Technical Report CC212, Department of Electrical Engineering, Technion, October 1997
Shorter version appeared in *Operations Research* 46, Dec. 1998, pp. 776–784.

Individual Equilibrium and Learning in Processor Sharing Systems [1]

EITAN ALTMAN[2] and NAHUM SHIMKIN[3]

Abstract: We consider a processor-sharing service system, where the service rate to individual customers decreases as the load increases. Each arriving customer may observe the current load, and should then choose whether to join the shared system. The alternative is a constant-cost option, modeled here for concreteness as a private server (e.g., a personal computer which serves as an alternative to a central mainframe computer). The customers wish to minimize their individual service time (or an increasing function thereof). However, the optimal choice for each customer depends on the decisions of subsequent ones, through their effect on the future load in the shared server. This decision problem is analyzed as a non-cooperative dynamic game between the customers. We first show that any Nash equilibrium point consists of threshold decision rules, and establish the existence and uniqueness of a *symmetric* equilibrium point. Computation of the equilibrium threshold is demonstrated for the case of Poisson arrivals, and some of its properties are delineated. We next turn attention to a plausible dynamic learning scheme, where arriving customers employ a simple decision rule solely based on observed performance history in the shared system. Convergence to the symmetric equilibrium point is established, thereby under-scoring the relevance of the Nash equilibrium solution for this system.

Subject classification: Queues, optimization: processor sharing. Games, stochastic: dynamic equilibrium and learning.

---

[1]This research was supported in part by a joint grant from the Israeli Ministry of Science, and the French Ministry of Research and Technology.

[2]INRIA, 2004 Route des Lucioles, B.P.93, 06902 Sophia-Antipolis Cedex, France.

[3]Electrical Engineering Department, Technion – Israel Institute of Technology, Haifa 32000, Israel.

Introduction

The quality of service experienced by an individual customer in a shared service system often depends on the current load. Thus, in order to estimate the expected service quality, a user must consider not only the current load at her arrival instant, but also how this load might develop throughout her service period. We study in this paper the implications of this observation for self-optimizing users, who consider joining the shared system. The model that we consider is motivated in part by the following scenario.

Potential computer users, each requiring the use of a computer to execute a given job, arrive sequentially at a computer facility. Each user, upon arrival, may choose between the following two options: either to connect to a central mainframe computer (MF), which is normally serving many users in parallel; or to use a personal computer (PC). Each user is solely interested in minimizing her own service time (which coincides here with the sojourn time).

Service at the MF computer is performed according to the Processor Sharing discipline (e.g., Jaiswal (1982), Ramaswami (1984)), where available computing power is equally divided among all users present. Consequently, the service rate to each user decreases as the load increases (although the *total* service rate may actually increase). An arriving user may observe the current load, namely the number of users already in MF. However, to evaluate her expected service time at MF she must take into account the possible load on this computer throughout her service time, which in turn is affected by the decisions of subsequent users. This leads us to consider the resulting decision problem in a game theoretic framework, and explore the Nash equilibrium solution for the resulting dynamic game.

Although we find it convenient to refer to this PC–MF application, the basic model is relevant to a wide range of application areas which involve shared service, ranging from transportation and recreation to computing and telecommunications. We mention here two relevant applications in the latter area.

(*i*) Consider a situation where users can communicate with each other either through a Local Area Network (LAN), or through the public network, e.g., by connecting to the telephone network via a modem, which is typically slower. However, the throughput available to each user on the LAN decreases as the total workload increases. This is especially the case in LANs where a single channel should be shared between all users, e.g. the FDDI (Fiber Distributed Data Interface). The LAN can thus be approximated by a processor sharing queue, whereas the public network may be viewed as assigning a private server to each session.

(*ii*) Consider a non real time application, such as data transfer, on an ATM (Asynchronous Transfer Mode) network (see ATM Forum (1996)). ATM networks support both guaranteed services as well as best-effort services. Guaranteed services are CBR (Constant Bit Rate), in which a fixed amount of bandwidth is assigned to a session, and VBR (Variable Bit Rate) in which some average and peak bit-rates are assigned to a session. Best effort services are ABR (Available Bit Rate) and UBR (Unspecified Bit Rate); in both cases, some available bandwidth is shared among the connections that use these services. At a session level, ABR and UBR services can be approximated by a processor sharing queue, whereas CBR and VBR services can be approximated by a single server, dedicated for one session.

Our study focuses on two main issues. In the first part of the paper, we explore the properties of the Nash equilibrium solution. The main results here are the existence, uniqueness, and structural characterization of a *symmetric* equilibrium policy. The required analysis of the processor sharing queue relies on stochastic coupling arguments, rather than explicit calculations, which facilitates the consideration of both general inter-arrival times and of state-dependent total service rates.

The second issue concerns the descriptive power of the Nash equilibrium solution. The question arises as to what extent does the Nash equilibrium actually describe the system operation under realistic conditions, and what is the mechanism that might lead to this equilibrium. We consider a dynamic learning scenario, where users make simple decisions based on past performance statistics. Since the observed statistics depend on past decisions, this leads to a closed-loop adaptive decision problem. It is established that the emerging decision rule for each user, and consequently the system operating characteristics, converge to the Nash equilibrium point. The analysis here uses the theory of stochastic approximations with controlled Markov state.

Dynamic control of queueing systems has been the subject of considerable research, and surveys may be found in Stidham (1985), Walrand (1988). Social optimization of a Processor Sharing queue has been considered in de Waal (1988). Individual optimality has been studied and compared to social optimality in various models, under the first-come-first-served (FCFS) service discipline; see Naor (1969), Yechiali (1972), Bartroli and Stidham (1992), and the above-mentioned surveys. Under a FCFS discipline the expected sojourn time of any single customer is completely determined by the queue length at her arrival, so that the individually optimal decision policy is trivial in that case.

There has been some work on the Nash equilibrium concept for the study of individually optimal dynamic control of queueing systems, see Glazer and Hassin (1986), Hassin and Haviv (1994). Game theoretical analysis has been applied to other queueing control problems, e.g., Kulkarni (1983), Lee and Cohen (1985), Bovopoulos and Lazar (1987), Shenker (1990), Hsiao and Lazar (1991), Altman (1992), Altman and Koole (1992), Altman and Hordijk (1995), Shimkin and Shwartz (1993), where the last five consider dynamic problems. Finally, results related to our work have been obtained in Assaf and Haviv (1990), Haviv (1991) and in Xu and Shantikumar (1993).

The organization of the paper is as follows. The model is presented in Section 1. In Section 2 the Nash equilibrium is studied, assuming i.i.d. inter-arrival times, possibly state-dependent service rate, and exponential service requirements. It is shown that in any equilibrium point, the decision rule of each user is a threshold rule. Existence and uniqueness of a *symmetric* equilibrium point are then established, using certain monotonicity and continuity properties of the service time at MF. The proofs of these results are established in Appendix A using stochastic-coupling arguments. Section 3 concerns the actual calculation of the symmetric equilibrium. Formulas are derived for the case of Poisson arrivals, and are illustrated by numerical examples. In Section 4 the (individually optimal) Nash solution is compared with the socially optimal one. Section 5 presents the proposed dynamic learning scheme, and establishes its convergence to the symmetric Nash equilibrium. The detailed proof of this result is given in Appendix B. We close with some concluding remarks in Section 6.

1 The Model

Consider a service system which consists of two service facilities, Q_{MF} and Q_{PC}. Customers (users) arrive at this system sequentially, with inter-arrival times which are independent and identically distributed, and have finite mean λ^{-1}. Simultaneous arrivals are excluded. Each arriving customer observes the number of customers in Q_{MF} at her arrival instant, and should choose whether to join Q_{MF} or Q_{PC}.

We assume that Q_{MF} has a buffer size B, which may be finite or infinite. A customer that arrives when the buffer is full cannot be admitted and must turn to Q_{PC}.

The service at Q_{MF} is exponential with rate $\mu(x)$, where $x \geq 1$ is the number of customers in Q_{MF}. The service discipline is Processor Sharing, so that the service intensity for each customer equals $\nu(x) \stackrel{\triangle}{=} \mu(x)/x$. The alternative Q_{PC} offers a fixed expected service time θ^{-1}. In both queues service commences immediately upon admission, so that the sojourn time coincides with the service time.

We make the following assumptions on the service rate at Q_{MF}:

(i) $0 < \mu(x) \leq \mu_{max}$ for every $x \geq 1$ (bounded service rate);

(ii) $\nu(x) = \mu(x)/x$ is strictly decreasing in x. Thus, the service rate applied to each customer decreases as the load increases.

Let $X(t)$ denote the number of customers at Q_{MF} at time t, and suppose that the system starts at $t = 0$ with initial state $X(0) = x_0$. Let $T_k,\ k \geq 0$, denote the arrival time of customer C_k, where $0 = T_0 < T_1 < T_2 < \ldots$. Hence $t = 0$ is the arrival time of customer C_0.

Upon arrival, customer C_k observes the current queue length $X(T_k)$, and should then decide which queue to join. A randomized decision rule for customer C_k is therefore defined by $u_k = \{u_k(x),\ 0 \leq x < \mathsf{B}\}$, where $u_k(x) \in [0, 1]$ is the probability of joining Q_{MF} if a queue length $X(T_k) = x$ is observed. Let U denote the collection of such decision rules, and let $\pi = (u_0, u_1, \ldots) \in \Pi \stackrel{\triangle}{=} U^\infty$ denote the vector of decision rules of all the customers, to which we refer as a policy.

Denote by W_k the service duration of customer C_k, and let $\bar{W}_k(x, \pi)$ be the expected value of W_k given that x customers are present at Q_{MF} upon arrival of C_k, and that all customers (including C_k) follow the policy π. Then

$$\bar{W}_k(x, \pi) = u_k(x)\bar{V}_k(x, \pi) + (1 - u_k(x))\theta^{-1} ,$$

where $\bar{V}_k(x, \pi)$ is the expected service duration of C_k at Q_{MF} under the same conditions. We observe that $\bar{V}_k$ depends on π through $\{u_l,\ l > k\}$, the decision rules of subsequent customers.

We assume that the customers are self-optimizing, so that each wishes to minimize her own service time. To this end, she should obviously evaluate her expected service time at the two queues, namely $\bar{V}_k(x,\pi)$ and θ^{-1}, and choose the lower one.

The dependence of $\bar{V}_k$ on the decisions of other users leads us to study this problem within a game theoretic framework.

2 Nash Equilibrium Solution

We now consider the system as a non-cooperative game, in which each customer wishes to minimize her own expected service time. The main results of this section concern the characterization of the equilibrium points of this game.

For each policy $\pi = (u_0, u_1, \ldots)$, let π^{-k} denote the collection of all decision rules in π excluding the u_k, and let $[\pi^{-k} \mid u'_k]$ be the policy which replaces u_k by u'_k.

Definition 1 *A decision rule u_k is an* optimal response *for C_k against a policy π if*

$$\bar{W}_k(x, [\pi^{-k} \mid u_k]) \leq \bar{W}_k(x, [\pi^{-k} \mid u'_k]) \ , \quad \forall\, x \geq 0, \;\; u'_k \in U \quad . \tag{1}$$

Definition 2 *A policy $\pi = (u_0, u_1, \ldots)$ is a* Nash equilibrium policy *if u_k is an optimal response for C_k against π, for every $k \geq 0$.*

Thus, in equilibrium no one can gain by a unilateral change of her decision rule. Since the queue length x is observed prior to decision, we require this to hold for every possible value of x.

In the sequel we shall take special interest in equilibrium policies which are *symmetric*, namely the decision rules of all customers are identical. Such policies are natural here, since the specifications of all customers are the same, and they all face the same decision problem. (In the terminology of Schelling (1960), symmetric equilibria are the natural candidates for the *focal equilibrium* of the game.)

We define next a special class of decision rules, namely threshold rules. For any $0 \leq q \leq 1$ and integer $L \geq 0$, the decision rule u is an $[L, q]$-threshold rule if

$$u(x) = \begin{cases} 1 & \text{if} \quad x < L \\ q & \text{if} \quad x = L \\ 0 & \text{if} \quad x > L \quad . \end{cases} \tag{2}$$

A customer which employs this rule joins Q_{MF} if the queue length x is smaller than L, while if $x = L$ she does so with probability q. Otherwise she joins Q_{PC}. An $[L, q]$ threshold rule will be denoted by $[L, q]$, or more compactly by $[g]$ with $g = L + q$. Note that $[L, 1]$ and $[L + 1, 0]$ are identical. When the buffer size is finite, any $[g]$ with $g >$ B is equivalent to [B].

We now turn to the main results of this section.

Theorem 1

(i) For any equilibrium policy $\pi^* = (u_0, u_1, \ldots)$, each decision rule u_k is a threshold rule.

(ii) A symmetric equilibrium policy $\pi^* = (u^*, u^*, \ldots)$ exists, is unique, and u^* is a threshold rule.

The proof proceeds through some lemmas. The first two establish basic monotonicity and continuity properties of the service time in Q_{MF}.

Recall that $\bar{V}_k(x, \pi)$ denotes the expected service time at Q_{MF}.

Lemma 1 *For every policy π and $k \geq 0$, $\bar{V}_k(x, \pi)$ is strictly increasing in x.*
In fact, for every $x \geq 0$, $\bar{V}_k(x+1, \pi) - \bar{V}_k(x, \pi) > \delta_x$ for some positive δ_x, independent of π.

A detailed proof, based on stochastic coupling arguments, is given in Appendix A. The idea is simple. Consider two cases, the first when a customer (say C_0) enters at queue length x, and the second when she enters at $x+1$. Since the queue lengths change by at most one customer at a time, the queue length in the second case will be higher than in the first one, until such a time τ (possibly infinite) when they coincide. From then on the queue lengths will remain equal, since the decision policies for incoming customers are the same in both cases. It then follows that the service rate applied to C_0 throughout her stay is higher in the first case than in the second, and strictly so up until τ. This implies a lower expected service time. □

Recall that $[g]$ stands for a threshold rule, and let $[g]^\infty$ denote the stationary policy $\pi = ([g], [g], \ldots)$. Thus $\bar{V}_k(x, [g]^\infty)$ is the expected service time of customer C_k if she joins Q_{MF} at queue length x, while all subsequent customers are using the threshold rule $[g]$.

Lemma 2 *For every $k \geq 0$ and $x \geq 0$, $\bar{V}_k(x, [g]^\infty)$ is:*

(i) strictly increasing in $g \in [0, \mathsf{B}]$, and

(ii) continuous in $g \in [0, \mathsf{B}]$.

The proof, again using stochastic coupling arguments, is presented in Appendix A. We note that in fact we show there that $\bar{V}_k(x, [g]^\infty)$ is uniformly Lipschitz continuous in g on compact subsets (hence, in particular, in the finite buffer case) – a property that will be required in Section 5. □

Lemma 3 *Let π be an arbitrary policy, and let U_k^* be the set of decision rules for C_k which are optimal against π.*

(i) Any $u_k^* \in U_k^*$ is a threshold rule, with finite threshold, and is given by:

$$u_k^*(x) = \begin{cases} 1 & \text{if} \quad \bar{V}_k(x, \pi) < \theta^{-1} \\ q_x & \text{if} \quad \bar{V}_k(x, \pi) = \theta^{-1} \\ 0 & \text{if} \quad \bar{V}_k(x, \pi) > \theta^{-1} \end{cases}, \tag{3}$$

where $0 \leq q_x \leq 1$ is arbitrary. (Recall that θ^{-1} is the expected service time in Q_{PC}.)

(ii) Consequently, the set U_k^* is given as follows. Let L^* be the smallest non-negative integer such that $\bar{V}_k(L^*,\pi) \geq \theta^{-1}$. If $\bar{V}_k(L^*,\pi) = \theta^{-1}$, then $U_k^* = \{[L^*,q] : 0 \leq q \leq 1\}$. Otherwise, U_k^* consists of the single threshold rule $[L^*,0]$.

Proof: Assume that C_k observes x customers in Q_{MF} at her arrival. She should now choose to join either Q_{MF}, where her expected service time would be $\bar{V}_k(x,\pi)$, or Q_{PC}, where her expected service time is θ^{-1}. Obviously, the optimal decision is to choose the lower one, and is thus given by (3). By Lemma 1, $\bar{V}_k(x,\pi)$ is strictly increasing in x, so that (3) is indeed a threshold rule. Finiteness of the threshold should be checked for $\mathsf{B} = \infty$; it follows from the easily verified fact that $\bar{V}_k(x,\pi) \to \infty$ as $x \to \infty$. Thus (i) is established, and (ii) follows immediately from (i) after noting again that $\bar{V}_k(x,\pi)$ is increasing in x. □

Proof of Theorem 1:

(i) Let $\pi^* = (u_0, u_1, \ldots)$ be an equilibrium policy. By Definition 2 each u_k must be optimal against π^*, and from Lemma 3 it follows that u_k is a threshold rule.

(ii) By (i), a symmetric equilibrium policy π^* must consist of identical *threshold* policies, i.e. $\pi^* = [g]^\infty$ for some $g \in [0,\mathsf{B}]$. It remains to establish existence and uniqueness of a threshold $g^* \in [0,\mathsf{B}]$ such that $[g^*]$ is optimal for C_0 (hence, by symmetry, for any customer C_k) against $[g^*]^\infty$.

Define the point-to-set mapping $G^* : [0,\mathsf{B}] \to 2^{[0,\mathsf{B}]}$, which associates with every threshold $g \in [0,\mathsf{B}]$ the set of "optimal thresholds" against $[g]^\infty$, namely

$$G^*(g) = \{g' \in [0,\mathsf{B}] : [g'] \text{ is optimal against } [g]^\infty\} \ . \tag{4}$$

Thus, it is required to prove that G^* possesses a unique fixed point, i.e., a unique $g^* \in [0,\mathsf{B}]$ such that $g^* \in G^*(g^*)$. In essence, the required existence and uniqueness follow, respectively, from continuity and monotonicity properties of G^*. In fact, existence may be deduced by applying the Kakutani fixed point theorem, which is commonly used to establish existence of Nash equilibria; see, e.g., Başar and Olsder (1995). However, it will be more instructive to construct explicitly the graph of G^*.

By Lemma 3(ii), the set $G^*(g)$ can be expressed as follows. Let $\ell(g)$ denote the minimal integer $L \geq 0$ for which $\bar{V}_0(L,[g]^\infty) > \theta^{-1}$ (note that $\ell(g)$ is finite since $\bar{V}_0(x,[g]^\infty) \to \infty$ as $x \to \infty$). If $\ell(g) \geq \mathsf{B}$ then $G^*(g) = \{\mathsf{B}\}$; otherwise,

$$G^*(g) = \begin{cases} \{\ell(g) + q \,:\, 0 \leq q \leq 1\} & \text{if} \quad \bar{V}_0(\ell(g),[g]^\infty) = \theta^{-1} \\ \{\ell(g)\} & \text{if} \quad \bar{V}_0(\ell(g),[g]^\infty) > \theta^{-1} \end{cases} \tag{5}$$

We next argue that the graph of $G^*(g)$ has the "staircase" form depicted in Figure 1. Recall from Lemmas 1 and 2 that $\bar{V}_0(x,[g]^\infty)$ is strictly increasing in x and g, and is continuous in g. This implies that, as g increases from 0 to B, the map $g \to \ell(g)$ is non-increasing, piecewise constant, left continuous, and its downward jumps are of exactly one unit. Let $J \geq 0$ be the number of jumps of $\ell(\cdot)$, and, provided $J \geq 1$, let $0 < g_1 < \cdots < g_J$ be the jump points of $\ell(\cdot)$. By the above-mentioned properties of $\bar{V}_0(x,[g]^\infty)$ it also follows that $\bar{V}_0(\ell(g_j),[g_j]^\infty) = \theta^{-1}$ at each jump point g_j. Defining

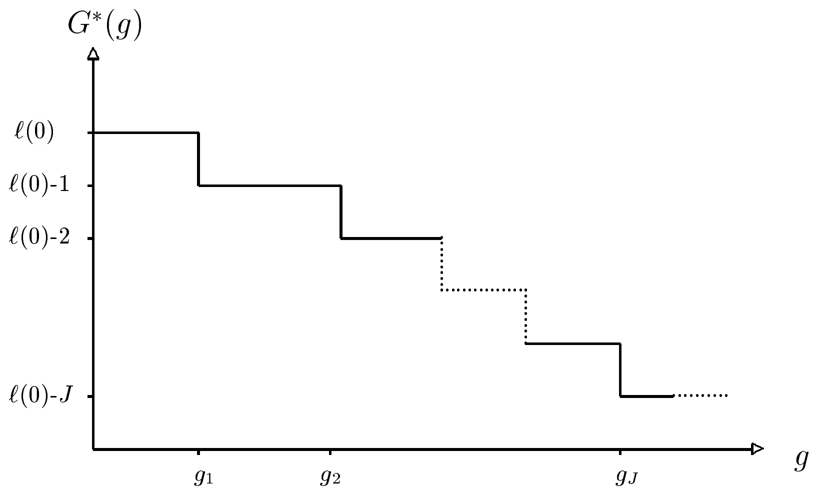

Figure 1: The optimal thresholds.

for notational convenience $g_0 = 0$ and $g_{J+1} = \infty$, $\ell(\cdot)$ is given by: $\ell(g) = \ell(0) - j$ for $g_j < g \leq g_{j+1}$, $j = 0, \ldots, J$. It then follows from (5) that (see Figure 1)

$$G^*(g) = \begin{cases} \{\ell(0)\} & \text{if} \quad g < g_1 \\ [\ell(0) - j, \ell(0) - j + 1] & \text{if} \quad g = g_j, \quad 1 \leq j \leq J \\ \{\ell(0) - j\} & \text{if} \quad g_j < g < g_{j+1}, \quad 1 \leq j \leq J \end{cases} \tag{6}$$

Obviously, the graph of G^* is intersected exactly once by the line with unit slope, say at point (g^*, g^*). It follows that g^* is the unique fixed point of $g \to G^*(g)$, and therefore $\pi^* \triangleq [g^*]^\infty$ is the unique symmetric equilibrium point. □

Remark: It is interesting to note that the equilibrium threshold g^* can be either an integer or a non-integer (corresponding to a deterministic or a randomized equilibrium policy), with neither case being generic. This observation, which is clearly evident in Figure 5, may be understood with the help of Figure 1, where the line of unit slope may intersect the graph of $G^*(g)$ in its horizontal (integer) or vertical (non-integer) part.

We close this section by pointing out some interesting generalizations.

1. For concreteness we used for the performance criterion the expected service time; however, our sample path proofs actually show that the waiting times in equilibrium are optimized in the sense of stochastic dominance, implying in particular that the cost could be defined as the expected value of any increasing function of the waiting time.

2. We can allow for a certain fraction (realized via random splitting) of the arrivals to be *uncontrolled*, in the sense they must join Q_{MF} whenever its buffer is not full. We shall find these useful in Section 5.

3. Lemma 1, and consequently part (i) of Theorem 1, hold even if the inter-arrival times are not identically distributed, and the expected service time in Q_{PC} are different for different customers.

3 Computation of the Equilibrium Threshold

Given the existence and uniqueness and the symmetric equilibrium, we consider in this section the computation of the corresponding equilibrium threshold. In the following subsection we characterize this threshold in terms of the expected waiting times in Q_{PC} under threshold policies. The computation of the equilibrium threshold, and in particular of the random part q^*, requires in general the numeric solution of an implicit equation. We then specialize our discussion to the case of Poisson arrivals and constant total service rate at Q_{MF}. Closed-form solutions are derived for this case, and the results are illustrated through a numerical example.

Our computations involve the expected service times $\bar{V}_k(x,[L,q]^\infty)$ in Q_{MF}. By symmetry these do not depend on the customer index k, which will be omitted.

3.1 Characterization

Lemma 4 *The equilibrium threshold $[L^*,q^*]$ may be determined by the following procedure.*
(a) If $\bar{V}(\mathsf{B}-1,[\mathsf{B}]^\infty) < \theta^{-1}$ then $L^* = \mathsf{B}$, $q^* = 0$. Otherwise,
(b) $L^* = \min\{L \geq 0 \,:\, \bar{V}(L,[L,1]^\infty) > \theta^{-1}\}$.

(b1) If $\bar{V}(L^*,[L^*,0]^\infty) \geq \theta^{-1}$, then the equilibrium threshold is $[L^*,0]$.

(b2) If $\bar{V}(L^*,[L^*,0]^\infty) < \theta^{-1}$, then the equilibrium threshold is $[L^*,q^*]$, where $0 < q^* < 1$ is the unique solution of

$$\bar{V}(L^*,[L^*,q^*]^\infty) = \theta^{-1}\,. \tag{7}$$

Proof: For (L^*,q^*) as defined above, it is required to show that $\pi^* := [L^*,q^*]^\infty$ is an equilibrium policy, namely that the threshold rule $[L^*,q^*]$ is an optimal response against π^*. Case (a) is straightforward, and follows from Lemma 3(ii). As for (b), note first that L^* is well defined even for $\mathsf{B}=\infty$, since $\lim_{L\to\infty}\bar{V}(L,[L,1]^\infty)=\infty$. In case *(b1)*, Lemma 3 implies that $[L^*,0]$ is an optimal response against $[L^*,0]^\infty$ if (i) $\bar{V}(L^*-1,[L^*,0]^\infty) \leq \theta^{-1}$ and (ii) $\bar{V}(L^*,[L^*,0]^\infty) \geq \theta^{-1}$. But (i) follows by the definition of L^* (note that $[L^*,0]$ and $[L^*-1,1]$ are the same), and (ii) is the defining condition for case *(b1)*. Concerning case $(b2)$, note first that existence and uniqueness of the solution q^* to equation (7) follows from the continuity and monotonicity of $\bar{V}$ in q. Now, Lemma 3 implies that $[L^*,q^*]$ is an optimal response against $[L^*,q^*]^\infty$ if (7) is satisfied. □

3.2 Some Explicit Calculations

We now specialize our discussion to the case of Poisson arrivals, and constant total service rate at Q_{MF} ($\mu(x) \equiv \mu$). For this case we determine the service times $\bar{V}(L, [L,q]^\infty)$, and obtain the solution q^* of equation (7).

Fix [L,q], and define $\bar{V}(x) := \bar{V}(x, [L,q]^\infty)$. The $\bar{V}(x)$, $0 \le x \le L$, is the solution of the following set of $L+1$ linear equations:

$$\begin{aligned}
\bar{V}(x) &= \frac{1}{\alpha} + \frac{\mu x}{\alpha(x+1)}\bar{V}(x-1) + \frac{\lambda}{\alpha}\bar{V}(x+1)\,, \quad 0 \le x \le L-2 && (8)\\
\bar{V}(L-1) &= \frac{1}{\alpha} + \frac{\mu(L-1)}{\alpha L}\bar{V}(L-2) + \frac{q\lambda}{\alpha}\bar{V}(L) + \frac{(1-q)\lambda}{\alpha}\bar{V}(L-1) && (9)\\
\bar{V}(L) &= \frac{1}{\mu} + \frac{L}{L+1}\bar{V}(L-1) && (10)
\end{aligned}$$

where $\alpha = \mu + \lambda$. These equations follow from the memoryless property of the system, which implies that V(x) equals the expected remaining service time of any customer present at queue length $x+1$. Thus, $\bar{V}(x)$ equals the expected time till the next transition, (α^{-1} in the first equations), plus the expected remaining service time after that transition. These equations may obviously be solved numerically for each given [L,q]; however, in order to obtain the optimal threshold in closed form we derive a more explicit solution. By (8), $\bar{V}(x)$ can be expressed as

$$\bar{V}(x) = a(x)\bar{V}(0) + b(x)\,, \quad 0 \le x < L\,, \tag{11}$$

where the coefficients $a(x)$ and $b(x)$ are obtained recursively by substituting (11) into (8), which yields

$$a(x+1) = \frac{1}{\lambda}\left[(\mu+\lambda)a(x) - \mu\frac{x}{x+1}a(x-1)\right]\,, \quad x \ge 1 \tag{12}$$

$$b(x+1) = \frac{1}{\lambda}\left[(\mu+\lambda)b(x) - \mu\frac{x}{x+1}b(x-1) - 1\right]\,, \quad x \ge 1 \tag{13}$$

with initial values $a(0) = 1$, $a(1) = (\mu+\lambda)/\lambda$, $b(0) = 0$, $b(1) = \lambda^{-1}$. Note that these coefficients *do not depend on L and q.* Next we obtain $\bar{V}(0, [L,q])$. By substituting $\bar{V}(L)$ from (10) into (9) and then substituting $\bar{V}(L-2)$ and $\bar{V}(L-1)$ from (11), we obtain after some algebraic manipulation

$$\bar{V}(0) = \frac{\mu^{-1} - \frac{1}{L+1}b(L-1) - q^{-1}\left(b(L) - b(L-1)\right)}{\frac{1}{L+1}a(L-1) + q^{-1}\left(a(L) - a(L-1)\right)}\,. \tag{14}$$

$\bar{V}(x)$ may now be obtained from (11)–(13).

We may now calculate the equilibrium threshold. We first compute L^* by Lemma 4. If we are in case (a) or $(b1)$ of Lemma 4 then $q^* = 0$. Otherwise, in case $(b2)$, we compute $0 < q^* < 1$ as the unique solution of (7). Using (10), (11) and (14) to express $\bar{V}(L^*, [L^*, q]^\infty)$ as a function of q, rearranging and canceling terms we obtain a linear equation for q^*, whose solution is

$$q^* = C^{-1}\left[\frac{a(L^*) - a(L^*-1)}{a(L^*-1)}\left(\theta^{-1} - \mu^{-1} - \frac{L^*}{L^*+1}b(L^*-1)\right) + \frac{L^*}{L^*+1}(b(L^*) - b(L^*-1))\right] , \tag{15}$$

where $C \triangleq \mu^{-1} - \theta^{-1}/(L^*+1)$.

3.3 A Numeric Example

We illustrate these results and some properties of the relevant quantities through a numerical example. We consider the parameters $\mathsf{B} = \infty$, $\theta = 10$ and $\mu = 100$.

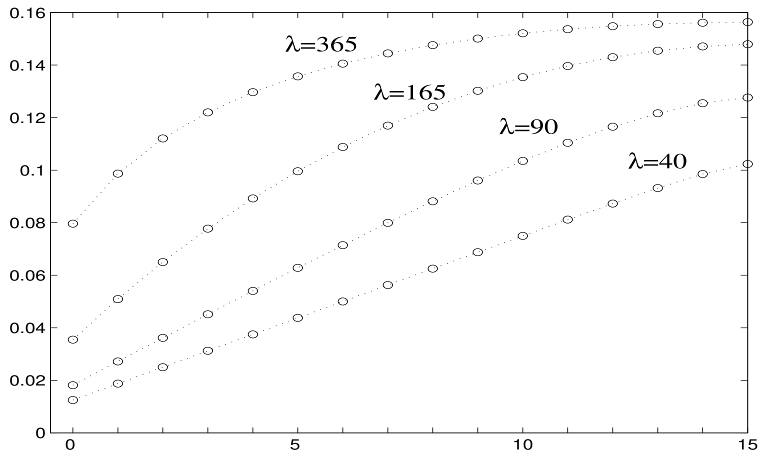

Figure 2: $\bar{V}(x, [15, 1])$ as a function of x, for different values of λ

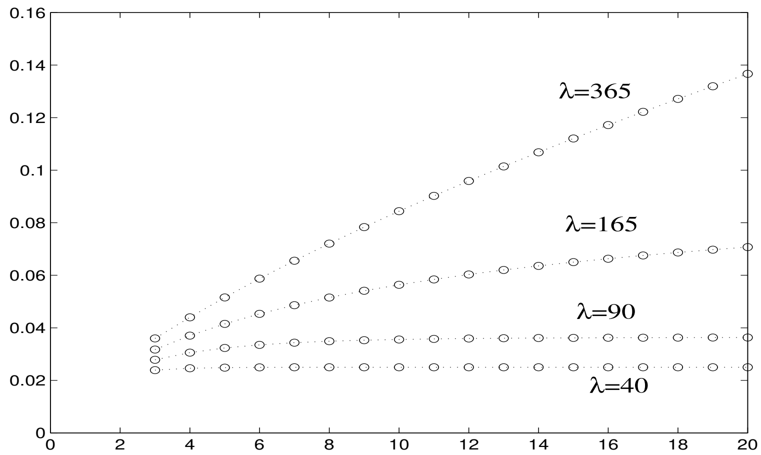

Figure 3: $\bar{V}(3, [L, 1])$ as a function of L, for different values of λ

In Figure 2 we depict $\bar{V}(x, [15, 1])$ as a function of x, for different values of λ. It can be seen that $\bar{V}(x, [15, 1])$ is bounded by 0.16 for any λ and x, and that it increases with both x and λ. Indeed,

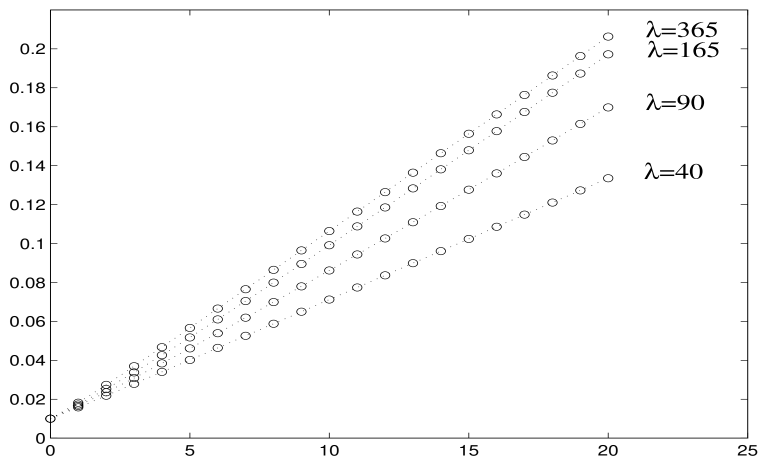

Figure 4: $\bar{V}(L,[L,1])$ as a function of L, for different values of λ

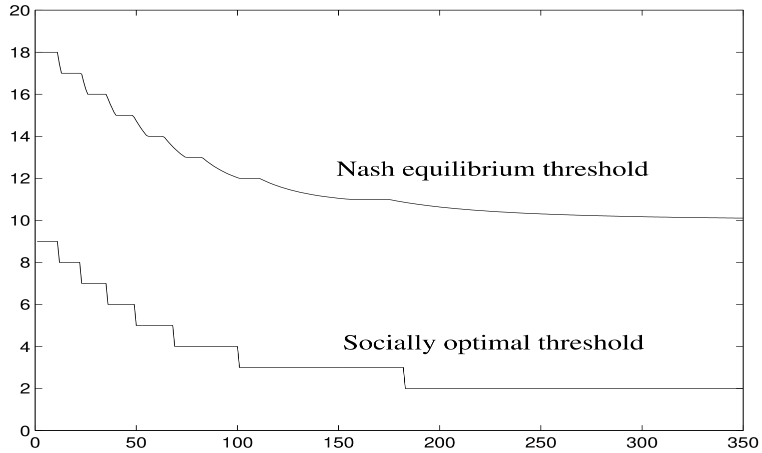

Figure 5: $g^* = L^* + q^*$ as a function of λ

for $\lambda \to \infty$ we have $\bar{V}(x,[15,1]) \to 0.16$ since in that case there will always be 16 customers in the queue, and therefore the expected service time of each one of them is $[\mu/16]^{-1} = 0.16$.

In Figures 3 and 4 we depict $\bar{V}(3,[L,1])$ and $\bar{V}(L,[L,1])$ as a function of L, for different values of λ. As expected, both are monotone increasing in L and in λ.

Finally, Figure 5 presents the equilibrium threshold $[L^*, q^*]$ as a function of λ. In that figure, a real number of the form 10.34 means $L^* = 10$ and $q^* = 0.34$. L^* may be obtained from Figure 4 as the first integer for which $\bar{V}(L,[L,1])$ exceeds $\theta^{-1} = 0.1$ (see Lemma 4). Note that L^* is a decreasing function of λ and approaches $\mu/\theta = 10$. Indeed, for $\lambda \to \infty$ and threshold $[L,0]$, the queue length approaches a constant L, and therefore the expected service time increases towards $(\mu/L)^{-1}$. Comparing with θ^{-1}, the individually optimal threshold decreases to $L^* = \mu\theta^{-1}$. Figure 5 also compares the equilibrium threshold to the socially optimal one. This will be discussed in the following Section.

4 Social, Individual, and Naïve Optimality

Depending on the nature and goals of the decision makers in the system, different solution concepts may be appropriate under different circumstances. Here we briefly compare the Nash equilibrium solution with the socially optimal one, and also touch upon individual optimality under simplified (naïve) assumptions.

Consider the social optimization problem, where the goal is to minimize the expected average sojourn time per customer. Here each customer's decision is evaluated not only for its affect on her own performance, but also for its affect on the others (the externality cost). Observing that this social cost does not depend on the service discipline in Q_{MF} (PS, FCFS, etc.), the problem becomes a standard one in queueing control, and is known to possess an optimal non-randomized threshold policy (see e.g. Stidham (1985)). Let L_{so} be the value of this optimal threshold.

We assert that L_{so} is not larger than the Nash equilibrium threshold g^*. This is a direct consequence of the presence of externality costs in Q_{MF} and their absence in Q_{PC}. Indeed, if we increase the threshold beyond g^*, then any customer entering Q_{MF} (rather than Q_{PC}) at queue size larger than g^* increases increases both her own sojourn time, by definition of g^*, and the sojourn time of others in Q_{MF} by increasing the load there. Therefore such a threshold is socially worse than g^*.

Let us calculate L_{so} for the case of Poisson arrivals and constant service rate μ. Suppose that the threshold $[L, 0]$ is used. Denoting $\rho = \lambda/\mu$, and using standard results for finite-buffer queues, the expected sojourn time in the system is given by

$$E[W_L] = \frac{\rho^L}{1-\rho^{L+1}}\left[(1-\rho)\theta^{-1} - (L+1)\mu^{-1}\right] + \frac{\mu^{-1}}{1-\rho}. \tag{16}$$

L_{so} is obtained as the integer that minimizes $E[W_L]$.

As discussed above, it is seen from Figure 5 that the socially optimal threshold is indeed lower than the Nash equilibrium threshold. The threshold L_{so} decreases as λ (and thus ρ) grows, and it reaches asymptotically $L_{so} = 1$; our calculations show that $L_{so} = 1$ for $\lambda \geq 815$. We already saw that as λ grows, the Nash equilibrium threshold tends to $[10, 0]$, i.e. 10 times higher than the socially optimal threshold. The expected sojourn time in Q_{MF} is thus 10 times higher when the individual (equilibrium) criterion is used. It is obvious that individual users will have a strong incentive to deviate from the socially optimal policy in this case.

Finally, it should be of interest to consider the case of naïve individual optimality. Here we have self-optimizing customers, who make the simplifying assumption that the load perceived at their arrival will not change throughout their service period. While false in general, this assumption may often be adopted in practice. Thus, a customer who sees x customers in Q_{PC} assumes that her expected sojourn time there will be $\mu(x+1)/(x+1)$, and compares this time to θ^{-1}. (Observe that for state-independent μ, this coincides with the case of a simple FCFS queue.) The resulting threshold is independent of the arrival rate λ, which indicates its deficiency for the processor sharing problem. In our example we obtain $L = \mu/\theta = 10$. It is not hard to verify that the naïve individual threshold will be lower than the individually-optimal equilibrium threshold (as long as $\mu(x)$ is non-decreasing in x, and there are no uncontrolled arrivals to Q_{PC}). This follows since for a customer who enters just below the threshold, the assumption that the queue length will not change is the worst possible one.

5 Learning and Equilibrium

The Nash equilibrium solution is defined from a normative viewpoint. It sets the "rational" choice for a sophisticated decision maker, who has global information about the system and can reason about the choices of others.

This section examines the relevance of the Nash equilibrium solution derived above, from a different, descriptive, viewpoint. We demonstrate that it naturally emerges in a dynamic learning scenario, which is a reasonable one for the system at hand. In the scenario considered here, customers base their decisions on statistical data which are accumulated by the server. No prior information is assumed regarding service in or arrivals to Q_{MF}, nor do customers employ game-theoretic considerations to arrive at their decisions.

Other dynamic schemes with partially rational behavior that give rise to the Nash equilibrium have been extensively studied within the game theoretic and engineering literature, mostly for the case where the underlying game is static. See, e.g., Li and Başar (1987), Hsiao and Lazar (1991), Lakshmivarahan (1981), and Fudenberg and Levine (1998).

5.1 The Learning Model

We consider the system of Section 1, and assume that the server monitors the average sojourn times of customers in Q_{MF}, depending on the queue length at their arrival instants. Newly arriving customers have access to this accumulated data, in addition to the queue length at their arrival, and may use it to assess their performance at Q_{MF} before deciding which queue to join.

The system starts at time 0 without any prior data or operating statistics regarding Q_{MF}. For every $t \geq 0$, let $N_t(x)$ denote the number of customers who had joined Q_{MF} at queue length x and already left it by time t. Let $\hat{V}_t(x)$ denote the (empirical) average service time of these $N_t(x)$ customers. A customer who arrives at time t may inspect the current vector $\hat{V}_t$.

Consider then a customer who arrives at time t, observes X_t customers in Q_{MF}, and has to choose between Q_{MF} and Q_{PC}. A natural decision rule for a customer who seeks to minimize her service time is:

$$\text{join } Q_{MF} \text{ if } \hat{V}_t(X_t) < \theta^{-1}, \text{ and join } Q_{PC} \text{ otherwise.} \tag{17}$$

It is our purpose to show that this learning decision rule, with some modifications that will be introduced below, converges to the one implied by the Nash equilibrium solution, as do the other performance measures such as average waiting times, number of customers in the queues, etc.

We introduce the following modifications to the basic model:

1. *Uncontrolled arrivals:* We assume that a nonzero fraction of the arrivals are uncontrolled, in the sense that they always join Q_{MF}. Uncontrolled arrivals, discussed at the end of Section 2, are a natural extension of the basic model; their importance for the proper convergence of the learning process is discussed below. The remaining arrivals are controlled, and may choose to join Q_{MF} or Q_{PC} as before.

   Specifically, we assume that each arrival has a probability $p_u > 0$ of being uncontrolled, independently of other arrivals. The resulting arrival rates for controlled and uncontrolled arrivals are then $\lambda_u = p_u \lambda$ and $\lambda_c = (1 - p_u)\lambda$, respectively.

   Uncontrolled arrivals, in this form or another, are essential to ensure that learning will not halt. Indeed, suppose the measured statistics at one time indicates that entering Q_{MF} at some queue length x is unfavorable. Then controlled arrivals at this queue size will cease, and no further data will be collected to update this (possibly misleading) statistics. Uncontrolled arrivals will provide the additional entries required to update the statistics for this queue size.

2. *Randomized decisions:* The decision rule (17) is completely deterministic – it prescribes a single choice based on the differences between the (estimated) waiting times in the two queues. It stands to reason that if this difference is small, some other factors may influence the actual decision. We model these factors by possibly randomized decisions, and consider the following decision rule in place of (17):

$$\text{join } Q_{MF} \text{ with probability } \; S_\epsilon(\theta^{-1} - \hat{V}_t(X_t)) \, , \tag{18}$$

   where ϵ is a small positive parameter, and S_ϵ is an increasing function, with $S_\epsilon(x) = 0$ for $x \leq -\epsilon$, and $S_\epsilon(x) = 1$ for $x \geq \epsilon$. More specifically, let $S_\epsilon(x) = S(x/\epsilon)$, where S is a monotone increasing, Lipschitz-continuous (i.e., with bounded slope) function, with $S(x) = 0$ for $x \leq -1$, and $S(x) = 1$ for $x \geq 1$.

   Besides its plausibility, the inclusion of randomized decisions facilitates a direct correspondence between the learning algorithm and the Nash equilibrium solution, in case that the latter is randomized. The deterministic decision rule (17) would then give rise to more complicated behavior, whereby a randomized threshold can best be approximated by time switching between two deterministic ones.

We further assume, for simplicity, that the buffer of the Q_{MF} is finite, i.e. $\mathsf{B} < \infty$.

It is easily verified that the results of Section 2 (including the lemmas and main theorem) remain valid in the case of additional uncontrolled arrivals. The *value* of the equilibrium threshold, which we shall denote by g^*, will of course be affected – it should decrease with the fraction p_u of uncontrolled arrivals.

5.2 Convergence

We now examine the asymptotic behavior of this dynamic learning process, and establish its convergence to the symmetric Nash equilibrium.

Theorem 2 *Let $\hat{u}_t$ denote the decision rule implied by (18) at time t, namely*

$$\hat{u}_t(x) = S_\epsilon(\theta^{-1} - \hat{V}_t(x)) \,. \tag{19}$$

(a) $\hat{u}_t$ converges (almost surely) to a decision rule u^ϵ which, for all ϵ small enough, is a threshold rule.

(b) For $\epsilon \to 0$, the limiting decision rule u^ϵ converges to the Nash equilibrium policy $[g^*]$.

An immediate consequence is that, for ϵ small enough, the long-term performance of the learning system coincides with the expected performance under the (stationary) Nash equilibrium policy. This pertains both to individual performance measures, namely the individual service times, and to system-oriented performance measure, such as the average load, server utilization, and arrival statistics in the different queues.

For the proof we shall on the theory of the stochastic approximations algorithm, and in particular the ODE (Ordinary Differential Equation) method for its convergence analysis (Kushner and Clark (1978), Benveniste, Metevier and Priouret (1982)). We present below a rough outline of this approach as it applies to our system, while a detailed proof is given in Appendix B.

The basic idea of the ODE method is to establish a correspondence between the asymptotic behavior of the random sequence $\hat{V}_t$ and the solution of an appropriate set of (deterministic) differential equations. The question of convergence of this random sequence then becomes a question of convergence of the solutions of the latter equations to an equilibrium point. We shall show below that our learning system may be recast in the form of a stochastic approximations iteration, and formulate the corresponding ODEs. We further show that the latter posses a unique equilibrium point which is globally stable, and, furthermore, corresponds (for small ϵ) to the symmetric Nash equilibrium point of the system. Put together, these arguments imply the convergence of the estimates $\hat{V}_t$ and the corresponding decision rules $\hat{u}_t$ to the Nash equilibrium.

We proceed here to clarify the relevance of the stochastic approximations framework to our model, and motivate the resulting ODE. Let $T_m(x)$ be the m-th update time of $\hat{V}_t(x)$, which is the time of the m-th departure from Q_{MF} of a customer who entered at queue size x, and let $W_m(x)$ be the service time of that customer. Denote $V_m(x) = \hat{V}_{T_m(x)}(x)$. Then $V_m(x) = \frac{1}{m}\sum_{k=1}^{m} W_k(x)$, or, equivalently,

$$V_m(x) = V_{m-1}(x) + \frac{1}{m}(W_m(x) - V_{m-1}(x)) \,. \tag{20}$$

We aim, roughly, to show that the vector V_m converges to a constant vector v_∞, for which the expected value of the last term in this equation is null. Note that the statistics of the measured service time $W_m(x)$ depends on past and present values of $\hat{V}$, since these determine the decisions of other customers through (18). As m becomes large, the "vanishing gain" term $\frac{1}{m}$ in (20) causes V_m to change slowly, so that the expected value of $W_m(x)$ approximately depends only on the present value of the entries of the vector $\hat{V}$. Using an appropriate time normalization and extrapolation, this dynamics may then be approximated by the ODE:

$$\frac{d}{dt}v_t(x) = \beta_t(x)[w^\epsilon(v_t,x) - v_t(x)]\,, \quad 0 \le x < \mathsf{B}\,, \tag{21}$$

where $w^\epsilon(v,\cdot)$ is the vector of expected waiting times which results when the "frozen" estimate $\hat{V}_t \equiv v$ is used in the decision rule (18), and $\beta_t(x)$ is an appropriately defined positive rate function.

The equilibrium point v of the latter ODE is obviously obtained as the solution of the vector equation $w^\epsilon(v) - v = 0$. Lemmas 5 to 7 of Appendix B establish the required properties of this equilibrium point, namely its global stability, and its relation to the Nash equilibrium point.

Two interesting features of our analysis should be noted. Observe that the different entries of the vector $\hat{V}_t$ (corresponding to different queue lengths x) are updated asynchronously, with distinct rates for each component, and these rates actually depend on the estimated vector $\hat{V}_t$ (through its affect on the decision rule, hence on the queue length process). This dependence gives rise to rate variables in the ODE ($\beta_t(x)$ in equation (21) above) which are updated dynamically as part of the system state. The precise equations are given in (A14), and the details of the construction (which transforms the problem to a standard stochastic approximations form by incorporating the rates as part of the estimated vector) are given in Lemma 8.

Our analysis requires to formulate the learning system as a controlled Markov chain (where the control variable is the estimated vector). As we shall see (in Lemma 8 below), the state of this process should include all data needed for the future updates of $\hat{V}_t$, and in particular it includes the waiting times so far of all customers currently in service. This gives rise to a continuous (non countable) state space, and requires the use of stochastic approximations results which apply to such processes.

5.3 A Simulation Example

We illustrate the learning behavior using a simulated example. This simulation was performed with a constant service rate $\mu = 100$ and buffer size $B = 20$ at Q_{MF}, $\theta^{-1} = 0.1$, and Poisson arrivals with rate $\lambda = 120$, of which $\lambda_c = 100$ are controlled arrivals and $\lambda_u = 20$ are uncontrolled. The Nash equilibrium threshold for these parameters equals $g^* = 11.288$, as calculated numerically using the obvious extension of equations (8) to include uncontrolled arrivals. The randomization interval parameter was chosen as $\epsilon = 0.001$, which means that randomized decisions will be implied by the decision rule (18) only if $\hat{V}$ is in the range $(\theta^{-1} - \epsilon, \theta^{-1} + \epsilon) = (0.099, 0.101)$, i.e., within one percent of θ^{-1}. We note that the equilibrium values $\bar{V}(x, [g^*]^\infty)$ for queue lengths x =10, 11 and 12 are (.9750, .1000, .1030) respectively. The equality $\bar{V}(11, [g^*]) = \theta^{-1}$ is of course expected because of the equilibrium threshold value. Note that the two neighboring values are outside the randomization band $\theta^{-1} \pm \epsilon$, so that as equilibrium is approached, randomized decisions will occur only at queue length $x = 11$.

The system was simulated over a time interval of 4000 time units (where each time unit corresponds to approximately 120 external arrivals). The behavior of the average waiting time statistics

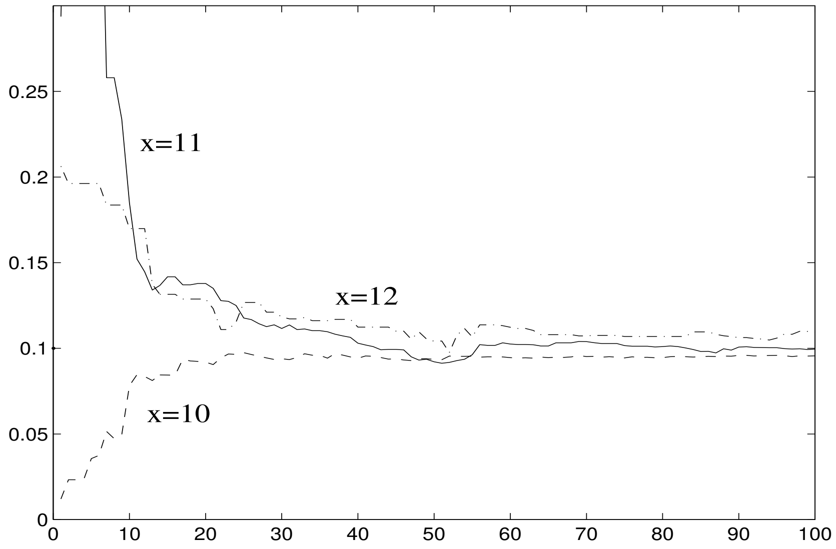

Figure 6: $\hat{V}_t(x)$ as a function of time

$\hat{V}_t(x)$ during the initial 100 time units is depicted in Figure 6, for entries at the queue lengths x =10, 11 and 12. We can see that fairly large average waiting times were obtained initially. This was caused mainly by the choice of null initial conditions for $\hat{V}_t$, which encouraged customers to initially join Q_{PC} at all queue sizes. However, as soon as this unfavorable statistics were observed, controlled arrivals temporarily stopped even at relatively low queue sizes, which caused low waiting times for those who did enter, and balanced the average statistics.

Beyond that initial period, reasonable convergence may be observed. The variable $\hat{V}_t(10)$ was the last to leave the band $\theta^{-1} \pm \epsilon$, around time 130, and beyond that time all arrivals at queue sizes other than $x = 11$ behaved exactly according to the Nash policy. As to the latter ($x = 11$), we can see in Figure 7 the value of the randomized decision $u_t(11) = S_\epsilon(\theta^{-1} - \hat{V}_t(11))$ over the full simulation period. The dashed line denotes the equilibrium value $q^* = 0.288$. Good agreement is seen here also.

6 Concluding Remarks

The fundamental issue that was considered in this paper is the effect of expected load buildup on individual user decisions, and consequently on system performance, in shared service facility. Assuming symmetric users, we have shown the existence of a unique Nash equilibrium point, and how this equilibrium emerges as a result of a simple learning scenario.

We conclude by pointing to some issues that deserve further investigation. An important extension of the model would be to the case of multiple customer classes (for example, corresponding to different cost parameters). In this case, customers of different classes are expected to employ different decision rules, and the question of uniqueness of the equilibrium policy becomes multi-dimensional and harder to resolve.

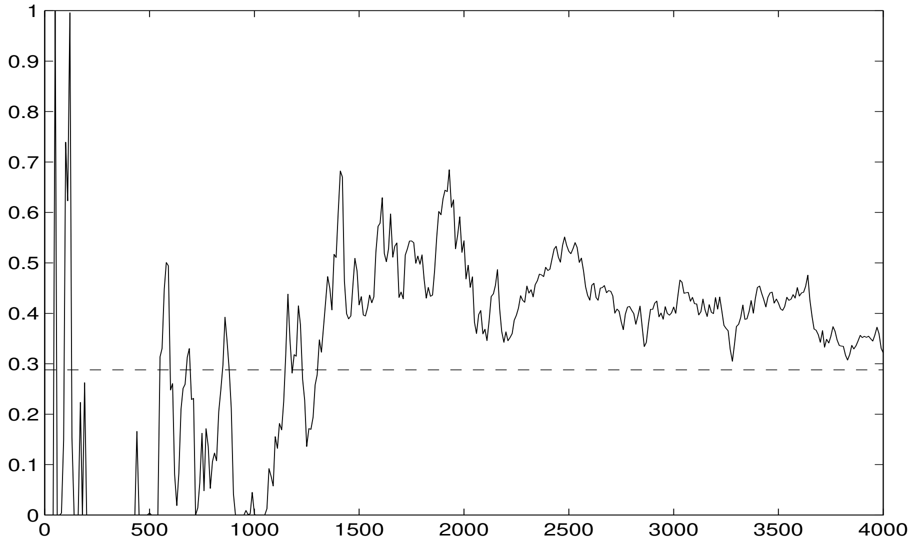

Figure 7: $u_t(11)$, the entry probability at queue size $x = 11$, as a function of time.

The learning framework suggested here seems quite general, and applicable to other similar models. In the present context, an important extension would be to the case of user-based learning. We have assumed that keeping record of the performance statistics is handled by a central entity (the server), which monitors all customers, and makes this information available to all. In certain situations it may be more appropriate to consider learning by (a finite number of) users who repeatedly use the same service facility, and each one learns out of its own personal experience. Convergence of this distributed learning scheme is currently under investigation. More complicated distributed learning scenarios which incorporate partial information sharing between users (see, e.g., Kushner and Yin (1987)) may be similarly considered and analyzed.

Appendix A

This Appendix presents the proofs of Lemmas 1 and 2. We consider customer C_0 (k=0) for concreteness. In order to compare the service times under different initial conditions and policies, we use stochastic coupling: the corresponding queueing processes are realized on a common probability space, as functions of the same primitive random variables, which enables sample-path comparisons to be made.

Let $\{T_k\}_{k\geq 0}\,, \{U_k\}_{k\geq 0}\,, \{S_\ell\}_{\ell\geq 1}\,, \{U'_\ell\}_{\ell\geq 1}$ denote independent random sequences, with the following statistics and interpretation:

- $\{T_k\}$ is the sequence of arrival times to the system, as described in section 1.
- $\{U_k\}$ and $\{U'_\ell\}$ are sequences of i.i.d. random variables, uniformly distributed on [0,1]. U_k is used for randomized decisions of customer C_k, and U'_ℓ determines which customer completes her service at time S_ℓ.

- $\{S_\ell\}$ is the sequence of 'potential service completion instants', such that the random variables $\{S_1, S_2 - S_1, S_3 - S_2, \ldots\}$ are i.i.d., distributed exponentially with expected value μ_{max}^{-1}.

Using these variables, we define the arrival and service processes at Q_{MF}.

Arrivals to Q_{MF}: At T_k, $k \geq 0$, customer C_k arrives at the system, and joins Q_{MF} with some probability p (determined by her decision rule). The actual decision is determined by U_k, as follows: C_k joins Q_{MF} if $U_k \leq p$, and otherwise joins Q_{PC}.

Service at Q_{MF}: The following description is statistically equivalent to the Processor Sharing service described in Section 1. Denote $x_\ell \triangleq X(S_\ell)$ and define $\mu(0) = 0$. At time S_ℓ, two randomized decisions are taken by the server. First, S_ℓ is an actual service completion instant with probability $\mu(x_\ell)/\mu_{max}$, and a dummy service completion otherwise. In a dummy service completion nothing happens (the queue length is unchanged). In an actual one the server chooses randomly, with equal probability, one of the x_ℓ customers in Q_{MF}, which completes her service and leaves the queue. U'_ℓ is used to implement these randomized decisions, as follows. The x_ℓ customers are assigned disjoint intervals in $[0, \mu(x_\ell)/\mu_{max}] \subset [0, 1]$, each of length $\nu(x_\ell)/\mu_{max}$ (recall that $\nu(x) = \mu(x)/x$). If U'_ℓ falls in $(\mu(x_\ell)/\mu_{max}, 1]$, then S_ℓ is a dummy service completion instant. Otherwise, the customer whose assigned interval contains U'_ℓ is the one to complete her service.

We choose the interval assignment so that customer C_0, if present, is always assigned the first interval. Thus, C_0 leaves the queue at S_ℓ if and only if $U'_\ell \leq \nu(x_\ell)/\mu_{max}$. The effect of this assignment is that service completion of C_0 at a given queue size implies the same for a smaller queue size.

For every initial state $X(0) = x$ and policy $\pi = \{u_k\}$ the above description defines the queue length process $\{X(t), t \geq 0\}$ and the service time of customer C_0. In particular, X is a left-continuous, piecewise-constant process, with $X(0) = x$, and potential jumps given by

$$X(T_k^+) = X(T_k) + 1\{U_k \leq u_k(X(T_k))\} \quad , \quad k \geq 0 \tag{A1}$$

$$X(S_\ell^+) = X(S_\ell) - 1\{X(S_\ell) > 0,\ U'_\ell \leq \mu(X(S_\ell))/\mu_{max}\} \quad , \quad \ell \geq 1 \quad . \tag{A2}$$

(We shall restrict attention to distinct arrival and service completion instants, which holds with probability 1.) Furthermore, if customer C_0 joins Q_{MF} at $T_0 = 0$ then her service time is given by:

$$W_0 = S_{\ell_0}, \quad \ell_0 \triangleq \min\{\ell \geq 1 : U'_\ell \leq \nu(X(S_\ell))/\mu_{max}\} \quad . \tag{A3}$$

Proof of Lemma 1:

Given policy π and state $x \geq 1$, it is required to prove that $\bar{V}_0(x+1, \pi) - \bar{V}_0(x, \pi) \geq \delta_x > 0$. We consider two processes on the probability space defined above. The first (the 'x-process') starts at initial state $X(0) = x$, and the second (the '$(x+1)$-process') at $x + 1$. The variables related to the $(x+1)$-process will be distinguished by a tilde accent, e.g., $\tilde{X}(0) = x + 1$. In both cases C_0 joins Q_{MF}, and all other customers follow the policy $\pi^{-0} = (u_1, u_2, \ldots)$. Let w_0 and $\tilde{w}_0$ denote the respective service times of C_0. Relations between random variables are to be interpreted here in a sample-path sense.

Let $\tau \triangleq \inf\left\{t \geq 0 : \tilde{X}(t) = X(t)\right\}$. Since $\tilde{X}(0) > X(0)$ and the queue length changes by one customer at a time, then $\tilde{X}(t) > X(t)$ for $0 \leq t \leq \tau$. Since the decision rule of each customer depends on the current state only, $\tilde{X}(t) = X(t)$ for $t > \tau$. Thus $\tilde{X}(t) \geq X(t)$ for all $t \geq 0$, and since $\nu(x) = \mu(x)/x$ is decreasing in x, it follows by (A3) that $\tilde{W}_0 \geq W_0$. Thus,

$$\bar{V}_0(x+1,\pi) - \bar{V}_0(x,\pi) = E(\tilde{W}_0 - W_0) \geq 0 \quad . \tag{A4}$$

To establish strict inequality (uniformly in π), consider for example the following event:

$$B = \{S_1 < T_1,\ U_1' \in (\nu(x+2)/\mu_{max}, \nu(x+1)/\mu_{max}]\}\,.$$

On B, in the x-process C_0 will complete her service at time S_1 while in the $(x+1)$-process she will have to wait at least until S_2, so that $\tilde{W}_0 - W_0 \geq S_2 - S_1$. Thus, $E(\tilde{W}_0 - W_0) \geq E(S_2 - S_1)\mathbf{1}\{B\} = \mu_{max}^{-1} P(B) > 0$. $\square$

Proof of Lemma 2:

(i) Let g and $\tilde{g}$ be two thresholds, with $\tilde{g} > g$. We need to show that $\bar{V}_0(x,\tilde{g}^\infty) > \bar{V}_0(x,g^\infty)$ for every x. The proof proceeds similarly to that of Lemma 1. The two coupled processes considered here correspond to $\pi = [g]^\infty$ and $\tilde{\pi} = [\tilde{g}]^\infty$ respectively, both with initial state x, and C_0 joining Q_{MF}. Since $\tilde{g} > g$, it is easily seen that $\tilde{X}(t) \geq X(t)$ for all $0 \leq t < \infty$, which (as in the proof of Lemma 1) implies that $\tilde{W}_0 \geq W_0$. Moreover, it is readily verified that $\tilde{W}_0 > W_0$ with positive probability (e.g., by noting that $\tilde{X}(S_1) > X(S_1)$ with positive probability), so that $\bar{V}_0(x,[\tilde{g}]^\infty) - \bar{V}_0(x,[g]^\infty) = E(\tilde{W}_0 - W_0) > 0$.

(ii) Fix x, and let $g < \tilde{g}$ be two thresholds with equal integer part L, so that $g = L+q$ and $\tilde{g} = L+\tilde{q}$ with $0 \leq q < \tilde{q} \leq 1$. Since, by (i), $\bar{V}_0(x,[g]^\infty)$ is increasing in g, to establish continuity it is sufficient to show that

$$\bar{V}_0(x,[\tilde{g}]^\infty) \leq \bar{V}_0(x,[g]^\infty) + (\tilde{g} - g)K_L \tag{A5}$$

where K_L is a finite constant, independent of q and $\tilde{q}$.

Assume the same coupling as in (i). Note that the threshold policies $[\tilde{g}]$ and $[g]$ differ only in the decision rule at queue size L. Thus, the service time of C_0 would be the same under $[\tilde{g}]^\infty$ and $[g]^\infty$, unless the following event B occurs: During the service period of C_0, an arriving customer (say C_k) who sees L customers in Q_{MF} (i.e. $X(T_k) = \tilde{X}(T_k) = L$) enters Q_{MF} under $[\tilde{g}]^\infty$ but not under $[g]^\infty$. In that case, the queue lengths just after T_k will be $L+1$ and L, respectively, and the expected *remaining* service time of C_0 from T_k onward will therefore be $\bar{V}_0(L,[\tilde{g}]^\infty)$ and $\bar{V}_0(L-1,[g]^\infty)$, respectively. More precisely, defining

$$B = \{\exists k \geq 1 \text{ s.t. } T_k \leq W_0,\ X(T_k) = L,\ q < U_k \leq \tilde{q}\}\ , \tag{A6}$$

we have $\tilde{W}_0 = W_0$ outside of B. Define further the stopping time k_0:

$$k_0 = \inf\{k \geq 1\ :\ T_k \leq W_0,\ X(T_k) = L,\ q < U_k \leq \tilde{q}\}\ , \tag{A7}$$

where $\inf \emptyset := \infty$. Note that $B = \{k_0 < \infty\}$, so that $\tilde{W}_0 = W_0$ on $\{k_0 = \infty\}$. Thus,

$$E(\tilde{W}_0 - W_0) = \sum_{k=1}^{\infty} E(\tilde{W}_0 - W_0 | k_0 = k) P\{k_0 = k\} . \tag{A8}$$

Now, as explained above,

$$\begin{aligned} E(\tilde{W}_0 - W_0 | k_0 = k) &= E(\tilde{W}_0 - T_k | k_0 = k) - E(W_0 - T_k | k_0 = k) \\ &= \bar{V}_0(L, [\tilde{g}]^\infty) - \bar{V}_0(L-1, [g]^\infty) \stackrel{\triangle}{=} \delta(\tilde{g}, g) > 0 . \end{aligned} \tag{A9}$$

In addition,

$$P\{k_0 = k) \le P\{T_k \le W_0,\ q \le U_k \le \tilde{q}\} = (\tilde{q} - q) P\{T_k \le W_0\} . \tag{A10}$$

Thus, we obtain from (A8) that

$$\begin{aligned} E(\tilde{W}_0 - W_0) &\le \delta(\tilde{g}, g)\,(\tilde{q} - q) \sum_{k=1}^{\infty} P\{T_k \le W_0\} \\ &= (\tilde{g} - g)\delta(\tilde{g}, g) N(g) , \end{aligned} \tag{A11}$$

where $N(g) = E \sum_{k=1}^{\infty} \mathbf{1}\{T_k \le W_0\}$ is the expected number of arrivals during the sojourn time of C_0. Finally, (A5) follows by noting that $\delta(\tilde{g}, g)$ and $N(g)$ are bounded by constants which depend only on L; i.e., $\delta(\tilde{g}, g) \le \delta(L+1, L)$ and $N(g) \le N(L+1)$, which both follow from part (i) of the lemma. $\square$

Appendix B

This appendix provides the proof for the convergence result of Theorem 2. The proof follows the outline given in Section 5.2. As mentioned there, the analysis is based on representing the dynamics of the system as a controlled Markov process. We proceed to identify this process, define more precisely the relevant ODE, and investigate its equilibrium properties and their relation to the Nash equilibrium of the system.

We first define $w^\epsilon(v, \cdot)$ more carefully. Let v be some possible value of the vector $\hat{V}_t$, and let $u^\epsilon(v)$ denote the decision rule derived from v through (19), namely,

$$u^\epsilon(v, x) = S_\epsilon(\theta^{-1} - v(x)) , \quad 0 \le x < \mathsf{B} . \tag{A12}$$

Consider the system operated with this stationary decision policies, applied by all controlled customers; this corresponds to freezing $\hat{V}_t$ in (19) at the fixed value v. Noting the finite buffer assumption and the existence of uncontrolled arrivals, it follows immediately that the queue length process in Q_{MF} is a recurrent, finite state Markov chain, with a unique stationary distribution. Then $w^\epsilon(v, x)$ denotes the expected waiting time, under this stationary distribution, of a customer which enters Q_{MF} at queue size x.

The following lemmas are valid for all $\epsilon > 0$ small enough. The first two, which are used to establish the third one, concern the properties of the simplified dynamical system:

$$\frac{d}{dt}v_t(x) = w^\epsilon(v_t, x) - v_t(x)\,, \quad 0 \le x < \mathsf{B}\,, \tag{A13}$$

or, more compactly, $\frac{d}{dt}v_t = w^\epsilon(v_t) - v_t$.

Lemma 5 *Consider the stationary points of the differential equation (A13), namely the solutions of $w^\epsilon(v) - v = 0$.*

(i) A solution exists and is unique. Denote it v^ϵ.

(ii) The corresponding decision rule $u^\epsilon(x) = S^\epsilon(\theta^{-1} - v^\epsilon(x))$, $x \ge 0$, is a threshold rule .

(iii) As $\epsilon \to 0$, the decision rule u^ϵ converges to the Nash equilibrium policy $[g^*]$.

Proof: Recall that the vector v determines $w^\epsilon(v)$ through the decision rule $u^\epsilon(v)$. We first show that at any candidate stationary point v, $u^\epsilon(v)$ is a threshold rule. In terms of the notation of Section 2, we have $w^\epsilon(v,x) = \bar{V}_0(x, \pi(v))$, where $\pi(v)$ is the stationary policy corresponding to the decision rule $u^\epsilon(v)$. It follows from Lemma 1 that $w^\epsilon(v,x)$ is increasing in x. In fact, noting the finite buffer size ($\mathsf{B} < \infty$) assumed in this section, it follows from that lemma that there exists a positive constant δ, independent of v, ϵ and x, such that $w^\epsilon(v, x+1) - w^\epsilon(v,x) > \delta$. Thus, if v satisfies $v = w^\epsilon(v)$, it must have the same monotonicity property; in particular, for ϵ small enough ($\epsilon < \delta/2$) we have $v(x+1) - v(x) > 2\epsilon$. But then it follows from $(A12)$ and the definition of S_ϵ that $u^\epsilon(v)$ is a threshold rule. Let $g^\epsilon(v)$ denote that threshold.

It follows that v is a solution of $v = w^\epsilon(v)$ if and only if there exists a threshold g such that the following two equalities hold: (a) $[g] = u^\epsilon(v)$, and (b) $v(x) = \bar{V}_0(x, [g]^\infty)$, $0 \le x < \mathsf{B}$. Existence and uniqueness of such a threshold may be established in a similar way to the proof of Theorem 1. That is, we first define a map $G^\epsilon(g)$, which assigns to every threshold g another threshold $G^\epsilon(g)$, by using (b) and (a) successively. It follows that $g \to G^\epsilon(g)$ is continuous and decreasing in g, hence has a unique fixed point, say g^ϵ. The corresponding vector v^ϵ, determined by (b), is then the unique solution of $v = w^\epsilon(v)$, and the corresponding policy is $u^\epsilon = [g^\epsilon]$.

It remains to show that g^ϵ converges to the Nash threshold g^*. This follows by noting, using the definition of S_ϵ, that $\lim_{\epsilon\to 0} G^\epsilon(g) \in G^*(g)$, where $G^*(g)$ is the point-to-set mapping defined in the proof of Theorem 1 (compare (5) with (A12)). It follows that the fixed point of $g \to G^\epsilon(g)$ (corresponding to g^ϵ) must converge to the fixed point of $g \to G^*(g)$ (which corresponds to g^*). $\square$

Lemma 6 *The stationary point v^ϵ of (A13) is globally exponentially stable; that is, $\lim_{t\to\infty} v_t = v^\epsilon$ for all initial conditions v_0, with an exponential convergence rate.*

Proof: Let $\delta > 0$ be as in the proof of the previous lemma, and again assume $\epsilon < \delta/2$. We first show that, after some finite time t_0, we have $v_t(x+1) > v_t(x) + 2\epsilon$ for all x. Define $\Delta v_t(x) = v_t(x+1) - v_t(x)$, and $\Delta w_t(x) = w^\epsilon(v_t, x+1) - w^\epsilon(v_t, x)$. From (A13), $\frac{d}{dt}\Delta v_t(x) = \Delta w_t(x) - \Delta v_t(x)$. Recalling that

$\Delta w_t(x) > \delta$ by definition of δ, this differential relation implies that $\Delta v_t(x) > 2\epsilon$ is reached within a finite time, which may be bounded by $(2\epsilon - \Delta v_0(x))/(\delta - 2\epsilon)$.

It follows that for all $t > t_0$, $u^\epsilon(v_t)$ will be a threshold rule. Let g_t denote the corresponding threshold, and assume $t > t_0$ hereafter.

As established in the previous lemma, $u^\epsilon := u^\epsilon(v^\epsilon)$ is a threshold rule with some threshold g^ϵ. Define $x^o = \min\{x \leq \mathsf{B} : v^\epsilon(x) \geq \theta^{-1} - \epsilon\}$, with $\min\{\emptyset\} := \mathsf{B}$. Since $v^\epsilon(x+1) - v^\epsilon(x) > 2\epsilon$, it may be seen that x^o alone is relevant for determining g^ϵ, which in fact equals $g^\epsilon = x^o - 1 + S_\epsilon(\theta^{-1} - v^\epsilon(x^o))$. We proceed to show that $v_t(x^o) \to v^\epsilon(x^o)$, from which we shall conclude that $g_t \to g^\epsilon$ and finally that $v_t(x) \to v^\epsilon(x)$ for all other x as well.

We first argue that the deviation of $w^\epsilon(v_t, x^o)$ from its stationary value is opposite in sign to that of $v_t(x^o)$, that is, $(v_t(x^o) - v^\epsilon(x^o))(w^\epsilon(v_t, x^o) - w^\epsilon(v^\epsilon, x^o)) \leq 0$. Indeed, if $v_t(x^o) - v^\epsilon(x^o) < 0$, then by the monotonicity of the components of v_t and the above definition of x^o, it follows that $g_t \geq g^\epsilon$, which implies that $w^\epsilon(v_t, x) - w^\epsilon(v^\epsilon, x) \geq 0$ for every x, and in particular for x^o. The opposite inequalities hold when $v_t(x^o) - v^\epsilon(x^o) > 0$, thereby establishing the required relation. This, together with the scalar differential equation $\frac{d}{dt} v_t(x^o) = w^\epsilon(v_t, x^o) - v_t(x^o)$, imply the exponential convergence of $v_t(x^o)$ to v^ϵ. Indeed, using the quadratic Lyapunov function $L_t = \frac{1}{2}(v_t(x^o) - v^\epsilon(x^o))^2$, we obtain

$$
\begin{aligned}
\frac{d}{dt} L_t &= (v_t(x^o) - v^\epsilon(x^o)) \frac{d}{dt} v_t(x^o) \\
&= (v_t(x^o) - v^\epsilon(x^o))[(w^\epsilon(v_t, x^o) - w^\epsilon(v^\epsilon, x^o)) - (v_t(x^o) - v^\epsilon(x^o))] \\
&\leq -(v_t(x^o) - v^\epsilon(x^o))^2 = -2L_t \,,
\end{aligned}
$$

where we have used the equilibrium relation $v^\epsilon = w^\epsilon(v^\epsilon)$. This implies that L_t converges to 0 at rate e^{-2t}, hence that $v_t(x^o)$ converges to $v^\epsilon(x^o)$ at rate e^{-t}.

Recalling that $g^\epsilon = x^o - 1 + S_\epsilon(\theta^{-1} - v^\epsilon(x^o))$, it follows by continuity arguments that g_t converges exponentially to g^ϵ. This, in turn, implies the exponential convergence of $w^\epsilon(v_t)$ to $v^\epsilon = w^\epsilon(v^\epsilon)$ (using continuity properties from Lemma 2. It then easily follows from (A13), considering the equation for every x independently, that v_t converges exponentially to v^ϵ. □

The following lemma introduces a modified version of the differential equation (A13), which accounts for the asynchronous updates of the components of the estimated vector.

Lemma 7 *Let $r(v, x)$ denote a Lipschitz-continuous function of v (for each x), which takes values in $[\beta, 1]$ for some $\beta > 0$. Consider the following set of differential equations, for $t \geq 0$:*

$$
\begin{aligned}
\frac{d}{dt} v_t(x) &= \frac{r(v_t, x)}{\alpha_t(x)} (w^\epsilon(v_t, x) - v_t(x)) \,, \quad x = 0 \ldots \mathsf{B} - 1 \,, \\
\frac{d}{dt} \alpha_t(x) &= r(v_t, x) - \alpha_t(x) \,; \quad \alpha_0(x) > 0 \,. \qquad \text{(A14)}
\end{aligned}
$$

Then the results of the previous two lemmas hold here as well. In particular, v^ϵ, appended by $\alpha^\epsilon = r(v^\epsilon)$, is a globally exponentially stable equilibrium point for these equations.

Proof: Since α_t remains positive by its equation, the stationary points of v_t must coincide with those of $(A13)$, and the stationary value for α follows trivially. Regarding stability, we note that after a finite time we obtain $\alpha_t(x) \in [\beta/2, 1]$, so that the ratio $r_t(x)/\alpha_t(x)$ is bounded in $[\beta, 2/\beta]$. This factor may then be viewed as a bounded time normalization of the equations for v_t, and the proof of the previous lemma applies here with minor modifications. $\square$

Lemma 8 *$\hat{V}_t$ converges (a.s.) to v^ϵ.*

Proof: Observe that $\hat{V}_t$ may change only at times $\{T_n\}$, where T_n is the time of the n-th departure from Q_{MF}. Let V_n denote the value of $\hat{V}_t$ at (just after) T_n. In order to write the update equation for V_n, define W_n – the sojourn time of the customer that left at T_n, $N_n(x)$ – the number of departures up to T_n which entered at queue length x, and $I_n(x)$ – the indicator of the event that the customer which left at T_n entered at queue size x. Obviously, $N_n = \sum_{k=1}^{n} I_n$. In this notation,

$$V_n(x) = V_{n-1}(x) + \frac{1}{N_n(x)} I_n(x)(W_n - V_{n-1}(x)) \, .$$

We would like to put that equation in the standard form of stochastic approximations, where the vanishing gain sequence is the same for all components of V_n. This may be done by introducing the auxiliary variables $\alpha_n(x) = N_n(x)/n$. Then

$$V_n(x) = V_{n-1}(x) + \frac{1}{n}\frac{I_n(x)}{\alpha_n(x)}(W_n - V_{n-1}(x)) \, , \tag{A15}$$

and

$$\alpha_n(x) = \alpha_{n-1}(x) + \frac{1}{n}(I_n(x) - \alpha_{n-1}(x)) \, . \tag{A16}$$

We proceed to establish that (V_n, α_n) converge to the stationary point of the differential equations (A14). Together with the previous lemma, this would imply the assertion of the present one. We employ the formulation and results in Benveniste, Metevier and Priouret (1982). Define first the *state* of the learning process at time T_n, denoted by ξ_n, which consists of: the number of customers in Q_{MF} at T_n; and for each of these customers as well as the customer that just left, the queue length at her arrival, and her service time in Q_{MF} so far.

We assume that decision rule (19) is in effect, with a fixed $\epsilon > 0$. It may then be seen that ξ_{n-1} and V_{n-1} determine the statistics of ξ_n, so that ξ_n has a controlled Markov dynamics. Furthermore, equations (A15) may be expressed as

$$V_n = V_{n-1} + \frac{1}{n} H(V_{n-1}, \alpha_{n-1}; \xi_n) \, ,$$

and similarly for α_n. This gives us the standard form of a stochastic approximations iteration, with (V_n, α_n) the estimated vector and ξ_n a controlled Markov state process.

Fixing $V_n \equiv v$ (and $\alpha_n \equiv \alpha$), the sequence ξ_n becomes a Markov chain. The times when Q_{MF} empties are renewal times for that sequence, and from the finite buffer assumption is easily follows that the duration of the busy period in Q_{MF} is *a.s.*-finite with bounded moments, uniformly in v. It then follows that ξ_n is recurrent, with a unique stationary distribution. Recall that $w^\epsilon(v,x)$ was defined as the expected sojourn time, under this distribution, of a customer which enters Q_{MF} at queue size x. Let $r(v,x)$ denote the expected value of $I_n(x)$ under this distribution; this is the fraction of customers who enter Q_{MF} at queue size x. Obviously $r(v,x) \leq 1$; also, because of the uncontrolled arrivals, it may be verified that $r(v,x) \geq \beta > 0$ uniformly in v. Now, the expected value of the perturbation term in (A15) becomes $\frac{r(v,x)}{\alpha(x)}(w^\epsilon(v,x) - v)$, and for the corresponding term in (A16) we obtain $(r(v,x) - \alpha(x))$. This identifies (A14) as the ODE corresponding to the stochastic approximation (A15)-(A16).

We may now employ Theorem 2.4 (see also Corollary 2.6) in Benveniste, Metevier and Priouret (1982) to deduce the convergence of (V_n, α_n) to the equilibrium point of (A14). Indeed, it is easy to verify directly that (V_n, α_n) converge (a.s.) to a bounded set (which follows, e.g., from the above-noted uniform bounded Ness of the expected busy periods), where the components of α are bounded away from zero (say, by $\beta/2$). The required convergence to equilibrium then follows from the recurrence properties noted above, and the continuity of the right hand side of the ODE. $\square$

Theorem 2 follows from the above lemmas. Indeed, the last lemma and (19) establish the convergence of $\hat{u}_t$ to u^ϵ, while the required properties of the latter follow from Lemma 5.

References

ATM Forum Technical Committee, *Traffic Management Specification*, Version 4.0, April 1996.

Altman, E. 1994. Flow Control Using the Theory of Zero-Sum Markov Games. *IEEE Trans. Automat. Contr.* **39**, No. 4, 814-818.

Altman E., and A. Hordijk. 1995. Zero-sum Markov games and worst-case optimal control of queueing systems, invited paper, *QUESTA* **21**, special issue on optimization of queueing systems, Ed. S. Stidham, 415-447.

Altman E., and G. Koole. 1993. Stochastic Scheduling Games with Markov Decision Arrival Processes, *Journal Computers and Mathematics with Appl.*, third special issue on "Differential Games", 141–148.

Assaf, D., and M. Haviv. 1990. Reneging from Processor Sharing Systems and Random Queues. *Math. Oper. Res.* **15**, 129–138.

Başar, T., and G. J. Olsder. 1995. *Dynamic Noncooperative Game Theory.* Academic Press, London, 2nd edition.

Benveniste, A., M. Metivier and P. Priouret. 1990. *Adaptive Algorithms and Stochastic Approximations.* Springer Verlag, Berlin.

BOVOPOULOS, A. D., AND A. A. LAZAR. 1987. Decentralized Algorithms for Optimal Flow Control. *Proc. of the 25th Allerton Conf. on Communications, Control and Computing,* University of Illinois at Urbana-Champaign, 979–987.

BARTROLI, M., AND S. STIDHAM. 1992. Admission to a General Stochastic Congestion System: Comparison of Individually and Socially Optimal Policies. *Queueing and Related Models,* ed. U. N. Bhat and I. V. Basawa. Oxford University Press, New York, 301–324.

DE WAAL, P. 1988. Stochastic Control of a Processor Sharing Queue with Partial State Information. *Proc. of the 27th IEEE Conf. on Decision and Control,* Austin, Texas, 1135–1136.

GLAZER, A., AND R. HASSIN. 1986. "Stable priority purchasing in queues", *Operations Research Letters*, **6**, pp. 285-288.

¿ FUDENBERG, D., AND D. K. LEVINE. ¿ 1998. ¿ *The Theory of Learning in Games.* ¿ MIT Press, Boston.

HASSIN R. AND M. HAVIV. 1994. "Equilibrium strategies and the value of information in a two line queueing system with threshold jockeying", *Commun. Statist. - Stochastic Models*, **10**, 415-435.

HAVIV. 1991. "Stable strategies for processor sharing systems", *European J. of Operations Research* **52**, 103-106.

HSIAO M. T., AND A. A. LAZAR. 1991. Optimal Decentralized Flow Control of Markovian Queueing Networks with Multiple Controllers. *Performance Evaluation* **13,** 181–204.

JAISWAL N. K. 1982. Performance Evaluation Studies for Time-Sharing Computer Systems. *Performance Evaluation* **2,** 223–236.

KULKARNI, V. G. 1983. A Game Theoretic Model for Two Types of Customers Competing for Service. *Operation Research Letters* **2,** 119–122.

KUSHNER, H. J., AND D. S. CLARK. 1978. *Stochastic Approximation Methods for Constrained and Unconstrained Systems.* Springer-Verlag, New York.

KUSHNER, H. J., AND G. YIN. 1987. Asymptotic Properties of Distributed and Communicating Stochastic Approximations Algorithms. *SIAM J. Control and Optimization* **25,** 1266–1290.

LAKSHMIVARAHAN, S. 1981. *Learning Algorithms Theory and Applications.* Springer, N.Y.

LEE, H. L., AND M. A. COHEN. 1985. Multi-Agent Customer Allocation in a Stochastic Service System. *Management Sci.* **31,** 752–763.

LI,S., AND T. BAŞAR. 1987. Distributed Algorithms for the Computation of Noncooperative Equilibria. *Automatica* **23,** 523–533.

LIPPMAN, S. A., AND S. STIDHAM. 1977. Individual Versus Social Optimization in Exponential Congestion Systems. *Operations Res.* **25,** 233–247.

NAOR, P. 1969. On the Regulation of Queue Size By Levying Tolls. *Econometrica* **37,** 15–24.

RAMASWAMI, V. 1984. The Sojourn Time in the GI/M/1 Queue with Processor Sharing. *J. Appl. Prob.* **21,** 437–442.

SCHELLING, T. C. 1960. *The Strategy of Conflict.* Harvard University Press, Cambridge, MA.

SHENKER, S. 1990. Efficient Network Allocation with Selfish Users. *Performance '90 - Proceedings of the 14th IFIP WG 7.3 International Symposium on Computer Performance Modeling,* Edinburgh, 279–285.

SHIMKIN, N., AND A. SHWARTZ. 1993. Guaranteed Performance Region in Markovian Systems with Competing Decision Makers. *IEEE Trans. Automat. Contr.* **90**, pp. 84–95.

STIDHAM, S. 1985. Optimal Control of Admission to a Queueing System. *IEEE Trans. Automat. Contr.* **30,** 705–713.

WALRAND, J. 1988. *An Introduction to Queueing Networks.* Prentice-Hall, N.J.

XU, S. H., AND J. G. SHANTIKUMAR. 1993. Optimal Expulsion Control – A Dual Approach to Admission Control of an Ordered Entry. *Oper. Res.* **41,** 1137–1152.

YECHIALI, U. 1972. Customers' Optimal Joining Rules for the GI/M/s Queue. *Management Science* **18,** 434–443.