History-Dependent Fluid Approximations and Performance Guarantees for Revenue Management with Markov-Modulated Demands

Weiyuan Li[1], Paat Rusmevichientong[2], Huseyin Topaloglu[1], Jingwei Zhang[3]

[1]School of Operations Research and Information Engineering, Cornell Tech, New York, NY 10044, USA
[2]Marshall School of Business, University of Southern California, Los Angeles, CA 90089, USA
[3]School of Data Science, The Chinese University of Hong Kong, Shenzhen 518172, PRC
wl425@cornell.edu, rusmevic@marshall.usc.edu, topaloglu@orie.cornell.edu, zhangjingwei@cuhk.edu.cn

January 21, 2025

A common approach for modeling the demand in revenue management systems is based on dividing the selling horizon into a number of time periods such that there is at most one customer arrival at each time period and the products requested by different customers are independent of each other. This approach rules out the possibility of modeling dependence between the demands from different customers. We study revenue management models that use Markov-modulated demands to incorporate dependence between the demands from different customers. We divide the selling horizon into multiple stages. The modulating Markov chain makes transitions at the beginning of each stage. The probability distribution for the product requested by each arriving customer depends on the state of the Markov chain. We can give dynamic programming formulations for such revenue management models by keeping track of the remaining resource capacities and state of the Markov chain in the state variable, but solving these dynamic programs is intractable. We seek to formulate the corresponding fluid approximation with the goal of obtaining approximate policies with performance guarantees. We give a family of history-dependent fluid approximations, where the probability of accepting a product request depends on the history of the Markov chain over past θ stages. Our fluid approximations yield upper bounds on the optimal total expected revenue and these upper bounds get tighter as we use longer histories. Letting K be the number of stages, $c_{\min}$ be the smallest resource capacity and θ be the number of stages kept in the history of the Markov chain, we use our fluid approximation to give an approximate policy with a performance guarantee of $1-\sqrt{\frac{\min\{M_1\,(2-(\theta/K))\,K, M_2\,c_{\min}\}\,\log c_{\min}}{c_{\min}^2}}$, where M_1 and M_2 are constants independent of K, $c_{\min}$ and θ. Thus, ignoring the logarithmic term, if the number of stages in the selling horizon and capacities of the resources scale proportional to β, then the optimality gap of our approximate policy vanishes with rate $1-\sqrt{\frac{M_1\,(2-(\theta/K))}{\beta}}$. To our knowledge, this policy is the first asymptotically optimal one under Markov-modulated demands. Our computational experiments show that the right fluid approximation under Markov-modulated demands can make a dramatic impact in practice.

1. Introduction

One of the most prevalent approaches for modeling the demand in revenue management systems is based on dividing the selling horizon into a number of time periods such that there is at most one customer arrival at each time period and the products requested by the successive customers are independent of each other. This approach is equivalent to using a discrete-time approximation to a Poisson process and it has been useful for building impactful revenue management models for multiple decades; see, for example, Talluri and van Ryzin (2005), Ozer and Phillips (2012), Gallego and Topaloglu (2019). However, such a discrete-time approximation to a Poisson process has shortcomings. A discrete-time approximation to a Poisson process rules out the possibility of modeling dependence between the demands from different customers. In practice, factors such as shifts in economic forecasts, weather events and changes in customer tastes can systematically alter the product demand, creating dependence between the demands from different customers. More subtly, perhaps, under a discrete-time approximation to a Poisson process, letting $\mathcal{T}$ be the set of time periods in the selling horizon, if there is a demand for a product at time period t with probability μ_t, then the mean and standard deviation of the total demand for the product are, respectively, $\sum_{t\in\mathcal{T}}\mu_t$ and $\sqrt{\sum_{t\in\mathcal{T}}\mu_t\,(1-\mu_t)}$, so the coefficient of variation is at most $1/\sqrt{\sum_{t\in\mathcal{T}}\mu_t}$. Thus, if the total demand for a product has a large mean, then it has to have a small coefficient of variation. One way to address these shortcomings is to use Markov-modulated demands, where a Markov chain makes transitions over the time periods and the probability distribution for the product requested by a customer depends on the state of the Markov chain. Markov-modulated demands naturally allow capturing dependence between the demands from different customers. Also, by using Markov-modulated demands, we can incorporate demand variance levels not possible under a discrete-time approximation to a Poisson process. Under Markov-modulated demands, the standard deviation for the total demand for a product can easily be as large as its mean.

In this paper, we study revenue management models with Markov-modulated demands. We divide the selling horizon into multiple stages. The Markov chain makes transitions at the beginning of each stage. The probability distribution of the product requested by a customer arriving at a time period depends on the current state of the Markov chain. Each stage can include multiple time periods, so that the transitions of the Markov chain may occur at a slower timescale than the arrivals of the customers. Observing the current state of the Markov chain, our goal is to find a policy to decide which customer requests to accept to maximize the total expected revenue over the selling horizon. It is not difficult to give a dynamic programming formulation for a revenue management model that incorporates Markov-modulated demands. The state variable in the dynamic program would keep track of the remaining capacities of the resources and state of the Markov chain, but such a dynamic program is computationally intractable to solve due to its high dimensional state variable. Fluid approximations have been a reliable workhorse to construct efficiently computable policies for revenue management problems. The question we seek to answer is whether we can formulate a fluid approximation under Markov-modulated demands and use the fluid approximation to construct approximate policies with performance guarantees.

Main Contributions: Our main contributions include giving a family of history-dependent fluid approximations and using them to construct policies with performance guarantees.

Family of History-Dependent Fluid Approximations: Under Markov-modulated demands, we give a family of history-dependent fluid approximations. There are two novel features of our fluid approximation. First, the probability of accepting a product request in the fluid approximation depends on the history of the Markov chain over past θ stages. In other words, the decision variables have dependence on the history. For any choice of θ, we show that our fluid approximation provides an upper bound on the optimal total expected revenue. In this way, we obtain a family of fluid approximations parameterized by θ. Also, we show that if the fluid approximation uses a larger number of stages in the history of the Markov chain, then the upper bound from the fluid approximation gets tighter. Therefore, larger values for θ allow approximating the optimal total expected revenue more accurately. Second, the capacity constraints in our fluid approximation keep track of the total expected capacity consumption of each resource up to a particular stage conditional on the state of the Markov chain at the stage. The conditional form of these constraints turns out to be critical to establish a performance guarantee for our approximate policy.

Approximate Policy with a Performance Guarantee: We use our fluid approximation to construct an approximate policy to decide which customer requests to accept. Letting K be the number of stages, $c_{\min}$ be the smallest resource capacity and θ be the number of stages in the history of the Markov chain in the fluid approximation, we show that our approximate policy has a performance guarantee of $1-\sqrt{\frac{\min\{M_1\,(2-(\theta/K))\,K, M_2\,c_{\min}\}\,\log c_{\min}}{c_{\min}^2}}$, where M_1 and M_2 are constants independent of K, $c_{\min}$ and θ. By the first term in the minimum operator, ignoring the logarithmic term, if both the number of stages in the selling horizon and smallest resource capacity get large proportional to β, then the optimality gap of our approximate policy vanishes with rate $1-\sqrt{\frac{M_1(2-(\theta/K))}{\beta}}$. The quantity θ/K is the fraction of the number of stages included in the history. By the second term in the minimum operator, if the smallest resource capacity gets large proportional to β, then the optimality gap of our approximate policy vanishes with rate $1-\sqrt{\frac{M_2}{\beta}}$.

The second scaling regime makes a milder assumption on the quantities that we scale, but the performance guarantee in the first scaling regime explicitly characterizes the dependence of the performance guarantee on the number of stages that we include in the history of the Markov chain. The proof of the performance guarantee uses novel ideas. We show that the random variable corresponding to the capacity consumption of a resource under our approximate policy satisfies the so-called average Lipschitz condition, which allows us to construct an upper bound on the moment generating function of the resource consumptions under our policy; see Dubhashi and Panconesi (2009). We use this bound to upper bound the exhaustion probability of a resource, which, in turn, allows us to compare the total expected revenue of the approximate policy with the optimal objective value of the fluid approximation. Our approximate policy, to our knowledge, is the *first* one with an asymptotic optimality guarantee under Markov-modulated demands.

Fluid Approximation with Varying Histories: In our fluid approximation, when making the decisions in different stages, the number of stages that we use in the history of the Markov chain is always the same. We give an alternative fluid approximation, where we use different numbers of stages in the history of the Markov chain when making the decisions in different stages. We refer to this fluid approximation as the fluid approximation with varying histories. We establish conditions to ensure that the fluid approximation with varying histories provides an upper bound on the optimal total expected revenue. Even though we use varying histories, we show that using longer histories yields tighter upper bounds. We show that we can use the fluid approximation with varying histories to construct an approximate policy that has the same performance guarantee as our original approximate policy. To achieve the same performance guarantee, however, the numbers of decision variables and constraints in the fluid approximation with varying histories are smaller than those in the original fluid approximation by a factor of K. Using varying histories, we can attain the same performance guarantee by solving a more compact fluid approximation.

Reductions and Computational Experiments: We give reductions that allow us to solve the fluid approximation with varying histories more efficiently under certain choices of the histories. Using these reductions as computational tools, we are able to decompose the fluid approximation with varying histories by groups of stages, as well as by the products, while making sure that the subproblems for different groups of stages and products still communicate with each other, ensuring no loss of optimality. Furthermore, we are able to formulate the fluid approximation as a dynamic program with a low dimensional, but continuous, state variable. We give computational experiments to test the tightness of the upper bounds and performance of the approximate policies from our fluid approximation. The practical behavior of our approximate policy is aligned with its performance guarantee in the sense that the total expected revenues that we obtain are larger when our fluid approximation uses a larger number of stages in the history of the Markov chain.

Related Literature: There is work on building fluid approximations for revenue management problems and using these fluid approximations to construct approximate policies with performance guarantees. Gallego and van Ryzin (1994) study fluid approximations for dynamic pricing problems with a single resource and give an approximate policy that is asymptotically optimal as the capacity of the resource and length of the selling horizon increase linearly with the same rate. Gallego and van Ryzin (1997) extend this result to multiple resources. Talluri and van Ryzin (1998) consider the setting where we decide whether to accept the product requests from the customers, rather than choosing the prices for the products. Gallego et al. (2004) and Liu and van Ryzin (2008) consider the case where the customers make a choice among the products offered to them. In the papers discussed in this paragraph, the authors use a discrete-time approximation to a Poisson process and solve the fluid approximation once at the beginning of the selling horizon.

Jasin and Kumar (2012) show that it is possible to attain stronger performance guarantees by solving the fluid approximation periodically over the selling horizon. Their policy is asymptotically optimal as the resource capacities and number of time periods in the selling horizon increase linearly with the same rate. Ma et al. (2021) and Feng et al. (2024) give asymptotically optimal policies with performance guarantees that depend only on the resource capacities, but not on the number of time periods in the selling horizon. Aouad and Ma (2022), Bai et al. (2023) and Li et al. (2024) allow arbitrary distributions for the numbers of customer arrivals, but the products requested by the successive customers are still independent of each other.

Considering problems with dependence between the random quantities at different time periods, Brown and Zhang (2022) study a project portfolio management problem, where the random amount of resource arriving at each time period has to be allocated to the projects. The resource availability at each time period is modulated by a Markov chain. The authors develop an approximate policy by relaxing the resource availability constraints through Lagrange multipliers that depend on the full history of the Markov chain. Their policy is asymptotically optimal as the number of projects gets large. Zhao et al. (2025) use a similar approach for a problem in web advertising to get asymptotically optimal policies as the number of contract types gets large. The asymptotic regimes in these papers are very different from ours, as we focus on the case where the capacities of the resources get large. Most related to us, Jiang (2023) considers a revenue management problem, where the types of the arriving customers are determined by a Markov chain. Letting L be the maximum number of resources consumed by a product, the paper gives a policy with a performance guarantee of $1/(1+L)$. This performance guarantee does not get arbitrarily close to one as the size of the problem, measured by the capacities of the resources, gets large. In his policy, the author uses a recursion to evaluate the opportunity cost of the capacities consumed by each product. Such an approach builds on Ma et al. (2020), but the opportunity costs in the Markov-modulated setting depend on the state of the Markov chain. In the same problem setting, Jin et al. (2024) show that one can use standard concentration inequalities to construct a policy with a performance guarantee of $1-O\Big(\frac{1}{c_{\min}^{1/3}}\Big)$, which is weaker than our performance guarantee of $1-O\Big(\sqrt{\frac{\log c_{\min}}{c_{\min}}}\Big)$. Lastly, Yuan et al. (2024) model demand dependence by using the Hawkes process. The use of Markov-modulated demands has regularly appeared in the inventory control literature; see, for example, Iglehart and Karlin (1960), Feldman (1978), Song and Zipkin (1993), Beyer and Sethi (1997), Chen and Song (2001). The focus in these papers is to characterize the optimal stocking policy through base stock levels, rather than giving fluid approximations or asymptotically optimal policies.

Outline: In Section 2, we formulate the revenue management problem with Markov-modulated demands. In Section 3, we give our fluid approximation and show that it yields an upper bound on the optimal total expected revenue. In Section 4, we give our approximate policy. In Section 5, we prove the performance guarantee for our approximate policy. In Section 6, we formulate the fluid approximation with varying histories. In Section 7, we give the reductions in this fluid approximation. In Section 8, we give our computational experiments. In Section 9, we conclude.

2. Problem Formulation

The set of resources is $\mathcal{L}$. The capacity of resource i is c_i. The set of products is $\mathcal{J}$. The revenue associated with product j is f_j. We use the vector $\boldsymbol{a}_j = (a_{ij} : i \in \mathcal{L}) \in \{0,1\}^{|\mathcal{L}|}$ to capture the set of resources used by product j, where $a_{ij} = 1$ if and only if product j uses resource i. There are K stages in the selling horizon indexed by $\mathcal{K} = \{1,\ldots,K\}$. The Markov chain that modulates the demand for the products makes transitions at the beginning of the stages. There are T time periods in each stage indexed by $\mathcal{T} = \{1,\ldots,T\}$. The time periods correspond to small enough intervals of time that there is at most one customer arrival at each time period. If the state of the Markov chain in stage k is s, then the customer arriving at time period t in stage k requests product j with probability $\lambda^t_{jk}(s)$. Noting that there is at most one customer arrival at each time period, we have $\sum_{j\in\mathcal{J}} \lambda^t_{jk}(s) \le 1$. The states of the Markov chain take values in the finite state space $\mathcal{X}$. We use the random variable X_k to denote the state of the Markov chain in stage k. We capture the probability law of the Markov chain by the transition probability $P_k(s,q) = \mathbb{P}\{X_{k+1} = q \,|\, X_k = s\}$. Note that the transition probability matrix $\{P_k(s,q) : s,q \in \mathcal{X}\}$ governing the evolution of the Markov chain over the different stages can be non-stationary.

We can use the Markov chain to capture the factors that can affect the demand for the products, such as customer tastes, weather conditions and economic environment. At the beginning of each stage, we observe the state of the Markov chain. We decide whether to accept the product request at each time period in the stage. At the end of the stage, the Markov chain transitions. Our goal is to find a policy to decide which customer requests to accept to maximize the total expected revenue over the selling horizon. We give a dynamic program to compute the optimal policy. We capture the remaining resource capacities with the vector $\boldsymbol{y} = (y_i : i \in \mathcal{L}) \in \mathbb{Z}_+^{|\mathcal{L}|}$, where y_i is the remaining capacity of resource i. We capture the decisions with the vector $\boldsymbol{u} = (u_j : j \in \mathcal{J}) \in \{0,1\}^{|\mathcal{J}|}$, where $u_j = 1$ if and only if we accept a request for product j. If the remaining resource capacities are given by the vector $\boldsymbol{y}$, then the set of feasible decisions is $\mathcal{F}(\boldsymbol{y}) = \{\boldsymbol{u} \in \{0,1\}^{|\mathcal{J}|} : a_{ij}\, u_j \le y_i \ \forall\, i \in \mathcal{L},\ j \in \mathcal{J}\}$, ensuring that if product j uses resource i, then we can accept a request for this product only if we have at least one unit of capacity for resource i. We use $(\boldsymbol{y}, s)$ as the state variable, where $\boldsymbol{y}$ corresponds to the remaining resource capacities and s corresponds to the state of the Markov chain. We can find the optimal policy by solving the dynamic program

$$J^t_k(\boldsymbol{y},s) \;=\; \max_{\boldsymbol{u}\in\mathcal{F}(\boldsymbol{y})} \left\{ \sum_{j\in\mathcal{J}} \lambda^t_{jk}(s) \Big\{ f_j\, u_j + J^{t+1}_k(\boldsymbol{y} - \boldsymbol{a}_j\, u_j, s) \Big\} \right\} + \Big\{ 1 - \sum_{j\in\mathcal{J}} \lambda^t_{jk}(s) \Big\}\, J^{t+1}_k(\boldsymbol{y},s) \qquad (1)$$

with boundary conditions $J^{T+1}_k(\boldsymbol{y},s) = \sum_{q\in\mathcal{X}} P_k(s,q)\, J^1_{k+1}(\boldsymbol{y},q)$ and $J^{T+1}_K(\boldsymbol{y},s) = 0$. The first boundary condition ensures that if the state of the Markov chain in stage k is s, then the Markov chain transitions to state q at the end of the stage with probability $P_k(s,q)$. The second boundary condition ensures that we do not collect revenue after the end of the selling horizon. In the dynamic program in (1), the state variable keeps track of both the remaining capacities of the resources and state of the Markov chain. If the state of the Markov chain in stage k is s, then the customer arriving at time period t in stage k requests product j with probability $\lambda^t_{jk}(s)$. If we accept this product request, then we generate the revenue f_j and consume the capacities of the resources used by product j. With probability $1 - \sum_{j \in \mathcal{J}} \lambda^t_{jk}(s)$, there is no customer arrival at time period t in stage k. Using $\boldsymbol{c} = (c_i : i \in \mathcal{L})$ to denote the capacities of the resources and noting that the initial state of the Markov chain is X_1, the optimal total expected revenue is $\mathsf{opt} = \mathbb{E}\{J^1_1(\boldsymbol{c}, X_1)\}$. In our model, the time periods correspond to small enough intervals of time that there is at most one customer arrival at each time period, whereas the stages correspond to large enough intervals of time that the factors that can affect the product demand have an opportunity to change. We can have as many as KT customer arrivals, but at most K Markov chain transitions from one stage to the next. Therefore, our model has two timescales, one fast timescale for the arrivals of the customers and one slow timescale for the transitions of the Markov chain, which, we believe, is a natural approach for incorporating demand modulation.

We make three assumptions for the Markov chain that modulates the demand. First, the evolution of the Markov chain is independent of the decisions that we make to accept or reject the product requests. In other words, the evolution of the Markov chain is exogenous. This assumption is reasonable when the Markov chain captures factors such as customer tastes, weather conditions and economic environment. An overwhelming majority of the papers in the existing literature that incorporate demand modulation, including those discussed in the introduction, proceed with the understanding that the evolution of the Markov chain is exogenous. Furthermore, we observe the state of the Markov chain at the beginning of each stage, which is also the prevalent assumption in the literature. Second, we assume that there exist some $\epsilon > 0$ such that we have $\mathbb{P}\{X_k = s\} \geq \epsilon$ for all $s \in \mathcal{X}$ and $k \in \mathcal{K}$. Thus, the Markov chain visits all states in all stages with strictly positive probability. Third, the total variation distance of the probability law for the Markov chain is given by $\mathsf{TV} = \max_{s,q \in \mathcal{X},\, k \in \mathcal{K}} \frac{1}{2} \sum_{p \in \mathcal{X}} |P_k(s,p) - P_k(q,p)|$. We assume that there exists some $\alpha > 0$ such that $\mathsf{TV} \leq 1 - \alpha$. This assumption is known as the Doeblin condition and it is commonly used to characterize the mixing time of a Markov chain; see Chapter 4 in Levin and Peres (2017). If all transition probabilities are strictly positive, then this assumption is satisfied, but having all strictly positive transition probabilities is not necessary for this assumption to hold.

Due to its high dimensional state variable, we cannot solve the dynamic program in (1) efficiently. We focus on building efficiently computable policies with performance guarantees.

3. Family of History-Dependent Fluid Approximations

We give a family of fluid approximations corresponding to the dynamic program in (1). The novel aspect of our fluid approximations is that they use decision variables that allow us to accept each product request with a probability that depends on the history of the Markov chain over a certain number of stages. Using decision variables that depend on longer histories of the Markov chain will allow us to obtain better approximations to the dynamic program in (1) in the sense that our fluid approximations will obtain tighter upper bounds on optimal total expected revenue and yield policies with stronger performance guarantees. For fixed $\theta \in \{1,\ldots,K\}$, we focus on history of the Markov chain over past θ stages. Let $\mathbf{X}_k^\theta = (X_{k-\theta+1},\ldots,X_k)$ be the history of the Markov chain in stage k over past θ stages. For notational uniformity, we use $\mathbf{X}_k^1$ to denote X_k throughout the rest of the paper. Let $\mathbf{s}_k^\theta = (s_{k-\theta+1},\ldots,s_k) \in \mathcal{X}^\theta$ be a realization of $\mathbf{X}_k^\theta = (X_{k-\theta+1},\ldots,X_k)$. In our fluid approximation, we use the decision variables $\boldsymbol{y} = (y_{jk}^t(\mathbf{s}_k^\theta) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}_k^\theta \in \mathcal{X}^\theta)$, where $y_{jk}^t(\mathbf{s}_k^\theta)$ is the probability that we accept a request for product j at time period t in stage k given that the history of the Markov chain is $\mathbf{s}_k^\theta$. We consider the linear program

$$
\begin{aligned}
Z_{\mathsf{LP}}^\theta \quad = \quad & \max_{\boldsymbol{y} \in \mathbb{R}_+^{|\mathcal{J}|TK|\mathcal{X}|^\theta}} \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} f_j\, \mathbb{E}\{y_{jk}^t(\mathbf{X}_k^\theta)\} \qquad (2) \\
& \text{st} \quad \sum_{\ell=1}^{k} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^\theta) \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\} \ \le\ c_i \quad \forall\, i \in \mathcal{L},\ k \in \mathcal{K},\ \mathbf{s}_k^\theta \in \mathcal{X}^\theta \\
& \qquad y_{jk}^t(\mathbf{s}_k^\theta) \ \le\ \lambda_{jk}^t(\mathbf{s}_k^1) \quad \forall\, j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}_k^\theta \in \mathcal{X}^\theta.
\end{aligned}
$$

The problem above is a linear program with finite numbers of decision variables and constraints. In particular, we can express the expectations in the objective function and first constraint, respectively, as $\sum_{\mathbf{s}_k^\theta \in \mathcal{X}^\theta} \mathbb{P}\{\mathbf{X}_k^\theta = \mathbf{s}_k^\theta\}\, y_{jk}^t(\mathbf{s}_k^\theta)$ and $\sum_{\mathbf{p}_\ell^\theta \in \mathcal{X}^\theta} \mathbb{P}\{\mathbf{X}_\ell^\theta = \mathbf{p}_\ell^\theta \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\}\, y_{j\ell}^t(\mathbf{p}_\ell^\theta)$. Computing the probabilities $\mathbb{P}\{\mathbf{X}_k^\theta = \mathbf{s}_k^\theta\}$ and $\mathbb{P}\{\mathbf{X}_\ell^\theta = \mathbf{p}_\ell^\theta \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\}$ a priori using the transition probabilities of the Markov chain, the objective function and constraints are linear in the decision variables. Also, the number of states of the Markov chain is finite, so the numbers of decision variables and constraints are finite. In the objective function, the expectation $\mathbb{E}\{y_{jk}^t(\mathbf{X}_t^\theta)\}$ is the expected sales for product j at time period t in stage k, so we account for the total expected revenue. In the first constraint, the expectation $\mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^\theta) \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\}$ is the expected sales for product j at time period t in stage ℓ, conditional on the fact that the history of the Markov chain in stage k is $\mathbf{s}_k^\theta$. By the first constraint, conditional on the history of the Markov chain in stage k being $\mathbf{s}_k^\theta$, the total expected consumption of resource i by the end of stage k does not exceed c_i. By the second constraint, the probability of accepting a product request does not exceed the request arrival probability.

In the second constraint above, considering the realization of the history $\mathbf{s}_k^\theta = (s_{k-\theta+1},\ldots,s_k)$, the realization of the history $\mathbf{s}_k^1 = (s_k)$ is the last element of the same vector. For $k < \theta$, the vector

$\mathbf{X}_k^\theta = (X_{k-\theta+1},\ldots,X_k)$ contains the states of the Markov chain before the beginning of the selling horizon, which are fixed and known. There are two novel aspects of problem (2). First, the first constraint ensures that the total expected capacity consumption of resource i until the end of stage k, conditional on the history in stage k, does not exceed the capacity of the resource. This conditioning is critical. Second, the probability of accepting a product request depends on the history of the Markov chain over past θ stages. By using decision variables that depend on the histories of the Markov chain with different lengths, we will obtain a family of fluid approximations. The number of decision variables and constraints in (2) increase exponentially in θ. In practice, we expect the transitions of the Markov chain to be significantly slower than the arrivals of the customers, so the number of stages would be relatively small. Also, using decision variables that depend on the history over the past few stages will be adequate to reap the benefits from our fluid approximation. Lastly, we will give performance guarantees for approximate policies obtained from our fluid approximation even when we use the history over a single stage.

In the dynamic program in (1), conditional on the current state, the evolution of the Markov chain is independent of the history, so we do not need to keep track of the history of the Markov chain in the state variable of the dynamic program. In the fluid approximation in (2), we only keep track of the expected capacity consumptions of the resources without capturing the evolution of the resource capacities over time periods. Thus, using decision variables that depend on the history of the Markov chain, which can give an indication of the evolution of the resource capacities, becomes useful. In particular, we will show that the linear program in (2) provides an upper bound on the optimal total expected revenue and this upper bound becomes tighter as we use longer histories. That is, noting that Z^θ_{LP} is the optimal objective value of (2) and opt is the optimal total expected revenue, we have $Z^1_{\mathsf{LP}} \geq Z^2_{\mathsf{LP}} \geq \ldots \geq Z^K_{\mathsf{LP}} \geq \mathsf{opt}$. Thus, by using longer histories in (2), this linear program approximates the optimal total expected revenue more accurately. Also, we will use our fluid approximation to give approximate policies with performance guarantees. The performance guarantees for our approximate policies are stronger when we use longer histories in (2).

In the next theorem, we show that the optimal objective value of problem (2) is an upper bound on the optimal total expected revenue for any choice of θ. We need one identity in the proof. Let the Bernoulli random variable Y^t_{jk} take value one if and only if the optimal policy accepts a request for product j at time period t in stage k. For stages ℓ and k that satisfy $\ell \leq k$, we have the identity $\mathbb{E}\{\mathbb{E}\{Y^t_{j\ell} \,|\, \mathbf{X}^\theta_\ell\} \,|\, \mathbf{X}^\theta_k\} = \mathbb{E}\{Y^t_{j\ell} \,|\, \mathbf{X}^\theta_k\}$. In particular, noting that $\ell \leq k$, conditional on $\mathbf{X}^\theta_\ell$, $(X_1,\ldots,X_\ell)$ are independent of $\mathbf{X}^\theta_k$. In this case, because the decisions of the optimal policy in stage ℓ depend on the history of the Markov chain up to and including stage ℓ, conditional on $\mathbf{X}^\theta_\ell$, $Y^t_{j\ell}$ is independent of $\mathbf{X}^\theta_k$, so we have $\mathbb{E}\{Y^t_{j\ell} \,|\, (\mathbf{X}^\theta_\ell, \mathbf{X}^\theta_k)\} = \mathbb{E}\{Y^t_{j\ell} \,|\, \mathbf{X}^\theta_\ell\}$. Taking the expectations of both sides in the last equality conditional on $\mathbf{X}_k^\theta$, we get $\mathbb{E}\{\mathbb{E}\{Y_{j\ell}^t \,|\, \mathbf{X}_\ell^\theta\} \,|\, \mathbf{X}_k^\theta\} = \mathbb{E}\{\mathbb{E}\{Y_{j\ell}^t \,|\, (\mathbf{X}_\ell^\theta, \mathbf{X}_k^\theta)\} \,|\, \mathbf{X}_k^\theta\}$. Because the filtration generated by $\mathbf{X}_k^\theta$ is a subset of the one generated by $(\mathbf{X}_\ell^\theta, \mathbf{X}_k^\theta)$, by the tower property of conditional expectation, we have $\mathbb{E}\{\mathbb{E}\{Y_{j\ell}^t \,|\, (\mathbf{X}_\ell^\theta, \mathbf{X}_k^\theta)\} \,|\, \mathbf{X}_k^\theta\} = \mathbb{E}\{Y_{j\ell}^t \,|\, \mathbf{X}_k^\theta\}$. In this case, the last two equalities yield $\mathbb{E}\{\mathbb{E}\{Y_{j\ell}^t \,|\, \mathbf{X}_\ell^\theta\} \,|\, \mathbf{X}_k^\theta\} = \mathbb{E}\{Y_{j\ell}^t \,|\, \mathbf{X}_k^\theta\}$. Conditioning arguments similar to the one in this paragraph will be critical at numerous points in our subsequent development.

In the next theorem, we use the identity in the previous paragraph to show that the optimal objective value of (2) is indeed an upper bound on the optimal total expected revenue.

Theorem 3.1 (Upper Bound) *Letting* opt *be the optimal total expected revenue and* Z_{LP}^θ *be the optimal objective value of problem* (2)*, we have* $Z_{\mathsf{LP}}^\theta \geq \mathsf{opt}$ *for all* $\theta = 1, \ldots, K$.

Proof: Let the Bernoulli random variable Y_{jk}^t take value one if and only if the optimal policy accepts a request for product j at time period t in stage k. For problem (2), we define the solution $\overline{\boldsymbol{y}} = (\overline{y}_{jk}^t(\mathbf{s}_k^\theta) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}_k^\theta \in \mathcal{X}^\theta)$ as $\overline{y}_{jk}^t(\mathbf{s}_k^\theta) = \mathbb{E}\{Y_{jk}^t \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\}$. We establish that the solution $\overline{\boldsymbol{y}}$ is feasible to (2). Using the definition of $\overline{y}_{j\ell}^t(\mathbf{s}_\ell^\theta)$, for stages ℓ and k that satisfy $\ell \leq k$, we have $\mathbb{E}\{\overline{y}_{j\ell}^t(\mathbf{X}_\ell^\theta) \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\} = \mathbb{E}\{\mathbb{E}\{Y_{j\ell}^t \,|\, \mathbf{X}_\ell^\theta\} \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\} = \mathbb{E}\{Y_{j\ell}^t \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\}$, where the last equality is by the identity just before the theorem. Under the optimal policy, the total capacity consumption of resource i until the end of stage k does not exceed its capacity, so $\sum_{\ell=1}^k \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} a_{ij}\, Y_{j\ell}^t \leq c_i$ with probability one. Taking the expectations of both sides conditional on $\mathbf{X}_k^\theta = \mathbf{s}_k^\theta$ and using the last equality, we get $\sum_{\ell=1}^k \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \mathbb{E}\{\overline{y}_{j\ell}^t(\mathbf{X}_\ell^\theta) \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\} \leq c_i$, so the solution $\overline{\boldsymbol{y}}$ satisfies the first constraint in (2). By the definition of $\overline{y}_{jk}^t(\mathbf{s}_k^\theta)$, we have $\mathbb{E}\{\overline{y}_{jk}^t(\mathbf{X}_k^\theta)\} = \mathbb{E}\{\mathbb{E}\{Y_{jk}^t \,|\, \mathbf{X}_k^\theta\}\} = \mathbb{E}\{Y_{jk}^t\}$, where the last equality uses the tower property of conditional expectation. Letting the Bernoulli random variable B_{jk}^t take value one if and only if we have a request for product j at time period t in stage k, we have $\mathbb{E}\{B_{jk}^t \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\} = \lambda_{jk}^t(\mathbf{s}_k^1)$. Under the optimal policy, we can accept a request for product j at time period t in stage k only when we have a request for the product, which implies that $Y_{jk}^t \leq B_{jk}^t$ with probability one. Taking the expectations of both sides in the last inequality conditional on $\mathbf{X}_k^\theta = \mathbf{s}_k^\theta$, we obtain $\boldsymbol{y}_{jk}^t(\mathbf{s}_k^\theta) = \mathbb{E}\{Y_{jk}^t \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\} \leq \mathbb{E}\{B_{jk}^t \,|\, \mathbf{X}_k^\theta = \mathbf{s}_k^\theta\} = \lambda_{jk}^t(\mathbf{s}_k^1)$. Thus, the solution $\overline{\boldsymbol{y}}$ satisfies the second constraint in (2). Recalling that $\mathbb{E}\{\overline{y}_{jk}^t(\mathbf{X}_k^\theta)\} = \mathbb{E}\{Y_{jk}^t\}$ by the discussion in this paragraph, by the definition of Y_{jk}^t, the optimal total expected revenue is

$$\mathsf{opt} = \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} f_j\, \mathbb{E}\{Y_{jk}^t\} = \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} f_j\, \mathbb{E}\{\overline{y}_{jk}^t(\mathbf{X}_k^\theta)\}.$$

Therefore, the feasible solution $\overline{\boldsymbol{y}}$ provides an objective value of opt for problem (2), which implies that the optimal objective value of this problem satisfies $Z_{\mathsf{LP}}^\theta \geq \mathsf{opt}$. ■

In the proof of the theorem, we use the decisions of the optimal policy to construct a feasible solution to problem (2) such that the objective value provided by this solution for problem (2) is equal to the optimal total expected revenue. This approach is useful for verifying that the optimal objective value of problem (2) is an upper bound, but it does not allow us to construct the form of the fluid approximation from scratch. It is possible to also give a constructive proof for the theorem. We can relax the capacity constraints in (1) by using Lagrange multipliers that depend on the history of the Markov chain over past θ stages. The value functions of the relaxed dynamic program has a closed form expression. For any choice of the Lagrange multipliers, the initial value function of the relaxed dynamic program yields an upper bound on the optimal total expected revenue. Thus, we can solve an auxiliary problem that chooses the Lagrange multipliers to make the upper bound from the relaxed dynamic program as tight as possible. The dual of the auxiliary problem is the fluid approximation in (2). Although the ideas in the non-constructive and constructive proofs have somewhat standard components, the non-constructive approach requires the careful conditioning argument, which does not appear in the earlier literature, whereas the constructive approach uses Lagrange multipliers that depend on the history of the Markov chain.

In the next theorem, we show that the upper bound from problem (2) gets tighter as we use longer histories of the Markov chain. In the proof, we use a variant of the conditioning argument that we gave just before Theorem 3.1. Let $\overline{\boldsymbol{y}} = (\overline{y}^t_{jk}(\mathbf{s}^{\theta+1}_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^{\theta+1}_k \in \mathcal{X}^{\theta+1})$ be an optimal solution to problem (2) with $\theta+1$ stages in the history. For stages ℓ and k that satisfy $\ell \leq k$, we have the identity $\mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^{\theta+1}_k\} \,|\, \mathbf{X}^\theta_k\} = \mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^\theta_\ell\} \,|\, \mathbf{X}^\theta_k\}$. In particular, noting that $\ell \leq k$, conditional on $\mathbf{X}^\theta_\ell$, $(X_1, \ldots, X_\ell)$ are independent of $\mathbf{X}^\theta_k$. In this case, we have $\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, (\mathbf{X}^\theta_\ell, \mathbf{X}^\theta_k)\} = \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^\theta_\ell\}$. Taking the expectations of both sides in the last equality conditional on $\mathbf{X}^\theta_k$, we get $\mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^\theta_\ell\} \,|\, \mathbf{X}^\theta_k\} = \mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, (\mathbf{X}^\theta_\ell, \mathbf{X}^\theta_k)\} \,|\, \mathbf{X}^\theta_k\} = \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^\theta_k\}$, where the second equality uses the fact that the filtration generated by $\mathbf{X}^\theta_k$ is a subset of the one generated by $(\mathbf{X}^\theta_\ell, \mathbf{X}^\theta_k)$ and the tower property of conditional expectation. The filtration generated by $\mathbf{X}^\theta_k$ is a subset of the one generated by $\mathbf{X}^{\theta+1}_k$, so by the tower property of conditional expectation, we have $\mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^{\theta+1}_k\} \,|\, \mathbf{X}^\theta_k\} = \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^\theta_k\}$ as well. The last two chains of equalities yield $\mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^{\theta+1}_k\} \,|\, \mathbf{X}^\theta_k\} = \mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell) \,|\, \mathbf{X}^\theta_\ell\} \,|\, \mathbf{X}^\theta_k\}$.

In the next theorem, we use the conditioning argument in the previous paragraph to show that the upper bound from problem (2) indeed gets tighter as we use longer histories.

Theorem 3.2 (Nested Approximations) *Letting Z^θ_{LP} be the optimal objective value of problem (2), we have $Z^\theta_{\mathsf{LP}} \geq Z^{\theta+1}_{\mathsf{LP}}$ for all $\theta = 1, \ldots, K-1$.*

Proof: Let $\overline{\boldsymbol{y}} = (\overline{y}^t_{jk}(\mathbf{s}^{\theta+1}_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^{\theta+1}_k \in \mathcal{X}^{\theta+1})$ be an optimal solution to problem (2) with $\theta+1$ stages in the history. Considering problem (2) with θ stages in the history, we define the solution $\overline{\boldsymbol{z}} = (\overline{z}^t_{jk}(\mathbf{s}^\theta_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^\theta_k \in \mathcal{X}^\theta)$ as $\overline{z}^t_{jk}(\mathbf{s}^\theta_k) = \mathbb{E}\{\overline{\boldsymbol{y}}^t_{jk}(\mathbf{X}^{\theta+1}_k) \,|\, \mathbf{X}^\theta_k = \mathbf{s}^\theta_k\}$. We establish that the solution $\overline{\boldsymbol{z}}$ is feasible to (2) with θ stages in the history. Using the definition of $\overline{z}^t_{j\ell}(\mathbf{s}^\theta_\ell)$, we have $\mathbb{E}\{\overline{z}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^\theta_k\} = \mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell)\,|\,\mathbf{X}^\theta_\ell\}\,|\,\mathbf{X}^\theta_k\} = \mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell)\,|\,\mathbf{X}^{\theta+1}_k\}\,|\,\mathbf{X}^\theta_k\}$ for stages ℓ and k that satisfy $\ell \le k$, where the second equality uses the identity given just before the theorem. Because $\overline{\boldsymbol{y}}$ is feasible to (2) with $\theta+1$ stages in the history, by the first constraint in (2), we have $\sum_{\ell=1}^k \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\,\mathbb{E}\{y^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell)\,|\,\mathbf{X}^{\theta+1}_k\} \le c_i$ with probability one. Noting that we have $\mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\theta+1}_\ell)\,|\,\mathbf{X}^{\theta+1}_k\}\,|\,\mathbf{X}^\theta_k = \mathbf{s}^\theta_k\} = \mathbb{E}\{\overline{z}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^\theta_k = \mathbf{s}^\theta_k\}$, taking the expectations of both sides of the last inequality conditional on $\mathbf{X}^\theta_k = \mathbf{s}^\theta_k$, it immediately follows that $\overline{\boldsymbol{z}}$ satisfies the first constraint in (2) with θ stages in the history. By the second constraint in (2), we have $\overline{y}^t_{jk}(\mathbf{X}^{\theta+1}_k) \le \lambda^t_{jk}(\mathbf{X}^1_k)$ with probability one. Conditional on $\mathbf{X}^\theta_k = \mathbf{s}^\theta_k$, we have $\mathbf{X}^1_k = \mathbf{s}^1_k$, because the vectors $\mathbf{X}^\theta_k$ and $\mathbf{X}^1_k$ share the same last component. In this case, taking the expectations of both sides of the last inequality conditional on $\mathbf{X}^\theta_k = \mathbf{s}^\theta_k$, we get $\overline{z}^t_{jk}(\mathbf{s}^\theta_k) = \mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^{\theta+1}_k)\,|\,\mathbf{X}^\theta_k = \mathbf{s}^\theta_k\} \le \lambda^t_{jk}(\mathbf{s}^1_k)$, so the solution $\overline{\boldsymbol{z}}$ satisfies the second constraint in (2) with θ stages in the history as well. For problem (2) with θ stages in the history, the feasible solution $\overline{\boldsymbol{z}}$ provides an objective value of

$$\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} f_j\,\mathbb{E}\{\overline{z}^t_{jk}(\mathbf{X}^\theta_k)\} \overset{(a)}{=} \sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} f_j\,\mathbb{E}\{\mathbb{E}\{\overline{\boldsymbol{y}}^t_{jk}(\mathbf{X}^{\theta+1}_k)\,|\,\mathbf{X}^\theta_k\}\} \overset{(b)}{=} \sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} f_j\,\mathbb{E}\{\overline{\boldsymbol{y}}^t_{jk}(\mathbf{X}^{\theta+1}_k)\} = Z^{\theta+1}_{\mathsf{LP}},$$

where (a) is by the definition of $\overline{z}^t_{jk}(\mathbf{s}^\theta_k)$ and (b) uses the tower property of conditional expectation. Thus, the optimal objective value of problem (2) with θ stages in the history is at least $Z^{\theta+1}_{\mathsf{LP}}$. ■

By the discussion so far, we can use our fluid approximation to get an upper bound on the optimal total expected revenue. We proceed to extracting policies from our fluid approximation.

4. Approximate Policy from the Fluid Approximation

We use our fluid approximation given in problem (2) to construct an approximate policy with a performance guarantee. The performance guarantee of our approximate policy gets arbitrarily close to one as we scale up the capacities of the resources and number of stages in the selling horizon linearly with the same rate. In our problem setup, adding additional stages to the selling horizon is a natural way to scale up the total expected demand for the products. Therefore, if we focus on an asymptotic regime where we scale up the capacities of the resources and total expected demand for the products linearly with the same rate, then the performance guarantee of our approximate policy gets arbitrarily close to one. Furthermore, our approximate policy has a performance guarantee irrespective of the number of stages that we use in the history of the Markov chain in problem (2), but the performance guarantee becomes stronger when we use a larger number of stages in the history. The construction of our approximate policy requires carefully aggregating the decision variables in an optimal solution to our fluid approximation, as we discuss next. In our approximate policy, we solve problem (2) once at the beginning of the selling horizon.

Letting $\overline{\boldsymbol{y}} = (\overline{y}^t_{jk}(\mathbf{s}^\theta_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^\theta_k \in \mathcal{X}^\theta)$ be an optimal solution, using $\gamma \in (0,1)$ to denote a tuning parameter, our approximate policy makes its decisions as follows.

Approximate Policy:

- Consider having a request for product j at time period t in stage k when the history of the Markov chain over past θ stages is $\mathbf{s}^\theta_k$.
- If the current stage satisfies $k \le K - \theta$, then we are willing to accept the product request with probability $\gamma\, \mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k) \,|\, \mathbf{X}^1_k = \mathbf{s}^1_k\} / \lambda^t_{jk}(\mathbf{s}^1_k)$.
- If the current stage satisfies $k \ge K - \theta + 1$, then we are willing to accept the product request with probability $\gamma\, \mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k) \,|\, \mathbf{X}^{k-(K-\theta)}_k = \mathbf{s}^{k-(K-\theta)}_k\} / \lambda^t_{jk}(\mathbf{s}^1_k)$.
- If we are willing to accept the product request and the remaining resource capacities allow accepting the product request, then we accept the product request. Otherwise, we reject.

Because $\theta \ge 1$, the knowledge of $\mathbf{s}^\theta_k$ implies the knowledge of $\mathbf{s}^1_k$. Thus, we can compute the expectation in the first case as $\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k) \,|\, \mathbf{X}^1_k = \mathbf{s}^1_k\} = \sum_{\mathbf{p}^\theta_k \in \mathcal{X}^\theta} \mathbb{P}\{\mathbf{X}^\theta_k = \mathbf{p}^\theta_k \,|\, \mathbf{X}^1_k = \mathbf{s}^1_k\}\, \overline{y}^t_{jk}(\mathbf{p}^\theta_k)$. On the other hand, because $k \le K$, we have $k - (K - \theta) \le \theta$, so the knowledge of $\mathbf{s}^\theta_k$ implies the knowledge of $\mathbf{s}^{k-(K-\theta)}_k$. Therefore, we can also compute the conditional expectation in the second case as $\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k) \,|\, \mathbf{X}^{k-(K-\theta)}_k = \mathbf{s}^{k-(K-\theta)}_k\} = \sum_{\mathbf{p}^\theta_k \in \mathcal{X}^\theta} \mathbb{P}\{\mathbf{X}^\theta_k = \mathbf{p}^\theta_k \,|\, \mathbf{X}^{k-(K-\theta)}_k = \mathbf{s}^{k-(K-\theta)}_k\}\, \overline{y}^t_{jk}(\mathbf{p}^\theta_k)$. By the preceding discussion, the knowledge of $\mathbf{s}^\theta_k$ is enough to compute the conditional expectations needed to come up with the probability of accepting a product request in either of the two cases. Noting the form of the two conditional expectations in this paragraph, our approximate policy aggregates the decision variables $(\overline{y}^t_{jk}(\mathbf{s}^\theta_k) : \mathbf{s}^\theta_k \in \mathcal{X}^\theta)$ through a convex combination to come up with the probability of accepting a request for product j at time period t in stage k.

To capture the two cases in our approximate policy, we define the *aggregation duration* as $\sigma_k = 1$ if $k \le K - \theta$, whereas $\sigma_k = k - (K - \theta)$ if $k \ge K - \theta + 1$. In this case, we can express both conditional expectations in our approximate policy as $\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k) \,|\, \mathbf{X}^{\sigma_k}_k = \mathbf{s}^{\sigma_k}_k\}$. By the definition σ_k, we have $1 \le \sigma_k \le \theta$. In the next section, to establish a performance guarantee for our approximation policy, we use the aggregation duration to succinctly capture the two conditional expectations in our approximate policy. If we choose $\theta = K$ so that we use the full history of the Markov chain in our fluid approximation in (2), then we never use the first case. For $\theta = K$, the histories $\mathbf{X}^{k-(K-\theta)}_k$ and $\mathbf{X}^\theta_k$ are equivalent, including the full history of the Markov chain up to and including stage k. Thus, the conditional expectation in the second case becomes $\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k) \,|\, \mathbf{X}^\theta_k = \mathbf{s}^\theta_k\} = \overline{y}^t_{jk}(\mathbf{s}^\theta_k)$, so our approximate policy does not aggregate the decision variables $(\overline{y}^t_{jk}(\mathbf{s}^\theta_k) : \mathbf{s}^\theta_k \in \mathcal{X}^\theta)$ when $\theta = K$.

Given that the history of the Markov chain over past θ stages is $\mathbf{s}^\theta_k$, a more natural approximate policy would, perhaps, be willing to accept a request for product j at time period t in stage k with probability $\gamma\,\overline{y}^t_{jk}(\mathbf{s}^\theta_k)/\lambda^t_{jk}(\mathbf{s}^1_k)$. This approximate policy does not aggregate the decision variables $(\overline{y}^t_{jk}(\mathbf{s}^\theta_k) : \mathbf{s}^\theta_k \in \mathcal{X}^\theta)$. For such an approximate policy that does not use aggregation, we can give a performance guarantee and this performance guarantee also gets arbitrarily close to one as we scale up the capacities of the resources and number of stages in the selling horizon linearly with the same rate, but despite our best efforts, we were not able to establish that this performance guarantee gets stronger as we use a larger number of stages in the history. The performance guarantee that we will give for our approximate policy with aggregation is at least as strong as the one we could establish for the approximate policy without aggregation and it gets stronger as we use a larger number of stages in the history. Nevertheless, our results do not preclude the possibility of establishing even stronger performance guarantees for an approximate policy without aggregation. We have to leave such an investigation as an open question for further research.

Lastly, by the second constraint in (2), we have $\overline{y}^t_{jk}(\mathbf{X}^\theta_k) \le \lambda^t_{jk}(\mathbf{X}^1_k)$ with probability one. Because $1 \le \sigma_k \le \theta$, taking the expectations of both sides in the last inequality conditional on $\mathbf{X}^{\sigma_k}_k = \mathbf{s}^{\sigma_k}_k$, we get $\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k) \,|\, \mathbf{X}^{\sigma_k}_k = \mathbf{s}^{\sigma_k}_k\} \le \lambda^t_{jk}(\mathbf{s}^1_k)$, so both acceptance probabilities in the description of our approximate policy are indeed less than one. In the next theorem, we give a performance guarantee for our approximate policy. We use $c_{\min} = \min_{i\in\mathcal{L}} c_i$ to denote the smallest resource capacity and $L = \max_{j\in\mathcal{J}} \sum_{i\in\mathcal{L}} a_{ij}$ to denote the largest number of resources used by a product. Recalling the assumptions regarding the Markov chain, we have $\mathbb{P}\{X_k = s\} \ge \epsilon$ for all $s \in \mathcal{X}$ and $k \in \mathcal{K}$, as well as $\mathsf{TV} \le 1-\alpha$, for some $\epsilon, \alpha > 0$, where TV refers to the total variation distance for the probability law of the Markov chain. For notational brevity, we set $\nu^\theta = 32\,\frac{T}{\epsilon^2\alpha^2}\,\min\{(K-\theta)\,T, c_{\min}\}$, which is decreasing in θ. As a function of the number of stages θ in the history of the Markov chain in problem (2), we have the following performance guarantee for our approximate policy.

Theorem 4.1 (Performance of Approximate Policy) *Letting* APX^θ *be the total expected revenue of the approximate policy as a function of the number of stages in the history of the Markov chain in* (2)*, there exist a choice of the tuning parameter* γ *such that we have*

$$\frac{\mathsf{APX}^\theta}{\mathsf{opt}} \;\ge\; \frac{\mathsf{APX}^\theta}{Z^\theta_{\mathsf{LP}}} \;\ge\; 1 - \sqrt{\frac{4\,(\nu^\theta + c_{\min})\log c_{\min}}{c^2_{\min}}} - \frac{L}{c_{\min}}.$$

We devote the next section to the proof of the theorem. To interpret the performance guarantee in the theorem, we fix T, L, ϵ and α. There are K stages in the selling horizon and T time periods in each stage, so the total demand for the capacity of any resource is at most KT, which implies that we can proceed with the understanding that $c_{\min} \le KT$. Because $\nu^\theta \le 32\,\frac{T}{\epsilon^2\alpha^2}(K-\theta)\,T$, we get $\nu^\theta + c_{\min} \le 32\,\frac{T^2}{\epsilon^2\alpha^2}(2K-\theta)$. In this case, noting that we fix T, L, ϵ and α, the expression on the right side of the chain of inequalities in the theorem is of the form $1 - O\Big(\sqrt{\frac{(2-\frac{\theta}{K})\,K\,\log c_{\min}}{c^2_{\min}}}\Big)$. Thus, ignoring the logarithmic term, if we scale up the smallest resource capacity and number of stages in the selling horizon linearly with the same rate β so that $c_{\min} = \beta\,\overline{C}$ and $K = \beta\,\overline{K}$ for fixed $\overline{C}, \overline{K} > 0$, then the relative gap between the total expected revenue from the approximate policy and optimal total expected revenue approaches to one with rate $1 - O\Big(\sqrt{\frac{2-\frac{\theta}{K}}{\beta}}\Big)$. The fraction of the number of stages included in the history of the Markov chain is $\frac{\theta}{K}$. This fraction controls the rate with which the relative optimality gap of the approximate policy diminishes. In this interpretation of the performance guarantee, we linearly scale up the smallest resource capacity and number of stages. The resources capacities that do not correspond to the smallest resource capacity could be scaling up or the problem parameters other than T, L, ϵ and α could be changing arbitrarily.

On the other hand, we have $\nu^\theta \le 32\,\frac{T}{\epsilon^2\alpha^2} c_{\min}$ as well, in which case, because $\alpha, \epsilon < 1$, we get $\nu^\theta + c_{\min} \le 33\,\frac{T}{\epsilon^2\alpha^2} c_{\min}$. Noting, once again, that we fix T, L, ϵ and α, the expression on the right side of the chain of inequalities in the theorem is of the form $1 - O\Big(\sqrt{\frac{\log c_{\min}}{c_{\min}}}\Big)$. In this case, ignoring the logarithmic term, if we scale up the smallest resource capacity with the rate β so that $c_{\min} = \beta\,\overline{C}$ for fixed $\overline{C} > 0$, then the relative gap between the total expected revenue of the approximate policy and optimal total expected revenue approaches to one with rate $1 - O\Big(\sqrt{\frac{1}{\beta}}\Big)$. In this interpretation of the performance guarantee, we do not scale the number of stages, so our assumptions are milder, but we do not capture the effect of the length of the history of the Markov chain on the rate with which the relative optimality gap of the approximate policy diminishes.

The result in Theorem 4.1 is, to our knowledge, the *first one* to give an asymptotically optimal policy under Markov-modulated demands. This result is arguably more nuanced than the existing asymptotic optimality guarantees in the sense that it gives an explicit characterization of the effect of the number of stages in the history of the Markov chain in problem (2) on the performance guarantee for the approximate policy.

5. Performance Guarantee

We give a proof for Theorem 4.1. The proof has two parts. In the first part, we show that if we can lower bound the probability that each resource is available at each time period in each stage in the selling horizon under our approximate policy, then we can upper bound the optimality gap of the approximate policy. The first part follows a standard argument. In the second part, we lower bound the probability that each resource is available at each time period in each stage in the selling horizon. To construct this lower bound, we show that the capacity consumption of a resource under our approximate policy satisfies the so-called average Lipschitz condition. Using the average Lipschitz condition, we upper bound the moment generating function of the capacity consumption of a resource, which, in turn, yields a lower bound on the probability that the resource is available at each time period in each stage in the selling horizon. The use of the average Lipschitz condition to upper bound the moment generating functions of the capacity consumptions of resources in a revenue management setting is, as far as we are aware, novel.

Upper Bound on the Optimality Gap of the Approximate Policy:

To upper bound the optimality gap of the approximate policy, we define a sequence of random variables, which will allow us to express the total expected revenue of the approximate policy. In particular, for each $j \in \mathcal{J}$, $t \in \mathcal{T}$ and $k \in \mathcal{K}$, we define the following random variables.

- *Product Request*. For each $\mathbf{s}_k^\theta \in \mathcal{X}^\theta$, the Bernoulli random variable $A_{jk}^t(\mathbf{s}_k^1)$ takes value one if have a request for product j at time period t in stage k given that the history of the Markov chain in stage k is $\mathbf{s}_k^\theta$. Therefore, we have $\mathbb{P}\{A_{jk}^t(\mathbf{s}_k^1) = 1\} = \lambda_{jk}^t(\mathbf{s}_k^1)$.

- *Decision*. For each $\mathbf{s}_k^\theta \in \mathcal{X}^\theta$, the Bernoulli random variable $Y_{jk}^t(\mathbf{s}_k^{\sigma_k})$ takes value one if our approximate policy is willing to accept a request for product j at time period t in stage k given that the history of the Markov chain in stage k is $\mathbf{s}_k^\theta$, so $\mathbb{P}\{Y_{jk}^t(\mathbf{s}_k^{\sigma_k}) = 1\} = \frac{\gamma}{\lambda_{jk}^t(\mathbf{s}_k^1)} \mathbb{E}\{\overline{y}_{jk}^t(\mathbf{X}_k^\theta) \,|\, \mathbf{X}_k^{\sigma_k} = \mathbf{s}_k^{\sigma_k}\}$.

- *Capacity Availability*. The Bernoulli random variable G_{jk}^t takes value one if we have capacity to accept a request for product j at time period t in stage k under our approximate policy. It is difficult to characterize the probability $\mathbb{P}\{G_{jk}^t = 1\}$, but we will lower bound it.

Other work employs analogues of the three random variables above to analyze performance guarantees for approximate policies from fluid approximations; see, for example, Feng et al. (2024). In the discussion above, $\overline{\boldsymbol{y}} = (\overline{y}_{jk}^t(\mathbf{s}_k^\theta) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}_k^\theta \in \mathcal{X}^\theta)$ is an optimal solution to problem (2). The random variables $A_{jk}^t(\mathbf{s}_k^1)$ and $Y_{jk}^t(\mathbf{s}_k^{\sigma_k})$ are independent of each other. Using the vector $\mathbf{A}_k^t = (A_{jk}^t(\mathbf{s}_k^1) : j \in \mathcal{J},\ \mathbf{s}_k^1 \in \mathcal{X})$, the vectors $\{\mathbf{A}_k^t : t \in \mathcal{T},\ k \in \mathcal{K}\}$ are independent of each other, which is to say that the product requests during the course of the selling horizon, for fixed trajectory of the underlying Markov chain, are independent of each other. Similarly, using the vector $\mathbf{Y}_k^t = (Y_{jk}^t(\mathbf{s}_k^{\sigma_k}) : j \in \mathcal{J},\ \mathbf{s}_k^{\sigma_k} \in \mathcal{X}^{\sigma_k})$, the vectors $\{\mathbf{Y}_k^t : t \in \mathcal{T},\ k \in \mathcal{K}\}$ are independent of each other. We argue that if we can lower bound the resource availability probabilities, then we can upper bound the optimality gap of our approximate policy. Under the approximate policy, we have a sale for product j at time period t in stage k if and only if there is a request for product j at time period t in stage k, the approximate policy is willing to accept the product request and there is capacity to accept the product request. Therefore, the sales for product j at time period t in stage k is given by the random variable $S_{jk}^t = A_{jt}^t(\mathbf{X}_k^1)\, Y_{jk}^t(\mathbf{X}_k^{\sigma_k})\, G_{jk}^t$.

Taking the expectation of the expression in the previous paragraph conditional on $\mathbf{X}_k^{\sigma_k} = \mathbf{s}_k^{\sigma_k}$, we get $\mathbb{E}\{S_{jk}^t \,|\, \mathbf{X}_k^{\sigma_k} = \mathbf{s}_k^{\sigma_k}\} = \mathbb{P}\{G_{jk}^t = 1 \,|\, \mathbf{X}_k^{\sigma_k} = \mathbf{s}_k^{\sigma_k}\}\, \mathbb{E}\{A_{jt}^t(\mathbf{X}_k^1)\, Y_{jk}^t(\mathbf{X}_k^{\sigma_k}) \,|\, \mathbf{X}_k^{\sigma_k} = \mathbf{s}_k^{\sigma_k},\ G_{jk}^t = 1\}$. In the last conditional expectation, the remaining capacities of the resources at time period t in stage k

depends only on the product requests and decisions of the approximate policy in the earlier time periods, so we can drop the condition $G^t_{jk}=1$. Also, given that $\mathbf{X}^{\sigma_k}_k=\mathbf{s}^{\sigma_k}_k$, the random variables $A^t_{jk}(\mathbf{X}^1_k)$ and $Y^t_{jk}(\mathbf{X}^{\sigma_k}_k)$ are independent of each other. Thus, the last conditional expectation is $\mathbb{E}\{A^t_{jk}(\mathbf{s}^1_k)\,|\,\mathbf{X}^{\sigma_k}_k=\mathbf{s}^{\sigma_k}_k\}\,\mathbb{E}\{Y^t_{jk}(\mathbf{s}^{\sigma_k}_k)\,|\,\mathbf{X}^{\sigma_k}_k=\mathbf{s}^{\sigma_k}_k\}=\lambda^t_{jk}(\mathbf{s}^1_k)\,\frac{\gamma}{\lambda^t_{jk}(\mathbf{s}^1_k)}\,\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\,|\,\mathbf{X}^{\sigma_k}_k=\mathbf{s}^{\sigma_k}_k\}$. Therefore, the expected sales for product j at time period t in stage k conditional on $\mathbf{X}^{\sigma_k}_k$ is given by $\mathbb{E}\{S^t_{jk}\,|\,\mathbf{X}^{\sigma_k}_k\}=\mathbb{P}\{G^t_{jk}=1\,|\,\mathbf{X}^{\sigma_k}_k\}\,\gamma\,\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\,|\,\mathbf{X}^{\sigma_k}_k\}$. If we can lower bound the product availability probabilities uniformly by β so that $\mathbb{P}\{G^t_{jk}=1\,|\,\mathbf{X}^{\sigma_k}_k\}\ge\beta$, then $\mathbb{E}\{S^t_{jk}\,|\,\mathbf{X}^{\sigma_k}_k\}\ge\gamma\,\beta\,\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\,|\,\mathbf{X}^{\sigma_k}_k\}$, in which case, taking the expectations of both sides of this inequality and using the tower property of conditional expectation, the expected sales for product j at time period t in stage k under our approximate policy is lower bounded as $\mathbb{E}\{S^t_{jk}\}\ge\gamma\,\beta\,\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\}$. Using APX^θ to denote the total expected revenue of our approximate policy, noting the definition of S^t_{jk}, we get

$$\mathsf{APX}^\theta\;=\;\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}f_j\,\mathbb{E}\{S^t_{jk}\}\;\ge\;\gamma\,\beta\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}f_j\,\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\}\;\overset{(a)}{=}\;\gamma\,\beta\,Z^\theta_{\mathsf{LP}}\;\overset{(b)}{\ge}\;\gamma\,\beta\,\mathsf{opt},\qquad(3)$$

where (a) uses the definition of $\overline{\boldsymbol{y}}$ and (b) uses Theorem 3.1. Therefore, if we can show that $\mathbb{P}\{G^t_{jk}=1\,|\,\mathbf{X}^{\sigma_k}_k\}\ge\beta$ with probability one for some $\beta>0$, then $\mathsf{APX}^\theta\ge\gamma\,\beta\,\mathsf{opt}$.

We focus on lower bounding $\mathbb{P}\{G^t_{jk}=1\,|\,\mathbf{X}^{\sigma_k}_k\}$. Noting that S^t_{jk} is the sales for product j at time period t in stage k, the capacity consumption of resource i at time period t in stage k is given by the random variable $\sum_{j\in\mathcal{J}}a_{ij}\,S^t_{jk}=\sum_{j\in\mathcal{J}}a_{ij}\,A^t_{jt}(\mathbf{X}^1_k)\,Y^t_{jk}(\mathbf{X}^{\sigma_k}_k)\,G^t_{jk}$. Dropping the random variable G^t_{jk} in the last sum, we upper bound the capacity consumption of resource i at time period t in stage k by $\sum_{j\in\mathcal{J}}a_{ij}\,A^t_{jt}(\mathbf{X}^1_k)\,Y^t_{jk}(\mathbf{X}^{\sigma_k}_k)$. We define the random variable $U^t_{ik}(\mathbf{s}^{\sigma_k}_k)=\sum_{j\in\mathcal{J}}a_{ij}\,A^t_{jk}(\mathbf{s}^1_k)\,Y^t_{jk}(\mathbf{s}^{\sigma_k}_k)$. By the definitions of $A^t_{jk}(\mathbf{s}^1_k)$ and $Y^t_{jk}(\mathbf{s}^{\sigma_k}_k)$, the random variable $U^t_{ik}(\mathbf{s}^{\sigma_k}_k)$ is Bernoulli with mean $\gamma\sum_{j\in\mathcal{J}}a_{ij}\,\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\,|\,\mathbf{X}^{\sigma_k}_k=\mathbf{s}^{\sigma_k}_k\}$. Also, using the vector $\mathbf{U}^t_k=(U^t_{ik}(\mathbf{s}^{\sigma_k}_k):i\in\mathcal{L},\;\mathbf{s}^{\sigma_k}_k\in\mathcal{X}^{\sigma_k})$, the vectors $\{\mathbf{U}^t_k:t\in\mathcal{T},\;k\in\mathcal{K}\}$ are independent of each other. Because $U^t_{ik}(\mathbf{X}^{\sigma_k}_k)$ is an upper bound on the capacity consumption of resource i at time period t in stage k, if $\sum_{\ell=1}^k\sum_{q\in\mathcal{T}}U^q_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)<c_i$, then we have remaining capacity for resource i at any time period until the end of stage k. Thus, using $\mathcal{L}_j=\{i\in\mathcal{L}:a_{ij}=1\}$ to denote the set of resources used by product j, we have

$$\begin{aligned}\mathbb{P}\{G^t_{jk}=1\,|\,\mathbf{X}^{\sigma_k}_k\}\;&\overset{(c)}{\ge}\;\mathbb{P}\bigg\{\sum_{\ell=1}^k\sum_{q\in\mathcal{T}}U^q_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)<c_i\quad\forall\,i\in\mathcal{L}_j\,|\,\mathbf{X}^{\sigma_k}_k\bigg\}\\&=1-\mathbb{P}\bigg\{\sum_{\ell=1}^k\sum_{q\in\mathcal{T}}U^q_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)\ge c_i\text{ for some }i\in\mathcal{L}_j\,|\,\mathbf{X}^{\sigma_k}_k\bigg\}\overset{(d)}{\ge}1-\sum_{i\in\mathcal{L}_j}\mathbb{P}\bigg\{\sum_{\ell=1}^k\sum_{q\in\mathcal{T}}U^q_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)\ge c_i\,|\,\mathbf{X}^{\sigma_k}_k\bigg\},\end{aligned}\qquad(4)$$

where (c) holds because $\sum_{\ell=1}^k\sum_{q\in\mathcal{T}}U^q_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)<c_i$ for all $i\in\mathcal{L}_j$ implies that we have enough capacity to accept a request for product j at time period t in stage k and (d) is the union bound.

By the chain of inequalities above, an upper bound on $\mathbb{P}\{\sum_{\ell=1}^k\sum_{q\in\mathcal{T}}U^q_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)\ge c_i\,|\,\mathbf{X}^{\sigma_k}_k\}$ yields a lower bound on $\mathbb{P}\{G^t_{jk}=1\,|\,\mathbf{X}^{\sigma_k}_k\}$. In the rest of this section, we focus on upper bounding the former probability. The discussion so far in this section follows standard arguments, but our approach for upper bounding $\mathbb{P}\{\sum_{\ell=1}^{k}\sum_{q\in\mathcal{T}} U_{i\ell}^{q}(\mathbf{X}_{\ell}^{\sigma_\ell}) \geq c_i \,|\, \mathbf{X}_{k}^{\sigma_k}\}$ will use novel ideas.

Average Lipschitz Condition for the Resource Capacity Consumptions:

To upper bound $\mathbb{P}\{\sum_{\ell=1}^{k}\sum_{q\in\mathcal{T}} U_{i\ell}^{q}(\mathbf{X}_{\ell}^{\sigma_\ell}) \geq c_i \,|\, \mathbf{X}_{k}^{\sigma_k}\}$, we establish that the capacity consumption of a resource under our approximate policy satisfies the average Lipschitz condition. In this case, we use the average Lipschitz condition to upper bound the moment generating function of the capacity consumption of a resource, which, in turn, yields an upper bound on the probability that the capacity consumption of a resource under our approximate policy exceeds its capacity. We define the average Lipschitz condition. Let $Z_1,\ldots,Z_n$ be a sequence of random variables. These random variables all have support $\mathcal{Z}$ and can be dependent on each other. Consider the function $g:\mathcal{Z}^n \to \mathbb{R}$ mapping the sequence of random variables to reals. For all $\ell = 1,\ldots,n$ and $p,q\in\mathcal{Z}$, if we have $\big|\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell = p, Z_{\ell+1},\ldots,Z_n\} - \mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell = q, Z_{\ell+1},\ldots,Z_n\}\big| \leq \delta_\ell$ with probability one for some deterministic $\delta_\ell \in \mathbb{R}_+$, then we say that the function g satisfies the *average Lipschitz condition*. Average Lipschitz condition yields a bound on the moment generating function of $g(Z_1,\ldots,Z_n)$. If the function g satisfies the average Lipschitz condition, then we have $\mathbb{E}\{\exp(\lambda\, g(Z_1,\ldots,Z_n))\,|\,Z_\ell,\ldots,Z_n\} \leq \exp(\frac{\lambda^2}{2}\sum_{k=1}^{\ell-1}\delta_k^2) \times \exp(\lambda\,\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell,\ldots,Z_n\})$ for all $\lambda \geq 0$ and $\ell = 1,\ldots,n$. This result follows by iteratively applying Lemma 5.1 in Dubhashi and Panconesi (2009). We give a proof in Appendix A. To use this result, we define the sequence of random variables $Z_1,\ldots,Z_n$ and function g in our problem setting. Recalling the random variable $U_{ik}^{t}(\mathbf{s}_k^{\sigma_k}) = \sum_{j\in\mathcal{J}} a_{ij}\, A_{jk}^{t}(\mathbf{s}_k^{1})\, Y_{jk}^{t}(\mathbf{s}_k^{\sigma_k})$, we set $\overline{u}_{ik}^{t}(\mathbf{s}_k^{\sigma_k}) = \mathbb{E}\{U_{ik}^{t}(\mathbf{s}_k^{\sigma_k})\}$, so $\overline{u}_{ik}^{t}(\mathbf{s}_k^{\sigma_k})$ is an upper bound on the expected capacity consumption of resource i at time period t in stage k given that the history of the Markov chain over the aggregation duration is $\mathbf{s}_k^{\sigma_k}$. To bound the total expected capacity consumption of resource i over the first k stages, define the function $f_{ik}:\mathcal{X}^k \to \mathbb{R}$ as

$$f_{ik}(X_1,\ldots,X_k) = \sum_{v=1}^{k}\sum_{t\in\mathcal{T}} \overline{u}_{iv}^{t}(\mathbf{X}_v^{\sigma_v}). \tag{5}$$

In the next lemma, for $k \leq K-\theta$, we show that the function f_{ik} satisfies the average Lipschitz condition. Verifying the average Lipschitz condition only for $k \leq K-\theta$ will be enough for us.

Lemma 5.1 (Average Lipschitz) *Letting* $\delta_{i\ell} = \sum_{v=1}^{\ell}\sum_{t\in\mathcal{T}} \frac{4}{\epsilon}(1-\alpha)^{\ell-v}\,\mathbb{E}\{\overline{u}_{iv}^{t}(\mathbf{X}_v^{\sigma_v})\}$, *for all* $k \leq K-\theta$, $\ell \leq k$, $i\in\mathcal{L}$ *and* $p,q\in\mathcal{X}$, *with probability one, we have*

$$\Big|\mathbb{E}\{f_{ik}(X_1,\ldots,X_k)\,|\,X_\ell = p, X_{\ell+1},\ldots,X_k\} - \mathbb{E}\{f_{ik}(X_1,\ldots,X_k)\,|\,X_\ell = q, X_{\ell+1},\ldots,X_k\}\Big| \leq \delta_{i\ell}.$$

Proof: To express the left side of the inequality in the lemma, we define $\Delta_{i\ell v}(p,q,X_{\ell+1},\ldots,X_k) = \sum_{t\in\mathcal{T}}\big[\mathbb{E}\{\overline{u}_{iv}^{t}(\mathbf{X}_v^{\sigma_v})\,|\,X_\ell = p, X_{\ell+1},\ldots,X_k\} - \mathbb{E}\{\overline{u}_{iv}^{t}(\mathbf{X}_v^{\sigma_v})\,|\,X_\ell = q, X_{\ell+1},\ldots,X_k\}\big]$ for all $v = 1,\ldots,k$,

$\ell = 1,\ldots,k$, $i \in \mathcal{L}$ and $p,q \in \mathcal{X}$. In this case, by (5), the left side of the inequality in the lemma is given by $|\sum_{v=1}^{k} \Delta_{i\ell v}(p,q,X_{\ell+1},\ldots,X_k)|$. We upper bound $\Delta_{i\ell v}(p,q,X_{\ell+1},\ldots,X_k)$. Consider the case $v \ge \ell+1$. Because $v \le k \le K-\theta$, we have $\sigma_v = 1$. Furthermore, noting that $v \ge \ell+1$, we have $\mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v) \,|\, X_\ell = p, X_{\ell+1},\ldots,X_k\} = \overline{u}^t_{iv}(\mathbf{X}^1_v) = \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v) \,|\, X_\ell = q, X_{\ell+1},\ldots,X_k\}$ for all $t \in \mathcal{T}$, which implies that $\Delta_{i\ell v}(p,q,X_{\ell+1},\ldots,X_k) = 0$. Consider the case $v \le \ell$. Because the total variation distance of the probability law for the Markov chain is upper bounded by $1-\alpha$, by a standard result in mixing times of Markov chains, we get $|\mathbb{P}\{X_\ell = p \,|\, X_v = s\} - \mathbb{P}\{X_\ell = p\}| \le 2\,(1-\alpha)^{\ell - v}$ for all $p,s \in \mathcal{X}$; see Theorem 4.6 in Levin and Peres (2017). Because $v \le \ell$, conditional on X_ℓ, X_v is independent of $X_{\ell+1},\ldots,X_k$, Thus, we get $\mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^1_v) \,|\, X_\ell = p, X_{\ell+1},\ldots,X_k\} = \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^1_v) \,|\, X_\ell = p\}$. In this case, using the fact that $\sigma_v = 1$ for $v \le K-\theta$, we have

$$
\begin{aligned}
\Big|\Delta_{i\ell v}(p,q,X_{\ell+1},\ldots,X_k)\Big| &= \Big|\sum_{t\in\mathcal{T}} \big[\mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^1_v) \,|\, X_\ell = p\} - \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^1_v) \,|\, X_\ell = q\}\big]\Big| \\
&\le \sum_{t\in\mathcal{T}} \Big|\mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^1_v) \,|\, X_\ell = p\} - \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^1_v) \,|\, X_\ell = q\}\Big| \\
&= \sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{X}} \overline{u}^t_{iv}(s) \Big|\mathbb{P}\{X_v = s \,|\, X_\ell = p\} - \mathbb{P}\{X_v = s \,|\, X_\ell = q\}\Big| \\
&\overset{(a)}{=} \sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{X}} \overline{u}^t_{iv}(s)\,\mathbb{P}\{X_v = s\} \left|\frac{\mathbb{P}\{X_\ell = p \,|\, X_v = s\}}{\mathbb{P}\{X_\ell = p\}} - \frac{\mathbb{P}\{X_\ell = q \,|\, X_v = s\}}{\mathbb{P}\{X_\ell = q\}}\right| \\
&\overset{(b)}{\le} \sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{X}} \overline{u}^t_{iv}(s)\,\mathbb{P}\{X_v = s\} \left\{\left|\frac{\mathbb{P}\{X_\ell = p \,|\, X_v = s\}}{\mathbb{P}\{X_\ell = p\}} - 1\right| + \left|\frac{\mathbb{P}\{X_\ell = q \,|\, X_v = s\}}{\mathbb{P}\{X_\ell = q\}} - 1\right|\right\} \\
&\overset{(c)}{\le} \sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{X}} \overline{u}^t_{iv}(s)\,\mathbb{P}\{X_v = s\}\,\frac{4}{\epsilon}\,(1-\alpha)^{\ell-v} \;=\; \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^1_v)\}\,\frac{4}{\epsilon}\,(1-\alpha)^{\ell-v},
\end{aligned} \tag{6}
$$

where (a) uses the Bayes rule, (b) follows by the triangle inequality and (c) holds because the Markov chain satisfies the assumption that $\mathbb{P}\{X_\ell = p\} \ge \epsilon$ for all $p \in \mathcal{X}$.

The left side of the inequality in the lemma is $|\sum_{v=1}^{k} \Delta_{i\ell v}(p,q,X_{\ell+1},\ldots,X_k)|$, which we upper bound by $\sum_{v=1}^{k} |\Delta_{i\ell v}(p,q,X_{\ell+1},\ldots,X_k)| = \sum_{v=1}^{\ell} |\Delta_{i\ell v}(p,q,X_{\ell+1},\ldots,X_k)|$ and use (6). ■

By Lemma 5.1, for all $k \le K-\theta$, f_{ik} satisfies the average Lipschitz condition, in which case, it follows that $\mathbb{E}\{\exp(\lambda\, f_{ik}(X_1,\ldots,X_k)) \,|\, X_k\} \le \exp(\frac{\lambda^2}{2}\sum_{\ell=1}^{k-1}(\delta_{i\ell})^2) \times \exp(\lambda\,\mathbb{E}\{f_{ik}(X_1,\ldots,X_k) \,|\, X_k\})$, where $\delta_{i\ell}$ is as in the lemma. Because $\exp(x)$ is convex in x, by Jensen inequality, we have $\mathbb{E}\{\exp(\lambda\, f_{ik}(X_1,\ldots,X_k)) \,|\, X_k\} \ge \exp(\lambda\,\mathbb{E}\{f_{ik}(X_1,\ldots,X_k) \,|\, X_k\})$. Thus, by Lemma 5.1, we also have the reverse inequality with a multiplicative margin of $\exp(\frac{\lambda^2}{2}\sum_{\ell=1}^{k-1}(\delta_{i\ell})^2)$. We proceed to giving an upper bound on $\delta_{i\ell}$ using only the problem data. Conditional on the history of the Markov chain over the aggregation duration σ_k, the random variables $A^t_{jk}(\mathbf{X}^1_k)$ and $Y^t_{jk}(\mathbf{X}^{\sigma_k}_k)$ are independent of each other. Noting that $U^t_{ik}(\mathbf{X}^{\sigma_k}_k) = \sum_{j\in\mathcal{J}} a_{ij}\, A^t_{jk}(\mathbf{X}^1_k)\, Y^t_{jk}(\mathbf{X}^{\sigma_k}_k)$, we get $\mathbb{E}\{U^t_{ik}(\mathbf{X}^{\sigma_k}_k) \,|\, \mathbf{X}^{\sigma_k}_k\} = \sum_{j\in\mathcal{J}} a_{ij}\,\mathbb{E}\{A^t_{jk}(\mathbf{X}^1_k) \,|\, \mathbf{X}^{\sigma_k}_k\}\,\mathbb{E}\{Y^t_{jk}(\mathbf{X}^{\sigma_k}_k) \,|\, \mathbf{X}^{\sigma_k}_k\}$. In this case, because $\mathbb{E}\{A^t_{jk}(\mathbf{X}^1_k) \,|\, \mathbf{X}^{\sigma_k}_k\} = \lambda^t_{jk}(\mathbf{X}^1_k)$ and $\mathbb{E}\{Y^t_{jk}(\mathbf{X}^{\sigma_k}_k) \,|\, \mathbf{X}^{\sigma_k}_k\} = \frac{\gamma}{\lambda^t_{jk}(\mathbf{X}^1_k)}\,\mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^{\theta}_k) \,|\, \mathbf{X}^{\sigma_k}_k\}$ by the definitions of $A^t_{jk}(\mathbf{X}^1_k)$ and $Y^t_{jk}(\mathbf{X}^{\sigma_k}_k)$, we obtain $\mathbb{E}\{U^t_{ik}(\mathbf{X}^{\sigma_k}_k)\,|\,\mathbf{X}^{\sigma_k}_k\} = \sum_{j\in\mathcal{J}} \gamma\, a_{ij}\, \mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\,|\,\mathbf{X}^{\sigma_k}_k\}$, In the next lemma, we use the last equality to give an upper bound on $\sum_{\ell=1}^k (\delta_{i\ell})^2$ that uses only the problem data.

Lemma 5.2 (Lipschitz Bound) *Letting* $\delta_{i\ell} = \sum_{v=1}^{\ell} \sum_{t\in\mathcal{T}} \frac{4}{\epsilon}(1-\alpha)^{\ell-v}\, \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\}$, *for all* $k \le K-\theta$ *and* $i\in\mathcal{L}$, *we have* $\sum_{\ell=1}^k (\delta_{i\ell})^2 \le 16\frac{T}{\epsilon^2\alpha^2}\min\{(K-\theta)\,T, c_i\}$.

Proof: We have $\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v) = \mathbb{E}\{U^t_{iv}(\mathbf{X}^{\sigma_v}_v)\,|\,\mathbf{X}^{\sigma_v}_v\} = \sum_{j\in\mathcal{J}} \gamma\, a_{ij}\, \mathbb{E}\{\overline{y}^t_{jv}(\mathbf{X}^\theta_v)\,|\,\mathbf{X}^{\sigma_v}_v\}$, where the second equality is by the discussion just before the lemma. Taking the expectations of both sides and using the tower property of conditional expectation, we get $\mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\} = \sum_{j\in\mathcal{J}} \gamma\, a_{ij}\, \mathbb{E}\{\overline{y}^t_{jv}(\mathbf{X}^\theta_v)\}$. By the first constraint in (2), $\sum_{v=1}^k \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{\overline{y}^t_{jv}(\mathbf{X}^\theta_v)\,|\,\mathbf{X}^\theta_k\} \le c_i$ with probability one, so taking the expectations of both sides of the inequality and using the tower property of conditional expectation, we obtain $c_i \ge \gamma\, c_i \ge \sum_{v=1}^k \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \gamma\, a_{ij}\, \mathbb{E}\{\overline{y}^t_{jv}(\mathbf{X}^\theta_v)\} = \sum_{v=1}^k \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\}$. By the definition of $U^t_{ik}(\mathbf{X}^{\sigma_k}_k)$, we have $U^t_{ik}(\mathbf{X}^{\sigma_k}_k) \le 1$ with probability one, which implies that $\mathbb{E}\{\overline{u}^t_{ik}(\mathbf{X}^{\sigma_k}_k)\} \le 1$, so $\sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}^t_{ik}(\mathbf{X}^{\sigma_k}_k)\} \le T$. Thus, we get $\delta_{i\ell} \le 4\frac{T}{\epsilon}\sum_{v=1}^{\ell}(1-\alpha)^{\ell-v} \le 4\frac{T}{\epsilon\,\alpha}$, where the last inequality uses an upper bound on the sum of the geometric series. On the other hand, we have $\sum_{\ell=1}^k \delta_{i\ell} = \frac{4}{\epsilon}\sum_{\ell=1}^k \sum_{v=1}^{\ell} \sum_{t\in\mathcal{T}} (1-\alpha)^{\ell-v}\mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\}$. Interchanging the order of the sums, the last expression is given by $\frac{4}{\epsilon}\sum_{v=1}^k \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\} \sum_{\ell=v}^k (1-\alpha)^{\ell-v} \le \frac{4}{\epsilon\,\alpha}\sum_{v=1}^k \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\}$. Thus, we have $\sum_{\ell=1}^k \delta_{i\ell} \le \frac{4}{\epsilon\,\alpha}\sum_{v=1}^k \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\}$. Noting that $\delta_{i\ell} \le 4\frac{T}{\epsilon\,\alpha}$, we get

$$\sum_{\ell=1}^k (\delta_{i\ell})^2 \;\le\; 4\frac{T}{\epsilon\,\alpha}\sum_{\ell=1}^k \delta_{i\ell} \;\le\; 16\frac{T}{\epsilon^2\,\alpha^2}\sum_{v=1}^k \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\}.$$

Because $k \le K-\theta$ and $|\mathcal{T}| = T$ and $\mathbb{E}\{\overline{u}^t_{iv}(\mathbf{X}^{\sigma_v}_v)\} \le 1$, the last sum is at most $(K-\theta)\,T$. By the discussion at the beginning of the proof, the last sum is also at most c_i. ■

Noting the discussion that follows (3), we establish the performance guarantee for our approximate policy by upper bounding the probability on the right side of (4). For notational brevity, we define the random variable $W_{ik} = \sum_{\ell=1}^k \sum_{t\in\mathcal{T}} U^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_t)$, so that the probability on the right side of (4) is given by $\mathbb{P}\{W_{ik} \ge c_i\,|\,\mathbf{X}^{\sigma_k}_k\}$. We upper bound this probability. By the first constraint in (2), $\sum_{\ell=1}^k \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^\theta_k\} \le c_i$ with probability one. Because $\sigma_k \le \theta$, the filtration generated by $\mathbf{X}^{\sigma_k}_k$ is a subset of the one generated by $\mathbf{X}^\theta$, so taking expectations in the last inequality conditional on $\mathbf{X}^{\sigma_k}_k$ and using the tower property of conditional expectation, it follows that $\sum_{\ell=1}^k \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^{\sigma_k}_k\} \le c_i$. Also, considering stages ℓ and k with $\ell \le k$, we have $\ell - \theta + 1 \le \ell - \sigma_\ell + 1 \le k - \sigma_k + 1$, which implies that conditional on $\mathbf{X}^{\sigma_\ell}_\ell$, the random vectors $\mathbf{X}^\theta_\ell$ and $\mathbf{X}^{\sigma_k}_k$ are independent of each other. Thus, we have $\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^{\sigma_\ell}_\ell\} = \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^{\sigma_\ell}_\ell, \mathbf{X}^{\sigma_k}_k\}$. Because we have $\mathbb{E}\{U^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)\,|\,\mathbf{X}^{\sigma_\ell}_\ell\} = \sum_{j\in\mathcal{J}} \gamma\, a_{ij}\, \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^{\sigma_\ell}_\ell\}$ by the discussion just before Lemma 5.2 and $\overline{u}^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell) = \mathbb{E}\{U^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)\,|\,\mathbf{X}^{\sigma_\ell}_\ell\}$ by the definition of $\overline{u}^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)$, we obtain the identity

$\overline{u}^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell) = \sum_{j\in\mathcal{J}} \gamma\, a_{ij}\, \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^{\sigma_\ell}_\ell, \mathbf{X}^{\sigma_k}_k\}$. Taking the expectations of both sides of the last equality conditional on $\mathbf{X}^{\sigma_k}_k$ and using the tower property of conditional expectation, for all $\ell \le k$, we get $\mathbb{E}\{\overline{u}^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)\,|\,\mathbf{X}^{\sigma_k}_k\} = \sum_{j\in\mathcal{J}} \gamma\, a_{ij}\, \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^{\sigma_k}_k\}$. Adding this equality over all $t\in\mathcal{T}$ and $\ell = 1,\ldots,k$ yields $\sum_{\ell=1}^k \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)\,|\,\mathbf{X}^{\sigma_k}_k\} = \sum_{\ell=1}^k \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} \gamma\, a_{ij}\, \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^\theta_\ell)\,|\,\mathbf{X}^{\sigma_k}_k\} \le \gamma\, c_i$, where the last inequality uses the discussion at the beginning of this paragraph. In the next lemma, we build the last inequality to give an upper bound on the probability $\mathbb{P}\{W_{ik} \ge c_i\,|\,\mathbf{X}^{\sigma_k}_k\}$, which gives an upper bound on the probability that we do not have remaining capacity for a resource at the end of stage k. Using this upper bound on the right side of (4) ultimately yields the performance guarantee for our approximate policy. In the lemma below, recall that we set $\nu^\theta = 32\, \frac{T}{\epsilon^2\alpha^2}\, \min\{(K-\theta)\,T, c_{\min}\}$ in the performance guarantee for our approximate policy in the previous section and γ is the tuning parameter in our approximate policy.

Lemma 5.3 (Resource Unavailability Probability) *Setting* $W_{ik} = \sum_{\ell=1}^k \sum_{t\in\mathcal{T}} U^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)$*, for all* $\gamma\in(0,1)$*,* $k\in\mathcal{K}$ *and* $i\in\mathcal{L}$*, with probability one, we have*

$$\mathbb{P}\{W_{ik} \ge c_i\,|\,\mathbf{X}^{\sigma_k}_k\} \;\le\; \exp\left(-\,(1-\gamma)^2\, \frac{c^2_{\min}}{4\,(\nu^\theta + c_{\min})}\right).$$

Proof: We define $V_{ik} = \sum_{\ell=1}^k \sum_{t\in\mathcal{T}} \overline{u}^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)$, so that $V_{ik} = \mathbb{E}\{W_{ik}\,|\,X_1,\ldots,X_k\}$, where we use the fact that $\overline{u}^t_{ik}(\mathbf{X}^{\sigma_k}_k) = \mathbb{E}\{U^t_{ik}(\mathbf{X}^{\sigma_k}_k)\,|\,\mathbf{X}^{\sigma_k}_k\}$. Given $X_1,\ldots,X_k$, for all $\ell = 1,\ldots,k$, the random variable $U^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell)$ is Bernoulli. Also, given $X_1,\ldots,X_k$, the random variables $\{U^t_{i\ell}(\mathbf{X}^{\sigma_\ell}_\ell) : t\in\mathcal{T},\ \ell = 1,\ldots,k\}$ are independent of each other. By a simple lemma, given as Lemma B.1 in Appendix B, if Z is a Bernoulli random variable with mean μ, then $\mathbb{E}\{\exp(\lambda\, Z)\} \le \exp((\lambda^2+\lambda)\,\mu)$ for all $\lambda\in[0,1]$. Thus, we get $\mathbb{E}\{\exp(\lambda\, W_{ik})\,|\,X_1,\ldots,X_k\} \le \exp((\lambda^2+\lambda)\,V_{ik})$, where we use the fact that W_{ik}, given $X_1,\ldots,X_k$, is a sum of independent Bernoullis and $\mathbb{E}\{W_{ik}\,|\,X_1,\ldots,X_k\} = V_{ik}$. Taking the expectations in the last inequality conditional on $\mathbf{X}^{\sigma_k}_k$, by the tower property of conditional expectation, we get $\mathbb{E}\{\exp(\lambda\, W_{ik})\,|\,\mathbf{X}^{\sigma_k}_k\} \le \mathbb{E}\{\exp((\lambda^2+\lambda)\,V_{ik})\,|\,\mathbf{X}^{\sigma_k}_k\}$. By the Markov inequality, we have $\mathbb{P}\{W_{ik} \ge c_i\,|\,\mathbf{X}^{\sigma_k}_k\} = \mathbb{P}\{\exp(\lambda\, W_{ik}) \ge \exp(\lambda\, c_i)\,|\,\mathbf{X}^{\sigma_k}_k\} \le \mathbb{E}\{\exp(\lambda(W_{ik} - c_i))\,|\,\mathbf{X}^{\sigma_k}_k\}$ for all $\lambda \ge 0$. First, consider the case $k \le K - \theta$, so $\sigma_k = 1$. By the discussion so far, for all $\lambda\in[0,1]$, we have

$$\begin{aligned}
\mathbb{P}\{W_{ik} \ge c_i\,|\,\mathbf{X}^{\sigma_k}_k\} \;&\le\; \mathbb{E}\{\exp(\lambda(W_{ik}-c_i))\,|\,\mathbf{X}^{\sigma_k}_k\} \;\le\; \mathbb{E}\{\exp((\lambda^2+\lambda)\,V_{ik} - \lambda\, c_i)\,|\,\mathbf{X}^{\sigma_k}_k\} \\
&\overset{(a)}{\le}\; \exp\left(\frac{1}{2}(\lambda^2+\lambda)^2 \sum_{\ell=1}^{k-1} (\delta_{i\ell})^2\right) \exp\left(\mathbb{E}\{(\lambda^2+\lambda)\,V_{ik} - \lambda\, c_i)\,|\,\mathbf{X}^1_k\}\right) \\
&\overset{(b)}{\le}\; \exp\left\{\left(32 \frac{T}{\epsilon^2\alpha^2}\, \min\{(K-\theta)\,T, c_i\} + c_i\right)\lambda^2 - (1-\gamma)\,\lambda\, c_i\right\}.
\end{aligned} \tag{7}$$

In (7), (a) holds by noting that the definition of V_{ik} matches that of $f_{ik}(X_1,\ldots,X_k)$ in (5) and using the discussion just after Lemma 5.1, whereas (b) holds by using the upper bound on $\sum_{\ell=1}^k (\delta_{i\ell})^2$ given in Lemma 5.2, noting that $\sum_{\ell=1}^k \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{u}_{i\ell}^t(\mathbf{X}_\ell^{\sigma_\ell}) \,|\, \mathbf{X}_k^{\sigma_k}\} \le \gamma\, c_i \le c_i$ by the discussion just before the lemma and the left side of this inequality corresponds to $\mathbb{E}\{V_{ik} \,|\, \mathbf{X}_k^1\}$, as well as using the fact that $\lambda \in [0,1]$, so $(\lambda^2+\lambda)^2 \le (2\lambda)^2$. Second, consider the case $k \ge K-\theta+1$ so $\sigma_k = k-(K-\theta)$. The chain of inequalities in (7) holds for $k \ge K-\theta+1$ as well. In particular, the first two inequalities in (7) did not use the fact that $k \le K-\theta$. Because $k \ge K-\theta+1$, we have $\mathbf{X}_k^{\sigma_k} = (X_{K-\theta+1},\ldots,X_k)$, in which case, we can compute $\mathbb{E}\{\exp((\lambda^2+\lambda)\,V_{ik}) \,|\, \mathbf{X}_k^{\sigma_k}\}$ by multiplying $\mathbb{E}\{\exp((\lambda^2+\lambda)\,V_{i,K-\theta}) \,|\, \mathbf{X}_{K-\theta+1}^1\}$ with $\exp((\lambda^2+\lambda)\sum_{\ell=K-\theta+1}^k \sum_{t\in\mathcal{T}} \overline{u}_{i\ell}^t(\mathbf{X}_\ell^{\sigma_\ell}))$, where we use the fact that knowing $\mathbf{X}_k^{\sigma_k}$ implies knowing $\mathbf{X}_\ell^{\sigma_\ell}$ for all $\ell = K-\theta+1,\ldots,k$. Noting that $V_{i,K-\theta} = \sum_{\ell=1}^{K-\theta}\sum_{t\in\mathcal{T}} \overline{u}_{i\ell}^t(\mathbf{X}_\ell^1)$, we can define $\upsilon_{i,K-\theta+1}(\mathbf{X}_{K-\theta+1}^1) = 0$ and follow precisely the same reasoning in the proof of Lemma 5.1 to show that $f_{i,K-\theta+1}(X_1,\ldots,X_{K-\theta+1}) = \sum_{\ell=1}^{K-\theta}\sum_{t\in\mathcal{T}} \overline{u}_{i\ell}^t(\mathbf{X}_\ell^1) + \upsilon_{i,K-\theta+1}(\mathbf{X}_{K-\theta+1}^1)$ satisfies the average Lipschitz condition with $\delta_{i\ell}$ as defined in this lemma. Therefore, we obtain the chain of inequalities

$$\begin{aligned}
\mathbb{E}\{\exp((\lambda^2+\lambda)\,V_{ik}) \,|\, \mathbf{X}_k^{\sigma_k}\} &= \mathbb{E}\{\exp((\lambda^2+\lambda)\,V_{i,K-\theta}) \,|\, \mathbf{X}_{K-\theta+1}^1\} \exp\left((\lambda^2+\lambda)\sum_{\ell=K-\theta+1}^k \sum_{t\in\mathcal{T}} \overline{u}_{i\ell}^t(\mathbf{X}_\ell^{\sigma_\ell})\right) \\
&\le \exp\left(\frac{1}{2}\,(\lambda^2+\lambda)^2 \sum_{\ell=1}^{K-\theta}(\delta_{i\ell})^2\right) \exp\Big(\mathbb{E}\{(\lambda^2+\lambda)\,V_{i,K-\theta} \,|\, \mathbf{X}_{K-\theta+1}^1\}\Big) \exp\left((\lambda^2+\lambda)\sum_{\ell=K-\theta+1}^k \sum_{t\in\mathcal{T}} \overline{u}_{i\ell}^t(\mathbf{X}_\ell^{\sigma_\ell})\right) \\
&= \exp\left(\frac{1}{2}\,(\lambda^2+\lambda)^2 \sum_{\ell=1}^{K-\theta}(\delta_{i\ell})^2\right) \exp\Big(\mathbb{E}\{(\lambda^2+\lambda)\,V_{ik} \,|\, \mathbf{X}_k^{\sigma_k}\}\Big).
\end{aligned}$$

In this case, (7) holds with no modification. Thus, (7) holds for all $\lambda \in [0,1]$ and $k \in \mathcal{K}$. Choosing $\lambda = \frac{(1-\gamma)\,c_i/2}{32\frac{T}{\epsilon^2\alpha^2}\,\min\{(K-\theta)\,T, c_i\} + c_i}$, the right side of (7) evaluates to $\exp\left(-\frac{(1-\gamma)^2\,c_i^2/4}{32\frac{T}{\epsilon^2\alpha^2}\,\min\{(K-\theta)\,T, c_i\}+c_i}\right)$.

It is simple to check that the last expression in the previous paragraph is decreasing in c_i, so replacing c_i with $c_{\min}$, this expression is upper bounded by $\exp\left(-(1-\gamma)^2\,\frac{c_{\min}^2}{4\,(\nu^\theta+c_{\min})}\right)$. ■

By (4), we have $\mathbb{P}\{G_{jk}^t = 1 \,|\, \mathbf{X}_k^{\sigma_k}\} \ge 1 - \sum_{i\in\mathcal{L}_j} \mathbb{P}\{W_{ik} \ge c_i \,|\, \mathbf{X}_k^{\sigma_k}\}$. By Lemma 5.3, we have a lower bound on the right side of the inequality. Using this observation, we give a proof for Theorem 4.1.

Proof of Theorem 4.1:

Setting the tuning parameter as $\gamma = 1 - \sqrt{\frac{4\,(\nu^\theta+c_{\min})\,\log c_{\min}}{c_{\min}^2}}$, we get $\mathbb{P}\{W_{ik} \ge c_i \,|\, \mathbf{X}_k^{\sigma_k}\} \le \frac{1}{c_{\min}}$ by Lemma 5.3, so $\mathbb{P}\{G_{jk}^t = 1 \,|\, \mathbf{X}_k^{\sigma_k}\} \ge 1 - \sum_{i\in\mathcal{L}_j}\mathbb{P}\{W_{ik}\ge c_i \,|\, \mathbf{X}_k^{\sigma_k}\} \ge 1 - \frac{L}{c_{\min}}$. Thus, by (3), we get

$$\frac{\mathsf{APX}^\theta}{Z_{\mathsf{LP}}^\theta} \;\ge\; \left(1 - \sqrt{\frac{4\,(\nu^\theta+c_{\min})\,\log c_{\min}}{c_{\min}^2}}\right)\left(1 - \frac{L}{c_{\min}}\right) \;\ge\; 1 - \sqrt{\frac{4\,(\nu^\theta+c_{\min})\log c_{\min}}{c_{\min}^2}} - \frac{L}{c_{\min}},$$

where the first inequality holds because $\mathbb{P}\{G_{jk}^t = 1 \,|\, \mathbf{X}_k^{\sigma_k}\} \ge 1 - \frac{L}{c_{\min}}$, so we can set $\beta = 1 - \frac{L}{c_{\min}}$ in (3) and the second inequality holds by arranging the terms.

The first inequality in the theorem holds because we have $Z_{\mathsf{LP}}^\theta \ge \mathsf{opt}$, in which case, $\frac{\mathsf{APX}^\theta}{\mathsf{opt}}$ is an upper bound on the left side of the chain of inequalities above. ■

6. Fluid Approximations with Varying Histories

In the fluid approximation in (2), we always use the history of the Markov chain over past θ stages in each stage in the selling horizon. We can construct a fluid approximation by using the history of the Markov chain over different numbers of stages in different stages in the selling horizon. In this way, we can get tighter upper bounds on the optimal total expected revenue by using a fluid approximation with fewer decision variables and constraints than the one in (2). Also, we can obtain a policy that provides the same performance guarantee as our approximate policy by using a fluid approximation with fewer decision variables and constraints. In particular, we use the history of the Markov chain over past η_k stages in stage k. We assume that if the stages ℓ and k satisfy $\ell \le k$, then $\ell - \eta_\ell \le k - \eta_k$. In other words, if stage k comes after stage ℓ, then the history of the Markov chain in stage k cannot reach out to a stage earlier than the history of the Markov chain in stage ℓ. We use the decision variables $\boldsymbol{y} = (y^t_{jk}(\mathbf{s}^{\eta_k}_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^{\eta_k}_k \in \mathcal{X}^{\eta_k})$ with the same interpretation as in (2), but we use the history of the Markov chain over past η_k stages in stage k. Using the vector $\boldsymbol{\eta} = (\eta_1, \dots, \eta_K) \in \mathbb{Z}^K_+$, we consider the linear program

$$Z^{\boldsymbol{\eta}}_{\mathsf{LP}} = \max_{\boldsymbol{y} \in \mathbb{R}_+^{|\mathcal{J}|T \sum_{k\in\mathcal{K}} |\mathcal{X}|^{\eta_k}}} \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} f_j\, \mathbb{E}\{y^t_{jk}(\mathbf{X}^{\eta_k}_k)\} \tag{8}$$

$$\text{st} \quad \sum_{\ell=1}^{k} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{y^t_{j\ell}(\mathbf{X}^{\eta_\ell}_\ell) \,|\, \mathbf{X}^{\eta_k}_k = \mathbf{s}^{\eta_k}_k\} \le c_i \quad \forall\, i \in \mathcal{L},\ k \in \mathcal{K},\ \mathbf{s}^{\eta_k}_k \in \mathcal{X}^{\eta_k}$$

$$y^t_{jk}(\mathbf{s}^{\eta_k}_k) \le \lambda^t_{jk}(\mathbf{s}^1_k) \quad \forall\, j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^{\eta_k}_k \in \mathcal{X}^{\eta_k}.$$

In the problem above, we use $\mathbf{X}^{\eta_k}_k = (X_{k-\eta_k+1}, \dots, X_k)$ to capture the history of the Markov chain over past η_k stages in stage k. The interpretation of the decision variables and constraints in (8) is identical to that in (2). Under the assumption that $\ell - \eta_\ell \le k - \eta_k$ for $\ell \le k$, we can show that the optimal objective value of (8) is an upper bound on the optimal total expected revenue. This result is the analogue of Theorem 3.1. The assumption that $\ell - \eta_\ell \le k - \eta_k$ for $\ell \le k$ is necessary for (8) to provide an upper bound on the optimal total expected revenue. In Appendix C, we demonstrate that if this assumption does not hold, then the optimal objective value of (8) may not be an upper bound. We use $\Theta = \{\boldsymbol{\eta} \in \mathbb{Z}^K_+ : k - \eta_k \le k + 1 - \eta_{k+1}\ \ \forall\, k = 1, \dots, K-1\}$ to capture the set of histories ensuring that the optimal objective value of (8) is an upper bound on the optimal total expected revenue. If $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ satisfy $\mu_k \ge \nu_k$ for all $k \in \mathcal{K}$, then the histories in $\boldsymbol{\mu}$ are longer than those in $\boldsymbol{\nu}$, in which case, we can show that using the histories $\boldsymbol{\mu}$ in (8) provides a tighter upper bound than using the histories in $\boldsymbol{\nu}$, so $Z^{\boldsymbol{\mu}}_{\mathsf{LP}} \le Z^{\boldsymbol{\nu}}_{\mathsf{LP}}$. This result is the analogue of Theorem 3.2.

In the next theorem, we collect the results discussed in the previous paragraph. The proof follows from arguments similar to those for Theorems 3.1 and 3.2, so we defer it to Appendix D.

Theorem 6.1 (Upper Bound with Varying Histories) *Letting* opt *be the optimal total expected revenue and* $Z_{\mathsf{LP}}^{\boldsymbol{\eta}}$ *be the optimal objective value of problem* (8) *as a function of* $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_K)$*, if* $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ *satisfy* $\mu_k \geq \nu_k$ *for all* $k \in \mathcal{K}$*, then we have* $Z_{\mathsf{LP}}^{\boldsymbol{\nu}} \geq Z_{\mathsf{LP}}^{\boldsymbol{\mu}} \geq \mathsf{opt}$.

In the proof of Theorem 6.1, the assumption that $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ allows using the tower property of conditional expectation as is done just before the proofs of Theorems 3.1 and 3.2. There are several uses of problem (8). First, by using different numbers of stages in the histories in different stages, we can obtain strong upper bounds on the optimal total expected revenue without using as many decision variables and constraints that would have been necessary with the same number of stages in the histories in all stages. Second, considering $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ with $\mu_k \geq \nu_k$ for all $k \in \mathcal{K}$, we can show that if we solve the dual of problem (8) with $\boldsymbol{\eta} = \boldsymbol{\nu}$, then we can use this optimal solution to construct a feasible solution to the dual of problem (8) with $\boldsymbol{\eta} = \boldsymbol{\mu}$. We give this result in Appendix E. When solving problem (8) with $\boldsymbol{\eta} = \boldsymbol{\mu}$, we can start from the feasible dual solution. In this way, we can use the dual solution obtained by using shorter histories to solve (8) with longer histories. Third, we can use problem (8) to develop policies with performance guarantees, as discussed next.

In our approximate policy, we used aggregation durations given by $\sigma_k = 1$ for $k \leq K - \theta$, whereas $\sigma_k = k - (K - \theta)$ for $k \geq K - \theta + 1$. Letting $\overline{\boldsymbol{y}} = (\overline{y}_{jk}^t(\mathbf{s}_k^\theta) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}_k^\theta \in \mathcal{X}^\theta)$ be an optimal solution to problem (2), if the history of the Markov chain in stage k is $\mathbf{s}_k^\theta$, then we accept a request for product j at time period t in stage k with probability $\gamma\, \mathbb{E}\{\overline{y}_{jk}^t(\mathbf{X}_k^\theta) \,|\, \mathbf{X}_k^{\sigma_k} = \mathbf{s}_k^{\sigma_k}\} / \lambda_{jk}^t(\mathbf{s}_k^1)$. The last probability has a conditional expectation. It turns out we can "embed" the aggregation durations into problem (8) to come up with an approximate policy that does not require a conditional expectation in the probability of accepting a product request. Fixing some $\theta = 1, \ldots, K$, we define $\overline{\boldsymbol{\eta}} = (\overline{\eta}_1, \ldots, \overline{\eta}_K)$ as $\overline{\eta}_k = 1$ for $k \leq K - \theta$, whereas $\overline{\eta}_k = k - (K - \theta)$ for $k \geq K - \theta + 1$. We solve problem (8) with $\boldsymbol{\eta} = \overline{\boldsymbol{\eta}}$ to obtain an optimal solution $\overline{\boldsymbol{y}} = (\overline{y}_{jk}^t(\mathbf{s}_k^{\overline{\eta}_k}) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}_k^{\overline{\eta}_k} \in \mathcal{X}^{\overline{\eta}_k})$. If the history of the Markov chain over past $\overline{\eta}_k$ stages in stage k is $\mathbf{s}_k^{\overline{\eta}_k}$, then we accept a request for product j at time period t in stage k with probability $\gamma\, \overline{y}_{jk}^t(\mathbf{s}_k^{\overline{\eta}_k}) / \lambda_{jk}^t(\mathbf{s}_k^1)$.

We refer to the policy in the previous paragraph as the approximate policy with varying histories. In the next theorem, we give a performance guarantee for this policy.

Theorem 6.2 (Approximate Policy with Varying Histories) *Letting* $\mathsf{APX}^{\overline{\boldsymbol{\eta}}}$ *be the total expected revenue of the approximate policy with varying histories as a function of the histories* $\overline{\boldsymbol{\eta}}$*, there exist a choice of the tuning parameter* γ *such that we have*

$$\frac{\mathsf{APX}^{\overline{\boldsymbol{\eta}}}}{\mathsf{opt}} \;\geq\; \frac{\mathsf{APX}^{\overline{\boldsymbol{\eta}}}}{Z_{\mathsf{LP}}^{\overline{\boldsymbol{\eta}}}} \;\geq\; 1 - \sqrt{\frac{4\,(\nu^\theta + c_{\min}) \log c_{\min}}{c_{\min}^2}} - \frac{L}{c_{\min}}.$$

We give the proof of the theorem in Appendix F. In Theorem 6.2, ν^θ is as defined just before Theorem 4.1. Thus, comparing Theorem 4.1 with the theorem above, the approximate policy with varying histories obtained from problem (8) has the same performance guarantee as our earlier approximate policy obtained from problem (2), although the decisions of the two policies can be different. The number of decision variables in (2) is $|\mathcal{J}|\,T\,K\,|\mathcal{X}|^\theta$, whereas the number of decision variables in (8) with $\boldsymbol{\eta}=\overline{\boldsymbol{\eta}}$ is $|\mathcal{J}|\,T\left(|\mathcal{X}|\,(K-\theta)+\frac{|\mathcal{X}|}{|\mathcal{X}|-1}\,(|\mathcal{X}|^\theta-1)\right)$. When K is large, the number of decision variables in problem (2) can exceed that in problem (8) by a factor arbitrarily close to K. A similar comparison holds for the numbers of constraints in the two problems.

7. Reductions in the Fluid Approximation with Varying Histories

We give a number of possible reductions in problem (8) to solve this problem more efficiently. While each reduction is useful in its own right, we can also combine these reductions.

Aggregation by Time Periods and Decomposition by Products:

In problem (8), we can aggregate the decision variables corresponding to different time periods in a stage. In particular, we can use the decision variable $z_{jk}(\mathbf{s}_k^{\eta_k})$ to capture $\sum_{t\in\mathcal{T}} y_{jk}^t(\mathbf{s}_k^{\eta_k})$. In this case, the objective function in (8) becomes $\sum_{k\in\mathcal{K}}\sum_{j\in\mathcal{J}} f_j\,\mathbb{E}\{z_{jk}(\mathbf{X}_k^{\eta_k})\}$, whereas the two constraints in (8) become $\sum_{\ell=1}^k\sum_{j\in\mathcal{J}} a_{ij}\,\mathbb{E}\{z_{j\ell}(\mathbf{X}_\ell^{\eta_\ell})\,|\,\mathbf{X}_k^{\eta_k}=\mathbf{s}_k^{\eta_k}\}\le c_i$ and $z_{jk}(\mathbf{s}_k^{\eta_k})\le\sum_{t\in\mathcal{T}}\lambda_{jk}^t(\mathbf{s}_k^1)$. In this way, we reduce the number of decision variables in (8) by a factor of T. On the other hand, we can allocate the capacity of a resource to different products and solve a separate linear program for each product. We use $\mathcal{L}_j=\{i\in\mathcal{L}: a_{ij}=1\}$ to capture the set of resources used by product j. Letting δ_{ij} be the capacity of resource i allocated to product j, consider the linear program

$$
\begin{aligned}
Z_j^{\boldsymbol{\eta}}(\boldsymbol{\delta}_j) \;=\; & \max_{\boldsymbol{y}\in\mathbb{R}_+^{T\sum_{k\in\mathcal{K}}|\mathcal{X}|^{\eta_k}}} \sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}} f_j\,\mathbb{E}\{y_{jk}^t(\mathbf{X}_k^{\eta_k})\} \qquad (9)\\
& \text{st} \quad \sum_{\ell=1}^{k}\sum_{t\in\mathcal{T}}\mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell})\,|\,\mathbf{X}_k^{\eta_k}=\mathbf{s}_k^{\eta_k}\} \;\le\; \delta_{ij} \quad \forall\, i\in\mathcal{L}_j,\ k\in\mathcal{K},\ \mathbf{s}_k^{\eta_k}\in\mathcal{X}^{\eta_k}\\
& \qquad\quad y_{jk}^t(\mathbf{s}_k^{\eta_k}) \;\le\; \lambda_{jk}^t(\mathbf{s}_k^1) \quad \forall\, t\in\mathcal{T},\ k\in\mathcal{K},\ \mathbf{s}_k^{\eta_k}\in\mathcal{X}^{\eta_k}.
\end{aligned}
$$

In the problem above, we use $\boldsymbol{\delta}_j=(\delta_{ij}: i\in\mathcal{L}_j)$ to capture the allocation of the capacities of the resources used by product j to this product. Problem (9) is a variant of problem (8), where we focus on the decisions for product j and the capacity of resource i is the allocation of the capacity of this resource to product j. Letting $\mathbf{1}(\cdot)$ be the indicator function, the optimal objective value of problem (8) is given by $Z_{\mathsf{LP}}^{\boldsymbol{\eta}}=\max_{\boldsymbol{\delta}\in\mathbb{R}_+^{\sum_{j\in\mathcal{J}}|\mathcal{L}_j|}}\left\{\sum_{j\in\mathcal{J}} Z_j^{\boldsymbol{\eta}}(\boldsymbol{\delta}_j): \sum_{j\in\mathcal{J}}\mathbf{1}(i\in\mathcal{L}_j)\,\delta_{ij}=c_i\ \forall\, i\in\mathcal{L}\right\}$, where we use the vector $\boldsymbol{\delta}=(\delta_{ij}: j\in\mathcal{J},\ i\in\mathcal{L}_j)$. In the last problem, we find the best possible allocation of the capacities of the resources to the products. By duality theory, $Z_j^{\boldsymbol{\eta}}(\boldsymbol{\delta}_j)$ is concave in $\boldsymbol{\delta}_j$, so we can use standard convex programming tools to solve the last problem.

In practical applications, even when the number of resources is large, the number of resources used by a particular product tends to be small. In airline revenue management applications, for example, even when the number of flight legs is large, the number of flight legs used by a particular itinerary rarely exceeds two, so $|\mathcal{L}_j| \leq 2$ for all $j \in \mathcal{J}$. Therefore, the number of constraints in problem (9) can be significantly smaller than that in problem (8), allowing us to solve problem (9) significantly more quickly than (8). The idea of aggregating the decision variables by time periods and decomposing the fluid approximation by products is applicable to fluid approximations other than ours, but the idea of decomposing the fluid approximation by products, to our knowledge, has not been used in the existing literature. In the remainder of this section, we focus on other possible reductions that are specifically designed for our fluid approximation.

Decomposition by Stages:

Consider the case where we can partition the stages into two groups of consecutive stages such that the histories that we use for the stages in the second group do not intersect with the stages in the first group. In this case, we can decompose problem (8) by the two groups of stages. In particular, assume that there exists a stage q such that we have $k - \eta_k + 1 \geq q$ for all $k = q, \ldots, K$. Therefore, if $k \geq q$, then the history that we use in stage k does not intersect with stages $1, \ldots, q-1$. In this case, we partition the stages into the two groups $\{1, \ldots, q-1\}$ and $\{q, \ldots, K\}$. Fixing the vector $\boldsymbol{z} = (z_{ik}(\mathbf{s}^1_{k-\eta_k+1}) : i \in \mathcal{L},\ k = q, \ldots, K,\ \mathbf{s}^1_{k-\eta_k+1} \in \mathcal{X}) \in \mathbb{R}_+^{|\mathcal{L}|(K-q+1)\,|\mathcal{X}|}$, considering the decisions that we make in stages $1, \ldots, q-1$, we solve the linear program

$$\Psi^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z}) \;=\; \max_{\boldsymbol{y} \in \mathbb{R}_+^{|\mathcal{J}|T\sum_{k=1}^{q-1}|\mathcal{X}|^{\eta_k}}} \sum_{k=1}^{q-1} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} f_j\, \mathbb{E}\{y^t_{jk}(\mathbf{X}^{\eta_k}_k)\} \tag{10}$$

$$\text{st} \quad \sum_{\ell=1}^{k} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} a_{ij}\, \mathbb{E}\{y^t_{j\ell}(\mathbf{X}^{\eta_\ell}_\ell) \,|\, \mathbf{X}^{\eta_k}_k = \mathbf{s}^{\eta_k}_k\} \leq c_i \quad \forall\, i \in \mathcal{L},\ k = 1, \ldots, q-1,\ \mathbf{s}^{\eta_k}_k \in \mathcal{X}^{\eta_k}$$

$$\sum_{\ell=1}^{q-1} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} a_{ij}\, \mathbb{E}\{y^t_{j\ell}(\mathbf{X}^{\eta_\ell}_\ell) \,|\, \mathbf{X}^1_{k-\eta_k+1} = \mathbf{s}^1_{k-\eta_k+1}\} \leq z_{ik}(\mathbf{s}^1_{k-\eta_k+1}) \quad \forall\, i \in \mathcal{L},\ k = q, \ldots, K,\ \mathbf{s}^1_{k-\eta_k+1} \in \mathcal{X}$$

$$y^t_{jk}(\mathbf{s}^{\eta_k}_k) \;\leq\; \lambda^t_{jk}(\mathbf{s}^1_k) \quad \forall\, j \in \mathcal{J},\ t \in \mathcal{T},\ k = 1, \ldots, q-1,\ \mathbf{s}^{\eta_k}_k \in \mathcal{X}^{\eta_k}.$$

If we drop the second constraint above, then problem (10) is a variant of problem (8), where we consider the decisions only in stages $1, \ldots, q-1$. In the second constraint above, we focus on stages $q, \ldots, K$. Given that the state of the Markov chain in stage $k - \eta_k + 1$ is $\mathbf{s}^1_{k-\eta_k+1}$, we ensure that the total expected capacity consumption of resource i over stages $1, \ldots, q-1$ is at most $z_{ik}(\mathbf{s}^1_{k-\eta_k+1})$. By our assumption, if $k \geq q$, then we have $k - \eta_k + 1 \geq q$. Therefore, the left side of the second constraint above corresponds to the total expected capacity consumption of resource i over stages $1, \ldots, q-1$ conditional on the state of the Markov chain in a stage after stage $q-1$. The vector $\boldsymbol{z}$ will help us link the decisions in the two groups of stages.

We focus on making the decisions in stages $q, \ldots, K$. To make the decisions in stages $q, \ldots, K$, we solve a variant of problem (8), where the initial capacities of the resources are set by using the vector $\boldsymbol{z}$. In particular, fixing the vector $\boldsymbol{z} = (z_{ik}(\mathbf{s}^1_{k-\eta_k+1}) : i \in \mathcal{L},\ k = q, \ldots, K,\ \mathbf{s}^1_{k-\eta_k+1} \in \mathcal{X})$, considering the decisions that we make in stages $q, \ldots, K$, we solve the linear program

$$\Gamma^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z}) = \max_{\boldsymbol{y} \in \mathbb{R}_+^{|\mathcal{J}|T \sum_{k=q}^K |\mathcal{X}|^{\eta_k}}} \sum_{k=q}^K \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} f_j \, \mathbb{E}\{y^t_{jk}(\mathbf{X}^{\eta_k}_k)\} \tag{11}$$

$$\text{st} \quad \sum_{\ell=q}^k \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} a_{ij} \, \mathbb{E}\{y^t_{j\ell}(\mathbf{X}^{\eta_\ell}_\ell) \,|\, \mathbf{X}^{\eta_k}_k = \mathbf{s}^{\eta_k}_k\} \le c_i - z_{ik}(\mathbf{s}^1_{k-\eta_k+1}) \quad \forall\, i \in \mathcal{L},\ k = q, \ldots, K,\ \mathbf{s}^{\eta_k}_k \in \mathcal{X}^{\eta_k}$$

$$y^t_{jk}(\mathbf{s}^{\eta_k}_k) \le \lambda^t_{jk}(\mathbf{s}^1_k) \quad \forall\, j \in \mathcal{J},\ t \in \mathcal{T},\ k = q, \ldots, K,\ \mathbf{s}^{\eta_k}_k \in \mathcal{X}^{\eta_k}.$$

In this case, we can show that the optimal objective value of problem (8) is given by $Z^{\boldsymbol{\eta}}_{\mathsf{LP}} = \max_{\boldsymbol{z} \in \mathbb{R}_+^{|\mathcal{L}|(K-q+1)|\mathcal{X}|}} \big\{ \Psi^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z}) + \Gamma^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z}) : z_{ik}(\mathbf{s}^1_{k-\eta_k+1}) \le c_i \quad \forall\, i \in \mathcal{L},\ k = q, \ldots, K,\ \mathbf{s}^1_{k-\eta_k+1} \in \mathcal{X} \big\}$. We give this result in Appendix G. By duality theory, both $\Psi^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z})$ and $\Gamma^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z})$ are concave in $\boldsymbol{z}$, so we can use standard convex programming tools to solve the last problem. Problems (10) and (11), respectively, focus on the decisions in stages $1, \ldots, q-1$ and $q, \ldots, K$, but we link the decisions in the two groups of stages by using the vector $\boldsymbol{z}$. In our approximate policy with varying histories, we solve problem (8) with $\eta_k = 1$ for $k \le K - \theta$ and $\eta_k = k - (K - \theta)$ for $k \ge K - \theta + 1$. We have $k - \eta_k + 1 = K - \theta + 1$ for all $k \ge K - \theta + 1$, so if we partition the stages into the two groups $\{1, \ldots, K - \theta\}$ and $\{K - \theta + 1, \ldots, K\}$, then the histories that we use for the stages in the second group do not intersect with the stages in the first group. Thus, we can use the idea of decomposition by stages when solving problem (8) for our approximate policy with varying histories.

Equivalent Dynamic Program:

Consider problem (8) with $\eta_k = k$ for all $k \in \mathcal{K}$ so that we use the longest possible history in each stage. In this case, we can solve problem (8) through a dynamic program, where the number of product requests at each time period in each stage takes on its expected value and we can accept a fraction of a product request. We proceed to formulating the dynamic program. We capture the remaining resource capacities with the vector $\boldsymbol{y} = (y_i : i \in \mathcal{L}) \in \mathbb{R}_+^{|\mathcal{L}|}$, where y_i is the remaining capacity of resource i. We capture the decisions with the vector $\boldsymbol{u} = (u_j : j \in \mathcal{J}) \in [0,1]^{|\mathcal{J}|}$, where u_j is the fraction of the requests for product j that we accept. If the remaining resource capacities are given by the vector $\boldsymbol{y}$ and the current state of the Markov chain is s, then the set of feasible decisions at time period t in stage k is $\mathcal{G}^t_k(\boldsymbol{y}, s) = \{\boldsymbol{u} \in [0,1]^{|\mathcal{J}|} : \sum_{j \in \mathcal{J}} a_{ij} \, \lambda^t_{jk}(s) \, u_j \le y_i \ \forall\, i \in \mathcal{L}\}$, so the total capacity of resource i consumed by the accepted fractions of the product requests does not exceed its remaining capacity. To solve problem (8), we can use the dynamic program

$$V^t_k(\boldsymbol{y}, s) = \max_{\boldsymbol{u} \in \mathcal{G}^t_k(\boldsymbol{y}, s)} \left\{ \sum_{j \in \mathcal{J}} f_j \, \lambda^t_{jk}(s) \, u_j + V^{t+1}_k\Big(\boldsymbol{y} - \sum_{j \in \mathcal{J}} \boldsymbol{a}_j \, \lambda^t_{jk}(s) \, u_j, s\Big) \right\} \tag{12}$$

with boundary conditions $V^{T+1}_k(\boldsymbol{y}, s) = \sum_{q \in \mathcal{X}} P_k(s, q) \, V^1_{k+1}(\boldsymbol{y}, q)$ and $V^{T+1}_K(\boldsymbol{y}, s) = 0$. The dynamic program in (12) is a variant of the one in (1), where if the state of the Markov chain is s, then we have $\lambda^t_{jk}(s)$ requests for product j at time period t in stage k and we can accept a fraction of these product requests, but the Markov chain evolves according to its transition probabilities.

We can show that the optimal objective value of problem (8) is given by $Z^{\boldsymbol{\eta}}_{\mathsf{LP}} = \mathbb{E}\{V^1_1(\boldsymbol{c}, X_1)\}$. We give this result in Appendix H. We can combine the reductions that we give in this section. We give four possible example combinations. First, if we have $\eta_k = k$ for all $k \in \mathcal{K}$, then we can solve problem (9) by using a dynamic program similar to the one in (12), but the state variable in this dynamic program would have $|\mathcal{L}_j| + 1$ dimensions. We can solve this dynamic program either by discretizing the state space or by using the primal-dual algorithm in Brown and Zhang (2022). Second, in each of problems (9), (10) and (11), we can aggregate the decision variables corresponding to different time periods in a stage, reducing the number of decision variables by a factor of T. Third, problem (11) has the same structure as (8), so if $\eta_k = k - q + 1$ for all $k = q, \ldots, K$ so that we use the longest possible history in each stage, then we can solve problem (11) through a dynamic program similar to the one in (12). Fourth, if we aggregate the decision variables in problem (8) corresponding to different time periods in a stage, then we can also aggregate the decision epochs in the dynamic program in (12) corresponding to different time periods in a stage.

8. Computational Experiments

We provide computational experiments to test the tightness of the upper bounds and performance of the approximate policies provided by our fluid approximation.

Experimental Setup: We consider an airline network with one hub and three spokes. There is a flight leg that connects each spoke to the hub, as well as a flight leg that connects the hub to each spoke. Therefore, there are $2 \times 3 = 6$ flight legs. There is a high-fare and a low-fare itinerary that connect each possible origin-destination pair. Therefore, there are $2 \times 4 \times 3 = 24$ itineraries. The itineraries between two spokes connect through the hub, each using two flight legs, whereas the itineraries between the hub and a spoke are direct, each using one flight leg. In this context, the flight legs correspond to the resources and the itineraries correspond to the products. To come up with the revenues associated with the products, we place the hub at the center of a 100×100 square. We generate the locations of the spokes randomly over the square. The revenue associated with a low-fare itinerary is the Euclidean distance between its origin and destination locations. The revenue associated with a high-fare itinerary is κ times the revenue associated with the corresponding low-fare itinerary. Next, we explain our approach for generating the modulating Markov chain, as well as the request arrival probabilities for the products.

The Markov chain is stationary and has three states. We denote the state space of the Markov chain by $\mathcal{X} = \{1, 2, 3\}$. The initial state of the Markov chain is uniformly distributed over the state space. The transition probability matrix of the Markov chain is given by $P_k(s,s) = 2\,\delta$ for all $s \in \mathcal{X}$, whereas $P_k(s,q) = \frac{1}{2} - \delta$ for all $s, q \in \mathcal{X}$ with $s \neq q$. We vary the parameter δ. Because the transition probability matrix is doubly stochastic and the initial state of the Markov chain is uniformly distributed over the state space, we have $\mathbb{P}\{X_k = s\} = \frac{1}{3}$ for all $s \in \mathcal{X}$ and $k \in \mathcal{K}$. Also, we have $\frac{1}{2}\sum_{p\in\mathcal{X}} |P_k(s,p) - P_k(q,p)| = |3\,\delta - \frac{1}{2}|$, so if $\delta \geq \frac{1}{6}$, then the total variation distance of the Markov chain is $\mathsf{TV} = 3\,\delta - \frac{1}{2}$. Recalling the assumptions that the Markov chain satisfies $\mathbb{P}\{X_k = s\} \geq \epsilon$ for all $s \in \mathcal{X}$ and $k \in \mathcal{K}$, as well as $\mathsf{TV} \leq 1 - \alpha$, for some $\epsilon, \alpha > 0$, we can satisfy these assumptions by setting $\epsilon = \frac{1}{3}$ and $\alpha = \frac{3}{2} - 3\,\delta$. In Theorem 4.1, we set $\nu^\theta = 32\,\frac{T}{\epsilon^2\alpha^2}\,\min\{(K-\theta)\,T, c_{\min}\}$. Therefore, if we choose a larger value for δ, then we need to set a smaller value for α, yielding a larger value for ν^θ, in which case, by Theorem 4.1, we obtain a smaller performance guarantee for our approximate policy. In our computational experiments, the probability that there is no customer arrival at a time period in a stage depends on the state of the Markov chain. In particular, we will generate the product request arrival probabilities such that if the Markov chain is in state 1, 2 and 3, then the probability that there is no customer arrival is, respectively, 0.1, 0.5 and 0.9. We use $\psi_0(s)$ to denote the probability that there is no customer arrival at a time period in a stage given that the Markov chain is in state s, so we have $\psi_0(1) = 0.1$, $\psi_0(2) = 0.5$ and $\psi_0(3) = 0.9$.

To come up with the product request arrival probabilities $\{\lambda^t_{jk}(s) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ s \in \mathcal{X}\}$, we assume that the requests for the high-fare itineraries tend to arrive towards the end of the selling horizon so that it becomes important to protect the capacities on the flight legs for the high-fare itinerary requests that arrive later. There are K stages in the selling horizon. Letting $\lfloor\cdot\rfloor$ be the round down function, for each origin-destination pair (f,g), we sample the stage τ_{fg} from the uniform distribution over $\{\lfloor\frac{1}{3}K\rfloor, \ldots, \lfloor\frac{1}{2}K\rfloor\}$. The probability that we have a request for the high-fare itinerary connecting origin-destination pair (f,g) is zero until we reach stage τ_{fg} and increases after stage τ_{fg}. The probability that we have a request for the corresponding low-fare itinerary decreases over the stages. In particular, the function $L(k) = \frac{K+1-k}{K}$ is decreasing for $k \in [1,K]$, whereas the function $H_{fg}(k) = \frac{[k-\tau_{fg}]^+}{K-\tau_{fg}}$ takes value zero for $k \in [1,\tau_{fg}]$ and is increasing for $k \in [\tau_{fg}, K]$. For each origin-destination pair (f,g), we sample ζ_{fg} from the uniform distribution over $[0,1]$. Letting $\mathcal{P}$ be the set of all origin-destination pairs, we normalize the samples $\{\zeta_{fg} : (f,g) \in \mathcal{P}\}$ to add up to one by setting $\gamma_{fg} = \zeta_{fg}/\sum_{(h,\ell)\in\mathcal{P}}\zeta_{h\ell}$. In this case, if product j corresponds to the low-fare itinerary connecting origin-destination pair (f,g), then we set $\lambda^t_{jk}(s) = (1-\psi_0(s))\,\gamma_{fg}\,\frac{L(k)}{L(k)+H_{fg}(k)}$, whereas if product j corresponds to the high-fare itinerary connecting origin-destination pair (f,g), then we set $\lambda^t_{jk}(s) = (1-\psi_0(s))\,\gamma_{fg}\,\frac{H_{fg}(k)}{L(k)+H_{fg}(k)}$.

By the discussion in the previous paragraph, given that the Markov chain is in state s, at any time period in any stage, we get a request for an itinerary connecting origin-destination pair

(f,g) with probability $(1-\psi_0(s))\,\gamma_{fg}$. Thus, the probability that we get a request for an itinerary connecting origin-destination pair (f,g) is always proportional to γ_{fg}, but the allocation of this probability between high-fare and low-fare itineraries depends on the stage. Furthermore, noting that $\sum_{(f,g)\in\mathcal{P}}\gamma_{fg}=1$, if the Markov chain is in state s, then we get a request for some itinerary at any time period in any stage with probability $1-\psi_0(s)$, so $\psi_0(s)$ is the probability that there is no customer arrival given that the Markov chain is in state s. Once we generate the probability law of the Markov chain and product request arrival probabilities, the total expected demand for the capacity on flight leg i is $\Lambda_i=\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\sum_{s\in\mathcal{X}}a_{ij}\,\mathbb{P}\{X_k=s\}\,\lambda^t_{jk}(s)$. We set the capacity of each flight leg such that the total expected demand for the capacity of the flight leg exceeds its capacity by a factor of ρ, so we set the capacity of flight leg i as $c_i=\lfloor\Lambda_i/\rho\rfloor$. We vary the parameter ρ. In all of our test problems, we fix the fare difference between high-fare and low-fare itineraries as $\kappa=6$ and the number of time periods in each stage as $T=100$. We tried different values for these parameters and our results remained qualitatively the same.

Noting that the number of stages is K, as well as recalling that the parameter ρ controls the tightness of the flight leg capacities and δ controls the total variation distance of the Markov chain, we vary $K\in\{4,8,12\}$, $\rho\in\{1.2,1.6,2\}$ and $\delta\in\{\frac{2}{9},\frac{3}{9},\frac{4}{9}\}$ to obtain 27 parameter combinations.

Benchmark Strategies: We use two benchmarks. In the first benchmark, we solve the fluid approximation in (2) with different values for the number of stages in the history of the Markov chain. We check the upper bound provided by the optimal objective value of problem (2), as well as the total expected revenue from the approximate policy provided by the optimal solution to problem (2). We refer to this benchmark as FP-θ to emphasize that the fluid approximation in (2) explicitly takes into account the probability law of the Markov chain and uses the history of the Markov chain over θ stages. In our computational experiments, we vary $\theta\in\{1,2,3,4\}$. In the second benchmark, we solve a fluid approximation that uses only the total expected number of requests for each product. We refer to this benchmark as FEX to emphasize that this benchmark uses only the total expected number of requests for each product. The total expected number of requests for product j over the selling horizon is given by $\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{X}}\mathbb{P}\{X_k=s\}\,\lambda^t_{jk}(s)$. We use the decision variables $\boldsymbol{u}=(u_j:j\in\mathcal{J})$, where u_j is the total expected number of request for product j that we accept over the selling horizon. We solve the linear program

$$\overline{Z}_{\mathsf{LP}}=\max_{\boldsymbol{u}\in\mathbb{R}^{|\mathcal{J}|}_+}\left\{\sum_{j\in\mathcal{J}}f_j\,u_j\,:\,\sum_{j\in\mathcal{J}}a_{ij}\,u_j\le c_i\quad\forall\,i\in\mathcal{L},\quad u_j\le\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{X}}\mathbb{P}\{X_k=s\}\,\lambda^t_{jk}(s)\quad\forall\,j\in\mathcal{J}\right\}.\qquad(13)$$

In the problem above, the first constraint ensures that the total expected capacity consumption of each resource does not exceed its capacity, whereas the second constraint ensures that the total expected number of requests that we accept for each product does not exceed the total expected number of requests. We can show that the optimal objective value of problem (13) is an upper bound on the optimal total expected revenue. In particular, letting $\overline{\boldsymbol{y}}$ be an optimal solution to problem (2), considering the first constraint in this problem with $k = K$, if we take the expectation of both sides of the constraint and use the tower property of conditional expectation, then we get $\sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\} \le c_i$. Similarly, if we take the expectation of both sides of the second constraint and add this constraint over all $t \in \mathcal{T}$ and $k \in \mathcal{K}$, we get $\sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\} \le \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \mathbb{E}\{\lambda^t_{jk}(\mathbf{X}^1_k)\} = \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{X}} \mathbb{P}\{X_k = s\}\, \lambda^t_{jk}(s)$. In this way, the solution $\overline{\boldsymbol{u}} = (\overline{u}_j : j \in \mathcal{J})$ with $\overline{u}_j = \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \mathbb{E}\{\overline{y}^t_{jk}(\mathbf{X}^\theta_k)\}$ is feasible to problem (13) and provides an objective value that is equal to the optimal objective value of problem (2). Thus, the optimal objective value of problem (13) is an upper bound on the optimal objective value of problem (2), which is, in turn, an upper bound on the optimal total expected revenue. We can use problem (13) to obtain an approximate policy as well. In particular, letting $\overline{\boldsymbol{u}} = (\overline{u}_j : j \in \mathcal{J})$ be an optimal solution to this problem, the approximate policy accepts a request for product j at any time period in any stage with probability $\overline{u}_j / \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \sum_{s\in\mathcal{X}} \mathbb{P}\{X_k = s\}\, \lambda^t_{jk}(s)$.

Computational Results: Varying $K \in \{4, 8, 12\}$, $\rho \in \{1.2, 1.6, 2\}$ and $\delta \in \{\frac{2}{9}, \frac{3}{9}, \frac{4}{9}\}$, we obtain 27 parameter combinations. Using the approach discussed earlier in this section, we generate one test problem for each parameter configuration. For each test problem, we solve the fluid approximation in (2) with $\theta \in \{1, 2, 3, 4\}$. We check the upper bound on the optimal total expected revenue provided by the optimal objective value of problem (2), as well as estimate the total expected revenue from the approximate policy provided by the optimal solution to problem (2). We use the approximate policy given in Section 4 by setting the tuning parameter as $\gamma = 1$. While it is necessary to choose the value of the tuning parameter away from one to give a performance guarantee for our approximate policy, the practical performance of our approximate policy turned out to be the strongest when we set the tuning parameter to one. We use Monte Carlo simulation with 2000 sample paths to estimate the total expected revenue of the approximate policy. Similarly, we solve problem (13) for each test problem. We check the upper bound on the optimal total expected revenue provided by the optimal objective value of problem (13), as well as estimate the total expected revenue from the approximate policy provided by the optimal solution to problem (13).

We give our results in Table 1. The first column in this table gives the parameter configuration for our test problems by using the tuple (K, ρ, δ). In the rest of the table, we have two blocks, each with four columns. The first column in the first block gives the upper bound provided by problem (2) with $\theta = 1$, normalized by the upper bound provided by problem (13). In particular, noting that Z^θ_{LP} and $\overline{Z}_{\mathsf{LP}}$ are, respectively, the optimal objective values of problems (2) and (13), the first column gives $100 \times \frac{Z^1_{\mathsf{LP}}}{\overline{Z}_{\mathsf{LP}}}$. Thus, we interpret the first column as the upper bound from FP-1 relative

Params.	Normalized Upp. Bnd.				Normalized Exp. Rev.			
(K, ρ, δ)	FP-1	FP-2	FP-3	FP-4	FP-1	FP-2	FP-3	FP-4
$(4, 1.2, 2/9)$	94.77	93.51	92.90	92.30	109.54	110.83	113.55	120.57
$(4, 1.2, 3/9)$	93.99	91.82	91.24	90.92	114.18	118.98	122.34	127.28
$(4, 1.2, 4/9)$	91.08	90.10	89.85	89.78	129.26	134.40	136.62	139.05
$(4, 1.6, 2/9)$	94.38	93.36	93.28	92.90	110.11	115.08	116.42	122.70
$(4, 1.6, 3/9)$	94.68	92.89	92.56	92.26	115.80	122.87	127.56	130.33
$(4, 1.6, 4/9)$	93.08	92.29	92.14	92.07	135.32	141.57	142.83	143.91
$(4, 2.0, 2/9)$	93.86	92.90	92.76	92.73	115.43	120.21	121.19	123.67
$(4, 2.0, 3/9)$	94.59	93.18	92.69	92.62	120.86	123.85	130.66	131.68
$(4, 2.0, 4/9)$	94.43	93.71	93.45	93.42	140.93	147.22	148.33	148.53
$(8, 1.2, 2/9)$	96.01	94.15	93.77	93.75	108.45	109.15	108.56	107.90
$(8, 1.2, 3/9)$	95.19	93.82	93.02	92.71	115.45	118.11	119.16	120.18
$(8, 1.2, 4/9)$	93.57	92.57	91.73	91.20	131.81	134.43	136.77	138.45
$(8, 1.6, 2/9)$	95.60	93.55	93.15	93.14	110.24	111.16	110.57	111.88
$(8, 1.6, 3/9)$	94.88	93.34	92.58	92.50	116.12	121.16	121.74	123.21
$(8, 1.6, 4/9)$	94.67	93.80	93.08	92.60	130.71	134.65	138.67	141.90
$(8, 2.0, 2/9)$	95.31	93.11	92.69	92.68	110.76	112.33	113.26	112.92
$(8, 2.0, 3/9)$	94.31	92.62	91.87	91.86	118.41	121.45	121.61	121.58
$(8, 2.0, 4/9)$	93.43	92.39	92.12	92.06	133.47	133.48	134.36	134.63
$(12, 1.2, 2/9)$	97.13	95.43	94.97	94.69	106.67	106.11	106.96	107.34
$(12, 1.2, 3/9)$	96.01	94.59	93.74	93.22	111.65	114.00	114.72	113.96
$(12, 1.2, 4/9)$	94.48	93.81	92.99	92.28	125.18	125.97	127.19	128.36
$(12, 1.6, 2/9)$	96.84	94.99	94.48	94.17	106.07	105.32	106.54	106.82
$(12, 1.6, 3/9)$	95.64	94.08	93.13	92.59	109.35	111.80	113.11	113.42
$(12, 1.6, 4/9)$	95.11	94.24	93.48	92.94	120.03	122.19	123.64	123.78
$(12, 2.0, 2/9)$	96.65	94.67	94.13	93.81	108.79	108.18	110.14	110.59
$(12, 2.0, 3/9)$	95.34	93.58	92.50	92.11	113.43	115.70	116.03	115.55
$(12, 2.0, 4/9)$	92.83	91.47	90.78	90.53	127.70	126.63	127.21	127.23
Average	94.74	93.33	92.78	92.51	118.36	120.99	122.58	123.98

Table 1 Upper bounds and total expected revenues obtained by the benchmarks.

to the upper bound from FEX. Similarly, the second, third and fourth columns give the upper bounds from FP-2, FP-3 and FP-4 relative to the upper bound from FEX. The first column in the second block gives the total expected revenue from the approximate policy provided by problem (2) with $\theta = 1$, normalized by the total expected revenue from the approximate policy provided by problem (13). In particular, letting APX^θ and $\overline{\mathsf{APX}}$, respectively, be the total expected revenues from the approximate policies provided by problems (2) and (13), the first column gives $100 \times \frac{\mathsf{APX}^1}{\overline{\mathsf{APX}}}$. Thus, we interpret the first column as the performance the approximate policy provided by FP-1 relative to the performance of the approximate policy provided by FEX. Similarly, the second, third and fourth columns give the performance of the approximate policies provided by FP-2, FP-3 and FP-4 relative to the performance of the approximate policy provided by FEX.

Our results indicate that the fluid approximation in (2) can provide significantly tighter upper bounds than the one in (13). On average, the percent gap between the upper bounds from FP-1 and FEX is 5.26%. The same average percent gap increases to 7.49% when we compare FP-4 and FEX. We can noticeably tighten the upper bounds provided by problem (2) by using a larger number of stages in the history of the Markov chain. On average, FP-4 improves the upper bounds

Params. (K, ρ, δ)	Exp. Rev. to Upp. Bnd. Ratio FEX	FP-1	FP-2	FP-3	FP-4
$(4, 1.2, 2/9)$	70.39	81.36	83.43	86.05	91.95
$(4, 1.2, 3/9)$	64.60	78.47	83.71	86.61	90.43
$(4, 1.2, 4/9)$	57.42	81.49	85.65	87.31	88.93
$(4, 1.6, 2/9)$	69.53	81.11	85.70	86.78	91.83
$(4, 1.6, 3/9)$	63.94	78.21	84.58	88.13	90.34
$(4, 1.6, 4/9)$	57.17	83.11	87.69	88.61	89.36
$(4, 2.0, 2/9)$	68.91	84.75	89.17	90.04	91.90
$(4, 2.0, 3/9)$	63.52	81.16	84.43	89.53	90.30
$(4, 2.0, 4/9)$	56.40	84.17	88.60	89.51	89.66
$(8, 1.2, 2/9)$	75.87	85.70	87.96	87.83	87.32
$(8, 1.2, 3/9)$	69.31	84.06	87.25	88.79	89.84
$(8, 1.2, 4/9)$	59.85	84.31	86.92	89.23	90.86
$(8, 1.6, 2/9)$	76.23	87.89	90.57	90.49	91.56
$(8, 1.6, 3/9)$	69.77	85.39	90.57	91.74	92.93
$(8, 1.6, 4/9)$	60.40	83.39	86.71	89.98	92.57
$(8, 2.0, 2/9)$	77.74	90.34	93.78	94.99	94.72
$(8, 2.0, 3/9)$	71.99	90.39	94.40	95.29	95.29
$(8, 2.0, 4/9)$	64.56	92.23	93.28	94.16	94.41
$(12, 1.2, 2/9)$	80.44	88.35	89.44	90.59	91.18
$(12, 1.2, 3/9)$	74.86	87.04	90.21	91.61	91.51
$(12, 1.2, 4/9)$	65.22	86.41	87.57	89.20	90.71
$(12, 1.6, 2/9)$	82.33	90.18	91.29	92.85	93.39
$(12, 1.6, 3/9)$	77.18	88.24	91.72	93.73	94.54
$(12, 1.6, 4/9)$	68.44	86.37	88.74	90.52	91.15
$(12, 2.0, 2/9)$	81.85	92.13	93.53	95.77	96.49
$(12, 2.0, 3/9)$	76.98	91.59	95.18	96.56	96.57
$(12, 2.0, 4/9)$	69.87	96.11	96.72	97.91	98.20
Average	69.44	86.07	89.21	90.88	92.15

Table 2 Ratios between the upper bounds and total expected revenues obtained by the benchmarks.

from FP-1 by 2.41%. There are test problems where the gap between the upper bounds from FP-4 and FP-1 can reach 3.51%. More importantly, perhaps, the performance of the approximate policy from the fluid approximation in (2) can be substantially stronger than the performance of the approximate policy from the fluid approximation in (13). On average, the percent gap between the total expected revenues from the approximate policies provided by FP-1 and FEX is 18.36%. The same average percent gap increases to 23.98% when we compare FP-4 and FEX. Also, using a larger number of stages in the history of the Markov chain in the fluid approximation in (2), we can obtain approximate policies with better performance. On average, the performance of the approximate policy from FP-4 improves the performance of the approximate policy from FP-1 by 4.75%. By Theorem 4.1, we obtain a stronger performance guarantee for our approximate policy when we use a larger number of stages in the history of the Markov chain. Aligned with this result, the practical performance of the approximate policy from the fluid approximation in (2) tends to get stronger as we use a larger number of stages in the history of the Markov chain.

There are two parameters that particularly affect the performance of the approximate policies from the fluid approximations in (2) and (13). First, as the value of the parameter δ increases, the gap between the performance of the policies from FP-1 and FEX increases. For the test problems with $\delta = \frac{2}{9}$, $\delta = \frac{3}{9}$ and $\delta = \frac{4}{9}$, the average percent gaps between the performance of the policies from FP-1 and FEX are, respectively, 9.56%, 15.03% and 30.49%. As the parameter δ gets larger, the total variation distance of the Markov chain increases. Markov chains with larger total variation distances take longer to mix; see Lemma 4.10 in Levin and Peres (2017). Therefore, if the Markov chain takes longer to mix, then using a fluid approximation that explicitly incorporates the probability law of the Markov chain pays off. Second, as the value of the parameter ρ increases, the gap between the performance of the policies from FP-1 and FEX increases. For the test problems with $\rho = 1.2$, $\rho = 1.6$ and $\rho = 2$, the average percent gaps between the performance of the policies from FP-1 and FEX are, respectively, 16.91%, 17.08% and 21.09%. As the parameter ρ gets larger, the capacities of the flight legs get tighter. If the capacities of the fight legs are tighter, then it becomes more important to carefully ration the capacity, in which case, using an approximate policy based on a more sophisticated fluid approximation appears to pay off.

By the chain of inequalities in Theorem 4.1, the gap between the total expected revenue of our approximate policy and the upper bound on the optimal total expected revenue diminishes as the capacities of the flight legs get large. In Table 2, we give the ratio between the total expected revenue of the approximate policy and the upper bound on the optimal total expected revenue for all of our benchmarks. The first column in this table gives the parameter configuration for our test problems. Recalling the definitions of Z^{θ}_{LP}, $\overline{Z}_{\mathsf{LP}}$, APX^{θ} and $\overline{\mathsf{APX}}$ earlier in this section, the second column gives $100 \times \frac{\overline{\mathsf{APX}}}{\overline{Z}_{\mathsf{LP}}}$, whereas the last four columns give $100 \times \frac{\mathsf{APX}^{\theta}}{Z^{\theta}_{\mathsf{LP}}}$ for all $\theta \in \{1, 2, 3, 4\}$. In our test problems, as the number of stages in the selling horizon gets larger, the capacities of the flight legs get larger. For the test problems with $K = 4$, $K = 8$ and $K = 12$, the average ratio between the total expected revenue of the approximate policy and the upper bound from FP-4 are, respectively, 90.52%, 92.17% and 93.75%. Thus, for test problems with $K = 12$, we can verify that the average optimality gap of the approximate policy from FP-4 is at most 6.25%. The same average ratios for FEX are, respectively, 63.54%, 69.52% and 75.24%, consistently staying away from one.

We carried our computational experiments in Python 3.11.3 and Gurobi 9.5.2 on MacOS with M1 CPU and 16 GB RAM. For our smallest test problems with $K = 4$, the average running times to solve problem (2) with $\theta = 1$ and $\theta = 4$ are, respectively, 0.29 and 32.15 seconds. For our largest test problems with $K = 12$, the average running times to solve problem (2) with $\theta = 1$ and $\theta = 4$ are, respectively, 2.57 seconds and 17.14 minutes. More than 98.31% of these running times is spent on constructing the linear program, rather than solving the linear program through the simplex algorithm. Lastly, in our approximate policy, letting $\overline{\boldsymbol{y}} = (\overline{y}^t_{jk}(\mathbf{s}^{\theta}_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^{\theta}_k \in \mathcal{X}^{\theta})$ be an optimal solution to problem (2), we aggregate the values of the decision variables in this solution to compute the probability of accepting a request for a product at each time period in each stage. We need the aggregation step to give a performance guarantee for our approximate policy. We can consider an alternative approximate policy that avoids the aggregation step. Given that the history of the Markov chain over past θ stages is $\mathbf{s}_k^\theta$, the alternative approximate policy accepts a request for product j at time period t in stage k with probability $\gamma\,\overline{y}_{jk}^t(\mathbf{s}_k^\theta)/\lambda_{jk}^t(\mathbf{s}_k^1)$. We can give an asymptotic optimality guarantee for the alternative approximate policy, but we do not have a characterization of the dependence of the performance guarantee on the number of stages in the history. In our test problems, the performance of the alternative approximate policy is very close to that of our approximate policy. Over all of our test problems, the average absolute percent gap between the performance of the two policies is 0.19%. In problem (8), we also give a fluid approximation with varying histories in different stages. Defining $\overline{\boldsymbol{\eta}} = (\overline{\eta}_1, \ldots, \overline{\eta}_K)$ as $\overline{\eta}_k = 1$ for $k \le K - \theta$, whereas $\overline{\eta}_k = k - (K - \theta)$ for $k \ge K - \theta + 1$, we can use an optimal solution to problem (8) with $\boldsymbol{\eta} = \overline{\boldsymbol{\eta}}$ to construct an approximate policy with varying histories. By Theorem 6.2, the approximate policy with varying histories has the same performance guarantee as our approximate policy. In our test problems, the performance of the approximate policy with varying histories is virtually identical to that of our approximate policy. Over all of our test problems, the average absolute percent gap between the performance of the two policies is 0.08%.

9. Conclusions

We developed a family of history-dependent fluid approximations for revenue management problems with Markov-modulated demands. We used our fluid approximation to construct an approximate policy with a performance guarantee. To our knowledge, our approximate policy is the first one to provide an asymptotic optimality guarantee. There are several avenues for future research. First, using Markov-modulated demands is only one way of incorporating demand dependence. We can explore other ways of incorporating demand dependence and develop corresponding fluid approximations. Second, a crucial step in the proof of our performance guarantee uses an upper bound on the moment generating function of the random variable $W_{ik} - c_i$ in (7) to upper bound the exhaustion probability of resource i. In this chain of inequalities, we can directly use the Markov inequality to obtain $\mathbb{P}\{W_{ik} \ge c_i \,|\, \mathbf{X}_k^{\sigma_k}\} \le \mathbb{E}\{W_{ik} \,|\, \mathbf{X}_k^{\sigma_k}\}/c_i \le \gamma$, where the last inequality follows from the discussion just before Lemma 5.3. Thus, by (4), we get $\mathbb{P}\{G_{jk}^t = 1 \,|\, \mathbf{X}_k^{\sigma_k}\} \ge 1 - L\gamma$, so we can choose $\beta = 1 - L\gamma$ in (3), yielding a performance guarantee of $\gamma(1 - L\gamma)$. In this case, setting the tuning parameter as $\gamma = \frac{1}{2L}$, we obtain a performance guarantee of $\frac{1}{4L}$, which is akin to a constant factor performance guarantee when the number of resources used by a product is uniformly bounded. It is useful to explore other performance guarantees in other asymptotic scaling regimes. Third, our use of the average Lipschitz condition to bound the availability probabilities can find applications in other revenue management problems.

References

Aouad, A., W. Ma. 2022. A nonparametric framework for online stochastic matching with correlated arrivals. Tech. rep., London Business School, London, UK.

Bai, Y., O. El Housni, B. Jin, P. Rusmevichientong, H. Topaloglu, D. P. Williamson. 2023. Fluid approximations for revenue management under high-variance demand. *Management Science* **69**(7) 4016–4026.

Beyer, D., S. P. Sethi. 1997. Average cost optimality in inventory models with Markovian demands. *Journal of Optimization Theory and Applications* **92**(3) 497–526.

Brown, D. B., J. Zhang. 2022. Dynamic programs with shared resources and signals: Dynamic fluid policies and asymptotic optimality. *Operations Research* **70**(5) 3015–3033.

Chen, F., J.-S. Song. 2001. Optimal policies for multiechelon inventory problems with Markov-modulated demand. *Operations Research* **49**(2) 226–234.

Dubhashi, D. P., A. Panconesi. 2009. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, Cambridge, UK.

Feldman, R. M. 1978. A continuous review (s,S) inventory system in a random environment. *Journal of Applied Probability* **15**(3) 654–659.

Feng, Y., R. Niazadeh, A. Saberi. 2024. Technical note – Near-optimal Bayesian online assortment of reusable resources. *Operations Research* **72**(5) 1861–1873.

Gallego, G., G. Iyengar, R. Phillips, A. Dubey. 2004. Managing flexible products on a network. CORC Technical Report TR-2004-01.

Gallego, G., H. Topaloglu. 2019. *Revenue Management and Pricing Analytics*. International Series in Operations Research & Management Science, Springer, New York, NY.

Gallego, G., G. van Ryzin. 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science* **40**(8) 999–1020.

Gallego, G., G. van Ryzin. 1997. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research* **45**(1) 24–41.

Iglehart, D., S. Karlin. 1960. Optimal policy for dynamic inventory process with non-stationary stochastic demands. Tech. rep., Stanford University, Stanford, CA. URL `https://purl.stanford.edu/kd300nw0124`.

Jasin, S., S. Kumar. 2012. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research* **37**(2) 313–345.

Jiang, J. 2023. Constant approximation for network revenue management with Markovian-correlated customer arrivals. Tech. rep., Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong.

Jin, B., H. Topaloglu, D. Williamson. 2024. Network revenue management under Markov-modulated demand. Tech. rep., Cornell University, Ithaca, NY. Work in progress.

Levin, D. A., Y. Peres. 2017. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI.

Li, W., P. Rusmevichientong, H. Topaloglu. 2024. Technical note – Revenue management with calendar-aware and dependent demands: Asymptotically tight fluid approximations. *Operations Research* (to appear). URL `https://doi.org/10.1287/opre.2023.0442`.

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

Ma, W., D. Simchi-Levi, J. Zhao. 2021. Dynamic pricing (and assortment) under a static calendar. *Management Science* **67**(4) 2292–2313.

Ma, Y., P. Rusmevichientong, M. Sumida, H. Topaloglu. 2020. An approximation algorithm for network revenue management under nonstationary arrivals. *Operations Research* **68**(3) 834–855.

Ozer, O, R. Phillips. 2012. *The Oxford Handbook of Pricing Management*. Oxford University Press, Oxford, UK.

Song, J.-S., P. Zipkin. 1993. Inventory control in a fluctuating demand environment. *Operations Research* **41**(2) 351–370.

Talluri, K., G. van Ryzin. 1998. An analysis of bid-price controls for network revenue management. *Management Science* **44**(11) 1577–1593.

Talluri, K. T., G. J. van Ryzin. 2005. *The Theory and Practice of Revenue Management*. Kluwer Academic Publishers, Boston, MA.

Yuan, Q., L. Du, M. Hu. 2024. Dynamic pricing under self-exciting arrival processes. Tech. rep., University of Toronto, Toronto, ON.

Zhao, Y., X. Li, L. Luo. 2025. Dynamic allocation of display advertising impressions in dual sales channels. *Omega* **131** 103213.

Online Appendix:
History-Dependent Fluid Approximations and Performance Guarantees for Revenue Management with Markov-Modulated Demands

Weiyuan Li, Paat Rusmevichientong, Huseyin Topaloglu, Jingwei Zhang
January 21, 2025

Appendix A: Moment Generating Function Bound under Average Lipschitz Condition

Let $Z_1,\ldots,Z_n$ be a sequence of random variables, each with support $\mathcal{Z}$. Consider the function $g:\mathcal{Z}^n\to\mathbb{R}$ mapping the sequence of random variables to reals. For all $\ell=1,\ldots,n$ and $p,q\in\mathcal{Z}$, if we have $\big|\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell=p,Z_{\ell+1},\ldots,Z_n\}-\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell=q,Z_{\ell+1},\ldots,Z_n\}\big|\le\delta_\ell$ with probability one for some deterministic $\delta_\ell\in\mathbb{R}_+$, then we say that the function g satisfies the average Lipschitz condition. In the next lemma, we give an upper bound on the moment generating function of $g(Z_1,\ldots,Z_n)$ when the function g satisfies the average Lipschitz condition.

Lemma A.1 (Moment Generating Function Bound) *Considering the function g satisfying the average Lipschitz condition, for all $\lambda\ge 0$ and $\ell=1,\ldots,n$, we have*

$$\mathbb{E}\{\exp(\lambda\,g(Z_1,\ldots,Z_n))\,|\,Z_\ell,\ldots,Z_n\}\le\exp\left(\frac{\lambda^2}{2}\sum_{k=1}^{\ell-1}\delta_k^2\right)\times\exp(\lambda\,\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell,\ldots,Z_n\}).$$

Proof: Define the random variable $M_\ell=\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell,\ldots,Z_n\}$ for all $\ell=1,\ldots,n$. Using the tower property of conditional expectation, we have the chain of equalities $\mathbb{E}\{M_\ell\,|\,Z_{\ell+1},\ldots,Z_n\}=\mathbb{E}\{\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell,\ldots,Z_n\}\,|\,Z_{\ell+1},\ldots,Z_n\}=\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_{\ell+1},\ldots,Z_n\}=M_{\ell+1}$. We have $-\delta_\ell\le\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell=p,Z_{\ell+1},\ldots,Z_n\}-\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell=q,Z_{\ell+1},\ldots,Z_n\}\le\delta_\ell$ for all $p,q\in\mathcal{Z}$ with probability one, where we use the fact that the function g satisfies the average Lipschitz condition. Multiplying this chain of inequalities by $\mathbb{P}\{Z_\ell=p\,|\,Z_{\ell+1},\ldots,Z_n\}$ and adding over all $p\in\mathcal{Z}$, we get $-\delta_\ell\le\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_{\ell+1},\ldots,Z_n\}-\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_\ell=q,Z_{\ell+1},\ldots,Z_n\}\le\delta_\ell$ for all $q\in\mathcal{Z}$. Thus, noting the definition of M_ℓ, the last chain of inequalities is equivalent to having $|M_\ell-M_{\ell+1}|\le\delta_\ell$ with probability one for all $\ell=1,\ldots,n-1$. Furthermore, by the discussion at the beginning of this paragraph, we have $\mathbb{E}\{M_\ell-M_{\ell+1}\,|\,Z_{\ell+1},\ldots,Z_n\}=0$, in which case, by Lemma 5.1, in Dubhashi and Panconesi (2009), we get $\mathbb{E}\{\exp(\lambda(M_\ell-M_{\ell+1}))\,|\,Z_{\ell+1},\ldots,Z_n\}\le\exp(\frac{\lambda^2}{2}\delta_\ell^2)$. We show the inequality in the lemma by using induction over $\ell=1,\ldots,n$. For $\ell=1$, observe that given $Z_1,\ldots,Z_n$, the quantity $g(Z_1,\ldots,Z_n)$ is deterministic. In this case, we immediately obtain the equality $\mathbb{E}\{\exp(\lambda\,g(Z_1,\ldots,Z_n))\,|\,Z_1,\ldots,Z_n\}=\exp(\lambda\,\mathbb{E}\{g(Z_1,\ldots,Z_n)\,|\,Z_1,\ldots,Z_n\})$ for all $\lambda\ge 0$. Therefore, it follows that the inequality in the lemma holds for $\ell=1$.

Assuming that the inequality in the lemma holds for ℓ, we proceed to showing that the inequality in the lemma also holds for $\ell+1$. By the tower property of conditional expectation, we have $\mathbb{E}\{\exp(\lambda\, g(Z_1,\dots,Z_n))\,|\,Z_{\ell+1},\dots,Z_n\} = \mathbb{E}\{\mathbb{E}\{\exp(\lambda\, g(Z_1,\dots,Z_n))\,|\,Z_\ell,\dots,Z_n\}\,|\,Z_{\ell+1},\dots,Z_n\}$, in which case, we obtain the chain of inequalities

$$\begin{aligned}
\mathbb{E}\{\exp(\lambda\, g(Z_1,\dots,Z_n))\,|\,Z_{\ell+1},\dots,Z_n\} &= \mathbb{E}\{\mathbb{E}\{\exp(\lambda\, g(Z_1,\dots,Z_n))\,|\,Z_\ell,\dots,Z_n\}\,|Z_{\ell+1},\dots,\dots,Z_n\} \\
&\overset{(a)}{\leq} \mathbb{E}\left\{\exp\left(\frac{\lambda^2}{2}\sum_{k=1}^{\ell-1}\delta_k^2\right)\times\exp(\lambda\,\mathbb{E}\{g(Z_1,\dots,Z_n)\,|\,Z_\ell,\dots,Z_n\})\,\middle|\,Z_{\ell+1},\dots,Z_n\right\} \\
&= \exp\left(\frac{\lambda^2}{2}\sum_{k=1}^{\ell-1}\delta_k^2\right)\times\mathbb{E}\{\exp(\lambda\, M_\ell)\,|\,Z_{\ell+1},\dots,Z_n\} \\
&\overset{(b)}{=} \exp\left(\frac{\lambda^2}{2}\sum_{k=1}^{\ell-1}\delta_k^2\right)\times\exp(\lambda\, M_{\ell+1})\times\mathbb{E}\{\exp(\lambda\,(M_\ell-M_{\ell+1}))\,|\,Z_{\ell+1},\dots,Z_n\} \\
&\overset{(c)}{\leq} \exp\left(\frac{\lambda^2}{2}\sum_{k=1}^{\ell}\delta_k^2\right)\times\exp(\lambda\, M_{\ell+1}) \\
&= \exp\left(\frac{\lambda^2}{2}\sum_{k=1}^{\ell}\delta_k^2\right)\times\exp(\lambda\,\mathbb{E}\{g(Z_1,\dots,Z_n)\,|\,Z_{\ell+1},\dots,Z_n\}),
\end{aligned}$$

where (a) uses the induction assumption, (b) holds because the quantity $M_{\ell+1}$, given $Z_{\ell+1},\dots,Z_n$, is deterministic and (c) uses the inequality that we give in the previous paragraph. ■

Appendix B: Moment Generating Function Bound for a Bernoulli Random Variable

In the next lemma, we give an upper bound on the moment generating function of a Bernoulli random variable. We use this lemma in the proof of Lemma 5.3.

Lemma B.1 (Bernoulli Moment Generating Function Bound) *If Z is a Bernoulli random variable with mean μ, then we have $\mathbb{E}\{\exp(\lambda\, Z)\}\leq\exp((\lambda^2+\lambda)\,\mu)$ for all $\lambda\in[0,1]$.*

Proof: Because Z is a Bernoulli random variable with mean μ, we have $\mathbb{E}\{e^{\lambda\, Z}\}=\mu\, e^\lambda+1-\mu=1+(e^\lambda-1)\,\mu\leq\exp((e^\lambda-1)\,\mu)$, where the last inequality uses the fact that $1+x\leq e^x$ for all $x\in\mathbb{R}$. Therefore, if we can show that $e^\lambda-1\leq\lambda^2+\lambda$ for all $\lambda\in[0,1]$, then the inequality in the lemma holds. We have $e^0=1$ and $e^1\leq 3$, so noting that e^x is convex in x, the function e^x lies below the line segment connecting the points $(0,1)$ and $(1,3)$ for all $x\in[0,1]$. Therefore, we have $e^x\leq 1+2x$ for all $x\in[0,1]$. In this case, for $\lambda\in[0,1]$, we obtain $e^\lambda-1=\int_0^\lambda e^x\,dx\leq\int_0^\lambda(1+2x)\,dx=\lambda^2+\lambda$, which is the desired inequality. ■

Appendix C: Losing the Upper Bound from the Fluid Approximation with Varying Histories

We give an example to demonstrate that if we do not have $\ell-\eta_\ell\leq k-\eta_k$ for all $\ell\leq k$, then the optimal objective value of problem (8) may not be an upper bound on the optimal total expected revenue. We consider a problem instance with a single resource and three products, so we have $\mathcal{L}=\{1\}$ and $\mathcal{J}=\{1,2,3\}$. The capacity of the resource is one. The revenues of the products are given by $f_1 = 3$, $f_2 = 2$ and $f_3 = 1$. There are three stages and there is one time period in each stage, so we have $K = 3$ and $T = 1$. The state space of the modulating Markov chain is $\mathcal{X} = \{0,1\}$. We have $\mathbb{P}\{X_1 = 0\} = \mathbb{P}\{X_1 = 1\} = \frac{1}{2}$, $\mathbb{P}\{X_2 = 1 \,|\, X_1 = 0\} = \mathbb{P}\{X_2 = 1 \,|\, X_1 = 1\} = 1$ and $\mathbb{P}\{X_3 = 1 \,|\, X_2 = 1\} = 1$. Thus, the initial state of the Markov chain is 0 or 1 with equality probabilities. In the second and third stages, the modulating Markov chain is always in state 1. We have $\lambda^1_{jk}(0) = 0$ for all $j = 1,2,3$ and $k = 1,2,3$, whereas $\lambda^1_{11}(1) = 1$, $\lambda^1_{22}(1) = 1$ and $\lambda^1_{33}(1) = 1$, which is to say that if the Markov is in state 0, then we do not have any product requests, whereas if the Markov chain is in stage 1 in the first, second and third stages, then we get requests, respectively, for products 1, 2 and 3. Consider problem (8) with $\eta_1 = 1$ $\eta_2 = 1$ and $\eta_3 = 3$. Note that we have $2 - \eta_2 = 1 > 0 = \eta_3 - 3$, so we do not have $\ell - \eta_\ell \le k - \eta_k$ for $\ell \le k$ for our choice of (η_1, η_2, η_3). For this choice of (η_1, η_2, η_3), there are two possible realizations of $\mathbf{X}^{\eta_1}_1$ given by (0) and (1), there is one possible realization of $\mathbf{X}^{\eta_2}_2$ given by (1) and there are two possible realizations of $\mathbf{X}^{\eta_3}_3$ given by $(0,1,1)$ and $(1,1,1)$. Thus, problem (8) for our problem instance is

$$\begin{aligned}
\max \quad & \frac{1}{2}\, 3\, y^1_{11}(1) + 2\, y^1_{22}(1) + \frac{1}{2} y^1_{33}(0,1,1) + \frac{1}{2} y^1_{33}(1,1,1) \\
\text{st} \quad & y^1_{11}(1) \le 1,\ \frac{1}{2} y^1_{11}(1) + y^1_{22}(1) \le 1,\ y^1_{11}(1) + y^1_{22}(1) + y^1_{33}(1,1,1) \le 1,\ y^1_{22}(1) + y^1_{33}(0,1,1) \le 1 \\
& 0 \le y^1_{11}(1) \le 1,\ 0 \le y^1_{22}(1) \le 1,\ 0 \le y^1_{33}(0,1,1) \le 1,\ 0 \le y^1_{33}(1,1,1) \le 1.
\end{aligned}$$

We can upper bound the objective function as $\frac{1}{2}\, 3\, y^1_{11}(1) + 2\, y^1_{22}(1) + \frac{1}{2} y^1_{33}(0,1,1) + \frac{1}{2} y^1_{33}(1,1,1) \le \frac{3}{2}(y^1_{11}(1) + y^1_{22}(1) + y^1_{33}(1,1,1)) + \frac{1}{2}\,(y^1_{22}(1) + y^1_{33}(0,1,1))$. Noting the third and fourth constraints in the problem above, the expression on the right side of the last inequality at any feasible solution is upper bounded by 2. Thus, the optimal objective value is upper bounded by 2, so $Z^{\boldsymbol{\eta}}_{\mathsf{LP}} \le 2$. We compute the optimal total expected revenue for our problem instance. The request for product 1 arrives before the request for product 2. Similarly, the request for product 2 arrives before the request for product 3. Furthermore, the revenue of product 1 is greater than the revenue of product 2, which is, in turn, greater than the revenue of product 3. Therefore, it is optimal to accept the product requests in a first come and first serve fashion without reserving the capacity of the resource for the future product requests. In this case, the optimal total expected revenue is given by $\mathsf{opt} = f_1\, \mathbb{P}\{X_1 = 1\} + f_2\, \mathbb{P}\{X_1 = 0,\ X_2 = 1\} = 3\, \frac{1}{2} + 2\, \frac{1}{2} = \frac{5}{2}$. Therefore, we have $\mathsf{opt} = \frac{5}{2} > 2 \ge Z^{\boldsymbol{\eta}}_{\mathsf{LP}}$, demonstrating that if we do not have $\ell - \eta_\ell \le k - \eta_k$ for all $\ell \le k$, then the optimal objective value of problem (8) may not be an upper bound on the optimal total expected revenue.

Appendix D: Upper Bound from the Fluid Approximation with Varying Histories

We give a proof for Theorem 6.1. The proof uses ideas similar to those in the proofs of Theorems 3.1 and 3.2. Therefore, we confine our attention to describing how the proof of Theorem 6.1

differs from the proofs of Theorems 3.1 and 3.2 without replicating the ideas in the proofs of Theorems 3.1 and 3.2. First, we focus on showing the inequality $Z^{\boldsymbol{\mu}}_{\mathsf{LP}} \geq \mathsf{opt}$ for all $\boldsymbol{\mu} \in \Theta$. Let the Bernoulli random variable Y^t_{jk} take value one if and only if the optimal policy accepts a request for product j at time period t in stage k. For stages ℓ and k that satisfy $\ell \leq k$, we establish the identity $\mathbb{E}\{\mathbb{E}\{Y^t_{j\ell} \,|\, \mathbf{X}^{\mu_\ell}_\ell\} \,|\, \mathbf{X}^{\mu_k}_k\} = \mathbb{E}\{Y^t_{j\ell} \,|\, \mathbf{X}^{\mu_k}_k\}$. This identity is the analogue of the one that we established just before the proof of Theorem 3.1. In particular, because $\boldsymbol{\mu} \in \Theta$, noting that $\ell \leq k$, we have $\ell - \mu_\ell \leq k - \mu_k$, so conditional on $\mathbf{X}^{\mu_\ell}_\ell$, $(X_1, \ldots, X_\ell)$ are independent of $\mathbf{X}^{\mu_k}_k$. In this case, because the decisions of the optimal policy in stage ℓ depends on the history of the Markov chain up to and including stage ℓ, conditional on $\mathbf{X}^{\mu_\ell}_\ell$, $Y^t_{j\ell}$ is independent of $\mathbf{X}^{\mu_k}_k$, so we have $\mathbb{E}\{Y^t_{j\ell} \,|\, (\mathbf{X}^{\mu_\ell}_\ell, \mathbf{X}^{\mu_k}_k)\} = \mathbb{E}\{Y^t_{j\ell} \,|\, \mathbf{X}^{\mu_\ell}_\ell\}$. Taking the expectations of both sides in the last equality conditional on $\mathbf{X}^{\mu_k}_k$, we get $\mathbb{E}\{\mathbb{E}\{Y^t_{j\ell} \,|\, \mathbf{X}^{\mu_\ell}_\ell\} \,|\, \mathbf{X}^{\mu_k}_k\} = \mathbb{E}\{\mathbb{E}\{Y^t_{j\ell} \,|\, (\mathbf{X}^{\mu_\ell}_\ell, \mathbf{X}^{\mu_k}_k)\} \,|\, \mathbf{X}^{\mu_k}_k\} = \mathbb{E}\{Y^t_{j\ell} \,|\, \mathbf{X}^{\mu_k}_k\}$, where the last equality uses the tower property of conditional expectation. Thus, the desired identity holds. Using this identity, we can follow the same argument in the proof of Theorem 3.1 to show that $Z^{\boldsymbol{\mu}}_{\mathsf{LP}} \geq \mathsf{opt}$.

Second, we focus on showing the inequality $Z^{\boldsymbol{\nu}}_{\mathsf{LP}} \geq Z^{\boldsymbol{\mu}}_{\mathsf{LP}}$ for all $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ with $\mu_k \geq \nu_k$ for all $k \in \mathcal{K}$. Let $\overline{\boldsymbol{y}} = (\overline{y}^t_{jk}(\mathbf{s}^{\mu_k}_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^{\mu_k}_k \in \mathcal{X}^{\mu_k})$ be an optimal solution to problem (8) when we solve this problem with $\boldsymbol{\eta} = \boldsymbol{\mu}$. For stages ℓ and k that satisfy $\ell \leq k$, we establish the identity $\mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\, \mathbf{X}^{\mu_k}_k\} \,|\, \mathbf{X}^{\nu_k}_k\} = \mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\, \mathbf{X}^{\nu_\ell}_\ell\} \,|\, \mathbf{X}^{\nu_k}_k\}$. This identity is the analogue of the one that we established just before Theorem 3.2. In particular, because $\boldsymbol{\nu} \in \Theta$, noting that $\ell \leq k$, we have $\ell - \nu_\ell \leq k - \nu_k$, so conditional on $\mathbf{X}^{\nu_\ell}_\ell$, $(X_1, \ldots, X_\ell)$ are independent of $\mathbf{X}^{\nu_k}_k$. In this case, we have $\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\, (\mathbf{X}^{\nu_\ell}_\ell, \mathbf{X}^{\nu_k}_k)\} = \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\, \mathbf{X}^{\nu_\ell}_\ell\}$. Taking the expectations of both sides in the last equality conditional on $\mathbf{X}^{\nu_k}_k$, we get $\mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\, \mathbf{X}^{\nu_\ell}_\ell\} \,|\, \mathbf{X}^{\nu_k}_k\} = \mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\, (\mathbf{X}^{\nu_\ell}_\ell, \mathbf{X}^{\nu_k}_k)\} \,|\, \mathbf{X}^{\nu_k}_k\} = \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\mathbf{X}^{\nu_k}_k\}$, where the last equality uses the tower property of conditional expectation. Also, because $\mu_k \geq \nu_k$, the filtration generated by $\mathbf{X}^{\nu_k}_k$ is a subset of the one generated by $\mathbf{X}^{\mu_k}_k$, so $\mathbb{E}\{\mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\, \mathbf{X}^{\mu_k}_k\} \,|\, \mathbf{X}^{\nu_k}_k\} = \mathbb{E}\{\overline{y}^t_{j\ell}(\mathbf{X}^{\mu_\ell}_\ell) \,|\, \mathbf{X}^{\nu_k}_k\}$. Thus, combining the last two chains of equalities yields the desired identity. Using this identity, we can follow the same argument in the proof of Theorem 3.2 to show that $Z^{\boldsymbol{\nu}}_{\mathsf{LP}} \geq Z^{\boldsymbol{\mu}}_{\mathsf{LP}}$.

Appendix E: Feasible Dual Solutions to the Fluid Approximation

Consider the two possible histories $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ with $\mu_k \geq \nu_k$ for all $k \in \mathcal{K}$. In this section, we show that if we solve the dual of problem (8) with $\boldsymbol{\eta} = \boldsymbol{\nu}$ to obtain the optimal objective value $Z^{\boldsymbol{\nu}}_{\mathsf{LP}}$, then we can use the optimal dual solution to construct a feasible solution to the dual of problem (8) with $\boldsymbol{\eta} = \boldsymbol{\mu}$. Furthermore, this feasible solution provides an objective value of $Z^{\boldsymbol{\nu}}_{\mathsf{LP}}$ to the dual of problem (8) with $\boldsymbol{\eta} = \boldsymbol{\mu}$. Therefore, when solving the dual of problem (8) with $\boldsymbol{\eta} = \boldsymbol{\mu}$, we can start the simplex algorithm with the dual feasible solution that we constructed, in which case, the solutions that we obtain in the successive iterations of the simplex algorithm can only provide upper bounds tighter than $Z^{\boldsymbol{\nu}}_{\mathsf{LP}}$. Associating the dual variables $\boldsymbol{\gamma} = (\gamma_{ik}(\mathbf{s}_k^{\eta_k}) : i \in \mathcal{L},\ k \in \mathcal{K},\ \mathbf{s}_k^{\eta_k} \in \mathcal{X}^{\eta_k})$ and $\boldsymbol{\sigma} = (\sigma^t_{jk}(\mathbf{s}_k^{\eta_k}) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}_k^{\eta_k} \in \mathcal{X}^{\eta_k})$ with the constraints, the dual of (8) is

$$Z^{\boldsymbol{\eta}}_{\mathsf{LP}} = \min_{(\boldsymbol{\gamma},\boldsymbol{\sigma}) \in \mathbb{R}_+^{(|\mathcal{L}|+|\mathcal{J}|T)\sum_{k\in\mathcal{K}}|\mathcal{X}|^{\eta_k}}} \sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{L}} c_i\,\mathbb{E}\{\gamma_{ik}(\mathbf{X}_k^{\eta_k})\} + \sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} \mathbb{E}\{\lambda^t_{jk}(\mathbf{X}^1_k)\,\sigma^t_{jk}(\mathbf{X}_k^{\eta_k})\} \tag{14}$$

$$\text{st} \quad \sum_{\ell=k}^{K}\sum_{i\in\mathcal{L}} a_{ij}\,\mathbb{E}\{\gamma_{i\ell}(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\} + \sigma^t_{jk}(\mathbf{s}_k^{\eta_k}) \ \geq\ f_j \quad \forall\, j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}_k^{\eta_k} \in \mathcal{X}^{\eta_k}.$$

One way to construct the dual of problem (8) is to express the expectations in the objective function and first constraint in this problem, respectively, as $\sum_{\mathbf{s}_k^{\eta_k} \in \mathcal{X}^{\eta_k}} \mathbb{P}\{\mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\}\, y^t_{jk}(\mathbf{s}_k^{\eta_k})$ and $\sum_{\mathbf{p}_\ell^{\eta_\ell} \in \mathcal{X}^{\eta_\ell}} \mathbb{P}\{\mathbf{X}_\ell^{\eta_\ell} = \mathbf{p}_\ell^{\eta_\ell} \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\}\, y^t_{j\ell}(\mathbf{p}_\ell^{\eta_\ell})$, in which case, problem (8) takes the form a standard linear program with finite numbers of decision variables and constraints. Noting the objective function in problem (14), we need to choose the value of the decision variable $\sigma^t_{jk}(\mathbf{s}_k^{\eta_k})$ as small as possible. Therefore, using $[x]^+ = \max\{x, 0\}$, using the constraint in problem (14), given the value of the decision variables $\boldsymbol{\gamma}$, we can obtain the optimal value of the decision variable $\sigma^t_{jk}(\mathbf{s}_k^{\eta_k})$ by setting $\sigma^t_{jk}(\mathbf{s}_k^{\boldsymbol{\eta}_k}) = [f_j - \sum_{\ell=k}^K \sum_{i\in\mathcal{L}} a_{ij}\,\mathbb{E}\{\gamma_{i\ell}(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\}]^+$. We turn our attention to the result that we would like to establish. Considering $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ that satisfy $\mu_k \geq \nu_k$ for all $k \in \mathcal{K}$, let $(\overline{\boldsymbol{\gamma}}, \overline{\boldsymbol{\sigma}})$ be an optimal solution to problem (14) with $\boldsymbol{\eta} = \boldsymbol{\nu}$. The corresponding optimal objective value is $Z^{\boldsymbol{\nu}}_{\mathsf{LP}}$. Focusing on problem (14) with $\boldsymbol{\eta} = \boldsymbol{\mu}$, we construct the solution $(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$ for this problem by setting $\widehat{\gamma}_{ik}(\mathbf{s}_k^{\mu_k}) = \overline{\gamma}_{ik}(\mathbf{s}_k^{\nu_k})$ and $\widehat{\sigma}^t_{jk}(\mathbf{s}_k^{\mu_k}) = [f_j - \sum_{\ell=k}^K \sum_{i\in\mathcal{L}} a_{ij}\,\mathbb{E}\{\overline{\gamma}_{i\ell}(\mathbf{X}_\ell^{\nu_\ell}) \,|\, \mathbf{X}_k^{\mu_k} = \mathbf{s}_k^{\mu_k}\}]^+$ for all $i \in \mathcal{L}$, $j \in \mathcal{J}$, $t \in \mathcal{T}$, $k \in \mathcal{K}$ and $\mathbf{s}_k^{\mu_k} \in \mathcal{X}^{\mu_k}$. Because $\mu_k \geq \nu_k$, if we know the value of $\mathbf{s}_k^{\mu_k}$, then we know the value of $\mathbf{s}_k^{\nu_k}$, so for each $\mathbf{s}_k^{\mu_k} \in \mathcal{X}^{\mu_k}$, we can deduce the corresponding value of $\mathbf{s}_k^{\nu_k}$ and set $\widehat{\gamma}_{ik}(\mathbf{s}_k^{\mu_k}) = \overline{\gamma}_{ik}(\mathbf{s}_k^{\nu_k})$. Thus, our approach for constructing the solution $(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$ by using the solution $(\overline{\boldsymbol{\gamma}}, \overline{\boldsymbol{\sigma}})$ is well defined. In the next lemma, we show that $(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$ is a feasible solution to problem (14) with $\boldsymbol{\eta} = \boldsymbol{\mu}$ and provides an objective value of $Z^{\boldsymbol{\nu}}_{\mathsf{LP}}$ for the latter problem.

Lemma E.1 (Constructing Dual Solutions) *Considering $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ that satisfy $\mu_k \geq \nu_k$ for all $k \in \mathcal{K}$, let $(\overline{\boldsymbol{\gamma}}, \overline{\boldsymbol{\sigma}})$ be an optimal solution to problem* (14) *with $\boldsymbol{\eta} = \boldsymbol{\nu}$, providing the optimal objective value of $Z^{\boldsymbol{\nu}}_{\mathsf{LP}}$. Defining the solution $(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$ for problem* (14) *with $\boldsymbol{\eta} = \boldsymbol{\mu}$ as*

$$\widehat{\gamma}_{ik}(\mathbf{s}_k^{\mu_k}) = \overline{\gamma}_{ik}(\mathbf{s}_k^{\nu_k}),$$

$$\widehat{\sigma}^t_{jk}(\mathbf{s}_k^{\mu_k}) = \Big[f_j - \sum_{\ell=k}^{K}\sum_{i\in\mathcal{L}} a_{ij}\,\mathbb{E}\{\overline{\gamma}_{i\ell}(\mathbf{X}_\ell^{\nu_\ell}) \,|\, \mathbf{X}_k^{\mu_k} = \mathbf{s}_k^{\mu_k}\}\Big]^+$$

for all $i \in \mathcal{L}$, $j \in \mathcal{J}$, $t \in \mathcal{T}$, $k \in \mathcal{K}$ and $\mathbf{s}_k^{\mu_k} \in \mathcal{X}^{\mu_k}$, this solution is feasible to problem (14) *with $\boldsymbol{\eta} = \boldsymbol{\mu}$ and provides an objective value of $Z^{\boldsymbol{\nu}}_{\mathsf{LP}}$.*

Proof: By the definition of the solution $(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$, we have $\widehat{\gamma}_{ik}(\mathbf{s}_k^{\mu_k}) = \overline{\gamma}_{ik}(\mathbf{s}_k^{\nu_k})$, so we immediately get $\mathbb{E}\{\widehat{\gamma}_{i\ell}(\mathbf{X}_\ell^{\mu_\ell}) \,|\, \mathbf{X}_k^{\mu_k} = \mathbf{s}_k^{\mu_k}\} = \mathbb{E}\{\overline{\gamma}_{i\ell}(\mathbf{X}_\ell^{\nu_\ell}) \,|\, \mathbf{X}_k^{\mu_k} = \mathbf{s}_k^{\mu_k}\}$. In this case, the fact that the solution

$(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$ is feasible to problem (14) with $\boldsymbol{\eta} = \boldsymbol{\mu}$ follows from the definition of the solution $(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$ in the lemma, along with the discussion just before the lemma. We proceed to checking the objective value that the solution $(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$ provides for problem (14) with $\boldsymbol{\eta} = \boldsymbol{\mu}$. Because $(\overline{\boldsymbol{\gamma}}, \overline{\boldsymbol{\eta}})$ is an optimal solution to problem (14) with $\boldsymbol{\eta} = \boldsymbol{\nu}$, by the discussion just before the lemma, we have $\overline{\sigma}^t_{jk}(\mathbf{s}^{\nu_k}_k) = [f_j - \sum_{\ell=k}^K \sum_{i \in \mathcal{L}} a_{ij} \, \mathbb{E}\{\overline{\gamma}_{i\ell}(\mathbf{X}^{\nu_\ell}_\ell) \,|\, \mathbf{X}^{\nu_k}_k = \mathbf{s}^{\nu_k}_k\}]^+$. Also, noting that $\boldsymbol{\mu}, \boldsymbol{\nu} \in \Theta$ and $\mu_k \geq \nu_k$, for stages k and ℓ that satisfy $k \leq \ell$, we have $k - \mu_k + 1 \leq k - \nu_k + 1 \leq \ell - \nu_\ell + 1$. Therefore, if stages k and ℓ satisfy $k \leq \ell$, then the distribution of $(X_{\ell-\nu_\ell+1}, \ldots, X_\ell)$ conditional on $(X_{k-\nu_k+1}, \ldots, X_k)$ is the same as the distribution of $(X_{\ell-\nu_\ell+1}, \ldots, X_\ell)$ conditional on $(X_{k-\mu_k+1}, \ldots, X_k)$. In this case, we obtain $\mathbb{E}\{\overline{\gamma}_{i\ell}(\mathbf{X}^{\nu_\ell}_\ell) \,|\, \mathbf{X}^{\nu_k}_k = \mathbf{s}^{\nu_k}_k\} = \mathbb{E}\{\overline{\gamma}_{i\ell}(\mathbf{X}^{\nu_\ell}_\ell) \,|\, \mathbf{X}^{\mu_k}_k = \mathbf{s}^{\mu_k}_k\}$ for $\ell \geq k$, so noting the definition of $\widehat{\sigma}^t_{jk}(\mathbf{s}^{\mu_k}_k)$ in the lemma, we get $\widehat{\sigma}^t_{jk}(\mathbf{s}^{\mu_k}_k) = \overline{\sigma}^t_{jk}(\mathbf{s}^{\nu_k}_k)$. Considering problem (14) with $\boldsymbol{\eta} = \boldsymbol{\mu}$, the solution $(\widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\sigma}})$ provides an objective value of

$$\begin{aligned}
&\sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{L}} c_i \, \mathbb{E}\{\widehat{\gamma}_{ik}(\mathbf{X}^{\mu_k}_k)\} + \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \mathbb{E}\{\lambda^t_{jk}(\mathbf{X}^1_k) \; \widehat{\sigma}^t_{jk}(\mathbf{X}^{\mu_k}_k)\} \\
&\qquad = \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{L}} c_i \, \mathbb{E}\{\overline{\gamma}_{ik}(\mathbf{X}^{\nu_k}_k)\} + \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} \mathbb{E}\{\lambda^t_{jk}(\mathbf{X}^1_k) \; \overline{\sigma}^t_{jk}(\mathbf{X}^{\nu_k}_k)\} \; = \; Z^{\boldsymbol{\nu}}_{\mathsf{LP}},
\end{aligned}$$

where the first equality holds because $\overline{\gamma}_{ik}(\mathbf{s}^{\nu_k}_k) = \widehat{\gamma}_{ik}(\mathbf{s}^{\mu_k}_k)$ by the definition of $\widehat{\gamma}_{ik}(\mathbf{s}^{\mu_k}_k)$ in the lemma and $\overline{\sigma}^t_{jk}(\mathbf{s}^{\nu_k}_k) = \widehat{\sigma}^t_{jk}(\mathbf{s}^{\mu_k}_k)$ by the discussion just before the chain of equalities above. ■

Thus, if we solve problem (14) with shorter histories, then we can obtain a feasible solution to problem (14) with longer histories, while matching the objective values of the two problems.

Appendix F: Approximate Policy with Varying Histories

In this section, we provide a proof for Theorem 6.2, where we give a performance guarantee for the approximate policy with varying histories obtained from problem (8). The proof almost line by line follows the same steps in Section 5, where we establish the performance guarantee for our original approximate policy obtained from problem (2). We focus on the minor changes necessary in the steps in Section 5. In our approximate policy with varying histories, fixing some $\theta = 1, \ldots, K$, we define $\overline{\boldsymbol{\eta}} = (\overline{\eta}_1, \ldots, \overline{\eta}_K)$ as $\overline{\eta}_k = 1$ for $k \leq K - \theta$, whereas $\overline{\eta}_k = k - (K - \theta)$ for $k \geq K - \theta + 1$. Letting $\overline{\boldsymbol{y}} = (\overline{y}^t_{jk}(\mathbf{s}^{\overline{\eta}_k}_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^{\overline{\eta}_k}_k \in \mathcal{X}^{\overline{\eta}_k})$ be an optimal solution to problem (8) with $\boldsymbol{\eta} = \overline{\boldsymbol{\eta}}$, if the history of the Markov chain over past $\overline{\eta}_k$ stages in stage k is $\mathbf{s}^{\overline{\eta}_k}_k$, then our approximate policy with varying histories is willing to accept a request for product j at time period t in stage k with probability $\gamma \, \overline{y}^t_{jk}(\mathbf{s}^{\overline{\eta}_k}_k) / \lambda^t_{jk}(\mathbf{s}^1_k)$. If we are willing to accept the product request and the remaining capacities allow accepting the product request, then we accept the product request.

We define the Bernoulli random variables $A^t_{jk}(\mathbf{s}^1_k)$ and G^t_{jk} as is done at the beginning of Section 5. Furthermore, we define the Bernoulli random variable $Y^t_{jk}(\mathbf{s}^{\overline{\eta}_k}_k)$ such that this random variable takes value one if and only if the approximate policy is willing to accept a request for product j at time period t in stage k given that the history of the Markov chain in stage k is $\mathbf{s}_k^{\overline{\eta}_k}$, so $\mathbb{P}\{Y_{jk}^t(\mathbf{s}^{\overline{\eta}_k}) = 1\} = \gamma\, \overline{y}_{jk}^t(\mathbf{s}_k^{\overline{\eta}_k})/\lambda_{jk}^t(\mathbf{s}_k^1)$. Letting opt be the optimal total expected revenue, by the same argument just before (3), if we can show that $\mathbb{P}\{G_{jk}^t = 1 \,|\, \mathbf{X}_k^{\overline{\eta}_k}\} \geq \beta$, then the total expected revenue of the approximate policy is at least $\gamma\,\beta\,\mathsf{opt}$. Therefore, we focus on obtaining a lower bound on the probability $\mathbb{P}\{G_{jk}^t = 1 \,|\, \mathbf{X}_k^{\overline{\eta}_k}\}$. Setting $U_{ik}^t(\mathbf{s}_k^{\overline{\eta}_k}) = \sum_{j\in\mathcal{J}} A_{jk}^t(\mathbf{s}_k^1)\, Y_{jk}^t(\mathbf{s}_k^{\overline{\eta}_k})$, we use the random variable $U_{ik}^t(\mathbf{s}_k^{\overline{\eta}_k})$ as an upper bound on the capacity consumption of resource i at time period t in stage k given that the history of the Markov chain in stage k is $\mathbf{s}_k^{\overline{\eta}_k}$. By the same argument in (4), to lower bound the probability $\mathbb{P}\{G_{jk}^t = 1 \,|\, \mathbf{X}_k^{\overline{\eta}_k}\}$, it is enough to upper bound the probability $\mathbb{P}\{\sum_{\ell=1}^k \sum_{q\in\mathcal{T}} U_{i\ell}^q(\mathbf{X}_\ell^{\overline{\eta}_\ell}) \geq c_i \,|\, \mathbf{X}_k^{\overline{\eta}_k}\}$. In this case, defining $\overline{u}_{jk}^t(\mathbf{s}_k^{\overline{\eta}_k}) = \mathbb{E}\{U_{ik}^t(\mathbf{s}_k^{\overline{\eta}_k})\}$, $f_{ik}(X_1,\ldots,X_k) = \sum_{v=1}^k \sum_{t\in\mathcal{T}} \overline{u}_{iv}^t(\mathbf{X}_v^{\overline{\eta}_v})$ and $\delta_{i\ell} = \sum_{v=1}^\ell \sum_{t\in\mathcal{T}} \frac{4}{\epsilon}\,(1-\alpha)^{\ell-v}\, \mathbb{E}\{\overline{u}_{iv}^t(\mathbf{X}_v^{\overline{\eta}_v})\}$, the proofs of Lemmas 5.1 and 5.2 go through with no modifications. Furthermore, we can follow the same argument just before Lemma 5.3 to show that $\sum_{\ell=1}^k \mathbb{E}\{\overline{u}_{i\ell}^t(\mathbf{X}_\ell^{\overline{\eta}_\ell}) \,|\, \mathbf{X}_k^{\overline{\eta}_k}\} \leq \gamma\, c_i$. Defining $W_{ik} = \sum_{\ell=1}^k \sum_{t\in\mathcal{T}} U_{i\ell}^t(\mathbf{X}_\ell^{\overline{\eta}_\ell})$ and using the last inequality, the proof of Lemma 5.3 goes through with no modifications. Lastly, we can choose the tuning parameter as in the proof of Theorem 4.1 and follow the same argument in the proof of this theorem to show that Theorem 6.2 holds.

Appendix G: Decomposition by Stages

Assuming that there exists a stage q such that $k - \eta_k + 1 \geq q$ for all $k = q,\ldots,K$, we argue that we can solve problem (8) by using (10) and (11). For stages ℓ and k that satisfy $\ell \leq q-1 < q \leq k$, note that we have $\mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\} = \mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_{k-\eta_k+1}^1 = \mathbf{s}_{k-\eta_k+1}^1\}$, where we use the fact that $k \geq k - \eta_k + 1 \geq q > \ell \geq \ell - \eta_\ell + 1$. Therefore, using the decision variable $z_{ik}(s_{k-\eta_k+1})$ to capture the sum $\sum_{\ell=1}^{q-1} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_{k-\eta_k+1}^1 = \mathbf{s}_{k-\eta_k+1}^1\}$, problem (8) is equivalent to

$$Z_{\mathsf{LP}}^{\boldsymbol{\eta}} = \max_{(\boldsymbol{y},\boldsymbol{z})\in\mathbb{R}_+^{|\mathcal{J}|T\sum_{k\in\mathcal{K}}|\mathcal{X}|^{\eta_k} + |\mathcal{L}|\,(K-q+1)\,|\mathcal{X}|}} \sum_{k\in\mathcal{K}} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} f_j\, \mathbb{E}\{y_{jk}^t(\mathbf{X}_k^{\eta_k})\} \tag{15}$$

$$\text{st} \quad \sum_{\ell=1}^{k} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\} \leq c_i \quad \forall\, i\in\mathcal{L},\ k = 1,\ldots,q-1,\ \mathbf{s}_k^{\eta_k}\in\mathcal{X}^{\eta_k}$$

$$\sum_{\ell=1}^{q-1} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_{k-\eta_k+1}^1 = \mathbf{s}_{k-\eta_k+1}^1\} = z_{ik}(\mathbf{s}_{k-\eta_k+1}^1) \quad \forall\, i\in\mathcal{L},\ k = q,\ldots,K,\ \mathbf{s}_{k-\eta_k+1}^1\in\mathcal{X}$$

$$\sum_{\ell=q}^{k} \sum_{t\in\mathcal{T}} \sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{y_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\} \leq c_i - z_{ik}(\mathbf{s}_{k-\eta_k+1}^1) \quad \forall\, i\in\mathcal{L},\ k = q,\ldots,K,\ \mathbf{s}_k^{\eta_k}\in\mathcal{X}^{\eta_k}$$

$$y_{jk}^t(\mathbf{s}_k^{\eta_k}) \leq \lambda_{jk}^t(\mathbf{s}_k^1) \quad \forall\, j\in\mathcal{J},\ t\in\mathcal{T},\ k\in\mathcal{K},\ \mathbf{s}_k^{\eta_k}\in\mathcal{X}^{\eta_k}.$$

To see that problems (8) and (15) are equivalent to each other, using the observation that we gave just before problem (15), the second and third constraints in problem (15) collectively correspond to the first constraint in problem (8) for stages $q, \ldots K$, whereas the first constraint in problem (15) corresponds to the first constraint in problem (8) for stages $1, \ldots, q-1$. Lastly, the fourth constraint in problem (15) corresponds to the second constraint in problem (8), establishing that problems (8) and (15) are equivalent to each other. We can replace the second constraint in problem (15) with an inequality constraint. In particular, if this constraint is not tight at an optimal solution, then we can decrease the value of the decision variable $z_{ik}(\mathbf{s}^1_{k-\eta_k+1})$ to obtain another optimal solution that satisfies the second constraint as equality. Because the right side of the third constraint is decreasing in $z_{ik}(\mathbf{s}^1_{k-\eta_k+1})$, the solution that we obtain is still feasible to problem (15). In this case, comparing problem (15) with problems (10) and (11), recalling that the optimal objective values of problems (10) and (11) are, respectively, $\Psi^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z})$ and $\Gamma^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z})$, if we fix the values of the decision variables $\boldsymbol{z} = (z_{ik}(\mathbf{s}^1_{k-\eta_k+1}) : i \in \mathcal{L},\ k = q, \ldots, K,\ \mathbf{s}^1_{k-\eta_k+1} \in \mathcal{X})$, then the optimal objective value of problem (15) is $\Psi^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z}) + \Gamma^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z})$. Thus, the optimal objective value of problem (15) is given by $\max_{\boldsymbol{z} \in \mathbb{R}_+^{|\mathcal{L}|(K-q+1)|\mathcal{X}|}} \left\{ \Psi^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z}) + \Gamma^{\boldsymbol{\eta}}_{\mathsf{LP}}(\boldsymbol{z}) : z_{ik}(\mathbf{s}^1_{k-\eta_k+1}) \le c_i \quad \forall\, i \in \mathcal{L},\ k = q, \ldots, K,\ \mathbf{s}^1_{k-\eta_k+1} \in \mathcal{X} \right\}$, which is, in turn, equal to the optimal objective value of problem (8).

By the preceding discussion, if there exists a stage q such that $k - \eta_k + 1 \ge q$ for all $k = q, \ldots, K$, then we can solve problem (8) by using problems (10) and (11).

Appendix H: Equivalent Dynamic Program

Considering the dynamic program in (12), noting that $Z^{\boldsymbol{\eta}}_{\mathsf{LP}}$ is the optimal objective value of problem (8) and recalling that the initial state of the Markov chain is X_1, we show that if $\eta_k = k$ for all $k \in \mathcal{K}$, then we have $Z^{\boldsymbol{\eta}}_{\mathsf{LP}} = \mathbb{E}\{V^1_1(\boldsymbol{c}, X_1)\}$. In the dynamic program in (12), the numbers of requests for the products at each time period in each stage are deterministic and we can accept a fraction of a product request, but the state of the Markov chain evolves according to its transition probabilities. In particular, if the state of the Markov chain in stage k is s, then the number of requests for product j at time period t in stage k is given by $\lambda^t_{jk}(s)$. Under the optimal policy characterized by the dynamic program in (12), we use $Y^t_{jk}(\mathbf{s}^{\eta_k}_k)$ to denote the fractional number of requests for product j that we accept at time period t in stage k, given that the history of the Markov chain up to and including stage k is $\mathbf{s}^{\eta_k}_k$. Because $\eta_k = k$, the history $\mathbf{s}^{\eta_k}_k$ captures the states of the Markov chain over stages $1, \ldots, k$. Noting that the numbers of product requests in the dynamic program in (12) are deterministic, given $\mathbf{s}^{\eta_k}_k$, the quantity $Y^t_{jk}(\mathbf{s}^{\eta_k}_k)$ is deterministic as well. The total expected revenue that we obtain under the optimal policy characterized by the dynamic program in (12) is given by $\mathbb{E}\{V^1_1(\boldsymbol{c}, X_1)\} = \sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} f_j\, \mathbb{E}\{Y^t_{jk}(\mathbf{X}^{\eta_k}_k)\}$.

First, we show that $Z^{\boldsymbol{\eta}}_{\mathsf{LP}} \ge \mathbb{E}\{V^1_1(\boldsymbol{c}, X_1)\}$. For problem (8) with $\eta_k = k$ for all $k \in \mathcal{K}$, we define the solution $\overline{\boldsymbol{y}} = (\overline{y}^t_{jk}(\mathbf{s}^{\eta_k}_k) : j \in \mathcal{J},\ t \in \mathcal{T},\ k \in \mathcal{K},\ \mathbf{s}^{\eta_k}_k \in \mathcal{X}^{\eta_k})$ as $\overline{y}^t_{jk}(\mathbf{s}^{\eta_k}_k) = Y^t_{jk}(\mathbf{s}^{\eta_k}_k)$. Under the optimal policy characterized by the dynamic program in (12), the capacity of each resource that we consume by the end of each stage cannot exceed the capacity of the resource. Therefore, we have $\sum_{\ell=1}^{k}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} a_{ij}\, Y_{j\ell}^t(\mathbf{s}_\ell^{\boldsymbol{\eta}_\ell}) \le c_i$ for all $\mathbf{s}_k^{\boldsymbol{\eta}_k} \in \mathcal{X}^{\eta_k}$. Because $\eta_k = k$ for all $k \in \mathcal{K}$, for stages ℓ and k that satisfy $\ell \le k$, if we know the value of $\mathbf{s}_k^{\eta_k}$, then we know the value of $\mathbf{s}_\ell^{\eta_\ell}$ as well, so we have $\mathbb{E}\{Y_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\} = Y_{j\ell}^t(\mathbf{s}_\ell^{\eta_\ell})$. Therefore, noting that $\overline{y}_{j\ell}^t(\mathbf{s}_\ell^{\eta_\ell}) = Y_{j\ell}^t(\mathbf{s}_\ell^{\eta_\ell})$, the last inequality is equivalent to $\sum_{\ell=1}^{k}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} a_{ij}\, \mathbb{E}\{\overline{y}_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\} \le c_i$. Thus, the solution $\overline{\boldsymbol{y}}$ satisfies the first constraint in problem (8). Similarly, under the optimal policy characterized by the dynamic program in (12), the number of requests that we accept for the products at each time period in each stage cannot exceed the expected number of product requests. Thus, we have $Y_{jk}^t(\mathbf{s}_k^{\eta_k}) \le \lambda_{jk}^t(\mathbf{s}_k^1)$, so noting that $\overline{y}_{jk}^t(\mathbf{s}_k^{\eta_k}) = Y_{jk}^t(\mathbf{s}_k^{\eta_k})$, the solution $\overline{\boldsymbol{y}}$ satisfies the second constraint in problem (8) as well. In this case, because the solution $\overline{\boldsymbol{y}}$ is feasible to problem (8), the optimal objective value of this problem satisfies $Z_{\mathsf{LP}}^{\boldsymbol{\eta}} \ge \sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} f_j\, \mathbb{E}\{\overline{y}_{jk}^t(\mathbf{X}_k^{\eta_k})\} = \sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} f_j\, \mathbb{E}\{Y_{jk}^t(\mathbf{X}_k^{\eta_k})\} = \mathbb{E}\{V_1^1(\boldsymbol{c}, X_1)\}$, where the first equality uses the definition of $\overline{\boldsymbol{y}}$ and the second equality is by the discussion at the end of the previous paragraph.

Second, we show that $Z_{\mathsf{LP}}^{\boldsymbol{\eta}} \le \mathbb{E}\{V_1^1(\boldsymbol{c}, X_1)\}$. Considering problem (8) with $\eta_k = k$ for all $k \in \mathcal{K}$, we use $\overline{\boldsymbol{y}} = (\overline{y}_{jk}^t(\mathbf{s}_k^{\eta_k}) : j \in \mathcal{J},\; t \in \mathcal{T},\; k \in \mathcal{K},\; \mathbf{s}_k^{\eta_k} \in \mathcal{X}^{\eta_k})$ to denote an optimal solution to this problem. In the dynamic program in (12), we follow the policy that accepts $\overline{y}_{jk}^t(\mathbf{s}_k^{\eta_k})$ requests for product j at time period t in stage k, given that the history of the Markov chain up to and including stage k corresponds to $\mathbf{s}_k^{\eta_k}$. We refer to this policy as the proxy policy. Because $\eta_k = k$ for all $k \in \mathcal{K}$, if stages ℓ and k satisfy $\ell \le k$, then we have $\mathbb{E}\{\overline{y}_{j\ell}^t(\mathbf{X}_\ell^{\eta_\ell}) \,|\, \mathbf{X}_k^{\eta_k} = \mathbf{s}_k^{\eta_k}\} = \overline{y}_{j\ell}^t(\mathbf{s}_\ell^{\eta_\ell})$. In this case, for each $\mathbf{s}_k^{\eta_k} \in \mathcal{X}^{\eta_k}$, by the first constraint in problem (8), we get $\sum_{\ell=1}^{k}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} a_{ij}\, \overline{y}_{jk}^t(\mathbf{s}_\ell^{\eta_\ell}) \le c_i$, so the capacity of each resource consumed by the proxy policy does not exceed the capacity of the resource. For each $\mathbf{s}_k^{\eta_k} \in \mathcal{X}^{\eta_k}$, by the second constraint in problem (8), we have $\overline{y}_{jk}^t(\mathbf{s}_k^{\eta_k}) \le \lambda_{jk}^t(\mathbf{s}_k^1)$, so the number of requests accepted by the proxy policy for each product at each time period in each stage does not exceed the expected number of product requests. Thus, the proxy policy is feasible. In the dynamic program in (12), if we follow the optimal policy, then we obtain a total expected revenue of $\mathbb{E}\{V_1^1(\boldsymbol{c}, X_1)\}$, whereas if we follow the proxy policy then we obtain a total expected revenue of $\sum_{k\in\mathcal{K}}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} f_j\, \mathbb{E}\{\overline{y}_{jk}^t(\mathbf{X}_k^{\eta_k})\} = Z_{\mathsf{LP}}^{\boldsymbol{\eta}}$. Therefore, we have $\mathbb{E}\{V_1^1(\boldsymbol{c}, X_1)\} \ge Z_{\mathsf{LP}}^{\boldsymbol{\eta}}$.