Branching Bandits and Klimov's Problem: Achievable Region and Side Constraints

Dimitris Bertsimas, Ioannis Ch. Paschalidis, *Student Member, IEEE*, and John N. Tsitsiklis, *Member, IEEE*

Abstract— **We consider the average cost branching bandits problem and its special case known as Klimov's problem. We consider the vector n whose components are the mean number of bandits (or customers) of each type that are present. We characterize fully the achievable region, that is, the set of all possible vectors n that can be obtained by considering all possible policies. While the original description of the achievable region involves exponentially many constraints, we also develop an alternative description that involves only $O(R^2)$ variables and constraints, where R is the number of bandit types (or customer classes). We then consider the problem of minimizing a linear function of n subject to L additional linear constraints on n. We show that optimal policies can be obtained by randomizing between $L+1$ strict priority policies that can be found efficiently (in polynomial time) using linear programming techniques.**

I. Introduction

CONSIDER a single-server multiclass M/GI/1 queue with Bernoulli feedback. In this context, one wishes to determine a policy which optimizes a linear combination of the mean number of customers of the different classes that are present in the system. This problem was posed and solved by Klimov [10], who established the optimality of strict priority rules. In addition, he developed a fairly simple and efficient one-pass algorithm that determines an optimal priority ordering. A shorter and simpler proof can be found in [14].

In the branching bandits problem, as defined by Weiss [18], there is again a single server who serves several customer classes and a similar performance criterion. At each service completion, however, the served customer is replaced by a random number of customers of every other class. This model is more general than Klimov's in that the random numbers of new customers need not correspond to Poisson arrival processes.

The branching bandits and Klimov's problems have important applications in many situations where a single server has to be optimally allocated among various customer classes. As an example, consider a machine in a job-shop manufacturing floor that processes a variety of parts. Klimov's model can be also viewed as being a network of queues with a single server in the network, where external arrivals are Poisson and the routing between the various queues Bernoulli.

Both problems can be extended by imposing some additional linear side constraints. For example, we might require that the mean queue length is the same for each customer class. Such side constraints are usually meant to represent fairness constraints.

Much of the work on the branching bandits and Klimov's problems views these problems as extensions of the classical multi-armed bandit problem [6], [17], [18]. In this paper, however, we take a philosophically very different approach. In particular, we consider the vector n whose components are the mean number of customers of each type that are present and characterize fully the achievable region, that is, the set of all possible vectors n that can be obtained by considering all possible policies. Our characterizations are polyhedral; that is, they are expressed in terms of linear equality and inequality constraints. We are thus able to convert a difficult stochastic control problem to one of optimizing a linear cost function over the achievable region, and this is a linear programming problem. There has already been a fair amount of research on such polyhedral characterizations, which we now discuss.

Gelenbe and Mitrani [7] used conservation laws to show that the performance region of a multiclass queue (without feedback) can be described as a polyhedron. Closer to the subject of this paper, Tsoucas [16] has derived a structural characterization of the achievable region for Klimov's problem, but without giving explicit formulas for some of the constants in his characterization. The idea of conservation laws was generalized by Federgruen and Groonvelt [5], Shantikumar and Yao [15], and Bertsimas and Niño-Mora [2]. In [2] also, an explicit characterization of the achievable region for Klimov's problem is obtained. Finally, the authors, in [3] and [4], have used quadratic potential functions to develop conservation laws for general controlled multiclass queueing networks with Poisson arrivals and exponential service times. In the network case, these conservation laws do not provide an exact characterization of the achievable region but lead to bounds for the achievable region which are often quite tight. For the special case of Klimov's problem in which service times are exponential and preemption is allowed, the potential method of [3] and [4] was shown to lead to an exact characterization of the achievable region.

Given that the achievable region is a polyhedron, the problem of finding an optimal policy amounts to a linear

Manuscript received April 22, 1994; revised March 8, 1995. Recommended by Associate Editor, P. Nain. This work was supported in part by a Presidential Young Investigator Award DDM-9158118 with matching funds from Draper Laboratory, by the Leaders for Manufacturing Program at MIT, and by ARO Grant DAAL-03-92-G-0115.

The authors are with the Laboratory for Information and Decision Systems and Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

IEEE Log Number 9415387.

programming problem. Since it is already known that optimal policies are strict priority rules, it is hardly surprising that the extreme points of the achievable region are the performance vectors of such priority rules. Note that if linear side constraints are imposed, the performance of an optimal policy is still a linear programming problem. In particular, an optimal policy can be expressed as a policy that randomizes between a number of strict priority rules. In addition, the problem of finding the probability with which each particular priority rule is to be used is the same as the problem of expressing an element of a polyhedron as a convex combination of its extreme points. This latter problem can be also solved, in principle, using linear programming techniques.

Unfortunately, the polyhedral characterizations discussed so far involve a number of constraints which is exponential in the number of customer classes. Therefore, even though linear programming problems are solvable in polynomial time, the naive application of the preceding ideas to the side-constrained problem leads to exponential time algorithms. For this reason, we use an alternative method developed by the authors [4] and Kumar and Kumar [9] whereby the achievable region is bounded in terms of a new polyhedron Q that involves a number of variables and constraints which is quadratic in the number of customer classes. We establish in this paper that the achievable region is equal to the image of such a polyhedron Q under a linear mapping into a lower-dimensional space. In particular, the side-constrained problem can be now solved efficiently as a linear programming problem involving the polyhedron Q. As will be shown later, some of the extreme points of Q do not correspond to strict priority rules. Thus, although we can express any element of Q as a combination of its extreme points, this does not solve for us the problem we are actually interested in: expressing an element of the achievable region as a combination of its extreme points. Later in this paper, we will manage to develop a polynomial time algorithm for the latter problem; as it turns out, this is much more complicated than it might have appeared at first sight.

We refer briefly to some earlier work on variations of the Klimov's problem involving side constraints. Nain and Ross [13] consider a multiclass M/GI/1 queue with a single side constraint and establish that an optimal policy randomizes between two priority policies. Makowski and Shwartz [11] derive similar structural results for the Klimov's problem; their methods are easily generalized to the branching bandits model as well. Nevertheless, in the absence of a polyhedral characterization of the achievable region, their methods do not seem to lead to usable algorithms for computing the optimal cost or an optimal policy, especially when more than one side constraints are present.

We wish to summarize at this point the technical contributions of this paper:

1) We derive a "parsimonious" characterization of the achievable region for the branching bandits problem, involving only a quadratic number of variables and constraints. This should be contrasted with all previous work in which the characterizations involve an exponential number of constraints.
2) We extend the methodology developed in [3] and then refined in [4] and [9] to characterize the achievable region of stochastic systems with general distributions; earlier work could only handle exponential distributions.
3) We give a polynomial time algorithm to solve the branching bandit problem with side constraints. More generally, we derive a polynomial time algorithm for expressing an element of a polyhedron as a convex combination of its extreme points, when the polyhedron is specified as the projection of a higher-dimensional polyhedron. This algorithm could be of independent interest.

The rest of the paper is organized as follows: In Section II, we formally define the problem and establish our notation. In Section III, we characterize the achievable region for the vector n^+ of mean queue lengths as observed on a typical service completion time. In Section IV, the same achievable region is described as a projection of a higher-dimensional polyhedron. In Section V, we provide analogs of the results of Sections III and IV, regarding the achievable region for the vector n of mean queue lengths. In Section VI, we discuss how to specialize the results of Section V to Klimov's problem. In Section VII, we bring side constraints into the picture and establish the structure of optimal policies. In addition, we develop a polynomial time algorithm for computing the coefficients needed to specify an optimal policy. Section VIII contains some concluding remarks.

II. Problem Formulation

In this section, we define the average cost branching bandits problem, as well as the special case known as Klimov's problem. We also define our notation and terminology.

Let there be given a set $\mathcal{R}_0 = \{0, 1, 2, \cdots, R\}$ of $R+1$ customer classes and a single server who keeps serving available customers. We assume that there is always an available customer. At any service completion time, the server chooses a customer, say of class i, to serve next. The duration of that customer's service is a positive, arbitrarily distributed, random variable T_i. At the time of the service completion, the customer just served disappears and is replaced by $N_{i0}, N_{i1}, \cdots, N_{iR}$, customers of classes $0, 1, \cdots, R$, respectively, with each N_{ij} a nonnegative integer, arbitrarily distributed, random variable. For any $i \in \mathcal{R}_0$, we assume that the joint distribution of the random variables $(T_i, N_{i1}, \cdots, N_{iR})$ is given and is the same each time a class i customer is served. We also assume that the realizations of the random vector $(T_i, N_{i1}, \cdots, N_{iR})$ corresponding to services of different customers (of the same or of different classes) are statistically independent.

The model just described assumes that the service of a customer cannot be interrupted, which means that we are only considering nonpreemptive policies. Finally, we assume that N_{00} is equal to 1, with probability 1, and that $N_{i0} = 0$ for every $i \neq 0$. Thus, if we start with a single customer of class 0, there will always be exactly one such customer; in particular, our assumption that there is always an available customer is satisfied.

We now define Klimov's problem and then argue that it is a special case of the branching bandits model. We have a single server who serves customers belonging to a set $\mathcal{R} = \{1, \cdots, R\}$ of different customer classes. Customers of each class $i \in \mathcal{R}$ arrive in the system according to an independent Poisson process with rate λ_i and require a random, arbitrarily distributed, service time with mean m_i and second moment σ_i^2. The service times of the customers of each class are independent and identically distributed. Service times of customers of different classes are independent. Finally, service times are independent of the arrival process. Upon service completion, a class i customer is fed back to the system as a class j customer, with probability p_{ij}, or leaves the system, with probability $p_{i0} = 1 - \sum_{j=1}^{R} p_{ij}$. We assume again that service is nonpreemptive. At any service completion time, the server can choose an available customer, if any, to be served next. It can also decide to stay idle. If it decides to stay idle, it is natural to stay idle until the "state" of the system changes, and this can only happen if there is a new arrival. We therefore impose the additional assumption that an idle period can only be terminated by a new arrival. We would like to point out that the M/GI/1 setting is the most general setting that one can hope obtaining results for the Klimov's problem. As a counterexample consider a multiclass G/GI/1 queue with class dependent service requirements and note that conservation laws (in the form of Theorem 5.3) do not hold.

We now indicate how Klimov's model can be obtained as a special case of our variant of the branching bandits model. We identify idling in Klimov's problem with serving a class 0 customer in the branching bandits model. Since idling is supposed to last until the next arrival, T_0 has an exponential distribution with parameter $\lambda = \lambda_1 + \cdots + \lambda_R$. In addition, the vector $(N_{01}, \cdots, N_{0R})$ is the jth unit vector with probability λ_j/λ. (This is the probability that the arriving customer that interrupts the idling period is of class j.) We also let $N_{00} = 1$ and $N_{i0} = 0$ for $i \neq 0$. If a class i customer is served, the mean service time is $E[T_i] = m_i$ and the second moment is σ_i^2. Finally, N_{ij}, for $i, j \neq 0$, is equal to the number of class j Poisson arrivals during the service time T_i, to which number we must add one if the customer served was transformed to a class j customer. In particular, we have

$$E[N_{ij}] = m_i\lambda_j + p_{ij}, \qquad i, j = 1, \cdots, R, \tag{1}$$

$$E[N_{ij}^2] = \lambda_j^2\sigma_i^2 + m_i\lambda_j + p_{ij} + 2m_i\lambda_j p_{ij}, \quad i, j = 1, \cdots, R. \tag{2}$$

(In deriving the last formula, we have used the fact that the second moment of the number of Poisson arrivals with rate λ_j, during the service time T_i is $\lambda_j^2\sigma_i^2 + m_i\lambda_j$.)

Here upon and for the rest of the paper we develop our theory for the more general model of branching bandits. We revisit Klimov's problem in Section VI to show how our results can be specialized to it. On a notational comment, all the vectors defined in this paper are assumed to be column vectors. Let $N_r(t)$ be the number of class r customers present in the system at time t, assumed to be a right-continuous function of time. In particular, if τ is a service completion time, then $N_r(\tau)$ refers to the number of customers right after the service completion. The vector $N(t) = (N_1(t), \cdots, N_R(t))$ will be called the state of the system at time t. (By our assumptions, $N_0(t)$ is the same for all times and, therefore, does not need to be included in the state vector.) Finally, let $\{\tau_k\}$ be the sequence of service completion times.

Definition 2.1:

a) We say that a policy gives priority to class i over class j if there is zero probability of choosing a class j customer to serve while class i customers are available.
b) We say that a policy is nonidling if it gives priority to class i over class 0, for all $i \neq 0$.
c) For any subset S of $\{1, \cdots, R\}$, we say that a policy is an S-priority policy if it gives priority to class i over class j for every $i \in S$ and every $j \notin S$.
d) We say that a policy is a priority policy if it is nonidling and there exists some ordering $(i_1, i_2, \cdots, i_R)$ of the set $\{1, \cdots, R\}$ such that the policy gives priority to class i_k over class i_{k+1}, for $k = 1, \cdots, R-1$.

Assumption A:

a) The $R \times R$ matrix $\boldsymbol{N}$ with entries $E[N_{ij}], i, j = 1, \cdots, R$, has spectral radius smaller than one.
b) The random variables N_{ij} and T_i are of exponential type for every i and j; that is, there exists some $\lambda > 0$ such that $E[e^{\lambda N_{ij}}] < \infty$ and $E[e^{\lambda T_i}] < \infty$.

Part b) of the above assumption is much stronger than needed, but we introduce it to avoid certain technical digressions. It intuitively states that the random variables involved in the model have a finite moment generating function in a neighborhood of zero. In the last section of the paper, we comment on how this assumption can be relaxed.

Assumption A guarantees that the stochastic process $N(\tau_k)$ is "stable" under all nonidling policies [2]. For a self-contained proof, let $w = (w_1, \cdots, w_R)$ be a positive vector and δ be a positive scalar satisfying

$$\sum_{j=1}^{R} E[N_{ij}]w_j \leq w_i - \delta, \qquad i = 1, \cdots, R.$$

[Such a vector exists by the Perron–Frobenius theorem and Assumption A-a).] It follows that for every nonidling policy and for every time τ_k for which $N(\tau_k) \neq 0$, we have

$$\sum_{i=1}^{R} E[N_i(\tau_{k+1}) \mid N(\tau_k)]w_i \leq \sum_{i=1}^{R} N_i(\tau_k)w_i - \delta.$$

Thus, $\sum_{i=1}^{R} N_i(\tau_k)w_i$ has negative drift away from the origin. In particular, if $N(\tau_k)$ is a Markov chain under the policy under consideration (in which case, we say that the policy is Markovian), this Markov chain is geometrically ergodic [8], [12] and all of its moments are finite under the corresponding ergodic distribution.

Let Π^+ be the set of all stationary policies that result into a discrete-time stochastic process $\{N(\tau_k)\}_{k=-\infty}^{\infty}$ with a unique stationary distribution satisfying $E[N_i^2(\tau_k)] < \infty$ for all $i \in \{1, \cdots, R\}$. According to the preceding discussion, Π^+ contains all nonidling stationary Markovian policies. For any

policy $\pi \in \Pi^+$, let n_i^+ be the expectation of $N_i(\tau_k)$ under the corresponding stationary distribution. Let $n^+ = (n_1^+, \cdots, n_R^+)$. Let X^+ (respectively, X_{ni}^+) be the set of all vectors n^+ that can be obtained by considering different policies in Π^+ (respectively, nonidling policies in Π^+). We will refer to X^+ (respectively, X_{ni}^+) as the achievable region for n^+ under all (respectively, nonidling) policies. A complete characterization of X^+ and X_{ni}^+ is obtained in the next section.

The performance vector n^+ refers to the average number of customers of each class that are present in the system at a typical completion time. Alternatively, we may be interested in n, the steady-state mean of $N(t)$. We let Π be the set of all stationary policies that result into a continuous time stochastic process $\{N(t)\}_{t=-\infty}^{\infty}$ with a unique stationary distribution satisfying $E[N_i^2(t)] < \infty$ for all $i \in \{1, \cdots, R\}$. Under Assumption A, every nonidling policy can be shown to belong to Π. This is shown in Lemma 5.2. The achievable region for n under policies in Π (respectively, under nonidling policies in Π) is denoted by X (respectively, by X_{ni}). These regions are studied in Section V.

Table I provides a brief summary of our notation.

TABLE I
NOTATION SUMMARY

n^+	vector of average number of customers at service completions
n	vector of steady-state mean number of customers
ρ^+	vector of traffic intensities at service completions
ρ	vector of steady-state traffic intensities
X^+ (resp. X_{ni}^+)	achievable region for n^+ (resp. under nonidling policies)
X (resp. X_{ni})	achievable region for n (resp. under nonidling policies)
P^+ (resp. P_{ni}^+)	exponential in size characterization of X^+ (resp. X_{ni}^+)
P (resp. P_{ni})	exponential in size characterization of X (resp. X_{ni})
U^+ (resp. U_{ni}^+)	polynomial in size characterization of X^+ (resp. X_{ni}^+)
U (resp. U_{ni})	polynomial in size characterization of X (resp. X_{ni})

III. Derivation of the Achievable Region for n^+

The line of development in this section is as follows. We first derive a set of linear inequalities that have to be satisfied by the vector n^+ under every policy. These constraints define a polyhedron, and we show that its extreme points are the vectors n^+ corresponding to priority policies. We then conclude that the achievable region is equal to this polyhedron.

We start with a few definitions. We use $\chi_i(t)$ to denote the indicator function of the event that at time t the server is serving a customer of class i. We assume that $\chi_i(\cdot)$ is a right-continuous function of time so that $\chi_i(\tau_k)$ is one if at time τ_k a class i customer starts being served. For any policy in Π^+, we let

$$\rho_i^+ = E[\chi_i(\tau_k)]$$

where the expectation is taken with respect to the stationary distribution. The next lemma states that ρ_i^+ is the same for all policies. The proof, as well the proofs of several other results, relies on the following formula that describes the evolution of the system

$$N_i(\tau_{k+1}) = N_i(\tau_k) + \sum_{j=0}^{R} \chi_j(\tau_k)(N_{ji} - \delta_{ij}) \tag{3}$$

where δ_{ij} is the Kronecker delta.[1]

Lemma 3.1: The value of ρ_i^+ is the same for all policies in Π^+ and can be obtained as the unique solution of the system of equations

$$\sum_{j=0}^{R} \rho_j^+ E[N_{ji}] = \rho_i^+, \qquad i = 1, \cdots, R \tag{4}$$

[1] Strictly speaking, we should have used a notation like $N_{ji}(\tau_k)$ instead of simply N_{ji} to indicate the fact that N_{ji} is selected independently after each service completion of a class j customer.

and

$$\sum_{i=0}^{R} \rho_i^+ = 1. \tag{5}$$

Proof: Fix a policy in Π^+. By taking expectations of both sides of (3) with respect to the stationary distribution, we obtain (4). Equation (5) follows from the definition of ρ_i^+.

Let $\rho = (\rho_1^+, \cdots, \rho_R^+)$, and let $u = (E[N_{01}], \cdots, E[N_{0R}])$. Then, (4) can be rewritten as

$$\rho' \boldsymbol{N} + \rho_0 u' = \rho'$$

where y' denotes the transpose of a vector y. Because of Assumption A-a), the matrix $\boldsymbol{I} - \boldsymbol{N}$ is invertible and $(\boldsymbol{I} - \boldsymbol{N})^{-1} = \boldsymbol{I} + \boldsymbol{N} + \boldsymbol{N}^2 + \cdots$ is a nonnegative matrix. We therefore have $\rho' = \rho_0 u'(\boldsymbol{I} - \boldsymbol{N})^{-1} = \rho_0 w'$, where w' is the nonnegative row vector $u'(\boldsymbol{I} - \boldsymbol{N})^{-1}$. Equation (5) can then be used to determine ρ_0 uniquely. ■

For the remainder of the paper, we impose the following assumption which is meant to exclude certain degenerate cases.

Assumption B: For every class $i \in \{0, 1, \cdots, R\}$, we have $\rho_i^+ > 0$.

Under Assumption A, the system is stable and we are guaranteed that $\rho_0^+ > 0$. We then see that Assumption B is guaranteed to hold if the vector u is nonzero and the matrix $\boldsymbol{I} + \boldsymbol{N} + \boldsymbol{N}^2 + \cdots$ is positive.

Let S be some nonempty subset of $\{1, \cdots, R\}$. We define a set of parameters f_{Si}^+, $i \in S$, by means of the system of equations

$$1 + \sum_{i \in S} E[N_{ji}] f_{Si}^+ = f_{Sj}^+, \qquad j \in S. \tag{6}$$

Notice that this is a linear system of the form $(\boldsymbol{I} - \boldsymbol{A})x = e$, where e is a vector with all entries equal to one. Here $\boldsymbol{A}$ is a square submatrix of the nonnegative matrix $\boldsymbol{N}$ which has been assumed to have spectral radius less than one. It follows that the spectral radius of $\boldsymbol{A}$ is also less than one, $\boldsymbol{I} - \boldsymbol{A}$ is invertible, and $(\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2 + \cdots$ is a nonnegative matrix. This establishes that the coefficients f_{Sj}^+ are uniquely defined and are nonnegative. We then use (6) once more to conclude that the coefficients f_{Sj}^+ are in fact positive. We note that f_{Sj}^+ can be interpreted as the expected number of customers served under an S-priority policy until we run out of customers whose class belongs to S and if we started with a single customer of class j.

Theorem 3.2: For every nonempty subset S of $\mathcal{R} = \{1, \cdots, R\}$, and any policy in Π^+, we have

$$\sum_{i \in S} f_{Si}^+ n_i^+ \geq G^+(S) \tag{7}$$

where

$$G^+(S) = \frac{1}{2} \sum_{j=0}^{R} \rho_j^+ E\left[\left(\sum_{i \in S} f_{Si}^+ (N_{ji} - \delta_{ij})\right)^2\right].$$

Inequality (7) holds with equality if and only if we have an S-priority policy.

Proof: Let $R_S(t) = \sum_{i \in S} f_{Si}^+ N_i(t)$. We use (3) and obtain

$$\begin{aligned} R_S(\tau_{k+1}) &= R_S(\tau_k) + \sum_{i \in S} f_{Si}^+ \sum_{j=0}^{R} \chi_j(\tau_k)(N_{ji} - \delta_{ij}) \\ &= R_S(\tau_k) + \sum_{j=0}^{R} \chi_j(\tau_k) \sum_{i \in S} f_{Si}^+ (N_{ji} - \delta_{ij}). \end{aligned} \tag{8}$$

We square both sides of (8), use the fact $\chi_i(\tau_k)\chi_j(\tau_k) = \delta_{ij}\chi_i(\tau_k)$, and take expectations with respect to the stationary distribution corresponding to the policy under consideration. Using also the fact $E[R_S^2(\tau_{k+1})] = E[R_S^2(\tau_k)]$, we obtain

$$\begin{aligned} 2E\left[R_S(\tau_k) \sum_{j=0}^{R} \chi_j(\tau_k) \sum_{i \in S} f_{Si}^+ (N_{ji} - \delta_{ij})\right] \\ + \sum_{j=0}^{R} \rho_j^+ E\left[\left(\sum_{i \in S} f_{Si}^+ (N_{ji} - \delta_{ij})\right)^2\right] = 0. \end{aligned} \tag{9}$$

Notice that the second term in the left-hand side of (9) is $2G^+(S)$, by definition. We now have

$$\begin{aligned} &E[R_S(\tau_k)] \\ &\geq E\left[R_S(\tau_k) \sum_{j \in S} \chi_j(\tau_k)\right] \\ &= -E\left[R_S(\tau_k) \sum_{j \in S} \chi_j(\tau_k) \sum_{i \in S} f_{Si}^+ (N_{ji} - \delta_{ij})\right] \\ &= G^+(S) + E\left[R_S(\tau_k) \sum_{j \notin S} \chi_j(\tau_k) \sum_{i \in S} f_{Si}^+ (N_{ji} - \delta_{ij})\right] \\ &= G^+(S) + E\left[R_S(\tau_k) \sum_{j \notin S} \chi_j(\tau_k) \sum_{i \in S} f_{Si}^+ N_{ji}\right] \\ &\geq G^+(S). \end{aligned}$$

The first inequality follows from $\sum_{j \in S} \chi_j(\tau_k) \leq 1$, the first equality from (6) since for $j \in S$ it holds $\sum_{i \in S} f_{Si}^+ (E[N_{ji}] - \delta_{ij}) = -1$, the second equality from (9), and the third equality because $i \in S$ and $j \notin S$ imply $\delta_{ij} = 0$.

Notice that the equality $E[R_S(\tau_k)] = G^+(S)$ is obtained if and only if

$$R_S(\tau_k) \sum_{j \notin S} \chi_j(\tau_k) = 0, \qquad \text{w.p.1}$$

equivalently, if and only if $N_i(\tau_k)\chi_j(\tau_k) = 0$ for all $i \in S$ and $j \notin S$. This is equivalent to having an S-priority policy. ■

Notice that nonidling policies are the same as $\mathcal{R}$-priority policies. It follows that the inequality $\sum_{i=1}^{R} f_{\mathcal{R}i}^+ n_i^+ \geq G^+(\mathcal{R})$ becomes an equality if and only if the policy is nonidling.

Theorem 3.2 provides us with $2^R - 1$ linear inequality constraints on the vector n^+, one for each nonempty subset of $\{1, \cdots, R\}$. These inequality constraints define a polyhedron in R-dimensional space, which we will denote by P^+. Let us also define P_{ni}^+ as the subset of P^+ on which the equality $\sum_{i=1}^{R} f_{\mathcal{R}i}^+ n_i^+ = G^+(\mathcal{R})$ holds. (Note that P_{ni}^+ is a bounded polyhedron while P^+ is unbounded.) Theorem 3.2 establishes that $X^+ \subset P^+$ and $X_{ni}^+ \subset P_{ni}^+$. We wish to show that $X^+ = P^+$ and $X_{ni}^+ = P_{ni}^+$; that is, that we have a complete characterization of the achievable region for the branching bandits problem under general (or nonidling, respectively) policies. Our first step is to characterize the extreme points of P_{ni}^+.

Theorem 3.3: A vector is an extreme point of the set P_{ni}^+ if and only if it is equal to the performance vector n^+ corresponding to some priority policy.

Proof: Given a set of inequality constraints that define a polyhedron, we say that a constraint is active at those points at which it is satisfied with equality. Recall that an element of a polyhedron in $\mathbb{R}^R$ is an extreme point if and only if there are R linearly independent constraints that are active at that point.

Consider the priority policy corresponding to the ordering $(1, 2, \cdots, R)$. This policy is an S-priority for every set S of the form $\{1, \cdots, i\}$ and the inequality $\sum_{i \in S} f_{Si}^+ n_i^+ \geq G^+(S)$ is satisfied with equality for every such S. Notice that the R equalities thus obtained form a triangular system of equations and are therefore linearly independent. It follows that the vector n^+ is an extreme point of P_{ni}^+. The same argument applies to any other priority policy.

To show that every extreme point corresponds to a priority policy, we observe that P_{ni}^+ satisfies the definition of an extended polymatroid (see [1] for a definition) and the result follows from Theorem 1 of [2]. We provide here an alternative self-contained proof.

Let us introduce the additional assumption that under any policy and for any i, j, there is a positive probability that customers of classes i and j may coexist. Consider an extreme point of P_{ni}^+ that corresponds to some priority policy, say the priority policy corresponding to the ordering $(1, 2, \cdots, R)$. Theorem 3.2 implies that the constraints $\sum_{i \in S} f_{Si}^+ n_i^+ = G^+(S)$ are active, for every S of the form $S = \{1, \cdots, i\}$. If there are more than R active constraints at n^+, we must also have $\sum_{i \in S} f_{Si}^+ n_i^+ = G^+(S)$ for some $S \subset \{1, \cdots, R\}$ which is not of this form; in particular, there exist i and j such that $i < j$, $i \notin S$ and $j \in S$. Thus, j must have priority over i. On the other hand, since $i < j$, i must also have priority over j. This can only happen if customers of classes i and j can

never coexist under the priority policy under consideration, which contradicts our earlier assumption. We conclude that there are exactly R active constraints at every extreme point corresponding to a priority policy.

We say that two extreme points are adjacent if there are $R-1$ constraints that are active at both points. Since the constraint corresponding to $S = \{1, \cdots, R\}$ is satisfied at all points, it follows that an extreme point can have up to $R-1$ adjacent extreme points. We say that two priority policies are adjacent if one can be obtained from the other by interchanging the order of two classes that are ordered consecutively. [For example, the priority ordering (1, 2, 3, 4) is adjacent to (1, 3, 2, 4) but is not adjacent to (1, 3, 4, 2).] It is seen that for adjacent priority policies there are $R-1$ common active constraints, and therefore the corresponding extreme points are adjacent. We conclude that if we have an extreme point that corresponds to a priority policy, all of its $R-1$ adjacent extreme points correspond to priority policies. It is well known that if we keep moving from an extreme point of a bounded polyhedron to an adjacent extreme point, every extreme point can be reached. Therefore, all extreme points of P^+_{ni} correspond to priority policies.

Let us now return to the general case in which we allow the possibility that two customer types may have zero probability of coexisting. Let us introduce a perturbed system, parameterized by a small positive parameter ϵ and for which the random number $N_{ij}(\epsilon)$ of type j customers due to a service completion of a type i customer is given by

$$N_{ij}(\epsilon) = \begin{cases} N_{ij}, \forall j, & \text{with probability } 1-\epsilon, \\ 1, \forall j, & \text{with probability } \epsilon \end{cases}$$

where the N_{ij} have the same distribution as in the original system. Given our assumption that the matrix $\mathbf{N}$ has spectral radius less than one and using the continuity of the spectral radius, it follows that the perturbed system also satisfies the same assumption. Note that if $\epsilon = 0$, we recover the original system.

Consider the coefficients $f^+_{Si}(\epsilon)$ defined for the perturbed system as in (6) and let $P^+_{ni}(\epsilon)$ be the associated polyhedron. It is easily seen that the moments of $N_{ij}(\epsilon)$ depend continuously on ϵ. Hence $f^+_{Si}(\epsilon)$ and $G^+(S, \epsilon)$ are continuous functions of ϵ. Thus, all of the coefficients involved in the linear constraints that define $P^+_{ni}(\epsilon)$ depend continuously on ϵ.

Consider an extreme point n^+ of P^+_{ni}. It is easily shown that n^+ is the limit, as $\epsilon \downarrow 0$, of an extreme point $n^+(\epsilon)$ of $P^+_{ni}(\epsilon)$. Given what we have proved earlier, it follows that for every $\epsilon > 0$, $n^+(\epsilon)$ is the performance vector associated to some priority policy. Since there are finitely many priority policies and by restricting to a sequence of ϵ's that converges to zero, we can assume that every $n^+(\epsilon)$ is the performance vector of the same priority policy, for the ϵ-perturbed system. Without loss of generality, let us assume that this is the priority policy corresponding to the ordering $1, \cdots, R$. Theorem 3.2 yields

$$\sum_{i \in S} f^+_{Si}(\epsilon) n^+_i(\epsilon) = G^+(S, \epsilon), \quad S = \{1, \cdots, k\}, k = 1, \cdots, n.$$

By taking the limit as ϵ converges to zero, we obtain

$$\sum_{i \in S} f^+_{Si} n^+_i = G^+(S), \qquad S = \{1, \cdots, k\}, \; k = 1, \cdots, n.$$

Using Theorem 3.2 once more, we conclude that n^+ is the performance vector associated to the same priority policy, for the original system. ∎

Corollary 3.4: There holds $X^+_{ni} = P^+_{ni}$.

Proof: From Theorem 3.2, we have $X^+_{ni} \subset P^+_{ni}$. Consider a collection of priority policies $\pi^1, \cdots, \pi^K$ whose performance vectors are $x^1, \cdots, x^K$. Consider also a policy that at the beginning of every busy period[2] decides with probability p_i that policy π^i will be followed for the entire duration of the busy period. It is then easily seen that this is a nonidling policy, and its performance vector is $\sum_{i=1}^K p_i x^i$. This establishes that every element of P^+_{ni} is the performance vector of some nonidling policy of this type. ∎

We note that in the preceding proof, a value of K larger than $R+1$ is never needed, by virtue of Caratheodory's theorem.

We now turn our attention to policies that are not necessarily nonidling. We first extend Theorem 3.3.

Theorem 3.5: The polyhedra P^+_{ni} and P^+ have the same set of extreme points.

Proof: At any extreme point of P^+_{ni} there are R linearly independent active constraints, and therefore we also have an extreme point of P^+. We now prove the converse. If P^+ has more extreme points than P^+_{ni}, then there are two adjacent extreme points of P^+ such that one, call it x, is an extreme point of P^+_{ni} and the other, call it y, is not. Assume for simplicity that x is associated to the priority ordering $(1, 2, \cdots, R)$. From the point x, we can move to an adjacent extreme point (along an edge) if exactly one of the active constraints becomes inactive. If any constraint other than the constraint $\sum_{i=1}^R f^+_{\mathcal{R}i} n^+_i \geq G^+(\mathcal{R})$ becomes inactive, we end up at another extreme point of P^+_{ni}. Therefore, to reach y, the constraint $\sum_{i=1}^R f^+_{\mathcal{R}i} n^+_i \geq G^+(\mathcal{R})$ must become inactive. Recall that the active constraints at the point x form a triangular system of equations. Therefore, by making the constraint $\sum_{i=1}^R f^+_{\mathcal{R}i} n^+_i \geq G^+(\mathcal{R})$ inactive, the variable n^+_R becomes free. The value of that variable can be increased without limit without violating any of the constraints associated with P^+. This means that the corresponding edge that starts at x does not end at another extreme point. ∎

We will next characterize the points that lie on infinite edges of P^+. We first need to define a set of policies pertinent to this problem. Consider an ordering σ of the classes $1, \cdots, R$, and relabel the classes such that $\sigma = (1, 2, \cdots, R)$. Let $\pi(p)$ be the policy under which:

a) Class i always has priority over class j, if $i < j \leq R$.
b) The policy never idles when there are available customers of some class $i < R$.
c) Whenever all available customers are of class R, there is a constant probability p of idling.[3]

We refer to all such policies as almost-priority policies.

[2] A busy period starts at a moment where a zero state vector becomes nonzero; it ends at the first time that the state becomes again zero.

[3] Note that this is the same as the Markovian policy that uses the priority ordering $(1, 2, \cdots, R, 0)$ with probability $1-p$ and the priority ordering $(1, 2, \cdots, R-1, 0, R)$ with priority p.

Recall that the vector n^+ associated to a priority policy can be obtained by solving a triangular system of linear equations. We will now describe a procedure for determining the vector n^+ associated with an almost-priority policy. Let us consider the almost priority policy $\pi(p)$ associated with the ordering $(1,\cdots,R)$. Under this policy, each time that there are only customers of class R available, we will have

$$N_i(\tau_{k+1}) = N_i(\tau_k) + (1-\chi)(N_{Ri} - \delta_{iR}) + \chi N_{0i}, \qquad i = 1,\cdots,R$$

where χ is a binary random variable which is independent of everything else and is equal to one with probability p. Equivalently

$$N_i(\tau_{k+1}) = N_i(\tau_k) + (1-\chi)N_{Ri} + \chi(N_{0i} + \delta_{Ri}) - \delta_{Ri}, \qquad i = 1,\cdots,R.$$

This implies that under policy $\pi(p)$, the system evolves exactly the same as if there were no idling and N_{Ri} were replaced by $\tilde{N}_{Ri} = (1-\chi)N_{Ri} + \chi(N_{0i} + \delta_{Ri})$, for $i = 1,\cdots,R$. Therefore, the vector n^+ associated with an almost priority policy can be found by evaluating the vector n^+ associated with a priority policy in a new branching bandits problem with a different distribution for the random variables N_{Ri}, $i = 1,\cdots,R$. In the new branching bandits problem, the matrix $\boldsymbol{N}$ is replaced by a matrix $\tilde{\boldsymbol{N}}(p)$ that differs from $\boldsymbol{N}$ only at the last row; in particular, the (R, j)th entry of $\tilde{\boldsymbol{N}}(p)$ is equal to $(1-p)E[N_{Rj}] + pE[N_{0j}] + p\delta_{Rj}$.

Let us define $p^* = \rho_0^+/(\rho_0^+ + \rho_R^+)$, where the coefficients ρ_i^+ are those corresponding to the original matrix $\boldsymbol{N}$, as in Lemma 3.1. We then have the following result.

Lemma 3.6: The spectral radius of $\tilde{\boldsymbol{N}}(p)$ is less than one for $p < p^*$ and equal to one for $p = p^*$.

Proof: We start from the fact that the coefficients ρ_i^+ satisfy (4), use the definitions of $\tilde{\boldsymbol{N}}(p)$ and p^*, and do some straightforward algebra to verify that the vector $(\rho_1^+, \rho_2^+,\cdots,\rho_{R-1}^+,\rho_0^+ + \rho_R^+)$ is a left eigenvector of $\tilde{\boldsymbol{N}}(p^*)$, with eigenvalue one. In addition, notice that the determinant of $\boldsymbol{I} - \tilde{\boldsymbol{N}}(p)$ is affine in p. Therefore, for every $p \neq p^*$, the determinant of $\boldsymbol{I} - \tilde{\boldsymbol{N}}(p)$ is nonzero and the spectral radius of $\tilde{\boldsymbol{N}}(p)$ is different from one. Since the spectral radius is less than one for $p = 0$ (Assumption A), a continuity argument implies the same for all values of p between zero and p^*. ■

Under the almost-priority policy $\pi(p)$, the values of ρ_i^+ and n_i^+ remain the same for $i = 1,\cdots,R-1$. It remains to determine how ρ_R^+ and n_R^+ vary with p, and we will be using the notation $\rho_R^+(p)$ and $n_R^+(p)$, to make this dependence explicit. In addition, we let $f_{\mathcal{R}i}^+(p)$, $i = 1,\cdots,R$, stand for the unique solution of (6) when N_{ij} is replaced by $\tilde{N}_{ij}(p)$ and when S is equal to $\mathcal{R} = \{1,\cdots,R\}$. Using Cramer's rule, we see that $f_{\mathcal{R}i}^+(p)$ is the ratio of two affine functions of p, with the denominator being the determinant of $\boldsymbol{I} - \tilde{\boldsymbol{N}}(p)$. Since the latter determinant becomes zero when $p = p^*$, we conclude that the denominator can be taken to be $p^* - p$.

We also note that $(1-p)\rho_R^+(p) = \rho_R^+$. (Intuitively, this expresses the fact that a fraction $1-p$ of all class R services in the modified model corresponds to class R services in the original model.) Concerning $n_R^+(p)$, it can be determined from the relation

$$\sum_{i=1}^{R} f_{\mathcal{R}i}^+(p) n_i^+(p) = \frac{1}{2}\sum_{j=0}^{R}\rho_j^+(p)E\left[\left(\sum_{i=1}^{R} f_{\mathcal{R}i}^+(p)(\tilde{N}_{ji} - \delta_{ij})\right)^2\right].$$

Using our earlier discussion on the dependence of $\rho_i^+(p)$ and $f_{\mathcal{R}i}^+(p)$ on p, we conclude that $n_R^+(p)$ is a rational function of p with a term of the form $p^* - p$ appearing in the denominator. This implies that $n_R^+(p)$ tends to infinity as p increases to p^*. In addition, p can be determined from $n_R^+(p)$ by solving a polynomial equation in p.

We summarize this discussion in the following theorem.

Theorem 3.7: Any point on an infinite edge of P^+ is the performance vector of some almost-priority policy. In addition, the value of p that corresponds to any given point can be determined by solving a polynomial equation.

Using the same argument as in the proof of Corollary 3.4, we conclude the following.

Corollary 3.8: There holds $X^+ = P^+$.

IV. A Parsimonious Representation of the Achievable Region

The polyhedra P^+ and P_{ni}^+ provide an exact representation of the achievable regions X^+ and X_{ni}^+, respectively. Their drawback is that they are specified in terms of an exponential number of constraints. In this section, we use the approach of [3] and [9] to obtain an equivalent but more compact representation. This new representation involves $R(R+1)$ variables but only $O(R^2)$ linear constraints.

The achievable region will be represented in terms of the auxiliary variables

$$I_{ji} = E[\chi_j(\tau_k)N_i(\tau_k)], \qquad i = 1,\cdots,R, \quad j = 0,\cdots,R. \tag{10}$$

Let I stand for the $R(R+1)$-dimensional vector with components I_{ij}. Notice that $I_{ji} = 0$ if and only if $N_i(\tau_k) > 0$ implies $\chi_j(\tau_k) = 0$; that is, if and only if class i has priority over class j. In particular, a policy is nonidling if and only if $I_{0i} = 0$ for all $i \neq 0$.

Theorem 4.1: For every policy in Π^+, the vector I belongs to the polyhedron Q^+ defined as the set of all nonnegative vectors z with components z_{ji}, $j = 0,\cdots,R$, $i = 1,\cdots,R$, that satisfy the following linear equality constraints

$$\sum_{j=0}^{R}\rho_j^+ E[(N_{ji} - \delta_{ji})^2] + 2\sum_{j=0}^{R} z_{ji}E[N_{ji} - \delta_{ji}] = 0, \qquad i = 1,\cdots,R \tag{11}$$

and

$$\sum_{j=0}^{R} z_{jr}E[N_{jr'} - \delta_{jr'}] + \sum_{j=0}^{R} z_{jr'}E[N_{jr} - \delta_{jr}] + \sum_{j=0}^{R}\rho_j^+ E[(N_{jr} - \delta_{jr})(N_{jr'} - \delta_{jr'}] = 0, \qquad r, r' = 1,\cdots,R. \tag{12}$$

Proof: Consider the evolution equation (3). We square both sides, we use the fact $\chi_i(\tau_k)\chi_j(\tau_k) = \delta_{ij}\chi_i(\tau_k)$, and we take expectations with respect to the stationary distribution corresponding to the policy under consideration. Using also the fact $E[N_i(\tau_{k+1})] = E[N_i(\tau_k)]$, we obtain (11). (In the derivation of this formula we have also used the independence of N_{ij} from the state of the system.)

To derive (12) we use (3) to derive a recursion for $N_r(\tau_{k+1})N_{r'}(\tau_{k+1})$ and proceed similarly. ■

Note that for every policy in Π^+, we have

$$n_i^+ = \sum_{j=0}^{R} I_{ji}, \qquad i = 1, \cdots, R.$$

In particular, n^+ belongs to the set U^+ defined by

$$U^+ = \left\{ x \geq 0 \mid \exists z \in Q^+ \text{ such that } x_i = \sum_{j=0}^{R} z_{ji} \right\}.$$

The set U^+ is the image of the polyhedron Q^+ under a particular linear mapping. Therefore, U^+ is also a polyhedron.

We have already shown that the achievable region X^+ is contained in U^+. It has been shown in [4], in much greater generality, that the use of auxiliary variables, as in the proof of Theorem 4.1, always provides a smaller polyhedron than the one obtained using the method of the preceding section; thus, $X^+ \subset U^+ \subset P^+$. Since we have shown earlier that $X^+ = P^+$, we have the following main result.

Theorem 4.2: There holds $P^+ = U^+ = X^+$.

Theorem 4.2 states that the achievable region X^+ is the image of the polyhedron Q^+. Given that Q^+ involves a much smaller (quadratic instead of exponential) number of constraints, this representation is much more suitable for the development of efficient algorithms.

A natural question to raise at this point is the following: is it true that every element of Q^+ is equal to the vector I associated to some policy in Π^+? Interestingly enough, the answer is negative, as explained in the Appendix. In other words, the set Q^+ is larger than the achievable region for the vector I, even though its image is exactly equal to the achievable region for the vector n^+. In particular, not every extreme point of Q^+ can be associated with an extreme point of P^+ and a priority policy.

If we are interested in nonidling policies, the preceding results are modified as follows. Notice that a policy is nonidling if and only if $I_{0i} = 0$ for all $i \neq 0$. We define Q_{ni}^+ as the subset of Q^+ in which the additional constraints $z_{0i} = 0$ hold for $i = 1, \cdots, R$. By using the same reasoning as before, we conclude that $X_{ni}^+ = U_{ni}^+ = P_{ni}^+$.

V. Achievable Region for the Mean Queue Lengths

In this section we characterize the achievable region X (respectively, X_{ni}) for the vector n of mean queue lengths, under policies in Π (respectively, under nonidling policies in Π). In fact, we obtain two different characterizations which are similar to the characterizations of X^+ in terms of the polyhedra P^+ and Q^+.

We first establish a connection between the steady-state mean number of customers n_i and the mean number n_i^+ of customers at a typical service completion time. Let us denote by m_j the expectation of the service time T_j for a customer of class $j \in \{0, \cdots, R\}$.

Lemma 5.1: For any policy in Π and for any $i \in \{1, \cdots, R\}$, we have

$$n_i = \frac{\sum_{j=0}^{R} m_j I_{ji}}{\sum_{j=0}^{R} m_j \rho_j^+}.$$

Proof: The general formula for passing from a Palm distribution to a stationary distribution (see, e.g., [17, p. 226]) states that n_i, the steady-state mean of $N_i(t)$, is given by

$$n_i = \frac{E\left[\int_{\tau_k}^{\tau_k+1} N_i(\sigma)\, d\sigma\right]}{E[\tau_{k+1} - \tau_k]}$$

where the expectations are taken with respect to the stationary distribution of the discrete-time Markov chain $N(\tau_k)$. We have $N_i(\sigma) = N_i(\tau_k)$ for $\sigma \in [\tau_k, \tau_{k+1})$, which leaves us with

$$\frac{E[(\tau_{k+1} - \tau_k)N_i(\tau_k)]}{E[\tau_{k+1} - \tau_k]}.$$

Note that $E[\tau_{k+1} - \tau_k] = \sum_{j=0}^{R} m_j \rho_j^+$. Furthermore

$$E[(\tau_{k+1} - \tau_k)N_i(\tau_k)] = \sum_{j=0}^{R} E[(\tau_{k+1} - \tau_k)N_i(\tau_k)\chi_j(\tau_k)] = \sum_{j=0}^{R} m_j I_{ji}$$

which completes the proof. ■

We next show that under Assumption A any nonidling stationary policy belongs to Π.

Lemma 5.2: Under Assumption A, any nonidling stationary policy results into a continuous-time stochastic process $\{N(t)\}_{t=-\infty}^{\infty}$ with a stationary distribution satisfying $E[N_i^2(t)] < \infty$ for all $i \in \{1, \cdots, R\}$.

Proof: We follow the same technique as in the preceding proof. We have

$$E[N_i^2(t)] = \frac{E\left[\int_{\tau_k}^{\tau_k+1} N_i^2(\sigma)\, d\sigma\right]}{E[\tau_{k+1} - \tau_k]}$$

where the expectations in the right-hand side are taken with respect to the stationary distribution of the discrete-time Markov chain $N(\tau_k)$ and the expectations in the left-hand side are taken with respect to the distribution of the continuous-time process $N(t)$. As in the preceding proof, we have $N_i(\sigma) = N_i(\tau_k)$ for $\sigma \in [\tau_k, \tau_{k+1})$, which implies

$$\begin{aligned} E[N_i^2(t)] &= \frac{E[(\tau_{k+1} - \tau_k)N_i^2(\tau_k)]}{E[\tau_{k+1} - \tau_k]} \\ &= \frac{\sum_{j=0}^{R} E[N_i^2(\tau_k)T_j\chi_j(\tau_k)]}{\sum_{j=0}^{R} E[T_j\chi_j(\tau_k)]} \\ &= \frac{\sum_{j=0}^{R} m_j E[N_i^2(\tau_k)\chi_j(\tau_k)]}{\sum_{j=0}^{R} m_j \rho_j^+} \\ &\leq \frac{\max_j m_j}{\sum_{j=0}^{R} m_j \rho_j^+} E[N_i^2(\tau_k)] < \infty \end{aligned}$$

since we have argued in Section II that the discrete-time Markov chain $N(\tau_k)$ is geometrically ergodic. ■

We now define a polyhedron U as the image of Q^+ under the linear mapping suggested by Lemma 5.1. That is

$$U = \left\{ x \geq 0 \mid \exists z \in Q^+ \text{ such that } x_i = \frac{\sum_{j=0}^{R} m_j z_{ji}}{\sum_{j=0}^{R} m_j \rho_j^+} \right\}.$$

If we are interested in nonidling policies only, we define U_{ni} similarly, except that Q^+ is replaced by Q_{ni}^+. Theorem 4.1 and Lemma 5.1 readily imply that the achievable region X (respectively, X_{ni}) is contained in U (respectively, U_{ni}). We intend to show that $U = X$ and $U_{ni} = X_{ni}$. Our first step in this direction is to derive polyhedra P and P_{ni} with structure similar to the polyhedra P^+ and P_{ni}^+ that were derived in Section III.

Let S be a nonempty subset of $\{1, \cdots, R\}$. We define a set of parameters f_{Si}, $i \in S$, by means of the system of equations

$$m_j + \sum_{i \in S} E[N_{ji}] f_{Si} = f_{Sj}, \qquad \forall j \in S. \tag{13}$$

This system of equations has a unique solution, which is positive, for the same reasons that were given when the coefficients f_{Si}^+ were defined.

Theorem 5.3: For every nonempty subset S of $\mathcal{R} = \{1, \cdots, R\}$ and any policy in Π, we have

$$\sum_{i \in S} f_{Si} n_i \geq G(S) \tag{14}$$

where

$$G(S) = \frac{\frac{1}{2}\rho_j^+ E\left[\left\{\sum_{r \in S} f_{Sr}(N_{jr} - \delta_{jr})\right\}^2\right]}{\sum_{w=0}^{R} m_w \rho_w^+}.$$

Inequality (14) holds with equality if and only if we have an S-priority policy.

Proof: Consider a policy $\pi \in \Pi$ and a subset S of $\mathcal{R}$. Then, the vector I, with components I_{ij} satisfies (11) and (12). We multiply (12) by $f_{Sr} f_{Sr'}$ and sum over all r, $r' \in S$ such that $r > r'$. We then obtain

$$\sum_{j=0}^{R} \Bigg(\sum_{r' \in S} f_{Sr'} E[N_{jr'} - \delta_{jr'}] \sum_{\{r \in S \mid r > r'\}} f_{Sr} I_{jr} + \sum_{r \in S} f_{Sr} E[N_{jr} - \delta_{jr}] \sum_{\{r' \in S \mid r' < r\}} f_{Sr'} I_{jr'} + \rho_j^+ \sum_{r \in S} \sum_{\{r' \in S \mid r' < r\}} f_{Sr} f_{Sr'} \cdot E[(N_{jr} - \delta_{jr})(N_{jr'} - \delta_{jr'})] \Bigg) = 0. \tag{15}$$

We also multiply (11) by $f_{Sr'}^2$ and sum over all $r' \in S$ to obtain

$$\sum_{j=0}^{R} \left(\sum_{r' \in S} f_{Sr'}^2 E[N_{jr'} - \delta_{jr'}] I_{jr'} + \frac{1}{2}\rho_j^+ \sum_{r' \in S} f_{Sr'}^2 E[(N_{jr'} - \delta_{jr'})^2] \right) = 0. \tag{16}$$

We interchange r and r' in the second term of (15) and add the result to (16) to obtain

$$\sum_{j=0}^{R} \left(\sum_{r' \in S} f_{Sr'} E[N_{jr'} - \delta_{jr'}] \sum_{r \in S} f_{Sr} I_{jr} + \frac{1}{2}\rho_j^+ E\left[\left\{\sum_{r \in S} f_{Sr}(N_{jr} - \delta_{jr})\right\}^2\right] \right) = 0$$

which yields

$$\sum_{j \in S} \sum_{r' \in S} f_{Sr'} E[N_{jr'} - \delta_{jr'}] \sum_{r \in S} f_{Sr} I_{jr} + \sum_{j \notin S} \sum_{r' \in S} f_{Sr'} E[N_{jr'} - \delta_{jr'}] \sum_{r \in S} f_{Sr} I_{jr} + AG(S) = 0$$

where A is defined by $A = \sum_{w=0}^{R} m_w \rho_w^+$. Using (13), we obtain

$$\sum_{r \in S} f_{Sr} \sum_{j \in S} m_j I_{jr} = \sum_{j \notin S} \sum_{r' \in S} f_{Sr'} E[N_{jr'} - \delta_{jr'}] \sum_{r \in S} f_{Sr} I_{jr} + AG(S). \tag{17}$$

We now recall Lemma 5.1 and observe that

$$\sum_{j \in S} m_j I_{jr} \leq A n_r. \tag{18}$$

Thus, we obtain

$$\sum_{r \in S} f_{Sr} n_r \geq \frac{1}{A} \sum_{j \notin S} \sum_{r' \in S} f_{Sr'} E[N_{jr'} - \delta_{jr'}] \cdot \sum_{r \in S} f_{Sr} I_{jr} + G(S) \geq G(S) \tag{19}$$

because I_{jr}, $f_{Sr'}$ are nonnegative and $\delta_{jr'} = 0$ for $j \notin S$ and $r' \in S$. It is easily checked that the inequalities in (19) hold with equality if and only if $I_{jr} = 0$ for $j \notin S$ and $r \in S$, that is, if and only if the policy under consideration is an S-priority. ■

Since nonidling policies are the same as $\mathcal{R}$-priority policies, the inequality $\sum_{i \in \mathcal{R}} f_{\mathcal{R}i} n_i \geq G(\mathcal{R})$ becomes an equality if and only if the policy is nonidling. Theorem 5.3 provides us with $2^R - 1$ linear inequality constraints on the vector $n = (n_1, \cdots, n_R)$. These constraints define a polyhedron in R-dimensional space which we denote by P. We also define P_{ni} to be the subset of P where the equality $\sum_{i \in \mathcal{R}} f_{\mathcal{R}i} n_i = G(\mathcal{R})$ holds. Theorem 5.3 asserts that $X_{ni} \subset P_{ni}$ and $X \subset P$.

The following is our main result.

Theorem 5.4:

a) A vector is an extreme point of the set P_{ni} if and only if it is equal to the performance vector n corresponding to a priority policy.
b) The polyhedra P and P_{ni} have the same set of extreme points.
c) Any point on an infinite edge of P is the performance vector of some almost-priority policy.
d) There holds $P = U = X$ and $P_{ni} = U_{ni} = X_{ni}$.

Proof: (Outline) The proof of parts a), b), and c) is identical to the proof of Theorems 3.3, 3.5, and 3.7, respectively.

Recall that we have already shown that $X \subset U$. Furthermore, in the course of the proof of Theorem 5.3, we showed that every element of U belongs to P. Therefore, we have $X \subset U \subset P$ and $X_{ni} \subset U_{ni} \subset P_{ni}$. On the other hand, part a) of this theorem states that the extreme points of P_{ni} belong to X_{ni}, and it follows that $X_{ni} = P_{ni}$. Similarly, parts b)–c) of this theorem imply that $X = P$. ■

VI. Klimov's Problem Revisited

In the branching bandits problem, the vector $N(t)$ changes only at service completion times. In contrast, in Klimov's problem, external arrivals are Poisson and will generically occur during a service interval. This makes no difference if we are only watching the system at service completion times. In particular, all of the results in Sections III and IV can be specialized to Klimov's problem by using (1) and (2).

Let us now consider the mean number of class i customers present in the system at some typical time t. This is equal to the mean number n_i, as determined from the branching bandits model, plus the expected number a_i of class i customers that have arrived since the last service completion, which occurred at some time τ. We have

$$a_i = \sum_{j=0}^{R} \Pr(\chi_j(t) = 1)\lambda_i E[t - \tau \mid \chi_j(t) = 1].$$

Notice that

$$\Pr(\chi_j(t) = 1) = \frac{m_j \rho_j^+}{\sum_{k=0}^{R} m_k \rho_k^+}.$$

In addition, $E[t - \tau \mid \chi_j(t) = 1] = \sigma_j^2/2m_j$, and this determines a_i completely. Notice that a_i is the same for all policies in II.

VII. Branching Bandits with Side Constraints

In this section, we consider the branching bandits problem in the presence of additional linear constraints on the vector n of mean queue lengths. Let these side constraints be of the form $\boldsymbol{A}n \geq b$, where $\boldsymbol{A}$ is a matrix of dimensions $L \times R$. To keep the discussion simple, we only consider nonidling policies. In view of our characterization of the achievable region (Theorem 5.3), the cost of an optimal policy obeying the side constraints can be found by solving the linear programming problem

$$\begin{aligned} &\text{minimize } c'x \\ &\text{subject to } x \in P_{ni} \\ &\qquad \boldsymbol{A}x \geq b. \end{aligned} \tag{20}$$

We assume that this problem has a feasible solution.

The linear programming problem (20) is hard to solve because the polyhedron P_{ni} is described by an exponential number of constraints. We recall, however, that we have available a parsimonious representation of P_{ni} of the form (Theorem 5.3)

$$P_{ni} = U_{ni} = \{Fz \mid z \in Q_{ni}^+\}$$

where Q_{ni}^+ is a polyhedron described in terms of a quadratic number of variables and constraints and where F is a known linear mapping. It follows that problem (20) is equivalent to the linear programming problem

$$\begin{aligned} &\text{minimize } c'x \\ &\text{subject to } x = Fz \\ &\qquad z \in Q_{ni}^+ \\ &\qquad \boldsymbol{A}x \geq b \end{aligned}$$

which is polynomial time solvable because it only has polynomial number of variables and constraints. We thus assume that we have computed, in polynomial time, an optimal solution x^* of problem (20).

Next, we express x^* as a convex combination of at most $R + 1$ extreme points of P_{ni}. This is always possible, by Caratheodory's theorem. (Later in this section, we show that this can be accomplished in polynomial time.) Let $u^1, \cdots, u^{R+1}$ be these extreme points. Consider the problem

$$\begin{aligned} &\text{minimize } \sum_{j=1}^{R+1} \zeta_j (c'u^j) \\ &\text{subject to } \sum_{j=1}^{R+1} \zeta_j = 1 \\ &\qquad \sum_{j=1}^{R+1} \zeta_j (\boldsymbol{A}u^j) \geq b \\ &\qquad \zeta_j \geq 0. \end{aligned}$$

Since there is a feasible solution of this problem for which $x^* = \sum_j \zeta_j u^j$, the optimal cost is the same as in problem (20), and any optimal solution of the new problem is also an optimal solution of the original problem (20). Consider an optimal basic feasible solution of the new problem, that is, at least $R + 1$ constraints are satisfied with equality. (Such an optimal basic feasible solution can be found in polynomial time because we have $O(R)$ variables and constraints.) In particular, at least $R+1-L-1$ of the constraints $\zeta_j \geq 0$ must be satisfied with equality, which means that at most $L + 1$ of the variables ζ_j are positive. Thus, an optimal solution of the original side-constrained problem (20) can be expressed as a convex combination of no more than $L + 1$ extreme points of P_{ni}. Equivalently, an optimal policy can be obtained by randomizing between no more than $L + 1$ priority policies. We summarize this discussion in the following theorem.

Theorem 7.1: If the side-constrained problem (20) is feasible, then there exists an optimal policy which at the beginning of each busy period selects one of $L + 1$ priority policies, according to some fixed probabilities, and follows this policy

throughout that busy period. Furthermore, such a policy can be found in polynomial time.

The only part of the proof of Theorem 7.1 that we have not yet presented is the fact that once an optimal solution x^* is available, it can be expressed as a convex combination of extreme points $u^1, \cdots, u^{R+1}$ of P_{ni}, in polynomial time. We now show how this can be accomplished.

Let u^1 be an extreme point of P_{ni}. Such an extreme point can be found by choosing an arbitrary priority policy and evaluating its performance vector. If $x^* = u^1$, we are done. If not, let us consider the line from u^1 to x^*, and let us consider the point at which this line exits the feasible set P_{ni}. Such a point exists because P_{ni} is bounded and can be found by solving the linear programming problem

$$\begin{aligned} &\text{maximize } t \\ &\text{subject to } x = x^* + t(x^* - u^1) \\ &\qquad x = Fz \\ &\qquad z \in Q_{ni}^+. \end{aligned} \tag{21}$$

This linear programming problem can be solved in polynomial time. Let (x^1, z^1, t^1) be its optimal solution, where x^1 is easily seen to be unique and different from u^1. The point x^1 lies on the boundary of P_{ni} and in particular, it must lie on a facet of P_{ni}. Furthermore, since $x^1 \neq u^1$, there exists a facet of P_{ni} such that x^1 lies on that facet but u^1 does not. We will now proceed to find such a facet.

One way of finding a facet of P_{ni} with the desired properties is to check each one of the constraints

$$\sum_{i \in S} f_{Si} n_i \geq G(S)$$

that define P_{ni} to see whether they are satisfied by x^1 and u^1. This would take exponential time, however, because there are exponentially many such constraints and a different approach is needed.

Consider the related to (21) linear programming problem in (22)

$$\begin{aligned} &\text{maximize } t \\ &\text{subject to } x = \hat{x}^* + t(x^* - u^1) \\ &\qquad x = Fz \\ &\qquad z \in Q_{ni}^+. \end{aligned} \tag{22}$$

Let us view the optimal solution $(\hat{x}^1, \hat{z}^1, \hat{t}^1)$ of the linear programming problem (22) as a function of the right-hand side vector $\hat{b} = (\hat{x}^*, b')$, where b' is the right-hand side vector corresponding to the constraints $x = Fz$ and $z \in Q_{ni}^+$. Let us consider small perturbations of $\hat{x}^*$ of the form $\hat{x}^* = x^* + \sum_{i=1}^R \epsilon_i e_i$, where $\epsilon_i > 0$, $i = 1, \cdots, R$ are small enough and e_i denotes the ith unit vector. Using the sensitivity analysis of linear programming, and in the absence of degeneracy, the optimal basis, denoted by $\boldsymbol{B}$ is unique and is not altered for the above small perturbations of $\hat{x}^*$. Thus, we have $(\hat{x}^1, \hat{z}^1, \hat{t}^1)' = \boldsymbol{B}^{-1}\hat{b}$, and by decomposing $\boldsymbol{B}^{-1}$, we obtain

$$\begin{bmatrix} \hat{x}^1 \\ \hat{z}^1 \\ \hat{t}^1 \end{bmatrix} = \boldsymbol{B}^{-1} \begin{bmatrix} \hat{x}^* \\ b' \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_1 & & \boldsymbol{B}_2 \\ & \boldsymbol{B}_3 & \end{bmatrix} \begin{bmatrix} \hat{x}^* \\ b' \end{bmatrix}.$$

From the above equation we obtain that

$$\hat{x}^1 = \boldsymbol{B}_1 \hat{x}^* + \boldsymbol{B}_2 b'.$$

Substituting $\hat{x}^* = x^* + \sum_{i=1}^R \epsilon_i e_i$ and using that the vector (x^1, z^1, t^1) optimally solves (21) we finally obtain

$$\hat{x}^1 = x^1 + \sum_{i=1}^R \epsilon_i \boldsymbol{B}_1 e_i. \tag{23}$$

In other words, $\hat{x}^1$ is locally a linear function of $\hat{x}^*$, and this linear function can be found very easily, as in (23). The range of this function is the desired facet. That is, the desired facet is spanned by the vectors $\{\boldsymbol{B}_1 e_i;\ i = 1, \cdots, R\}$. Given this, it is not hard to obtain a constraint of the form $\sum_{i=1}^R a_i x_i = \beta$ which is satisfied by all the points of the facet. In the case where (x^1, z^1, t^1) is a degenerate optimal solution of (21), we first do a preliminary perturbation of x^* to come back to the nondegenerate case and then use the above outlined approach. It is not hard to verify that all of the above can be accomplished in polynomial time. Let, for example, $R = 3$ and assume that the above outlined procedure yields the facet $a_1 x_1 + a_3 x_3 = \beta$. By the structure of P_{ni} (see Theorem 5.3), this facet corresponds to $\{1, 3\}$-priority policies.

Once we have found a facet of P_{ni} to which x^1 belongs, we now proceed to express x^1 as a convex combination of R extreme points of that facet. This is a problem of the same type as the one we were trying to solve but in one dimension less. For the example given above ($R = 3$), we let u^2 be an extreme point lying on the facet $a_1 x_1 + a_3 x_3 = \beta$. Such an extreme point can be found by choosing an arbitrary $\{1, 3\}$-priority policy, say the ordering (3, 1, 2). We thus have a recursive algorithm, consisting of R stages. Each stage only takes polynomial time, and the desired result has been established. That is, we have expressed x^* as a convex combination of extreme points $u^1, \cdots, u^{R+1}$ of P_{ni}, in polynomial time. Moreover, from the above discussion it is clear how policies are associated with these extreme points.

VIII. Concluding Remarks

We have presented a generalization of the potential function method developed in [4] to describe the achievable region of stochastic systems with exponential distributions to systems with general distributions. A challenging open question is to extend the method further to queueing networks with general distributions.

Our main result in the paper is a polynomial reformulation of the branching bandit problem. An exponential characterization of the achievable region has been known partially through the work of Tsoucas [16] and explicitly through the work of Bertsimas and Niño-Mora [2]. In particular, the achievable region is characterized as an extended polymatroid. This raises the question of whether an arbitrary extended polymatroid is always a projection of a higher dimensional polyhedron involving a polynomial number of variables and

constraints. Since polymatroids and extended polymatroids appear in several applications in combinatorial optimization, such a reformulation will be very useful for combinatorial problems with side constraints.

We finally indicate how to relax Assumption A-b) which required the probability distributions of the random variables of interest (N_{ij} and T_i) to be of exponential type (i.e., with finite moment generating function in a neighborhood of zero). Let us only assume that each T_i has finite mean and each N_{ij} has finite mean and variance. If these random variables are not of exponential type, let us approximate them by random variables of exponential type with the same means and variances, and let us take the limit as this approximation becomes better and better. For each approximant, the results we have proved establish that the achievable region will be the same; this is because the constraints that define the achievable region only depend on the means and variances of N_{ij} and the mean of T_i. Taking the limit, and using a continuity argument, the same achievable region is obtained in the limit.

APPENDIX

We show here that not every point in the polyhedron Q_{ni}^+ is equal to the vector I associated to some nonidling policy in Π^+.

Consider a problem in which $R = 3$, and suppose that there is a positive probability that customers of all three classes may coexist, no matter what policy is used. (For this, it is sufficient to assume that $E[N_{01}N_{02}N_{03}] > 0$.) The polyhedron Q_{ni}^+ is described in terms of nine variables z_{ij}, $i, j = 1, 2, 3$, and six constraints.

If we impose the additional constraints $z_{21} = 0$, $z_{31} = 0$, and $z_{32} = 0$, we obtain an extreme point z^* of Q_{ni}^+. This extreme point is in fact the vector I associated with the priority policy corresponding to the ordering (1, 2, 3).

Let us now consider the following policy. We follow the priority ordering (1, 2, 3) except that whenever $N_2 = 0$, class 3 gets priority over class 1. With this policy, we will still have $z_{21} = 0$ and $z_{32} = 0$ but z_{31} will be positive. This shows that the set of points $\{z \in Q_{ni}^+ \mid z_{21} = 0, z_{32} = 0\}$ is an edge of Q_{ni}^+. Given that Q_{ni}^+ is bounded, if we start at z^* and move that edge, we must eventually hit another extreme point. At that extreme point, at least one of the variables z_{11}, z_{22}, z_{33}, z_{12}, z_{23}, or z_{13} is equal to zero. We will argue such a vector cannot be the vector I corresponding to a policy.

Indeed, if $z_{12} = 0$, then the extreme point can only be achieved by a policy that simultaneously satisfies $I_{21} = 0$ and $I_{12} = 0$. Such a policy must give priority to class 1 over class 2 and to class 2 over class 1, which is impossible given our assumption that customers of these two classes will sometimes coexist. If $z_{23} = 0$, the extreme point is not achievable for similar reasons. If $z_{13} = 0$, the extreme point can only be achieved by a policy that satisfies $I_{21} = 0$, $I_{32} = 0$, and $I_{13} = 0$. Such a policy would reach an impasse at times when customers of all three classes are present. Finally note that $I_{ii} > 0$ for every policy because otherwise class i customers would be never served. Thus, extreme points at which $z_{ii} = 0$ for some i are not achievable either, and this concludes the argument.

REFERENCES

[1] P. P. Bhattacharya, L. Georgiadis, and P. Tsoucas, "Extended polymatroids: Properties and optimization," IBM Research Division, T. J. Watson Research Center, Yorktown Heights, NY, Res. Rep., 1992.

[2] D. Bertsimas and J. Nino-Mora, "Conservation laws, extended polymatroids and the multi-armed bandit problem: A unified polyhedral approach," Operations Research Center, MIT, Cambridge, MA, Working Paper, 1992.

[3] D. Bertsimas, I. Ch. Paschalidis, and J. N. Tsitsiklis, "Scheduling of multiclass queueing networks: Bounds on achievable performance," in *Proc. Workshop Hierarchical Contr. Real Time Scheduling Manufacturing Syst.*, Lincoln, NH, Oct. 16–18 1992, extended abstract.

[4] ——, "Optimization of multiclass queueing networks: Polyhedral and nonlinear characterizations of achievable performance," *Annals Applied Prob.*, vol. 4, no. 1, pp. 43–75, 1994.

[5] A. Federgruen and H. Groenevelt, "Characterization and optimization of achievable performance in queueing systems," *Op. Res.*, vol. 36, pp. 733–741, 1988.

[6] J. C. Gittins, *Multi-Armed Bandit Allocation Indices.* New York: Wiley, 1989.

[7] E. Gelenbe and I. Mitrani, *Analysis and Synthesis of Computer Systems.* London: Academic, 1980.

[8] B. Hajek, "Hitting-time and occupation-time bounds implied by drift analysis with applications," *Advances Applied Prob.*, vol. 14, pp. 502–525, 1982.

[9] S. Kumar and P. R. Kumar, "Performance bounds for queueing networks and scheduling policies," *IEEE Trans. Automat. Contr.*, vol. 39, no. 8, pp. 1600–1611, 1994.

[10] G. P. Klimov, "Time-sharing service systems I," *Theory Prob. Applicat.*, vol. XIX, no. 3, 1974.

[11] A. M. Makowski and A. Shwartz, "On constrained optimization of the Klimov network and related Markov decision processes," *IEEE Trans. Automat. Contr.*, vol. 38, no. 2, pp. 354–359, 1993.

[12] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability.* New York: Springer-Verlag, 1993.

[13] P. Nain and K. W. Ross, "Optimal priority assignment with hard constraint," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 883–888, 1986.

[14] P. Nain, P. Tsoucas, and J. Walrand, "Interchange arguments in stochastic scheduling," *J. Applied Prob.*, vol. 27, pp. 815–826, 1989.

[15] J. G. Shantikumar and D. D. Yao, "Multiclass queueing systems: Polymatroid structure and optimal scheduling control," *Op. Res.*, vol. 40, no. 2, pp. 293–299, 1992.

[16] P. Tsoucas, "The region of achievable performance in a model of Klimov," IBM Research Division, T. J. Watson Research Center, Yorktown Heights, NY, Res. Rep., 1991.

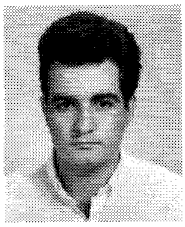
[17] J. Walrand, *An Introduction to Queueing Networks.* Englewood Cliffs, NJ: Prentice-Hall, 1988.

[18] G. Weiss, "Branching bandit processes," *Prob. Eng. Inform. Sci.*, vol. 2, pp. 269–278, 1988.

Dimitris Bertsimas was born in Greece in 1962. He received the B.S. degree in electrical engineering and computer science from the National Technical University of Athens, Greece, in 1985, the M.S. degree in operations research at the Massachusetts Institute of Technology (MIT), Cambridge, MA, in 1987, and the Ph.D. degree in applied mathematics and operations research at MIT in 1988.

Since 1988, he has been with MIT, where he is presently Professor of Operations Research. His research interests include optimization theory and the analysis and control of stochastic systems and finance.

Dr. Bertsimas received the Nicholson prize in 1988 and the Presidential Young Investigator Award in 1991. He is Associate Editor of *Operations Research* and of *Queueing Systems and Applications.*



Ioannis Ch. Paschalidis (S'94) was born in Athens, Greece, in 1968. He received the Professional Diploma degree in electrical and computer engineering from the National Technical University of Athens, Greece, in 1991 and the S.M. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, in 1993.

He is currently a Ph.D. candidate at the Department of Electrical Engineering and Computer Science, MIT, and Research Assistant at the Laboratory for Information and Decision Systems, MIT. His research interests include the analysis and control of stochastic systems with main applications in communication networks and manufacturing systems.

Mr. Paschalidis is a member of INFORMS (former ORSA) and the Technical Chamber of Greece.

John N. Tsitsiklis (S'80–M'83) was born in Thessaloniki, Greece, in 1958. He received the B.S. degree in mathematics (1980), and the B.S. (1980), M.S. (1981) and Ph.D. (1984) degrees in electrical engineering, all from the Massachusetts Institute of Technology (MIT), Cambridge, MA.

During the academic year 1983–84, he was an Acting Assistant Professor of Electrical Engineering at Stanford University, Stanford, CA. Since 1984, he has been with MIT, where he is currently Professor of Electrical Engineering. His research interests are in the areas of systems and control theory, and operations research.

Dr. Tsitsiklis is a coauthor (with D. Bertsekas) of *Parallel and Distributed Computation: Numerical Methods* (1989). He has been a recipient of an IBM Faculty Development Award (1983), an NSF Presidential Young Investigator Award (1986), an Outstanding Paper Award by the IEEE Control Systems Society (for a paper coauthored with M. Athans, 1986), and of the Edgerton Faculty Achievement Award by M.I.T. (1989). He was a plenary speaker at the 1992 *IEEE Conference on Decision and Control*. He is an Associate Editor of *Applied Mathematics Letters* and has been an Associate Editor of IEEE TRANSACTIONS ON AUTOMATIC CONTROL and *Automatica*.