Optimal Flow Control Allocation Policies in Communication Networks with Multiple Message Classes

Redha M. Bournas, Frederick J. Beutler, and Demosthenis Teneketzis

Abstract—**We consider M (≥ 2) transmitting stations sending packets to a single receiver over a slotted time-multiplexed link. For each phase consisting of T consecutive slots, the receiver dynamically allocates these slots among the M transmitters. The cost per slot for holding a packet may vary among the transmitters, and may be interpreted in terms of multiple classes of messages. Our objective is to characterize policies that minimize the discounted and long-term average costs due to holding packets at the M stations, based on delayed information on the numbers of packets being held at the respective transmitters.**

We derive properties of optimal (discounted) policies that reduce the computational complexity of the optimal flow control algorithm. For $M = 2$, we show that the minimal total cost is convex and submodular in the state, and we prove the following properties of optimal policies: 1) when the state at transmitter i increases by unity while the state at the other transmitter j is fixed, the optimal allocation is either unchanged, or increases by one at transmitter i and decreases by one at transmitter j; and 2) the optimal policy is of the threshold type. We use these properties to show that the optimization reduces to the calculation of optimal allocations for a *finite* number of states. In addition, for each such state (excluding the origin), property 1) implies a significant reduction in the computation of optimal allocations. As an application, we further characterize optimal policies when the message generation at the transmitter of higher priority is stochastically larger than the message generation at the other. Under additional restrictions on the average arrival rate and the second moment of the number of arrivals per slot, similar results are derived for optimal policies with time-average costs.

I. Introduction

The flow control problem addressed in this paper arises in the performance modeling of the "hop-by-hop" layer of computer communication networks. The hop-by-hop scheme considered here is the same as the one in [3]–[6]. Its purpose is to maintain a smooth flow of traffic between M transmitting stations sending packets to a single receiver over a communication channel. The channel is assumed to be slotted, that is, the channel time is divided into equal segments called *slots*. All messages consist of packets of equal length; the transmission time of a packet is one slot, and a packet transmission may only begin on a slot boundary. Each transmitter has an independent generally distributed arrival process of packets per slot and a buffer of infinite size. We assume that the arrival processes to distinct transmitters are mutually independent. Only one station is allowed to transmit during any particular slot. It is desired to determine policies that allocate slots among the respective transmitters to minimize holding costs for the M transmitting stations.

T consecutive slots form a *phase*. Prior to the beginning of each phase, the receiver informs each transmitter of the number of packets (referred to as a *window* size) that it is prepared to accept, and the particular slots in which each transmitter is allowed to transmit. For the purpose of making a decision on the window sizes for the current phase k, the receiver possesses the following information: the knowledge of the arrival statistics, the history of previous allocations, and the number of packets queued at each transmitter at the beginning of phase $k - 1$. The allocation at phase k must be predicated on delayed information (from phase $k - 1$) since transmission delays, together with computational requirements, make it impossible for data generated in one location to be instantly accessible elsewhere. Because a random number of new packets arrive at each transmitter during the course of phase $k - 1$, the allocation algorithm can alternatively be described as depending on partial information. Generally, advising the receiver of the number of packets requires a capacity of insignificant size as compared to the transmitted packets themselves; hence, each transmitter can send the receiver this information separately, or as an extra packet of small size over the channel.

At each station, there is a fixed cost per slot for holding a packet. Our cost per phase is the expectation of a linear combination of the number of untransmitted packets at the respective stations. A policy for the receiver is any function of the above-mentioned information that allocates the slots ($\leq T$) among the stations at each phase.

The results reported in [3]–[6] investigate discounted and time-average optimal policies when the cost per slot for holding a packet is equal at all transmitters. Here, we turn our attention to the general case where the holding cost may vary among the respective transmitters, which

Manuscript received February 12, 1991; revised December 16, 1991 and July 23, 1992. Paper recommended by Associate Editor, C. G. Cassandras.

R. M. Bournas is with IBM Networking Systems, Research Triangle Park, NC 27709.

F. J. Beutler and D. Teneketzis are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109.

IEEE Log Number 9207143.

0018-9286/93$03.00 © 1993 IEEE

may be interpreted in terms of multiple classes of messages. Thus, we can view this problem as one of optimal resource allocation in a queueing system with different classes of messages.

While the derivation of the explicit form of optimal policies appears very difficult, we are able to obtain qualitative properties that reduce the computational complexity of the optimal flow control algorithm. By defining the state to be the most recent (delayed) information on the number of packets awaiting transmission at the M stations, we find that the qualitative properties of optimal allocation policies have the following computational implications: 1) the state space can be partitioned into regions S_i, $0 \leq i < M$; for each state in S_i, an optimal policy allocates all the slots among $i+1$ transmitters; 2) for each state in S_i, there is a further reduction in the search for an optimal allocation among the $(i+1)$ transmitters. Since we relax the hypothesis that the holding costs are equal, the properties of optimal policies and their implications are weaker than those in [3].

When $M = 2$, we find that the addition of one packet at a transmitter either leaves the allocation of slots unchanged, or increases the allocation by unity in favor of that transmitter. This property is stronger than monotonicity in the state, and implies that the optimal policy is of the threshold type. It is reasonable to suppose that such a property significantly reduces the complexity of computation of the optimal allocation policy.

We formalize the model and formulate the problem as a discounted Markov decision process in Section II. In Section III, we derive structural properties of optimal policies for the problem with $M \geq 2$ transmitters. In Section IV, we let $M = 2$ and derive the properties of optimal policies stated in the preceding paragraph, and we show that the minimal total discounted cost is convex and submodular in the state. In Section V, we further characterize optimal policies when the message generation at the transmitter of higher holding cost is stochastically larger than the message generation at the other. In Section VI, we point out that the time-average optimal policies have properties similar to those derived in [4] for equal holding costs. Conclusions are presented in Section VII.

II. Problem Formulation

The operation of the hop-by-hop scheme is as described in the Introduction, as well as in [3]. Two constraints are placed on the model: 1) no holding costs are assessed for packets in the phase in which they are being transmitted, and 2) packets arriving in a particular phase may not be transmitted in that phase. Constraint 2) leads to a simpler implementation of the receiver without significant sacrifice in performance; indeed, 2) is analogous to a *gated* reservation system, as described in [1, sect. 3.5.2]. Moreover, relaxing these two constraints results in an optimization problem whose action space consists of not only the window sizes allocated to each transmitter, but also of the order in which the slots are scheduled for transmission. This is a considerably more difficult problem involving combinatorics, and is left as a topic for future investigation.

The processes of message generation at each transmitter are stochastic with known statistics. The number of packets generated at transmitter j during slot i, $\xi_i^{(j)}$, $j = 1, 2, \cdots, M$, $i = 1, 2, \cdots$, are assumed to be independent random variables. For fixed j, $\xi_i^{(j)}$, $i = 1, 2, \cdots$, are identically distributed (i.i.d.) with finite first moment $\lambda^{(j)}$. For the remainder of the paper, for $z \in Z_+^M$, $z^{(j)}$ will denote the jth component of z, unless stated otherwise.

Let $Y_k^{(j)}$ be the number of packets generated at transmitter j during phase k. For fixed j, these random variables are independent and identically distributed random variables, and the random variables $\{Y_k^{(j)}, 1 \leq j \leq M, k = 0, 1, \cdots\}$ are independent. We will denote by $N_k^{(j)}$ the number of packets at transmitter j at the beginning of phase k, and by $w_k^{(j)}$ the window size allocated to transmitter j during phase k. Assume that $w_0^{(j)}$, $N_0^{(j)}$, $j = 1, \cdots, M$, are given.

Prior to the beginning of each phase, the transmitters are informed of the particular slots during which they are allowed to transmit. The receiver computes $w_k^{(j)}$ before the beginning of phase k based on the following information: the knowledge of the arrival statistics, the history of previous window sizes, and the number of messages queued at each transmitter at the beginning of phase $(k-1)$.

Corresponding to this description, we define $X_k \triangleq (X_k^{(1)}, \cdots, X_k^{(M)})$ where

$$X_k^{(j)} = \max\{0, N_{k-1}^{(j)} - w_{k-1}^{(j)}\} \triangleq (N_{k-1}^{(j)} - w_{k-1}^{(j)})^+ \tag{2.1}$$

as the state of the system at the start of phase k. The values of $\{X_i, 1 \leq i \leq k\}$ will be used to compute $w_k = (w_k^{(1)}, w_k^{(2)}, \cdots, w_k^{(M)})$. The most recent value of X_k can be calculated by the receiver before the end of phase $(k-1)$ for the following reasons: 1) each transmitter j sends $N_{k-1}^{(j)}$ to the receiver just before the beginning of phase $(k-1)$, and 2) as T is sufficiently large, the receiver is guaranteed to receive the $N_{k-1}^{(j)}$ before the start of phase k.

Since packets arriving during phase $(k-1)$ are not allowed to be transmitted during this phase, then

$$N_k = X_k + Y_{k-1}, \qquad k \geq 1 \tag{2.2}$$

where $N_k = (N_k^{(1)}, \cdots, N_k^{(M)})$ and $Y_k = (Y_k^{(1)}, \cdots, Y_k^{(M)})$. Combining (2.1) and (2.2) yields the dynamic evolution equations for the state sequence $\{X_n\}$:

$$X_{k+1}^{(j)} = (X_k^{(j)} + Y_{k-1}^{(j)} - w_k^j)^+. \tag{2.3}$$

Observe from (2.3) that the transmission of N_k by the transmitters during phase k enables the receiver to deduce X_{k+1} since the receiver has previously computed w_k.

Fig. 1 illustrates the operation of the modeled flow control scheme in accordance with (2.2) and (2.3). In terms of the preceding notation, we have at the beginning of phase n the values $X_n = (2, 1, 0)$ and unsent packets

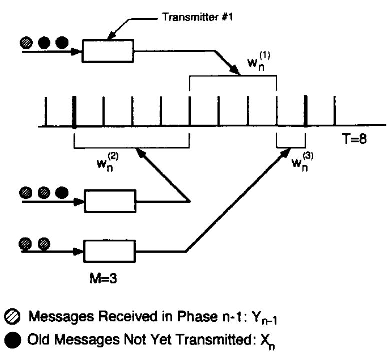


Fig. 1. Flow control scheme.

received during the preceding phase $Y_{n-1} = (1,2,2)$; this implies that $N_n = (3,2,2)$. With the specified value of w_n, we then obtain $X_{n+1} = (0,0,1)$, regardless of the arrival vector Y_n during this phase.

We let $c^{(j)} > 0$ be a weighting on the cost of holding a packet at transmitter j, and without loss of generality, assume that

$$c^{(1)} \geq c^{(2)} \geq \cdots \geq c^{(M)}. \tag{2.4}$$

These costs could be interpreted as measures of the transmission *priorities* of the respective transmitters. In the setting of (2.4), the transmitters are arranged in descending priority order.

For the N-step finite horizon problem, the objective is to minimize over the window vectors $w_i \in A$, $1 \leq i \leq N$, the expected total β-discounted cost

$$J_N(x) = \sum_{i=1}^{N} \beta^{i-1} \sum_{j=1}^{M} c^{(j)} E\left[(X_i^{(j)} + Y_{i-1}^{(j)} - w_i^{(j)})^+ | X_1 = x \right]. \tag{2.5}$$

Here, A is the set of admissible slot allocations in a phase, consisting of M-vectors w with nonnegative integer components satisfying $\sum_{j=1}^{M} w^{(j)} \leq T$. As in [3], we model this problem as a Markov decision process, and derive the optimality equations of dynamic programming. We first define the expected cost per phase as

$$L(x,w) \triangleq \sum_{j=1}^{M} c^{(j)} E\left[(Y^{(j)} + x^{(j)} - w^{(j)})^+ \right]. \tag{2.6}$$

For future reference, we alternatively write

$$L(x,w) = \sum_{i=1}^{M} c^{(i)} \sum_{j=0}^{\infty} P[Y^{(i)} > w^{(i)} - x^{(i)} + j]. \tag{2.7}$$

Let $V_k^\beta(x)$ be the minimal achievable total expected β-discounted cost when the system is in state x and there are k phases to go. The optimality equations of dynamic programming for the N-phase finite horizon problem yield

$$\begin{cases} V_0^\beta(x) = 0 \\ V_k^\beta(x) = \min_{w \in A} \{ L(x,w) + \beta E[V_{k-1}^\beta([Y + x - w]^+)] \}, \\ \qquad 1 \leq k \leq N. \end{cases} \tag{2.8}$$

Since β is fixed, we shall set

$$V_k \equiv V_k^\beta \tag{2.9}$$

to simplify the notation.

For the infinite horizon problem, the objective is to minimize over the windows $w_i \in A$ the expected total β-discounted ($\beta < 1$) cost:

$$J_\infty(x) = \sum_{i=1}^{\infty} \beta^{i-1} \sum_{j=1}^{M} c^{(j)} \cdot E\left[(X_i^{(j)} + Y_{i-1}^{(j)} - w_i^{(j)})^+ | X_1 = x \right]. \tag{2.10}$$

As in [3], $V(x) \triangleq \lim_{N \to \infty} V_N(x)$ exists, is finite, and is the minimal total expected discounted cost for the infinite horizon problem. We only show that $V(x)$ is finite, and we refer the reader to [3] for the proof that it exists and that it is the minimal total expected discounted infinite horizon cost.

For any policy $\pi \in P$, let $V_N(\pi, x)$ be the total expected β-discounted cost given that the system is in state x and there are N phases to go, i.e.,

$$V_N(\pi, x) = \sum_{i=1}^{N} \beta^{i-1} E_x^\pi [L(X_i, w_i)] \tag{2.11}$$

and

$$V_\infty(\pi, x) \triangleq \lim_{N \to \infty} V_N(\pi, x) \tag{2.12}$$

which exists by the nonnegativity of the one-step costs. Moreover, from [3, lemma 2.2], $V_\infty(\pi, x)$ is finite for every π and every $0 < \beta < 1$.

The optimality equation of dynamic programming for the infinite horizon problem is given by (see [3])

$$V(x) = \min_{w \in A} \{ L(x,w) + \beta E[V([Y + x - w]^+)] \}. \tag{2.13}$$

The result $V(x) = \lim_{N \to \infty} V_N(x)$ will enable us to derive the qualitative properties of optimal allocation policies for the infinite horizon problem by a standard limiting argument. Therefore, we first study the finite horizon problem.

III. Qualitative Properties of Optimal Policies for $M \geq 2$

In this section, we derive qualitative properties of $V_k(x)$ and $V(x)$ that will be used to partially characterize the structure of a set of optimal allocation policies. We remind the reader that the stations are arranged in descending order of holding costs [cf. (2.4)]. We derive the following properties for optimal discounted policies. Let x be the initial system state, and fix i, $1 \leq i \leq M$. Then,

P1) if $\sum_{j=1}^{i} x^{(j)} \geq T$, there exists an optimal allocation $w_*(x)$ such that $\sum_{j=1}^{i} w_*^{(j)}(x) = T$;

P2) if $\sum_{j=1}^{i} x^{(j)} \leq T$, there exists an optimal allocation $w_*(x)$ such that $\sum_{j=1}^{i} w_*^{(j)}(x) \geq \sum_{j=1}^{i} x^{(j)}$.

Property P1) assures that all of the slots are allocated to the first i transmitters when the sum of their known queue lengths exceeds T. Property P2) asserts that the first i transmitters are allocated at least as many slots as the sum of their known queue lengths. We remark that for the case of equal holding costs, the analogous of property P2) is the stronger result: $w_*^{(j)} \geq x^{(j)}$ for each j; see [3, lemma 3.6].

As a consequence of properties P1) and P2), we transform this optimal resource allocation problem into an equivalent one with a smaller action space. This then reduces the computational complexity of the optimal control algorithm. In particular, for large values of M and/or T, this reduction may be significant.

To proceed with the analysis, we first establish some preliminary properties of $V_k(x)$ and $V(x)$.

Lemma 3.1:

a) $V_k(x)$ is a nondecreasing function in each $x^{(j)}$, $1 \leq j \leq M$.

b) $V(x)$ is a nondecreasing function in each $x^{(j)}$, $1 \leq j \leq M$.

Proof: Mimic the proof of [3, lemma 3.1]. ■

Lemma 3.2:

a) $V_k(x)$ is achieved by an allocation $w_k(x) \in A$ satisfying $\sum_{j=1}^{M} w_k^{(j)}(x) = T$.

b) $V(x)$ is achieved by an allocation $w(x) \in A$ satisfying $\sum_{j=1}^{M} w^{(j)}(x) = T$.

Proof: Mimic the proof of [3, lemma 3.2]. ■

By Lemma 3.2, there is no loss of optimality in restricting attention to those Markov policies whose action space is the set

$$\bar{A} \triangleq \left\{ w = (w^{(1)}, w^{(2)}, \cdots, w^{(M)}) \in A \colon \sum_{j=1}^{m} w^{(j)} = T \right\}. \tag{3.1}$$

If we define

$$G_k(x,w) \triangleq L(x,w) + \beta E\left[V_{k-1}([Y + x - w]^+)\right] \tag{3.2}$$

$$G(x,w) \triangleq L(x,w) + \beta E[V([Y + x - w]^+)] \tag{3.3}$$

then the dynamic programming equations (2.8) and (2.13) become, respectively,

$$V_k(x) = \min_{w \in \bar{A}} \left\{ L(x,w) + \beta E\left[V_{k-1}([Y + x - w]^+)\right] \right\} = \min_{w \in \bar{A}} \{G_k(x,w)\} \tag{3.4}$$

$$V(x) = \min_{w \in \bar{A}} \{ L(x,w) + \beta E[V([Y + x - w]^+)] \} = \min_{w \in \bar{A}} \{G(x,w)\}. \tag{3.5}$$

We let e_i denote the M-dimensional row vector with one in the ith entry and zero in all other entries. In the next lemma, we prove that if the holding cost at transmitter i is higher than the one at j, then the minimal total cost starting from state x is higher than the one starting from $x - e_i + e_j$. That is, transferring a packet from the queue of transmitter i to the one of transmitter j results in a smaller minimal total cost since it costs more to hold a packet at i than at j. We remark that if the holding costs at transmitters i and j are equal, then by a symmetry argument, the minimal total cost is unaltered by a packet transfer from i to j (see [3, lemma 3.4]).

Lemma 3.3: If $x^{(i)} \geq 1$, then

$$V_k(x - e_i + e_j) \leq V_k(x) \qquad \text{for } j > i \tag{3.6}$$

$$V(x - e_i + e_j) \leq V(x) \qquad \text{for } j > i. \tag{3.7}$$

Proof: The proof of (3.6) is by induction on k. The assertion trivially holds for $k = 0$. Suppose the assertion is true for k, and let $V_{k+1}(x) = G_{k+1}(x, w_*)$ for some $w_* \in \bar{A}$. We consider the two cases $w_*^{(i)} = 0$ and $w_*^{(i)} \geq 1$. We begin first with the case $w_*^{(i)} = 0$. We clearly have

$$\begin{aligned} V_{k+1}&(x - e_i + e_j) \\ &\leq G_{k+1}(x - e_i + e_j, w_*) = L(x - e_i + e_j, w_*) \\ &\quad + \beta E\left[V_k\left([Y + x - e_i + e_j - w_*]^+\right)\right]. \end{aligned} \tag{3.8}$$

Since when $w_*^{(i)} = 0$ and $x^{(i)} \geq 1$

$$[Y^{(i)} + x^{(i)} - 1 - w_*^{(i)}]^+ = [Y^{(i)} + x^{(i)}] - 1, \tag{3.9}$$

and for any realization $Y^{(j)}$

$$(Y^{(j)} + x^{(j)} + 1 - w_*^{(j)})^+ \leq (Y^{(j)} + x^{(j)} - w_*^{(j)})^+ + 1, \tag{3.10}$$

we then obtain

$$\begin{aligned} L(x - e_i + e_j, w_*) &= c^{(i)} E[Y^{(i)} + x^{(i)} - 1] \\ &\quad + c^{(j)} E\left[(Y^{(j)} + x^{(j)} + 1 - w_*^{(j)})^+\right] \\ &\quad + \sum_{l \neq i, l \neq j} c^{(l)} E\left[(Y^{(l)} + x^{(l)} - w_*^{(l)})^+\right] \\ &\leq c^{(i)} E[Y^{(i)} + x^{(i)}] + c^{(j)} E\left[(Y^{(j)} + x^{(j)} - w_*^{(j)})^+\right] \\ &\quad + \sum_{l \neq i, l \neq j} c^{(l)} E\left[(Y^{(l)} + x^{(l)} - w_*^{(l)})^+\right] + c^{(j)} - c^{(i)}. \end{aligned} \tag{3.11}$$

As $c^{(j)} \leq c^{(i)}$ for $j > i$ [cf. (2.4)], we then get

$$L(x - e_i + e_j, w_*) \leq L(x, w_*). \tag{3.12}$$

Using (3.9), (3.10), the monotonicity of $V_k(x)$ in the jth component, and the induction hypothesis, respectively, then

$$\begin{aligned} E&\left[V_k\left([Y + x - e_i + e_j - w_*]^+\right)\right] \\ &\leq E\left[V_k\left([Y + x - w_*]^+ - e_i + e_j\right)\right] \\ &\leq E\left[V_k([Y + x - w_*]^+)\right]. \end{aligned} \tag{3.13}$$

Combining (3.12) and (3.13) results in $V_{k+1}(x - e_i + e_j) \le V_{k+1}(x)$. For the case $w_*^{(j)} \ge 1$, we have

$$V_{k+1}(x - e_i + e_j) \le G_{k+1}(x - e_i + e_j, w_* - e_i + e_j) = G_{k+1}(x, w_*) = V_{k+1}(x) \quad (3.14)$$

where the first equality on the RHS of (3.14) is immediate from the definitions of $G_{k+1}(\cdot,\cdot)$ and $L(x,w)$ [cf. (3.2) and (2.6)]. This concludes the proof of (3.6). To prove (3.7), take the limit on both sides of (3.6) as $k \to \infty$. ■

As a corollary, we prove the following result. Suppose that a slot is allocated to a packet with a certain holding cost according to an optimal policy. Then if this packet is replaced by one of higher holding cost, it is optimal to assign this slot to the new packet without altering the allocation of the other slots.

Corollary 3.4:

a) Suppose $j > 1$, $x^{(j)} \ge 1$ and let $V_k(x) = G_k(x, w_*)$. Then if $w_*^{(j)} \ge 1$,

$$V_k(x + e_i - e_j) = G_k(x + e_i - e_j, w_* + e_i - e_j) = V_k(x) \quad \text{for all } i < j.$$

b) Property a) above holds for the infinite horizon cost $V(x)$.

Proof: We prove a) only because the proof of b) follows in a similar fashion. By the minimality of $V_k(x + e_i - e_j)$,

$$V_k(x + e_i - e_j) \le G_k(x + e_i - e_j, w_* + e_i - e_j) = G_k(x, w_*) = V_k(x). \quad (3.15)$$

But by Lemma 3.3, for $i < j$,

$$V_k(x) \le V_k(x + e_i - e_j). \quad (3.16)$$

The combination of (3.15) and (3.16) leads to the desired result. ■

The above property is very useful in reducing the number of computations of state-dependent allocation schemes. Given an optimal allocation for x, it enables us to obtain, in some cases, an optimal allocation for $x + e_i - e_j$ without the requirement of any calculations. We now prove the existence of an optimal allocation policy satisfying properties P1) and P2).

Theorem 3.5: For any i, $1 \le i \le M$, there exists an optimal allocation $w_* \in \underline{A}$ such that $\Sigma_{j=1}^{i} w_*^{(j)} \ge \min\{\Sigma_{j=1}^{i} x^{(j)}, T\}$, that is,

a) if $\Sigma_{j=1}^{i} x^{(j)} \ge T$, then there exists an optimal allocation $w_* \in A$ such that $\Sigma_{j=1}^{i} w_*^{(j)} = T$, i.e., $V_k(x) = G_k(x, w_*)$;

b) if $\Sigma_{j=1}^{i} x^{(j)} \le T$, then there exists an optimal allocation $w_* \in \underline{A}$ such that $\Sigma_{j=1}^{i} w_*^{(j)} \ge \Sigma_{j=1}^{i} x^{(j)}$, i.e., $V_k(x) = G_k(x, w_*)$;

c) there exist optimal policies for the infinite horizon total expected discounted cost problem with properties a) and b).

Proof: We prove a) and b) first. Assume that $i < M$, for if $i = M$, the statements of Theorem 3.5 are trivially true. The proof is by construction; that is, for every optimal policy that does not satisfy the stated property, there exists an optimal policy that satisfies it. Let $V_k(x) = G_k(x, w_*)$ for some $w_* \in \underline{A}$, and suppose that $\Sigma_{j=1}^{i} w_*^{(j)} < \min\{\Sigma_{j=1}^{i} x^{(j)}, T\}$. This then implies that $w_*^{(m)} \ge 1$ for some $m > i$, and $w_*^{(n)} < x^{(n)}$ for some $n \le i$. We next consider the allocation $v_* = w_* + e_n - e_m$, and show that

$$G_k(x, v_*) \le V_k(x). \quad (3.17)$$

By a straightforward calculation,

$$\begin{aligned} L(x, v_*) - L(x, w_*) &= c^{(n)}\Big\{E\Big[\big(Y^{(n)} + x^{(n)} - w_*^{(n)} - 1\big)^+ - \big(Y^{(n)} + x^{(n)} - w_*^{(n)}\big)^+\Big]\Big\} \\ &\quad + c^{(m)}\Big\{E\Big[\big(Y^{(m)} + x^{(m)} - w_*^{(m)} + 1\big)^+ - \big(Y^{(m)} + x^{(m)} - w_*^{(m)}\big)^+\Big]\Big\}. \end{aligned} \quad (3.18)$$

As $w_*^{(n)} < x^{(n)}$, the first term in { } on the RHS of (3.18) is equal to -1. Moreover, since the second term in { } on the RHS of (3.18) cannot exceed 1,

$$L(x, v_*) - L(x, w_*) \le c^{(m)} - c^{(n)} \le 0, \quad (3.19)$$

the last inequality following from (2.4) as $m > n$. We show now that

$$E\big[V_{k-1}([Y + x - v_*]^+)\big] \le E\big[V_{k-1}([Y + x - w_*]^+)\big]. \quad (3.20)$$

$[Y^{(n)} + x^{(n)} - w_*^{(n)} - 1]^+ = [Y^{(n)} + x^{(n)} - w_*^{(n)}] - 1$ for every realization $Y^{(n)}$ because $x^{(n)} > w_*^{(n)}$. Using this, together with the monotonicity of $V_{k-1}(\cdot)$ in the mth component,

$$\begin{aligned} E\big[V_{k-1}([Y + x - v_*]^+)\big] &= E\big[V_{k-1}\big([Y + x - w_* - e_n + e_m]^+\big)\big] \\ &\le E\big[V_{k-1}\big([Y + x - w_*]^+ - e_n + e_m\big)\big]. \end{aligned} \quad (3.21)$$

Applying (3.6) to the RHS of (3.21), we obtain, as $m > n$,

$$E\big[V_{k-1}\big([Y + x - w_*]^+ - e_n + e_m\big)\big] \le E\big[V_{k-1}([Y + x - w_*]^+)\big]. \quad (3.22)$$

Inequalities (3.21) and (3.22) lead to (3.20), and the combination of (3.19) and (3.20) produces (3.17). If $\Sigma_{j=1}^{i} v_*^{(j)} = \min\{\Sigma_{j=1}^{i} x^{(j)}, T\}$, the proof is then complete; otherwise, $\Sigma_{j=1}^{i} v_*^{(j)} < \min\{\Sigma_{j=1}^{i} x^{(j)}, T\}$ implies the existence of an $m_1 > i$ such that $v_*^{(m_1)} \ge 1$ and an $n_1 \le i$ such that $v_*^{(n_1)} < x^{(n_1)}$. By the same argument as above,

$$G_k(x, v_* + e_{n_1} - e_{m_1}) \le G_k(x, v_*) \le V_k(x). \quad (3.23)$$

By proceeding in this fashion, we get

$$G_k(x, u_*) \le V_k(x) \quad (3.24)$$

for some $u_* = w_* + \Sigma_{j=1}^{l}(e_{n_j} - e_{m_j})$ and integers l, $n_j \le i$, $m_j > i$, $1 \le j \le l$ such that $\Sigma_{j=1}^{i} u_*^{(j)} = \min\{\Sigma_{j=1}^{i} x^{(j)}, T\}$. This completes the proof of a) and b). To prove c), use the dynamic programming equation (3.5) for $V(x)$, inequality (3.7), and proceed in the same way. ■

As a corollary, we show that if the ordering (2.4) is strict, i.e.,

$$c^{(1)} > c^{(2)} > \cdots > c^{(M)}, \tag{3.25}$$

then any optimal policy must allocate at least $\min\{\Sigma_{j=1}^{i} x^{(j)}, T\}$ to the first i transmitters.

Corollary 3.6: Suppose (3.25) holds. Then for any i, $1 \le i \le M$, every optimal allocation $w_* \in \bar{A}$ must satisfy $\Sigma_{j=1}^{i} w_*^{(j)} \ge \min\{\Sigma_{j=1}^{i} x^{(j)}, T\}$, that is,

a) if $\Sigma_{j=1}^{i} x^{(j)} \ge T$, every optimal allocation $w_* \in \bar{A}$ is such that $\Sigma_{j=1}^{i} w_*^{(j)} = T$;

b) if $\Sigma_{j=1}^{i} x^{(j)} \le T$, every optimal allocation $w_* \in \bar{A}$ is such that $\Sigma_{j=1}^{i} w_*^{(j)} \ge \Sigma_{j=1}^{i} x^{(j)}$;

c) any optimal policy for the infinite horizon total expected discounted cost problem has properties a) and b).

Proof: We follow the same proof as in Theorem 3.5. Under assumption (3.25), the second inequality on the RHS of (3.19) is strict. We then obtain $G_k(x, v_*) < G_k(x, w_*)$, which contradicts the optimality of w_*. ■

We next use Theorem 3.5 to show that the search of optimal policies is reduced, as stated at the beginning of the section. We first define

$$\alpha(i) \triangleq \sum_{j=1}^{i} x^{(j)} \tag{3.26}$$

$$\gamma(i) \triangleq T - \alpha(i). \tag{3.27}$$

For every initial state, a brute-force optimization over $\bar{A}$ will require the computation of $\binom{M+T-1}{M-1}$ allocation schemes (see [10]). However, if we take advantage of properties a) and b) of Theorem 3.5, the computational complexity of the optimal control algorithm will be reduced. Indeed, if we suppose that the sum of the queue lengths of the first $(i+1)$ transmitters is not less than T ($\alpha(i+1) \ge T$), then by Theorem 3.5 a), it is optimal to allocate all of the T slots to transmitters $(1, 2, \cdots, i+1)$. Then, a brute-force optimization will require the computation of $\binom{i+T}{i}$ allocation schemes. In addition, if $\alpha(i) < T$, then by Theorem 3.5 b), it is optimal to allocate at least $\alpha(l)$ slots to the first l transmitters for each $l = 1, 2, \cdots, i$. Hence, we only need to search for the optimal additional number of slots $\sigma_*^{(1)}, \sigma_*^{(2)}, \cdots, \sigma_*^{(i)}$ to allocate to transmitter 1, transmitters 1 and 2, $\cdots$, transmitters$\langle 1, 2, \cdots, i\rangle$, respectively. This then leads to a second computational reduction of the optimal allocation scheme because $(\sigma_*^{(1)}, \cdots, \sigma_*^{(i)})$ lies in a set of size smaller than $\binom{i+T}{i}$. This is made explicit in Theorem 3.7, which also gives a formulation for the reduced equations of optimality.

Theorem 3.7:

a) Let $i < M$ be such that $\alpha(i) \le T$ and $\alpha(i+1) \ge T$. If $\sigma \triangleq (\sigma^{(1)}, \sigma^{(2)} - \sigma^{(1)}, \cdots, \sigma^{(i+1)} - \sigma^{(i)}, 0, \cdots, 0)$ and $y \triangleq (0, \cdots, 0, x^{(i+2)}, \cdots, x^{(M)})$, then

$$V_k(x) = \min_{\sigma}\{G_k(y, \sigma)\} \tag{3.28}$$

subject to the constraints

$$\sigma^{(i+1)} = \gamma(i+1) \tag{3.29}$$

$$0 \le \sigma^{(l)} \le x^{(l+1)} + \sigma^{(l+1)}, \qquad 1 \le l \le i. \tag{3.30}$$

Furthermore, if σ_* minimizes the RHS of (3.28), then

$$w_* \triangleq x + \sigma_* - y \tag{3.31}$$

is an optimal allocation, i.e., $V_k(x) = G_k(x, w_*)$.

b) Equations (3.28)–(3.31) hold when $\alpha(M) \le T$. This reduced set of optimality equations is exactly the same when $\alpha(M-1) < T$ and $\alpha(M) \ge T$.

c) Properties a) and b) hold for the infinite horizon problem, i.e., with y and σ as defined in a):

$$V(x) = \min_{\sigma}\{G(y, \sigma)\} \tag{3.32}$$

subject to (3.29)–(3.31).

Proof: We prove a) only as the proof of b) and c) are conducted in a similar manner. By Theorem 3.5, there is no loss of optimality in restricting attention to allocations $w \in \bar{A}$ such that

$$\sum_{j=1}^{l} w^{(j)} \ge \sum_{j=1}^{l} x^{(j)}, \qquad 1 \le l \le i \tag{3.33}$$

$$\sum_{j=1}^{i+1} w^{(j)} = T. \tag{3.34}$$

Then there exist nonnegative integers $\sigma^{(1)}, \cdots, \sigma^{(i)}$ such that

$$\sum_{j=1}^{l} w^{(j)} = \sum_{j=1}^{l} x^{(j)} + \sigma^{(l)}, \qquad 1 \le l \le i. \tag{3.35}$$

If we define $\sigma^{(i+1)} \triangleq \gamma(i+1)$ and use (3.34)–(3.35), we then obtain

$$w^{(l)} = \begin{cases} x^{(l)} + \sigma^{(l)} & \text{if } l = 1 \\ x^{(l)} + \sigma^{(l)} - \sigma^{(l-1)} & \text{if } 1 < l \le i+1 \\ 0 & \text{if } i+1 < l \le M. \end{cases} \tag{3.36}$$

The constraints on the RHS of (3.30) follow directly from (3.36) because $w^{(l)} \ge 0$.

We next verify (3.28). With y and σ as defined in statement a) of Theorem 3.7, we write (3.36) in the vector form

$$w = x + \sigma - y. \tag{3.37}$$

From the definitions of $L(x, w)$ and $G_k(x, w)$ [cf. (2.6) and (3.2)], one checks that

$$G_k(x, x + \sigma - y) = G_k(y, \sigma). \tag{3.38}$$

Since

$$V_k(x) = \min_{w \in \bar{A}}\{G_k(x, w)\}, \tag{3.39}$$

(3.28) is now immediate from (3.37) and (3.38). Finally, (3.31) follows from (3.28), (3.37), and (3.39). ■

To compute optimal policies, we first partition the state space into sets $S_0 = \{x\colon \alpha(1) \geq T\}$, $S_i = \{x\colon \alpha(i) < T, \alpha(i+1) \geq T\}$ for $1 \leq i \leq M-1$, and $S_M = \{x\colon \alpha(M) < T\}$. Optimal policies over S_0 are trivial; all of the T slots are allocated to transmitter 1 [cf. theorem 3.5 a)]. We then use the simplifications of Theorem 3.7 to compute an optimal allocation for each state in S_i, $1 \leq i \leq M$.

We believe that these properties may reduce the computational difficulty of the optimization problem, even though the computation of an optimal policy still requires knowledge of the value function for all states.

IV. Further Results for $M = 2$

In this section, we let $M = 2$ and show that $V_k(x)$ and $V(x)$ are convex and submodular in x, and that optimal allocation policies have (as in the case of equal holding costs) the following property: if w_* is an optimal allocation when the state is x, then either w_* or $w_* + e_i - e_j$ ($i \neq j$) is optimal when the state is $x + e_i$. This property is stronger than monotonicity in the state. It implies that optimal policies possess the threshold property, that is, for every state $x^{(1)}$ of transmitter one, there exist an allocation $v(x^{(1)}) \in \bar{A}$ and a threshold state $t(x^{(1)})$ such that $v(x^{(1)})$ is optimal for all states $(x^{(1)}, x^{(2)})$ for which $x^{(2)} \geq t(x^{(1)})$. Combining these properties with those of Section III, we finally show that for some $x_*^{(2)}$, the optimization is reduced to the calculation of optimal allocations for states in the set $\{(0, x^{(2)})\colon 0 \leq x^{(2)} \leq x_*^{(2)}\}$. Moreover, by the first above-mentioned property, optimal allocations over this set are calculated recursively using an optimal allocation for state $(0,0)$, thus reducing the computational complexity of the optimal flow control algorithm significantly. We first define convexity, submodularity, and prove some preliminary lemmas.

The following definitions are as in [2]. Let g be a function of $x = (x_1, \cdots, x_n)$, where $x_1, \cdots, x_n \in Z$. We say g is convex in Z^n if

$$g(x + e_i + e_j) - g(x + e_i) \geq g(x + e_j) - g(x) \qquad \forall 1 \leq i, j \leq n \quad (4.1)$$

and g is submodular in Z^n if

$$g(x + e_i - e_j) - g(x) \geq g(x - e_j) - g(x - e_i) \qquad \forall 1 \leq i \neq j \leq n. \quad (4.2)$$

The following two lemmas are proved in [3] (see Lemma 4.1 and Lemma 4.2, respectively). We restate them here for completeness.

Lemma 4.1: Let g be a function of $x = (x_1, x_2) \in Z_+^2$. If g is monotonically nondecreasing in each variable and convex, then $g(x_1^+, x_2^+)$ is convex in Z^2, i.e., for any i, j,

$$g(x^+) - g([x - e_i]^+) \geq g([x - e_j]^+) - g([x - e_i - e_j]^+). \quad (4.3)$$

Lemma 4.2: Let g, a function of $x = (x_1, x_2) \in Z_+^2$, satisfy the three properties:

a) $g(x_1, x_2)$ is monotonically nondecreasing in each variable,

b) $g(x_1, x_2)$ is convex in each variable,

c) $g(x_1, x_2)$ is submodular.

Then $g(x_1^+, x_2^+)$ is submodular in Z^2, i.e., for $i \neq j$,

$$g([x + e_i - e_j]^+) - g(x^+) \geq g([x - e_j]^+) - g([x - e_i]^+). \quad (4.4)$$

In the next theorem, we prove that $V_k(x)$ is convex and submodular in x, and that it satisfies the first property of the optimal control policy stated at the beginning of the section.

Theorem 4.3: a) Let $V_k(x) = G_k(x, w_*)$ for some $w_* \in \bar{A}$. Then for $i = 1, 2$,

$$V_k(x + e_i) = \begin{cases} \min_{j=1,2} G_k(x + e_i, w_* + e_i - e_j) \\ \qquad \text{if } w_*^{(i)} < T \\ G_k(x + e_i, w_*) \\ \qquad \text{if } w_*^{(i)} = T. \end{cases} \quad (4.5)$$

b) $V_k(x)$ is convex.

c) $V_k(x)$ is submodular.

Proof: The theorem is trivial for $k = 0$. Assuming that $V_k(x)$ is convex and submodular, we shall show that a), b), and c) hold for $V_{k+1}(x)$; thus, by way of induction, the theorem holds for all $k \geq 0$.

Proof of a): Suppose $V_k(x)$ is convex and submodular. Combining these two properties with the monotonicity of $V_k(x)$ in each component (cf. Lemma 3.1), then by Lemma 4.1 and Lemma 4.2, respectively,

P3) $V_k(x^+)$ is convex in Z^2,

P4) $V_k(x^+)$ is submodular in Z^2.

We will invoke these properties later in the proof. Let

$$V_{k+1}(x) = G_{k+1}(x, w_*) = L(x, w_*) + \beta E\left[V_k([Y + x - w_*]^+)\right]$$

for some $w_* \in \bar{A}$. For $j = 1, 2$, define

$$B(j) \triangleq \{w \in \bar{A}\colon w^{(j)} \leq w_*^{(j)}\}. \quad (4.6)$$

Our goal is to prove the following two inequalities:

$$G_{k+1}(x + e_i, w_*) \leq G_{k+1}(x + e_i, w) \qquad \forall w \in B(i); \quad (4.7)$$

if $w_*^{(i)} < T$,

$$G_{k+1}(x + e_i, w_* + e_i - e_j) \leq G_{k+1}(x + e_i, w + e_i - e_j) \qquad \forall w \in B(j). \quad (4.8)$$

Suppose for now that (4.7)–(4.8) hold. If $w_*^{(i)} = T$, then (4.7) implies that $V_{k+1}(x + e_i) = G_{k+1}(x + e_i, w_*)$. For the other case, $w_*^{(i)} < T$, pick $w \in \bar{A}$. If $w^{(i)} \leq w_*^{(i)}$, then by (4.7), $G_{k+1}(x + e_i, w_*) \leq G_{k+1}(x + e_i, w)$. On the other hand, if $w^{(i)} > w_*^{(i)}$, then for the other index j, $w^{(j)} < w_*^{(j)}$ because $w_*, w \in \bar{A}$. Hence, $w - e_i + e_j \in B(j)$, and by (4.8), $G_{k+1}(x + e_i, w_* + e_i - e_j) \le G_{k+1}(x + e_i, w)$. We then conclude that in all cases,

$$V_{k+1}(x+e_i) = \begin{cases} \min_{j=1,2} G_{k+1}(x+e_i, w_* + e_i - e_j) \\ \qquad \text{if } w_*^{(i)} < T \\ G_{k+1}(x+e_i, w_*) \\ \qquad \text{if } w_*^{(i)} = T, \end{cases} \tag{4.9}$$

and thus a) holds for $k+1$. We next proceed to prove (4.7)–(4.8). By a straightforward calculation,

$$\begin{aligned} G_{k+1}&(x+e_i, w+e_i-e_j) - G_{k+1}(x,w) \\ &= c^{(j)}\Big\{E\big[(Y^{(j)}+x^{(j)}+1-w^{(j)})^+\big] \\ &\quad - E\big[(Y^{(j)}+x^{(j)}-w^{(j)})^+\big]\Big\} \\ &\quad + \beta\Big\{E\big[V_k([Y+x-w+e_j]^+)\big] \\ &\quad - E\big[V_k([Y+x-w]^+)\big]\Big\}, \end{aligned} \tag{4.10}$$

and by the convexity of $(\cdot)^+$,

$$\begin{aligned} &E\big[(Y^{(j)}+x^{(j)}+1-w^{(j)})^+ - (Y^{(j)}+x^{(j)}-w^{(j)})^+\big] \\ &\ge E\big[(Y^{(j)}+x^{(j)}+1-w_*^{(j)})^+ - (Y^{(j)}+x^{(j)}-w_*^{(j)})^+\big] \\ &\qquad \forall w^{(j)} \le w_*^{(j)}. \end{aligned} \tag{4.11}$$

If we show that $\forall w^{(j)} \le w_*^{(j)}$

$$\begin{aligned} E\big[V_k&([Y+x-w+e_j]^+) - V_k([Y+x-w]^+)\big] \\ &\ge E\big[V_k([Y+x-w_*+e_j]^+) \\ &\quad - V_k([Y+x-w_*]^+)\big], \end{aligned} \tag{4.12}$$

then combining (4.10)–(4.12), we obtain $\forall w^{(j)} \le w_*^{(j)}$

$$\begin{aligned} G_{k+1}&(x+e_i, w+e_i-e_j) - G_{k+1}(x,w) \\ &\ge G_{k+1}(x+e_i, w_*+e_i-e_j) - G_{k+1}(x,w_*). \end{aligned} \tag{4.13}$$

Using the optimality of w_*, i.e., $G_{k+1}(x,w) \ge G_{k+1}(x,w_*)$ for all $w \in \bar{A}$, and (4.13), we then obtain (4.7)–(4.8). To complete the proof, it remains to demonstrate (4.12). Without loss of generality, assume $j = 1$ and let $w^{(1)} = w_*^{(1)} - m$, $w^{(2)} = w_*^{(2)} + m$. If $m = 0$, (4.12) holds with equality; so suppose $m \ge 1$. For any realization $(Y^{(1)}, Y^{(2)})$, let $z_1 = Y^{(1)} + x^{(1)} - w^{(1)}$ and $z_2 = Y^{(2)} + x^{(2)} - w^{(2)}$. Then by repeated applications of P4), we obtain

$$\begin{aligned} V_k&([z_1+1]^+, z_2^+) - V_k(z_1^+, z_2^+) \\ &\ge V_k(z_1^+, [z_2+1]^+) - V_k([z_1-1]^+, [z_2+1]^+) \\ &\ge V_k([z_1-1]^+, [z_2+2]^+) \\ &\quad - V_k([z_1-2]^+, [z_2+2]^+) \\ &\ge \cdots \ge V_k([z_1-m+1]^+, [z_2+m]^+) \\ &\quad - V_k([z_1-m]^+, [z_2+m]^+) \end{aligned} \tag{4.14}$$

or, equivalently,

$$\begin{aligned} V_k&([Y+x-w+e_1]^+) - V_k([Y+x-w]^+) \\ &\ge V_k([Y+x-w_*+e_1]^+) - V_k([Y+x-w_*]^+). \end{aligned} \tag{4.15}$$

Taking expectations on both sides of (4.15), and repeating the above argument for $j = 2$, we get (4.12). This completes the proof of a).

Proof of b): We take advantage of the structure of the optimal policy to prove that $V_{k+1}(x)$ is convex. Since the optimal policy takes different forms depending on whether $x^{(1)} \ge T$ or not, we break the proof in three cases: $x^{(1)} \ge T$, $x^{(1)} \le T-2$, and the boundary case $x^{(1)} = T-1$. Throughout this proof, we let

$$\begin{aligned} \Delta \triangleq V_{k+1}&(x+e_i+e_j) - V_{k+1}(x+e_i) \\ &- V_{k+1}(x+e_j) + V_{k+1}(x), \end{aligned} \tag{4.16}$$

and we remind the reader that our objective is to show that $\Delta \ge 0$.

Case 1: $x^{(1)} \ge T$.

Applying Theorem 3.5a), we obtain

$$\begin{aligned} V_{k+1}(x+e_i+e_j) &= G_{k+1}(x+e_i+e_j, Te_1) \\ V_{k+1}(x+e_i) &= G_{k+1}(x+e_i, Te_1) \\ V_{k+1}(x+e_j) &= G_{k+1}(x+e_j, Te_1) \\ V_{k+1}(x) &= G_{k+1}(x, Te_1). \end{aligned}$$

One finds that

$$\begin{aligned} \Delta = \beta E\big[&V_k(Y+x+e_i+e_j-Te_1) \\ &- V_k(Y+x+e_i-Te_1) \\ &- V_k(Y+x+e_j-Te_1) + V_k(Y+x-Te_1)\big] \end{aligned}$$

so that by the convexity of $V_k((\cdot)^+)$ [cf. property P3)], $V_{k+1}(x)$ is convex.

Case 2: $x^{(1)} \le T-2$.

Let $V_{k+1}(x+e_i+e_j) = G_{k+1}(x+e_i+e_j, w_*)$ for some $w_* \in \bar{A}$. Then by the remark following the proof of Theorem 3.7,

$$V_{k+1}(x+e_i+e_j) = G_{k+1}(0, \sigma_*) \tag{4.17}$$

for some $\sigma_* \triangleq (\sigma_*^{(1)}, T - x^{(1)} - x^{(2)} - 2 - \sigma_*^{(1)})$ such that

$$w_* = x + e_i + e_j + \sigma_*. \tag{4.18}$$

To simplify the notation, we define $\sigma_*^{(2)} \triangleq T - x^{(1)} - x^{(2)} - 2 - \sigma_*^{(1)}$. Our immediate goal is to show that

$$V_{k+1}(x+e_i) = G_{k+1}(0, \sigma_* + e_m) \tag{4.19}$$

for some m. Using the optimality of w_* when the state is $x + e_i + e_j$ and applying a), we get

$$V_{k+1}(x+e_i) = G_{k+1}(x+e_i, w_* - e_j + e_m) \tag{4.20}$$

for some m such that $w_* - e_j + e_m$ is well defined, i.e., $m = j$ if $w_*^{(j)} = 0$. By Theorem 3.7 a) and b), we also obtain

$$V_{k+1}(x+e_i) = G_{k+1}(0, \delta_*) \tag{4.21}$$

for some $\delta_* \triangleq (\delta_*^{(1)}, T - x^{(1)} - x^{(2)} - 1 - \delta_*^{(1)})$ such that

$$w_* - e_j + e_m = x + e_i + \delta_*. \tag{4.22}$$

From (4.18) and (4.22), we deduce that $\delta_* = \sigma_* + e_m$. Equation (4.19) is now immediate from (4.21). We next establish an inequality that will be useful in the sequel. From (4.19), $G_{k+1}(0, \sigma_* + e_m) \le G_{k+1}(0, \sigma_* + e_l)$ for all l or, equivalently,

$$\begin{aligned} c^{(m)}E&\left[(Y^{(m)} - \sigma_*^{(m)} - 1)^+ - (Y^{(m)} - \sigma_*^{(m)})^+\right] \\ &+ \beta E\left[V_k([Y - \sigma_* - e_m]^+)\right] \\ &\le c^{(l)}E\left[(Y^{(l)} - \sigma_*^{(l)} - 1)^+ - (Y^{(l)} - \sigma_*^{(l)})^+\right] \\ &\quad + \beta E\left[V_k([Y - \sigma_* - e_l]^+)\right] \qquad \forall l. \end{aligned} \tag{4.23}$$

By a similar argument as above, we also obtain

$$V_{k+1}(x+e_j) = G_{k+1}(0, \sigma_* + e_n) \tag{4.24}$$

$$V_{k+1}(x) = G_{k+1}(x, \sigma_* + e_n + e_l) \tag{4.25}$$

for some n and l. We now show that $\Delta \ge 0$ by considering all the possible combinations of m, n, and l. In each of the cases below, the calculation of Δ is based on (4.17), (4.19), (4.24), and (4.25).

i) $m = n = l$.

In this case,

$$\begin{aligned} \Delta = c^{(m)}&\Big\{E\Big[(Y^{(m)} - \sigma_*^{(m)})^+ - 2(Y^{(m)} - \sigma_*^{(m)} - 1)^+ \\ &+ (Y^{(m)} - \sigma_*^{(m)} - 2)^+\Big]\Big\} \\ &+ \beta\Big\{E\Big[V_k([Y - \sigma_*]^+) - 2V_k([Y - \sigma_* - e_m]^+) \\ &+ V_k([Y - \sigma_* - 2e_m]^+)\Big]\Big\}. \end{aligned}$$

The first term in { } on the RHS of the above is nonnegative by the convexity of $(\cdot)^+$, and the second term in { } is nonnegative by the convexity of $V_k((\cdot)^+)$. Thus, $\Delta \ge 0$.

ii) $m = n$ and $l \ne m$.

One checks that

$$\begin{aligned} \Delta = c^{(m)}&E\left[(Y^{(m)} - \sigma_*^{(m)})^+ - (Y^{(m)} - \sigma_*^{(m)} - 1)^+\right] \\ &- c^{(l)}E\left[(Y^{(l)} - \sigma_*^{(l)})^+ - (Y^{(l)} - \sigma_*^{(l)} - 1)^+\right] \\ &+ \beta E\Big[V_k([Y - \sigma_*]^+) - 2V_k([Y - \sigma_* - e_m]^+) \\ &+ V_k([Y - \sigma_* - e_m - e_l]^+)\Big]. \end{aligned}$$

This difference can be written in the equivalent form

$$\begin{aligned} \Delta = &\Big\{\Big(c^{(l)}E\left[(Y^{(l)} - \sigma_*^{(l)} - 1)^+ - (Y^{(l)} - \sigma_*^{(l)})^+\right] \\ &+ \beta E\left[V_k([Y - \sigma_* - e_l]^+)\right]\Big) \\ &- \Big(c^{(m)}E\left[(Y^{(m)} - \sigma_*^{(m)} - 1)^+ - (Y^{(m)} - \sigma_*^{(m)})^+\right] \\ &+ \beta E\left[V_k([Y - \sigma_* - e_m]^+)\right]\Big)\Big\} \\ &+ \beta\Big\{E\Big[V_k([Y - \sigma_*]^+) - V_k([Y - \sigma_* - e_m]^+) \\ &- V_k([Y - \sigma_* - e_l]^+) \\ &+ V_k([Y - \sigma_* - e_m - e_l]^+)\Big]\Big\}. \end{aligned}$$

$\Delta \ge 0$ because the first term in { } on the RHS of the above is nonnegative by (4.23), and the second term in { } is nonnegative by the convexity of $V_k((\cdot)^+)$.

iii) $m \ne n$ and $l = m$.

By a straightforward calculation,

$$\begin{aligned} \Delta = \beta E&\Big[V_k([Y - \sigma_*]^+) - V_k([Y - \sigma_* - e_m]^+) \\ &- V_k([Y - \sigma_* - e_n]^+) \\ &+ V_k([Y - \sigma_* - e_m - e_n]^+)\Big] \end{aligned}$$

so that $\Delta \ge 0$ by the convexity of $V_k((\cdot)^+)$.

iv) $m \ne n$ and $l = n$.

In this case,

$$\begin{aligned} \Delta = c^{(m)}&E\left[(Y^{(m)} - \sigma_*^{(m)})^+ - (Y^{(m)} - \sigma_*^{(m)} - 1)^+\right] \\ &- c^{(l)}E\left[(Y^{(l)} - \sigma_*^{(l)} - 1)^+ - (Y^{(l)} - \sigma_*^{(l)} - 2)^+\right] \\ &+ \beta E\Big[V_k([Y - \sigma_*]^+) - V_k([Y - \sigma_* - e_m]^+) \\ &- V_k([Y - \sigma_* - e_l]^+) + V_k([Y - \sigma_* - 2e_l]^+)\Big]. \end{aligned}$$

We write this difference in the equivalent form

$$\begin{aligned} \Delta = &\Big\{\Big(c^{(l)}E\left[(Y^{(l)} - \sigma_*^{(l)} - 1)^+ - (Y^{(l)} - \sigma_*^{(l)})^+\right] \\ &+ \beta E\left[V_k([Y - \sigma_* - e_l]^+)\right]\Big) \\ &- \Big(c^{(m)}E\left[(Y^{(m)} - \sigma_*^{(m)} - 1)^+ - (Y^{(m)} - \sigma_*^{(m)})^+\right] \\ &+ \beta E\left[V_k([Y - \sigma_* - e_m]^+)\right]\Big)\Big\} \\ &+ c^{(l)}\Big\{E\Big[(Y^{(l)} - \sigma_*^{(l)})^+ - 2(Y^{(l)} - \sigma_*^{(l)} - 1)^+ \\ &+ (Y^{(l)} - \sigma_*^{(l)} - 2)^+\Big]\Big\} \\ &+ \beta\Big\{E\Big[V_k([Y - \sigma_*]^+) - 2V_k([Y - \sigma_* - e_l]^+) \\ &+ V_k([Y - \sigma_* - 2e_l]^+)\Big]\Big\}. \end{aligned}$$

The first term in { } on the RHS of the above is nonnegative by (4.23), the second term in { } is nonnegative by the convexity of $(\cdot)^+$, and the third term in { } is nonnegative by the convexity of $V_k((\cdot)^+)$. Thus, $\Delta \geq 0$. This finishes the proof of Case 2.

Case 3: $x^{(1)} = T - 1$.

We only consider the case $i = j = 1$ in (4.16) because all of the other possible combinations of i and j are covered in Case 2 above. By Theorem 3.5 a),

$$V_{k+1}(x + 2e_1) = G_{k+1}(x + 2e_1, Te_1)$$
$$V_{k+1}(x + e_1) = G_{k+1}(x + e_1, Te_1). \quad (4.26)$$

Using (4.26) and applying a), we obtain

$$V_{k+1}(x) = G_{k+1}(x, Te_1 - e_1 + e_m)$$

for some m. If $m = 1$, then

$$\Delta = c^{(1)}\left\{1 - E\left[Y^{(1)} - (Y^{(1)} - 1)^+\right]\right\} + \beta\{E[V_k(Y + x - (T-2)e_1) - 2V_k(Y+x-(T-1)e_1) + V_k([Y+x-Te_1]^+)]\}.$$

Since for any realization $Y^{(1)}, Y^{(1)} - (Y^{(1)} - 1)^+ \leq 1$, the first term in { } on the RHS of the above is nonnegative. In addition, since the second term in { } is nonnegative by the convexity of $V_k((\cdot)^+)$, then $\Delta \geq 0$. If $m = 2$, then

$$\Delta = \left\{c^{(1)}\text{-}c^{(2)}E\left[(Y^{(2)} + x^{(2)}) - (Y^{(2)} + x^{(2)} - 1)^+\right]\right\} + \beta\left\{E\left[V_k(Y + x - (T-2)e_1) - 2V_k(Y + x - (T-1)e_1) + V_k\left([Y + x - (T-1)e_1 - e_2]^+\right)\right]\right\}. \quad (4.27)$$

Since for any realization $Y^{(2)}$, $(Y^{(2)} + x^{(2)}) - (Y^{(2)} + x^{(2)} - 1)^+ \leq 1$, and $c^{(1)} \geq c^{(2)}$, then the first term in { } on the RHS of (4.27) is nonnegative. For Δ to be nonnegative, it is then sufficient to show that the second term in { } on the RHS of (4.27) is nonnegative. In fact, for any realization $(Y^{(1)}, Y^{(2)})$, we claim that

$$V_k(Y^{(1)} + 1, Y^{(2)} + x^{(2)}) - V_k(Y^{(1)}, Y^{(2)} + x^{(2)}) \geq V_k(Y^{(1)}, Y^{(2)} + x^{(2)}) - V_k\left(Y^{(1)}, [Y^{(2)} + x^{(2)} - 1]^+\right). \quad (4.28)$$

The nonnegativity of the second term in { } on the RHS of (4.27) then immediately follows from (4.28). We break the proof of (4.28) into cases. If $Y^{(2)} + x^{(2)} = 0$, then (4.28) is immediate by the monotonicity of $V_k(\cdot)$ in the first component. We next consider the other possibility, $Y^{(2)} + x^{(2)} \geq 1$. Suppose first that $Y^{(1)} = 0$. Then by the convexity of $V_k(\cdot)$ in the second component,

$$V_k(0, Y^{(2)} + x^{(2)}) - V_k(0, Y^{(2)} + x^{(2)} - 1) \leq V_k(0, Y^{(2)} + x^{(2)} + 1) - V_k(0, Y^{(2)} + x^{(2)}). \quad (4.29)$$

Since by Lemma 3.3, $V_k(0, Y^{(2)} + x^{(2)} + 1) \leq V_k(1, Y^{(2)} + x^{(2)})$, then (4.28) follows from this observation and (4.29). Finally, if $Y^{(1)} \geq 1$, then by the convexity of $V_k(\cdot)$ in the first component,

$$V_k(Y^{(1)} + 1, Y^{(2)} + x^{(2)}) - V_k(Y^{(1)}, Y^{(2)} + x^{(2)}) \geq V_k(Y^{(1)}, Y^{(2)} + x^{(2)}) - V_k(Y^{(1)} - 1, Y^{(2)} + x^{(2)}). \quad (4.30)$$

By Lemma 3.3, $V_k(Y^{(1)} - 1, Y^{(2)} + x^{(2)}) \leq V_k(Y^{(1)}, Y^{(2)} + x^{(2)} - 1)$. Combining this result with (4.30) yields (4.28). This concludes the proof of Case 3, and hence b).

Proof of c): As in the proof of b), we take advantage of the structural properties of the optimal policy to show that $V_{k+1}(x)$ is submodular. Since the optimal policy takes different forms depending on whether $x^{(1)} \geq T$ or not, we break the proof into three cases: $x^{(1)} > T$, $x^{(1)} < T$, and the boundary case $x^{(1)} = T$. Throughout this proof, we let

$$D \triangleq V_{k+1}(x + e_i - e_j) - V_{k+1}(x - e_j) - V_{k+1}(x) + V_{k+1}(x - e_i) \quad (4.31)$$

and remind the reader that our objective is to show that $D \geq 0$ whenever $i \neq j$.

Case 1: $x^{(1)} > T$. Applying Theorem 3.5a), we obtain

$$V_{k+1}(x + e_i - e_j) = G_{k+1}(x + e_i - e_j, Te_1)$$
$$V_{k+1}(x - e_j) = G_{k+1}(x - e_j, Te_1)$$
$$V_{k+1}(x - e_i) = G_{k+1}(x - e_i, Te_1)$$
$$V_{k+1}(x) = G_{k+1}(x, Te_1)$$

so that

$$D = c^{(i)}\left\{1 - E\left[(Y^{(i)} + x^{(i)}) - (Y^{(i)} + x^{(i)} - 1)^+\right]\right\} + \beta\left\{E\left[V_k\left([Y + x - Te_1 + e_i - e_j]^+\right) - V_k\left([Y + x - Te_1 - e_j]^+\right) - V_k\left([Y+x-Te_1]^+\right) + V_k\left([Y+x-Te_1-e_i]^+\right)\right]\right\}.$$

Since for any realization $Y^{(i)}$, $(Y^{(i)} + x^{(i)}) - (Y^{(i)} + x^{(i)} - 1)^+ \leq 1$, the first term in { } on the RHS of the above is nonnegative. In addition, the second term in { } on the RHS of the above is nonnegative by the submodularity of $V_k((\cdot))$ [cf. property P4)]. Thus, $D \geq 0$.

Case 2: $x^{(1)} < T$ or ($x^{(1)} = T$ and $i = 2$).

By Theorem 3.7 a) and b), there exists $\sigma_* = (\sigma_*^{(1)}, \sigma_*^{(2)})$ such that

$$V_{k+1}(x + e_i - e_j) = G_{k+1}(0, \sigma_*). \quad (4.32)$$

Using (4.32) and applying a), then by the argument in the first paragraph of the proof of Case 2b),

$$V_{k+1}(x - e_j) = G_{k+1}(0, \sigma_* + e_m) \quad (4.33)$$
$$V_{k+1}(x) = G_{k+1}(0, \sigma_* + e_m - e_n) \quad (4.34)$$
$$V_{k+1}(x - e_i) = G_{k+1}(0, \sigma_* + e_m - e_n + e_l) \quad (4.35)$$

for some m, n, and l. In addition, by the optimality of $\sigma_* + e_m$ when the state is $x - e_j$, m is such that $G_{k+1}(0, \sigma_* + e_m) \leq G_{k+1}(0, \sigma_* + e_l)$ for all l or, equivalently,

$$c^{(m)}E\left[(Y^{(m)}-\sigma_*^{(m)}-1)^+-(Y^{(m)}-\sigma_*^{(m)})^+\right]+\beta E\left[V_k([Y-\sigma_*-e_m]^+)\right]\le c^{(l)}E\left[(Y^{(l)}-\sigma_*^{(l)}-1)^+-(Y^{(l)}-\sigma_*^{(l)})^+\right]+\beta E\left[V_k([Y-\sigma_*-e_l]^+)\right]\quad\forall l. \tag{4.36}$$

We next show that $D\ge 0$ by considering all the possible combinations of m, n, and l. In each of the cases below, the calculation of D is based on (4.32)–(4.35).

i) $m=n=l$.

D is equal to zero in this case.

ii) $m=n$ and $l\ne m$.

By a straightforward calculation,

$$D=\left\{c^{(l)}E\left[(Y^{(l)}-\sigma_*^{(l)}-1)^+-(Y^{(l)}-\sigma_*^{(l)})^+\right]+\beta E\left[V_k([Y-\sigma_*-e_l]^+)\right]\right\}-\left\{c^{(m)}E\left[(Y^{(m)}-\sigma^{(m)}{}_*-1)^+-(Y^{(m)}-\sigma_*^{(m)})^+\right]+\beta E\left[V_k([Y-\sigma_*-e_m]^+)\right]\right\}$$

so that $D\ge 0$ by (4.36).

iii) $m\ne n$ and $l=m$.

One checks that

$$D=c^{(m)}\left\{E\left[(Y^{(m)}-\sigma_*^{(m)})^+-2(Y^{(m)}-\sigma_*^{(m)}-1)^++(Y^{(m)}-\sigma_*^{(m)}-2)^+\right]\right\}+\beta\left\{E\left[V_k([Y-\sigma_*]^+)-V_k([Y-\sigma_*-e_m]^+)-V_k([Y-\sigma_*-e_m+e_n]^+)+V_k([Y-\sigma_*-2e_m+e_n]^+)\right]\right\}.$$

The first term in { } on the RHS of the above is nonnegative by the convexity of $(\cdot)^+$, while the second term in { } is nonnegative by the submodularity of $V_k((\cdot)^+)$. Thus, $D\ge 0$.

iv) $m\ne n$ and $l=n$.

In this case,

$$D=c^{(m)}E\left[(Y^{(m)}-\sigma_*^{(m)})^+-(Y^{(m)}-\sigma_*^{(m)}-1)^+\right]-c^{(l)}E\left[(Y^{(l)}-\sigma_*^{(l)}+1)^+-(Y^{(l)}-\sigma_*^{(l)})^+\right]+\beta E\left[V_k([Y-\sigma_*]^+)-V_k([Y-\sigma_*-e_m-e_l]^+)\right].$$

We write D in the equivalent form

$$D=\left\{\left(c^{(l)}E\left[(Y^{(l)}-\sigma_*^{(l)}-1)^+-(Y^{(l)}-\sigma_*^{(l)})^+\right]+\beta E\left[V_k([Y-\sigma_*-e_l]^+)\right]\right)-\left(c^{(m)}E\left[(Y^{(m)}-\sigma_*^{(m)}-1)^+-(Y^{(m)}-\sigma_*^{(m)})^+\right]+\beta E\left[V_k([Y-\sigma_*-e_m]^+)\right]\right)\right\}+\left\{c^{(l)}E\left[2(Y^{(l)}-\sigma_*^{(l)})^+-(Y^{(l)}-\sigma_*^{(l)}+1)^+-(Y^{(l)}-\sigma_*^{(l)}-1)^+\right]\right\}+\beta\left\{E\left[V_k([Y-\sigma_*]^+)-V_k([Y-\sigma_*-e_m-e_l]^+)-V_k([Y-\sigma_*-e_l]^+)+V_k([Y-\sigma_*-e_m]^+)\right]\right\}.$$

The first term in { } on the RHS of the above is nonnegative by (4.36), the second term in { } is nonnegative by the convexity of $(\cdot)^+$, and the third term in { } is nonnegative by the convexity of $V_k((\cdot)^+)$. Thus, $D\ge 0$, and the proof of Case 2 is complete.

Case 3: $x^{(1)}=T$ and $i=1$.

By Theorem 3.5a),

$$V_{k+1}(x+e_1-e_2)=G_{k+1}(x+e_1-e_2,Te_1) \tag{4.37}$$

$$V_{k+1}(x-e_2)=G_{k+1}(x-e_2,Te_1) \tag{4.38}$$

$$V_{k+1}(x)=G_{k+1}(x,Te_1). \tag{4.39}$$

Using (4.39) and applying Theorem 4.3a), we obtain

$$V_{k+1}(x-e_1)=G_{k+1}(x-e_1,Te_1-e_1+e_m) \tag{4.40}$$

for some m. If $m=1$, then calculating D using (4.37)–(4.40), we find that

$$D=c^{(1)}\left\{1-E\left[Y^{(1)}+(Y^{(1)}-1)^+\right]\right\}+\beta\left\{E\left[V_k([Y+x-Te_1+e_1-e_2]^+)-V_k([Y+x-Te_1-e_2]^+)-V_k(Y+x-Te_1)+V_k([Y+x-Te_1-e_1]^+)\right]\right\}.$$

The first term in { } on the RHS of the above is nonnegative because for any realization $Y^{(1)}$, $Y^{(1)}-(Y^{(1)}-1)^+\le 1$. Since by the submodularity of $V_k((\cdot)^+)$ the second term in { } on the RHS of the above is nonnegative also, then $D\ge 0$. If $m=2$, then

$$D=\left\{c^{(1)}-c^{(2)}E\left[(Y^{(2)}+x^{(2)})-(Y^{(2)}+x^{(2)}-1)^+\right]\right\}+\beta\left\{E\left[V_k([Y+x-Te_1+e_1-e_2]^+)-V_k(Y+x-Te_1)\right]\right\}. \tag{4.41}$$

Since $E[(Y^{(2)}+x^{(2)})-(Y^{(2)}+x^{(2)}-1)^+]\le 1$ and $c^{(1)}\ge c^{(2)}$, the first term in { } on the RHS of (4.41) is nonnegative. Then $D\ge 0$ if we show that the second term on the RHS of (4.41) is nonnegative. In fact, we claim that for any realization $(Y^{(1)}, Y^{(2)})$,

$$V_k\left(Y^{(1)} + 1, [Y^{(2)} + x^{(2)} - 1]^+\right) \geq V_k(Y^{(1)}, Y^{(2)} + x^{(2)}). \tag{4.42}$$

The nonnegativity of the second term in { } on the RHS of (4.41) is then immediate from (4.42). To prove (4.42), we first suppose that $Y^{(2)} + x^{(2)} = 0$. In this case, (4.42) follows from the monotonicity of $V_k(\cdot)$ in the first component. If $Y^{(2)} + x^{(2)} \geq 1$, (4.42) follows from Lemma 3.3. This completes the proof of c) and the theorem. ■

As for the finite horizon, we have the following results for the infinite horizon when $\beta < 1$.

Corollary 4.4: a) $V(x)$ is convex and submodular. b) Let $V(x) = G(x, w_*)$ for some $w_* \in \bar{A}$. Then

$$V(x + e_i) = \begin{cases} \min_{j=1,2} G(x + e_i, w_* + e_i - e_j) & \\ \quad \text{if } w_*^{(i)} < T & \\ G(x + e_i, w_*) & \\ \quad \text{if } w_*^{(i)} = T. & \end{cases} \tag{4.43}$$

Proof: As the limit of convex (submodular) functions is convex (submodular), a) is immediate from Theorem 4.3b) and c). The proof of b) follows in the exact same manner as the proof Theorem 4.3a) using the analogous properties of $V(x)$. ■

We next use Theorem 4.3a) and Corollary 4.4b) to strengthen the result of Corollary 3.4.

Theorem 4.5: If $w_*(x)$ denotes an optimal allocation when the state is x for either the finite or infinite horizon problem and $x^{(2)} \geq 1$, then

$$w_*(x + e_1 - e_2) = \begin{cases} w_*(x) + e_1 - e_2 & \text{if } w_*^{(2)}(x) \geq 1 \\ w_*(x) & \text{if } w_*^{(2)}(x) = 0. \end{cases} \tag{4.44}$$

Proof: If $w_*^{(2)}(x) \geq 1$, then by Corollary 3.4, $w_*(x + e_1 - e_2) = w_*(x) + e_1 - e_2$. If $w_*^{(2)}(x) = 0$, then applying Theorem 4.3a) (respectively Corollary 4.4b) for the infinite horizon case) twice, we obtain $w_*(x - e_2) = w_*(x)$ and $w_*(x - e_2 + e_1) = w_*(x - e_2)$, thus establishing (4.44). ■

As stated in Theorem 3.5a), when $x^{(1)} \geq T$, it is optimal to allocate all of the T slots to transmitter one. The determination of an optimal policy is thus reduced to the subset of the state space, $\{x\colon x^{(1)} < T\}$. Since the optimal allocation for each transmitter is monotone in its state and the allocation space is finite, then for every $x^{(1)} < T$, there exists a threshold $t(x^{(1)}) \in Z_+$ such that it is optimal to make the same allocation of slots for all states $(x^{(1)}, x^{(2)})$ for which $x^{(2)} \geq t(x^{(1)})$.

We now show that it is sufficient to calculate $t(0)$ and $w_*(0, x^{(2)})$ for all $x^{(2)} \leq t(0)$ to determine all of the optimal allocations over $\{x\colon x^{(1)} < T\}$. By Theorem 4.5, optimal allocations $w_*(x^{(1)}, x^{(2)})$ are simply derived from $w_*(0, x^{(1)} + x^{(2)})$. Indeed, if $t(0)$ and $\{w_*(0, x^{(2)}),\ x^{(2)} \leq t(0)\}$ are known, then given any $(x^{(1)}, x^{(2)})$, $w_*(x^{(1)}, x^{(2)})$ is derived as follows. If $x^{(1)} \geq T$, then $w_*(x^{(1)}, x^{(2)}) = (T, 0)$; otherwise, for $1 \leq x^{(1)} < T$, $w_*(x^{(1)}, x^{(2)})$ is obtained recursively from the known values of $w_*(0, x^{(1)} + x^{(2)})$ using (4.44). We conclude this section with the following remarks. 1) The computation of $\{w_*(0, x^{(2)})\colon x^{(2)} \geq 1\}$ is reduced using Theorem 4.3a) (Corollary 4.4b) for the infinite horizon case). Given $w_*(0,0)$, $\{w_*(0, x^{(2)})\colon x^{(2)} \geq 1\}$ are recursively computed, using the property that $w_*(0, x^{(2)} + 1)$ equals either $w_*(0, x^{(2)})$ or $w_*(0, x^{(2)}) + e_2 - e_1$. 2) The computation of an optimal policy requires knowledge of the value function for all states. 3) The computation of $t(0)$ may not result from a finite calculation; $t(0) = \infty$ is possible.

V. A Further Property of Optimal Policies for Two Stochastically Ordered Arrival Streams

In this section, we further characterize optimal policies when the message generation at the transmitter (transmitter one) of higher holding cost is stochastically larger than the message generation at the other. We prove the intuitive result: when the state of transmitter one is no lower than the state of transmitter two, it is optimal to allocate at least as many slots to transmitter one as to transmitter two. The application of this result is a reduction by a half of the allocation space to determine $w_*(0,0)$.

Theorem 5.1:

a) If $Y^{(1)} \overset{st}{\geq} Y^{(2)}$ and $x^{(1)} \geq x^{(2)}$, then there exists $w_* \in \bar{A}$ satisfying $w_*^{(1)} \geq w_*^{(2)}$ such that $V_k(x) = G_k(x, w_*)$.

b) Property a) above holds for the infinite horizon cost $V(x)$.

Proof: We only prove a) as the proof of b) follows the same lines using the analogous properties of $V(x)$. The method of proof is as follows: for every optimal policy that does not satisfy the property of a), there exists an optimal policy with this property.

Suppose that $V_k(x) = G_k(x, w_*)$, but $w_*^{(1)} < w_*^{(2)}$. We first consider the allocation $v_* = w_* + e_1 - e_2$ and show that

$$V_k(x) \leq G_k(x, v_*). \tag{5.1}$$

By a straightforward calculation using (2.7),

$$L(x, v_*) - L(x, w_*) = c^{(2)} P[Y^{(2)} > w_*^{(2)} - x^{(2)} - 1] - c^{(1)} P[Y^{(1)} > w_*^{(1)} - x^{(1)}]. \tag{5.2}$$

Since $w_*^{(1)} < w_*^{(2)}$ and $x^{(1)} \geq x^{(2)}$, then $w_*^{(1)} - x^{(1)} \leq w_*^{(2)} - x^{(2)} - 1$. This then implies that $P[Y^{(1)} > w_*^{(1)} - x^{(1)}] \geq P[Y^{(2)} > w_*^{(2)} - x^{(2)} - 1]$ because $Y^{(1)} \overset{st}{\geq} Y^{(2)}$. Combining this result with the hypothesis $c^{(2)} \leq c^{(1)}$, we obtain

$$L(x, v_*) \leq L(x, w_*). \tag{5.3}$$

We now show that

$$E\left[V_{k-1}([Y + x - v_*]^+)\right] \leq E\left[V_{k-1}([Y + x - w_*]^+)\right]. \tag{5.4}$$

Define

$$K \triangleq E\left[V_{k-1}([Y+x-v_*]^+) - V_{k-1}([Y+x-w_*]^+)\right] \quad (5.5)$$

$$W(x_1) \triangleq \begin{cases} V_{k-1}(x_1,0) - V_{k-1}(x_1-1,0) & \text{if } x_1 \geq 1 \\ 0 & \text{if } x_1 \leq 0 \end{cases} \quad (5.6)$$

$$U(x_2) \triangleq \begin{cases} V_{k-1}(0,x_2) - V_{k-1}(0,x_2-1) & \text{if } x_2 \geq 1 \\ 0 & \text{if } x_2 \leq 0. \end{cases} \quad (5.7)$$

We first note that both $W(\cdot)$ and $U(\cdot)$ are nonnegative functions because $V_{k-1}(\cdot,\cdot)$ is monotone in each of its arguments. Moreover, both $W(\cdot)$ and $U(\cdot)$ are nondecreasing functions because $V_{k-1}(\cdot,\cdot)$ is convex [cf. Theorem 4.3b)]. Since for all realizations $(Y^{(1)}, Y^{(2)})$ such that $Y^{(1)} \geq w_*^{(1)} + 1 - x^{(1)}$ and $Y^{(2)} \geq w_*^{(2)} - x^{(2)}$, $V_{k-1}([Y+x-v_*]^+) \leq V_{k-1}([Y+x-w_*]^+)$ by Lemma 3.3, then

$$K \leq P[Y^{(1)} \leq w_*^{(1)} - x^{(1)}]E[U(Y^{(2)} + x^{(2)} - w_*^{(2)} + 1)] - P[Y^{(2)} \leq w_*^{(2)} - x^{(2)} - 1]E[W(Y^{(1)} + x^{(1)} - w_*^{(1)})]. \quad (5.8)$$

By an earlier observation, $P[Y^{(1)} \leq w_*^{(1)} - x^{(1)}] \leq P[Y^{(2)} \leq w_*^{(2)} - x^{(2)} - 1] \leq 1$, so that

$$K \leq E[U(Y^{(2)} + x^{(2)} - w_*^{(2)} + 1)] - E[W(Y^{(1)} + x^{(1)} - w_*^{(1)})]. \quad (5.9)$$

By repeated applications of the submodularity of $V_{k-1}(\cdot,\cdot)$ [cf. Theorem 4.3c)], we obtain for $x_1 \geq 1$,

$$V_{k-1}(x_1,0) - V_{k-1}(x_1-1,0) \geq V_{k-1}(1,x_1-1) - V_{k-1}(0,x_1-1). \quad (5.10)$$

Since $V_{k-1}(1,x_1-1) \geq V_{k-1}(0,x)$ by Lemma 3.3, then

$$V_{k-1}(x_1,0) - V_{k-1}(x_1-1,0) \geq V_{k-1}(0,x_1) - V_{k-1}(0,x_1-1). \quad (5.11)$$

Thus,

$$W(x_1) \geq U(x_1) \qquad \forall x_1. \quad (5.12)$$

Since $Y^{(1)} + x^{(1)} - w^{(1)} \overset{st}{\geq} Y^{(2)} + x^{(2)} - w^{(2)} + 1$ because $Y^{(1)} \overset{st}{\geq} Y^{(2)}$, $x^{(1)} \geq x^{(2)}$, and $w^{(1)} < w^{(2)}$, then by the monotonicity $\uparrow$ of $W(\cdot)$,

$$W(Y^{(1)} + x^{(1)} - w^{(1)}) \overset{st}{\geq} W(Y^{(2)} + x^{(2)} - w^{(2)} + 1) \quad (5.13)$$

so that, by (5.12),

$$W(Y^{(1)} + x^{(1)} - w^{(1)}) \overset{st}{\geq} U(Y^{(2)} + x^{(2)} - w^{(2)} + 1). \quad (5.14)$$

Hence, $E[U(Y^{(2)} + x^{(2)} - w^{(2)} + 1) - W(Y^{(1)} + x^{(1)} - w^{(1)})] \leq 0$, and thus by (5.9), $K \leq 0$. This establishes (5.4).

Combining (5.3)–(5.4) and the optimality of w_*, then $V_k(x) = G_k(x, w_* + e_1 - e_2)$. If $w_*^{(1)} + 1 \geq w_*^{(2)} - 1$, the claim is proved; otherwise, by the same argument as above, we obtain $V_k(x) = G_k(x, w_* + 2e_1 - 2e_2)$. Proceeding in this fashion, we get $V_k(x) = G_k(x, w_* + ne_1 - ne_2)$ where n is the smallest integer such that $w_*^{(1)} + n \geq w_*^{(2)} - n$. ■

VI. Existence and Properties of Time-Average Optimal Policies

In [4], we investigated the time-average holding cost with equal weighting at the respective nodes. The time-average cost for phase length T is defined as

$$\bar{V}_T(\pi, x) \triangleq \lim_{n\to\infty} \sup n^{-1} \sum_{k=1}^{n} \sum_{j=1}^{M} E^{\pi}\left[X_k^{(j)} \mid X_0 = x\right] \quad (6.1)$$

where π is the control (allocation) policy and E^{π} denotes the corresponding expectation. By comparing this cost with the waiting time for the $G/G/1$ queue and using results in [9], we found that the necessary conditions for the existence of finite cost policies were phrased in terms of the first two moments of the arrival stream by

$$\sum_{j=1}^{M} \lambda^{(j)} < 1 \quad \text{and} \quad E\left[(\xi^{(j)})^2\right] < \infty. \quad (6.2)$$

Sufficiency of the same conditions (6.2) was proved by exhibiting a pure policy for which the time-average cost (6.1) is finite. We further showed there exist pure policies[1] such that the average cost (6.1) converges to zero as the phase length $T \to \infty$.[2] In each case, the results hold for any initial state x.

For differing holding costs, the time average generalizing (6.1) is

$$\bar{V}_T(\pi, x) \triangleq \lim_{n\to\infty} \sup n^{-1} \sum_{k=1}^{n} \sum_{j=1}^{M} c_k^{(j)} E^{\pi}\left[X_k^{(j)} \mid X_0 = x\right]. \quad (6.3)$$

Since there are a finite number M of holding cost coefficients $c^{(j)}$, we have

$$0 < \min c^{(j)} < \max c^{(j)} < \infty. \quad (6.4)$$

It follows from (6.4) that (6.3) is finite (respectively, converges to 0 with T) *iff* the same is true of (6.1). Thus, (6.2) is also necessary and sufficient for the existence of finite time-average cost when the holding costs differ among the transmitters. Finally, there exist for unequally weighted holding costs pure policies such that the cost (6.3) converges to zero as $T \to \infty$.

The characterization of optimal policies for equally weighted holding costs on all transmitters implies that there is a finite set of allocations such that an optimal

[1] These policies are not only pure, but even static.
[2] This ignores any holding costs that accumulate in the phase in which a packet arrives.

allocation can be chosen from the finite set for every discounting factor β, $0 < \beta < 1$; this is shown in [3]. We were then able to argue in [4] that there is a sequence $\{\beta_n\}$ such that $\beta_n \to 1$ with $f_n = f$ for all n, where f_n is an optimal stationary allocation policy corresponding to the discount factor β_n. Moreover, f is then a time-average optimal policy possessed of all the properties proved in [3] for discounted cost optimal policies.

When the holding cost coefficients are unequal, we can no longer assert the existence of a finite set of optimal allocations. Nevertheless, the arguments in [4] continue to apply without modification, so that we again obtain from [11] a somewhat weaker version of the results stated above. These are summarized by the following.

Theorem 6.1: Every sequence of discount factors β converging to unity has a subsequence $\{\beta_n\}$ such that corresponding optimal stationary policies $\{f_n\}$ converge, i.e., there exists an integer $N(x)$ such that $f_n(x) = f(x)$ for all $n \geq N(x)$. The stationary policy f is average cost optimal with average cost

$$g = \lim_{\beta \to 1} (1 - \beta) V^{\beta}(x) \tag{6.5}$$

where the limit does not depend on x. Furthermore, f satisfies all of the properties of the optimal discounted policies derived in Section III and in Section IV for $M = 2$. ■

VII. Conclusions

We have generalized the optimal allocation problem of [3]–[6] to the case of nonidentical holding costs at the M transmitters. We derived qualitative properties of optimal discounted and time-average policies that reduce the computational complexity of the M-dimensional optimal flow control algorithm. For $M = 2$, we established a simple relationship between optimal allocations for states x and $x + e_i$ $(i = 1, 2)$ that leads to significant computational savings in the optimal algorithm. We have been unable to prove this desirable relationship for $M > 2$, and leave this problem open for future investigation.

References

[1] D. P. Bertsekas and R. G. Gallagher, *Data Networks*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[2] F. J. Beutler and D. Teneketzis, "Routing in queueing networks under imperfect information: Stochastic dominance and thresholds," *Stochastics Stochastics Rep.*, vol. 26, pp. 81–100, 1989.

[3] R. M. Bournas, F. J. Beutler, and D. Teneketzis, "Properties of optimal hop-by-hop allocation policies in networks with multiple transmitters and linear equal holding costs," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1450–1463, Dec. 1991.

[4] ——, "Time-average and asymptotically optimal flow control policies in networks with multiple transmitters," *Ann. Oper. Res.*, vol. 35, pp. 327–355, Apr. 1992.

[5] D. H. Cansever and R. A. Milito, "Optimal hop-by-hop flow control with multiple transmitters," in *Proc. 26th IEEE Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 1858–1862.

[6] ——, "Optimal hop-by-hop flow control with multiple heterogeneous transmitters," in *Proc. 27th IEEE Conf. Decision Contr.*, Austin, TX, Dec. 1988, pp. 1291–1296.

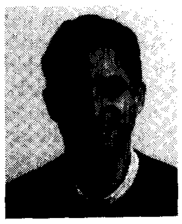
[7] M. Gerla and L. Kleinrock, "Flow control: A comparative survey," *IEEE Trans. Commun.*, vol. COM-28, pp. 533–574, 1980.

[8] ——, "Flow control protocols," in *Computer Network Architectures and Protocols*, P. E. Green, Jr., Ed. New York: Plenum, 1982, pp. 361–412.

[9] J. Kieffer and J. Wolfowitz, "On the theory of queues with many servers," *Trans. Amer. Math. Soc.*, vol. 78, pp. 147–161, 1956.

[10] L. Kleinrock, *Queueing Systems, Vol. 1, Theory*. New York: 1975.

[11] L. I. Sennott, "Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs," *Oper. Res.*, vol. 37, pp. 626–633, 1989.



Redha Bournas was born in Annaba, Algeria, on April 15, 1958. He received the B.S. degree in computer science and mathematics with honors, in 1980, the M.S. degree in electrical engineering from the University of Pittsburgh, PA, in 1981, and the Ph.D. degree in electrical engineering systems from the University of Michigan, Ann Arbor, in 1990.

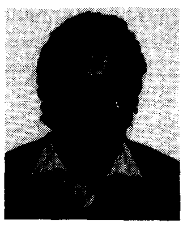
Since 1981, he has been with IBM. Currently he is an advisory engineer in Networking Systems, Research Triangle Park, NC. His research interests include multidimensional queueing systems, and performance modeling and analysis of communication networks.

Dr. Bournas was awarded a three-year IBM graduate fellowship from 1987 to 1990. He received the graduate Distinguished Achievement Award in electrical engineering systems at the University of Michigan, Ann Arbor, in April of 1990.

Frederick J. Beutler (M'53–M'67–SM'78–F'80–LF'92) received the Ph.D. degree in engineering science and mathematics from the California Institute of Technology, Pasadena, in 1957.

Since 1957, he has been on the faculty at the University of Michigan, Ann Arbor, where he is now Professor Emeritus; formerly, he served as Chair of the Electrical Engineering Systems Graduate Program, and the Graduate Program in Computer, Information, and Control Engineering.

From 1970 to 1976, Prof. Beutler was Managing Editor of the *SIAM Journal on Applied Mathematics* and served SIAM in other editorical capacities, and as a member of the governing Council of SIAM. He was Co-Chair of the *1986 International Symposium on Information Theory* and in 1981 was named Eminent Engineer by Tau Beta Pi.



Demosthenis Teneketzis (M'87) received the B.S. degree in electrical engineering from the University of Patras, Greece, in 1974, and the M.S., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1976, 1977, and 1979, respectively.

From 1979 to 1980 he worked for Systems Control Inc., Palo Alto, CA, and from 1980 to 1984 he was with Alphatech Inc., Burlington, MA. Since September 1984 he has been with the University of Michigan, Ann Arbor, where he is presently an Associate Professor of Electrical Engineering and Computer Science. In the winter 1992, he was a Visiting Professor at the Institute for Signal and Information Processing, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland. His current research interests include stochastic systems and control, decentralized systems, queueing networks, communication networks, stochastic scheduling, and resource allocation problems.