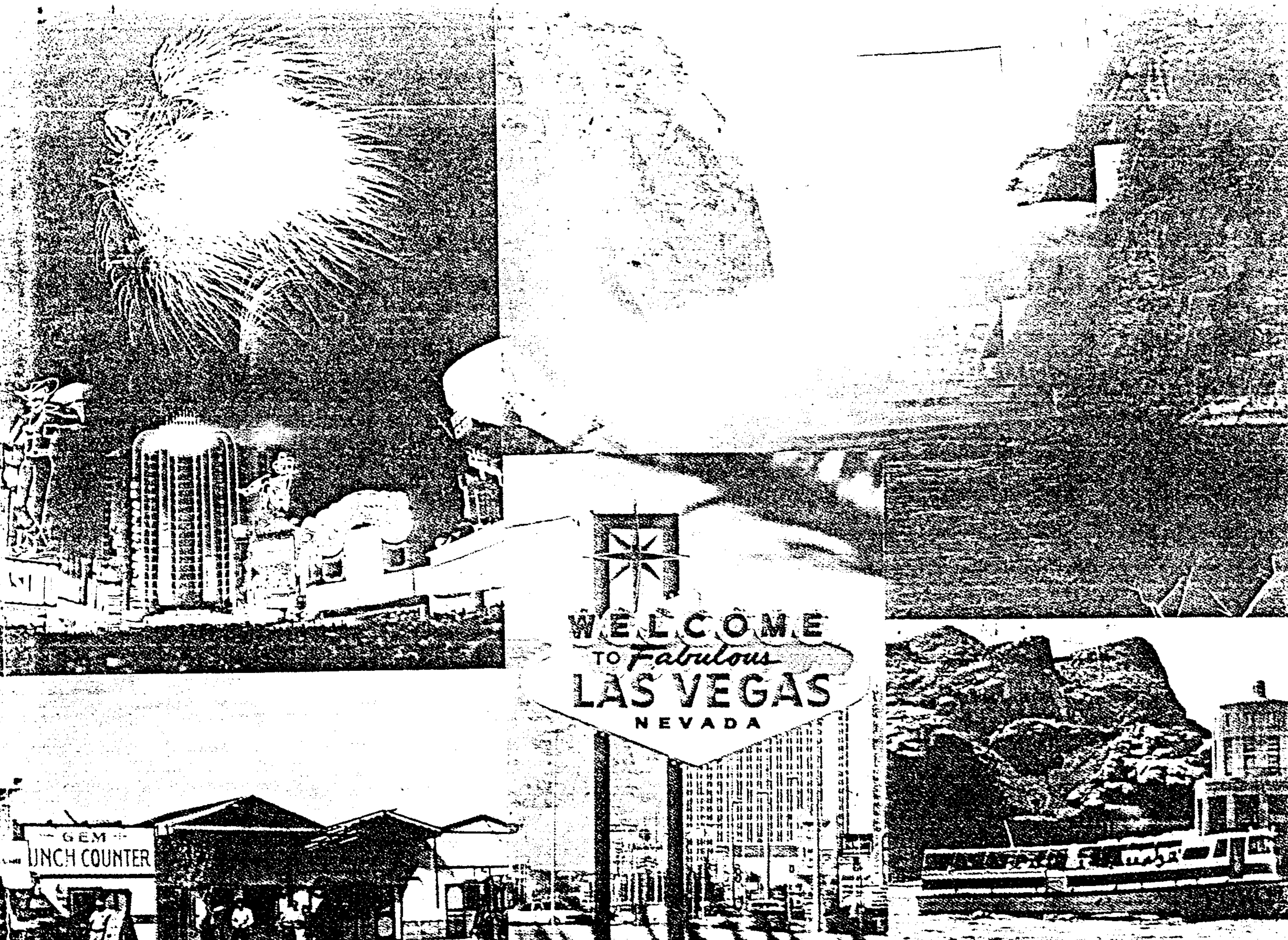


THE 23rd IEEE CONFERENCE ON DECISION AND CONTROL

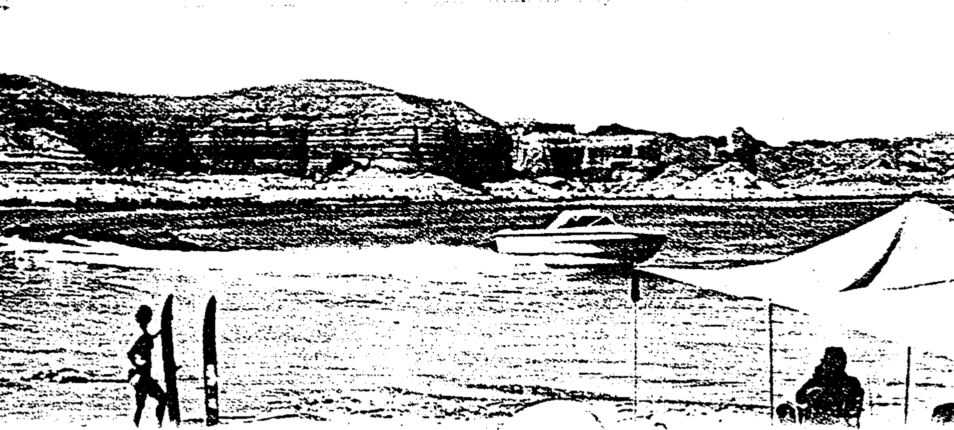
DECEMBER 12-14, 1984 • LAS VEGAS, NEVADA

1884 · 1984
A CENTURY OF ELECTRICAL PROGRESS
IEEE

Volume 2 of 3

84CH2093-3

CSS



Proceedings of 23rd Conference
on Decision and Control
Las Vegas, NV, December 1984

STABILITY, PARAMETER ESTIMATION AND ADAPTIVE CONTROL FOR DISCRETE-TIME COMPETING QUEUES

J.S. Baras[1], A.J. Dorsey[2] and A.M. Makowski[1]

[1]Electrical Engineering Dept., University of Maryland, College Park, MD 20742
[2]IBM-Federal Systems Div., 21 Firstfield Rd., Gaithersburg, MD 20748

ABSTRACT

In this paper, several discrete-time queues compete at the beginning of each time slot for the attention of a single server with infinite capacity. The service requirements are geometric with queue-dependent parameters, independent from customer to customer and independent from the arrivals. The stability of such systems is characterized under the action of an arbitrary, not necessarily Markovian, service allocation policy. Maximum likelihood estimates are obtained in explicit form for the service rates and their strong consistency is established. The results are then applied to the problem of adaptively controlling such systems when the service rates are unknown, under the long-run average performance criterion with instantaneous cost linear in the queue sizes.

I. INTRODUCTION

In the context of resource sharing environments, a great many situations can be modelled by the following scenario: A natural time unit exists and divides the time horizon into contiguous slots of unit length. The system is composed of a single server with the capability of providing several grades of service, each such grade of service being characteristic of a customer class (or equivalently of a queue). New customers arrive on a slot-per-slot basis and await service in an infinite capacity waiting room. The service requirements are geometrically distributed, with class dependent parameters, and are statistically independent from customer to customer as well as from the arrivals. At the beginning of a time slot, a customer class is selected to receive service attention during that time slot on the basis of past decisions, service completions and arrival data. This assignment of service attention may be pre-emptive as a customer may experience an interruption of service before completion of the service requirement; the allocation of effort may also fail to be work-conserving for the server may give service attention to a customer class with an empty queue.

A cost, linear in the queue sizes, is incurred for operating the system over one time slot and the service discipline is selected to minimize the corresponding expected long-run average criterion over an infinite horizon. The policy that allocates service attention to the non-empty queue with the largest expected cost decrease per slot defines the μc-rule. It is known that the μc-rule is optimal among all admissible allocation policies when the arrival streams have arbitrary statistics but are statistically independent of the service requirements [3,8].

The μc-rule is a fixed prioritization scheme determined solely by the service parameters and the cost coefficients, independently of the arrival statistics. However, in many applications, the service parameters are not available to the decision-maker as the system is initially put in operation, and the simple μc-rule is thus not implementable in its given form. Indeed, a learning capability needs to be included into the decision-making mechanism while steering the system to optimality. This suggests that in many instances, a non-Bayesian adaptive formulation might be more realistic and appropriate in studying the problem of optimally controlling such competing queues systems.

The present paper deals with a particular version of this adaptive control problem and discusses various results of system stability, parameter identification and performance optimality under the long-run average criterion. Here, the Certainty Equivalence design philosophy is adopted in conjunction with maximum likelihood estimators. More specifically, at any given time, a maximum likekihood estimate of the service rates is computed on the basis of available information; the decision to be implemented is then generated according to the μc-rule policy for the model that corresponds to these most recent estimate values of the system parameters. This procedure defines the adaptive μc-rule, and the study of its estimation and control performance constitutes the main motivation behind the results reported here.

The maximum likekihood estimates are based on the knowledge of the initial queue sizes and of the past control actions, arrivals and service completion data, as are the control decisions. The discussion is given under the assumption that the arrivals of new customers are independently and identically distributed over successive time slots, with possible correlations between the customer types in a particular time slot. This assumption, when combined with the other model features, allows for the explicit evaluation of the maximum likelihood estimators of the service parameters given an arbitrary control policy.

The strong consistency of these maximum likelihood estimates is investigated in detail and under simple conditions on the system parameters, the property is shown to hold uniformly in the work-conserving control policies. One of these conditions relates to system stability of the competing queues model and provides a complete characterization of (in)stability, as the system is operated with general non-anticipative (non-Markovian) strategies. This stability criterion explicitly involves the mean statistics of the arrival and service processes, and is established by complementing ideas from drift analysis with martingale-theoretic arguments. The method is powerful enough to yield the exact statistics of the passage times to the empty state under arbitrary non-idling policies. This methodology seems to have wide applicability for studying the stability of many queueing systems and elsewhere, and is particularly useful when the model admits a state-space representation via a difference

This work was supported in part by the Department of Energy under Grant DOE-AC01-78ET29244, A5 and the National Science Foundation under Grant ECS-82-04451.

CH2093-3/84/0000-1149 $1.00 © 1984 IEEE

equation as in this work and in [1,16].

The conditions under which strong consistency is shown to hold uniformly in the non-idling control policies, are essentially sufficient to ensure that the adaptive μc-rule is optimal for the long-run average criterion.

In recent years, stochastic adaptive control has received considerable attention, as evidenced by the extensive review of the subject given by Kumar in a recent survey paper [14]. One of the formulations reported there does pertain to the present work; it originates, at least in its general form, in the seminal papers of Mandl [18-20] and deals with the class of controlled Markov chains with a non-Bayesian uncertainty in the model parameters, under a long-run average performance measure. The situation treated here naturally fits into this framework as the multi-queue size process plays the role of the controlled Markov chain and the service rates constitute the unknown model parameters.

In most papers concerned with this class of problems, the corresponding dual control problem of Fel'dbaum [11] is approached by the Certainty Equivalence Principle and various results on performance optimality and parameter identification are available under a variety of technical conditions [6,7,9,13,18-20]. Unfortunately, owing to the special characteristics of the problem at hand, these earlier results do not apply directly and alternate arguments are needed. The main difference between the present paper and previous works on the adaptive control of Markov chains lies in the allowed information pattern for estimation and control purposes. Earlier results all assume the information pattern customarily adopted in the context of Markov decision processes, whereby past states and control values are available. Here, the state of the controlled chain is the queue size vector and knowing initial queue sizes, and past controls, service completion records and arrival data thus defines a much richer information pattern due to the discrete nature of time.

The results reported here (without proofs) extend preliminary work by Baras and Dorsey [2, 10] on the adaptive control of two competing queues with geometric service requirements. A detailed discussion of the results is available in [5].

II. THE MODEL

In this section, a simple mathematical model is formulated to capture the evolution of the multi-queue system described in the introduction. Throughout this paper, the number of competing queues is fixed and is denoted by the positive integer K.

An underlying sample space Ω is given with a σ-field $\mathcal{F}$ of events and is assumed to simultaneously carry an $\mathbb{N}^K$-valued random variable (RV) Ξ, a sequence $\{A(t)\}_1^\infty$ of $\mathbb{N}^K$-valued RV's and a sequence $\{B(n)\}_1^\infty$ of $\{0,1\}^K$-valued RV's. As a notational convention, the k-th component of an $\mathbb{R}^K$-valued RV (resp. element in $\mathbb{R}^K$) is always denoted by the same symbol as the RV (resp. element) subscripted by k. The initial size (at time t=1) of the k-th queue is represented by Ξ_k and the RV $A_k(t)$ quantifies the arrivals to the queue during the time slot [t,t+1), whereas $B_k(n)$ records service completion in the slot during which the k-th queue is non-empty and is given service attention for the n-th time.

The assignment of service attention in the slot t,t+1) will be based on knowledge of the initial queue sizes, and of past control values, service completion and arrival data over the horizon [1,t). Here, an admissible policy π is thus any sequence $\{\pi_t\}_1^\infty$ of mappings π_t from $\mathbb{N}^K \times (\{1,\dots,K\} \times \{0,1\}^K \times \mathbb{N}^K)^{t-1}$ into $\{1,\dots,K\}$, with the domain of definition of π_1 being simply $\mathbb{N}^K$. The collection of all such policies is denoted by $\mathcal{P}$ in the sequel.

For every policy π in $\mathcal{P}$, the sequences $\{X^\pi(t)\}_1^\infty$, $\{U^\pi(t)\}_1^\infty$, $\{N^\pi(t)\}_1^\infty$ and $\{D^\pi(t)\}_1^\infty$ of RV's are defined below with values in $\mathbb{N}^K$, $\{1,\dots,K\}$, $\mathbb{N}^K$ and $\{0,1\}^K$, respectively. The RV $X_k^\pi(t)$ represents the number of customers present in the k-th queue at the end of the horizon [1,t), and $U^\pi(t)$ specifies the customer class to receive service attention in the slot [t, t+1); the RV $N_k^\pi(t)$ counts the number of slots over the time period [1,t) during which the k-th queue was non-empty and was given service attention, while $D_k^\pi(t)$ encodes the departure of a type k customer from the system at the end of the time slot [t,t+1). To initialize the recursion, define $X^\pi(1) := \Xi$ and $U^\pi(1) := \pi_1[\Xi]$, and for $1 \le k \le K$, set $N_k^\pi(1) := 0$ and $D_k^\pi(1) := 1[\Xi_k \neq 0]\delta[k, U^\pi(1)] B_k(N_k^\pi(1)+1)$. Here, the expression $1[A]$ denotes the characteristic function of the event A in $\mathcal{F}$ and $\delta[\,,\,]$ stands for the standard Kronecker symbol. Then, for all $t \neq 0$ in $\mathbb{N}$ and $1 \le k \le K$, set

$$X_k^\pi(t+1) := X_k^\pi(t) + A_k(t) - D_k^\pi(t), \tag{2.1}$$

$$U^\pi(t+1) := \pi_{t+1}[\xi; U^\pi(s), D^\pi(s), A(s),\ 1 \le s \le t], \tag{2.2}$$

$$N_k^\pi(t+1) := N_k^\pi(t) + 1[X_k^\pi(t) \neq 0]\delta[k, U^\pi(t)], \tag{2.3}$$

$$D_k^\pi(t+1) := 1[X_k^\pi(t+1) \neq 0]\delta[k, U^\pi(t+1)] B_k(N_k^\pi(t+1)+1). \tag{2.4}$$

An admissible policy π in $\mathcal{P}$ is idling at time t if there exists k and ℓ, $1 \le k \neq \ell \le K$, such that for some ω in Ω, $X_k^\pi(t,\omega) > 0$, $X_\ell^\pi(t,\omega) = 0$ but $U^\pi(t,\omega) = \ell$. An admissible policy π in $\mathcal{P}$ which is not idling at any time is called a non-idling or work conserving policy.

To capture the randomness affecting the multi-queue system studied here, a collection $\{P^\mu,\ \mu \in [0,1]^K\}$ of probability measures is postulated on $(\Omega, \mathcal{F})$ with the property that under each measure P^μ, μ in $[0,1]^K$, the following statistical statements hold true:

(A1): The sequences $\{B_k(n)\}_1^\infty$ are mutually independent Bernoulli sequences with parameter μ_k, $1 \le k \le K$;

(A2): The RV Ξ, the arrival sequence $\{A(t)\}_1^\infty$ and the Bernoulli sequences $\{B(n)\}_1^\infty$ are mutually independent;

(A3): The RV's $\{A(t)\}_1^\infty$ form a sequence of independent identically distributed $\mathbb{N}^K$-valued RV's, with known probability distribution q_A given by

$$q_A(a) := P^\mu[A(t) = a] \tag{2.5}$$

for all a in $\mathbb{N}^K$ and all $t \neq 0$ in $\mathbb{N}$;

(A4): The RV Ξ has a known probability distribution q_Ξ given by

$$q_{\Xi}(x) := P^{\mu}[\Xi = x] \qquad (2.6)$$

for all x in N^K.

The described model is to be put in contrast with the single probability triple model adopted in previous studies 3,4 of the competing queues system. The approach taken here incorporates the fact that the statistical system characteristics may be known to the decision-maker only up to a parametrization. Specifically, the situation captured in the assumptions (A1)-(A4) is one where initial conditions and arrival patterns have a known and fully specified statistical description, whereas the service completion mechanism is specified only within the class of memoryless or Bernoulli mechanisms, parametrized by the service rates μ. Under the assumptions made, the service duration of any given customer in the k-th queue is clearly a geometric RV with parameter $_K\mu_k$, $1\le k\le K$, and each probability measure P^{μ}, μ in $[0,1]^K$, thus fully characterizes a statistical model for the competing queues system described in the introduction. Throughout this paper, the true service parameters are denoted by μ^o_k, $1\le k\le K$.

III. THE MAXIMUM LIKELIHOOD ESTIMATES

When the service rate vector is unknown to the decision-maker, learning has to take place in conjunction with the control process. Here the principle of maximum likelihood is invoked to generate simple and explicit estimates for these quantities.

III.1 The Principle of Maximum Likelihood

The estimation procedure relies on an information pattern described by the following simple chain of events: As the admissible policy π in P operates over the horizon $[1,t+1)$, the initial condition Ξ and the data $\{(U^{\pi}(s),D^{\pi}(s),A(s))\}$, $1\le s\le t$ become available by time t+1 through perfect monitoring of the system. On the basis of this information, an estimate $\mu^{\pi}(t+1)$ is generated for the unknown service rate vector μ^o by maximizing a likelihood functional evaluated on the observed data trajectory and is used on the interval $[t+1, t+2)$. At time t=1, only the initial condition Ξ is assumed known and an initial guess is made to produce the initial vector estimate $\mu^{\pi}(1)$ which is used on the interval $[1,2)$.

For all t 0 in N, the likelihood functional L^{π}_{t+1} is defined by the formula

$$L^{\pi}_{t+1}[\mu;x,u(s),d(s),a(s),1\le s\le t] \qquad (3.1)$$
$$:= P^{\mu}[\Xi=x;U^{\pi}(s)=u(s),D^{\pi}(s)=d(s),A(s)=a(s),1\le s\le t]$$

with μ in $[0,1]^K$, x in N^K and $\{(u(s),d(s),a(s)),\ 1\le s\le t\}$ in $\{1,\dots,K\}\times\{0,1\}^K\times N^K$, and the obvious convention that $L^{\pi}_1[\mu;x] := P^{\mu}[\Xi=x] = q_{\Xi}(x)$.

For each policy π in P and every parameter μ in $[0,1]^K$, the R_+-valued RV's $\{L^{\pi}(t;\mu)\}_1^{\infty}$ are defined by

$$L^{\pi}(t+1;\mu) := L^{\pi}_{t+1}[\mu;\ \Xi,U^{\pi}(s),D^{\pi}(s),A(s),1\le s\le t] \qquad (3.2)$$

for all $t\ne0$ in N, with $L^{\pi}(1;\mu) := L^{\pi}_1[\mu;\Xi]$. In other words, $L^{\pi}(t+1;\mu)$ is the RV obtained by composing the likelihood functional evaluated at μ with the initial condition Ξ and the data $\{(U^{\pi}(s),D^{\pi}(s),A(s))\},\ 1\le s\le t\}$. The Principle of Maximum Likelihood states that the parameter estimate $\mu^{\pi}(t)$ of the true service rates μ^o should be selected so as to maximize $L^{\pi}(t;\mu)$, i.e., for all $t\ne0$ in N,

$$L^{\pi}(t;\mu^{\pi}(t)) = \max_{\mu\in[0,1]^K}\{L^{\pi}(t;\mu)\}. \qquad (3.3)$$

III.2 An Expression for the Likelihood Functionals

As a first step of the methodology outlined above, an explicit expression is now derived for the likelihood functionals. To do this, it is convenient to develop some additional notation:

For every admissible policy π in P, the sequences $\{V^{\pi}(t)\}_1^{\infty}$ and $\{B^{\pi}(t)\}_1^{\infty}$ of $\{0,1\}^K$-valued RV's are defined componentwise for all $t\ne0$ in N by

$$V^{\pi}_k(t) := 1[X^{\pi}_k(t)\ne0]\,\delta[k,U^{\pi}(t)],\quad 1\le k\le K, \qquad (3.4)$$

$$B^{\pi}_k(t) := B_k(N^{\pi}_k(t)+1),\quad 1\le k\le K. \qquad (3.5)$$

Moreover, define the R_+-valued mapping by G

$$G[\mu;d,v] := \prod_{k=1}^{K}[\mu_k d_k v_k + (1-d_k)(1-\mu_k v_k)] \qquad (3.6)$$

for all μ in $[0,1]^K$, d in $\{0,1\}^K$ and v in $\{0,1\}^K$.

Given these preliminaries, explicit expressions for the likelihood functionals can now be compactly stated:

Theorem 3.1. For every admissible policy π in P and all $t\ne0$ in N, the likelihood functionals (3.2) are given by the formula

$$L^{\pi}(t+1;\ ) = [\prod_{s=1}^{t} G[\mu;D^{\pi}(s),V^{\pi}(s)]][\prod_{s=1}^{t} q_A(A(s))]L^{\pi}(1;\mu) \qquad (3.7)$$

for all rate vectors μ in $[0,1]^K$.

III.3 The Maximum Likelihood Estimation Scheme

With the formula (3.7) available for the likelihood functionals, an explicit expression can now be obtained for the maximum likelihood estimators. The result is summarized in the following proposition.

Theorem 3.2. For any admissible policy π in P, the maximum likelihood estimates $\{\mu^{\pi}(t)\}_1^{\infty}$ are given by

$$\mu^{\pi}_k(t+1) = \begin{cases} \dfrac{\sum_{s=1}^{t} D^{\pi}_k(s)}{\sum_{s=1}^{t} V^{\pi}_k(s)} & \text{if } \sum_{s=1}^{t} V^{\pi}_k(s)>0 \quad (3.8a) \\ \text{arbitrary in } [0,1] & \text{if } \sum_{s=1}^{t} V^{\pi}_k(s)=0 \quad (3.8b) \end{cases}$$

for all $t\ne0$ in N and $1\le k\le K$, with $\mu^{\pi}(1)$ chosen arbitrarily in $[0,1]$.

The exact expressions (3.8) for the maximum likelihood estimates readily suggest a recursive implementation for the sequential estimation process. To see this observe from (2.3), (3.4) and (3.5) that for every admissible policy π in P, the equality

$$N_k^\pi(t+1) = \sum_{s=1}^{t} V_k^\pi(s), \quad 1\le k\le K, \tag{3.9}$$

holds for all $t\neq 0$ in $\mathbb{N}$. It is then easy to conclude from (3.8) that the recursion

$$\mu_k^\pi(t+1) = \frac{\mu_k^\pi(t) + \dfrac{D_k^\pi(t)}{N_k^\pi(t)}}{1 + \dfrac{V_k^\pi(t)}{N_k^\pi(t)}} \tag{3.10}$$

holds provided $N_k^\pi(t)>0$.

It is noteworthy that for the system at hand, expressions such as (3.8) and (3.10) are not available as soon as the information pattern is modified. This fact was clearly illustrated by Dorsey in his doctoral dissertation [10] where the two competing queues situation was studied and maximum likelihood estimates computed for the service rates under the standard state feedback information structure. The reason behind the simplifications observed here is easily traced. Indeed, under the information pattern described in Section III.1, the parameter estimation problem is really one of estimating the rate of the Bernoulli sequences $\{B_k(n)\}_1^\infty$ sampled under the policy π at the times t where $V_k^\pi(t)\neq 0$. The results of Theorem 3.2 are thus not surprising in that light.

IV. THE STABILITY RESULTS

The results of this section are concerned with system stability and are valid under any one measure (or model assumption) P^μ, μ in $[0,1]^K$, with the properties described in Section II. Throughout this section, a given service parameter μ in $[0,1]^K$ is held fixed and all probabilistic statements are made under the probability measure P^μ, with corresponding expectation operator denoted by E^μ.

As the discussion shows, the stability of the competing queues system (2.1)-(2.4) is completely characterized by a single quantity ρ^μ defined by the relation

$$\rho^\mu := \sum_{k=1}^{K} \frac{\alpha_k}{\mu_k}, \tag{4.1}$$

where α_k denotes the (possibly infinite) mean of each one of the RV's $\{A_k(t)\}_1^\infty$, $1\le k\le K$.

Under the action of an arbitrary policy π in $\mathcal{P}$, the queue sizes process does not usually enjoy the Markov property and standard methods from the theory of Markov chains are thus seemingly not applicable to obtain a characterization of system stability. Here, instead, ideas from drift analysis and direct martingale-theoretic arguments are used. Although drift analysis was originally formulated for Markov chains [15,17,21,22], it has also proved successful in handling some non-Markovian systems as recently demonstrated by Hajek [12]. Systems with explicit state-space dynamics such as (2.1)-(2.4) appear to constitute a large natural class of models to which this circle of ideas are applicable [1,16]. Indeed, this explicit state-space representation of the queue dynamics is most helpful in generating (super)martingales of interest in the study of (in)stability.

The point of departure lies in the intuitive idea that stability should be related to the frequency with which the system visits the empty state: For every admissible policy π in $\mathcal{P}$, the filtration $\{F_t^\pi\}_1^\infty$ is defined on (Ω,F) by setting

$$F_{t+1}^\pi := \sigma\{\Xi, U^\pi(s), B^\pi(s), A(s), 1\le s\le t\} \tag{4.2}$$

for all $t\neq 0$ in $\mathbb{N}$, with $F_1^\pi := \sigma\{\Xi\}$. Observe from the assumptions (A1)-(A3) that for every admissible policy π in $\mathcal{P}$, the relations

$$E^\mu[X_k^\pi(t+1)\,|\,F_t^\pi] = X_k^\pi(t) + \alpha_k - \mu_k V_k^\pi(t), \quad 1\le k\le K, \tag{4.3}$$

hold for all $t\neq 0$ in $\mathbb{N}$. Moreover, whenever the policy π is non-idling at time t, the identity

$$\sum_{k=1}^{K} V_k^\pi(t) = 1 - 1[X_k^\pi(t)\neq 0,\ 1\le k\le K] \tag{4.4}$$

holds. With drift analysis in mind, the relations (4.3) and (4.4) together suggest that the information as to how often the system empties itself under an admissible policy π, is carried by the sequence $\{Z^\pi(t)\}_1^\infty$ of $\mathbb{R}_+$-valued RV's, with

$$Z^\pi(t) := \sum_{k=1}^{K} \frac{X_k^\pi(t)}{\mu_k} \tag{4.5}$$

for all $t\neq 0$ in $\mathbb{N}$. This quantity will act as a drift functional and satisfies the following recursion

$$Z^\pi(t+1) = Z^\pi(t) + \sum_{k=1}^{K} \frac{A_k(t)}{\mu_k} - \sum_{k=1}^{K} \frac{V_k^\pi(t)}{\mu_k} B_k^\pi(t). \tag{4.6}$$

As pointed out by Hajek [12], the evolution of the drift sequence (4.5) is best understood by studying the corresponding conditional probability generating function. Under the assumptions (A1)-(A4), it is an easy exercise to conclude from (4.6) that

$$E[z^{Z^\pi(t+1)}\,|\,F_t^\pi] = a(z)b(z;V^\pi(t))z^{Z^\pi(t)} \quad \text{a.s.} \tag{4.7}$$

with the argument z restricted to the interval $(0,1]$. In (4.7), the quantities $a(z)$ and $b(z;v)$ are defined for all $0<z\le 1$ and v in $\{0,1\}^K$, to be

$$a(z) := E^\mu[z^{\sum_{k=1}^{K} \frac{A_k(t)}{\mu_k}}], \tag{4.8}$$

$$b(z;v) := E^\mu[z^{-\sum_{k=1}^{K} \frac{v_k}{\mu_k} B_k(n)}], \tag{4.9}$$

where the right handsides of (4.8) and (4.9) do not depend on t and n as a result of the assumptions (A1)-(A3).

IV.1 The Case $\rho^\mu>1$: an Instability Result

The region of system instability is determined first in the following proposition.

Theorem 4.1. Under the foregoing assumptions with $\rho^\mu>1$, the queueing system is unstable under the action of any admissible policy π in $\mathcal{P}$, in the sense that

$$\lim_{t\uparrow\infty} Z^\pi(t) = \infty \qquad P^\mu\text{-a.s.} \tag{4.10}$$

This proposition is indeed a statement of system instability, since it amounts to saying that at least one of the K queues explodes in time and grows beyond any limit. The key fact leading to (4.10) lies in the observation that when $\rho^\mu > 1$, the sequence of RV's $\{z^{Z^\pi(t)}\}_1^\infty$ is a (P^μ, F_t^π)-supermartingale.

IV.2 The Case $\rho^\mu < 1$: The Stability Results

Let π be an admissible policy in P. To study the stability of the competing queues system under its action, consider an arbitrary F_t^π-stopping time σ and define the N-valued RV $\nu(\sigma)$ by

$$\nu(\sigma) := \begin{cases} \inf\{n > 1:\ X_k^\pi(\sigma + n) = 0,\ 1<k<K\} & \text{if } \sigma<\infty \text{ and this set is non-empty,} \\ \infty & \text{otherwise .} \end{cases} \qquad (4.11)$$

On the set $[\sigma < \infty]$, the RV $\nu(\sigma)$ counts the number of slots it takes for the system to empty for the first time after σ. Also, set $\tau(\sigma) := \sigma + \nu(\sigma)$ and observe that $\tau(\sigma)$ is an F_t^π-stopping time.

The key result for characterizing stability is contained in the following proposition.

Theorem 4.2. Under the foregoing assumptions with $\rho^\mu < 1$ and an arbitrary non-idling policy π in P, the relation

$$E^\mu\Big[1[\sigma<\infty, \nu(\sigma)<\infty]\Big[a(z)^{\nu(\sigma)} \prod_{r=\sigma}^{\tau(\sigma)-1} b(z; V^\pi(r))\Big]^{-1} \Big| F_\sigma^\pi\Big] = 1[\sigma<\infty] z^{Z^\pi(\sigma)} \qquad P^\mu\text{-a.s. } (4.12)$$

holds for every F_t^π-stopping time σ, as z ranges over $(0,1)$.

For all $0<z<1$, the sequence $\{M^\pi(t;z)\}_1^\infty$ of R_+-valued RV's defined by

$$M^\pi(t+1;z) := \begin{cases} z^{Z^\pi(1)} & \text{for } t = 0 \\ \dfrac{z^{Z^\pi(t+1)}}{a(z)^t \prod_{s=1}^{t} b(z; V^\pi(s))} & \text{for } t \neq 0 \text{ in } N \end{cases} \qquad (4.13)$$

is a positive (P^μ, F_t^π)-martingale owing to (4.7). Technicalities aside, the relation (4.12) follows from Doob's Optional Sampling Theorem applied to the martingale (4.13) with the stopping times σ and $\tau(\sigma)$, if this martingale were uniformly integrable. Unfortunately, this does not seem to be the case and a straightforward application of the Sampling Theorem is thus not possible. The required argument is somewhat technical and is given in [5, Appendix B]. Interesting consequences readily follow from the relation (4.12) and are given in the next corollary.

Corollary 4.2.1. Under the foregoing assumptions, with $\rho^\mu < 1$ and an arbitrary non-idling policy π in P,

$$P^\mu[\nu(\sigma) < \infty | F_\sigma^\pi] = 1[\sigma<\infty] \qquad P^\mu\text{-a.s. } (4.14)$$

for every F_t^π-stopping time σ. In particular, if $\sigma<\infty$ a.s., then $\nu(\sigma)<\infty$ a.s.

Another case of interest arises when $\sigma<\infty$ a.s. with $X_k^\pi(\sigma) = 0$, $1<k<K$, on the event $[\sigma<\infty]$. This situation holds promise for obtaining statistical information on the busy periods via Theorem 4.2. However, it is possible for each $\ell \neq 0$ in N to have $X_k^\pi(\sigma+\ell) = 0$, $1<k<K$, with a positive probability and the RV $\nu(\sigma)$ is thus not quite adequate to represent busy periods, as usually understood, unless some care is taken. With this is mind, set

$$\beta(\sigma) := \begin{cases} \inf\{n > 0:\ \sum_{k=1}^{K} A_k(\sigma + n) > 0\} & \text{if } \sigma<\infty \text{ and this set is non-empty} \\ \infty & \text{otherwise} \end{cases} \qquad (4.15)$$

and observe that the RV $\gamma(\sigma) := \sigma + [1 + \beta(\sigma)]$ is an F_t^π-stopping time. On the set $[\sigma<\infty]$, the RV $\beta(\sigma)$ counts the number of slots before customers again arrive into the system after time σ, while the RV $\gamma(\sigma)$ represents the left-boundary of the slot during which there is at least one customer present in the system for the first time since σ.

At this stage, it is useful to make an additional assumption on the arrival stream, so as to avoid limiting cases of little interest. Specifically, it is assumed that with a positive probability, new customers always enter the system during a slot. This is formalized in the following assumption (A5), hereafter enforced:

(A5): The probability distribution q_A satisfies the constraint $q_A(0)<1$, where 0 denotes the element in N^K with all zero components.

Observe that the RV $\beta(\sigma)$ is geometrically distributed, with

$$P^\mu[\beta(\sigma) = n] = [1 - q_A(0)] q_A(0)^n \qquad (4.16)$$

for all n in N, owing to the statistical assumptions (A1) - (A3) enforced on the data and the assumed fact $\sigma<\infty$ a.s. Now, if $q_A(0)<1$, then $\beta(\sigma)<\infty$ a.s. and so $\gamma(\sigma)<\infty$ a.s. obviously, whence $\nu(\gamma(\sigma))<\infty$ a.s. by Corollary 4.2.1.

If $\tau(\gamma(\sigma)) := \gamma(\sigma) + \nu(\gamma(\sigma))$ as before, it is now straightforward to interprete the interval $[\gamma(\sigma), \tau(\gamma(\sigma)))$ as the first busy period after time σ. Statistical information concerning its length, namely $\nu(\gamma(\sigma))-1$, can be obtained from Theorem 4.2 as the forthcoming result shows.

Theorem 4.3. Under the foregoing assumptions, with $\rho^\mu < 1$ and an arbitrary non-idling policy π in P,

$$E^\mu[\nu(\gamma(\sigma))] = \begin{cases} \dfrac{1}{1-\rho^\mu} & \text{if } \rho^\mu < 1 \qquad (4.17a) \\ \infty & \text{if } \rho^\mu = 1 \qquad (4.17b) \end{cases}$$

whenever $\sigma<\infty$ a.s. and $X_k^\pi(\sigma) = 0$, $1<k<K$, on $[\sigma<\infty]$.

If the (possibly infinite) mean of the RV Ξ_k is denoted by ξ_k, $1<k<K$, and define the coefficient r^μ to be

$$r^\mu := \sum_{K=1}^{K} \frac{\xi_k}{\mu_k}. \qquad (4.18)$$

A result similar to Theorem 4.3 is then also available for $\nu(1)$.

Theorem 4.4. Under the foregoing assumptions, with $\rho^{\mu}<1$ and an arbitrary non-idling policy π in P, the RV $\nu(1)$ is a.s. finite and

$$E^{\mu}[\nu(1)] = \begin{cases} \dfrac{1+r^{\mu}}{1-\rho^{\mu}} & \text{if } \rho^{\mu}<1 \quad (4.19a) \\ \infty & \text{if } \rho^{\mu}=1. \quad (4.19b) \end{cases}$$

It is now possible to completely characterize the stability behavior of the system (2.1)-(2.4) in terms of the quantity ρ^{μ}. For any admissible policy π in P, recursively define the sequences $\{\tau_n\}_1^{\infty}$ and $\{\gamma_n\}_1^{\infty}$ of F_t^{π}-stopping times by setting $\tau_1 := \tau(1)$ and

$$\gamma_n := \gamma(\tau_n), \quad \tau_{n+1} := \tau(\gamma_n) \qquad (4.20)$$

for all $n\neq 0$ in N, with the notation introduced earlier. The epochs $\{\tau_n\}_1^{\infty}$ are the clearing times for the system, and it is natural to interpret the intervals $[\tau_n,\gamma_n)$ and $[\gamma_n,\tau_{n+1})$ as idle and busy periods, respectively, and to view the interval $[\tau_n,\tau_{n+1})$ as the corresponding busy cycle. Their properties are summarized below:

Theorem 4.5. Under the foregoing assumptions, with $\rho^{\mu}<1$ and an arbitrary non-idling policy π in P, the stopping times $\{\tau_n\}_1^{\infty}$ and $\{\gamma_n\}_1^{\infty}$ are all a.s. finite and their first moments obey the following relations: First,

$$E^{\mu}[\tau_1] = \begin{cases} 1+\dfrac{1+r^{\mu}}{1-\rho^{\mu}} & \text{if } \rho^{\mu}<1 \quad (4.21a) \\ \infty & \text{if } \rho^{\mu}=1 \quad (4.21b) \end{cases}$$

while for all $n\neq 0$ in N,

$$E^{\mu}[\gamma_n - \tau_n] = \frac{1}{1-q_A(0)} \qquad (4.22)$$

and

$$E^{\mu}[\tau_{n+1} - \gamma_n] = \begin{cases} \dfrac{1}{1-\rho^{\mu}} & \text{if } \rho^{\mu}<1 \quad (4.23a) \\ & \text{if } \rho^{\mu}=1. \quad (4.23b) \end{cases}$$

This last result is clearly a description of system stability in terms of the mean busy periods; the terminology of null and positive recurrence for the case $\rho^{\mu}=1$ and $\rho^{\mu}<1$ is readily justified on the grounds of the expressions (4.21) and (4.23), in complete analogy with the situation for Markov chains.

That the formulae (4.21) and (4.23) are independent of the policy π may seem puzzling at first. However, a moment of reflection should convince the reader that in fact a much stronger property holds for the system (2.1)-(2.4). Indeed, two different non-idling policies generate identical busy periods, i.e., if π and $\tilde{\pi}$ are two non-idling policies in P and their corresponding clearing times (4.20) are denoted by $\{\tau_n^{\pi}\}_1^{\infty}$ and $\{\tau_n^{\tilde{\pi}}\}_1^{\infty}$, respectively, then

$$\tau_n^{\pi} = \tau_n^{\tilde{\pi}} \qquad (4.24)$$

for all $n\neq 0$ in N. Of course, this is no longer true if the policies idles. This property explains also why the stability of the controlled system was characterizable by a single parameter for all non-idling policies.

At this point, the reader may want to argue that the stability results of Theorem 4.5 could have been obtained more directly by first observing (4.24) and then by studying the stability of the queue size process for a given non-idling Markov stationary feedback policy. Indeed in that case the queue size process would be a Markov chain [5, Thm. 6.1] and standard ergodicity results [15,17,21,22] seemingly would apply to yield the obtained results. Unfortunately, the results in the cited references are not powerful enough to cover the level of generality adopted here, as no growth conditions or bounds are imposed on the arrival stream and a direct argumentation would have been needed anyway. The discussion sketched in this section shows that such arguments do indeed exist and that they apply to Markovian and non-Markovian control policies alike!

In the next sections, all statements are made under the probability measure P^{o} that characterizes the true model assumptions; the corresponding mathematical expectation operator and stability coefficient are denoted by E^{o} and ρ^{o}, respectively.

V. STRONG CONSISTENCY

The performance of sequential estimates is often evaluated through their convergence properties to the true parameter value. In this section, the consistency of the maximum likelihood estimates is investigated in some detail; a preliminary result on this question is given in the following proposition.

Theorem 5.1. For any admissible policy π in P, the sequence of estimates $\{\mu^{\pi}(t)\}_1^{\infty}$ has the property that

$$\lim_{t\uparrow\infty} \mu_k^{\pi}(t) = \mu_k^{o} \qquad P^{o}\text{-a.s.} \quad (5.1)$$

on the event $\Omega_k^{\pi} := [\liminf_{t\uparrow\infty} \frac{1}{t}\sum_{s=1}^{t} V_k^{\pi}(s) > 0]$.

For future use, it is desirable to understand when the sequences of estimates $\{\mu_k^{\pi}(t)\}_1^{\infty}$ converge to the true parameter value μ_k^{o}, $1\le k\le K$, under the true and correct model assumptions. Clearly, as a consequence of Theorem 5.1, this will happen if and only if $P^{o}(\Omega_k^{\pi}) = 1$. However, such a statement constitutes a rather weak and implicit version of the expected sufficiency result, insofar as a direct characterization in terms of system parameters and policy structure is sought! This task is undertaken in the remainder of this section, under the additional assumption (A6), where

(A6): The arrival sequence $\{A(t)\}_1^{\infty}$ has the property that

$$P^{\mu}[A_k(t)=0] = \sum_{a:a_k=0} q_A(a) < 1, \quad 1\le k\le K, \qquad (5.2)$$

for all $t\neq 0$ in N and μ in $[0,1]^K$.

This assumption implies (A5) and is a condition of non-degeneracy on the system, in the sense that each queue is required to have arrivals in every slot with positive probability. This is clearly necessary if parameter identification results are sought; otherwise, a given queue may become empty at some time, stay empty thereafter with a positive probability and since the service process will remain inactivated, parameter identification will not take place due to lack of relevant information.

Theorem 5.2. Under the enforced assumptions, with $\rho^o<1$ and an arbitrary non-idling policy π in P, $P^o(\Omega_k^\pi) = 1$, $1\le k\le K$, and therefore,

$$\lim_{t\uparrow\infty} \mu_k^\pi(t) = \mu_k^o \qquad P^o\text{-a.s.} \quad (5.3)$$

VI. THE LONG-RUN AVERAGE COST CONTROL PROBLEM

VI.1 Problem Formulation:

Simple performance measures are associated with the operation of the queueing system (2.1)-(2.4) by imposing an instantaneous cost proportional to queue sizes. With c_k, $1\le k\le K$, being positive constants held fixed throughout the discussion, the total cost per slot is defined through the mapping c from N^K into R, where

$$c(x) := \sum_{k=1}^{K} c_k x_k \quad (6.1)$$

for all x in N^K. For every admissible policy π in P, set

$$J_{\beta,t}(\pi) := E^o[\sum_{s=1}^{t} \beta^s c(X^\pi(s))], \quad (6.2)$$

and

$$J_{av}(\pi) := \liminf_{t\uparrow\infty} \frac{1}{t} J_{1,t}(\pi) = \liminf_{t\uparrow\infty} E^o[\frac{1}{t} \sum_{s=1}^{t} c(X^\pi(s))] \quad (6.3)$$

where β is a discount factor in $[0,1]$, $t\neq 0$ is in N and the RV's $\{X^\pi(s)\}_1^\infty$ are generated via the dynamics (2.1)-(2.4). The quantity $J_{\beta,t}(\pi)$ is the expected β-discounted costs associated with the admissible policy π in P over the finite horizon $[1,t+1)$, whereas $J_{av}(\pi)$ is the corresponding expected long-run average cost.

As in previous work on the competing queues problem [3,8], the following optimal control problems $(P_{\beta,t})$ and (P_{av}) are considered hereafter. They are simultaneously defined below as problem (P), with the convention that $J(\pi)$ represents anyone of the cost functions (6.2)-(6.3) and that in each case, the parameters range is the one for the corresponding cost function:

(P): Minimize $J(\pi)$ over the class P of all admissible control policies π.

A feedback policy π is now defined as a sequence $\{\pi_t\}_1^\infty$ of mappings π_t from $N^K x(N^K x\{1,\dots,K\})^{t+1}$ into $\{1,\dots K\}$, with the convention that the domain of definition of π_1 is simply N^K. Denote the collection of all such feedback policies by P_F. Under the action of a feedback policy π in P_F, the dynamics of the competing queues system is still given by the equations (2.1)-(2.4), but with (2.2) now replaced by

$$U^\pi(t+1) := \pi_{t+1}[\Xi;\ U^\pi(s),\ X^\pi(s+1),\ 1\le s\le t] \quad (6.4)$$

for all $t\neq 0$ in N. It is easy to see that to every such feedback policy, say π, there corresponds an admissible policy $\bar\pi$ in P with queue dynamics identical to the one generated under the action of π. Notwithstanding a slight abuse of notation, the inclusion $P_F\subseteq P$ thus hold, and this inclusion is strict owing to the discrete nature of the time parameter.

A feedback policy π is said to be a Markov policy if each one of its mappings $\{\pi_t\}_1^\infty$ reduces to a mapping from N^K into $\{1,\dots,K\}$; in that case $U^\pi(t) = \pi_t[X^\pi(t)]$ for all t 0 in N. A Markov stationary policy π is then defined as a Markov policy for which the mappings $\{\pi_t\}_1^\infty$ are all identical. For sake of notational simplicity, the mapping from N^K into $\{1,\dots,K\}$ that defines a Markov stationary policy π will also be denoted by π in what follows.

The μc-rule corresponding to the service parameter vector μ in $[0,1]^K$ is the Markov stationary policy $\pi^*(\mu)$ in P defined by

$$\pi^*[x;\mu] := \operatorname{Arg} \max_{1\le k\le K} \{\mu_k c_k[1 - \delta[x_k,0]]\} \quad (6.5)$$

for all x in N^K, with a tie breaker, when needed. Clearly, the μc-rule is a static prioritization scheme and it can be interpreted as the one that incurs the largest cost decrease per slot [3,4].

VI.2 Optimality of the μc-Rule for Problem (P_{av})

As a result of earlier investigations by the authors [3] and by Buyukkoc, Varaiya and Walrand [8], it is now known that the μc-rule provides a solution to the discounted problems $(P_{\beta,t})$, β in (0,1) and $t\neq 0$ in N, under the enforced assumptions (A1)-(A3), as it minimizes (6.2) over the class P of all admissible policies.

Theorem 6.1. Under the assumptions (A1)-(A3), the μc-rule solves problems $(P_{\beta,t})$ for all β in (0,1) and $t\neq 0$ in N, in the sense that

$$J_{\beta,t}(\pi^*(\mu^o)) \le J_{\beta,t}(\pi) \quad (6.6)$$

for every admissible policy π in P.

The main result of this section is concerned with the long-run average cost problem and is now easily derived.

Theorem 6.2. Under the assumptions (A1)-(A3), the μc-rule is optimal for problem (P_{av}).

As in [3], it is convenient at this stage to introduce the following finite mean assumption (A7), where

$$(A7):\ \xi_k<\infty,\ \alpha_k<\infty,\ 1\le k\le K. \quad (6.7)$$

Routine calculations [3, Section III] show that for every $t\neq 0$ in N and every β in $(0,1]$, the discounted cost $J_{\beta,t}(\pi)$ is finite for all admissible policies π in P if and only if the condition (A7) holds; consequently, the long-run average cost function (6.3) is then identically infinite when (A7) fails to hold and the problem (P_{av}) is meaningful only when the finite mean condition holds.

VII. THE ADAPTIVE LONG-RUN AVERAGE COST CONTROL PROBLEM

VII.1 Problem Formulation

The results of the previous section show that the minimum long-run average cost is achieved by implementing the μc-rule; this, of course, requires knowledge of the true value of the service parameters in the model assumptions. However, this information may not always be available to the decision-maker and a natural question thus arises: In the absence of knowledge of the true model parameters, how should service attention be allocated so as to minimize the long-run average cost? This question defines the adaptive long-run average cost control problem (P_{ad}) to be studied in this section.

The optimization problem (P_{av}) for the fully specified model has a known solution, the μc-rule, which involves the unknown service parameters μ^o in a very simple manner. In view of this, it is thus natural to invoke the so-called Certainty Equivalence Principle for generating easily implementable policies: At time t, on the basis of accumulated data, compute an estimate of the unknown parameters and generate a control action by making use of the optimal strategy for the system with true service parameters equal to the estimate value.

Here, the parameter estimate to be used is the maximum likelihood estimate derived in Section III. The admissible policy produced by applying the Certainty Equivalence Principle is called the adaptive μc-rule and is denoted by α. It is defined recursively by the following chain of events: At time t=1, only the initial queue size is known to the decision maker and the corresponding maximum likelihood estimate of μ^o consists of a value chosen arbitrarily in $[0,1]^K$, say $\bar{\mu}$. The control action to be made at time t=1 is thus

$$U^\alpha(1) := \text{Arg} \max_{1<k<K} \{\bar{\mu}_k c_k 1[\Xi_k \neq 0]\} = \pi^*[\Xi;\bar{\mu}] \qquad (7.1)$$

with a tie breaker. Now at time t, the data Ξ and $\{(U^\alpha(s), D^\alpha(s), A(s)),\ 1<s<t\}$ become available as α is used on the interval $[1,t)$ and the maximum likelihood estimate $\mu^\alpha(t)$ is computed according to Theorem 3.2 while the state $X^\alpha(t)$ is generated by the dynamics (2.1)-(2.4). By the Certainty Equivalence Principle, the service assignment in the time slot $[t,t+1)$ is determined to be $U^\alpha(t)$, where

$$U^\alpha(t) := \text{Arg} \max_{1<k<K} \{ {}^\alpha(t) c_k 1[X^\alpha(t) \neq 0]\}$$

$$= \pi^*[X^\alpha(t), \mu^\alpha(t)] \qquad (7.2)$$

with the usual tie breaker.

VII.2 Optimality of the Adaptive μc-Rule

It is clear that the adaptive μc-rule α is a non-idling policy but it is no longer a feedback policy as understood in this paper, thus no longer a Markov stationary policy. Nevertheless, under reasonable conditions, it shares some of the properties associated with of the μc-rule $\pi^*(\mu^o)$.

Theorem 7.1 Under the foregoing assumptions, with $\rho^o<1$ and the finite mean assumption (A7), the adaptive μc-rule α enjoys the following properties:

$$\lim_{t\uparrow\infty} \mu_k^\alpha(t) = \mu_k^0,\ 1<k<K, \qquad P^o\text{-a.s.}\ (7.3)$$

$$J_{av}(\alpha) = \lim_{t\uparrow\infty} \frac{1}{t} \sum_{s=1}^{t} c(X^\alpha(s)) \qquad P^o\text{-a.s.}\ (7.4)$$

$$= J_{av}(\pi^*(\mu^0)), \qquad (7.5)$$

i.e., α solves (P_{ad}) and parameter identification takes place.

REFERENCES

[1] F. Baccelli and A.M. Makowski, Direct martingale arguments for stability: The M|G|1 case, submitted to Stochastic Processes and Their Applications, 1984.

[2] J.S. Baras and A.J. Dorsey, Adaptive control of two competing queues, in Proceedings of the Second IEEE Annual Joint Conference Infocom, April 1983, San Diego, California, pp. 427-435.

[3] J.S. Baras, A.J. Dorsey and A.M. Makowski, Two competing queues with geometric service requirements and linear costs: The μc-rule is often optimal, Technical Research Report SRR-83-1, Electrical Engineering Department, University of Maryland, College Park. Also Advances in Applied Probability, forthcoming 1985.

[4] J.S. Baras, D.-J. Ma and A.M. Makowski, Competing queues with geometric service requirements and linear costs: the μc-rule is always optimal, Technical Research Report SRR-83-9, Electrical Engineering Department, University of Maryland, College Park.

[5] J.S. Baras, A.J. Dorsey and A.M. Makowski, Discrete-time competing queues with geometric service requirements: Stability, parameter estimation and adaptive control, Technical Research Report SRR-84-17, Electrical Engineering Department, University of Maryland, College Park. Submitted to SIAM J. Control and Optimization, 1984.

[6] V. Borkar and P.V. Varaiya, Adaptive control of Markov chains I: Finite parameter set, IEEE Trans. Automat. Contr., Vol. AC-24(1979), pp. 953-958.

[7] V. Borkar and P.V. Varaiya, Identification and adaptive control of Markov chains, SIAM J. Control and Optimization, Vol. 20(1982), pp. 470-489.

[8] C. Buyukkoc, P.V. Varaiya and J. Walrand, The $c\mu$-rule revisited, Advances in Applied Probability, forthcoming 1985.

[9] B. Doshi and S.E. Shreve, Strong consistency of a modified maximum likelihood estimator for controlled Markov chains, Journal of Applied Probability, Vol. 17(1980), pp. 726-734.

[10] A.J. Dorsey, Adaptive Control of a Simple Queueing System, Ph.D. Thesis, Electrical Engineering Department, University of Maryland, College Park, 1983.

[11] A.A. Fel'dbaum, Optimal Control Systems, Academic Press, New York, 1965.

[12] B. Hajek, Hitting-time and occupation-time bounds implied by drift analysis with applications, Advances in Applied Probability, Vol. 14(1982), pp. 502-525.

[13] P.R. Kumar and A. Becker, A new family of optimal adaptive controllers for Markov chains, IEEE Trans. Automat. Control, Vol. AC-27(1982), pp. 137-146.

[14] P.R. Kumar, A survey of some results in stochastic adaptive control, preprint.

[15] J. Lamperti, Criteria for the recurrence or transience of stochastic processes I, Journal of Mathematical Analysis and its Applications, Vol. 1(1960), pp. 314-330.

[16] A.M. Makowski, Simple control problems for multiple-access broadcast systems, in Proceedings of the 18th Annual Conference on Information Sciences and

Systems, Princeton University, Princeton, New Jersey, March 1984.
chains, Advances in Applied Probability, Vol. 6(1974), pp. 40-60.

[19] P. Mandl, On the adaptive control of finite state Markov processes, Zeitschrift Wahscheinlichkeits Theorie verw. Geb., Vol. 27(1973), pp. 263-276.

[20] P. Mandl, On the control of a Markov chain in the presence of unknown parameters, Transactions of the Sixth Prague Conference on Information Theory, Random Processes and Statistical Decision Functions, Prague, 1971, pp. 601-612.

[21] A.G. Pakes, Some conditions for ergodicity and recurrence of Markov chains, Operations Research, Vol. 17(1969), pp. 1058-1061.

[22] R.L. Tweedie, Criteria for classifying general Markov chains, Advances in Applied Probability, Vol. 8(1976), pp. 737-771.