follows as $N(k) \neq \emptyset$ and there exists $j \in N(k)$ such that $\tilde{x}_j \geq \epsilon$. Now, we construct an independent set, $\mathcal{I}$ as follows.

$$\mathcal{I} = \{v_k \in V \mid \tilde{x}_k \geq \epsilon, \, N(k) = \emptyset\}.$$

Since for all k such that $v_k \in \mathcal{I}$, $N(k) = \emptyset$, we know that $\mathcal{I}$ is an independent set (using an argument similar to proof of Theorem 4). From the above case analysis, we know

$$\sum_{k: v_k \in \mathcal{I}} \tilde{z}_k \leq \tilde{z} \leq \sum_{k: v_k \in \mathcal{I}} \tilde{z}_k + \frac{2\theta}{n}, \tag{19}$$

where the second inequality follows from the fact that $\tilde{z}_k \leq 2\theta/n^2$ if $v_k \notin \mathcal{I}$. We also know that $\tilde{z} \geq z^*/2 \geq \theta/4$ and $\tilde{z}_k \leq \theta/2$ for all $k : v_k \in \mathcal{I}$. Therefore,

$$z^* \leq 2\tilde{z} \leq 2\left(\sum_{k: v_k \in \mathcal{I}} \frac{\theta}{2} \right) + \frac{4\theta}{n} \leq |\mathcal{I}| \cdot \theta + \frac{8z^*}{n},$$

which implies

$$|\mathcal{I}| \geq \frac{\left(1-\frac{8}{n}\right)}{\theta} \cdot z^* \geq \frac{1}{2\theta} \cdot z^*.$$

Therefore,

$$\frac{1}{2\theta} \cdot z^* \leq t^* \leq \frac{2}{\theta} \cdot z^*.$$

Recall that θ is a constant in $[1/2, 1]$. Therefore, an α-approximation for the continuous relaxation of MMNL-Assort implies an $O(\alpha)$-approximation for the maximum independent set problem. Since the maximum independent set is hard to approximation within a factor better than $O(1/n^{1-\delta})$ (where $|V| = n = K$) for any constant $\delta > 0$ (see Feige et al. (1996)), so must be the continuous relaxation of MMNL-Assort. □

Appendix D: Proof of Theorem 7

Our proof uses the ellipsoid algorithm. Let Π denote a true separation oracle for (12) and let Π^α denote a separation oracle which uses an α-approximation to solve (14) where $0 < \alpha \leq 1$. Moreover, let $\mathsf{OPT}(\Pi)$ (resp. $\mathsf{OPT}(\Pi^\alpha)$) denotes the optimal value returned by the ellipsoid algorithm in polynomial time using Π (resp. Π^α) as a separation oracle (Grötschel et al. (1981)). We prove that

$$\mathsf{OPT}(\Pi^\alpha) \leq \mathsf{OPT}(\Pi) \leq \frac{1}{\alpha} \cdot \mathsf{OPT}(\Pi^\alpha).$$

The first part of the above inequality follows from the fact that using an α-approximation to solve (14) gives a relaxation of the true problem (12).

To prove the second inequality, let $(x^\alpha, \boldsymbol{y}^\alpha)$ be the optimal solution returned by the ellipsoid algorithm with Π^α as a separation oracle. Therefore, for all $j \in [n]$, we have

$$y_j^\alpha \geq \alpha \cdot z_j(x^\alpha)$$

where for all x

$$z_j(x) = \max_{S_k \subseteq [n]} \left\{ U_k(S_k)^{\gamma_k} (R_k(S_k) - x) \,\middle|\, \sum_{i \in S_k} w_{i,k} \leq W_k \right\}.$$

Consider the following solution for (12):

$$\tilde{x} = \frac{x^\alpha}{\alpha}, \text{ and } \tilde{\boldsymbol{y}} = \frac{1}{\alpha} \cdot \boldsymbol{y}^\alpha.$$

We have for all $j \in [n]$,

$$z_j(\tilde{x}) = \max_{S_k \subseteq [n]} \left\{ U_k(S_k)^{\gamma_k} (R_k(S_k) - \tilde{x}) \,\middle|\, \sum_{i \in S_k} w_{i,k} \leq W_k \right\}$$

$$
\begin{aligned}
&= \max_{S_k \subseteq [n]} \left\{ U_k(S_k)^{\gamma_k}(R_k(S_k) - x^\alpha) - U_k(S_k)^{\gamma_k}((1/\alpha) - 1)x^\alpha \,\middle|\, \sum_{i \in S_k} w_{i,k} \le W_k \right\} \\
&\le \max_{S_k \subseteq [n]} \left\{ U_k(S_k)^{\gamma_k}(R_k(S_k) - x^\alpha) \,\middle|\, \sum_{i \in S_k} w_{i,k} \le W_k \right\} \\
&= z_j(x^\alpha).
\end{aligned}
$$

Therefore, for all $j \in [n]$,

$$\tilde{y}_j = \frac{y_j^\alpha}{\alpha} \ge z_j(x^\alpha) \ge z_j(\tilde{x}),$$

which implies that $(\tilde{x}, \tilde{\boldsymbol{y}})$ is a feasible for (12) and $\tilde{x} \ge \mathsf{OPT}(\Pi)$. Therefore,

$$\mathsf{OPT}(\Pi) \le \frac{1}{\alpha} \cdot \mathsf{OPT}(\Pi^\alpha).$$

This concludes the proof. □