Access Control to Two Multiserver Loss Queues in Series

Cheng-Yuan Ku and Scott Jordan

***Abstract*—We consider admission policies to two multiserver loss queues in series with two types of traffic. Both are generated according to independent Poisson processes with constant arrival rates. The first type requires service at the first queue and with a positive probability enters the second queue; the second type requires service at only the second queue. The service time distribution is exponential at either station. We show that under appropriate conditions the optimal admission policy that maximizes the expected total discounted reward over an infinite horizon is given by a switching curve. We characterize the form and shape of this curve and its variation with system parameters.**

***Index Terms*—Dynamic programming, loss network, optimal control.**

I. Introduction

Queueing networks are often used to model manufacturing systems and telecommunication systems and have been successful in analysis of resource allocation in these systems. Three trends, however, are challenging the effectiveness of traditional network models. First, both manufacturing and telecommunication systems are evolving toward flexible systems in which many heterogeneous products or services are produced. Second, the trend is toward minimizing work in progress and ensuring low delay. Third, control of such networks is increasingly accomplished by limiting the work entering the network, rather than using flow control within the network. Current mathematical models and their corresponding algorithms are inadequate in addressing resource allocation, pricing, and control of these systems. In this paper, we study a basic building block for a multiple service, multiple resource network. Such networks encompass a wide class of nonstorable production systems and broadband integrated service telecommunication systems.

In manufacturing, the first trend is toward flexible production systems. Mass customization demands that systems be able to produce a variety of products instead of a single product. Efficiency encourages that the production capacity be shared in an integrated fashion, rather than partitioned between separate production lines. Problems of allocation of capacity to accomplish desired throughputs of various products are common. The second trend, toward systems minimizing work in progress, has also become an area of intense interest. The benefits are seen as both minimizing inventory and decreasing the time to produce a product, especially in systems that must produce customized goods. Increasingly, we view production as nonstorable, and networks are designed with just-in-time or other push-or-pull mechanisms to accomplish these goals. The third trend in manufacturing is a direct result of the first two. Control of these nonstorable flexible manufacturing systems is increasingly implemented by limitations on product entering the network. Access control is viewed as more effective than flow control when minimal work in progress and minimal delay in producing a customized product are the goals.

In telecommunication systems, these three trends are also evident. The first trend, toward integrated service networks, is made possible by the increase in intelligence of switches and the separation of control from transmission. The merging of telephone and computer networks will continue to push the demand for telecommunication systems that can multiplex heterogeneous services and guarantee a variety of definitions of performance. The second trend, toward more tightly constrained use of buffers in the network, is caused by the requirements of synchronous services, such as real-time video or audio, for bounded loss within a guaranteed delay. When the burstiness of even a few services in the network becomes significant, buffering must be tightly controlled or these delay and loss requirements will not be met. Increasingly, guaranteed performance for a wide variety of services is the critical issue. The third trend is again a function of the first two. The increase in speed results in a corresponding increase in the number of packets in transit, and traditional flow control techniques no longer work well. Control of broadband telecommunication systems is tending toward access control, with the goal of limiting packets entering the network according to characterizations of burstiness and desired quality of service.

These three trends in both manufacturing and telecommunication systems pose severe challenges to traditional modeling and analysis techniques. Existing queueing theory does not cope well with nonstorable integrated service networks. Queueing models typically use delay as the measure of performance. In contrast, these new networks are likely to be engineered to guarantee low delay, and the measure of performance has become throughput (or blocking probability) of each service type. The optimization criteria is maximization of revenue, perhaps with constraints on minimal quality of service. Traditional queueing theory has provided us with useful information about resource allocation in homogeneous service networks, but fails us when multiple services compete for multiple resources and when blocking is the principal concern. In previous work, we have studied multiple service, multiple resource loss networks with simultaneous resource usage in Jordan [6]. In this paper, we study the simplest multiple service, multiple resource loss network with sequential resource usage. We hope this simple model will give us insight into sequential resource allocation policies that may be of use in larger multiple service, multiple resource loss networks.

II. Related Research

Throughout the years, many papers have addressed optimal control to a single queue. Apparently, it began with Naor [12] who studied critical-number policies in steady state in a M/M/1 queue. This approach has been extended to a M/D/1 queue Adler [1], a M/M/c queue with state-dependent benefit Knudsen [8], two heterogeneous servers with a common queue Lin [11], a general birth–death congestion model Knudsen [9], GI/M/c queue Simonovits [13], and a general input–output system Johansen [5]. A few papers have addressed optimal control to multiple queues. Davis [3] studied admission control for two exponential servers in parallel with separate queues and renewal arrivals. The optimal policy was shown to be monotonic and always assigns an accepted arrival to the shortest queue. Ghoneim [4] considered admission control on each of two exponential servers in series. Under a metric consisting of random

Manuscript received May 19, 1995; revised February 15, 1996.

C.-Y. Ku is with the Department of Information Engineering, Kaohsiung Polytechnic Institute, Taiwan, ROC (e-mail: cooperku@nas05.kpi.edu.tw).

S. Jordan is with the Department of Electrical & Computer Engineering, Northwestern University, Evanston, IL 60208-3118 USA.

Publisher Item Identifier S 0018-9286(97)04283-9.

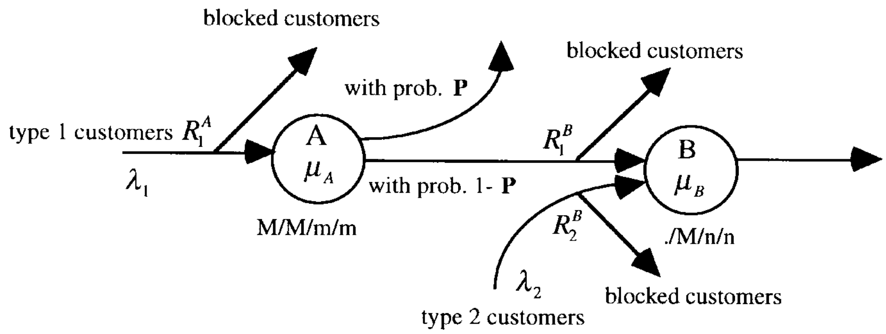

Fig. 1. Tandem multiserver queues.

rewards for entering customers and holding costs for customers at each node, the optimal policy was shown to be given by a pair of two-dimensional monotonic switching curves. A good survey of research on this topic until 1993 can be found in Stidham [14]. These multiple queue studies, however, have generally assumed infinite queues, with holding costs as the performance metric. There has been relatively little work done on admission control in loss networks, where the performance metric is a loss rate. Recently, Blanc [2] analyzed optimal control of admission to a multiserver queue with two arrival streams.

In this paper, we focus on multiserver loss queues in series. By concentrating on loss rather than delay or holding costs, on multiple customer types, and on access control, we hope to better understand control of evolving telecommunication and manufacturing systems. Our primary technique, stochastic dynamic programming, is similar to many of those used in previous characterizations of access control in delay networks. Additional complexity is presented by boundary states due to finite buffers, by the shape of the switching curve, and by the variability of total service rate due to the use of multiple servers. The key to the analysis lies in the combination of properties used in the value iteration. It is of note that the structural properties of the optimal access control shown here can be extended to parallel first stage queues [10].

III. Tandem Multiserver Loss Queues

Suppose there are two customer types in the system. Customer type 1 requires use of a resource A. When he is finished with A, with a positive probability he requires use of a resource B. Customer type 2 only requires use of a resource B. If there are finite many A's and B's, how should these resources be managed to maximize revenue? We model this system as a tandem multiserver loss network as pictured in Fig. 1.

We adopt the following notation: λ_1 as Poisson arrival rate of type-1 customers going to station A, μ_A as exponential service rate of servers at station A, R_1^A as revenue for service to a type-1 customer at station A, $\mathbf{P}$ as probability that customers leave the system after receiving service at station A, λ_2 as Poisson arrival rate of type-2 customers going into station B, μ_B as exponential service rate of servers at station B, R_1^B as revenue for service to a type-1 customer at station B, and R_2^B as revenue for service to a type-2 customer at station B.

A principal factor in the admission control system is the relative value of customers at station B. We wish to investigate the situation where type-1 customers are more valuable, in order to analyze the effect of knowledge of the existing number of type-1 customers at station A on the admission control decision at station B. Therefore, we assume $R_1^B > R_2^B$, $1 > \mathbf{P} \geq 0$, and positive revenue is collected

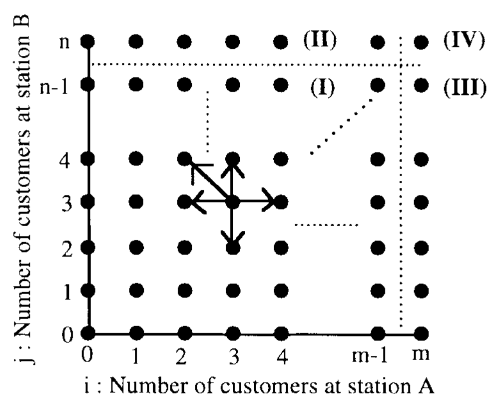

Fig. 2. The state space.

when a customer enters for service. This system can be modeled as a two-dimensional continuous-time Markov chain, with state (i, j) defined as the number of customers at stations A and B, respectively. The admission control policy becomes a decision to accept or deny the arrivals, as a function of the state, the customer type, and the station at which the customer is arriving. Uniformization (see, e.g., Kumar [7]) results in an equivalent discrete-time Markov chain by allowing fictitious transitions from a state to itself. Choose a $Q \in \mathbf{R}$ s.t. $Q > \lambda_1 + \lambda_2 + m\mu_A + n\mu_B$ and let $p_1 = \frac{\lambda_1}{Q}$, $p_2 = \frac{\lambda_2}{Q}$, $q_A = \frac{\mu_A}{Q}$, and $q_B = \frac{\mu_B}{Q}$. The equivalent discrete-time system has corresponding parameters p_1, p_2, q_A, q_B and the appropriate discount factor $\alpha < 1$.

For our system, if there is a customer at station A, we know there may be one respective request at station B in the near future. So, the number of customers at station A can be regarded as an amount of reservation for type B resources. How should this knowledge be incorporated into the admission decision for type-2 customers at station B? We address this problem using a stochastic dynamic programming framework. The two-dimensional state space can be divided into four regions according to boundaries as in Fig. 2.

We consider the objective of maximizing discounted revenue over an infinite horizon. We define $V(i, j)$ as expected discounted revenue starting in state (i, j), and $0 < \alpha < 1$ is the discount factor. An admission policy is given by a mapping $a : \{0, 1, \cdots, m\} \times \{0, 1, \cdots, n\} \rightarrow \{0, 1\}^3$, where $a = (a_1, a_2, a_3)$ represents whether each type of customer at each station should be admitted when the system is in state (i, j). $a_1 = 1$ iff we admit customer type 1 at station A, $a_2 = 1$ iff we admit customer type 2 at station B, and $a_3 = 1$ iff we admit customer type 1 at station B. The optimal admission policy is chosen in each state to maximize the future expected discounted revenue, as given by the following set of equations, stated by region.

1) For $0 \le i \le m-1,\ 0 \le j \le n-1$

$$\begin{aligned} V(i,j) = \max_{a_1,a_2,a_3} \alpha\big\{&a_1p_1\big[V(i+1,j)+R_1^A\big] \\ &+ a_2p_2\big[V(i,j+1)+R_2^B\big] \\ &+ a_3(1-\mathbf{P})iq_A\big[V(i-1,j+1)+R_1^B\big] \\ &+ (1-a_3)(1-\mathbf{P})iq_AV(i-1,j) \\ &+ \mathbf{P}iq_AV(i-1,j)+jq_BV(i,j-1) \\ &+ (1-a_1p_1-a_2p_2-iq_A-jq_B)V(i,j)\big\}. \end{aligned}$$

2) For $0 \le i \le m-1,\ j = n$

$$\begin{aligned} V(i,j) = \max_{a_1} \alpha\big\{&a_1p_1\big[V(i+1,j)+R_1^A\big] \\ &+ iq_AV(i-1,j)+jq_BV(i,j-1) \\ &+ (1-a_1p_1-iq_A-jq_B)V(i,j)\big\}. \end{aligned}$$

3) For $i = m,\ 0 \le j \le n-1$

$$\begin{aligned} V(i,j) = \max_{a_2,a_3} \alpha\big\{&a_2p_2\big[V(i,j+1)+R_2^B\big] \\ &+ a_3(1-\mathbf{P})iq_A\big[V(i-1,j+1)+R_1^B\big] \\ &+ (1-a_3)(1-\mathbf{P})iq_AV(i-1,j) \\ &+ \mathbf{P}iq_AV(i-1,j)+jq_BV(i,j-1) \\ &+ (1-a_2p_2-iq_A-jq_B)V(i,j)\big\}. \end{aligned}$$

4) For $i = m,\ j = n$

$$\begin{aligned} V(i,j) = \alpha\big\{&iq_AV(i-1,j)+jq_BV(i,j-1) \\ &+ (1-iq_A-jq_B)V(i,j)\big\}. \end{aligned}$$

We present notation that will be useful in analyzing the optimal policy. We use the value iteration approach. Choose an arbitrary initial value function, V_0, then define the step h value function, V_h, by choosing actions that maximize future expected discounted revenue, assuming that transitions more than one time step in the future generate a total discounted revenue according to the step $h-1$ value function.

Definition 3.1: Let $V_0(i,j)$ be an arbitrary bounded function, and for $h \in \mathbf{N}$ and $h \ne 0$, define

$$V_h(i,j) = \max_a \left\{ \alpha \sum_k \sum_l P_{(i,j)\to(k,l)}(a) \times \left[R((i,j),(k,l)) + V_{h-1}(k,l)\right] \right\} \qquad (1)$$

where $P_{(i,j)\to(k,l)}(a)$ is the one-step transition probability and $R((i,j),(k,l))$ is the revenue associated with a change of state from (i,j) to (k,l), if any. It is known that the optimal value function is defined by $V(i,j) = \lim_{h\to\infty} V_h(i,j)$.

Definition 3.2: Let $\Delta(i,j) \equiv V(i,j) - V(i,j+1)$

$$\Delta_h(i,j) \equiv V_h(i,j) - V_h(i,j+1) \quad \text{for } h,i,j \in \mathbf{N},\ 0 \le i \le m \text{ and } 0 \le j \le n-1.$$

Definition 3.3: Let $\nabla(i,j) \equiv V(i,j) - V(i+1,j)$

$$\nabla_h(i,j) \equiv V_h(i,j) - V_h(i+1,j) \quad \text{for } h,i,j \in \mathbf{N},\ 0 \le i \le m-1 \text{ and } 0 \le j \le n.$$

Then, we can rewrite the optimality equations 1)–4) using the notation above. For example, if we rewrite 1), we obtain

$$\begin{aligned} V(i,j) = \max_{a_1,a_2,a_3} \alpha\big\{&a_1p_1\big[R_1^A - \nabla(i,j)\big] + a_2p_2\big[R_2^B - \Delta(i,j)\big] \\ &+ a_3(1-\mathbf{P})iq_A\big[R_1^B - \Delta(i-1,j)\big] + iq_AV(i-1,j) \\ &+ jq_BV(i,j-1) + (1-iq_A-jq_B)V(i,j)\big\}. \end{aligned}$$

The optimal admission policy admits a customer if the immediate revenue generated by that customer exceeds the expected loss in future discounted revenue caused by future blocking due to this customer. However, from the rewriting of optimality equations, the optimal policy, in state (i,j), is obvious: accepts a customer of type 1 at station A iff $\nabla(i,j) \le R_1^A$, accepts a customer of type 2 at station B iff $\Delta(i,j) \le R_2^B$, and accepts a customer of type 1 at station B iff $\Delta(i-1,j) \le R_1^B$.

Definition 3.4: An admission control policy for type-2 customers is called a threshold policy if there exists a switching curve $j^s(i)$ so that customers are admitted iff $j < j^s(i)$. For such policies, define:

A) $j^s(i) = \min(j \mid \Delta(i,j) > R_2^B)$ and $j_h^s(i) = \min(j \mid \Delta_h(i,j) > R_2^B)$. Let $j^s(i) = n$ and $j_h^s(i) = n$, if $\Delta(i,j) \le R_2^B$ for $0 \le j \le n$;

B) $i^s(j) = \max(i \mid \Delta(i,j) \le R_2^B)$ and $i_h^s(j) = \max(i \mid \Delta_h(i,j) \le R_2^B)$. Let $i^s(j) = -1$ and $i_h^s(j) = -1$, if $\Delta(i,j) > R_2^B$ for $0 \le i \le m$.

Likewise, the optimal threshold is defined by $j^s(i) = \lim_{h\to\infty} j_h^s(i)$.

IV. Optimal Access Control

The form of the optimal admission control policy is stated in this section in a series of theorems. Theorems 4.2 and 4.3 state that, under appropriate conditions, only type-2 traffic need be controlled. Theorems 4.4 and 4.5 state that optimal admission policy of type-2 customers follows a monotonically decreasing switching curve. The following lemma details the structure of the value function under optimal policy, using value iteration. It will be used repeatedly to prove that the optimal policy, given by the corresponding limit, has similar desirable properties to those described here.

Lemma 4.1: If:

a) $\Delta_h(i,j)$ is monotonically increasing on j for fixed i (i.e., $V_h(i,j)$ is concave on j);

b) $\Delta_h(i,j)$ is monotonically increasing on i for fixed j (i.e., $V_h(i,j)$ is submodular);

c) $\Delta_h(i,j) < R_1^B$ for $0 \le i \le m$ and $0 \le j \le n-1$;

d) $\nabla_h(i,j) < R_1^A$ for $0 \le i \le m-1$ and $0 \le j \le n$;

then:

A) $\Delta_{h+1}(i,j)$ is monotonically increasing on j for fixed i (i.e., $V_{h+1}(i,j)$ is concave on j);

B) $\Delta_{h+1}(i,j)$ is monotonically increasing on i for fixed j (i.e., $V_{h+1}(i,j)$ is submodular);

C) $\Delta_{h+1}(i,j) < R_1^B$ for $0 \le i \le m$ and $0 \le j \le n-1$;

D) $\nabla_{h+1}(i,j) < R_1^A$ for $0 \le i \le m-1$ and $0 \le j \le n$.

Proof: The proof proceeds by expanding Definition 3.1 in each region of the state space. First, similar terms are grouped. Second, remaining terms, often due to boundaries, are bounded by others. Finally, the hypotheses are invoked on each group to prove the conclusions.

A): By hypotheses c) and d), $\Delta_h(i,j) < R_1^B$ and $\nabla_h(i,j) < R_1^A$ for all (i,j), we know $a_1 = 1$ and $a_3 = 1$ achieve the maximization in (1). The control parameter a_2 is set according to the subregion of the state space. In order to prove that $\Delta_{h+1}(i,j)$ is monotonically increasing on j for fixed i, we consider two states (i,j_1) and (i,j_2)

with $j_1 < j_2$. We show that $\Delta_{h+1}(i, j_2) - \Delta_{h+1}(i, j_1) > 0$. We initially consider the case $0 \le i < m$, $0 \le j_1 < j_2 < j_h^s(i) - 1 \le n - 1$.

Using Definition 3.2, $\Delta_{h+1}(i, j_2) - \Delta_{h+1}(i, j_1) \equiv [V_{h+1}(i, j_2) - V_{h+1}(i, j_2+1)] - [V_{h+1}(i, j_1) - V_{h+1}(i, j_1+1)]$.

Using (1) and grouping terms, we get

$$\begin{aligned}
&\Delta_{h+1}(i, j_2) - \Delta_{h+1}(i, j_1)\\
&\quad = \alpha\{p_1[[V_h(i+1, j_2) - V_h(i+1, j_2+1)]\\
&\qquad - [V_h(i+1, j_1) - V_h(i+1, j_1+1)]]\\
&\qquad + p_2[[V_h(i, j_2+1) - V_h(i, j_2+2)]\\
&\qquad - [V_h(i, j_1+1) - V_h(i, j_1+2)]]\\
&\qquad + (1-\mathbf{P})iq_A[[V_h(i-1, j_2+1)\\
&\qquad - V_h(i-1, j_2+2)] - [V_h(i-1, j_1+1)\\
&\qquad - V_h(i-1, j_1+2)]] + \mathbf{P}iq_A[[V_h(i-1, j_2)\\
&\qquad - V_h(i-1, j_2+1)] - [V_h(i-1, j_1)\\
&\qquad - V_h(i-1, j_1+1)]] + j_2 q_B[V_h(i, j_2-1)\\
&\qquad - V_h(i, j_2)] - j_1 q_B[V_h(i, j_1-1)\\
&\qquad - V_h(i, j_1)] + (1 - p_1 - p_2 - iq_A - (j_2+1)q_B)\\
&\qquad \times [V_h(i, j_2) - V_h(i, j_2+1)]\\
&\qquad - (1 - p_1 - p_2 - iq_A - (j_1+1)q_B)\\
&\qquad \times [V_h(i, j_1) - V_h(i, j_1+1)]\}.
\end{aligned}$$

Now expand one term

$$\begin{aligned}
&-(1 - p_1 - p_2 - iq_A - (j_1+1)q_B)[V_h(i, j_1) - V_h(i, j_1+1)]\\
&\quad = -(1 - p_1 - p_2 - iq_A - (j_2+1)q_B)[V_h(i, j_1)\\
&\qquad - V_h(i, j_1+1)] - (j_2 q_B - j_1 q_B)[V_h(i, j_1) - V_h(i, j_1+1)].
\end{aligned}$$

Using hypothesis a) and $j_2 > j_1$, we can bound part of this term: $V_h(i, j_1) - V_h(i, j_1+1) \le V_h(i, j_2-1) - V_h(i, j_2)$.

By grouping terms, and using Definition 3.2, we get

$$\begin{aligned}
&\Delta_{h+1}(i, j_2) - \Delta_{h+1}(i, j_1)\\
&\quad \ge \alpha\{p_1[\Delta_h(i+1, j_2) - \Delta_h(i+1, j_1)] + p_2[\Delta_h(i, j_2+1)\\
&\qquad - \Delta_h(i, j_1+1)] + (1-\mathbf{P})iq_A[\Delta_h(i-1, j_2+1)\\
&\qquad - \Delta_h(i-1, j_1+1)] + \mathbf{P}iq_A[\Delta_h(i-1, j_2)\\
&\qquad - \Delta_h(i-1, j_1)] + j_1 q_B[\Delta_h(i, j_2-1)\\
&\qquad - \Delta_h(i, j_1-1)] + (1 - p_1 - p_2 - iq_A - (j_2+1)q_B)\\
&\qquad \times [\Delta_h(i, j_2) - \Delta_h(i, j_1)]\}.
\end{aligned}$$

Finally, hypothesis a) guarantees that each term in brackets is greater than zero, and hence $\Delta_{h+1}(i, j_2) - \Delta_{h+1}(i, j_1) > 0$.

Consideration of other sections of the state space results in similar proofs. In some sections, knowledge of the location of the threshold is required to bound some terms, e.g., $\Delta_h(i, j) > R_2^B$ if $j > j_h^s(i) - 1$, $\Delta_h(i, j) \le R_2^B$ if $j = j_h^s(i) - 1$, and $\Delta_h(i, j) < R_2^B$ if $j < j_h^s(i) - 1$. Near boundaries, hypothesis c) is used to bound other terms. Details are omitted. For B), this proof also requires hypothesis b), but is otherwise substantially similar to A) and is hence omitted. For C), by hypotheses c) and d), $\Delta_h(i, j) < R_1^B$ and $\nabla_h(i, j) < R_1^A$ for all (i, j), we know $a_1 = 1$ and $a_3 = 1$ achieve the maximization in (1). From A) and B), we know $\Delta_{h+1}(m, n-1)$ is the largest difference. So, it suffices to prove that $\Delta_{h+1}(m, n-1) < R_1^B$.

Using Definition 3.2, $\Delta_{h+1}(m, n-1) \equiv V_{h+1}(m, n-1) - V_{h+1}(m, n)$. Using (1) and grouping terms, we get

$$\begin{aligned}
&\Delta_{h+1}(m, n-1)\\
&\quad = \max_{a_2} \alpha\big\{a_2 p_2\big[V_h(m, n) + R_2^B\big]\\
&\qquad + (1-\mathbf{P})mq_A\big[V_h(m-1, n) + R_1^B\big]\\
&\qquad + \mathbf{P}mq_A V_h(m-1, n-1) + (n-1)q_B V_h(m, n-2)\\
&\qquad + (1 - a_2 p_2 - mq_A - (n-1)q_B)V_h(m, n-1)\big\}\\
&\qquad - \alpha\{mq_A V_h(m-1, n) + nq_B V_h(m, n-1)\\
&\qquad + (1 - mq_A - nq_B)V_h(m, n)\}.
\end{aligned}$$

By grouping terms, we have

$$\begin{aligned}
&\Delta_{h+1}(m, n-1)\\
&\quad = \max_{a_2} \alpha\big\{a_2 p_2\big[V_h(m, n) + R_2^B\big]\\
&\qquad - p_2 V_h(m, n) + (1-\mathbf{P})mq_A\big[V_h(m-1, n) + R_1^B\big]\\
&\qquad - (1-\mathbf{P})mq_A V_h(m-1, n) + \mathbf{P}mq_A V_h(m-1, n-1)\\
&\qquad - \mathbf{P}mq_A V_h(m-1, n) + (n-1)q_B V_h(m, n-2)\\
&\qquad - (n-1)q_B V_h(m, n-1) + (1 - a_2 p_2 - mq_A - nq_B)\\
&\qquad \times V_h(m, n-1) - (1 - p_2 - mq_A - nq_B)V_h(m, n)\big\}.
\end{aligned}$$

Using hypotheses a) and b) and $\alpha < 1$

$$\begin{aligned}
\Delta_{h+1}(m, n-1) < \max_{a_2} &\big\{a_2 p_2\big[V_h(m, n) + R_2^B\big]\\
&- p_2 V_h(m, n) + (1-\mathbf{P})mq_A R_1^B\\
&+ \mathbf{P}mq_A[V_h(m, n-1) - V_h(m, n)]\\
&+ nq_B[V_h(m, n-1) - V_h(m, n)]\\
&+ (1 - a_2 p_2 - mq_A - nq_B)V_h(m, n-1)\\
&- (1 - p_2 - mq_A - nq_B)V_h(m, n)\big\}.
\end{aligned}$$

Case 1: for $a_2 = 1$ this gives

$$\begin{aligned}
\Delta_{h+1}(m, n-1) &< \big\{p_2 R_2^B + (1-\mathbf{P})mq_A R_1^B\\
&\quad + (1 - p_2 - (1-\mathbf{P})mq_A)\Delta_h(m, n-1)\big\}\\
&< \big\{p_2 R_1^B + (1-\mathbf{P})mq_A R_1^B\\
&\quad + (1 - p_2 - (1-\mathbf{P})mq_A)R_1^B\big\}\\
&\Rightarrow \Delta_{h+1}(m, n-1) < R_1^B.
\end{aligned}$$

Case 2: for $a_2 = 0$ this gives

$$\begin{aligned}
\Delta_{h+1}(m, n-1) &< \big\{(1-\mathbf{P})mq_A R_1^B\\
&\quad + (1 - (1-\mathbf{P})mq_A)\Delta_h(m, n-1)\big\}\\
&< \big\{(1-\mathbf{P})mq_A R_1^B\\
&\quad + (1 - (1-\mathbf{P})mq_A)R_1^B\big\}\\
&\Rightarrow \Delta_{h+1}(m, n-1) < R_1^B.
\end{aligned}$$

From the cases above, we can conclude that $\Delta_{h+1}(m, n-1) < R_1^B$, and the result follows.

For D), this proof is substantially similar to C) and is hence omitted.

Theorem 4.2: $\nabla(i, j) \le R_1^A$ for $0 \le i \le m-1$ and $0 \le j \le n$. Consequently, it is optimal to always admit type-1 customers at station A, i.e., $a_1 = 1$ in all states.

Proof: Proof is by induction. It is always possible to choose an initial value function, V_0, such that it satisfies the hypotheses of Lemma 4.1. Suppose these hypotheses hold for the step h value function, V_h. Then, the lemma states that they hold for the step $h+1$ value function V_{h+1}. The optimal value function is defined by the limit in Definition 3.1, and therefore statement D) in Lemma 4.1 implies that $\nabla(i,j) \le R_1^A$.

Theorem 4.3: $\Delta(i,j) \le R_1^B$ for $0 \le i \le m$ and $0 \le j \le n-1$. Consequently, it is optimal to always admit type-1 customers at station B, i.e., $a_3 = 1$ in all states.

Proof: The proof is by induction as in Theorem 4.2, except that it follows from statement C) in Lemma 4.1.

From Theorems 4.2 and 4.3, we understand that the optimal policy always admits type-1 customers when space permits. As we will see, the optimal policy on type-2 customers is a threshold type: type-2 customers are only accepted while the system state is below some threshold. The next two theorems characterize the form of the optimal control upon type-2 customers.

Theorem 4.4: $\Delta(i,j)$ is monotonically increasing on j for fixed i. Consequently, the optimal policy is threshold-type.

Proof: The proof is by induction as in Theorem 4.2, except it follows from statement A) in Lemma 4.1.

Recall that the optimal policy accepts a type-2 customer if and only if $\Delta(i,j) \le R_2^B$. For a fixed number of customers, i, at stage A, the threshold lies at the smallest j such that $\Delta(i,j) > R_2^B$. Since $\Delta(i,j)$ is monotonically increasing on j for fixed i, if the policy blocks a type-2 customer when $j = j^s(i)$, it also blocks when $j > j^s(i)$.

Theorem 4.5: $\Delta(i,j)$ is monotonically increasing on i for fixed j. Consequently, $j^s(i)$ is a nonincreasing function on i.

Proof: The proof is by induction as in Theorem 4.2, except it follows from statement B) in Lemma 4.1.

These previous propositions characterize optimal admission control to the pair of queues, when type-1 customers are more valuable than type 2. Type-1 customers are always admitted. Type-2 customers are admitted only if the number of free servers at station B exceeds some threshold. We view $n - j^s(i)$, the minimum number of free servers at station B, given i customers at station A, as the reservation at B for upstream traffic. This reservation is monotonically increasing in i.

V. Variation of System Parameters

In the previous section, we characterized the optimal admission policy on type-2 customers in tandem multiserver loss queues. We found that for fixed parameters $(\lambda_1,\ \lambda_2,\ \alpha, \mathbf{P},\ \mu_A,\ \mu_B,\ R_1^A,\ R_2^B,\ R_1^B,\ m,\ n)$ the optimal policy is given by a threshold $j^s(i)$. In this section, we investigate the variation of this threshold with variations in system parameters.

Theorem 5.1: The admission control threshold on type-2 customers, $j^s(i)$, is, for $0 \le i \le m$ and $R_1^B > R_2^B$:

A) monotonically decreasing in $\lambda_1,\ \alpha,\ \lambda_2,\ R_1^B$, and m;
B) monotonically increasing in $\mu_B,\ \mathbf{P}$, and R_2^B;
C) insensitive to R_1^A;
D) $n - j^s(i)$ is monotonically decreasing in n, on $i \le i^s(0)$.

Proof for Decreasing in λ_1: Choose $\infty > \bar\lambda_1 > \lambda_1 > 0$. With λ_1 and all other system parameters given, use value iteration upon appropriate initial rewards to obtain $V(i,j,\lambda_1)$ for all possible i's and j's. Now, replace λ_1 by $\bar\lambda_1$ and keep all other system parameters unchanged. Choosing $V(i,j,\lambda_1)$ for $0 \le i \le m$ and $0 \le j \le n$ as initial rewards, use value iteration again to calculate $V(i,j,\bar\lambda_1)$. Since $\lim_{h\to\infty} \Delta_h(i,j,\bar\lambda_1) = \Delta(i,j,\bar\lambda_1)$ and $\Delta_0(i,j,\bar\lambda_1) = \Delta(i,j,\lambda_1)$, to obtain $\Delta(i,j,\bar\lambda_1) > \Delta(i,j,\lambda_1)$, it is sufficient to show that $\Delta_{h+1}(i,j,\bar\lambda_1) > \Delta_h(i,j,\bar\lambda_1)$ for all $h \in \mathbf{N}$, $h \ne 0$, and $\Delta_1(i,j,\bar\lambda_1) \ge \Delta_0(i,j,\bar\lambda_1)$. Consequently, it will follow that $j^s(i,\bar\lambda_1) \le j^s(i,\lambda_1)$ for all $0 \le i \le m$.

By Lemma 4.1 and Theorems 4.2, 4.3, 4.4, and 4.5, we know $a_1 = 1$ and $a_3 = 1$ achieve the maximization in (1). The control parameter a_2 is set according to the subregion of the state space. Let $Q > \bar\lambda_1 + \lambda_2 + m\mu_A + n\mu_B$ and $\bar p_1 = \frac{\bar\lambda_1}{Q}$, then $\bar\lambda_1 > \lambda_1 \Rightarrow \bar p_1 > p_1$. The base step is to show that for $i \ne m$, $\Delta_1(i,j,\bar\lambda_1) > \Delta_0(i,j,\bar\lambda_1)$, and for $i = m$, $\Delta_1(i,j,\bar\lambda_1) = \Delta_0(i,j,\bar\lambda_1)$. We consider here the case $0 \le i \le m-1$ and $0 \le j < j_0^s(i,\bar\lambda_1) - 1 = j^s(i,\lambda_1) - 1$. Other cases with $i \ne m$ are similar.

Using Definition 3.2, (1), and $\Delta_0(i,j,\bar\lambda_1) = \Delta(i,j,\lambda_1)$

$$\begin{aligned}
\Delta_1(i,j,\bar\lambda_1)& \\
&= \alpha\{\bar p_1 \Delta(i+1,j,\lambda_1) + p_2\Delta(i,j+1,\lambda_1)\\
&\quad + (1-\mathbf{P})iq_A\Delta(i-1,j+1,\lambda_1) + \mathbf{P}iq_A\Delta(i-1,j,\lambda_1)\\
&\quad + jq_B\Delta(i,j-1,\lambda_1) + (1-\bar p_1 - p_2 - iq_A\\
&\quad - (j+1)q_B)\Delta(i,j,\lambda_1)\}.
\end{aligned}$$

Grouping terms, using Definition 3.2 and (1)

$$\begin{aligned}
\Delta_1(i,j,\bar\lambda_1) &= \Delta(i,j,\lambda_1) + \alpha(\bar p_1 - p_1)\\
&\quad \times [\Delta(i+1,j,\lambda_1) - \Delta(i,j,\lambda_1)].
\end{aligned}$$

From Theorem 4.5, we know $\Delta(i+1,j,\lambda_1) - \Delta(i,j,\lambda_1) > 0$. Since $\alpha(\bar p_1 - p_1)$ is positive, then we get $\Delta_1(i,j,\bar\lambda_1) > \Delta(i,j,\lambda_1) = \Delta_0(i,j,\bar\lambda_1)$. For $i = m$, it is straightforward to show that $\Delta_1(i,j,\bar\lambda_1) = \Delta_0(i,j,\bar\lambda_1) = \Delta(i,j,\lambda_1)$.

The induction step supposes that for $i \ne m$, $\Delta_h(i,j,\bar\lambda_1) > \Delta_{h-1}(i,j,\bar\lambda_1)$, and for $i = m$, $\Delta_h(i,j,\bar\lambda_1) \ge \Delta_{h-1}(i,j,\bar\lambda_1)$ and shows that $\Delta_{h+1}(i,j,\bar\lambda_1) > \Delta_h(i,j,\bar\lambda_1)$ for all (i,j). We consider here the case $0 \le i \le m-1$ and $0 \le j = j_h^s(i,\bar\lambda_1) - 1 < j_{h-1}^s(i,\bar\lambda_1) - 1$. Other cases are similar.

Using Definition 3.2 and (1):

$$\begin{aligned}
\Delta_{h+1}(i,j,\bar\lambda_1)& \\
&= \alpha\big\{\bar p_1\Delta_h(i+1,j,\bar\lambda_1) + p_2R_2^B\\
&\quad + (1-\mathbf{P})iq_A\Delta_h(i-1,j+1,\bar\lambda_1)\\
&\quad + \mathbf{P}iq_A\Delta_h(i-1,j,\bar\lambda_1) + jq_B\Delta_h(i,j-1,\bar\lambda_1)\\
&\quad + (1-\bar p_1 - p_2 - iq_A - (j+1)q_B)\Delta_h(i,j,\bar\lambda_1)\big\}.
\end{aligned}$$

Imposing the hypothesis that for $i \ne m$, $\Delta_h(i,j,\bar\lambda_1) > \Delta_{h-1}(i,j,\bar\lambda_1)$ and for $i = m$, $\Delta_h(i,j,\bar\lambda_1) \ge \Delta_{h-1}(i,j,\bar\lambda_1)$ and using $R_2^B \ge \Delta_{h-1}(i,j+1,\bar\lambda_1)$, we obtain

$$\begin{aligned}
\Delta_{h+1}(i,j,\bar\lambda_1)& \\
&> \alpha\big\{\bar p_1\Delta_{h-1}(i+1,j,\bar\lambda_1) + p_2\Delta_{h-1}(i,j+1,\bar\lambda_1)\\
&\quad + (1-\mathbf{P})iq_A\Delta_{h-1}(i-1,j+1,\bar\lambda_1)\\
&\quad + \mathbf{P}iq_A\Delta_{h-1}(i-1,j,\bar\lambda_1) + jq_B\Delta_{h-1}(i,j-1,\bar\lambda_1)\\
&\quad + (1-\bar p_1 - p_2 - iq_A - (j+1)q_B)\Delta_{h-1}(i,j,\bar\lambda_1)\big\}\\
&= \Delta_h(i,j,\bar\lambda_1).
\end{aligned}$$

Consequently $\Delta(i,j,\bar\lambda_1) > \Delta(i,j,\lambda_1)$ and $j^s(i)$ is monotonically decreasing in λ_1.

Proof for Decreasing in α, λ_2, and R_1^B; Increasing in μ_B and $\mathbf{P}$; and Insensitive to R_1^A: These proofs are substantially similar to those above and are hence omitted.

Proof for Increasing in R_2^B: Consider a system with identical proportional rewards $\bar R_1^B = kR_1^B$, $\bar R_2^B = kR_2^B$. Choose initial values such that

$$\frac{\Delta_0(i,j,\bar R_2^B,\bar R_1^B)}{\Delta_0(i,j,R_2^B,R_1^B)} = k$$

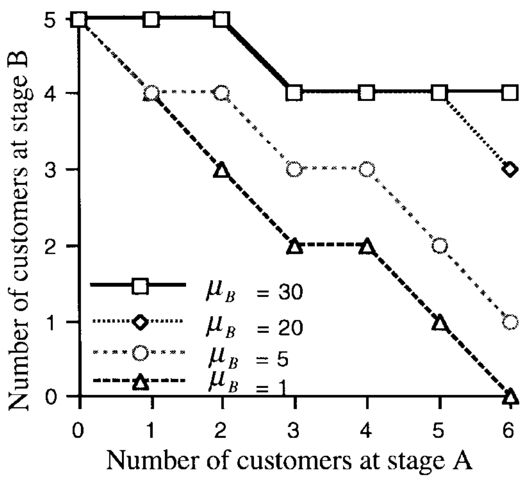

Fig. 3. Varying μ_B.

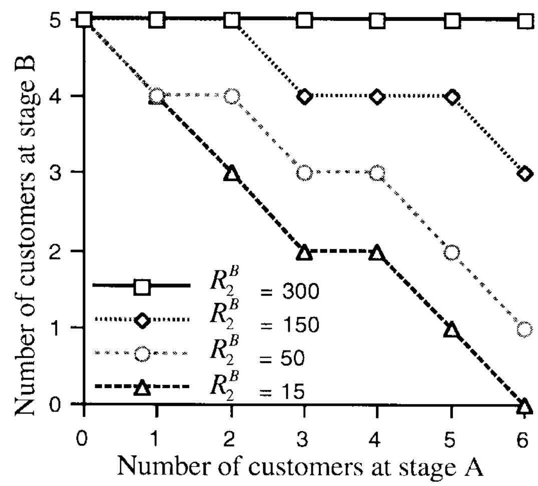

Fig. 4. Varying R_2^B.

and satisfy the hypotheses of Lemma 4.1. It is easy to show that

$$\frac{\Delta_h(i,j,\bar{R}_2^B,\bar{R}_1^B)}{\Delta_h(i,j,R_2^B,R_1^B)} = k, \qquad \forall h \in \mathbf{N}.$$

Therefore

$$\frac{\Delta(i,j,\bar{R}_2^B,\bar{R}_1^B)}{\bar{R}_2^B} = \frac{\Delta(i,j,R_2^B,R_1^B)}{R_2^B}$$

and thus $j^s(i,\bar{R}_2^B,\bar{R}_1^B) = j^s(i,R_2^B,R_1^B)$. Then consider another system with

$$\bar{\bar{R}}_1^B = R_1^B \geq \bar{\bar{R}}_2^B > R_2^B > 0.$$

Choose

$$k = \frac{\bar{\bar{R}}_2^B}{R_2^B} > 1$$

so that $\bar{\bar{R}}_2^B = \bar{R}_2^B$ and $\bar{R}_1^B = kR_1^B > R_1^B = \bar{\bar{R}}_1^B$. By above, $j^s(i,\bar{R}_2^B,\bar{R}_1^B) = j^s(i,R_2^B,R_1^B)$. However, by A), optimal policy is monotonically decreasing in R_1^B, so $j^s(i,\bar{\bar{R}}_2^B,\bar{\bar{R}}_1^B) \geq j^s(i,\bar{R}_2^B,\bar{R}_1^B)$. Therefore, $j^s(i,\bar{\bar{R}}_2^B,R_1^B) = j^s(i,\bar{\bar{R}}_2^B,\bar{\bar{R}}_1^B) \geq j^s(i,R_2^B,R_1^B)$.

Proof for Decreasing in m: Choose $\infty > \bar{m} > m > 0$. The proof is similar to the proof for λ_1 in A), except for the initial values $V_0(i,j,\bar{m}) = V(i,j,m)$ for $0 \leq i \leq m$ and $V_0(i,j,\bar{m}) = V(m,j,m)$ for $m < i \leq \bar{m}$.

Proof for D): Choose $\infty > \bar{n} > n > 0$. Define $d_n = \bar{n} - n$. We use initial values $V_0(i,j,\bar{n}) = V(i,j-d_n,n)$ for $d_n \leq j \leq \bar{n}$ and $V_0(i,j,\bar{n}) = V(i,0,n) + (d_n - j)\Delta(i,0,n)$ for $0 \leq j < d_n$ so that $\Delta_0(i,j,\bar{n}) = \Delta(i,j-d_n,n)$ for $d_n \leq j \leq \bar{n}-1$ and $\Delta_0(i,j,\bar{n}) = \Delta(i,0,n)$ for $0 \leq j < d_n$. The proof is then similar to the proof for λ_1 in A) to show that $\Delta(i,j,\bar{n}) < \Delta(i,j-d_n,n)$ for $0 \leq i \leq m$ and $d_n \leq j \leq \bar{n}-1$. Hence $j^s(i,\bar{n}) - (\bar{n}-n) \geq j^s(i,n)$ on $j^s(i,n) > 0$ and the result directly follows.

Now, we further investigate the variation of the optimal admission threshold on type-2 customers with changes in system parameters, through a couple of examples. The optimal value function is found by successive approximation, and from that the optimal policy is inferred. For both experiments, we show the threshold switching curve, $j^s(i)$, for the optimal policy. The nominal system parameters are: $m = 6$, $n = 5$, $\alpha = 0.9999$, $Q = 100000$, $\lambda_1 = 1$, $\lambda_2 = 0.8$, $\mu_A = 10$, $\mu_B = 5$, $R_1^A = 100$, $R_1^B = 300$, $R_2^B = 50$, and $\mathbf{P} = 0$. Example 1 (Fig. 3) demonstrates the variation of μ_B. We find that the number of spaces reserved at station B for type-1 traffic, given a fixed number of customers at A, decreases steadily as the service rate at station B increases. Faster service at B decreases the likelihood that a type-2 customer, if admitted, will block a type–1 customer. Consequently, it is more favorable to admit type-2 customers. Example 2 (Fig. 4) demonstrates the variation of R_2^B. We can observe that the reservation at B for type-1 customers decreases steadily as the revenue generated by type 2 approaches that of type-1 customers at station B. As we expected, if these two revenues are equal, then type-2 customers are always admitted.

VI. Conclusion

We have analyzed access control policies in the tandem multiserver loss queues. The optimal policy was found to be a monotonically decreasing threshold. Furthermore, monotonic variation of the threshold was proven for most system parameters. We hope this simple model will give us insight into sequential resource allocation in larger multiple service, multiple resource loss networks, which are increasingly common in telecommunication systems. It is of note that the structural properties proven here can be extended to an arbitrary number of parallel first stage queues [10]. It would be of value to extend these results to more general multiserver loss networks or to study methods to use these results to model the congested segments of such networks.

References

[1] I. Adler and P. Naor, "Social optimization versus self-optimization in waiting lines," Dept. Operat. Res., Stanford Univ., Stanford, CA, Tech. Rep. 4, 1969.
[2] J. P. C. Blanc, P. R. de Waal, P. Nain, and D. Towsley, "Optimal control of admission to a multiserver queue with two arrival streams," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 785–797, 1992.
[3] E. Davis, "Optimal control of arrivals to a two-server queueing system with separate queues," Ph.D. dissertation, Program Operat. Res., North Carolina State Univ., Raleigh, 1977.
[4] H. Ghoneim and S. Stidham, "Optimal control of arrivals to two queues in series," *Europe. J. Operat. Res.*, 1985.
[5] S. G. Johansen and S. Stidham, "Control of arrivals to a stochastic input-output system," *Adv. Appl. Prob.*, vol. 12, pp. 972–999, 1980.
[6] S. Jordan and P. P. Varaiya, "Throughput in multiple service multiple resource communication networks," *IEEE Trans. Commum.*, vol. 39, no. 8, pp. 1216–1222, 1991.
[7] P. R. Kumar and P. Varaiya, *Stochastic Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1986.
[8] N. C. Knudsen, "Individual and social optimization in a multi-server queue with a general cost-benefit structure," *Econometrica*, vol. 40, pp. 515–528, 1972.

[9] N. C. Knudsen and S. Stidham, "Individual and social optimization in a birth-death congestion system with a general cost-benefit structure," Dept. Indust. Eng., North Carolina State Univ., Raleigh, Tech. Rep., pp. 76–78, 1976.
[10] C.-Y. Ku and S. Jordan, "Access control of parallel multiserver loss queues," preprint.
[11] W. Lin and P. R. Kumar, "Optimal control of a queueing system with two heterogeneous servers," *IEEE Trans. Automat. Contr.*, vol. AC-29, no. 8, pp. 696–703, 1984.
[12] P. Naor, "On the regulation of queue size by levying tolls," *Econometrica*, vol. 37, pp. 15–24, 1969.
[13] A. Simonovits, "Self and social optimization in queues," *Studia Sci. Math. Hungarica*, vol. 11, pp. 131–138, 1976.
[14] S. Stidham and R. Weber, "A survey of Markov decision models for control of networks of queues," *Queueing Systems 13*, pp. 291–314, 1993.

Convergence Rate for RLS-Based Direct Self-Tuning Minimum-Variance Regulation of ARMAX Minimum Phase Plants

A. Niederliński

***Abstract*—An upper bound for the error convergence rate of recursive least squares (RLS)-based direct self-tuning minimum-variance (DSTMV) regulation of minimum phase multiple time-delay AutoRegressive Moving Average with eXogenous input (ARMAX) plant is derived. The RLS algorithm used is k-interlaced, with k being the plant time-delay. The bound is derived for known fixed b_0 by extending a recently proposed methodology. The bound provides a joint explanation of DSTMV regulation stability and parameter estimate convergence. The paper demonstrates that self-tuning is based on convergence properties of RLS as well as on the excitation quality of plant white noise, which generates (*via* controller feedback) the plant input.**

***Index Terms*—Minimum-variance regulation, recursive least squares, self-tuning.**

I. Introduction

The oldest self-tuning controller is undoubtedly the direct (implicit) self-tuning minimum-variance (DSTMV) controller, based on recursive least squares (RLS) estimation. Its main feature is that RLS determines *directly* the controller parameters, which *implicitly* establish a plant model. This is in contrast with indirect (explicit) self-tuning, where *explicitly* estimated plant parameters are used to *indirectly* determine the controller parameters.

The idea of DSTMV regulation was formulated by Peterka [12] and developed by Åström and Wittenmark [1], who also presented the first results of its asymptotic properties. DSTMV as applied to a k-delayed AutoRegressive Moving Average with eXogenous input (ARMAX) minimum phase plant is based on the idea of restructuring the plant model so that the k-step predicted plant output is a linear function of all controller parameters; they may therefore be estimated using RLS. It follows that: 1) self-tuning may be done *via* simple and well-understood RLS, even for ARMAX plants, for which the *indirect* approach requires nonlinear estimation algorithms and 2) controller parameters being estimated, there is no need to design the control law by solving a Diophantine equation at each step of the self-tuning recursions.

The price paid for these advantages seems to be minor: 1) for direct self-tuning a few more parameters need to be estimated than the number of parameters in the usual input–output model and 2) any change of control law (e.g., control input weighting) must be accommodated by repeating the estimation for a differently restructured plant model, whereas an estimated input–output model may be used to design a broad range of control laws.

Although simple, the DSTMV idea is notorious for resisting attempts at theoretical justifications. It is relatively straightforward to demonstrate that the k-step predicted plant output converges toward the minimum variance (MV) prediction error given by a moving average (MA) time-series of order $k-1$, provided the plant is controlled by a controller converging to the target MV controller. However, to demonstrate that the CE DSTMV controller is in fact converging to the target MV controller seems to be a problem considered by many as not yet solved, even for AutoRegressive with eXogenous input (ARX) plants; see, e.g., [6] and [14].

The oldest approach to self-tuning convergence analysis is the *ODE-approach* proposed by Ljung; see [7]. Its essence is to analyze local convergence points of an averaged nonstochastic *Ordinary Differential Equation,* which approximates the stochastic discrete-time system in a compressed time scale. The (implicit) stability assumption of this approach forfeits a complete explanation of the self-tuning mechanism.

Another approach is based on martingale theory. Its first broad exposition was presented by Goodwin and Sin (see [3]) and later developed by Chen and Guo (see, e.g., [2]). The main advantages of this approach are its fundamental nature, the power of already existing martingale theory, and the fact that both stability and convergence are analyzed simultaneously.

The lack of satisfactory theoretical results was perhaps partially responsible for the decline of interest in RLS-based DSTMV regulation for ARMAX plants, in favor of the more obvious (but numerically more demanding) approach relying upon modified extended RLS (see, e.g., [3]) or extended RLS (see, e.g., [2]). This is to be regretted because DSTMV regulation with RLS parameter estimation is a robust technique (see, e.g., [8]), easily accommodating all types of MV and pole-zero placement control laws for minimum phase as well as for nonminimum phase plants.

The aim of the paper is to present a result guaranteeing stability and parameter convergence for RLS-based DSTMV regulation of minimum phase ARMAX plants with multiple time-delay. The main result is an almost sure (a.s.) upper bound on the convergence rate of the estimation error. It is shown that the mechanism of getting stability and parameter convergence is based on RLS convergence properties as well as on excitation properties of plant white noise, which generates (*via* a generally nonstationary pole-zero filter feedback) the plant excitation in the process of self-tuning. The result supports an old conjecture attributing self-tuning mainly to some properties of RLS; see, e.g., [3].

II. DSTMV Regulation—A Summary

The word *plant* means in the sequel an open-loop asymptotically stable entity, which responds to the input time series $\{u(t)\}$ with

Manuscript received February 21, 1996; revised July 11, 1996. This work was supported in part by the Polish Committee of Scientific Research (KBN) under Grant 8 T11A 006 13.
The author is with the Instytut Automatyki, Politechnika Ślaska, PL-44-100 Gliwice, Poland (e-mail: aniederlinski@ia.polsl.gliwice.pl).
Publisher Item Identifier S 0018-9286(97)04280-3.