RC24452 (W0712-011) December 4, 2007
Mathematics

IBM Research Report

Provably Near-Optimal LP-Based Policies for Revenue Management in Systems with Reusable Resources

Retsef Levi
Sloan School of Management
MIT
Cambridge, MA 02139

Ana Radovanovic
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Provably Near-Optimal LP-Based Policies for Revenue Management in Systems with Reusable Resources

Retsef Levi* Ana Radovanovic†

21 February 2007

Abstract

Motivated by emerging applications in workforce management, we consider a class of revenue management problems in systems with reusable resources. The corresponding applications are modeled using the well-known *loss network systems*.

We use an extremely simple linear program (LP) that provides an upper bound on the best achievable expected long-run revenue rate. The optimal solution of the LP is used to devise a conceptually simple control policy that we call the *class selection policy* (CSP). Moreover, the LP is used to analyze the performance of the CSP policy. We obtain the first control policy with uniform performance guarantees. In particular, for the model with single resource and uniform resource requirements, the CSP policy is guaranteed to have expected long-run revenue rate that is at least half of the best achievable. More generally, as the ratio between the capacity of the system and the maximum resource requirement grows to infinity, the CSP policy is asymptotically optimal, regardless of *any* other parameter of the problem. The asymptotic performance analysis that we obtain is more general than existing results in several important dimensions. It is based on several novel ideas that we believe will be useful in other settings.

*retsef@mit.edu. Sloan School of Management, MIT, Cambridge, MA, 02139. Part of this research was conducted while the author was a postdoctoral fellow at the IBM, T. J. Watson Research Center.

†aradovan@us.ibm.com. IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598.

1 Introduction

In this paper, we consider a class of revenue management problems that arise in systems with reusable resources. The paper is motivated by several application domains, and, in particular, by several emerging applications in *workforce management*. In many industries, a significant part of the workforce is hired ad-hoc to perform a specific project. Thus, professional manpower services is a growing market that brings up new challenges in workforce revenue management. Similar problems arise in large corporations, such as IBM that need to manage their internal workforce in the face of dynamic and evolving tasks. The major issue in all of these scenarios is how to manage capacitated resources over time in dynamic environments with many uncertainties, specifically, how to choose the most profitable customers/projects to maximize the resulting revenue. Other notable applications are hotel room booking and car rentals.

Typically, these systems consist of several capacitated resources that are used to serve multiple classes of customers, each of which has different characteristics, such as arrival rate, price, resource and service time requirements. The goal is to devise a policy that selects profitable customers and maximizes the resulting revenue. There are three key characteristics of these systems. The first characteristic is the reusability of resources. That is, resources that are allocated to serve a certain customer/project will become available to serve other customers after the service/project is over. The second characteristic is that the decision whether to serve customers should be made upon their arrival. In particular, if a customer is not served upon arrival, either because the system decides she/he is not profitable enough, or because the available capacity in the system is not sufficient to satisfy her/him, she/he is assumed to be lost and leaves the system. (In many of the corresponding applications customers are not willing to wait or only willing to wait a very short time relative to the service times.) The third characteristic is that the arrival process of customers, as well as their service time requirements, are stochastic. This generates stochastic optimization models that are usually computationally challenging.

In this paper, we model the corresponding revenue management problems as *loss network models*. These are well-known models that have been introduced over four decades ago, and have been studied extensively in the context of communication networks (see, for example, the survey paper by Kelly ([13])). The classical loss network model consists of a system with several capacitated resources that faces multiple classes of customers. Customers of different classes arrive according to mutually independent homogenous Poisson processes, each of which requires a certain combination of resources for a time that is a-priori random (with finite mean), and is willing to pay a certain price per unit of service time. Customers must be served upon their arrival, or otherwise, they leave the system. If a customer is served, the required combination of resources must be engaged for the (random) duration of the service time, and can not be used by other customers until the service is over. The system may deny service from customers in order to keep the capacity free for more profitable future customers. A customer can be served only if at the moment of arrival the available capacity in the system is sufficient to satisfy her/his specific requirements. The goal is to find an admission policy that maximizes the long-run revenue rate. Like many stochastic optimization models, one can formulate the problem using a dynamic programming approach. However, even in special cases (e.g., with exponentially distributed service times), the resulting dynamic program seems computationally intractable as the corresponding state-space grows very fast. (This is known as the 'curse of dimensionality'.) Thus, finding provably good policies is a very challenging task.

We first focus on *revenue management model with single reusable resource*, where there is only a single resource in the system that is used to serve multiple classes of customers as described above. (In the literature on loss network models this is sometimes called *the stochastic knapsack problem*.) We use a simple knapsack-type linear program (LP) that provides an upper bound on the expected long-run revenue rate. The LP can be easily solved, and the optimal solution is used to construct a conceptually simple admission control policy for the original model; the policy is called the *class selection policy* (CSP). The LP optimal solution guides the policy to select the more profitable classes. The CSP policy admits all the customers of the selected (profitable) classes as long as capacity permits, and always rejects customers from other classes.

Moreover, the LP is used to analyze the performance of the CSP policy. We use the fact that the CSP policy induces a stochastic process that can be reduced to a classical loss network model. Facilitating the results from [21, 11, 10, 4, 23], which characterize the stationary distributions of the corresponding loss network models, we are able to develop explicit expressions for the resulting *blocking probabilities* induced by the CSP policy. That is, for each one of the profitable classes, we derive an exact expression for the stationary probability that a customer arrives at some random time, and the available capacity in the system is not sufficient to satisfy her/his requirement. We then bound the customer blocking probabilities and analyze their asymptotic behavior as the capacity of the system grows large. In particular, the bounds on the blocking probabilities are used to obtain uniform and asymptotic performance guarantees.

For the case, where all the classes requirements are identical, we obtain an explicit lower bound on the ratio between the expected long-run revenue rate of the CSP policy and the best achievable rate. The bound is a function *only* of the capacity of the system, regardless of the other parameters, such as arrival rates, number of classes, prices and service time distributions. It is shown that this bound is at least 0.5, uniformly for all capacity values, and that it approaches 1 as the capacity of the system grows to infinity. That is, the CSP policy is guaranteed to have expected long-run revenue rate that is at least half of the best achievable, and it is asymptotically optimal as the capacity of the system grows large. To the best of our knowledge, this is the first proof of uniform performance guarantees that hold for all capacity values.

These results are then extended to the more general case with arbitrary, possibly non-identical, resource requirements. In this case, the underlying combinatorics of the blocking states is more complex, and hence, it is much harder to bound the corresponding blocking probabilities. In particular, the lower bounds on the ratio between the long run average revenue rate of the CSP policy and the best achievable rate depend *only* on the *ratio* between the capacity of the system and the *maximum resource requirement of a class* raised to the power of seven. Moreover, if the ratio between the capacity and the maximum resource requirement raised to the power of seven grows to infinity, the CSP policy is asymptotically optimal, regardless of the other parameters of the problem. For each fixed value of maximum resource requirement, it is possible to derive uniform performance guarantees that depend only on the system capacity value.

The CSP policy and the asymptotic performance analysis can be extended to the *revenue management model with multiple reusable resources* as long as the number of resources is bounded. In this case, the resulting linear program is a packing-type LP that is again very easy to solve.

Finally, we incorporate *static pricing* to the single resource model. In this model we first determine the respective prices that each class is charged. The respective arrival rate of each class depends on the price it is charged. After the prices are set, we wish to find the best admission control policy that maximizes the expected long-run revenue rate. The CSP policy and its performance analysis can be extended to this more general model. However, the policy is derived based on a non-linear program (NLP). We show how to simplify the resulting NLP, and discuss several scenarios in which it can be solved efficiently.

As we already mentioned loss networks have been studied extensively in the context of communication networks, and there is a huge body of literature. The study of loss networks has been focused on two major issues, the *study of heuristics* and *sensitivity analysis*.

Since it is apparent that computing optimal policies is likely to be intractable, researchers have proposed different heuristics, studied their properties and analyzed their performance (see, for example, [17, 19, 14, 13, 8, 18, 7]). The knapsack-type LP used in this paper has been discussed by several researchers (see for example, [14, 8]). In fact variants of the CSP policy have been discussed by Key [14] and Kelly [13], who proposed the *randomized thinning policy*. Moreover, Key [14] has shown that the variant of the CSP policy for the single resource case is asymptotically optimal, but in a very specific *heavy traffic regime*. (We discuss this further in the next paragraph.) Iyengar and Sigman [9] have also used an LP identical to the one used in this paper to devise a heuristic for the same model. However, the policy they have proposed is very different than ours. Specifically, they have used the LP to generate a 'desirable' target performance mode, and, then, exploit exponential penalty functions to maintain the system as close as possible to the target mode. Another policy that has been studied extensively is the *trunk reservation policy*. According to this policy, each class of customers is associated with a *trunk reservation level*, and a customer of that class is admitted to the system only if upon arrival the available capacity in the system exceeds the corresponding trunk reservation level. Key [14] has shown that for the single resource model this policy is asymptotically optimal in the corresponding heavy traffic regime. (It is interesting to note that the CSP policy can be viewed as a special trunk reservation policy, where the trunk reservation of a class is either 0 or the capacity of the system.) Other mathematical-programming-based approximation have been used to study models similar to the one discussed in this paper (see, for example, Adelman [2]).

However, the performance analysis of the policies described above and their asymptotic optimality are obtained only in very specific regimes, usually called *heavy traffic regimes*. Specifically, the capacity and the arrival rates are scaled simultaneously at the same (linear) rate, while all the other parameters of the problem, such as service time distributions, resource requirements, number of classes and price rates are kept fixed. In some scenarios one also needs to assume that the service times are exponentially distributed. Moreover, in models with multiple resources it was usually required to assume a very specific structure of the different class resource requirements. The performance of the corresponding policies is then analyzed in the resulting limiting regimes. While these might be reasonable assumptions in the context of communication networks, they are less likely to hold in the application domains that motivate this paper. In contrast, our analysis provides uniform performance guarantees that hold for *any capacity value*. Our asymptotic performance analysis holds under very general assumptions. Specifically, the asymptotic analysis holds for general service time distributions, and only requires that the ratio between the capacity of the system and the maximum resource requirement (raised to the power of seven) grows to infinity, allowing all the other parameters of the problem to change arbitrarily. In addition, we can easily characterize the corresponding rate of convergence. Thus, as a by-product of our work, we obtain generalizations for some of the results in [14, 13].

The second major issue that has been studied is the sensitivity of the corresponding loss system to changes in various parameters, especially the capacity and the arrival rates (see, for example, [20]). The main effort has been to study changes in the resulting blocking probabilities. (By blocking we refer to the event that a customer arrives at some random time, and can not be served upon arrival because the available capacity in the system is not sufficient.) Since computing blocking probabilities is known to be $\#P$-Hard [16], there have been efforts to propose methods to approximately compute blocking probabilities and bound them (see, for example, [6, 21, 10, 4, 23, 12, 13, 25]). We note that there have been several approaches that use linear and non-linear programs to bound blocking probabilities (see, for example, [15, 3]). One of the key features in our performance analysis is the bounds that we develop on the corresponding blocking probabilities induced by the CSP policy. The techniques that we use are significantly different than the ones used in the existing literature, and we believe that they will have applications in other settings. It is interesting to note that our asymptotic analysis in the multiple resources case captures the specific regime analyzed by Kelly in [13], where he analyzed the blocking probabilities in loss network models with multiple resources.

The rest of the paper is organized as follows. In Section 2, we provide the mathematical formulation of the revenue management model with a single reusable resource. In Section 3, we develop the LP and describe the CSP policy. In Section 4, we discuss the performance analysis of the CSP policy. Finally, in Section 5, we discuss the extensions to multiple resources and static pricing.

2 Model Formulation

In this section, we provide a mathematical formulation of the revenue management model with a single reusable resource discussed in this paper. Consider a system with a single resource pool of integer capacity $C < \infty$ that is facing demands from M different classes of customers. The customers of each class $i = 1, \ldots, M$, arrive according to an independent Poisson process with rate λ_i. Each class-i customer requests $A_i \in \mathbb{Z}_+$ units of the resource for a certain period of time that is a-priori random and has finite mean μ_i. During the time a class-i customer is served, the requested A_i units can not be used by other customers; after the service is over, the units become available again to serve other customers. (While a customer is served we only know the conditional distribution of the residual service time of this customers.) In particular, we allow generally distributed service times, and assume that service times are independent of the customer arrival process and among different customers. If served, a class-i customer is willing to pay r_i dollars per time unit of service. A customer can be served only if the available capacity in the system at the moment of her/his arrival is sufficient to satisfy its requirement. That is, a class-i customer can be served only if there are currently at least A_i units available in the system. However, customers can be rejected even if the available capacity is sufficient to serve them. (Rejecting a customer now possibly enables serving more profitable customers in the future.) The assumption is that customers that are not satisfied upon arrival, either due lack of sufficient capacity or because they are rejected, are lost and leave the system immediately. The goal is to find an admission policy that maximizes the expected long-run revenue rate.

For each policy π, let $R_\pi(T)$ be the revenue achieved by policy π over the interval $[0, T]$. Next define

$\mathcal{R}(\pi)$, the expected long-run revenue rate of a policy π as

$$\mathcal{R}(\pi) =: \lim_{T\to\infty} \frac{\mathbb{E}_\pi[R_\pi(T)]}{T}, \tag{1}$$

where the expectation $\mathbb{E}_\pi$ is taken with respect to probability measure induced by policy π.

At any point of time t, the state of the system is specified by the class of each customer currently being served, as well as the time that elapsed from the moment of her/his arrival. Without loss of generality, we restrict attention to *state-dependent policies*. That is, policies that are represented as measurable functions from the state-space defined above to actions. (The actions are whether to accept a class-i customer in the current state, for each $i = 1, \ldots, N$.) It is straightforward to verify that for each feasible policy, there exists a state-dependent policy that achieves at least the same expected long-run average revenue. Note that each state-dependent policy induces a Markov process over the state-space. Moreover, using analogous arguments to those used in Theorems 4, 5 and 6 of [21], one can show that conditions in the statement of Theorem 1 in [21] hold and, therefore, the Markov process homogeneous in time that is induced by the state dependent policy has a unique stationary distribution which is ergodic. Thus, for each state-dependent policy π, the limit in (1) above is well-defined and the expectation in the numerator of (1) can be omitted.

As in most stochastic control optimization models, dynamic programming framework is the most common way to formulate the problem. However, it is straightforward to see that the corresponding state-space of the underlying dynamics program is very large even for simple special cases. In particular, it seems computationally intractable to solve the dynamic program and compute an optimal policy. For example, consider the special case with service times that follow exponential distributions. The state in this case is specified merely by the number of customers of each class currently being served. However, the corresponding state-space grows exponentially fast in the capacity C, and becomes computationally intractable.

3 LP-Based Approach

In this section, we construct a simple linear program (LP) that provides an upper bound on the achievable long-run average revenue rate. Our LP is identical to the one used by Key [14] and Iyengar and Sigman [9], and it is also similar to the one used by Adelman [2] in the queueing networks framework with unit resource requirements. We shall show how to use the optimal solution of the LP to construct a simple admission control policy that is called *class selection policy* (CSP). Moreover, the LP will be used to analyze the performance of the proposed CSP policy. In particular, we shall show that the expected long-run revenue rate of the CSP policy is guaranteed to be near-optimal for any capacity value C (above a certain threshold) and that the policy is asymptotically optimal. More specifically, we shall show that

$$\mathcal{R}(CSP) \geq \beta(C/A^7)\mathcal{R}(OPT), \tag{2}$$

where OPT denotes the optimal control policy, A denotes the maximum resource requirement (i.e., $A = \max_{i=1,\dots,M} A_i$) and $\beta(C/A^7)$ is a positive scalar for each value of C/A^7. (If no optimal policy exists, we think about $\mathcal{R}(OPT)$ as the corresponding supreme of the achievable expected revenue rate.) Furthermore, if the resource requirements of all classes are identical, i.e., $A_i = 1$ for each $i = 1, \dots, M$, then $\beta(C) \geq 1/2$ for all $C \geq 1$. Moreover, as C grows larger the CSP policy becomes asymptotically optimal. That is, $\beta(C)$ approaches 1 as C grows to infinity, and this occurs *irrespective* of other model parameters, such as the number of classes, arrival rates, service durations and price rates. For the model with non-uniform resource requirements, we establish a similar results. The CSP policy is asymptotically optimal as C/A^7 grows large. That is, we show that $\beta(C/A^7)$ approaches 1 as C/A^7 grows to infinity. In addition, for each fixed value A, it is possible to compute a threshold $\bar{C}$, such that for each capacity values $C \geq \bar{C}$, the CSP policy is guaranteed to have expected long-run revenue rate at least half of the best achievable.

3.1 An LP

We have already discussed in Section 2 that any state-dependent policy induces a Markov process on the state-space of the system with a unique stationary distribution that is ergodic. In particular, for each class i and a given state-dependent policy π, there exists a stationary probability $\alpha_i^{(\pi)}$ for accepting a class-i customer, which is equal to the long-run proportion of accepted customers of class i while running the policy π. Thus, any state-dependent policy π is associated with the stationary probabilities $\alpha_1^{(\pi)}, \alpha_2^{(\pi)}, \dots, \alpha_M^{(\pi)}$. Furthermore, by applying Little's law, we can use the stationary probability $\alpha_i^{(\pi)}$ to express the expected number of class-i customers being served in the system under state-dependent policy π as $\lambda_i \mu_i \alpha_i^{(\pi)} = \rho_i \alpha_i^{(\pi)}$. (Note that $\rho_i = \lambda_i \mu_i$ is the expected number of class-i customers being served in the system with infinite capacity and no rejections; in the context of communication networks it is usually called the *traffic intensity*.) It follows that the expected long-run revenue rate of policy π can be expressed as

$$\sum_{i=1}^{M} r_i \alpha_i^{(\pi)} \lambda_i \mu_i = \sum_{i=1}^{M} r_i \alpha_i^{(\pi)} \rho_i.$$

Similarly, the overall expected long-run number of resource units being engaged to serve customers can be expressed as

$$\sum_{i=1}^{M} \alpha_i^{(\pi)} \rho_i A_i.$$

The physical constrains of the system discussed in this paper imply that, for any feasible policy, it is not possible to find more than C units being used to serve customers. We conclude that $\sum_{i=1}^{M} \alpha_i^{(\pi)} \rho_i A_i \leq C$ for any feasible state-dependent policy π. This suggests the following LP:

$$\max_{\alpha_1,\ldots,\alpha_M} \sum_{i=1}^{M} r_i \alpha_i \rho_i \tag{3}$$

$$\text{s.t.} \sum_{i=1}^{M} \alpha_i \rho_i A_i \leq C \tag{4}$$

$$0 \leq \alpha_i \leq 1, \ \forall 1 \leq i \leq M. \tag{5}$$

Note that, for each feasible state-dependent policy π, the corresponding vector $\alpha^{(\Pi)} = (\alpha_1^{(\pi)}, \alpha_2^{(\pi)}, \ldots, \alpha_M^{(\pi)})$ is a feasible solution for the LP defined by (3)-(5) above, and has objective value that is equal to the expected long-run revenue rate of policy π. In fact, the LP enforces the capacity constraint of the system only in expectation, while in the original problem this constraint has to hold for every sample path. It follows that the LP defined by (3)-(5) relaxes the original problem and provides an upper bound on the optimal expected long-run revenue rate. Moreover, the LP defined above is a *knapsack* LP. Thus, it can be solved optimally by applying the following greedy rule. Without loss of generality, assume that classes are renumbered such that $r_1/A_1 \geq r_2/A_2 \geq \cdots \geq r_M/A_M$. Then, for each $i = 1, \ldots, M$, we set $\alpha_i = 1$ as long as constraint (4) is satisfied. In particular, the optimal solution has the following structure: $\alpha_1 = \alpha_2 = \cdots = \alpha_{M'-1} = 1$, for some $1 \leq M' \leq M$; for M' the corresponding value of $\alpha_{M'}$ is possibly a fraction, i.e., $0 < \alpha_{M'} \leq 1$; and, for $i = M'+1, \ldots, M$, we have $\alpha_i = 0$. Next we shall use the optimal solution of the knapsack LP defined above to construct an extremely simple admission policy, and show that its expected long-run revenue rate is guaranteed to be near-optimal in the sense defined in (2) above.

3.2 The Class Selection Policy

Let $\alpha^* = (\alpha_1^*, \ldots, \alpha_M^*)$ be the optimal solution of the knapsack LP defined by (3)-(5) above. We propose the following policy that we call *class selection policy*:

Consider an arrival of a class-i customer ($i = 1, \ldots, M$):

- For each $i = 1, \ldots, M'-1$, *accept* the customer as long as the available capacity in the system upon arrival is sufficient (i.e., greater than A_i units);
- If $i = M'$, *accept* with probability $0 < \alpha_{M'} \leq 1$ and as long as there available capacity in the system upon arrival is sufficient (i.e., greater than $A_{M'}$ units);

- For each $i = M' + 1, \ldots, M$, *reject*.

The CSP policy has a very simple structure. It always admits customers from the classes for which the corresponding value α_i^* in the optimal LP solution equals to 1, as long as capacity permits; it never admits customers from classes for which the corresponding value α_i^* equals to 0; it flips a coin for the possibly one class with fractional value $\alpha_{M'}^*$. The CSP policy is conceptually very intuitive in that it splits the classes into profitable and non-profitable that should be ignored.

4 Performance Analysis

In this section, we analyze the performance of the CSP policy. The special properties of the CSP policy induce a well-structured stochastic process. Each class $i = 1, \ldots, M$ generates an independent Poisson arrival stream with respective rate $\alpha_i^* \lambda_i$. Thus, each class i with $\alpha_i^* = 1$ generates the original process, each class i with $\alpha_i^* = 0$ can be ignored, and possibly one class with fractional $0 < \alpha_{M'}^* < 1$ generates a thinned Poisson process. Moreover, the induced stochastic process can be described as a classical loss network model with a single resource. There are C servers that are used to serve M' independent Poisson streams of requests. The requests of stream (class) i arrive at rate λ_i, and each requires A_i servers for some random service time with mean μ_i; whenever a request arrives and the number of idle servers is not sufficient to serve it, the request is lost and leaves the system. It can be easily verified that the loss network model described above is identical to the stochastic process induced by the CSP policy.

One of the natural questions studied in the context of loss networks is what is the stationary *blocking probability* of a given request. That is, what is the stationary probability that a certain request arrives at some random time and the number of idle servers in the system is not sufficient to serve it. The latter question is directly related to the performance analysis of the CSP policy. Focus on the classes with positive α_i^*, say there are M' of them. Without loss of generality, assume that there is no fractional variable in the optimal solution $\alpha*$, i.e., for each $i = 1, \ldots, M'$, $\alpha_i^* = 1$. (If $\alpha_{M'}^*$ is fractional, we think of class M' as having an arrival rate $\lambda'_{M'} = \alpha_{M'}^* \lambda_i$ and then eliminate the fractional variable from α^*.) For each $i = 1, \ldots, M'$, let P_i be the stationary probability of rejecting a class-i customer under the CSP policy. It is straightforward to verify that probability P_i is equal to the stationary blocking probability of a stream-i request in the corresponding loss network model described above. Specifically, let X_i be the the stationary (random) number of class-i (stream-i) customers being served in the system at some random time under the

CSP policy. Then by the PASTA property (see [24]), it follows that, for each $i = 1, \dots, M'$:

$$P_i = \mathbb{P}\left(\sum_{k=1}^{M'} X_k A_k > C - A_i\right). \tag{6}$$

Moreover, since the corresponding stochastic process is ergodic, it follows that the long-run proportion of class-i customers being served equals to $1 - P_i$. Thus, the expected long-run revenue rate of the CSP policy can be expressed as

$$\sum_{i=1}^{M'} r_i \alpha_i^* \rho_i (1 - P_i) = \sum_{i=1}^{M'} r_i \rho_i (1 - P_i). \tag{7}$$

However, $\sum_{i=1}^{M'} r_i \alpha_i^* \rho_i$ is the optimal value of the LP, which is an upper bound on the achievable expected long-run revenue rate for any feasible policy. Thus, a key aspect of the performance analysis of the CSP policy is lower bounding probabilities $1 - P_i$, or equivalently, upper bounding the probabilities P_i. In particular, (7) above implies that any uniform constant bound on these probabilities can be directly translated to a performance guarantee of the CSP policy. Specifically, if $1 - P_i \geq \beta$ for each $i = 1, \dots, M'$, it follows that

$$\mathcal{R}(CSP) = \sum_{i=1}^{M'} r_i \alpha_i^* \rho_i (1 - P_i) \geq \sum_{i=1}^{M'} r_i \alpha_i^* \rho_i \beta \geq \beta \mathcal{R}(OPT).$$

In the rest of the analysis, we shall establish upper bounds on the corresponding blocking probabilities $P_1, \dots, P_{M'}$, and analyze their asymptotic behavior.

Recall that the CSP policy induces a stochastic process that can be reduced to a classical loss network with a single resource. A central tool in our analysis is the result of Burman et. al. [4], who characterized the stationary probabilities for general loss network models (see Theorem 2 in [4]). In fact, the result of Burman et. al. [4] implies that the stationary probabilities of the corresponding loss network model can be expressed through the counterpart system with no capacity constrains. That is, consider an infinite capacity system that faces Poisson streams of requests/customers of class $1, \dots, M'$, with respective rates $\lambda_1, \dots, \lambda_{M'}$ and service time distributions with respective means $\mu_1, \dots, \mu_{M'}$, and accept all the requests/customers. In particular, for each $i = 1, \dots, M'$, let Y_i be the stationary number of class-i customers being served in the infinite capacity system described above. Then, for each $n = 0, \dots, C$, we have

$$\mathbb{P}\left(\sum_{k=1}^{M'} A_k X_k = n\right) = \frac{\mathbb{P}\left(\sum_{k=1}^{M'} A_k Y_k = n\right)}{\mathbb{P}\left(\sum_{k=1}^{M'} A_k Y_k \leq C\right)}. \tag{8}$$

Equation (8) is very useful, since the variables $Y_1, \dots, Y_{M'}$ are independent of each other, and for each $i = 1, \dots, M'$, the random variable Y_i follows a Poisson distribution with parameter ρ_i. Thus, in conjunction with (6) above, we get that, for each $i = 1, \dots, M'$

$$P_i = \frac{\sum_{n=C-A_i}^{C} \sum_{y \in \mathcal{Y}(n)} \frac{\rho_1^{y_1}}{y_1!} \cdots \frac{\rho_{M'}^{y_{M'}}}{y_{M'}!}}{\sum_{n=0}^{C} \sum_{y \in \mathcal{Y}(n)} \frac{\rho_1^{y_1}}{y_1!} \cdots \frac{\rho_{M'}^{y_{M'}}}{y_{M'}!}}, \tag{9}$$

where we define $\mathcal{Y}(n) = \{y \in \mathbb{Z}_+^{M'} \ : \ \sum_{j=1}^{M'} A_j y_j = n\}$.

Next we use the explicit expression in (9) and the LP Constraint (4) to uniformly bound customer rejection probabilities, and obtain uniform and asymptotic performance guarantees of the CSP policy.

4.1 Identical Resource Requirements

In this section, we discuss the special case, in which the resource requirements of all classes are identical; without loss of generality, $A_i = 1$, for each $i = 1, \dots, M$. In this special case, in view of (6), the rejection probabilities are identical for all classes $1, \dots, M'$. Specifically, $P_i = \mathbb{P}(\sum_{k=1}^{M'} X_k = C)$, where X_k, as before, denotes the stationary (random) number of class-k customers being served in the system under the CSP policy.

Lemma 4.1 *Consider the revenue management model with a single reusable resource and identical resource requirements, i.e.,* $A_1 = A_2 = \cdots = A_M = 1$. *Without loss of generality, assume that there is no fractional variable in the solution* α^**, and that for each class* $i = 1, \dots, M'$ *(*$M' \leq M$*), we have* $\alpha_i^* = 1$. *Then, for each* $i = 1, \dots, M'$*, the blocking probabilities* P_i *has the following properties:*

(i) $P_i \leq 0.5$ *for all capacity values* $C \geq 1$.

(ii) The probability P_i *diminishes to 0 as* C *grows to infinity, regardless of other parameters of the problem such as the price rates, service time distributions and number of classes.*

Proof: Recall that, in view of discussion above, probabilities P_is are the same and equal to blocking probabilities in the corresponding loss network model described above. That is, $P_i = \mathbb{P}(\sum_{k=1}^{M'} X_k = C)$. We use the following identity

$$\sum_{y \in \mathcal{Y}(n)} \frac{\rho_1^{y_1}}{y_1!} \cdots \frac{\rho_{M'}^{y_{M'}}}{y_{M'}!} = \frac{(\rho_1 + \cdots + \rho_{M'})^n}{n!},$$

that holds for each $n = 0, \dots, C$. In conjunction with (9) above we obtain

$$P_i = \frac{\frac{(\rho_1 + \cdots + \rho_{M'})^C}{C!}}{\sum_{n=0}^{C} \frac{(\rho_1 + \cdots + \rho_{M'})^n}{n!}}.$$

Consider the function $f(z) = \frac{z^C/C!}{\sum_{k=0}^{C} z^k/k!}$. By looking on the derivative it is straightforward to check that $f(z)$ is increasing in z on $(0, C]$. This and Constraint (4) imply that

$$P_i \leq \frac{\frac{C^C}{C!}}{\sum_{n=0}^{C} \frac{C^n}{n!}}. \tag{10}$$

However, from the properties of the Poisson distribution, it follows that function $g(C) \triangleq \frac{e^{-C}\frac{C^C}{C!}}{\sum_{n=0}^{C} e^{-C}\frac{C^n}{n!}}$ is decreasing in C. One way to show this is by replacing $x!$ using identity 6.1.38 from [1] first, and, then, showing that function $\log(g(C))$ is decreasing since its derivative with respect to C is negative for all $C \geq 1$. More specifically, for any $x > 0$,

$$x! = \sqrt{2\pi} x^{x+1/2} e^{-x+\frac{\theta}{12x}}, \tag{11}$$

for some $\theta \in (0, 1)$. Then, it is not hard to show that, for all $C \geq 1$,

$$\frac{d}{dC} \log(g(C)) \leq -\frac{1}{2C} + \frac{1}{12C^2} < 0.$$

Thus, in view of the previous observation, the maximum value of $g(C)$ is $g(1) = 1/2$. Moreover, expression $\sum_{n=0}^{C} \frac{C^n}{n!} e^{-C}$ approaches 0 as C grows to infinity (see 6.5.34 of [1]). Using Stirling approximation, we can characterize the rate of convergence. Specifically, we know that

$$\frac{\frac{C^C}{C!}}{\sum_{k=0}^{C} \frac{C^k}{k!}} \sim \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{C}}$$

for large values of C. This concludes the proof of the lemma. ■

We have obtained the following theorem.

Theorem 4.2 *Consider the revenue management model with a single reusable resource and identical resource requirements, i.e.,* $A_1 = A_2 = \cdots = A_M = 1$. *Then, for each capacity value* C *the long-run expected revenue rate of the CSP policy is guaranteed to be at least* $g(C) = \frac{e^{-C}\frac{C^C}{C!}}{\sum_{n=0}^{C} e^{-C}\frac{C^n}{n!}}$ *of the best achievable long-run expected revenue rate. In particular, the long-run expected revenue rate of the CSP policy is guaranteed to be at least half of the best achievable, for all capacity values* $C \geq 1$. *Moreover, the CSP policy is asymptotically optimal as capacity* C *grows to infinity.*

In light of Theorem 4.2 above, we note that if $C = 1$, it is usually straightforward to compute the optimal policy. In fact, one can improve the performance guarantee of the CSP policy by solving the problem exactly for small values of C, and apply the LP-based policy only for values of C, where it becomes computationally intractable to compute the optimal policy. Taking this strategy will improve the overall performance guarantee.

We also note that the bound in Theorem 4.2 above is tight with respect to the LP defined by (3)-(5) above. Specifically, consider the case where there is a single class with $\rho_1 = 1 = C$. Clearly, the optimal policy is to accept every customer as long as capacity permits. It can be verified that the expected long-run revenue rate of the optimal policy is half the optimal value of the LP. Thus, there is no hope of proving uniform performance guarantees stronger than half using the LP as the only upper bound.

Finally, note that the performance analysis presented above still holds if the price rate of each class is a random variable with mean r_i.

4.2 Non-identical Resource Requirements

In this section, we consider the case, where each class may have different resource requirement $A_i \in \mathbb{Z}_+$. Again, we assume without loss of generality that there is no fractional variable in α^*, the optimal solution of the LP, and that there are $M' \leq M$ classes with positive α_i^* variable. That is, $\alpha_i^* = 1$, for each $i = 1, \ldots, M'$. It is now clear that the respective rejection (blocking) probabilities $P_1, \ldots, P_{M'}$ are not identical. In particular, by Theorem 2 of [4], for each $i = 1, \ldots, M'$, the rejection probability P_i can be expressed as

$$P_i = \mathbb{P}\left(\sum_{k=1}^{M'} X_k A_k > C - A_i\right) = \frac{\mathbb{P}\left(C - A_i < \sum_{k=1}^{M'} Y_k A_k \leq C\right)}{\mathbb{P}\left(\sum_{k=1}^{M'} Y_k A_k \leq C\right)}. \quad (12)$$

It follows that if $A_i \geq A_j$ then $P_i \geq P_j$.

However, the analysis in this case is more involved compared to the analysis in the case with identical resource requirements. The main difficulty comes from the fact that if there are several classes with different resource requirements, then $\sum_{k=1}^{M'} Y_k A_k$ follows a compound Poisson distribution, whereas in the case of identical requirements it follows a Poisson distribution. (In the latter case we can assume that all the resource requirements are equal 1.) This makes the analysis significantly harder. Nevertheless, we will show that the CSP policy is asymptotically optimal, even if the resource requirements are not identical.

Let $A = \max_{i=1,\ldots,M'} A_i$ be the maximum resource requirement by a class. We shall show that all blocking probabilities $P_1, \ldots, P_{M'}$ diminish to zero as ratio C/A^7 grows to infinity. In light of (12) and the monotonicity of the blocking probabilities in the respective class resource requirement, it is sufficient to show that $\frac{\mathbb{P}(C - A < \sum_{k=1}^{M'} Y_k A_k \leq C)}{\mathbb{P}(\sum_{k=1}^{M'} Y_k A_k \leq C)}$ diminishes to 0 as the ratio C/A^7 grows to the infinity. Specifically, we will show that $\mathbb{P}(C - A < \sum_{k=1}^{M'} Y_k A_k \leq C)$ diminishes to 0, and that $\mathbb{P}(\sum_{k=1}^{M'} Y_k A_k \leq C)$ is asymptotically at least 0.25 as the ratio C/A^7 grows to infinity. Indeed, this implies that the blocking probabilities $P_1, \ldots, P_{M'}$ diminish to 0 as the ratio C/A^7 grows to infinity. We first focus on the case where $\lambda = \sum_{i=1}^{M'} \rho_i \geq \frac{C}{2A}$. That is, the total traffic intensity is relatively high.

Lemma 4.3 *Assume that $\lambda \geq \frac{C}{2A}$. Then probability $\mathbb{P}(C - A < \sum_{k=1}^{M'} Y_k A_k \leq C)$ diminishes to 0 as the ratio C/A^4 grows to infinity.*

Proof: Without loss of generality, we assume that $M' \leq A$. Otherwise, there are two classes with the same resource requirement, and we can consider them in the analysis as one class with intensity equal to the sum of the intensities of the original classes. (This will not change the probability $\mathbb{P}(C - A < \sum_{k=1}^{M'} Y_k A_k \leq C)$.) Thus, $A \max_i \rho_i \geq \lambda \geq \frac{C}{2A}$ and, therefore, there exists at least one class with $\rho_i \geq \frac{C}{2A^2}$; without loss of generality, assume that M' is that class. Let $\mathcal{S}$ be the set of all feasible states. That is, $\mathcal{S} = \{y \in \mathbb{Z}_+^{M'} : 0 \leq \sum_{i=1}^{M'} y_i A_i \leq C\}$. Let $\mathcal{B} \subseteq \mathcal{S}$ be the set of all blocking states. That is,

$$\mathcal{B} = \left\{ y \in \mathbb{Z}_+^{M'} \; : \; C - A < \sum_{i=1}^{M'} y_i A_i \leq C \right\}.$$

Let $\mathcal{S}^{M'}$ be the projection of $\mathcal{S}$ onto the space of the variables $(y_1, \ldots, y_{M'-1})$. Specifically,

$$\mathcal{S}^{M'} = \left\{ y \in \mathbb{Z}_+^{M'-1} \; : \; 0 \leq \sum_{i=1}^{M'-1} y_i A_i \leq C \right\}.$$

For each $y \in \mathcal{S}^{M'}$, let $P(y) = Pr(Y_1 = y_1, \ldots, Y_{M'-1} = y_{M'-1})$. Focus now on $y' \in \mathcal{S}^{M'}$, and consider the set of values of $y_{M'}$, for which state $(y', y_{M'})$ is a blocking state, i.e., $(y', y_{M'}) \in \mathcal{B}$. Denote this set by $\mathcal{B}^{M'}(y')$. By definition $\mathcal{B}^{M'}(y') = \{y_{M'} \in \mathbb{Z}_+ : \; C - A < \sum_{i=1}^{M'-1} y'_i A_i + y_{M'} A_{M'} \leq C\}$. We conclude that the set $\mathcal{B}^{M'}(y')$ consists of at most $\lceil \frac{A}{A_{M'}} \rceil \leq A$ consecutive integers. Moreover, since the random variables $Y_1, \ldots, Y_{M'}$ are independent, it follows that

$$\begin{aligned} \mathbb{P}\left(C - A < \sum_{i=1}^{M'} Y_i A_i \leq C \right) &= \sum_{y \in \mathcal{S}^{M'}} P(y) \sum_{y_{M'} \in \mathcal{B}^{M'}(y)} \mathbb{P}(Y_{M'} = y_{M'}) \\ &\leq \max_{z \geq 0} \mathbb{P}(z - A < Y_{M'} \leq z) \leq A \frac{\rho_{M'}^{\rho_{M'}}}{\rho_{M'}!} e^{-\rho_{M'}}. \end{aligned} \tag{13}$$

The first inequality follows from the fact that for each $y \in \mathcal{S}^{M'}$, the set $\mathcal{B}^{M'}(y)$ consists of at most A consecutive integers and $\sum_{y \in \mathcal{S}^{M'}} P(y) \leq 1$. (This is the sum of probabilities of disjoint events.) The second inequality follows from the properties of the Poisson distribution. In particular, the maximum probability assigned to a specific value is at most $\frac{\rho_{M'}^{\rho_{M'}}}{\rho_{M'}!} e^{-\rho_{M'}}$, and as we have already seen, the set $\mathcal{B}^{M'}(y)$ contains at most A values.

Finally, it is straightforward to check that function $h(z) \triangleq \frac{z^z}{z!} e^{-z}$ is monotone decreasing in z. This would follow from identical arguments used to justify monotonicity of function $g(z)$ in the proof of Lemma 4.1. Then, since $\rho_{M'} \geq \frac{C}{2A^2}$, we obtain

$$\mathbb{P}\left(C - A < \sum_{i=1}^{M'} Y_i A_i \leq C \right) \leq A \frac{\frac{C}{2A^2}^{\frac{C}{2A^2}}}{\frac{C}{2A^2}!} e^{-\frac{C}{2A^2}}, \tag{14}$$

where, for any positive x, $x!$ satisfies identity (11). Moreover, using Stirling approximation, we know that the function $\frac{\frac{C}{2A^2}^{\frac{C}{2A^2}}}{\frac{C}{2A^2}!}e^{-\frac{C}{2A^2}}$ behaves like $\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\frac{C}{2A^2}}}$ as C/A^2 grows to infinity. Therefore, as C/A^4 grows to infinity, probability $\mathbb{P}(C - A < \sum_{i=1}^{M'} Y_i A_i \leq C)$ diminishes to 0. This concludes the proof of the lemma. ■

Next, we show that as C/A^7 grows to infinity, probability $\mathbb{P}(0 \leq \sum_{i=1}^{M'} Y_i A_i \leq C)$ is asymptotically at least 0.25.

Lemma 4.4 *Assume that* $\lambda \geq \frac{C}{2A}$. *Then probability* $\mathbb{P}(0 \leq \sum_{k=1}^{M'} Y_k A_k \leq C)$ *is asymptotically at least 0.25 as* C/A^7 *grows to infinity.*

Proof: Recall that random variable $\sum_{k=1}^{M'} Y_k A_k$ follows a compound Poisson distribution. Let W be a Poisson random variable with parameter $\lambda = \sum_{i=1}^{M'} \rho_i$, and let Q be a discrete random variable that takes value A_k with probability ρ_k/λ, for each $k = 1, \ldots, M'$. Next, we express random variable $\sum_{k=1}^{M'} Y_k A_k$ as $\sum_{i=0}^{W} Q_i$, where $\{Q_i\}_{i=1}^{\infty}$ is a sequence of independent and identically distributed random variables equal in distribution to Q, and are independent of W. (It is easy to verify that random variable $\sum_{i=0}^{W} Q_i$ has the same distribution as $\sum_{k=1}^{M'} Y_k A_k$.)

In order to prove the lemma, we develop two lower bounds for probability $\mathbb{P}(0 \leq \sum_{i=0}^{W} Q_i \leq C)$; one of the bounds is based on the Central Limit Theorem and the other on Chebyshev inequality. Then, we show that, as the ratio C/A^7 grows to infinity, at least one of the two lower bounds is asymptotically at least 0.25.

Clearly,

$$\begin{aligned} \mathbb{P}\left(\sum_{i=0}^{W} Q_i \leq C\right) &\geq \mathbb{P}(W \leq \lambda)\mathbb{P}\left(\sum_{i=0}^{W} Q_i \leq C | W \leq \lambda\right) \\ &\geq \mathbb{P}(W \leq \lambda)\mathbb{P}\left(\sum_{i=0}^{\lfloor\lambda\rfloor} Q_i \leq C | W \leq \lambda\right) \\ &= \mathbb{P}(W \leq \lambda)\mathbb{P}\left(\sum_{i=0}^{\lfloor\lambda\rfloor} Q_i \leq C\right). \end{aligned} \quad (15)$$

The first inequality follows from the fact that we restrict attention only to event $[W \leq \lambda]$. The second inequality follows from the fact that Q_i's are nonnegative, W is an integer-valued random variable and we restrict attention to event $[W \leq \lambda]$. The equality follows from the fact that W is independent of $\sum_{i=0}^{\lfloor\lambda\rfloor} Q_i$.

First, consider the last expression in (15) above. Focus on term $\mathbb{P}(W \leq \lambda)$. By the properties of the median of the Poisson distribution (see [22]), we know that $\mathbb{P}(W \leq \lambda) \geq 0.5 - 0.5\mathbb{P}(W = \lceil\lambda\rceil)$. However, we have already seen that probability $\mathbb{P}(W = \lceil\lambda\rceil)$ diminishes to 0 as $\lambda \geq \frac{C}{2A}$ grows to infinity. Thus, it follows that as $\frac{C}{2A}$ grows to infinity, probability $\mathbb{P}(W \leq \lambda)$ is asymptotically at least 0.5.

Next, focus on term $\mathbb{P}(\sum_{i=0}^{\lfloor\lambda\rfloor} Q_i \le C)$. From Constraint (4) it follows that $C \ge \lambda E[Q_1] \ge \lfloor\lambda\rfloor E[Q_1]$, and we obtain

$$\mathbb{P}\left(\sum_{i=0}^{\lfloor\lambda\rfloor} Q_i \le C\right) \ge \mathbb{P}\left(\sum_{i=0}^{\lfloor\lambda\rfloor}(Q_i - E[Q_1]) \le 0\right).$$

However, the random variables Q_is are independent and identically distributed and $\lfloor\lambda\rfloor \ge \frac{C}{A}$ grows to infinity. Thus, a Central Limit argument can be applied. In particular, we use the well-known Berry-Essen bound to express deviation from the Normal distribution (see, for example, [5]). Specifically, let $\sigma^2 = Var[Q]$ be the variance of Q and $\gamma = E[Q^3]$ its third moment. (Both exist and are finite.) Then,

$$\begin{aligned}\mathbb{P}\left(\sum_{i=0}^{\lfloor\lambda\rfloor}(Q_i - E[Q_1]) \le 0\right) &\ge \Phi(0) - \tau_0 \frac{\gamma}{\sigma^3\sqrt{\lfloor\lambda\rfloor}} \\ &= 0.5 - \tau_0 \frac{\gamma}{\sigma^3\sqrt{\lfloor\lambda\rfloor}},\end{aligned} \tag{16}$$

where $\tau_0 > 0$ is a universal constant independent of Q, and $\Phi(x)$ is the cumulative distribution function of the standard normal distribution. Since $\gamma \le A^3$, we can extend (16) above to get $\mathbb{P}(\sum_{i=0}^{\lfloor\lambda\rfloor}(Q_i - E[Q_1]) \le 0) \ge 0.5 - \frac{\tau_0 A^3}{\sigma^3\sqrt{\lfloor\lambda\rfloor}}$. Since $\lambda \ge \frac{C}{2A}$ it follows that, as C/A^7 grows to infinity, and unless σ^3 is going to 0 faster than $(C/A^7)^{-1/2}$, the lower bound developed in (15) above approaches 0.25.

Note that we only assume that the ratio C/A^7 grows to infinity, irrespective of other parameters of the model (which can be arbitrary). In particular, we can not assume that σ is a bounded from below. Indeed the bound in (16) above becomes meaningless when σ^3 approaches zero faster than $(C/A^7)^{-1/2}$. Thus, we develop the second lower bound for $\mathbb{P}(\sum_{i=0}^{W} Q_i \le C)$ that is based on Chebyshev inequality, and is strong when σ is close to 0. Using similar arguments to those in (15) above, we can condition on event $[W \le \lambda - \sqrt{\lambda}]$, and obtain

$$\mathbb{P}\left(\sum_{i=0}^{W} Q_i \le C\right) \ge \mathbb{P}\left(W \le \lambda - \sqrt{\lambda}\right)\mathbb{P}\left(\sum_{i=0}^{\lfloor\lambda-\sqrt{\lambda}\rfloor} Q_i \le C\right). \tag{17}$$

Focus on the first term $\mathbb{P}(W \le \lambda - \sqrt{\lambda})$. We claim that, as $\lambda \ge \frac{C}{2A}$ grows to infinity, the probability $\mathbb{P}(W \le \lambda - \sqrt{\lambda})$ approaches $\Delta = 0.5 - \frac{1}{\sqrt{2\pi}}$. We have already seen that $\mathbb{P}(W \le \lambda)$ approaches 0.5. Thus, it is sufficient to show that $\mathbb{P}(\lambda - \sqrt{\lambda} < W \le \lambda)$ is asymptotically at most $\frac{1}{\sqrt{2\pi}}$. However, by arguments identical to the one used in the proof of Lemma 4.3 above, we obtain that $\mathbb{P}(\lambda-\sqrt{\lambda} < W \le \lambda) \le \frac{\lambda^\lambda}{\lambda!}e^{-\lambda}\sqrt{\lambda}$. Using Stirling approximation, we conclude that, indeed, $\mathbb{P}(\lambda - \sqrt{\lambda} < W \le \lambda)$ approaches $\frac{1}{\sqrt{2\pi}}$ as $\lambda \ge \frac{C}{2A}$ grows to infinity.

Now, we focus on the second term $\mathbb{P}\left(\sum_{i=0}^{\lfloor\lambda-\sqrt{\lambda}\rfloor} Q_i \le C\right)$. Observe that $C \ge \lambda E[Q]$, the expectation $E\left[\sum_{i=0}^{\lfloor\lambda-\sqrt{\lambda}\rfloor} Q_i\right] = \lfloor\lambda - \sqrt{\lambda}\rfloor E[Q]$ and the variance $Var\left[\sum_{i=0}^{\lfloor\lambda-\sqrt{\lambda}\rfloor} Q_i\right] = \lfloor\lambda - \sqrt{\lambda}\rfloor\sigma^2$. Thus, by applying Chebyshev inequality, we derive

$$\begin{aligned}\mathbb{P}\left(\sum_{i=0}^{\lfloor \lambda-\sqrt{\lambda}\rfloor} Q_i > C\right) &\leq \mathbb{P}\left(\sum_{i=0}^{\lfloor \lambda-\sqrt{\lambda}\rfloor} (Q_i - E[Q_1]) > \lfloor\sqrt{\lambda}\rfloor E[Q_1]\right) \\ &\leq \frac{(\lambda-\sqrt{\lambda})\sigma^2}{(\sqrt{\lambda}-1)^2(E[Q])^2} \\ &\leq \sqrt{\lambda}\sigma^2/(\sqrt{\lambda}-1),\end{aligned} \tag{18}$$

where the last inequality follows from the fact that $E[Q] \geq 1$.

Now, if $\sqrt{\lambda}\sigma^2/(\sqrt{\lambda}-1) \leq 1 - 0.25/\Delta$, it follows that the lower bound developed in (17) above approaches 0.25 as C/A grows to infinity. Otherwise, we get

$$\sigma^2 > (1 - 0.25/\Delta)\frac{\sqrt{\lambda}-1}{\sqrt{\lambda}},$$

and the lower bound developed in (15) above is asymptotically at least 0.25 as C/A^7 grows to infinity. This concludes the proof of the lemma. ■

Lemmas 4.3 and 4.4 imply the following theorem.

Theorem 4.5 *Assume that* $\lambda \geq \frac{C}{2A}$. *Then the blocking probabilities* $P_1, \ldots, P_{M'}$ *diminish to* 0 *as the ratio* C/A^7 *grows to infinity.*

Finally, we discuss the case where $\lambda = \sum_{i=1}^{M'} \rho_i < \frac{C}{2A}$. Using the notation introduced above, we wish to show that probability $\mathbb{P}(\sum_{i=1}^{W} Q_i > C - A)$ diminishes to 0 as ratio C/A grows to infinity. Indeed, $\mathbb{P}(\sum_{i=1}^{W} Q_i > C - A) \leq \mathbb{P}(W > \frac{C-A}{A})$. (If $W \leq \frac{C-A}{A}$, it is clear that $\sum_{i=1}^{W} Q_i \leq C - A$ since $Q \leq A$.) However, by the properties of Poisson random variable (see 6.5.34 of [1]), it follows that probability $\mathbb{P}(W > \frac{C-A}{A})$ diminishes to 0 as $\frac{C}{A}$ grows to infinity. (The mean of W is at most $\frac{C}{2A}$.) We have obtained the following theorem.

Theorem 4.6 *The CSP policy is asymptotically optimal for the revenue management model with single reusable resource as the ratio* C/A^7 *grows to infinity, regardless of other parameters of the problem.*

We note that, for each fixed value of A one can use Lemmas 4.3 and 4.4 to derive uniform performance guarantees similar to the one developed in Theorem 4.2.

4.3 A Comparison of the Performance Analysis with Existing Approaches

Next we would like to contrast our performance analysis with existing literature. As already mentioned, there are several results that establish the asymptotic optimality of different policies. However, most if not all of the existing results assume at least one of the following: (i) simultaneous scaling of the arrival rates and the capacity (That is, we set $C_n = nC$ and $\lambda_n = n\lambda$, and let n grow to infinity.); (ii) other parameters of the problem, such as the number of classes and resource requirements and service time distributions, are kept fixed; (iii) (in some cases) exponentially distributed service time.

In contrast, our analysis holds for general service time distributions, and it only requires that the ratio C/A^7 grows to infinity, letting other parameters of the problem, such as arrival rates, number of classes and resource requirements, be arbitrary. The analysis highlights the fact that the most important characteristic of the problem is the ratio between the capacity and the resource requirements of different classes. Moreover, the fact that our performance analysis relies on the LP-based upper bound enables us to compute explicit and uniform performance bounds. As we shall show in the next section this analysis extends to models with multiple resources.

5 Extensions

5.1 Revenue Management Model with Multiple Reusable Resources

In this section, we show how to use ideas analogous to those used in the single resource case to extend the CSP policy and the performance analysis to models with multiple resources. We present a packing-type LP that provides an upper bound on the best achievable revenue rate and use its optimal solution to construct the CSP policy and analyze its performance. As before we can use general Erlang formulas (see, for example, Burman et. al. [4]) to express the stationary class rejection (blocking) probabilities. This is again used to show the asymptotic optimality of the CSP policy in a way similar to the analysis of the single resource case discussed in Section 4 above.

Let C_j be the capacity of resource $j = 1, \ldots, N$, and let $A_{ij} \in \mathbb{Z}_+$ be the number of units of resource j requested by class i. One can write a packing-type LP that provides a similar upper bound on the best achievable revenue rate. Specifically, the objective is the same as (3), but we have a constraint $\sum_{i=1}^{M} \alpha_i \rho_i A_{ij} \leq C_j$, for each $j = 1, \ldots, N$. This LP can be solved and one can derive the CSP policy that accepts class-i customer with probability α_i as long as there is sufficient amount of available resources to satisfy her/his requirement. In view of the discussion at the beginning of Section 4, one can extend the same arguments to this case as well, and express the rejection probability of a class $k \leq M$ customer as

$$P_k = \frac{\mathbb{P}\left(\cup_{j=1}^{N}\{C_j - A_{kj} < \sum_{k=1}^{M} Y_k A_{kj} \leq C_j\}\right)}{\mathbb{P}\left(\cap_{j=1}^{N}\{\sum_{k=1}^{M} Y_k A_{kj} \leq C_j\}\right)}, \tag{19}$$

where Y_k, $1 \leq k \leq M$, are independent Poisson random variables with mean values $\alpha_k \rho_k$, $1 \leq k \leq$

M. Next, note that for every $1 \le j \le N$, random variable $\sum_{k=1}^{M} Y_k A_{kj}$ follows a compound Poisson distribution. Let W be a Poisson random variable with parameter $\lambda = \sum_{i=1}^{M} \rho_i \alpha_i$, and let $Q^{(j)}$, $1 \le j \le N$, be a discrete random variable that takes value A_{kj} with probability ρ_k/λ. Next, we express random variable $\sum_{k=1}^{M} Y_k A_{kj}$, $1 \le j \le N$, as $\sum_{i=1}^{W} Q_i^{(j)}$, where for each $1 \le j \le N$, $\{Q_i^{(j)}\}_{i=1}^{\infty}$ is a sequence of independent and identically distributed random variables equal in distribution to $Q^{(j)}$ and independent of W.

First, we show that the numerator in (19) diminishes to zero as $\min_j C_j/A_j^4$ approaches infinity, where $A_j \triangleq \max_{1 \le k \le M} A_{kj}$. By union bound,

$$\mathbb{P}\left(\cup_{j=1}^{N}\left\{C_j - A_{ij} < \sum_{k=1}^{M} Y_k A_{kj} \le C_j\right\}\right) \le \sum_{j=1}^{N} \mathbb{P}\left(C_j - A_{ij} < \sum_{l=0}^{W} Q_l^{(j)} \le C_j\right).$$

Using arguments analogous to one used in the proof of Lemma 4.3, each summand on the right hand side of the previous expression approaches zero as $\min_{1\le j\le N}(C_j/A_j^4)$ approaches infinity. Since the number of resources N is a finite constant, we obtain that $\mathbb{P}(\cup_{j=1}^{N}\{C_j - A_{ij} < \sum_{k=1}^{M} Y_k A_{kj} \le C_j\})$ diminishes to zero as $\min_{1\le j\le N}(C_j/A_j^4)$ grows to infinity.

Next, we show that the denominator in (19) can be bounded by some positive constant as $\min_{1\le j\le N} C_j/A_j^7$ approaches infinity. In view of the previous discussions,

$$\mathbb{P}\left(\cap_{j=1}^{N}\left\{\sum_{k=1}^{M} Y_k A_{kj} \le C_j\right\}\right) = \mathbb{P}\left(\cap_{j=1}^{N}\left\{\sum_{l=0}^{W} Q_l^{(j)} \le C_j\right\}\right),$$

and, similarly as in (15), we lower bound this expression by conditioning on event $[W \le \lambda - f(N)\sqrt{\lambda}]$, with $f(N) = N^a$, where a is the smallest positive real number such that $N \max_j \bar{\Phi}\left(\frac{f(N)}{A_j}\right) < 1$. (We use $\bar{\Phi}(x)$ to denote the tail of the standard normal distribution at x.) Therefore, we obtain

$$\mathbb{P}\left(\cap_{j=1}^{N}\left\{\sum_{l=0}^{W} Q_l^{(j)} \le C_j\right\}\right) \ge \mathbb{P}(W \le \lambda - f(N)\sqrt{\lambda})\mathbb{P}\left(\cap_{j=1}^{N}\left\{\sum_{i=0}^{\lfloor \lambda - f(N)\sqrt{\lambda}\rfloor} Q_i^{(j)} \le C_j\right\}\right). \quad (20)$$

Furthermore, by applying the union bound, we further lower bound the left hand side of (20) and get

$$\mathbb{P}\left(\cap_{j=1}^{N}\left\{\sum_{l=0}^{W} Q_l^{(j)} \le C_j\right\}\right) \ge \mathbb{P}(W \le \lambda - f(N)\sqrt{\lambda})\left(1 - \sum_{j=1}^{N} \mathbb{P}\left(\sum_{i=0}^{\lfloor \lambda - f(N)\sqrt{\lambda}\rfloor} Q_i^{(j)} > C_j\right)\right). \quad (21)$$

Next, we show that both terms on the right hand side of (21) can be lower bounded by positive constants as $\min_j C_j/A_j^7$ approaches infinity. Before proceeding, we will differentiate between two cases: (i) $\lambda \ge \min_j C_j/2A_j$, and (ii) $\lambda < \min_j C_j/2A_j$.

We first focus on the first case. Since W is a Poisson random variable, the probability $\mathbb{P}(W \le \lambda - f(N)\sqrt{\lambda}) \ge \mathbb{P}(\sum_{i=1}^{\lceil \lambda - f(N)\sqrt{\lambda}\rceil} \epsilon_i \ge \lambda)$, where $\{\epsilon_i\}_{i\ge 1}$ is a sequence of independent exponentially distributed random variables with mean 1. Thus,

$$\mathbb{P}(W \leq \lambda - f(N)\sqrt{\lambda}) \geq \mathbb{P}\left(\sum_{i=1}^{\lceil \lambda - f(N)\sqrt{\lambda}\rceil} \epsilon_i \geq \lambda\right)$$

$$\geq \mathbb{P}\left(\sum_{i=1}^{\lceil \lambda - f(N)\sqrt{\lambda}\rceil} (\epsilon_i - 1) \geq f(N)\sqrt{\lambda}\right)$$

$$= \mathbb{P}\left(\frac{\sum_{i=1}^{\lceil \lambda - f(N)\sqrt{\lambda}\rceil}(\epsilon_i - 1)}{\sqrt{\lceil \lambda - f(N)\sqrt{\lambda}\rceil}} \geq f(N)\frac{\sqrt{\lambda}}{\sqrt{\lceil \lambda - f(N)\sqrt{\lambda}\rceil}}\right)$$

$$\geq \bar{\Phi}\left(\frac{f(N)\sqrt{\lambda}}{\sqrt{\lambda - f(N)\sqrt{\lambda}}}\right) - \tau_0 \frac{1}{\sqrt{\lambda - f(N)\sqrt{\lambda} - 1}},$$

where in the last inequality we applied Berry-Essen inequality analogously as in (16). Now, if $\min_j C_j/A_j^7$ grows to infinity, then $\lambda \geq \min_j C_j/2A_j$ grows to infinity as well, implying that $\mathbb{P}(W \leq \lambda - f(N)\sqrt{\lambda})$ approaches $\bar{\Phi}(f(N))$. Finally, since the number of resources N is finite, $\bar{\Phi}(f(N)) > 0$.

Next, we shall show that the sum in the second term of (21) can be upper bounded by a constant less than 1 as $\min_j C_j/A_j^7$ approaches infinity. In particular, we upper bound each summand $\mathbb{P}(\sum_{i=1}^{\lfloor \lambda - f(N)\sqrt{\lambda}\rfloor} Q_i^{(j)} > C_j)$, $1 \leq j \leq N$. Define $\sigma_j^2 \triangleq Var[Q^{(j)}]$. Then, similarly as in the proof of Lemma 4.4, depending on whether σ_j^3 approaches zero faster than $(C_j/A_j^7)^{-1/2}$, or not, the upper bound is based on Chebyshev inequality, or Berry-Essen bound. In particular, by Berry-Essen inequality and $\lambda\mathbb{E}[Q^{(j)}] \leq C_j$, we obtain

$$\mathbb{P}\left(\sum_{i=0}^{\lfloor \lambda - f(N)\sqrt{\lambda}\rfloor} Q_i^{(j)} > C_j\right) \leq \mathbb{P}\left(\sum_{i=1}^{\lfloor \lambda - f(N)\sqrt{\lambda}\rfloor} (Q_i^{(j)} - \mathbb{E}[Q^{(j)}]) > f(N)\sqrt{\lambda}\mathbb{E}[Q^{(j)}]\right) \tag{22}$$

$$\leq \mathbb{P}\left(\frac{\sum_{i=1}^{\lfloor \lambda - \sqrt{f}(N)\lambda\rfloor}(Q_i^{(j)} - \mathbb{E}[Q_i^{(j)}])}{\sigma_j\lfloor\sqrt{\lambda - f(N)\sqrt{\lambda}}\rfloor} > \frac{f(N)\sqrt{\lambda}\mathbb{E}[Q^{(j)}]}{\sigma_j\sqrt{\lambda - f(N)\sqrt{\lambda}}}\right)$$

$$\leq \bar{\Phi}\left(\frac{f(N)\sqrt{\lambda}\mathbb{E}[Q^{(j)}]}{\sigma_j\sqrt{\lambda - f(N)\sqrt{\lambda}}}\right) + \tau_0\frac{A_j^3}{\sigma_j^3\sqrt{\lambda - f(N)\sqrt{\lambda}}}.$$

Now, when $\lambda \geq C_j/2A_j$ and σ_j^3 does not go to zero faster than $(C_j/A_j^7)^{-1/2}$, we apply the bound in (22) and obtain that $\mathbb{P}\left(\sum_{i=0}^{\lfloor \lambda - f(N)\sqrt{\lambda}\rfloor} Q_i^{(j)} > C_j\right)$ does not exceed $\bar{\Phi}(f(N)/A_j)$ as C_j/A_j^7 approaches infinity.

On the other hand, if σ_j^3 approaches zero faster than $(C_j/A_j^7)^{-1/2}$, we apply Chebyshev inequality (using the facts that $\lambda\mathbb{E}[Q^{(j)}] \leq C_j$ and $\mathbb{E}[Q^{(j)}] \geq 1$) to obtain

$$\mathbb{P}\left(\sum_{i=0}^{\lfloor \lambda - f(N)\sqrt{\lambda}\rfloor} Q_i^{(j)} > C_j\right) \leq \mathbb{P}\left(\sum_{i=1}^{\lfloor \lambda - f(N)\sqrt{\lambda}\rfloor} (Q_i^{(j)} - \mathbb{E}[Q^{(j)}]) > f(N)\sqrt{\lambda}\mathbb{E}[Q^{(j)}]\right) \tag{23}$$

$$\leq \frac{\sqrt{\lambda} - f(N)}{(f(N))^2\sqrt{\lambda}}\sigma_j^2.$$

It follows that if σ_j^2 satisfies

$$\frac{(\sqrt{\lambda} - f(N))\sigma_j^2}{(f(N))^2\sqrt{\lambda}} \leq \bar{\Phi}\left(\frac{f(N)}{A_j}\right),$$

then bound (23) implies that the probability $\mathbb{P}\left(\sum_{i=0}^{\lfloor \lambda - f(N)\sqrt{\lambda} \rfloor} Q_i^{(j)} > C_j\right)$ does not exceed $\bar{\Phi}(f(N)/A_j)$ as C_j/A_j^7 approaches infinity. Otherwise, the same asymptotic upper bound holds by (22).

Finally, we show that when $\lambda \geq \min_j C_j/2A_j$ and $\min_j C_j/A_j^7$ approaches infinity, then for each $1 \leq j \leq N$, the probability $\mathbb{P}\left(\sum_{i=0}^{\lfloor \lambda - f(N)\sqrt{\lambda} \rfloor} Q_i^{(j)} > C_j\right)$ does not exceed asymptotically $\bar{\Phi}(f(N)/A_j)$. Given the assumption that $N \max_j \bar{\Phi}(f(N)/A_j) < 1$, the asymptotic lower bound for $\mathbb{P}(W \leq \lambda - f(N)\sqrt{\lambda})$ derived above, and (21) above, it follows that $\mathbb{P}\left(\cap_{j=1}^N \left\{\sum_{l=0}^W Q_l^{(j)} \leq C_j\right\}\right)$ can be lower bounded by a positive constant as $\min_j C_j/A_j^7$ approaches infinity.

In a way similar to one used in the case with a single resource with non-uniform requirements, we show that when $\lambda < \min_j C_j/2A_j$, then for each $1 \leq j \leq N$, the probability $\mathbb{P}(\sum_{l=1}^W Q_l^{(j)} > C_j) \leq \mathbb{P}(W > C_j/A_j)$ diminishes to zero as C_j/A_j approaches infinity. This implies that the probability $\mathbb{P}\left(\cap_{j=1}^N \left\{\sum_{l=1}^W Q_l^{(j)} \leq C_j\right\}\right)$ $1 - \sum_{j=1}^N \mathbb{P}\left(\sum_{l=1}^W Q_l^{(j)} \leq C_j\right)$ approaches 1 as $\min_j C_j/A_j^7$ approaches infinity. This concludes the analysis in the multiple resource case.

5.2 Price-Driven Customer Arrivals

In this section, we consider an extension of the model discussed in Section 2, in which the arrival rates of the different classes of customers are affected by prices. Specifically, consider a two stage decision. At the first stage we set the respective prices $r_1, \ldots, r_M$ for each class. This determines the respective arrival rates $\lambda_1(r_1), \ldots, \lambda_M(r_M)$. (The rate of class i is affected only by price r_i.) Then, given the arrival rates, we wish to find the optimal admission policy that maximizes the expected long-run revenue rate. In particular, we assume that $\lambda_i(r_i)$ is nonnegative, differentiable and decreasing in r_i for every $1 \leq i \leq M$. In addition, there exists a price, say r_∞, such that for each $i = 1, \ldots, M$, we have $\lambda_i(r) = 0$ for $r \geq r_\infty$.

Using arguments analogous to the discussion in Section 3, we construct an upper bound on the achievable expected long-run revenue rate through the following nonlinear program (NLP1):

$$\max_{\alpha_1,\ldots,\alpha_M,r_1,\ldots,r_M} \sum_{i=1}^M r_i\alpha_i\rho_i(r_i) \tag{24}$$

$$\text{s.t.} \sum_{i=1}^M \alpha_i\rho_i(r_i)A_i \leq C \tag{25}$$

$$0 \leq \alpha_k \leq 1, \ \ \forall\, 1 \leq i \leq M. \tag{26}$$

As before, for each $i = 1, \ldots, M$, define $\rho_i(r_i) = \lambda_i(r_i)\mu_i$. In particular, it can be verified that any optimal solution of (NLP1) has only nonnegative prices. Also, observe that for any fixed prices $r_1, \ldots, r_M$, the corresponding solution of $\alpha_1, \ldots, \alpha_M$ has the same knapsack structure defined above. We renumber the classes according to decreasing order of ratio r_i/A_i, $1 \leq i \leq M$, and denote the optimal prices by $r^* = (r_1^*, \ldots, r_M^*)$. Let $\alpha^* = (\alpha_1^*, \ldots, \alpha_M^*)$ be the corresponding optimal α values. Then, for some $M' \leq M$, the vector α^* has the following structure:

$$\alpha_1^* = 1, \ldots, \alpha_{M'-1}^* = 1, 0 < \alpha_{M'}^* \leq 1, \alpha_{M'+1}^* = 0, \ldots, \alpha_M^* = 0. \tag{27}$$

Note that if one can solve (NLP1) and obtain the solution (r^*, α^*) then one can construct a similar CSP policy that will be amenable to the same performance analysis discussed in Section 4 above. However, solving (NLP1) directly may be computationally hard. Next, we show that under relatively mild conditions imposed on the functions $\lambda_1(r_1), \ldots, \lambda_M(r_M)$, one can reduce (NLP1) to an equivalent nonlinear program that is more tractable and we denote by (NLP2). (By equivalent we mean that they have the same set of optimal solutions.) consider (NLP2) as follows:

$$\max_{r_1,\ldots,r_M} \sum_{i=1}^{M} r_i \rho_i(r_i) \tag{28}$$

$$\text{s.t.} \sum_{i=1}^{M} \rho_i(r_i) A_i \leq C \tag{29}$$

It can be readily verified that as long as $\rho_i(r_i)$ is nonnegative (and decreasing) it is always optimal to have nonnegative prices.

Lemma 5.1 *The programs (NLP1) and (NLP2) are equivalent.*

Proof: First, we show that for each solution $r = (r_1, \ldots, r_M)$ of (NLP 2), we can construct a solution of (NLP1) with the same objective value. Specifically, consider solution (r', α'), such that $r' = r$ and $\alpha_i' = 1$ if and only if $r_i \rho_i(r_i) > 0$. It can be verified that the resulting solution is feasible for (NLP1) and has the same objective value.

Next, we show how to map optimal solution (r^*, α^*) of (NLP1) to a feasible solution of (NLP2) with the same objective function. For each $i = 1, \ldots, M' - 1$, set $r_i = r_i^*$, and for each $i = M' + 1, \ldots, M$ set $r_i = r_\infty$. It is clear that, for each $i \neq M' - 1$, the resulting contributions to the objective value and Constraint (29) are the same as in (NLP1). Consider now possibly fractional $\alpha_{M'}$. The respective contribution of class M' to the objective value is $\alpha_{M'}^* r_{M'}^* \rho_{M'}(r_{M'}^*)$. Similarly, the contribution to Constraint (29) is $\alpha_{M'}^* \rho_{M'}(r_{M'}^*) A_{M'}$. Thus, it is sufficient to show that there exists a price $r_{M'}$ such that $r_{M'} \rho_{M'}(r_{M'}) \geq \alpha_{M'}^* r_{M'}^* \rho_{M'}(r_{M'}^*)$ and $\rho_{M'}(r_{M'}) A_{M'} \leq \alpha_{M'}^* \rho_{M'}(r_{M'}^*) A_{M'}$.

Since $r^*_{M'}\rho_{M'}(r^*_{M'}) \geq \alpha^*_{M'}r^*_{M'}\rho_{M'}(r^*_{M'})$, by the continuity and monotonicity of $\lambda_{M'}(r_{M'})$, we know that there exists $\bar{r} \in [r^*_{M'}, r_\infty)$ such that $\bar{r}\rho_{M'}(\bar{r}) = \alpha^*_{M'}r^*_{M'}\rho_{M'}(r^*_{M'})$. Note that $\bar{r} \geq r^*_{M'}$, and, therefore, we obtain

$$r^*_{M'}\rho_{M'}(\bar{r}) \leq \bar{r}\rho_{M'}(\bar{r}) = \alpha^*_{M'}r^*_{M'}\rho_{M'}(r^*_{M'}).$$

We conclude that $\rho_{M'}(\bar{r}) \leq \alpha^*_{M'}\rho_{M'}(r^*_{M'})$, which concludes the proof of this lemma. ■

Lemma 5.1 above implies that instead of solving (NLP1) we can, instead, solve (NLP2). However, (NLP2) is computationally more tractable, and can be solved relatively easy in many scenarios. Specifically, if we Lagrangify Constraint (29) with some Lagrange multiplier Θ, then, the resulting problem is separable in $r_1, \ldots, r_{M'}$. Specifically, we obtain

$$\max_{r_i \in [\Theta A_i, r_\infty)} \sum_{1 \leq i \leq M} (r_i - \Theta A_i)\rho_i(r_i).$$

If the separable maximization problem from above can be solved and the resulting solution is feasible with respect to Constraint (29), we obtain the solution to the KKT conditions. (Observe that the linear qualification constrains always hold in this problem.) In fact, one aims to find the minimal Θ for which the resulting solution satisfies Constraint (29), and this can be done by applying bi-section search on interval $[0, r_\infty]$. The complexity of this procedure depends on the complexity of maximizing $(r_i - \Theta A_i)\rho_i(r_i)$ for each $1 \leq i \leq M$. It is not hard to check that there are at least two tractable cases are:

- $\rho_i(r_i)$ is a concave function on $[0, r_\infty]$ for every $1 \leq i \leq M$;
- $\rho_i(r_i)$ is convex, but $r_i\rho_i(r_i)$ is a concave function on $[0, r_\infty]$, for every $1 \leq i \leq M$.

In both of the two previous cases, objective functions $(r_i - \Theta A_i)\rho_i(r_i)$, $1 \leq i \leq M$, are concave and could be solved using standard methods.

Acknowledgments

The authors would like to thank Dan Adelman, Garud Iyengar and Karl Sigman for fruitful discussions.

References

[1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions*. Dover Publications, Inc., New York, 1974.

[2] D. Adelman. Price-directed control of a closed logistics queueing network. *Operations Research*, 2006.

[3] D. Adelman. A simple algebraic approximation to the erlang loss system. Working paper, submitted, 2006.

[4] D. Y. Burman, J. P. Lehoczky, and Y. Lim. Insensitivity of Blocking Probabilities in a Circuit-Switching Network. *Journal of Applied Probability*, 21:850–859, 1984.

[5] K. L. Chung. *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, New York, 1960.

[6] A. K. Erlang. Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Elektrotkeknikeren*, 13:5–13, 1917.

[7] X. Fan-Orzechowski and E. A. Feinberg. Optimality of randomized trunk reservation for a problem with a single constraint. *Advances in Applied Probability*, 38, 2006.

[8] P. J. Hunt and C. N. Laws. Optimization via trunk reservation in single resource loss systems under heavy traffic. *Annals of Applied Probability*, 7:1058–1079, 1997.

[9] G. Iyengar and K. Sigman. Exponential penalty function control of loss networks. *Annals of Applied Probability*, 14:1698–1740, 2004.

[10] J. S. Kaufman. Blocking in a shared resources environment. *IEEE Transactions on Communications*, 29:1474–1481, 1981.

[11] F. P. Kelly. *Reversability and Stochastic Networks*. John Wiley & Sons, Inc., 1979.

[12] F. P. Kelly. Blocking probabilities in large circuit-switched networks. *Advances in Applied Probability*, 18:473–505, 1986.

[13] F. P. Kelly. Effective bandwidths at multi-class queues. *Queueing Systems*, 9:5–16, 1991.

[14] P. Key. Optimal Control and Trunk Reservation in Loss Networks. *Probability in Engineering and Informations Sciences*, 4:203–242, 1990.

[15] S. Kumar, R. Srikant, and P. Kumar. Bounding blocking probabilities and throughput in queueing networks with buffer capacity constraints. *Queueing Systems*, 28:55–77, 1998.

[16] G. Louth, M. Mitzenmacher, and F. Kelly. Computational complexity of loss networks. *Theoretical Computer Science*, 125:45–59, 1994.

[17] B. Miller. A queueing reward system with several customer classes. *Management Science*, 16:234–245, 1969.

[18] A. A. Puhalskii and M. I. Reiman. A critically loaded multirate link with trunk reservation. *Queueing Systems*, 28:157–190, 1998.

[19] K. Ross and D. Tsang. The stochastic knapsack problem. *IEEE Transactions on Communications*, 37:740–747, 1989.

[20] K. Ross and D. Yao. Monotonicity properties for the stochastic knapsack. *IEEE Transactions on Information Theory*, 36:1173–1179, 1990.

[21] B. A. Sevastyanov. An ergodic theorem for Markov processes and its application to telephone systems with refusals. *Theory of probability and its applications*, 2:104–112, 1957.

[22] A. Siegel. Median bounds and their application. In *Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 776–785, 1999.

[23] W. Whitt. Blocking when service is required from several facilities simultaniously. *AT&T Technical Journal*, 64:1807–1856, 1985.

[24] R. W. Wolff. *Stochastic Modeling and Theory of Queues*. Prentice Hall, 1989.

[25] S. Zachary. On blocking in loss networks. *Advances in Applied Probability*, 23:355–372, 1991.