SRC TR 87-102

A Simple Problem of Flow Control
I: Optimality Results

by

Dye-Jyun Ma, and A.M. Makowski

DRAFT - February 1987

A SIMPLE PROBLEM OF FLOW CONTROL I: OPTIMALITY RESULTS

by

Dye-Jyun Ma[1] and Armand M. Makowski[2]

Electrical Engineering Department and Systems Research Center
University of Maryland, College Park, Maryland 20742

ABSTRACT

This paper presents a problem of optimal flow control for discrete $M|M|1$ queues. The problem is cast as a constrained Markov decision process, where the throughput is maximized with a bound on the average queue size. By Lagrangian arguments, the optimal strategy is shown to be of threshold type and to saturate the constraint. The method of analysis proceeds through the discounted version of the Lagrangian problems for which the corresponding value functions are shown to be integer-concave. Dynamic Programming and stochastic comparison ideas constitute the main ingredients of the solution.

[1] The work of this author was supported partially through NSF Grant ECS-83-51836 and partially through NSF Grant NSFD CDR-85-00108.

[2] The work of this author was supported partially through ONR Grant N00014-84-K-0614 and partially through a grant from AT&T Bell Laboratories. Please use the name of this author for any correspondence concerning this manuscript.

1. Introduction

Consider a *synchronous* communication channel between two entities, a transmitter and a receiver which are both equipped with buffers of infinite capacity. Information is formatted in packets and time is slotted so that the duration of a time slot coincides with the transmission time of a packet. Packet transmissions are initiated at the beginning of a slot. The channel is assumed noisy in that a packet transmission may not be successful with probability $1-\mu$ in which case retransmission is attempted in the next slot. This scenario is repeated until successful transmission occurs, at which time the packet is deleted from the transmitter's buffer. The transmission failures are assumed independent from slot to slot, and independent of the arrival process. Packets arrive at the transmitter one at a time according to a Bernoulli sequence with rate λ, i.e., λ is the probability that a packet will arrive in any time slot.

The system described above may experience congestion and it may be desirable to take certain actions in order to guarantee an expected performance level. One possible approach consists in restricting access to the communication system, i.e., new packets which are about to enter the transmitter's buffer may be denied entrance on the basis of information reflecting system congestion. This is often referred to as *flow control* and should be done on the basis of some performance criterion [6]. Here, an approach similar to the one of Lazar [8] is adopted in that a flow control strategy is sought that maximizes the channel throughput with a constraint on the long-run average number of packets in the system.

Under the statistical assumptions given earlier, the uncontrolled system can be modelled as a discrete-time $M|M|1$ queue, and the problem of finding good flow control schemes can be cast as a Markov decision process (MDP) with *constraint.* Analysis shows that this constrained flow control problem admits a solution within the class of *threshold* policies which are parametrized by an integer-valued threshold level L $(=0,1,\cdots)$ and an acceptance probability η $(0 \leq \eta \leq 1)$. A threshold policy (L,η) has a simple structure in that at the beginning of each time slot, a new packet is accepted (resp. rejected) if the buffer content is strictly below L (resp. strictly above L), while if there are *exactly* L packets in the buffer, this new packet is accepted (resp. rejected) with probability η (resp. $1-\eta$).

The constrained MDP considered here has a *countably infinite* state space, since the problem is formulated for an open system, as opposed to the approach taken by Lazar [8] or by Beutler and Ross [3]. Therefore, the results on constrained MDP's given by Ross and co-workers [2,3,14] are not directly applicable for they were derived under the assumption that the state space is *finite*. However, the optimality result obtained here is in the same spirit as Corollary 3.5 of [3, pp.353] since the optimal threshold policy (L^*, η^*) can be naturally interpreted as a simple randomization with bias η^* between the *pure* policies $(L^*, 0)$ and $(L^*, 1)$ which are identical in all but one state, the state where L^* packets are present in the system.

As in other constrained MDP's treated in the literature, the discussion proceeds along a standard Lagrangian argument, and most of the paper deals with a thorough study of the corresponding Lagrangian problems which are *unconstrained* MDP's with a long-run average cost criterion. These auxiliary problems are analyzed through their discounted counterpart by a standard Tauberian argument. The technical contributions of this study concern the discounted problem and lie in two areas: Use of *stochastic comparison* ideas is made to show that the search of optimal discounted policies need to be performed within a much smaller subset of admissible flow control policies, thus in essence reducing the problem to a finite-state one. Moreover, the *integer-concavity* of the value function for the discounted problems is established by showing that integer-concavity *cum* growth conditions propagates under the Dynamic Programming operator.

The work presented here provides ample evidence of the usefulness of several ideas and techniques for solving certain MDP's. Moreover, the thorough analysis given in the forthcoming sections should be viewed as a necessary step towards the discussion of some aspects of *adaptive* flow control reported in the companion papers [10,11].

Throughout the years, several authors have studied problems of flow control (or control of arrivals) in the context of simple queueing systems, and a good discussion of such work can be found in the survey paper of Stidham [18]. It should be pointed out that previous papers dealt exclusively with continuous-time models, and that concavity of the value function for the single node situation could be obtained fairly easily through standard arguments. Here, establishing the concavity of the value functions of interest turns out to be a much more cumbersome task; this can possibly be explained by the fact that multiple transitions can be realized, a phenomenon which precludes use of the homogeneization technique for the discrete-time situation [7].

Amongst the models covered in Stidham's survey paper, only the work of Lazar [8] formulates the problem as a *constrained* problem. However, the approach taken here is different from the one used by Lazar in that he considers a *closed* system from the onset with a *fixed* number of packets, while the work discussed here assumes an open system. Of course, both approaches lead to similar results, as expected.

The paper is organized as follows. The model is described in Section 2 and the constrained flow control problem is posed in Section 3, where the optimality results are summarized and the necessary Lagrangian are briefly outlined. Section 4 is devoted to the study of the discounted version of the Lagrangian problems, for which threshold policies are identified to be optimal. Their properties are discussed in Section 5, and used in Section 6 to find the solution to the long-run version of the Lagrangian problems. A useful comparison result is given in Appendix I, while the propagation of integer-concavity in the backward induction of Dynamic Programming is studied in Appendix II.

A word on the notation: The set of real numbers is denoted by $I\!R$, while $I\!N$ denotes the set of all non-negative integers. For any x in $I\!R$, it is convenient to pose $\bar{x} = 1 - x$. The Kronecker delta $\delta(\bullet,\bullet)$ is defined as usual by $\delta(a,b) = 1$ if $a = b$ and $\delta(a,b) = 0$ otherwise. The characteristic function of any set E is denoted simply by $1[E]$.

2. Model

In order to formally define a flow control model for discrete-time $M|M|1$ systems, start with the sample space $\Omega := I\!N \times \left(\{0,1\}^3\right)^\infty$ which acts as the canonical space for the Markov decision problem under consideration. The information spaces $\{I\!H_n\}_0^\infty$ are recursively generated by $I\!H_0 := I\!N$ and $I\!H_{n+1} := I\!H_n \times \{0,1\}^3$ for all $n = 0, 1, \cdots$, and with a slight abuse of notation, Ω is naturally identified with $I\!H_\infty$.

An element ω of Ω is viewed as a sequence $(x, \omega_0, \omega_1, \cdots)$ with x in $I\!N$ and ω_n in $\{0,1\}^3$ for all $n = 0, 1, \cdots$. Each block component ω_n is written in the form (u_n, a_n, b_n), with u_n, a_n and b_n being all elements in $\{0,1\}$. An element h_n in $I\!H_n$ is uniquely associated with the sample ω by $h_n := (x, \omega_0, \cdots, \omega_{n-1})$ with $h_0 := x$.

Let the sample $\omega = (x, \omega_0, \omega_1, \cdots)$ be realized. The initial queue size is set at x. During each time slot $[n, n+1)$, $a_n = 1$ (resp. a_n=0) indicates that a customer (resp. no customer) has arrived into the queue, $b_n = 1$ (resp. b_n=0) encodes a successful (resp. unsuccessful) completion of service in that slot, whereas control action u_n is selected at the beginning of the time slot $[n, n+1)$, with $u_n=1$ (resp. $u_n=0$) for admitting (resp. rejecting) the incoming customer during that slot. If x_n denotes the queue size at the beginning of the slot $[n, n+1)$, its successive values are determined through the recursion

$$x_{n+1} = x_n + u_n a_n - 1[x_n \neq 0] b_n \qquad n = 0, 1, \cdots$$

with $x_0 := x$.

The coordinate mappings $\Xi, \{U(n)\}_0^\infty, \{A(n)\}_0^\infty$ and $\{B(n)\}_0^\infty$ are defined on the sample space Ω by posing

$$\Xi(\omega) := x, \;\; U(n, \omega) := u_n, \;\; A(n, \omega) := a_n \text{ and } B(n, \omega) := b_n \qquad n = 0, 1, \cdots \quad (2.1)$$

for every ω in Ω, with the information mappings $\{H(n)\}_0^\infty$ given by

$$H(n, \omega) := (x, \omega_0, \omega_1, \cdots, \omega_{n-1}) := h_n. \qquad n = 0, 1, \cdots \quad (2.2)$$

For each $n = 0, 1, \cdots$, let $\mathbb{F}_n$ be the σ-field generated by the mapping $H(n)$ on the sample space Ω. Clearly, $\mathbb{F}_n \subset \mathbb{F}_{n+1}$, and with standard notation, $\mathbb{F} := \vee_{n=0}^\infty \mathbb{F}_n$ is simply the natural σ-field on the infinite cartesian product $\mathbb{H}_\infty$ generated by the mappings Ξ and $\{U(n), A(n), B(n)\}_0^\infty$. Thus, on the space $(\Omega, \mathbb{F})$, the mappings $\Xi, \{U(n)\}_0^\infty$, $\{A(n)\}_0^\infty$, $\{B(n)\}_0^\infty$ and $\{H(n)\}_0^\infty$ are all random variables (RV) taking values in $\mathbb{N}$, $\{0, 1\}$, $\{0, 1\}$, $\{0, 1\}$ and $\mathbb{H}_n$, respectively. The queue sizes $\{X(n)\}_0^\infty$ are $\mathbb{N}$-valued RV's which are defined recursively by

$$X(n+1) = X(n) + U(n)A(n) - 1[X(n) \neq 0]B(n) \qquad n = 0, 1, \cdots \quad (2.3)$$

with $X(0) := \Xi$. Each RV $X(n)$ is clearly $\mathbb{F}_n$-measurable.

Since randomization is allowed, an admissible policy π is defined as any collection $\{\pi_n\}_0^\infty$ of mappings $\pi_n : \mathbb{H}_n \to [0, 1]$, with the interpretation that the potential arrival during the slot $[n, n+1)$ is admitted (resp. rejected) with probability $\pi_n(h_n)$ (resp. $1 - \pi_n(h_n)$) whenever the information h_n is available to the decision-maker. In the sequel, denote the collection of all such admissible policies by $\mathcal{P}$.

Let $q(\bullet)$ be a probability distribution on $\mathbb{N}$, and let λ and μ be fixed constants in $(0, 1)$. Given any policy π in $\mathcal{P}$, there exists an unique probability measure P^π on $\mathbb{F}$, with corresponding expectation operator E^π, satisfying the requirements (R1)-(R3), where

(R1): For all x in $\mathbb{N}$,

$$P^\pi[\Xi = x] := q(x),$$

(R2): For all a and b in $\{0,1\}$,

$$\begin{aligned} P^\pi[A(n) = a, B(n) = b | \mathbb{F}_n \vee \sigma\{U(n)\}] &:= P^\pi[A(n) = a] P^\pi[B(n) = b] \\ &:= (a\lambda + \bar{a}\bar{\lambda})(b\mu + \bar{b}\bar{\mu}) \end{aligned} \qquad n = 0, 1, \cdots$$

(R3):

$$P^\pi[U(n) = 1 | \mathbb{F}_n] := P^\pi[U(n) = 1 | H(n)] := \pi_n(H(n)). \qquad n = 0, 1, \cdots$$

This notation is specialized to P^π_x and E^π_x, respectively, when $q(\bullet)$ is the point mass distribution at x in $\mathbb{N}$; it is plain that $P^\pi[A | X(0) = x] = P^\pi_x[A]$ for every A in $\mathbb{F}$.

It readily follows from (R1)-(R3) that under each probability measure P^π,

(P1): The $\mathbb{N}$-valued RV Ξ is independent of the sequences of RV's $\{A(n)\}_0^\infty$ and $\{B(n)\}_0^\infty$,

(P2): The sequences $\{A(n)\}_0^\infty$ and $\{B(n)\}_0^\infty$ of $\{0,1\}$-valued RV's are mutually independent Bernoulli sequences with parameters λ and μ, respectively, and

(P3): The transition probabilities take the form

$$P^\pi[X(n+1) = y | \mathbb{F}_n] = p[X(n), y; \pi_n(H(n))] \qquad n = 0, 1, \cdots \quad (2.4)$$

where

$$p[x, y; \eta] := \eta Q^1(x, y) + \bar{\eta} Q^0(x, y) \quad (2.5)$$

with

$$Q^i(x, y) := P^\pi[x + iA(n) - 1(x \neq 0)B(n) = y], \quad i = 0, 1 \quad (2.6)$$

for all x and y in $\mathbb{N}$, and all η in [0,1].

The right-hand sides of (2.6) depend neither on n nor on the policy π owing to the assumptions (R1)-(R3) made earlier. It is assumed throughout this paper that for every π in $\mathcal{P}$,

$$E^\pi[\Xi] = \sum_{x=0}^\infty x q(x) < \infty. \quad (2.7)$$

Several subclasses of policies in $\mathcal{P}$ will be of interest in the sequel.

A policy π in $\mathcal{P}$ is said to be a *Markov* policy if there exists a family $\{g_n\}_0^\infty$ of mappings $g_n: \mathbb{N} \to [0,1]$ such that

$$\pi_n(H(n)) = g_n(X(n)) \quad P^\pi - a.s. \qquad n = 0, 1, \cdots \quad (2.8)$$

In the event $g_n = g$ for all $n = 0, 1, \cdots$, the Markov policy π is called *stationary* and can be identified with the mapping g itself.

A policy π in $\mathcal{P}$ is said to be a *pure* (or *non-randomized*) policy if there exists a family $\{f_n\}_0^\infty$ of mappings $f_n: \mathbb{H}_n \to \{0,1\}$ such that

$$\pi_n(H(n)) = \delta(1, f_n(H(n))) \quad P^\pi - a.s. \qquad n = 0, 1, \cdots \quad (2.9)$$

A *pure Markov stationary* policy π is thus fully characterized by a single mapping $f: \mathbb{N} \to \{0,1\}$.

A stationary policy g is said to be of *threshold* type if there exists a pair (L, η), with L an integer in $\mathbb{N}$ and η in $[0,1]$, such that

$$g(x) = \begin{cases} 1 & \text{if } x < L; \\ \eta & \text{if } x = L; \\ 0 & \text{if } x > L. \end{cases} \qquad (2.10)$$

Such a *threshold* policy is denoted by (L, η), and note that $(L, 1) \equiv (L+1, 0)$. For convenience, the Markov stationary policy that admits every single customer, i.e., $g(x) = 1$ for all x in $\mathbb{N}$, is simply denoted by $(\infty, 1)$.

3. The optimal control problems

For any admissible policy π in $\mathcal{P}$, pose

$$T(\pi) := \liminf_{n \uparrow \infty} \frac{1}{n+1} E^\pi \sum_{t=0}^{n} \mu 1[X(t) \neq 0] \qquad (3.1)$$

and

$$N(\pi) := \limsup_{n \uparrow \infty} \frac{1}{n+1} E^\pi \sum_{t=0}^{n} X(t). \qquad (3.2)$$

These quantities $T(\pi)$ and $N(\pi)$ are readily interpreted as the *throughput* and the long-run average *queue size*, respectively, when the policy π is used.

Given $V > 0$, denote by $\mathcal{P}_V$ the collection of all admissible policies π in $\mathcal{P}$ which satisfy the constraint

$$N(\pi) \le V. \tag{3.3}$$

The problem discussed in this paper is the following *constrained* optimization problems (P_V), where

$$(P_V): \quad \text{maximize } T(\pi) \text{ over } \mathcal{P}_V.$$

If the constraint (3.3) is satisfied when admitting every single customer, then $\mathcal{P}_V = \mathcal{P}$. In that case, the constrained optimization problem (P_V) reduces to an unconstrained optimization problem and has a *trivial* solution as shown below.

Theorem 3.1 *If $N((\infty, 1)) \le V$, then $\mathcal{P}_V = \mathcal{P}$ and the policy $(\infty, 1)$ solves problem (P_V).*

Proof: For any policy π in $\mathcal{P}$, the relations $\pi_n(H(n)) \le 1 = (\infty, 1)(X(n))$ hold for $n = 0, 1, \cdots$, and therefore $(\{X(t)\}_0^\infty, P^\pi) \le_{st} (\{X(t)\}_0^\infty, P^{(\infty,1)})$ by Theorem I.3 of Appendix I. Use of $(I.2)$ now shows that $N(\pi) \le N((\infty, 1)) \le V$ and $T(\pi) \le T((\infty, 1))$, and the optimality of the policy $(\infty, 1)$ follows. $\square$

If $N((\infty, 1)) > V$, the solution to the constrained problem (P_V) is no longer trivial and it is the main objective of this paper to identify its structure. The main result is summarized in

Theorem 3.2 *If $N((\infty, 1)) > V$, then there exists a threshold policy (L^*, η^*) which solves problem (P_V) with $N((L^*, \eta^*)) = V$.*

The proof of Theorem 3.2 is given in Section 6. The solution method for these constrained optimization problems uses Lagrangian arguments similar to the ones given in [2,3,12,13]. The appropriate Lagrangian functional is defined for any admissible policy π in $\mathcal{P}$ to be

$$J^\gamma(\pi) := \liminf_{n \uparrow \infty} \frac{1}{n+1} E^\pi \sum_{t=0}^{n} \mu 1[X(t) \ne 0] - \gamma X(t) \tag{3.4}$$

with $\gamma > 0$ denoting the Lagrange multiplier. The corresponding *Lagrangian* problem (LP^γ) is then the *unconstrained* problem

$$(LP^\gamma): \quad \text{maximize } J^\gamma(\pi) \text{ over } \mathcal{P}.$$

Under the properties (P1)-(P3), each unconstrained problem $(LP^\gamma), \gamma > 0$, can be viewed as a Markov decision problem under the long-run average cost criterion, with state process

$\{X(n)\}_0^\infty$, cost per stage $c^\gamma : \mathbb{N} \to \mathbb{R}$ given by

$$c^\gamma(x) := \mu 1[x \neq 0] - \gamma x \tag{3.5}$$

for all x in $\mathbb{N}$, and information pattern $\{H(n)\}_0^\infty$. This information pattern $H(n)$ is richer than the state feedback information pattern $\{\Xi, U(k), X(k+1), 0 \leq k < n\}$, used in the standard formulation of Markov decision processes [15]. The richer information pattern incorporates the RV's $\{(A(k), B(k))\}_0^{n-1}$ into the state feedback information pattern [7, Chap. 4].

The following result indicates in what sense the Lagrangian problems (LP^γ), $\gamma > 0$, are useful for solving the constrained problem (P_V).

Theorem 3.3 *Any policy* π^* in $\mathcal{P}$ *which*

(C1): *yields the expressions* $T(\pi^*)$ *and* $N(\pi^*)$ *as limits,*

(C2): *meets the constraint with* $N(\pi^*) = V$, *and*

(C3): *solves the unconstrained problem* (LP^γ) *for some value* $\gamma = \gamma(V) > 0$,

necessarily solves the constrained problem (P_V).

Its proof is elementary and is omitted for sake of brevity. Details are available in [12].

4. The discounted problems

Solving problem (P_V) reduces to the search of a policy in $\mathcal{P}$ satisfying conditions (C1)-(C3) of Theorem 3.3. Since this involves the solutions to the long-run average problems (LP^γ), $\gamma > 0$, it is natural to investigate the corresponding discounted problems, for they often provide the key to solving the long-run average cost problems. Let $\gamma > 0$ and $0 < \beta < 1$ held fixed throughout this section. The expected β-*discounted* Lagrangian cost $J_\beta^\gamma(\pi)$ associated with an admissible policy π in $\mathcal{P}$ is then defined by

$$J_\beta^\gamma(\pi) := E^\pi \sum_{t=0}^{\infty} \beta^t c^\gamma(X(t)), \tag{4.1}$$

and the corresponding discounted optimization problem (LP_β^γ) is simply

$$(LP_\beta^\gamma): \quad \text{maximize } J_\beta^\gamma(\pi) \text{ over } \mathcal{P}.$$

Since at most one arrival can be admitted in each time slot, the pathwise bound $X(n) \leq \Xi + n$ holds for all $n = 0, 1, \cdots$ and yields the estimate

$$|J_\beta^\gamma(\pi)| \leq \frac{\mu + \gamma E^\pi[\Xi]}{1-\beta} + \frac{\gamma\beta}{(1-\beta)^2} < \infty. \tag{4.2}$$

by elementary calculations. The bound (4.2) is independent of the policy π in $\mathcal{P}$, and the quantity $J_\beta^\gamma(\pi)$ is thus well-defined and *uniformly* bounded over $\mathcal{P}$.

As customary with the Dynamic Programming methodology, the *β-discounted cost-to-go* associated with any policy π in $\mathcal{P}$ is the mapping $J_\beta^{\gamma,\pi}: \mathbb{N} \to \mathbb{R}$ defined by

$$J_\beta^{\gamma,\pi}(x) := E_x^\pi[\sum_{t=0}^{\infty} \beta^t c^\gamma(X(t))] \tag{4.3}$$

for all x in $\mathbb{N}$, while the corresponding *value function* $V_\beta^\gamma: \mathbb{N} \to \mathbb{R}$ is given by

$$V_\beta^\gamma(x) := \sup_{\pi \in \mathcal{P}} J_\beta^{\gamma,\pi}(x).$$

Let the RV's A and B be generic elements in $\{A(n)\}_0^\infty$ and $\{B(n)\}_0^\infty$, respectively, and for all x in $\mathbb{N}$, define the $\mathbb{N}$-valued RV's $A^0(x)$ and $A^1(x)$ by

$$A^i(x) = x + iA - 1[x \neq 0]B, \quad i = 0, 1. \tag{4.4}$$

For any mapping $f: \mathbb{N} \to \mathbb{R}$, define the mapping $T_\beta^\gamma f: \mathbb{N} \to \mathbb{R}$ by

$$(T_\beta^\gamma f)(x) = c^\gamma(x) + \beta \max_{0 \le \eta \le 1} \{\eta E[f(A^1(x))] + \bar{\eta} E[f(A^0(x))]\} \tag{4.5}$$

for all x in $\mathbb{N}$. Here, for each $i = 0, 1$, $E[f(A^i(x))] := E^\pi[f(A^i(x))]$ for all π in $\mathcal{P}$ owing to (2.6), with

$$E[f(A^1(x))] = \begin{cases} \lambda f(1) + \bar{\lambda} f(0) & \text{if} \quad x = 0; \\ \bar{\lambda}\mu f(x-1) + (\lambda\mu + \bar{\lambda}\bar{\mu}) f(x) + \lambda\bar{\mu} f(x+1) & \text{if} \quad x \ge 1 \end{cases} \tag{4.6a}$$

and

$$E[f(A^0(x))] = \begin{cases} f(0) & \text{if} \quad x = 0; \\ \mu f(x-1) + \bar{\mu} f(x) & \text{if} \quad x \ge 1. \end{cases} \tag{4.6b}$$

For future reference, for any mapping $f: \mathbb{N} \to \mathbb{R}$, pose

$$\triangledown f(x) = \begin{cases} f(1) - f(0) & \text{if} \quad x = 0; \\ \mu(f(x) - f(x-1)) + \bar{\mu}(f(x+1) - f(x)) & \text{if} \quad x \ge 1 \end{cases} \tag{4.7}$$

and observe that

$$E[f(A^1(x))] - E[f(A^0(x))] = \lambda \triangledown f(x) \tag{4.8}$$

for all x in $\mathbb{N}$.

The backward induction of Dynamic Programming produces the sequence $\{V_\beta^n\}_0^\infty$ of mappings $V_\beta^n: \mathbb{N} \to \mathbb{R}$ through the recursion

$$V_\beta^{n+1} = T_\beta^\gamma V_\beta^n \qquad n = 0, 1, \cdots \tag{4.9}$$

with $V_\beta^0 := c^\gamma$. The cost c^γ being bounded above by $\max(0, \mu - \gamma)$, the discounted problem (LP_β^γ) is essentially covered by Assumption P of Bertsekas [1, pp. 251], and the following theorem is now readily obtained by specializing results from Section 6.4 of Bertsekas [1].

Theorem 4.1 *The value function V_β^γ satisfies the Dynamic Programming equation*

$$V_\beta^\gamma = T_\beta^\gamma V_\beta^\gamma \tag{4.10}$$

and is obtained as the pointwise limit

$$\lim_{n\uparrow\infty} V_\beta^n(x) = V_\beta^\gamma(x) \tag{4.11}$$

for all x in $\mathbb{N}$. Moreover, the Markov stationary policy g^* in $\mathcal{P}$ defined by the mapping $g^*: \mathbb{N} \to [0,1]$ is optimal for problem (LP_β^γ) if

$$g^*(x) = \begin{cases} 1 & \text{if } \nabla V_\beta^\gamma(x) > 0; \\ \text{arbitrary in } [0,1] & \text{if } \nabla V_\beta^\gamma(x) = 0; \\ 0 & \text{if } \nabla V_\beta^\gamma(x) < 0 \end{cases} \tag{4.12}$$

for every x in $\mathbb{N}$.

The value iteration method implicit in Theorem 4.1 constitutes a powerful tool to further characterize the structure of the optimal policy. Lemma 4.3 below already sheds some light on the form of the optimal policy and leads to some interesting consequences for problem (LP_β^γ).

Choose L^γ in $\mathbb{N}$ such that $\mu - \gamma L^\gamma > 0$ and $\mu - \gamma(L^\gamma + 1) \le 0$, i.e., $L^\gamma := \max\{l \in \mathbb{N}: \mu - \gamma l > 0\}$. The quantity L^γ is clearly finite and induces an obvious partition of the state space $\mathbb{N}$ with $c^\gamma(x) \ge 0$ for $0 \le x \le L^\gamma$ and $c^\gamma(x) \le 0$ for $L^\gamma < x$. The following fact will be useful in what follows.

Lemma 4.2 *For every $\gamma > 0$, if $\mu - \gamma \le 0$,* i.e., *$L^\gamma = 0$, then $V_\beta^\gamma(L^\gamma) = 0$, while if $\mu - \gamma > 0$,* i.e., *$L^\gamma > 0$, then $V_\beta^\gamma(L^\gamma) > 0$.*

Proof: For $L^\gamma = 0$, c^γ being non-positive, the relation $J_\beta^{\gamma,\pi}(L^\gamma) \leq 0$ for all π in $\mathcal{P}$ implies $V_\beta^\gamma(L^\gamma) \leq 0$, while it is plain that $J_\beta^{\gamma,(0,0)}(L^\gamma) = 0$. For $L^\gamma > 0$, direct inspection shows that $J_\beta^{\gamma,(0,0)}(L^\gamma) \geq c^\gamma(L^\gamma) > 0$ by the definition of L^γ, and the result follows. $\square$

Lemma 4.3 *For the discounted problem (LP_β^γ), an optimal Markov stationary policy g^* can always be chosen so that $g^*(x) = 0$ for all $x > L^\gamma$.*

Proof: Define the $\mathbb{F}_n$-stopping time τ by

$$\tau := \begin{cases} \inf\{\ k \geq 0 : \ X(k) = L^\gamma\} & \text{if the set is non-empty;} \\ \infty & \text{otherwise} \end{cases} \tag{4.13}$$

with the obvious interpretation that τ is the first passage time into the state L^γ. For any admissible policy π in $\mathcal{P}$, pose $B_\beta^\pi(x) := E_x^\pi[\beta^\tau]$ and

$$I_\beta^{\gamma,\pi}(x) := E_x^\pi[\sum_{t=0}^{\tau-1} \beta^t c^\gamma(X(t))] = E_x^\pi[\sum_{t=0}^{\infty} \beta^t 1[\tau > t] c^\gamma(X(t))]$$

for all x in $\mathbb{N}$. Since an optimal stationary policy exists by Theorem 4.1, it readily follows from the Markov property that the value function V_β^γ satisfies the relation

$$V_\beta^\gamma(x) = \max_g \{I_\beta^{\gamma,g}(x) + V_\beta^\gamma(L^\gamma) B_\beta^g(x)\} \tag{4.14}$$

for all x in $\mathbb{N}$, with the maximization being taken over all *stationary* policies g in $\mathcal{P}$.

Take any arbitrary stationary policy g in $\mathcal{P}$ and construct from it a new policy $\tilde{g}$ which generates actions according to

$$\tilde{g}(x) := \begin{cases} g(x) & \text{if } \ 0 \leq x \leq L^\gamma \ ; \\ 0 & \text{if } \ x > L^\gamma. \end{cases} \tag{4.15}$$

Since $V_\beta^\gamma(L^\gamma) \geq 0$ by Lemma 4.2, Lemma 4.3 will now be established by showing that $I_\beta^{\gamma,g}(x) \leq I_\beta^{\gamma,\tilde{g}}(x)$ and $B_\beta^g(x) \leq B_\beta^{\tilde{g}}(x)$ for all x in $\mathbb{N}$.

By Theorem I.3 of Appendix I, the very form of $\tilde{g}$ implies the ordering

$$(\{X(t)\}_0^\infty, P_x^{\tilde{g}}) \leq_{st} (\{X(t)\}_0^\infty, P_x^g) \tag{4.16}$$

for all x in $\mathbb{N}$. If $0 \leq x \leq L^\gamma$, the probability measures P_x^g and $P_x^{\tilde{g}}$ coincide on the σ-field $\mathbb{F}_\tau$ and obviously $I_\beta^{\gamma,g}(x) = I_\beta^{\gamma,\tilde{g}}(x)$ and $B_\beta^g(x) = B_\beta^{\tilde{g}}(x)$. For $x > L^\gamma$, the reader will readily check that almost surely

$$1[\tau > t] = 1[X(s) > L^\gamma, 1 \leq s \leq t] = f(X(1), \cdots, X(t)) \qquad t = 1, 2, \cdots \quad (4.17)$$

and

$$\begin{aligned} 1[\tau > t]c^\gamma(X(t)) &= -1[X(s) > L^\gamma, 1 \leq s \leq t][\gamma X(t) - \mu]^+ \\ &= -h(X(1), \cdots, X(t)) \end{aligned} \qquad t = 1, 2, \cdots \quad (4.18)$$

under both probability measures P_x^g and $P_x^{\tilde{g}}$, where the mappings $f, h \colon \mathbb{N}^t \to \mathbb{R}$ are monotone *non-decreasing*. It is now immediate from (4.16) that $(\tau, P_x^{\tilde{g}}) \leq_{st} (\tau, P_x^g)$, or equivalently,

$$(\beta^\tau, P_x^g) \leq_{st} (\beta^\tau, P_x^{\tilde{g}}) \quad (4.19)$$

since $0 < \beta < 1$, and that

$$(1[\tau > t]c^\gamma(X(t)), P_x^g) \leq_{st} (1[\tau > t]c^\gamma(X(t)), P_x^{\tilde{g}}). \qquad t = 0, 1, \cdots \quad (4.20)$$

The inequalities $B_\beta^g(x) \leq B_\beta^{\tilde{g}}(x)$ and $I_\beta^{\gamma,g}(x) \leq I_\beta^{\gamma,\tilde{g}}(x)$ are now readily obtained from (4.19)-(4.20). □

The threshold value L^γ given in Lemma 4.3 is *independent* of the policy π and of the discount factor β. This simple property already leads to a series of interesting facts for the discounted problem.

For the case $L^\gamma = 0$, the optimal policy for problem (LP_β^γ) is easily specified.

Theorem 4.4 *Assume $\mu - \gamma \leq 0$, i.e., $L^\gamma = 0$ and $c^\gamma(x) \leq 0$ for all x in $\mathbb{N}$. The Markov stationary policy* $(0, 0)$ *is optimal.*

Proof: As pointed out in the proof of Lemma 4.2, $V_\beta^\gamma(0) = J_\beta^{\gamma,(0,0)}(0) = 0$, and Lemma 4.3 yields the result. □

Thus, only the case $L^\gamma > 0$, or equivalently $\mu - \gamma > 0$, needs to be considered. By virtue of Lemma 4.3, the Dynamic Programming equation (4.10) reduces to

$$V_\beta^\gamma(x) = c^\gamma(x) + \beta E[V_\beta^\gamma(A^0(x))] \quad (4.21)$$

for all $x > L^\gamma$, and therefore,

$$V_\beta^\gamma(x+1) = c^\gamma(x+1) + \beta(\mu V_\beta^\gamma(x) + \bar{\mu} V_\beta^\gamma(x+1)) \tag{4.22}$$

for all $x \geq L^\gamma$ upon using (4.6b). The following property is immediate.

Lemma 4.5 *Assume* $\mu - \gamma > 0$, i.e., $L^\gamma > 0$. *The value function* V_β^γ *satisfies the inequality*

$$V_\beta^\gamma(L^\gamma + 1) - V_\beta^\gamma(L^\gamma) < 0. \tag{4.23}$$

Proof: Substract $V_\beta^\gamma(L^\gamma)$ from both sides of (4.22) evaluated at $x = L^\gamma$. Easy algebraic manipulations give the result via the fact $V_\beta^\gamma(L^\gamma) > 0$ from Lemma 4.2. $\square$

The value iteration method of Theorem 4.1 is now used to establish the ***integer-concavity*** of the value function V_β^γ by showing the concavity of each one of the iterates $\{V_\beta^n\}_0^\infty$ given by (4.9). However, this is a non-trivial task as several situations need to be discussed separately. The difficulty seems to stem from the fact that multiple transitions are possible here owing to the discrete nature of time in this system. This is in contrast with the continuous-time version of this problem for which concavity of the value function is more readily obtained through some of the arguments of [18].

The next result shows in what sense integer-concavity is preserved at each step of the backward induction of Dynamic Programming. It will be convenient to say that a mapping $f : \mathbb{N} \to \mathbb{R}$ satisfies the property (Ai), $i = 1, \cdots, 4$, if

(A1): f is integer-concave with $0 \leq f(1) - f(0) \leq \mu - \gamma$,

(A2): f is integer-concave with $\mu - \gamma \leq f(1) - f(0) \leq 1$,

(A3): f is integer-concave with $f(2) - f(1) \leq -\gamma$,

(A4): f is integer-concave with $f(2) - f(1) \geq -\gamma$.

Theorem 4.6 *Assume that* $\mu - \gamma > 0$. (i) *Suppose* $\lambda + \mu \leq 1$. *If* f *satisfies* (A2), *so does* $T_\beta^\gamma f$. (ii) *Suppose* $\lambda + \mu > 1$ *and* $\mu^2 < \gamma$. *If* f *satisfies* (A1) *and* (A3), *so does* $T_\beta^\gamma f$. (iii) *Suppose* $\lambda + \mu > 1$ *and* $\frac{\gamma}{\mu} \leq \bar{\lambda} \leq \mu$. *If* f *satisfies* (A2) *and* (A4), *so does* $T_\beta^\gamma f$. (iv) *Suppose* $\lambda + \mu > 1$ *and* $\bar{\lambda} < \frac{\gamma}{\mu} \leq \mu$. *If* f *satisfies* (A1) *with* $\nabla f(1) < 0$, *then* $T_\beta^\gamma f$ *satisfies* (A1) *and* (A3). *If* f *satisfies* (A1) *and* (A3) *with* $\nabla f(1) \geq 0$, *then* $T_\beta^\gamma f$ *satisfies* (A1) *and* (A4). *If* f *satisfies either* (A1) *and* (A4) *with* $\nabla f(1) \geq 0$, *or* (A2) *and* (A4) (*whence* $\nabla f(1) \geq 0$ *necessarily*), *then* $T_\beta^\gamma f$ *satisfies either* (A1) *and* (A4), *or* (A2) *and* (A4).

A complete discussion of this key result can be found in Appendix II. It is noteworthy that concavity alone does *not* propagate through the induction and that additional *growth* conditions are needed. Unfortunately there does not seem to be a natural interpretation for these conditions.

Theorem 4.7 *Assume* $\mu - \gamma > 0$, i.e., $L^\gamma > 0$. *The value function* V_β^γ *is integer-concave, with*

$$\frac{\beta\lambda}{1-\beta\bar{\lambda}} V_\beta^\gamma(1) = V_\beta^\gamma(0) < V_\beta^\gamma(1) \quad \text{and} \quad V_\beta^\gamma(L^\gamma + 1) < V_\beta^\gamma(L^\gamma). \tag{4.24}$$

Proof: The argument is standard and inductively uses Theorem 4.6 on the successive iterates $\{V_\beta^n\}_0^\infty$. This is made possible by observing that the 0th iterate V_β^0 is the concave mapping c^γ which satisfies $c^\gamma(1) - c^\gamma(0) = \mu - \gamma$ and $c^\gamma(2) - c^\gamma(1) = -\gamma$. The reader will now check that each one of the four situations discussed in Theorem 4.6 applies to yield the integer-concavity of V_β^n with $0 \leq V_\beta^n(1) - V_\beta^n(0) \leq 1$ for all $n = 1, 2, \cdots$.

In the limit, by Theorem 4.1, the value function V_β^γ is thus integer-concave with $0 \leq V_\beta^\gamma(1) - V_\beta^\gamma(0) \leq 1$. Consequently, $\nabla V_\beta^\gamma(0) \geq 0$ and $g^*(0) = 1$ is an optimal action by (4.12), whence $V_\beta^\gamma(0) = \beta[\lambda V_\beta^\gamma(1) + \bar{\lambda} V_\beta^\gamma(0)]$ by virtue of the Dynamic Programming equation (4.10). The first part of (4.24) now follows via the fact that $V_\beta^\gamma(1) \geq J_\beta^{\gamma,(0,0)}(1) \geq \mu - \gamma > 0$, while the second part is nothing but (4.23). □

Theorem 4.8 *Assume* $\mu - \gamma > 0$, i.e., $L^\gamma > 0$. *For every* $0 < \beta < 1$, *there exists a threshold policy* $(L_\beta^\gamma, 1)$ *which solves problem* (LP_β^γ), *where the optimal threshold value* L_β^γ *satisfies the relation* $0 \leq L_\beta^\gamma \leq L^\gamma$.

Proof: Integer-concavity of V_β^γ implies that ∇V_β^γ is monotone decreasing, whence the quantity ∇V_β^γ changes sign at most once, from positive to negative since $\nabla V_\beta^\gamma(0) > 0$ and $\nabla V_\beta^\gamma(L^\gamma + 1) < 0$ by Theorem 4.7. Consequently, there exists a level L_β^γ, with $0 \leq L_\beta^\gamma \leq L^\gamma$, such that $\nabla V_\beta^\gamma(x) \geq 0$ for $0 \leq x \leq L_\beta^\gamma$ and $\nabla V_\beta^\gamma(x) < 0$ for all $x > L_\beta^\gamma$. The threshold policy $(L_\beta^\gamma, 1)$ is then clearly optimal by Theorem 4.1. □

5. Properties of threshold policies

For each threshold policy (L, η) with L in $\mathbb{N}$ and $0 \leq \eta \leq 1$, the sequence $\{X(n)\}_0^\infty$ is a time-homogeneous Markov chain with state space $\mathbb{N}$ under the probability measure $P^{(L,\eta)}$.

This chain has a single ergodic set, namely, $\{0,1,\cdots,L\}$ if $\eta=0$ or $\{0,1,\cdots,L{+}1\}$ if $0<\eta\leq 1$, with all the other states being transient. Consequently, the Markov chain $\{X(n)\}_0^\infty$ admits under $P^{(L,\eta)}$ a unique invariant measure, which is denoted by $\mathbb{P}^{(L,\eta)}$ with corresponding expectation operator $\mathbb{E}^{(L,\eta)}$. This invariant measure $\mathbb{P}^{(L,\eta)}$ is computed by solving the equations

$$\mathbb{P}^{(L,\eta)}=\mathbb{P}^{(L,\eta)}Q((L,\eta)) \quad \text{and} \quad \sum_{x=0}^{\infty}\mathbb{P}^{(L,\eta)}(x)=1. \tag{5.1}$$

Routine calculations yield the solution of (5.1) in the form

$$\mathbb{P}^{(L,\eta)}(x)=W_L^\eta(x)\mathbb{P}^{(L,\eta)}(0)=\frac{W_L^\eta(x)}{1+\sum_{x=1}^{\infty}W_L^\eta(x)} \qquad x=0,1,\cdots \tag{5.2}$$

where upon defining $\rho:=\frac{\lambda\bar{\mu}}{\mu\bar{\lambda}}$, for $L\geq 1$,

$$W_L^\eta(x)=\begin{cases}1 & \text{if } x=0;\\ \frac{\rho^x}{\bar{\mu}} & \text{if } 1\leq x<L;\\ \frac{\rho^L}{\bar{\mu}}\frac{\bar{\lambda}}{1-\lambda\eta} & \text{if } x=L;\\ \frac{\rho^{L+1}}{\bar{\mu}}\frac{\bar{\lambda}^2\eta}{1-\lambda\eta} & \text{if } x=L+1;\\ 0 & \text{if } x>L+1\end{cases} \tag{5.3a}$$

while for $L=0$,

$$W_0^\eta(0)=1,\ W_0^\eta(1)=\frac{\lambda\eta}{\mu} \quad \text{and} \quad W_0^\eta(x)=0,\ x>1. \tag{5.3b}$$

Let X denote a generic $\mathbb{N}$-valued random variable. For any mapping $d:\mathbb{N}\rightarrow\mathbb{R}$, the quantity $\mathbb{E}^{(L,\eta)}d(X)$ is always *finite* since $\mathbb{P}^{(L,\eta)}$ has finite support, and the first passage time to the set of ergodic states being almost surely finite, the following characterization is obtained [4, Thm. I.15.2, pp. 92] [5, Thm. 4.5.4, pp. 97].

Lemma 5.1 *For any mapping $d:\mathbb{N}\rightarrow\mathbb{R}$, the convergence*

$$\lim_{n\rightarrow\infty}\frac{1}{n+1}\sum_{t=0}^{n}d(X(t))=\mathbb{E}^{(L,\eta)}d(X) \qquad P^{(L,\eta)}-a.s. \tag{5.4}$$

takes place, independently of the initial distribution. Furthermore, if the RV's $\{d(X(n))\}_0^\infty$ are uniformly integrable under $P^{(L,\eta)}$, then the convergence (5.4) also holds in $L^1(\Omega,\mathbb{F},P^{(L,\eta)})$.

When d is bounded, the second part of Lemma 5.1 follows from the Bounded Convergence Theorem, whereas if d is unbounded but monotone, uniform integrability of the RV's $\{d(X(n)\}_0^\infty$ under $P^{(L,\eta)}$ is determined by the integrability of the RV $d(\Xi)$.

Lemma 5.2 *If the mapping $d\colon \mathbb{N} \rightarrow \mathbb{R}$ is monotone and the RV $d(\Xi)$ is integrable, then the RV's $\{d(X(n))\}_0^\infty$ form an uniformly integrable sequence under $P^{(L,\eta)}$, and the convergence* (5.4) *thus takes place both $P^{(L,\eta)}$ – a.s. and in $L^1(\Omega, \mathbb{F}, P^{(L,\eta)})$.*

Proof: Under the threshold policy (L, η), the queue sizes $\{X(n)\}_0^\infty$ satisfy the inequality

$$X(n) \leq \Xi \vee (L+1) \quad P^{(L,\eta)} - a.s. \qquad n = 0, 1, \cdots \tag{5.5}$$

and the monotonicity of d implies $|d(X(n))| \leq |d(\Xi)| + |d(L+1)| + |d(0)|$ for all $n = 0, 1, \cdots$. Therefore, the sequence $\{d(X(n))\}_0^\infty$ is uniformly integrable under $P^{(L,\eta)}$ whenever the RV $d(\Xi)$ is integrable. Lemma 5.1 concludes the proof. $\square$

The estimate (5.5) implies that for each x in $\mathbb{N}$, the Markov chain $\{X(n)\}_0^\infty$ visits only finitely many states, say $\{0, 1, \cdots, x \vee (L+1)\}$, $P_x^{(L,\eta)}$-a.s. The chain is thus equivalent to a *finite-state* Markov chain under $P_x^{(L,\eta)}$, and the Bounded Convergence Theorem yields

Lemma 5.3 *For any mapping $d\colon \mathbb{N} \rightarrow \mathbb{R}$, the convergence* (5.4) *always takes place both $P_x^{(L,\eta)}$ – a.s. and in $L^1(\Omega, \mathbb{F}, P_x^{(L,\eta)})$ for every x in $\mathbb{N}$.*

The next lemma will be useful in proving Theorem 6.2 which constitutes the main result of Section 6. The results obtained will also be applicable to various situations discussed in the companion paper [10].

Lemma 5.4 *For any mapping $d\colon \mathbb{N} \rightarrow \mathbb{R}$ and any threshold policy (L, η), there always exist a scalar J and a mapping $h\colon \mathbb{N} \rightarrow \mathbb{R}$ such that*

$$h(x) + J = d(x) + (L,\eta)(x) E[h(A^1(x))] + \overline{(L,\eta)(x)} E[h(A^0(x))] \tag{5.6}$$

for all x in $\mathbb{N}$. The quantity J is given by

$$J = \lim_{n \uparrow \infty} \frac{1}{n+1} E_x^{(L,\eta)} [\sum_{t=0}^{n} d(X(t))] = \mathbb{E}^{(L,\eta)} d(X), \tag{5.7}$$

whereas the mapping $h\colon \mathbb{N} \rightarrow \mathbb{R}$ is unique up to an additive constant and is given by

$$h(x) = E_x^{(L,\eta)} [\sum_{t=0}^{\tau-1} d(X(t))] - E_x^{(L,\eta)}[\tau] J \tag{5.8}$$

for all x in $I\!N$, *under the constraint* $h(L) = 0$. *Here τ is the $I\!F_n$-stopping time defined by* (4.13) *but with L instead of L^γ.*

Equation (5.6) is sometimes referred as the *Poisson* equation associated with the cost function d under the threshold policy (L, η), and Lemma 5.4 asserts that the pair (h, J) given by (5.7)-(5.8) is a solution to this Poisson equation. Its proof is by now standard and is omitted for sake of brevity. Crucial to the argument is the fact that the chain $\{X(n)\}_0^\infty$ reduces to a *finite-state* Markov chain under $P_x^{(L,\eta)}$ for every x in $I\!N$. The interested reader is invited to consult the monograph by Ross [16] or the work by Shwartz and Makowski [17] for a typical approach.

It is clear from Lemma 5.2 that the equalities $T((L,\eta)) = \mu I\!P^{(L,\eta)}[X \neq 0]$, $N((L,\eta)) = I\!E^{(L,\eta)} X$ and

$$J^\gamma((L,\eta)) = T((L,\eta)) - \gamma N((L,\eta)) = \mu I\!P^{(L,\eta)}[X \neq 0] - \gamma I\!E^{(L,\eta)} X \tag{5.9}$$

are obtained under the threshold policy (L, η) for every L in $I\!N$ and $0 \leq \eta \leq 1$. The following property is immediate.

Lemma 5.5 *For every L in $I\!N$, the mappings $\eta \to T((L,\eta))$ and $\eta \to N((L,\eta))$ are continuous and strictly monotone increasing on the interval* $[0,1]$, *whence, the quantities $T((L,0))$ and $N((L,0))$ increase as L increases.*

Proof: From (5.2)-(5.3), continuity of the mappings $\eta \to T((L,\eta))$ and $\eta \to N((L,\eta))$ is a direct consequence of the continuity of the mappings $\eta \to I\!P^{(L,\eta)}(x)$ for all x in $I\!N$, whereas strict monotonicity follows from the strict monotonicity of each one of the mappings $\eta \to I\!P^{(L,\eta)}[X \geq k]$ for $1 \leq k \leq L+1$. The second part of the lemma is now immediate since $(L+1, 0) = (L, 1)$. $\square$

With the notation $a = b\frac{\eta}{1-\lambda\eta}$ and $b = \frac{\rho^{L-1}\lambda^2}{\mu^2}$, (5.2)-(5.3) yield the expressions

$$T((L,\eta)) = \frac{\mu(V_{L-1} + a)}{1 + V_{L-1} + a} = \frac{\mu((b - \lambda V_{L-1})\eta + V_{L-1})}{(b - \lambda(1 + V_{L-1}))\eta + 1 + V_{L-1}} \tag{5.10}$$

and

$$N((L,\eta)) = \frac{U_{L-1} + a(L + \bar{\mu})}{1 + V_{L-1} + a} = \frac{(b(L+\bar{\mu}) - \lambda U_{L-1})\eta + U_{L-1}}{(b - \lambda(1 + V_{L-1}))\eta + 1 + V_{L-1}} \tag{5.11}$$

for $L \geq 1$, with $T(0,\eta) = \mu N(0,\eta) = \frac{\mu\lambda\eta}{\mu+\lambda\eta}$. Here U_{L-1} and V_{L-1} are defined for $L \geq 1$ by

$$U_{L-1} = \begin{cases} \sum_{k=1}^{L-1} k\frac{\rho^k}{\bar{\mu}} + L\bar{\lambda}\frac{\rho^L}{\bar{\mu}} & \text{if } L > 1; \\ \frac{\lambda}{\mu} & \text{if } L = 1 \end{cases} \tag{5.12}$$

and

$$V_{L-1} = \begin{cases} \sum_{k=1}^{L-1} \frac{\rho^k}{\bar{\mu}} + \bar{\lambda}\frac{\rho^L}{\bar{\mu}} & \text{if } L > 1; \\ \frac{\lambda}{\mu} & \text{if } L = 1. \end{cases} \tag{5.13}$$

The next lemma shows that for each $\gamma > 0$, the quantity $J^\gamma((L,0))$ has a *single* relative maximum.

Lemma 5.6 *For each $\gamma > 0$, the mapping $L \to J^\gamma((L,0))$ is discretely unimodal, with the global maximum being achieved at at most two adjacent levels.*

Proof: For each $L = 0, 1, \cdots$, pose $\Delta(L) := J^\gamma((L+1,0)) - J^\gamma((L,0))$, and observe from the relations (5.10)-(5.11) that

$$\Delta(L) = \begin{cases} \frac{\rho^L\lambda}{\mu\bar{\mu}(1+V_L)(1+V_{L-1})}(\mu - \gamma C_L) & \text{if } L \geq 1; \\ \frac{V_0}{1+V_0}(\mu - \gamma C_0) & \text{if } L = 0 \end{cases} \tag{5.14}$$

where $C_L = (L+\bar{\mu})(1+V_{L-1}) - U_{L-1}$ for $L \geq 1$ and $C_0 = 1$. The property will be established if the quantities C_L can be shown to be positive and strictly increasing in L. Positivity is obvious since $C_L > 0$ if and only if $L + \bar{\mu} > N((L-1,1))$, an inequality which obviously holds since the chain can never go above level L under the threshold policy $(L-1,1)$. Monotonicity follows from the relations

$$C_{L+1} - C_L = 1 + V_L + (L+\mu)(V_L - V_{L-1}) - (U_L - U_{L-1}) = 1 + V_L > 0$$

for $L \geq 1$, with $C_1 - C_0 = \bar{\mu}(1+V_0) > 0$, by making use of (5.12)-(5.13). $\Box$

6. The long-run average problems

The long-run average problem (LP^γ) is now solved by standard Tauberian arguments applied to the discounted problem (LP^γ_β).

Theorem 6.1 *For each $\gamma > 0$, there always exists a threshold policy $(L^*_\gamma, \eta^*_\gamma)$, with $0 \leq L^*_\gamma \leq L^\gamma$, which solves the long-run average problem (LP^γ) and yields the optimal cost J^γ as*

$J^\gamma = I\!E^{(L^*_\gamma,\eta^*_\gamma)}c^\gamma(X)$. *When* $\mu - \gamma \le 0$, i.e., $L^\gamma = 0$, *then necessarily* $\eta^*_\gamma = 0$ *and* $J^\gamma = 0$, *while if* $\mu - \gamma > 0$, i.e., $L^\gamma > 0$, *then* η^*_γ *can always be chosen to be* 1.

Proof: If $L^\gamma = 0$, the invariant measure $I\!P^{(0,0)}$ has all its mass concentrated on $\{0\}$ and the cost c^γ is non-positive. The relation (5.9) readily implies $J^\gamma((0,0)) = I\!E^{(0,0)}c^\gamma(X) = 0 \ge J^\gamma(\pi)$ for all π in $\mathcal{P}$, and the optimality of the policy (0,0) thus trivially follows.

To study the non-trivial case $L^\gamma > 0$, consider $\mathcal{P}_f$ to be the collection of all policies π in $\mathcal{P}$ which incur a *finite* long-run average cost $J^\gamma(\pi)$, or equivalently, $\mathcal{P}_f := \{\pi \in \mathcal{P}:\ J^\gamma(\pi) > -\infty\}$. It should be clear that in solving (LP^γ), only those policies in $\mathcal{P}_f$ need to be considered. By virtue of (4.2), a version of the Tauberian Theorem [9, Lemma 1] [20] applies to the sequence $\{E^\pi c^\gamma(X(n))\}_0^\infty$ to give

$$J^\gamma(\pi) = \liminf_{n\uparrow\infty} \frac{1}{n+1} E^\pi \sum_{t=0}^{n} c^\gamma(X(t)) \le \liminf_{\beta\uparrow 1}(1-\beta)J^\gamma_\beta(\pi) \tag{6.1}$$

for every policy π in $\mathcal{P}_f$, with *equality* in (6.1) when $J^\gamma(\pi)$ is defined as a limit. As shown in Lemma 5.2, this is clearly the case when π is of threshold type, with

$$J^\gamma((L,\eta)) = I\!E^{(L,\eta)}c^\gamma(X) = \lim_{n\uparrow\infty} \frac{1}{n+1} E^{(L,\eta)} \sum_{t=0}^{n} c^\gamma(X(t)) = \lim_{\beta\uparrow 1}(1-\beta)J^\gamma_\beta((L,\eta)) \tag{6.2}$$

for every L in $I\!N$ and $0 \le \eta \le 1$. If $(L^\gamma_\beta, 1)$ is the optimal threshold policy for problem (LP^γ_β) given by Theorem 4.8, then obviously the inequality

$$(1-\beta)J^\gamma_\beta(\pi) \le (1-\beta)J^\gamma_\beta((L^\gamma_\beta, 1)) \tag{6.3}$$

holds for all β in (0,1) and every policy π in $\mathcal{P}_f$. Since the threshold L^γ_β never exceeds the level L^γ, which is independent of the discount factor β, there are only $L^\gamma + 1$ possible threshold values for L^γ_β. Consequently, for any sequence $\{\beta_n\}_1^\infty$ in (0,1), with $\beta_n \uparrow 1$ as $n \uparrow \infty$, there exists a further subsequence $\{\beta_m\}_1^\infty$, with $\beta_m \uparrow 1$, such that $\lim_{m\uparrow\infty} L^\gamma_{\beta_m} = L^*_\gamma$. The threshold values $\{L^\gamma_{\beta_m}\}_1^\infty$ being discrete, for m large enough, they must all be identical to L^*_γ and the optimal policy $(L^\gamma_{\beta_m}, 1)$ necessarily coincides with the threshold policy $(L^*_\gamma, 1)$. Upon taking the limits in (6.3) along the subsequence $\{\beta_m\}_1^\infty$, the relations (6.1)-(6.3) readily imply that $J^\gamma(\pi) \le J^\gamma((L^*_\gamma, 1))$ for all π in $\mathcal{P}_f$, i.e., the policy $(L^*_\gamma, 1)$ is optimal. $\square$

For the case $\mu - \gamma > 0$, i.e., $L^\gamma > 0$, the search for an optimal policy can therefore be restricted to the class of all *pure* threshold policies with threshold below level L^γ, and the optimal cost J^γ of problem (LP^γ) can simply be written as

$$J^\gamma = \max_{0 \leq L \leq L^\gamma} J^\gamma((L,1)) = \max_{0 \leq L \leq L^\gamma} I\!E^{(L,1)} c^\gamma(X) = \max_{L \in I\!N} I\!E^{(L,1)} c^\gamma(X). \tag{6.4}$$

In Section 5, explicit expressions were developed for the cost generated by threshold policies, and can now be used to identify the optimal cost and threshold policy through (6.4). This idea is now exploited to produce a somewhat strengthened result, which is key to solving the constrained problem (P_V).

Theorem 6.2 *For each threshold value L in $I\!N$, there always exists $\gamma(L) > 0$, with $L^{\gamma(L)} \geq L$, so that any admissible policy π in $\mathcal{P}$ given by*

$$\pi_n(H(n)) = \begin{cases} 1 & \text{if } X(n) < L; \\ \text{arbitrary in } [0,1] & \text{if } X(n) = L; \\ 0 & \text{if } X(n) > L \end{cases} \tag{6.5}$$

solves the long-run average problem $(LP^{\gamma(L)})$.

The following result will be useful in the proof of Theorem 6.2.

Lemma 6.3 *Let the pair (h, J) be obtained in Lemma 5.4. If the sequence $\{d(X(n))\}_0^\infty$ is uniformly integrable under $P^{(L,\eta)}$, then the convergence*

$$\begin{aligned} &\lim_{n \uparrow \infty} \frac{1}{n+1} [E^{(L,\eta)}[h(X(n+1))] - E^{(L,\eta)}[h(\Xi)]] \\ &= \lim_{n \uparrow \infty} \frac{1}{n+1} [E^{(L,\eta)}[1(\tau > n+1) h(X(n+1))] - E^{(L,\eta)}[h(\Xi)]] = 0 \end{aligned} \tag{6.6}$$

takes place.

Proof: A standard argument based on (5.6) [16,17] readily leads to the relation

$$E^{(L,\eta)} h(\Xi) + (n+1) J = E^{(L,\eta)} h(X(n+1)) + E^{(L,\eta)} \sum_{t=0}^{n} d(X(t)) \qquad n = 0, 1, \cdots$$

and therefore

$$\lim_{n \uparrow \infty} \frac{1}{n+1} [E^{(L,\eta)} h(X(n+1)) - E^{(L,\eta)} h(\Xi)] = 0 \tag{6.7}$$

owing to (5.7) and to Lemma 5.1. The quantity $E^{(L,\eta)}[1(\tau \leq n+1)h(X(n+1))]$ is bounded by virtue of (5.5), and the last part of relation (6.6) is now obtained. $\square$

A proof of Theorem 6.2.

For each L in $\mathbb{N}$, pose

$$\gamma(L) := \frac{T((L,1)) - T((L,0))}{N((L,1)) - N((L,0))}, \tag{6.8}$$

and observe that $\gamma(L) > 0$ owing to Lemma 5.5. The relations (5.9) and (6.8) easily imply $J^{\gamma(L)}((L,1)) = J^{\gamma(L)}((L,0))$. It then follows from Lemma 5.6 and (6.4) that

$$J^{\gamma(L)}((L,1)) = J^{\gamma(L)}((L,0)) = \max_{l \in \mathbb{N}} J^{\gamma(L)}((l,0)) = J^{\gamma(L)}, \tag{6.9}$$

and both policies $(L,1)$ and $(L,0)$ solve problem $(LP^{\gamma(L)})$. The reader will check from the expressions (5.10)-(5.11) that necessarily $L \leq L^{\gamma(L)}$, as expected from Theorem 6.1. To prove Theorem 6.2, i.e., that any policy π of the form (6.5) is optimal for problem $(LP^{\gamma(L)})$, it only remains to show that $J^{\gamma(L)}(\pi) = J^{\gamma(L)}$. Note that the policies $(L,1)$ and $(L,0)$ are clearly amongst these policies.

Take the mapping $d = c^{\gamma(L)}$ in Lemma 5.4, and let (h_i, J_i) be the corresponding solution to the Poisson equation (5.6) associated with the threshold policy (L,i), $i = 0,1$, when $h_1(L) = h_0(L) = 0$. The relation (6.9) yields $J_1 = J_0 = J^{\gamma(L)}$, whereas direct inspection of (5.8) reveals $h_1(x) = h_0(x)$ for all $x \neq L$ by the very definition of the stopping time τ. The condition $h_1(L) = h_0(L) = 0$ immediately implies that $h_1 \equiv h_0 := h$. This last fact, when substituted into the Poisson equations associated with the two policies $(L,1)$ and $(L,0)$, readily shows that they must coincide with $E[h(A^1(L))] = E[h(A^0(L))]$, and that they are of the form

$$h(x) + J^{\gamma(L)} = c^{\gamma(L)}(x) + p(x)E[h(A^1(x))] + \overline{p(x)}E[h(A^0(x))], \tag{6.10}$$

for all x in $\mathbb{N}$, where

$$p(x) = \begin{cases} 1 & \text{if } 0 \leq x < L; \\ \text{arbitrary in } [0,1] & \text{if } x = L; \\ 0 & \text{if } x > L. \end{cases}$$

Consider the policy π in $\mathcal{P}$ be defined by (6.5). As in the proof of Lemma 6.3, (6.10) again leads to the relation

$$J^{\gamma(L)} = \frac{1}{n+1} E^{\pi} \sum_{t=0}^{n} c^{\gamma(L)}(X(t)) + \frac{1}{n+1}[E^{\pi} h(X(n+1)) - E^{\pi} h(\Xi)] \qquad n = 0,1,\cdots \tag{6.11}$$

with

$$E^{\pi}[h(X(n+1))] = E^{\pi}[1(\tau \leq n+1)h(X(n+1))] + E^{\pi}[1(\tau > n+1)h(X(n+1))]. \quad (6.12)$$

By the very form (6.5) assumed for π and the definition of τ, it is easy to see that the first term on the right-hand side of (6.12) is bounded while the second term satisfies the relation

$$E^{\pi}[1(\tau > n+1)h(X(n+1))] = E^{(L,\eta)}[1(\tau > n+1)h(X(n+1))] \quad (6.13)$$

for any $0 \leq \eta \leq 1$, since both probability measures $P^{(L,\eta)}$ and P^{π} coincide on $\mathbb{F}_{\tau}$. As a result,

$$\begin{aligned} &\lim_{n\uparrow\infty} \frac{1}{n+1}\left[E^{\pi}[h(X(n+1))] - E^{\pi}[h(\Xi)]\right] \\ &= \lim_{n\uparrow\infty} \frac{1}{n+1}\left[E^{\pi}[1(\tau > n+1)h(X(n+1))] - E^{\pi}[h(\Xi)]\right] \\ &= \lim_{n\uparrow\infty} \frac{1}{n+1}\left[E^{(L,\eta)}[1(\tau > n+1)h(X(n+1))] - E^{(L,\eta)}[h(\Xi)]\right] = 0 \end{aligned} \quad (6.14)$$

upon invoking Lemma 6.3. The relation $J^{\gamma(L)}(\pi) = J^{\gamma(L)}$ is now obtained by taking the limit in (6.11) and making use of (6.14), and this completes the proof. □

A proof of Theorem 3.2.

It should be clear from Theorem 6.2 that any threshold policy (L,η), with η arbitrary in [0,1], satisfies conditions (C1) and (C3) of Theorem 3.3 for some $\gamma(L) > 0$. Thus, in order to solve problem (P_V), it only remains to find one such threshold policy that saturates the constraint.

Since $N((0,0)) = 0$, Lemma 5.5 readily implies the existence and uniqueness of the pair (L^*,η^*) such that $N((L^*,\eta^*)) = V$ if $N((\infty,1)) > V$. The optimal threshold and bias values L^* and η^* are uniquely defined by solving

$$\mathbb{E}^{(l,\eta)}X = V, \quad 0 \leq \eta \leq 1 \text{ and } l = 0,1,\cdots \quad (6.15)$$

With the help of Lemma 5.5 and equations (5.10)-(5.13), this is equivalent to the following: Either $0 < V \leq \frac{\lambda}{\lambda+\mu}$, and then

$$L^* = 0 \quad \text{and} \quad \eta^* = \frac{\mu V}{\lambda(1-V)}, \quad (6.16)$$

or $V > \frac{\lambda}{\lambda+\mu}$ and then

$$U_{L^*-1} - VV_{L^*-1} < V \leq U_{L^*} - VV_{L^*} \tag{6.17}$$

and

$$\eta^* = \frac{\mu^2(V(1+V_{L^*-1}) - U_{L^*-1})}{\rho^{L^*-1}\lambda^2(L^* + \bar{\mu} - V) + \lambda\mu^2(V(1+V_{L^*-1}) - U_{L^*-1})}, \tag{6.18}$$

in which case $L^* \geq 1$. Theorem 3.2 is proved. $\square$

APPENDIX I

A comparison result

The notion of *stochastic ordering* is useful for comparing the performance of various policies. This is achieved through a stochastic comparison result for the underlying queue size process among certain class of admissible policies. The reader is invited to consult the monographs by Ross [15, Chap. 8] and Stoyan [19] for further information on stochastic ordering.

Let P^1 and P^2 be two probability measures defined on $\mathbb{F}$, with the corresponding expectation operators E^1 and E^2, respectively. If Y is a $\mathbb{R}^n$-valued RV defined on $(\Omega, \mathbb{F})$, then the RV (Y, P^2) is said to be *stochastic larger* than (Y, P^1) if and only if $E^1[f(Y)] \leq E^2[f(Y)]$ for all *increasing* functions $f: \mathbb{R}^n \to \mathbb{R}$ for which these expectations exist; this is customarily denoted by $(Y, P^1) \leq_{st} (Y, P^2)$.

This notion extends naturally to *sequences* of $\mathbb{R}$-valued RV's defined on $(\Omega, \mathbb{F})$. The sequence $(\{Y(t)\}_0^\infty, P^2)$ is said to be *stochastic larger* than $(\{Y(t)\}_0^\infty, P^1)$ if and only if

$$((Y(0), Y(1), \cdots, Y(n)), P^1) \leq_{st} ((Y(0), Y(1), \cdots, Y(n)), P^2) \qquad n = 0, 1, \cdots \quad (I.1)$$

This is denoted simply by $(\{Y(t)\}_0^\infty, P^1) \leq_{st} (\{Y(t)\}_0^\infty, P^2)$ and is equivalent to

$$E^1[f(Y(0), Y(1), \cdots, Y(n))] \leq E^2[f(Y(0), Y(1), \cdots, Y(n))] \qquad n = 0, 1, \cdots \quad (I.2)$$

for all increasing functions $f: \mathbb{R}^{n+1} \to \mathbb{R}$ for which the expectations exist.

While the relation $(I.2)$ is usually hard to verify directly in practice, sufficient conditions are available in the literature, and one such condition, due to Veinott [19, pp. 29], is given below for easy reference. Throughout the discussion, the $\mathbb{R}^{n+1}$-valued RV $(Y(0), \cdots, Y(n))$ and the element $(y_0, \cdots, y_n)$ of $\mathbb{R}^{n+1}$ are denoted by $Y^{(n)}$ and $y^{(n)}$, respectively.

Lemma I.1 *Let $\{Y(t)\}_0^\infty$ be a sequence of $\mathbb{R}$-valued RV's on $(\Omega, \mathbb{F})$. If*

$$(Y(0), P^1) \leq_{st} (Y(0), P^2) \qquad (I.3a)$$

and for every a in $\mathbb{R}$

$$P^1[Y(n+1) > a | Y^{(n)} = x^{(n)}] \leq P^2[Y(n+1) > a | Y^{(n)} = y^{(n)}] \qquad n = 0, 1, \cdots \quad (I.3b)$$

whenever $x^{(n)} \leq y^{(n)}$ componentwise in $\mathbb{N}^{n+1}$, then $(\{Y(t)\}_0^\infty, P^1) \leq_{st} (\{Y(t)\}_0^\infty, P^2)$.

Here, this result is used as follows. For every admissible policy π in $\mathcal{P}$, introduce the sequence $\{\hat{\pi}_n\}_0^\infty$ of mappings $\hat{\pi}_n : \mathbb{N}^{n+1} \to [0,1]$ defined by

$$\hat{\pi}_n(x^{(n)}) := P^\pi[U(n) = 1 | X^{(n)} = x^{(n)}] = E^\pi[\pi_n(H(n)) | X^{(n)} = x^{(n)}] \qquad n = 0, 1, \cdots (I.4)$$

for all $x^{(n)}$ in $\mathbb{N}^{n+1}$.

Theorem I.2 *Consider two admissible policies π^1 and π^2 in $\mathcal{P}$. If the relations*

$$\hat{\pi}_n^1(x^{(n)}) \leq \hat{\pi}_n^2(y^{(n)}) \qquad n = 0, 1, \cdots (I.5)$$

hold for all $x^{(n)} \leq y^{(n)}$ with $x_n = y_n$, then $(\{X(t)\}_0^\infty, P^{\pi^1}) \leq_{st} (\{X(t)\}_0^\infty, P^{\pi^2})$.

Proof: Since the probability distribution of Ξ is independent of the policy, the relation $(I.3a)$ trivially holds. It suffices to show that the conditions $(I.5)$ imply $(I.3b)$.

Routine calculations first imply via (2.4)-(2.6) that for every policy π in $\mathcal{P}$,

$$P^\pi[X(n+1) > a | X^{(n)} = x^{(n)}] = \begin{cases} 1 & \text{if } x_n > a+1; \\ \bar{\mu} + \lambda\mu\hat{\pi}_n(x^{(n)}) & \text{if } x_n = a+1; \\ \lambda\bar{\mu}\hat{\pi}_n(x^{(n)}) & \text{if } x_n = a; \\ 0, & \text{if } x_n < a \end{cases} \qquad n = 0, 1, \cdots (I.6)$$

for all $x^{(n)}$ in $\mathbb{N}^{n+1}$. It is plain that $0 \leq \lambda\bar{\mu}\hat{\pi}_n(x^{(n-1)}, a) \leq \bar{\mu} + \lambda\mu\hat{\pi}_n(y^{(n-1)}, a+1) \leq 1$ for *all* $x^{(n-1)}$ and $y^{(n-1)}$ in $\mathbb{R}^n$, and $(I.3b)$ thus holds whenever $x^{(n)} \leq y^{(n)}$ *with* $x_n \leq y_n$ for any two *arbitrary* policies in $\mathcal{P}$. It thus suffices to show that $(I.3b)$ holds for $x^{(n)} \leq y^{(n)}$ with $x_n = y_n$ for the policies π^1 and π^2 considered here. But assumption $(I.5)$ and the relation $(I.6)$ readily combine to yield $(I.3b)$ whenever $x^{(n)} \leq y^{(n)}$ with $x_n = y_n$, and the conclusion now follows from Lemma I.1. $\square$

Theorem I.3 *Consider two admissible policies π^1 and π^2 in $\mathcal{P}$. If there exists a sequence $\{f_n\}_0^\infty$ of mappings $f_n : \mathbb{N} \to [0,1]$ such that*

$$\pi_n^1(h_n) \leq f_n(x_n) \leq \pi_n^2(h_n) \qquad n = 0, 1, \cdots (I.7)$$

for all h_n in $\mathbb{H}_n$, then $(\{X(t)\}_0^\infty, P^{\pi^1}) \leq_{st} (\{X(t)\}_0^\infty, P^{\pi^2})$.

Proof: It is plain from $(I.4)$ and $(I.7)$ that for all $x^{(n)}$ and $y^{(n)}$ in $\mathbb{N}^{n+1}$, $\hat{\pi}_n^1(x^{(n)}) \leq f_n(x_n)$ and $\hat{\pi}_n^2(y^{(n)}) \geq f_n(y_n)$. Condition $(I.5)$ is now easily justified with $x_n = y_n$, and the result follows from Theorem I.2. $\square$

APPENDIX II

A proof of Theorem 4.6

Let $\gamma > 0$ and $0 < \beta \leq 1$ held fixed throughout the discussion and pose $g := T_\beta^\gamma f$ for notational simplicity.

Lemma II.1 *If the mapping $f: \mathbb{N} \to \mathbb{R}$ is integer-concave over $\mathbb{N}$, then the mapping g is integer-concave over $\mathbb{N} - \{0\}$.*

Proof: From (4.4)-(4.5), it follows that

$$g(x) = c^\gamma(x) + \bar{\lambda}\beta E[f(A^0(x))] + \lambda\beta \max\{E[f(A^0(x)+1)], E[f(A^0(x))]\} \qquad (II.1)$$

for all x in $\mathbb{N}$. With the notation $h(x) := E[f(A^0(x))]$ for all x in $\mathbb{N}$, $(II.1)$ takes the form

$$g(x) = c^\gamma(x) + \bar{\lambda}\beta h(x) + \lambda\beta \max\{h(x+1), h(x)\} \qquad (II.2)$$

for all $x \geq 1$ in $\mathbb{N}$ since then $A^0(x) + 1 = A^0(x+1)$. The integer-concavity of f over $\mathbb{N}$ readily implies that the mapping h is integer-concave over $\mathbb{N} - \{0\}$, and thus so is the mapping $x \to \max\{h(x+1), h(x)\}$ by direct inspection. The cost c^γ being integer-concave, g is the sum of mappings which are integer-concave over $\mathbb{N} - \{0\}$ and is therefore integer-concave over $\mathbb{N} - \{0\}$. □

Owing to Lemma II.1, the study of the integer-concavity of g when f is integer-concave reduces to the study of the inequality

$$g(1) - g(0) \geq g(2) - g(1). \qquad (II.3)$$

The following facts are obtained by direct inspection of (4.5)-(4.7) and are useful in the discussion. If f is integer-concave with $\nabla f(0) = f(1) - f(0) \geq 0$, then $g(0) = \beta(\bar{\lambda} f(0) + \lambda f(1))$ and

$$\begin{aligned} g(1) - g(0) &= \mu - \gamma \\ &\quad + \beta \max\{(\bar{\mu} - \lambda)(f(1) - f(0)), \bar{\mu}[\bar{\lambda}(f(1) - f(0)) + \lambda(f(2) - f(1))]\}. \end{aligned} \qquad (II.4)$$

More precisely, one of two cases occurs:

(1) $\nabla f(1) < 0$: By concavity of f, $\nabla f(2) < 0$ with

$$g(1) - g(0) = \mu - \gamma + \beta(\bar{\mu} - \lambda)(f(1) - f(0)) \qquad (II.5a)$$

and

$$g(2)-g(1)=-\gamma+\beta\bigtriangledown f(1). \qquad (II.5b)$$

(2) $\bigtriangledown f(1)\geq 0$:

$$g(1)-g(0)=\mu-\gamma+\beta\bar{\mu}[\bar{\lambda}(f(1)-f(0))+\lambda(f(2)-f(1))] \qquad (II.6a)$$

and

$$g(2)-g(1)=\begin{cases}-\gamma+\beta\bar{\lambda}\bigtriangledown f(1) & \text{if } \bigtriangledown f(2)<0;\\ -\gamma+\beta(\bar{\lambda}\bigtriangledown f(1)+\lambda\bigtriangledown f(2)) & \text{if } \bigtriangledown f(2)\geq 0.\end{cases} \qquad (II.6b)$$

From $(II.4)$-$(II.6)$ and the concavity of f, the reader will readily check the following estimates.

Lemma II.2 *Assume that f is integer-concave with $f(1)-f(0)\geq 0$.*

(a) *If $f(1)-f(0)\leq 1$, the upper bound $g(1)-g(0)\leq\mu-\gamma+\beta\bar{\mu}\leq 1-\gamma<1$ holds, and* $\bigtriangledown f(1)\leq 1$ and $\bar{\lambda}(f(1)-f(0))+\lambda(f(2)-f(1))\leq 1$.

(b) *Suppose $\lambda+\mu\leq 1$. The lower bound $g(1)-g(0)\geq\mu-\gamma$ holds, and $\bar{\lambda}(f(1)-f(0))+\lambda(f(2)-f(1))\geq\bigtriangledown f(1)$.*

(c) *Suppose $\lambda+\mu>1$. The relation $\bar{\lambda}(f(1)-f(0))+\lambda(f(2)-f(1))\leq\bigtriangledown f(1)$ holds. If $f(1)-f(0)\leq\mu-\gamma$, then $g(1)-g(0)\geq(\mu-\gamma)(1-\beta\mu)>0$, whereas $\bar{\lambda}(f(1)-f(0))+\lambda(f(2)-f(1))\geq$ (resp. $\leq$)0 if and only if $g(1)-g(0)\geq$ (resp. $\leq$)$\mu-\gamma$.*

(d) *$\bigtriangledown f(1)\geq$ (resp. $\leq$)0 if and only if $g(2)-g(1)\geq$ (resp. $\leq$) $-\gamma$.*

The next lemma shows that the relation $(II.3)$ is easily obtained in certain cases.

Lemma II.3 *Assume that f is integer-concave with $f(1)-f(0)\geq 0$. If $g(1)-g(0)\geq 0$ with $\bigtriangledown f(1)<0$, or if $f(1)-f(0)\leq 1$ with $\bigtriangledown f(1)\geq 0$ and $\bigtriangledown f(2)\geq 0$, then $(II.3)$ holds.*

Proof: The first part of lemma readily follows from (d) of Lemma II.2. To prove the second part, note from the concavity of f that $f(1)-f(0)\geq\bigtriangledown f(1)$ and $f(2)-f(1)\geq\bigtriangledown f(2)$. Consequently,

$$(g(1)-g(0))-(g(2)-g(1))\geq\mu-\beta\mu[\bar{\lambda}(f(1)-f(0))+\lambda(f(2)-f(1))]\geq\mu-\beta\mu\geq 0$$

by making use of $(II.6)$ and (a) of Lemma II.2. □

A proof of Theorem 4.6.

First, the growth condition on g is seen to hold in each case (i)-(iv) from Lemma II.2. For (i), (a) and (b) of Lemma II.2 imply $\mu - \gamma \leq g(1) - g(0) \leq 1$. In (ii) $\nabla f(1) \leq \mu(\mu - \gamma) + \bar{\mu}(-\gamma) = \mu^2 - \gamma < 0$, while in the first part of (iv) $\nabla f(1) < 0$, and in both cases, $0 \leq g(1) - g(0) \leq \mu - \gamma$ and $g(2) - g(1) \leq -\gamma$ by (c) and (d) of Lemma II.2, respectively. In (iii), $\nabla f(1) \geq \mu(\mu - \gamma) + \bar{\mu}(-\gamma) = \mu^2 - \gamma \geq 0$, and $\bar{\lambda}(f(1) - f(0)) + \lambda(f(2) - f(1)) \geq \bar{\lambda}(\mu - \gamma) + \lambda(-\gamma)) \geq 0$, and thus by (a), (c) and (d) of Lemma II.2, $\mu - \gamma \leq g(1) - g(0) \leq 1$ and $g(2) - g(1) \geq -\gamma$. In the second part of (iv), $\nabla f(1) \geq 0$ and $\bar{\lambda}(f(1) - f(0)) + \lambda(f(2) - f(1)) \leq \bar{\lambda}(\mu - \gamma) + \lambda(-\gamma)) \leq 0$, whence $0 \leq g(1) - g(0) \leq \mu - \gamma$ and $g(2) - g(1) \geq -\gamma$ by (c) and (d) of Lemma II.2, respectively. Finally, in the last part of (iv), the growth condition on g is nothing but $0 \leq g(1) - g(0) \leq 1$ and $g(2) - g(1) \geq -\gamma$, while the condition $\nabla f(1) \geq 0$ always holds. Since $g(1) - g(0) \leq 1$ and $-\gamma \leq g(2) - g(1)$ by (a) and (d) of Lemma II.2, respectively, it suffices to show that $g(1) - g(0) \geq 0$. That this is indeed the case, observe from $(II.6a)$ that $g(1) - g(0) \geq \mu - \gamma + \beta\bar{\mu}(-\lambda\gamma) \geq \mu\bar{\mu}(1 - \beta\lambda\mu) \geq 0$, since here $\mu^2 \geq \gamma$ and $f(2) - f(1) \geq -\gamma$.

The proof of Theorem 4.6 will be complete by proving $(II.3)$. Since $0 \leq f(1) - f(0) \leq 1$ and $g(1) - g(0) \geq 0$ in all four cases (i)-(iv), it follows from Lemma II.3 that only the situation where $\nabla f(1) \geq 0$ and $\nabla f(2) < 0$ needs to be considered, in which case

$$(g(1) - g(0)) - (g(2) - g(1)) = \mu + \beta\bar{\mu}[\bar{\lambda}(f(1) - f(0)) + \lambda(f(2) - f(1))] - \beta\bar{\lambda} \nabla f(1) \quad (II.7)$$

by $(II.6)$. In (ii) and in the first part of (iv), $\nabla f(1) < 0$ and $(II.3)$ follows from Lemma II.3. For (i), $(II.7)$ and (a)-(b) of Lemma II.2 imply

$$(g(1) - g(0)) - (g(2) - g(1)) \geq \mu + \beta\bar{\mu} \nabla f(1) - \beta \nabla f(1) = \mu - \beta\mu \nabla f(1) \geq \mu - \beta\mu \geq 0.$$

For (iii), $\bar{\lambda}(f(1) - f(0)) + \lambda(f(2) - f(1)) \geq 0$ as shown earlier, and $(II.7)$ yields $(g(1) - g(0)) - (g(2) - g(1)) \geq \mu - \beta\bar{\lambda} \nabla f(1) \geq \mu - \beta\mu \nabla f(1) \geq 0$ since $\bar{\lambda} < \mu$. For the last part of (iv), it is plain that $\nabla f(1) \leq 1$ so that when $\bar{\lambda}(f(1) - f(0)) + \lambda(f(2) - f(1)) \geq 0$, the right-hand side of $(II.7)$ is obviously non-negative since $\mu - \beta\bar{\lambda} \nabla f(1) \geq \mu - \beta\mu \nabla f(1) \geq 0$. On the other hand, if $\bar{\lambda}(f(1) - f(0)) + \lambda(f(2) - f(1)) < 0$, then $\bar{\lambda}(f(1) - f(0)) < \lambda(f(1) - f(2)) \leq \lambda\gamma$ and thus $\bar{\lambda} \nabla f(1) \leq \bar{\lambda}\mu(f(1) - f(0)) \leq \mu\lambda\gamma$ since obviously $-\gamma \leq f(2) - f(1) < 0$. From $(II.7)$ and the fact $f(2) - f(1) \geq -\gamma$, it follows that

$$(g(1) - g(0)) - (g(2) - g(1)) \geq \mu + \beta\bar{\mu}\lambda(f(2) - f(1)) - \beta\mu\lambda\gamma \geq \mu - \beta\bar{\mu}\lambda\gamma - \beta\mu\lambda\gamma \geq 0.$$

Finally, for the second part of (iv), note that now $\bar{\lambda}(f(1) - f(0)) + \lambda(f(2) - f(1)) \leq 0$. It is convenient to use the estimate obtained for $g(1) - g(0)$ in (c) of Lemma II.2 to conclude that

$$(g(1) - g(0)) - (g(2) - g(1)) \geq (\mu - \gamma)(1 - \beta\mu) - (-\gamma + \beta\bar{\lambda} \bigtriangledown f(1)). \qquad (II.8)$$

Since by assumption $\bar{\lambda} \bigtriangledown f(1) \leq \bar{\lambda}[\mu(\mu-\gamma) + \bar{\mu}(-\gamma)] \leq \frac{\gamma}{\mu}(\mu^2 - \gamma)$, and $(\mu-\gamma)(1-\beta\mu) - (-\gamma + \beta\frac{\gamma}{\mu}(\mu^2 - \gamma)) = \mu - \beta\mu^2 + \frac{\beta\gamma^2}{\mu} > 0$, $(II.8)$ readily implies $(g(1) - g(0)) - (g(2) - g(1)) > 0$. The proof of Theorem 4.6 is now complete. □

REFERENCES

[1] D. P. Bertsekas, *Dynamic Programming and Stochastic Control*, Academic Press, New York, 1976.

[2] F. J. Beutler and K. W. Ross, "Optimal policies for controlled Markov chains with a constraint," *J. Math. Analysis and Appl.*, vol. 112, pp. 236-252 (1985).

[3] F. J. Beutler and K. W. Ross, "Time-average optimal constrained semi-Markov decision processes," *Adv. Appl. Prob.*, vol. 18, pp. 341-359 (1986).

[4] K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, Second Edition, Springer-Verlag, New York, 1967.

[5] K. L. Chung, *A Course in Probability Theory*, Second Edition, Academic Press, New York, 1974.

[6] M. Gerla and L. Kleinrock, "Flow control: A comparative survey," *IEEE Trans. Commun.*, vol. COM-28, pp. 553-574 (1980).

[7] D. Heyman and M. Sobel, *Stochastic Models in Operations Research, Volume II: Stochastic Optimization*, MacGraw-Hill, New York, 1984.

[8] A. A. Lazar, "The throughput time delay function of an $M|M|1$ queue," *IEEE Trans. Info. Theory*, vol. IT-29, pp. 914-918 (1983).

[9] S. A. Lippman, "Semi-Markov decision processes with unbounded rewards," *Management Science*, vol. 19, pp. 717-731 (1973)

[10] D.-J. Ma and A. M. Makowski, "A simple problem of flow control II: Implementation of threshold policies via Stochastic Approximations," *IEEE Trans. Auto. Control*, submitted (1987).

[11] D.-J. Ma and A. M. Makowski, "Parameter estimation, identification and adaptive implementation of threshold policies for a system of flow control," in preparation (1987).

[12] D.-J. Ma, A. M. Makowski and A. Shwartz, "Estimation and optimal control for constrained Markov chains," *Proceedings of the 25th IEEE Conference on Decision and Control*, Athens, Greece, pp. 994-999 (1986).

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

[14] K. W. Ross, *Constrained Markov decision processes with queueing applications*, Ph.D. thesis, CICE Program, University of Michigan, Ann Arbor, Michigan (1985).

[15] S. M. Ross, *Stochastic Processes*, John Wiley & Sons, New York, 1982.

[16] S. M. Ross, *Introduction to Stochastic Dynamic Programming*, Academic Press, New York, 1984.

[17] A. Shwartz and A. M. Makowski, "Adaptive policies for a system of competing queues I: Convergence results for the long-run average cost," *Adv. Appl. Prob.*, submitted (1987).

[18] S. Stidham, Jr., "Optimal control of admission to a queueing system," *IEEE Trans. Auto. Control*, vol. AC-30, pp. 705-713 (1985).

[19] D. Stoyan, *Comparison Methods for Queues and Other Stochastic Models*, John Wiley & Sons, New York, 1983.

[20] D. V. Widder, *The Laplace Transform*, Princeton University Press, Princeton, New Jersey, 1946.