

Properties of Optimal Hop-by-Hop Allocation Policies in Networks with Multiple Transmitters and Linear Equal Holding Costs

Redha M. Bournas, Frederick J. Beutler, *Fellow, IEEE*, and Demosthenis Teneketzis, *Member, IEEE*

Abstract—We model flow control between one receiving node and its adjacent transmitting nodes in a computer network as a Markov decision problem. Given that the control action is to dynamically allocate a fixed number of time slots among $M \geq 2$ transmitters, the objective is to characterize policies that minimize the total number of messages awaiting service at the transmitting nodes subject to the evolution of the state. We partially characterize a set of optimal allocation policies, and show that finding a finite number of dynamic optimal allocations suffices to completely describe an optimal allocation policy. Moreover, by taking advantage of these structural properties of an optimal allocation policy, we show that the number of computations required to find an optimal policy may be significantly reduced. For $M = 2$, we show that the optimal policy is a monotone function of the state and that the total cost is convex. When the process of message generation at one transmitter is stochastically larger than the process of message generation at the other, we further characterize an optimal allocation. These last two properties of an optimal policy imply an additional reduction in the number of computations. Finally, when the processes of message generation at the M transmitters are independent and identically distributed, the explicit form of an optimal allocation policy is found.

I. INTRODUCTION

WE address a flow control problem that arises in the performance design of the hop-by-hop layer of computer communication networks. For a detailed overview on the architectural layers and flow control mechanisms, the reader is referred to [1]–[3]. The hop-by-hop scheme studied in this paper is the same as the one in [4], [5], its purpose being to maintain a smooth flow of traffic between M nodes, termed as *transmitters*, attempting to send messages through a single communication channel to an adjacent node, termed as the *receiver*. The time axis is divided into equal segments called slots. All messages consist of packets of equal length and the transmission time of a packet is one slot. A packet transmission may only begin on a slot boundary. Each transmitter has an independent arrival process of packets and a

buffer of infinite size. Only one transmitter is allowed to transmit during a particular slot.

T consecutive slots form a *phase*. Prior to the beginning of each phase, the receiver informs each transmitter of the number of messages (referred to as a *window size*) it is prepared to accept and the particular slots during which the transmitter is allowed to transmit. In making a decision on the current window sizes, the receiver uses the knowledge of the arrival statistics, the history of previous window sizes, and the information on the number of messages queued at each transmitter at the beginning of the previous phase. This latter information is sent by each transmitter to the receiver in a message of smaller size than a packet at some time between the end of the current phase and the beginning of the next one. Due to propagation delays and the arrival of new packets, the number of queued messages at each transmitter will have changed by the time this information reaches the receiver.

The problem is then to find a policy for the receiver to allocate the slots among the transmitters at each phase of the decision process to minimize a total discounted cost composed of the number of messages awaiting service at each transmission node. We note that this is a discrete-time stochastic control problem with partial information. It can also be viewed as one of optimal resource allocation. Special cases of this problem were addressed by Rosberg and Gopal [6] and Cansever and Milito [4], [5]. Rosberg and Gopal [6] considered a single transmitter with a cost function composed of the number of queued packets and the number of unutilized transmission slots that could be allocated to other links. Cansever and Milito [4], [5] studied a problem with two transmitters and simplified the cost structure in [6]. They also conjectured that their results hold for $M > 2$ transmitters.

In this paper, we consider the general problem with $M \geq 2$ transmitters and messages of equal priority. We do not find an optimal allocation policy in analytical form as this task is very difficult. However, by deriving several structural properties of an optimal policy, we show that finding a finite number (i.e., T) of dynamic optimal allocations suffices to completely describe an optimal allocation policy. For $M = 2$, we discover additional properties of the optimal allocation policy not found in [4], [5]. We prove that an optimal

Manuscript received February 23, 1990; revised March 18, 1991. Paper recommended by Associate Editor At Large, P. R. Kumar. The work of R. Bournas was supported by IBM Endicott; the work of D. Teneketzis was supported by ONR under Grant N00014-87-K-0540.

The authors are with the Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109.
IEEE Log Number 9103544.

allocation policy is monotone in the state and that the total cost is convex. Using this property, we show that to obtain an optimal policy it is enough to find optimal allocations for only T states; for all but one of these states there are only two allocations that are candidates for optimality. The convexity property of the total cost also enables us to further characterize an optimal allocation policy when the arrival processes at the two transmitters are stochastically ordered. The situation where the arrival processes at the $M > 2$ transmitters are identical follows as a particular case, and unlike [4], [5], we find the optimal allocation policy in closed form.

Systems involving transmitters whose outputs embody different priorities (e.g., voice and data) are not considered here; analysis of optimal allocation policies for models with priorities are reserved for future publication [14]. A limited amount of work has been published on optimization of protocols for allocation with priority, but for schemes different from ours. For example, in one of their models, [15], [16] grant absolute priority to voice transmission up to some maximum number of slots, and handle the remaining slots on a random access basis.

The paper is organized as follows. In Section II, we formalize the model and formulate the problem as a discounted Markov decision process. In Section III, we partially characterize an optimal allocation policy for the problem with $M \geq 2$ transmitters by deriving several structural properties. In Section IV, we let $M = 2$ and show that an optimal allocation is monotone in the state and that the total cost is convex. The convexity property of the total cost enables us to further characterize an optimal allocation policy when the process of message generation at one transmitter is stochastically larger than the process of message generation at the other. As a corollary, the explicit form of an optimal allocation policy is found when the message generation processes at the two transmitters are independent and identically distributed. In Section V, we study the special case wherein the message generation processes at the $M \geq 2$ transmitters are independent and identically distributed. The explicit form of an optimal allocation policy is found and the total cost is shown to be convex. Finally, we conclude the paper in Section VI.

II. PROBLEM FORMULATION

Consider a hop-by-hop scheme that operates as follows. There are M transmitting nodes attempting to send messages to a single receiver. All messages consist of packets of fixed length and time is divided into equal slots, one slot being long enough to transmit a packet. A packet transmission may begin only on a slot boundary. The window allocation proceeds in phases, a *phase* being a fixed predetermined number of slots, say T slots. Only one transmitter is allowed to transmit during a particular slot. We also assume that each transmitter has a buffer of infinite size. We shall place two further restrictions on the model: 1) packets arriving in a particular phase may not be transmitted in that phase, and 2) packets that are being transmitted during a phase are not penalized. These assumptions may be considered as a restric-

tion on the model. Relaxing them results in an optimization problem whose action space consists of not only the window sizes allocated to each transmitter but of the *order* in which the slots are scheduled for transmission also. This is a considerably more difficult problem which will be the topic of future investigation.

The processes of message generation at each transmitter are stochastic processes with known statistics. The number of packets generated at transmitter j during slot i , $\xi_i^{(j)}$, $j = 1, 2, \dots, M$, $i = 1, 2, \dots$, are assumed to be independent random variables. For fixed j , $\xi_i^{(j)}$, $i = 1, 2, \dots$, are identically distributed with a general distribution function $F^{(j)}$ and finite first moment $\lambda^{(j)}$.

For every $k = 0, 1, 2, \dots$, let $Y_k^{(j)}$ be the number of packets generated at each transmitter j during phase k . Clearly for fixed j , $1 \leq j \leq M$, $Y_k^{(j)}$, $k = 0, 1, \dots$, are independent and identically distributed random variables with general distribution, say $F_j^{(j)}$, the T -convolution of $F^{(j)}$. In addition, the random variables $\{Y_k^{(j)}, 1 \leq j \leq M, k = 0, 1, \dots\}$ 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 $w_0^{(j)}$, $N_0^{(j)}$, $j = 1, 2, \dots, M$, 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 information on the number of messages queued at each transmitter at the beginning of the previous phase. This latter information is sent by each transmitter to the receiver in a message of smaller size than a packet at some time between the end of the current phase and the beginning of the next one. This information is not sent in the header of the first message that each transmitter transmits to the receiver as in [4]–[6], in order to avoid the problem of a possible zero allocation window to some transmitter.

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

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

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 $w_k = \{w_k^{(1)}, \dots, w_k^{(M)}\}$. The most recent value of this set X_k is calculated by the receiver well 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 $N_{k-1}^{(j)}$'s before the start of phase k .

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

$$N_k = X_k + Y_{k-1}, \quad k \geq 1 \quad (2.2)$$

where $N_k = (N_k^{(1)}, \dots, N_k^{(M)})$ and $Y_k = (Y_k^{(1)}, \dots, Y_k^{(M)})$. Let $p_i^{(j)} = P(Y_k^{(j)} = i)$, $i = 0, 1, \dots$, $j = 1, \dots, M$. The

distribution of $N_k^{(j)}$ given $X_k^{(j)} = x$ is then

$$P(N_k^{(j)} = n | X_k^{(j)} = x) = \begin{cases} P(Y_{k-1}^{(j)} = n - x) = p_{n-x}^{(j)} & \text{if } n \geq x \\ 0 & \text{otherwise.} \end{cases} \quad (2.3)$$

We assume that all packets arriving during a slot join the buffer at slot end. This is a reasonable assumption since an arriving packet at a transmitter is usually placed in a temporary area upon arrival and joins the queue when the transmitter is ready to accept it (at the end of the current packet transmission). For $0 < \nu \leq 1$, we then define the ν -discounted cost incurred at phase i when there are $N_i^{(j)}$ packets at the beginning of the phase, and a window $w_i^{(j)}$ is allocated to transmitter j during the phase as

$$C^\nu(N_i, w_i) \triangleq \sum_{j=1}^M \sum_{k=0}^{T-1} \nu^{(i-1)T+k} \cdot \left((N_i^{(j)} - w_i^{(j)})^+ + \sum_{m=1}^k \xi_{i,m}^{(j)} \right). \quad (2.4)$$

Here, $N_i = (N_i^{(1)}, \dots, N_i^{(M)})$, $w_i = (w_i^{(1)}, \dots, w_i^{(M)})$ and $\{\xi_{i,m}^{(j)}; 1 \leq m \leq T-1\}$ denotes the number of packets arriving during slot m of phase i at transmitter j . The above cost function shows that all packets (except the ones being transmitted during a phase) awaiting transmission at the beginning of every slot are penalized. By a straightforward calculation, we obtain

$$E[C^\nu(N_i, w_i)] = A\nu^{(i-1)T} \cdot \sum_{j=1}^M E[(N_i^{(j)} - w_i^{(j)})^+] + B\nu^{(i-1)T} \quad (2.5)$$

where $A = \sum_{k=0}^{T-1} \nu^k$ and $B = (\sum_{j=1}^M \lambda^{(j)}) (\sum_{k=1}^{T-1} k \nu^k)$ are constants for fixed ν and T (recall $\lambda^{(j)} = E[\xi_{i,m}^{(j)}]$).

By (2.1) and (2.2) the system dynamical and observable equations are written respectively as

$$\begin{cases} N_{k+1} = (N_k - w_k)^+ + Y_k, & k \geq 0 \\ X_k = N_k - Y_{k-1}, & k \geq 1. \end{cases} \quad (2.6)$$

The state space is $S = Z_+^M$ (M -dimensional vectors with nonnegative integer valued components) and the action space is

$$A = \left\{ w = (w^{(1)}, w^{(2)}, \dots, w^{(M)}) \in Z_+^M: \sum_{j=1}^M w^{(j)} \leq T \right\}. \quad (2.7)$$

Without loss of optimality, we restrict attention to Markov policies (see [7, Chapter 3], [12, Sections 6.4-6.7])

$$g_k: S \rightarrow A: x_k \mapsto w_k = g_k(x_k). \quad (2.8)$$

Let $X_1 = x = (x^{(1)}, \dots, x^{(M)})$, the initial system state, be given. 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 ν -discounted cost

$$K_N(x) \triangleq \sum_{i=1}^N E[C^\nu(X_i + Y_{i-1}, w_i) | X_1 = x]. \quad (2.9)$$

Letting $\beta = \nu^T$ and using (2.5), we obtain for $\beta < 1$

$$K_N(x) = A \sum_{i=1}^N \beta^{i-1} \cdot \sum_{j=1}^M E[(X_i^{(j)} + Y_{i-1}^{(j)} - w_i^{(j)})^+ | X_1 = x] + B \frac{(1 - \beta^N)}{(1 - \beta)}. \quad (2.10)$$

For fixed T , since A and B are constants, (2.10) shows that there is no loss of optimality in restricting attention to the expected total β -discounted cost.

$$J_N(x) = \sum_{i=1}^N \beta^{i-1} \sum_{j=1}^M \cdot E[(X_i^{(j)} + Y_{i-1}^{(j)} - w_i^{(j)})^+ | X_1 = x]. \quad (2.11)$$

According to (2.11), we only need to consider the cost of the untransmitted packets at the beginning of each phase. We then define the expected cost per phase when the system is in state x and a window w is allocated as

$$L(x, w) \triangleq \sum_{j=1}^M E[(Y^{(j)} + x^{(j)} - w^{(j)})^+] \quad (2.12)$$

where each $Y^{(j)}$ is an integer random variable with distribution $F_i^{(j)}$. For future reference, we let $Y = (Y^{(1)}, Y^{(2)}, \dots, Y^{(M)})$. By (2.3) and (2.12)

$$\begin{aligned} L(x, w) &= \sum_{i=1}^M \sum_{j=w^{(i)} - x^{(i)} + 1}^{\infty} (x^{(i)} - w^{(i)} + j) p_j^{(i)} \\ &= \sum_{i=1}^M \sum_{j=0}^{\infty} P[Y^{(i)} > w^{(i)} - x^{(i)} + j]. \end{aligned} \quad (2.13)$$

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

$$\begin{cases} V_0^\beta(x) = 0, \\ V_k^\beta(x) = \min_{w \in A} \{ L(x, w) \\ + \beta E[V_{k-1}^\beta([Y + x - w]^+)] \}, \quad 1 \leq k \leq N \end{cases} \quad (2.14)$$

where $V_k^\beta([Y + x - w]^+) \triangleq V_k^\beta([Y^{(1)} + x^{(1)} - w^{(1)}]^+, \dots, [Y^{(M)} + x^{(M)} - w^{(M)}]^+)$. From (2.9) and the definition of $V_k^\beta(x)$

$$V_k^\beta(x) = \sum_{i=1}^k \beta^{i-1} E_x[L(X_i, w_i^*)], \quad 1 \leq k \leq N \quad (2.15)$$

where $E_x[L(X_i, w_i^*)] \triangleq E[L(X_i, w_i^*) | X_1 = x]$ and

$w_i^* \in \mathcal{A}$, $1 \leq i \leq k$, are the optimal window vectors. Since β is fixed, we let

$$V_k \equiv V_k^\beta \quad (2.16)$$

to simplify the notation.

III. QUALITATIVE PROPERTIES OF THE OPTIMAL POLICY FOR $M \geq 2$

In this section, we derive qualitative properties of $V_k(x)$ that will be used to partially characterize the structure of a set of optimal allocation policies. Our first main result (Lemma 3.4) is that the minimal achievable total expected cost is the same for all initial system states whose components sums are equal. This is due to the linearity and the equality of the holding costs. As a consequence of this result, we shall derive the following properties of the optimal allocation policy (Theorems 3.8, 3.10):

(1) With x being the initial system state, if $\sum_{j=1}^M x^{(j)} \geq T$, then any allocation policy g satisfying

$$g(x) = w, \quad \sum_{j=1}^M w^{(j)} = T \quad \text{and} \quad w^{(j)} \leq x^{(j)} \quad \text{for each } j,$$

$$1 \leq j \leq M$$

is optimal. If $\sum_{j=1}^M x^{(j)} \leq T$, then there exists an optimal allocation policy, say π , satisfying

$$\pi(x) = w, \quad \sum_{j=1}^M w^{(j)} = T \quad \text{and} \quad w^{(j)} \geq x^{(j)} \quad \text{for each } j,$$

$$1 \leq j \leq M.$$

This property has the following interpretation: it is optimal to transmit the maximum number possible of queued messages that are known to the receiver because the holding costs at all the transmitters are equal. We note that in the case where $\alpha = \sum_{j=1}^M x^{(j)} < T$, the optimal allocation of the $(T - \alpha)$ remaining slots depends on the statistics of the arrival processes.

(2) The optimal allocation for any initial state is easily derived from the optimal allocation for the problem with initial state the M -dimensional vector with $\sum_{j=1}^M x^{(j)}$ in any i th entry and zeroes in all other entries. This result together with property (1) imply that it suffices to compute the T optimal allocations corresponding to the initial states $\{(0, 0, \dots, 0), (1, 0, \dots, 0), \dots, (T-1, 0, \dots, 0)\}$ to completely solve the finite horizon problem for each epoch.

(3) Because of property (1), the problem of fully characterizing the optimal allocation of policy reduces to characterizing it when $\sum_{j=1}^M x^{(j)} < T$. In this case, because of property (2), the optimization problem is equivalent to one with zero as the initial state and an action space of smaller size.

The above properties reduce the computational effort required to determine an optimal allocation policy. To proceed with the analysis, we first establish some preliminary properties of $V_k(x)$.

Lemma 3.1: $V_k(x)$ is a nondecreasing function in each $x^{(j)}$, $1 \leq j \leq M$.

Proof: The proof is by induction on k . ■

Lemma 3.2: $V_k(x)$ is achieved by an allocation $w_k(x) \in \mathcal{A}$

satisfying

$$\sum_{j=1}^M w_k^{(j)}(x) = T. \quad (3.1)$$

Proof: Define

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

For any i , $1 \leq i \leq M$, we have from (2.13)

$$L(x, w) - L(x, w + e_i) = P[Y^{(i)} > w^{(i)} - x^{(i)}] \quad (3.3)$$

which implies that $L(x, w + e_i) \leq L(x, w)$. Moreover, by Lemma 3.1

$$E[V_{k-1}([Y + x - w - e_i]^+)] \leq E[V_{k-1}([Y + x - w]^+)] \quad (3.4)$$

so that

$$G_k(x, w + e_i) \leq G_k(x, w). \quad (3.5)$$

Thus if $V_k(x) = G_k(x, w_*)$ and $\sum_{j=1}^M w_*^{(j)} < T$, then by (3.5) there exists $v_* \in \mathcal{A}$ such that $\sum_{j=1}^M v_*^{(j)} = T$ and $V_k(x) = G_k(x, v_*)$. ■

In the following corollary we show that if the probability of at least T arrivals within a phase is positive, then every optimal policy must allocate all the T slots, and conversely, if every optimal policy allocates all the T slots, then it must be that the probability of at least T arrivals within a phase is positive.

Corollary 3.3: $P[\sum_{j=1}^M Y^{(j)} \geq T] > 0$ if and only if $\forall x \in \mathcal{S}$, $\forall w_k(x) \in \mathcal{A} \ni V_k(x) = G_k(x, w_k(x))$, it must be that $\sum_{j=1}^M w_k^{(j)}(x) = T$.

Proof:

(\Rightarrow): $P[\sum_{j=1}^M Y^{(j)} \geq T] > 0$ implies the existence of an integer $n \geq T$ such that $P[\sum_{j=1}^M Y^{(j)} = n] > 0$. As $\{Y^{(j)}, j = 1, 2, \dots, M\}$ are independent, this, in turn, implies the existence of nonnegative integers n_1, \dots, n_M such that $\sum_{j=1}^M n_j = n$ and $p_{n_j}^{(j)} > 0$ for each j , $1 \leq j \leq M$. Let $V_k(x) = G_k(x, w_*)$ for some $w_* \in \mathcal{A}$ and assume $\sum_{j=1}^M w_*^{(j)} < T$. As $\sum_{j=1}^M (w_*^{(j)} - x^{(j)}) < T$ and $\sum_{j=1}^M n_j \geq T$, there exists an i such that $n_i > w_*^{(i)} - x^{(i)}$. Then $p_{n_i}^{(i)} > 0$ and $n_i > w_*^{(i)} - x^{(i)}$ imply that $P[Y^{(i)} > w_*^{(i)} - x^{(i)}] > 0$. Using this result in (3.3) and combining with (3.4), we obtain $G_k(x, w_* + e_i) < G_k(x, w_*)$ and this contradicts the optimality of w_* .

(\Leftarrow): The proof is by contradiction. We assume that $P[\sum_{j=1}^M Y^{(j)} < T] = 1$ and show that for some system state $x = (x^{(1)}, \dots, x^{(M)})$, there is an optimal allocation $w_* \in \mathcal{A}$ with $\sum_{j=1}^M w_*^{(j)} < T$. If $P[\sum_{j=1}^M Y^{(j)} < T] = 1$, then $P[\sum_{j=1}^M Y^{(j)} = k] = 0$ for all $k \geq T$. Then as the random variables $\{Y^{(j)}\}_{j=1}^M$ are independent, there exist nonnegative integers n_1, \dots, n_M such that $\sum_{j=1}^M n_j < T$, and $P[Y^{(j)} \leq n_j] = 1$ for each j , $1 \leq j \leq M$. With $n \triangleq (n_1, \dots, n_M)$, let $\pi = \prod_{i=1}^{\infty} \pi_i$ be the stationary policy $\pi_i(x) = n$ for all $x \in \mathcal{S}$ and all $i \geq 1$. We will then show by induction that the total cost achieved by policy π as defined by (2.15), is equal to

zero, i.e.,

$$\forall k \geq 0, \quad V_k(\pi, 0) = 0. \quad (3.6)$$

For $k = 0$, (3.6) is trivial. Suppose (3.6) holds for k . We then have the following recursive equations:

$$\forall k \geq 0, \\ V_{k+1}(\pi, 0) = L(0, n) + \beta E[V_k(\pi, [Y - n]^+)].$$

Since $P[Y^{(j)} \leq n_j] = 1$, then $P[(Y^{(j)} - n_j)^+ = 0] = 1$ for each j , $1 \leq j \leq M$. Hence $L(0, n) = 0$ and $P[V_k(\pi, [Y - n]^+) = V_k(\pi, 0)] = 1$. By using the induction hypothesis, we then have $V_{k+1}(\pi, 0) = 0$. Hence π minimizes the k -step cost. ■

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

$$\bar{A} \triangleq \left\{ w = (w^{(1)}, w^{(2)}, \dots, w^{(M)}) \in A : \sum_{j=1}^M w^{(j)} = T \right\}. \quad (3.7)$$

The dynamic programming equations (2.14) then become

$$V_k(x) = \min_{w \in \bar{A}} \{ L(x, w) + \beta E[V_{k-1}([Y + x - w]^+)] \} \\ = \min_{w \in \bar{A}} \{ G_k(x, w) \}. \quad (3.8)$$

In the next lemma we show that the minimal achievable total expected cost is invariant under packet transfers of the individual initial states of the transmitters. As a corollary, we show that this cost is the same for all initial system states whose components sums are equal.

Lemma 3.4: For any i, j , if $x^{(i)} \geq 1$, then

$$V_k(x - e_i + e_j) = V_k(x). \quad (3.9)$$

Proof: The proof of (3.9) is by induction on k . The assertion certainly holds for $k = 0$. Suppose $V_k(x - e_i + e_j) = V_k(x) \forall x \in S$. For some $w_* \in \bar{A}$, write

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

If $w_*^{(i)} \geq 1$, then

$$V_{k+1}(x - e_i + e_j) \leq G_{k+1}(x - e_i + e_j, w_* - e_i + e_j) \\ = G_{k+1}(x, w_*) = V_{k+1}(x).$$

If $w_*^{(i)} = 0$, then as $x^{(i)} \geq 1$, we have $(Y^{(i)} + x^{(i)} - 1)^+ = Y^{(i)} + x^{(i)} - 1$ for all $Y^{(i)}$. Moreover, for all $Y^{(j)}$, $(Y^{(j)} + x^{(j)} + 1 - w_*^{(j)})^+ \leq (Y^{(j)} + x^{(j)} - w_*^{(j)})^+ + 1$. Thus by the monotonicity of $L(\cdot, w)$ and $V_k(\cdot)$

$$L(x - e_i + e_j, w_*) \leq L(x, w_*), \quad (3.10)$$

$$E[V_k([Y + x - e_i + e_j - w_*]^+)] \\ \leq E[V_k([Y + x - w_*]^+ - e_i + e_j)]. \quad (3.11)$$

The term on the RHS of (3.11) is equal to

$E[V_k([Y + x - w_*]^+)]$ by the induction hypothesis. Adding (3.10) and (3.11), we obtain

$$V_{k+1}(x - e_i + e_j) \leq V_{k+1}(x). \quad (3.12)$$

If $x^{(j)} \geq 1$, then by the above inequality

$$V_{k+1}(x + e_i - e_j) \leq V_{k+1}(x).$$

Replacing x by $x - e_i + e_j$ in the above, we get

$$V_{k+1}(x) \leq V_{k+1}(x - e_i + e_j). \quad (3.13)$$

Combining (3.12) and (3.13) completes the proof of (3.9). ■

For $x \in S$ define

$$\alpha(x) \triangleq \sum_{j=1}^M x^{(j)}, \quad (3.14)$$

$$\gamma(x) \triangleq T - \alpha(x), \quad \text{for } \alpha(x) \leq T. \quad (3.15)$$

To simplify the notation, whenever $x \in S$ is fixed, we let $\alpha \triangleq \alpha(x)$ and $\gamma \triangleq \gamma(x)$.

Corollary 3.5:

$$\forall x, y \in S \quad \alpha(x) = \alpha(y) \Rightarrow V_k(x) = V_k(y) \quad (3.16)$$

Proof: Define

$$U \triangleq \{u_{ij} = e_i - e_j, 1 \leq i, j \leq M\}.$$

There exist finitely many elements $v_1, v_2, \dots, v_n \in U$ such that the vectors defined by $y_0 \triangleq y$ and $y_l \triangleq y_{l-1} + v_l$, $1 \leq l \leq n$, are well defined (they have nonnegative components) and $x = y_n$. Then by (3.9)

$$V_k(y) = V_k(y_0) = V_k(y_1) = \dots = V_k(y_n) = V_k(x). \quad \blacksquare$$

In the next lemma and theorem, we prove the properties of an optimal allocation policy stated at the beginning of the section.

Lemma 3.6: If $\alpha \leq T$, then $\exists w_k(x) \in \bar{A}$ satisfying $w_k^{(j)}(x) \geq x^{(j)}$ for each j , $1 \leq j \leq M$ such that

$$V_k(x) = G_k(x, w_k(x)).$$

Proof: Let $V_k(x) = G_k(x, w_*) = L(x, w_*) + \beta E[V_{k-1}([Y + x - w_*]^+)]$ for some $w_* \in \bar{A}$. Assume $w_*^{(i)} < x^{(i)}$ for some i . Then as $\sum_{l=1}^M (w_*^{(l)} - x^{(l)}) \geq 0$, we have $\sum_{l \neq i} (w_*^{(l)} - x^{(l)}) > 0$. Hence there exists $j_1 \neq i$ such that $w_*^{(j_1)} > x^{(j_1)}$. Define a policy π_1 that allocates a window $v_* = w_* + e_i - e_{j_1}$ at the first phase when the system state is x and then continues for the remaining $(k-1)$ phases according to an optimal policy. Then

$$V_k(\pi_1, x) = G_k(x, v_*) \\ = L(x, v_*) + \beta E[V_{k-1}([Y + x - v_*]^+)].$$

As $x^{(i)} > w_*^{(i)}$, then for all $Y^{(i)}$, $(Y^{(i)} + x^{(i)} - w_*^{(i)} - 1)^+ = (Y^{(i)} + x^{(i)} - w_*^{(i)})^+ - 1$. Also $(Y^{(j_1)} + x^{(j_1)} - w_*^{(j_1)} + 1)^+ \leq (Y^{(j_1)} + x^{(j_1)} - w_*^{(j_1)})^+ + 1$ for all $Y^{(j_1)}$. Thus by the monotonicity of $V_k(\cdot)$, $E[V_{k-1}([Y + x - v_*]^+)] \leq E[V_{k-1}([Y + x - w_*]^+ - e_i + e_{j_1})]$. By (3.16), $E[V_{k-1}([Y + x - w_*]^+ - e_i + e_{j_1})] = E[V_{k-1}([Y + x - w_*]^+)]$.

Hence

$$E[V_{k-1}([Y + x - v_*]^+)] \leq E[V_{k-1}([Y + x - w_*]^+)]. \quad (3.17)$$

Moreover by (2.13)

$$\begin{aligned} L(x, v_*) - L(x, w_*) &= E[(Y^{(j_1)} + x^{(j_1)} - w_*^{(j_1)} + 1)^+] \\ &\quad - E[(Y^{(j_1)} + x^{(j_1)} - w_*^{(j_1)})^+] - 1 \\ &= P[Y^{(j_1)} \geq w_*^{(j_1)} - x^{(j_1)}] - 1 \end{aligned} \quad (3.18)$$

and clearly $L(x, v_*) \leq L(x, w_*)$. Thus $V_k(\pi_1, x) = V_k(x)$. If $w_*^{(j_1)} + 1 = x^{(j_1)}$, then the claim is proved. If $w_*^{(j_1)} + 1 < x^{(j_1)}$, then define a policy π_2 that allocates a window vector $w_* + 2e_i - e_{j_2}$ (where j_2 is such that $w_*^{(j_2)} > x^{(j_2)}$) at the first phase when the system state is x and then continues for the remaining $(k-1)$ phases according to an optimal policy. Then by the same argument as above $V_k(\pi_2, x) = V_k(x)$. By proceeding in this fashion, we end up with a policy π_l that allocates a window $w_*^{(i)} + l = x^{(i)}$ to transmitter i at the first phase when the system state is x and then continues for the $(k-1)$ remaining phases according to an optimal policy and $V_k(\pi_l, x) = V_k(x)$. Since i is arbitrary, the proof is complete. ■

We next show that if in addition to the hypothesis of Lemma 3.6, the probability of no arrival at each transmitter is positive, then the optimal allocation policy must allocate at least $x^{(j)}$ slots to each transmitter j . The intuitive reasoning behind this result is as follows. Suppose the receiver allocates $w^{(i)} < x^{(i)}$ to transmitter i . Then it must allocate $w^{(j)} > x^{(j)}$ to some other transmitter j . As the probability of no arrivals at transmitter j is positive, then with some positive probability, transmitter j will not use all its allocated slots while there certainly are $(x^{(i)} - w^{(i)}) > 0$ packets awaiting service at transmitter i . Since i and j have equal holding costs and are arbitrary, it is then optimal to transmit all messages currently awaiting service.

Corollary 3.7: Suppose $p_0^{(j)} > 0$ for all j , $1 \leq j \leq M$. With the same hypothesis of Lemma 3.6, i.e., $\alpha \leq T$, $\forall w_k(x) \in \bar{A}$ such that $V_k(x) = G_k(x, w_k(x))$, it must be that $w_k^{(j)}(x) \geq x^{(j)}$ for each j , $1 \leq j \leq M$.

Proof: As in the proof of Lemma 3.6, let

$$\begin{aligned} V_k(x) &= G_k(x, w_*) \\ &= L(x, w_*) + \beta E[V_{k-1}([Y + x - w_*]^+)] \end{aligned}$$

for some $w_* \in \bar{A}$. Assume $w_*^{(i)} < x^{(i)}$ for some i . Then there exists $j_1 \neq i$ such that $w_*^{(j_1)} > x^{(j_1)}$. Define a policy π_1 that allocates a window $v_* = w_* + e_i - e_{j_1}$ at the first phase when the system state is x and then continues for the remaining $(k-1)$ phases according to an optimal policy. Then $w_*^{(j_1)} > x^{(j_1)}$ and $p_0^{(j_1)} > 0$ imply in (3.18) that $L(x, v_*) < L(x, w_*)$. Combining with (3.17), we have $V_k(\pi_1, x) < V_k(x)$. But this is a contradiction to the optimality of the allocation w_* . ■

Next, we use the results of Lemma 3.4 and Lemma 3.6 to obtain the following theorem.

Theorem 3.8:

a) If $\alpha \geq T$, then any allocation $w_k(x) \in \bar{A}$ such that $w_k^{(l)}(x) \leq x^{(l)}$ for each l is optimal, i.e., $V_k(x) = G_k(x, w_k(x))$.

b) If $\alpha \leq T$, then for any i , $1 \leq i \leq M$, there exists $w_k(\alpha e_i) \in \bar{A}$ such that $w_k^{(i)}(\alpha e_i) \geq \alpha$ and

$$V_k(\alpha e_i) = G_k(\alpha e_i, w_k(\alpha e_i)),$$

$$V_k(x) = G_k(x, x + w_k(\alpha e_i) - \alpha e_i).$$

Proof: Let $\alpha \geq T$, and pick i , $1 \leq i \leq M$. For any realization $Y = (Y^{(1)}, \dots, Y^{(M)})$ and any $w = (w^{(1)}, \dots, w^{(M)}) \in \bar{A}$

$$\begin{aligned} \sum_{j=1}^M (Y^{(j)} + x^{(j)} - w^{(j)})^+ &\geq \left(\sum_{j=1}^M (Y^{(j)} + x^{(j)} - w^{(j)}) \right)^+ \\ &= \alpha - T + \sum_{j=1}^M Y^{(j)}. \end{aligned} \quad (3.19)$$

Hence

$$\begin{aligned} L(x, w) &= E \left[\sum_{j=1}^M (Y^{(j)} + x^{(j)} - w^{(j)})^+ \right] \\ &\geq \alpha - T + \sum_{j=1}^M E[Y^{(j)}]. \end{aligned} \quad (3.20)$$

By (3.16), $V_{k-1}([Y + x - w]^+) = V_{k-1}(\{\sum_{j=1}^M (Y^{(j)} + x^{(j)} - w^{(j)})^+ \} e_i)$. Using the monotonicity of $V_{k-1}(\cdot)$ and (3.19) in the above, we obtain

$$\begin{aligned} E[V_{k-1}([Y + x - w]^+)] &\geq E \left[V_{k-1} \left(\left\{ \alpha - T + \sum_{j=1}^M Y^{(j)} \right\} e_i \right) \right]. \end{aligned} \quad (3.21)$$

Combining (3.20) and (3.21), we get

$$\begin{aligned} G_k(x, w) &\geq \alpha - T + \sum_{j=1}^M E[Y^{(j)}] \\ &\quad + \beta E \left[V_{k-1} \left(\left\{ \alpha - T + \sum_{j=1}^M Y^{(j)} \right\} e_i \right) \right]. \end{aligned} \quad (3.22)$$

Pick any $\bar{w} \in \bar{A}$ with $\bar{w}^{(j)} \leq x^{(j)}$ for each j . Noting that $(Y^{(j)} + x^{(j)} - \bar{w}^{(j)})^+ = Y^{(j)} + x^{(j)} - \bar{w}^{(j)}$ for each j , and using (3.16), then

$$\begin{aligned} G_k(x, \bar{w}) &= \alpha - T + \sum_{j=1}^M E[Y^{(j)}] \\ &\quad + \beta E \left[V_{k-1} \left(\left\{ \alpha - T + \sum_{j=1}^M Y^{(j)} \right\} e_i \right) \right]. \end{aligned}$$

The lower bound on the RHS of (3.22) is then achieved by \bar{w} , and thus \bar{w} is optimal. This proves a). Next, we prove b). Suppose $\alpha \leq T$ and let

$$\begin{aligned} V_k(\alpha e_i) &= G_k(\alpha e_i, w(\alpha e_i)) \\ &= L(\alpha e_i, w(\alpha e_i)) \\ &\quad + \beta E[V_{k-1}([Y + \alpha e_i - w(\alpha e_i)]^+)] \end{aligned}$$

for some $w(\alpha e_i) \in \bar{A}$. By Lemma 3.6, such an allocation with $w^{(i)} \geq \alpha$ exists. Thus the allocation $v(x) = x + w(\alpha e_i) - \alpha e_i$ is well defined and $v(x) \in \bar{A}$. Moreover,

$$\begin{aligned} L(\alpha e_i, w(\alpha e_i)) &= L(\alpha e_i, v(x) + \alpha e_i - x) \\ &= L(x, v(x)) \\ E[V_{k-1}(\{Y + \alpha e_i - w(\alpha e_i)\}^+)] \\ &= E[V_{k-1}(\{Y + \alpha e_i - (v(x) + \alpha e_i - x)\}^+)] \\ &= E[V_{k-1}(\{Y + x - v(x)\}^+)]. \end{aligned}$$

Hence $V_k(\alpha e_i) = G_k(x, v(x))$. Also by (3.16), $V_k(x) = V_k(\alpha e_i)$. Thus $V_k(x) = G_k(x, v(x)) = G_k(x, x + w(\alpha e_i) - \alpha e_i)$. ■

The theorem asserts that for $\alpha \geq T$ optimal allocations are easily specified. When $\alpha < T$, then for each $\alpha \in \{0, 1, 2, \dots, T-1\}$ it suffices to compute an optimal allocation, say $w(x)$, for one x whose component sum is α ; optimal allocations for other x with the same component sum are easily determined from this $w(x)$. Nevertheless, the relevant dynamic programming equations still require knowledge of the value function for all states. This requirement becomes less formidable when it is realized that the value function depends only on the component sum α .

The next lemma helps to prove that an optimal allocation for any state x such that $\sum_{j=1}^M x^{(j)} = \alpha < T$ can be determined from an optimization problem with an action space smaller than the original one. This is the third property of an optimal allocation policy stated at the beginning of the section.

Lemma 3.9:

$$G_k(x, x + w) = G_k(0, w). \quad (3.23)$$

Proof: The identity follows directly from the definition (2.12) of $L(x, w)$ and the definition (3.2) of $G_k(x, w)$. ■

At this juncture, we remind the reader that for $\alpha \leq T$, $\gamma \triangleq T - \alpha$ (cf. (3.15)) is the number of slots that we are interested in allocating optimally among the transmitters. We shall next parametrize our allocation problem in terms of γ . We first define

$$B(\gamma) \triangleq \left\{ \sigma = (\sigma^{(1)}, \dots, \sigma^{(M)}) \in \mathbf{Z}_+^M : \sum_{j=1}^M \sigma^{(j)} = \gamma \right\}, \quad (3.24)$$

$$\bar{V}_k(0, \gamma) \triangleq \min_{\sigma \in B(\gamma)} G_k(0, \sigma). \quad (3.25)$$

We shall use these definitions frequently throughout the remainder of this paper.

Theorem 3.10: Let $\alpha \leq T$. Then for some $\sigma_{k,*} \in B(\gamma)$

$$V_k(x) = G_k(x, x + \sigma_{k,*}) = \bar{V}_k(0, \gamma). \quad (3.26)$$

Proof: Let $V_k(x) = G_k(x, w_*)$ for some $w_* \in B(T)$. By Theorem 3.8, such a w_* with $w_*^{(j)} \geq x^{(j)}$, for each j ,

exists. Then

$$V_k(x) = G_k(x, w_*) = \min_{\sigma \in B(\gamma)} G_k(x, x + \sigma).$$

Using (3.23) in the above, (3.26) is now immediate. ■

This theorem states that when $\alpha \leq T$, the search for optimal allocations over the action space \bar{A} is equivalent to the search for optimal allocations over the action space $B(\gamma)$, $0 \leq \gamma \leq T$. With $|X|$ denoting the size of set X , then $|B(\gamma)|$ is the number of nonnegative integer solutions of the equation

$$\sum_{j=1}^M x^{(j)} = \gamma$$

and it is given by (see [9, p. 150])

$$|B(\gamma)| = \binom{M + \gamma - 1}{M - 1}.$$

Then

$$|\bar{A}| = |B(T)| = |B(\gamma)| \prod_{i=1}^{T-\gamma} \frac{M + T - i}{T - i + 1}$$

and clearly $|B(\gamma)| < |\bar{A}|$. For a given γ , the number of allocations that could be optimal is thus reduced by a factor of $|\bar{A}| / |B(\gamma)|$.

Remark 1: The results obtained in this section can be easily extended to hold for the infinite horizon expected discounted cost problem. By induction on k , the value function satisfies $V_k(x) \leq V_{k+1}(x)$. Moreover, since the average cost cannot increase with time faster than the average number of arrivals per unit time, and discounting reduces future costs at a geometric rate, it can be shown that for every allocation policy, the corresponding infinite horizon expected discounted cost is bounded. Then, by the standard arguments of [8] (Theorem 1) the minimal achievable infinite horizon expected discounted cost is the limit of the minimal achievable finite horizon expected discounted cost as the horizon increases indefinitely. Hence, by a limiting argument, the results of Lemmas 3.1, 3.2, 3.4, 3.6, 3.9, Corollaries 3.3, 3.5, 3.7, and Theorems 3.8, 3.10 are true for the infinite horizon expected discounted cost problem.

IV. MONOTONICITY OF THE OPTIMAL POLICY WHEN $M = 2$

In this section, we let $M = 2$ and show that $V_k(x)$ is convex in x and that the optimal allocation policy is monotone in x . Submodularity of the total cost function in its arguments is shown to be a sufficient condition in the derivation of these results. Moreover, when the arrival process at one transmitter is stochastically larger than the arrival process at the other, we use the convexity property of the value function to show that if the initial state $(x^{(1)}, x^{(2)})$ is such as $x^{(1)} + x^{(2)} < T$, then for some $\sigma_*^{(1)}, \sigma_*^{(2)}$ satisfying $\sigma_*^{(1)} \geq \sigma_*^{(2)}$ and $\sigma_*^{(1)} + \sigma_*^{(2)} = T - (x^{(1)} + x^{(2)})$, $(x^{(1)} + \sigma_*^{(1)}, x^{(2)} + \sigma_*^{(2)})$ is an optimal allocation policy.

A. Structure of the Optimal Policy When $Y^{(1)}$ and $Y^{(2)}$ are General

In this section, we derive the monotonicity property of the optimal allocation policy and the convexity of the value

function when the arrival processes at the transmitters are general. We first define convexity, submodularity and prove some preliminary lemmas.

The following definitions are as in [13]. Let g be a function of $x = (x_1, \dots, x_n)$, where $x_1, \dots, x_n \in \mathbb{Z}$. We say g is convex in \mathbb{Z}^n if

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

and g is submodular in \mathbb{Z}^n if

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

The following two lemmas are easy to prove; therefore, their proof is omitted.

Lemma 4.1: Let g be a function of $x = (x_1, x_2) \in \mathbb{Z}_+^2$. If g is monotonically nondecreasing in each variable, and convex, then $g(x_1^+, x_2^+)$ is convex in \mathbb{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 $x = (x_1, x_2) \in \mathbb{Z}_+^2$, satisfy the three properties:

- $g(x_1, x_2)$ is monotonically nondecreasing in each variable,
- $g(x_1, x_2)$ is convex in each variable,
- $g(x_1, x_2)$ is submodular.

Then $g(x_1^+, x_2^+)$ is submodular in \mathbb{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 two theorems we show that $V_k(x)$ is convex and that the optimal allocation policy is monotone in x .

Theorem 4.3:

- Let $0 \leq \gamma < T$ and $\bar{V}_k(0, \gamma) = G_k(0, \sigma_*)$ for some $\sigma_* \in B(\gamma)$. Then for some i ,

$$\bar{V}_k(0, \gamma + 1) = G_k(0, \sigma_* + e_i). \quad (4.5)$$

- $V_k(x)$ is convex.

Proof: Clearly, the theorem holds for $k = 0$. By assuming that $V_k(x)$ is convex, we shall show that a) and b) hold for $V_{k+1}(x)$; thus by way of induction the theorem holds for all $k \geq 0$. We first note that $V_k(x^{(1)}, x^{(2)})$ satisfies the hypothesis of Lemma 4.2, namely, property a) follows from Lemma 3.1, property b) follows from the induction hypothesis, and property c) follows by a direct calculation from Lemma 3.4 (also note that this property holds with equality). Hence $V_k((x^{(1)})^+, (x^{(2)})^+)$ is submodular in \mathbb{Z}^2 , a property that we shall use shortly in this proof.

Proof of a): Suppose $V_k(x)$ is convex, and let $x^{(1)} + x^{(2)} = T - \gamma > 0$. Then by (3.26)

$$\begin{aligned} V_{k+1}(x) &= \bar{V}_{k+1}(0, \gamma) \\ &= L(0, \sigma_*) + \beta E[V_k([Y - \sigma_*]^+)] \end{aligned}$$

for some $\sigma_* \in B(\gamma)$. Define

$$B(\gamma, j) \triangleq \{\sigma \in B(\gamma) : \sigma^{(j)} \geq \sigma_*^{(j)}\}. \quad (4.6)$$

Our goal is to show that

$$\forall \sigma \in B(\gamma, j), \quad G_{k+1}(0, \sigma_* + e_j) \leq G_{k+1}(0, \sigma + e_j). \quad (4.7)$$

For any $\sigma \in B(\gamma)$, write

$$\begin{aligned} &G_{k+1}(0, \sigma + e_j) - G_{k+1}(0, \sigma) \\ &= \{E[(Y^{(j)} - \sigma^{(j)} - 1)^+] - E[(Y^{(j)} - \sigma^{(j)})^+]\} \\ &\quad + \beta \{E[V_k([Y - \sigma - e_j]^+)] \\ &\quad - E[V_k([Y - \sigma]^+)]\}. \end{aligned}$$

If $\sigma \in B(\gamma, j)$, then $\sigma^{(j)} \geq \sigma_*^{(j)}$. Hence by the convexity of $(\cdot)^+$

$$\begin{aligned} &E[(Y^{(j)} - \sigma^{(j)} - 1)^+] - E[(Y^{(j)} - \sigma^{(j)})^+] \\ &\geq E[(Y^{(j)} - \sigma_*^{(j)} - 1)^+] - E[(Y^{(j)} - \sigma_*^{(j)})^+]. \quad (4.8) \end{aligned}$$

We next show that

$$\begin{aligned} &E[V_k([Y - \sigma - e_j]^+)] - E[V_k([Y - \sigma]^+)] \\ &\geq E[V_k([Y - \sigma_* - e_j]^+)] - E[V_k([Y - \sigma_*]^+)]. \quad (4.9) \end{aligned}$$

Without loss of generality, assume $j = 1$. If $\sigma^{(1)} = \sigma_*^{(1)}$, then (4.9) is trivial; so suppose $\sigma^{(1)} > \sigma_*^{(1)}$. Then $\sigma^{(1)} = \sigma_*^{(1)} + m$, $\sigma^{(2)} = \sigma_*^{(2)} - m$ for some $m \geq 1$. For any realization $(Y^{(1)}, Y^{(2)})$, let $x_1 = Y^{(1)} - \sigma_*^{(1)} - 1$ and $x_2 = Y^{(2)} - \sigma_*^{(2)}$. Then by repeated applications of Lemma 4.2 to $V_k(x^{(1)}, x^{(2)})$

$$\begin{aligned} &V_k([x_1 + 1]^+, x_2^+) - V_k(x_1^+, x_2^+) \\ &\geq V_k(x_1^+, [x_2 + 1]^+) \\ &\quad - V_k([x_1 - 1]^+, [x_2 + 1]^+) \\ &\geq V_k([x_1 - 1]^+, [x_2 + 2]^+) \\ &\quad - V_k([x_1 - 2]^+, [x_2 + 2]^+) \geq \dots \\ &\geq V_k([x_1 - m + 1]^+, [x_2 + m]^+) \\ &\quad - V_k([x_1 - m]^+, [x_2 + m]^+) \end{aligned}$$

or equivalently

$$\begin{aligned} &V_k([Y^{(1)} - \sigma^{(1)} - 1]^+, [Y^{(2)} - \sigma^{(2)}]^+) \\ &\quad - V_k([Y^{(1)} - \sigma^{(1)}]^+, [Y^{(2)} - \sigma^{(2)}]^+) \\ &\geq V_k([Y^{(1)} - \sigma_*^{(1)} - 1]^+, [Y^{(2)} - \sigma_*^{(2)}]^+) \\ &\quad - V_k([Y^{(1)} - \sigma_*^{(1)}]^+, [Y^{(2)} - \sigma_*^{(2)}]^+). \end{aligned}$$

Taking expectations on both sides of the above, we get (4.9).

Combining (4.8) and (4.9), we obtain

$$\forall \sigma \in B(\gamma, j),$$

$$\begin{aligned} G_{k+1}(0, \sigma + e_j) - G_{k+1}(0, \sigma) \\ \geq G_{k+1}(0, \sigma_* + e_j) - G_{k+1}(0, \sigma_*). \end{aligned} \quad (4.10)$$

Since by the optimality of σ_*

$$\forall \sigma \in B(\gamma, j), \quad G_{k+1}(0, \sigma) \geq G_{k+1}(0, \sigma_*) \quad (4.11)$$

then (4.10) and (4.11) lead to (4.7). Let $v = (v^{(1)}, v^{(2)}) \in B(\gamma + 1)$. Then there exists j such that $v^{(j)} \geq \sigma_*^{(j)} + 1$. Thus $v - e_j \in B(\gamma, j)$ and by (4.7)

$$G_{k+1}(0, \sigma_* + e_j) \leq G_{k+1}(0, v).$$

We thus conclude that

$$\begin{aligned} \overline{V}_{k+1}(0, \gamma + 1) &= \min_{v \in B(\gamma+1)} G_{k+1}(0, v) \\ &= \min_{j=1,2} G_{k+1}(0, \sigma_* + e_j). \end{aligned}$$

This completes the proof of a). Let $\overline{V}_{k+1}(0, \gamma + 1) = G_{k+1}(0, \sigma_* + e_m)$. Then by the optimality of $\sigma_* + e_m$, it must be that

$$\forall l, \quad G_{k+1}(0, \sigma_* + e_m) \leq G_{k+1}(0, \sigma_* + e_l)$$

or equivalently, $\forall l$

$$\begin{aligned} E[(Y^{(m)} - \sigma_*^{(m)} - 1)^+] - E[(Y^{(m)} - \sigma_*^{(m)})^+] \\ + \beta E[V_k([Y - \sigma_* - e_m]^+)] \\ \leq E[(Y^{(l)} - \sigma_*^{(l)} - 1)^+] - E[(Y^{(l)} - \sigma_*^{(l)})^+] \\ + \beta E[V_k([Y - \sigma_* - e_l]^+)]. \end{aligned} \quad (4.12)$$

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

Case 1: $\alpha \geq T$. Let $V_{k+1}(x) = G_{k+1}(x, w_*)$ for some $w_* \in B(T)$. By Theorem 3.8, there exists an optimal w_* such that $w_*^{(l)} \leq x^{(l)}$ for each l . Then

$$\begin{aligned} V_{k+1}(x) &= \alpha - T + \sum_{l=1}^2 E[Y^{(l)}] \\ &\quad + \beta E[V_k(Y + x - w_*)]. \end{aligned}$$

Applying Theorem 3.8 again, i.e., w_* is also optimal when the system state is $x + e_i$ or $x + e_j$ or $x + e_i + e_j$, we obtain

$$\begin{aligned} V_{k+1}(x + e_i) &= \alpha + 1 - T + \sum_{l=1}^2 E[Y^{(l)}] \\ &\quad + \beta E[V_k(Y + x + e_i - w_*)], \end{aligned}$$

$$\begin{aligned} V_{k+1}(x + e_j) &= \alpha + 1 - T + \sum_{l=1}^2 E[Y^{(l)}] \\ &\quad + \beta E[V_k(Y + x + e_j - w_*)], \\ V_{k+1}(x + e_i + e_j) &= \alpha + 2 - T + \sum_{l=1}^2 E[Y^{(l)}] \\ &\quad + \beta E[V_k(Y + x + e_i + e_j - w_*)]. \end{aligned}$$

Thus by the induction hypothesis, i.e., the convexity of $V_k(x)$, $V_{k+1}(x)$ is convex.

Case 2: $\alpha \leq T - 2$. By (3.26)

$$V_{k+1}(x + e_i + e_j) = \overline{V}_{k+1}(0, \gamma - 2) = G_{k+1}(0, \sigma_*)$$

for some $\sigma_* \in B(\gamma - 2)$. Also by Theorem 3.10 and part a), respectively, we have for some m and n

$$V_{k+1}(x + e_i) = \overline{V}_{k+1}(0, \gamma - 1) = G_{k+1}(0, \sigma_* + e_m),$$

$$V_{k+1}(x) = \overline{V}_{k+1}(0, \gamma) = G_{k+1}(0, \sigma_* + e_m + e_n).$$

Moreover by (3.9), $V_{k+1}(x + e_j) = V_{k+1}(x + e_i)$ and by the optimality of $\sigma_* + e_m$, m satisfies (4.12). Then

$$\begin{aligned} V_{k+1}(x + e_i + e_j) - V_{k+1}(x + e_i) \\ = E[(Y^{(m)} - \sigma_*^{(m)})^+] - E[(Y^{(m)} - \sigma_*^{(m)} - 1)^+] \\ + \beta \{ E[V_k([Y - \sigma_*]^+)] \\ - E[V_k([Y - \sigma_* - e_m]^+)] \}. \end{aligned}$$

If $m = n$, then

$$\begin{aligned} V_k(x + e_j) - V_k(x) \\ = E[(Y^{(m)} - \sigma_*^{(m)} - 1)^+] - E[(Y^{(m)} - \sigma_*^{(m)} - 2)^+] \\ + \beta \{ E[V_k([Y - \sigma_* - e_m]^+)] \\ - E[V_k([Y - \sigma_* - 2e_m]^+)] \} \end{aligned}$$

and in this case, by the convexity of $(\cdot)^+$ and the convexity of $V_k(x^+)$ (as it is implied by the convexity of $V_k(x)$ and Lemma 4.1), it follows that $V_{k+1}(x)$ is convex.

If $m \neq n$, then

$$\begin{aligned} V_{k+1}(x + e_i + e_j) - V_{k+1}(x + e_i) \\ - V_{k+1}(x + e_j) + V_{k+1}(x) \\ = \left\{ E[(Y^{(n)} - \sigma_*^{(n)} - 1)^+] - E[(Y^{(n)} - \sigma_*^{(n)})^+] \right. \\ \left. + \beta E[V_k([Y - \sigma_* - e_n]^+)] \right. \\ \left. - \left(E[(Y^{(m)} - \sigma_*^{(m)} - 1)^+] - E[(Y^{(m)} - \sigma_*^{(m)})^+] \right) \right. \\ \left. + \beta E[V_k([Y - \sigma_* - e_m]^+)] \right\} \\ + \beta \{ (E[V_k([Y - \sigma_*]^+)] \\ - E[V_k([Y - \sigma_* - e_m]^+)] \\ - (E[V_k([Y - \sigma_* - e_n]^+)] \\ - E[V_k([Y - \sigma_* - e_m - e_n]^+)])) \}. \end{aligned}$$

By (4.12), the first expression in $\{ \}$ on the RHS of the above identity is nonnegative. By the convexity of $V_k(x^+)$, the second expression in $\{ \}$ on the RHS of the above identity is nonnegative. Thus $V_{k+1}(x)$ is convex.

Case 3: $\alpha = T - 1$. Applying Theorem 3.8, we obtain

$$\begin{aligned} V_{k+1}(x + e_i) &= G_{k+1}(x + e_i, x + e_i) \\ &= \sum_{l=1}^2 E[Y^{(l)}] + \beta E[V_k(Y)], \\ V_{k+1}(x + e_j) &= V_{k+1}(x + e_i), \\ V_{k+1}(x + e_i + e_j) &= G_{k+1}(x + e_i + e_j, x + e_j), \\ &= 1 + \sum_{l=1}^2 E[Y^{(l)}] \\ &\quad + \beta E[V_k(Y + e_j)]. \end{aligned}$$

Moreover by (3.26), we have for some n

$$\begin{aligned} V_{k+1}(x) &= \overline{V_{k+1}}(0, 1) \\ &= E[(Y^{(n)} - 1)^+] + \sum_{l \neq n} E[Y^{(l)}] \\ &\quad + \beta E[V_k([Y - e_n]^+)] \end{aligned}$$

so that

$$\begin{aligned} &V_{k+1}(x + e_i + e_j) - V_{k+1}(x + e_i) \\ &\quad - V_{k+1}(x + e_j) + V_{k+1}(x) \\ &= \{E[(Y^{(n)} - 1)^+] - E[Y^{(n)}] + 1\} \\ &\quad + \beta \{E[V_k([Y - e_n]^+)] - E[V_k(Y)]\} \\ &\quad - (E[V_k(Y)] - E[V_k(Y + e_j)]). \end{aligned}$$

As we have shown earlier, the first expression in $\{ \}$ on the RHS of the above identity is nonnegative. Noting that $E[V_k(Y + e_j)] = E[V_k(Y + e_n)]$, and using the convexity of $V_k(x^+)$, the second expression in $\{ \}$ on the RHS of the above identity is nonnegative. Therefore $V_{k+1}(x)$ is convex. This completes the proof of b) and the theorem. ■

Theorem 4.4: Let $V_k(x) = G_k(x, w_*)$. Then

$$V_k(x + e_i) = \begin{cases} \min_{j=1,2} G_k(x + e_i, w_* + e_i - e_j) \\ \text{if } w_*^{(i)} < T \\ G_k(x + e_i, w_*) & \text{if } w_*^{(i)} = T. \end{cases} \quad (4.13)$$

Proof: Let $x^{(1)} + x^{(2)} = \alpha$, and $\gamma = T - \alpha$. If $\alpha \geq T$, then by Theorem 3.8, $V_k(x + e_i)$ is achieved by w_* and the desired result is immediate. Suppose then that $\alpha < T$. By (3.26)

$$V_k(x + e_i) = \overline{V_k}(0, \gamma - 1) = G_k(x + e_i, x + e_i + \sigma_*)$$

for some $\sigma_* \in B(\gamma - 1)$. Applying Theorem 4.3, we have for some j

$$V_k(x) = \overline{V_k}(0, \gamma) = G_k(x, x + \sigma_* + e_j).$$

Letting $w_* = x + \sigma_* + e_j$, we deduce (4.13). ■

It can be shown by limiting arguments (cf. Remark 1) that the results of Theorems 4.3 and 4.4 are also true for the infinite horizon expected discounted cost problem.

Note that the proof of Theorem 4.3 was based on the submodularity of $V_k(x_1^+, x_2^+)$ in Z^2 as it is implied by the monotonicity, convexity, and submodularity properties of $V_k(x_1, x_2)$ in Z_+^2 , and Lemma 4.2. This technique however, does not work in the proof of Theorem 4.3 for $M > 2$ because $V_k(x_1^+, \dots, x_M^+)$ is not multimodular (see [10] for the definition) in Z^M . More specifically, if part c) of Lemma 4.2 is changed to multimodularity of $g(\cdot)$ (a property that the value function satisfies) in Z_+^M , then multimodularity of $g((\cdot)^+)$ in Z^M is in general not true when $M > 2$.

We conclude this section by noting the practical usefulness of Theorem 4.4. By Theorem 3.8, for each epoch we only need to compute an optimal allocation for each state $(x^{(1)}, x^{(2)}) \in \{(0, 0), (1, 0), \dots, (T - 1, 0)\}$. By Theorem 4.4, after an optimal allocation is found for state $(0, 0)$, the optimal allocation for each sequent state is either the same or differs by one allocated slot. This certainly is a reduction in the computational effort required to determine an optimal allocation policy.

B. Structure of the Optimal Policy When $Y^{(1)}$ is Stochastically Larger than $Y^{(2)}$

In this section, we derive the structure of an optimal allocation policy (as described in the beginning of the section) when the process of message generation at Station 1 is stochastically larger than the process of message generation at Station 2. We prove the intuitive result due to the linearity and equality of the holding costs, that when $x^{(1)} + x^{(2)} = \alpha \leq T$, it is optimal to: 1) allocate $x^{(j)}$ to each transmitter j , and 2) allocate the remaining $(T - \alpha)$ slots in such a way that transmitter 1 gets at least as many slots as transmitter 2.

Theorem 4.5: Let $0 \leq \gamma \leq T$ and suppose $Y^{(1)} \stackrel{st}{\geq} Y^{(2)}$. Then for some $\sigma_* = (\sigma_*^{(1)}, \sigma_*^{(2)}) \in B(\gamma)$ satisfying $\sigma_*^{(1)} \geq \sigma_*^{(2)}$ we have

$$V_k(x) = G_k(x, x + \sigma_*).$$

Proof: Write $V_k(x) = G_k(0, \delta_*) = L(0, \delta_*) + \beta E[V_{k-1}([Y - \delta_*]^+)]$ for some $\delta_* \in B(\gamma)$ and suppose $\delta_*^{(1)} < \delta_*^{(2)}$. We will show that the allocation $\delta_1 = \delta_* + e_1 - e_2$ is also optimal. We have

$$\begin{aligned} G_k(0, \delta_1) - V_k(x) &= L(0, \delta_1) - L(0, \delta_*) \\ &\quad + \beta \{E[V_{k-1}([Y - \delta_1]^+)] - E[V_{k-1}([Y - \delta_*]^+)]\}, \end{aligned} \quad (4.14)$$

$$\begin{aligned} &L(0, \delta_1) - L(0, \delta_*) \\ &= E[(Y^{(2)} - \delta_*^{(2)} + 1)^+ - (Y^{(2)} - \delta_*^{(2)})^+] \\ &\quad - E[(Y^{(1)} - \delta_*^{(1)})^+ - (Y^{(1)} - \delta_*^{(1)} - 1)^+]. \end{aligned}$$

Define the function $G(x)$ of the integer variable x as

$$G(x) \triangleq \begin{cases} x^+ - (x - 1)^+ & \text{if } x \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$L(0, \delta_1) - L(0, \delta_*) = E[G(Y^{(2)} - \delta_*^{(2)} + 1)] \\ - E[G(Y^{(1)} - \delta_*^{(1)})]. \quad (4.15)$$

The convexity of x^+ implies that $G(x)$ is monotonically nondecreasing. Moreover, $\delta_*^{(1)} \leq \delta_*^{(2)} - 1$ and $Y^{(1)} \stackrel{st}{\geq} Y^{(2)}$ imply that $Y^{(1)} - \delta_*^{(1)} \stackrel{st}{\geq} Y^{(2)} - \delta_*^{(2)} + 1$. Hence by [11, p. 252], $G(Y^{(1)} - \delta_*^{(1)}) \stackrel{st}{\geq} G(Y^{(2)} - \delta_*^{(2)} + 1)$ and this entails $E[G(Y^{(1)} - \delta_*^{(1)})] \geq E[G(Y^{(2)} - \delta_*^{(2)} + 1)]$. Using this result in (4.15), we get

$$L(0, \delta_1) \leq L(0, \delta_*). \quad (4.16)$$

We next show that

$$E[V_{k-1}([Y - \delta_1]^+)] \leq E[V_{k-1}([Y - \delta_*]^+)]. \quad (4.17)$$

By (3.9), for all $z = (z_1, z_2) \in \mathbb{Z}^2$

$$V_k([z - e_1 + e_2]^+) - V_k(z^+) \\ = \begin{cases} V_k(z_1 - 1, 0) - V_k(z_1, 0) & \text{if } z_1 \geq 1 \text{ and } z_2 \leq -1 \\ V_k(z_2 + 1, 0) - V_k(z_2, 0) & \text{if } z_1 \leq 0 \text{ and } z_2 \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Letting $z_i = Y^{(i)} - \delta_*^{(i)}$, $i = 1, 2$, we obtain

$$E[V_{k-1}([Y - \delta_1]^+)] - E[V_{k-1}([Y - \delta_*]^+)] \\ = P[Y^{(1)} \leq \delta_*^{(1)}] E[W(Y^{(2)} - \delta_*^{(2)} + 1)] \\ - P[Y^{(2)} \leq \delta_*^{(2)} - 1] E[W(Y^{(1)} - \delta_*^{(1)})] \quad (4.18)$$

where

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

By the monotonicity of $V_{k-1}(x)$, $W(x_1) \geq 0$ and by the convexity of $V_{k-1}(x)$, $W(x_1)$ is monotonically nondecreasing. Moreover, $\delta_*^{(1)} \leq \delta_*^{(2)} - 1$ implies that $Y^{(1)} - \delta_*^{(1)} \stackrel{st}{\geq} Y^{(2)} - \delta_*^{(2)} + 1$. Thus $W(Y^{(1)} - \delta_*^{(1)}) \stackrel{st}{\geq} W(Y^{(2)} - \delta_*^{(2)} + 1)$ and therefore

$$E[W(Y^{(2)} - \delta_*^{(2)} + 1)] \leq E[W(Y^{(1)} - \delta_*^{(1)})]. \quad (4.19)$$

In addition, $Y^{(1)} \stackrel{st}{\geq} Y^{(2)}$ implies that $P[Y^{(1)} \leq \delta_*^{(1)}] \leq P[Y^{(2)} \leq \delta_*^{(2)} - 1]$. Substituting this result and (4.19) in (4.18), we obtain (4.17). Combining (4.16) and (4.17) in (4.14), and using the optimality of δ_* , we obtain $G_k(0, \delta_1) = V_k(x)$. If $(\delta_*^{(1)} + 1) \geq (\delta_*^{(2)} - 1)$, the claim is proved; otherwise let $\delta_2 = \delta_1 + e_1 - e_2$. Then by the same argument $G_k(0, \delta_2) = G_k(0, \delta_1)$. Proceeding in this fashion, we get $G_k(0, \delta_n) = V_k(x)$ where $\delta_n = \delta_* + ne_1 - ne_2$ and n is the smallest integer such that $(\delta_*^{(1)} + n) \geq (\delta_*^{(2)} - n)$. ■

Remark 2: The convexity of the value function $V_k(x)$ for $M = 2$ (Theorem 4.3) is a crucial result in establishing Theorem 4.5. We note that if one shows that the value

function is convex for $M > 2$, then Theorem 4.5 easily generalizes by the same technique.

By taking advantage of the convexity property of the value function, we have been able to further characterize an optimal allocation policy when the arrival processes at the two transmitters are stochastically ordered. The property of an optimal allocation policy, as described by the above theorem, could be used to decrease the number of computations in the search of optimal allocations even further.

An immediate application of Theorem 4.5 is when the arrival processes are i.i.d. In this case, we show that an optimal allocation policy is expressed in closed form.

Corollary 4.6: Let $0 \leq \gamma \leq T$ and suppose $Y^{(1)}$ and $Y^{(2)}$ are i.i.d. Then for any $\sigma_* = (\sigma_*^{(1)}, \sigma_*^{(2)}) \in B(\gamma)$ satisfying $|\sigma_*^{(1)} - \sigma_*^{(2)}| \leq 1$, we have

$$V_k(x) = G_k(x, x + \sigma_*). \quad (4.20)$$

Proof: We first establish the existence of an allocation as stated above. $Y^{(1)}$ and $Y^{(2)}$ are i.i.d. is equivalent to $Y^{(1)} \stackrel{st}{\geq} Y^{(2)}$ and $Y^{(2)} \stackrel{st}{\geq} Y^{(1)}$. By Theorem 4.5, $Y^{(1)} \stackrel{st}{\geq} Y^{(2)}$ implies the existence of an optimal allocation $\delta_* \in B(\gamma)$ such that $\delta_*^{(1)} \geq \delta_*^{(2)}$. If $\delta_*^{(1)} = \delta_*^{(2)}$, then the corollary is immediate; otherwise $\delta_*^{(2)} < \delta_*^{(1)}$ and $Y^{(2)} \stackrel{st}{\geq} Y^{(1)}$ imply, by the same argument in the proof of Theorem 4.5, that the allocation $\sigma_* = \delta_* - ne_1 + ne_2$, where n is the smallest integer such that $(\delta_*^{(2)} + n) - (\delta_*^{(1)} - n) \in \{0, 1\}$, is also optimal. The existence proof is thus established. Next, we show that any allocation as described by the corollary is optimal. Let $\sigma_* = (\sigma_*^{(1)}, \sigma_*^{(2)})$ be optimal and such that $|\sigma_*^{(1)} - \sigma_*^{(2)}| \leq 1$. If $\sigma_*^{(1)} = \sigma_*^{(2)}$, there is nothing to prove, otherwise without loss of generality, suppose $\sigma_*^{(2)} = \sigma_*^{(1)} + 1$. Note that the only other allocation satisfying the hypothesis of the corollary is $\delta_* = (\sigma_*^{(2)}, \sigma_*^{(1)})$. Then clearly, $L(0, \delta_*) = L(0, \sigma_*)$ and

$$E[V_k([Y - \delta_*]^+)] \\ = E[V_k([Y^{(1)} - \sigma_*^{(1)} - 1]^+, [Y^{(2)} - \sigma_*^{(1)}]^+)] \\ = E[V_k([Y^{(2)} - \sigma_*^{(1)}]^+, [Y^{(1)} - \sigma_*^{(1)} - 1]^+)] \\ = E[V_k([Y^{(1)} - \sigma_*^{(1)}]^+, [Y^{(2)} - \sigma_*^{(1)} - 1]^+)] \\ = E[V_k([Y - \sigma_*]^+)]$$

the second and third equalities following from the symmetry of $V_k(\cdot, \cdot)$ and the fact that $Y^{(1)}$ and $Y^{(2)}$ are i.i.d., respectively. Hence $G_k(x, x + \sigma_*) = G_k(x, x + \delta_*)$. ■

For $M = 2$, we have been able to show that an optimal allocation policy is monotone in the state and the value function is convex for general arrival processes. For $M > 2$, these properties are difficult to prove when the arrival processes are general. Motivated by the result of Corollary 4.6, and the problem simplification when the arrival processes are i.i.d., we generalize Corollary 4.6 for $M > 2$ in the next section.

V. THE OPTIMAL POLICY WHEN THE MESSAGE GENERATION PROCESSES AT THE M TRANSMITTERS ARE I.I.D.

In this section, we derive the optimal policy when the processes of message generation at the M transmitters are

independent and identically distributed. We show that there exists an optimal allocation of the form: if $x^{(1)}, x^{(2)}, \dots, x^{(M)}$, the initial states at transmitters 1, 2, \dots , M , respectively, are such that $\sum_{j=1}^M x^{(j)} < T$, then it is optimal to: 1) allocate to each transmitter j $x^{(j)}$ slots, and 2) allocate the remaining $[T - \sum_{j=1}^M x^{(j)}]$ slots to the transmitters as evenly as possible. Cansever and Mito [4], [5] studied this problem but they did not solve it in explicit form.

Theorem 5.1: Assume that $\{Y_i^{(j)}, j = 1, 2, \dots, M\}_{i=0}^\infty$ are i.i.d., and $\alpha \leq T$. Then

a) any allocation scheme $\sigma_* \in B(\gamma)$ such that $|\sigma_*^{(i)} - \sigma_*^{(j)}| \leq 1$ for all i, j is optimal, i.e.,

$$V_k(x) = \bar{V}_k(0, \gamma) = G_k(0, \sigma_*) = G_k(x, x + \sigma_*), \quad (5.1)$$

2) $V_k(x)$ is convex.

Proof: The theorem clearly holds for $k = 0$. By assuming that $V_k(x)$ is convex, we shall show that a) and b) 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. We prove a) for $V_{k+1}(x)$. We first show that there exists an optimal allocation scheme $\sigma_* \in B(\gamma)$ satisfying $|\sigma_*^{(i)} - \sigma_*^{(j)}| \leq 1$ for all i and j . We know that $V_{k+1}(x)$ is achieved by some allocation scheme; so write

$$\begin{aligned} V_{k+1}(x) &= G_{k+1}(0, \delta_*) \\ &= L(0, \delta_*) + \beta E[V_k([Y - \delta_*]^+)] \end{aligned}$$

for some $\delta_* \in B(\gamma)$ and suppose that for some i and j , $\delta_*^{(i)} - \delta_*^{(j)} > 1$. We will show that the policy $\delta_1 = \delta_* - e_i + e_j$ is also optimal. To simplify the notation, rename the variables i and j by 1 and 2, respectively. We have

$$\begin{aligned} G_{k+1}(0, \delta_1) - V_{k+1}(x) &= L(0, \delta_1) - L(0, \delta_*) \\ &\quad + \beta \{E[V_k([Y - \delta_1]^+)] - E[V_k([Y - \delta_*]^+)]\}, \end{aligned} \quad (5.2)$$

$$\begin{aligned} L(0, \delta_1) - L(0, \delta_*) &= \{E[(Y^{(1)} - \delta_*^{(2)} - 1)^+] - E[(Y^{(1)} - \delta_*^{(2)})^+]\} \\ &\quad - \{E[(Y^{(1)} - \delta_*^{(1)})^+] - E[(Y^{(1)} - \delta_*^{(1)} + 1)^+]\}. \end{aligned}$$

As $\delta_*^{(2)} + 1 < \delta_*^{(1)}$, then by the convexity of $(\cdot)^+$, the RHS of the above equality is nonpositive. Hence

$$L(0, \delta_1) \leq L(0, \delta_*). \quad (5.3)$$

We next show that

$$E[V_k([Y - \delta_1]^+)] \leq E[V_k([Y - \delta_*]^+)]. \quad (5.4)$$

We have

$$\begin{aligned} E[V_k([Y - \delta_1]^+) - V_k([Y - \delta_*]^+)] &= E\{E[V_k([Y - \delta_* + e_1 - e_2]^+) \\ &\quad - V_k([Y - \delta_*]^+) | Y^{(l)}, l \neq 1, l \neq 2]\}. \end{aligned}$$

(5.4) will hold if we show that

$$\begin{aligned} A \triangleq E[V_k([Y - \delta_* + e_1 - e_2]^+) \\ - V_k([Y - \delta_*]^+) | Y^{(l)}, l \neq 1, l \neq 2] \leq 0. \end{aligned}$$

Using (3.9), then for all $z = (z_1, z_2, \bar{z})$, where $\bar{z} = (z_3, \dots, z_M)$

$$\begin{aligned} V_k([z + e_1 - e_2]^+) - V_k(z^+) &= \begin{cases} V_k(z_2 - 1, 0, \bar{z}^+) - V_k(z_2, 0, \bar{z}^+) \\ \quad \text{if } z_1 \leq -1 \text{ and } z_2 \geq 1 \\ V_k(z_1 + 1, 0, \bar{z}^+) - V_k(z_1, 0, \bar{z}^+) \\ \quad \text{if } z_1 \geq 0 \text{ and } z_2 \leq 0 \\ 0 \quad \text{otherwise.} \end{cases} \end{aligned}$$

Letting $z_l = Y^{(l)} - \delta_*^{(l)}$ for each l , we obtain

$$\begin{aligned} A &= P[Y^{(1)} \leq \delta_*^{(2)}] E[W(Y^{(1)} - \delta_*^{(1)} + 1)] \\ &\quad - P[Y^{(1)} \leq \delta_*^{(1)} - 1] E[W(Y^{(1)} - \delta_*^{(2)})] \end{aligned} \quad (5.5)$$

where

$$W(x_1) \triangleq \begin{cases} V_k(x_1, 0, \bar{z}^+) - V_k(x_1 - 1, 0, \bar{z}^+) \\ \quad \text{if } x_1 \geq 1 \\ 0 \quad \text{if } x_1 \leq 0. \end{cases}$$

By the monotonicity of $V_k(x)$, $W(x_1) \geq 0$, and by the convexity of $V_k(x)$, $W(x_1)$ is monotonically nondecreasing. Since $\delta_*^{(2)} \leq \delta_*^{(1)} - 1$, then

$$P[Y^{(1)} \leq \delta_*^{(2)}] \leq P[Y^{(1)} \leq \delta_*^{(1)} - 1], \quad (5.6)$$

$$E[W(Y^{(1)} - \delta_*^{(1)} + 1)] \leq E[W(Y^{(1)} - \delta_*^{(2)})]. \quad (5.7)$$

Substituting (5.6) and (5.7) in (5.5), we obtain $A \leq 0$, and thus (5.4) holds. Combining (5.3) and (5.4) in (5.2), and using the optimality of δ_* , we obtain $G_{k+1}(0, \delta_1) = V_{k+1}(x)$. If $(\delta_*^{(1)} - 1) - (\delta_*^{(2)} + 1) \in \{0, 1\}$, the claim is proved; otherwise $((\delta_*^{(1)} - 1) - (\delta_*^{(2)} + 1) > 1)$, let $\delta_2 = \delta_1 - e_1 + e_2$. Then by the same argument $G_{k+1}(0, \delta_2) = G_{k+1}(0, \delta_1) = V_{k+1}(x)$. Proceeding in this fashion, we get $G_{k+1}(0, \delta_n) = V_{k+1}(x)$ where $\delta_n = \delta_* - ne_1 + ne_2$ and n is the smallest integer such that $(\delta_*^{(1)} - n) - (\delta_*^{(2)} + n) \in \{0, 1\}$. By repeating this procedure for all i and j such that $\delta_*^{(i)} - \delta_*^{(j)} > 1$, we thus have showed how to construct an allocation scheme σ_* satisfying $|\sigma_*^{(i)} - \sigma_*^{(j)}| \leq 1$ for all i and j such that $V_{k+1}(x) = G_{k+1}(0, \sigma_*)$. This finishes the existence proof.

We now show that any allocation $\sigma_* \in B(\gamma)$ satisfying $|\sigma_*^{(i)} - \sigma_*^{(j)}| \leq 1$ for all i and j is optimal. Let $\sigma_* = (\sigma_*^{(1)}, \dots, \sigma_*^{(M)}) \in B(\gamma)$ be such an optimal allocation. If $\sigma_*^{(i)} = \sigma_*^{(j)}$ for all i and j , then σ_* is unique and a) is proved. Suppose then $\sigma_*^{(i)} \neq \sigma_*^{(j)}$ for some i and j . Let $p \geq 1$ be the number of components of σ_* that are equal, say to a , and $q = M - p$ be the number of components of σ_* that are equal to $a + 1$. Then $\gamma = pa + (M - p)(a + 1)$ or equivalently $\gamma = Ma + M - p$. By Euclid's division, then a is unique. It then follows that any $\delta_* \in B(\gamma)$ satisfying $|\delta_*^{(i)} - \delta_*^{(j)}| \leq 1$ for all i and j is a permutation of σ_* , i.e., $\delta_* = (\sigma_*^{(i_1)}, \sigma_*^{(i_2)}, \dots, \sigma_*^{(i_M)})$ for some distinct indexes

i_1, i_2, \dots, i_M . Then

$$\begin{aligned} L(0, \delta_*) &= \sum_{l=1}^M E[(Y^{(l)} - \sigma_*^{(i_l)})^+] \\ &= \sum_{l=1}^M E[(Y^{(l)} - \sigma_*^{(l)})^+] = L(0, \sigma_*). \end{aligned}$$

By the symmetry of $V_k(x)$

$$\begin{aligned} V_k([Y - \delta_*]^+) &= V_k([Y^{(1)} - \sigma_*^{(i_1)}]^+, \dots, [Y^{(M)} - \sigma_*^{(i_M)}]^+) \\ &= V_k([Y^{(j_1)} - \sigma_*^{(1)}]^+, \dots, [Y^{(j_M)} - \sigma_*^{(M)}]^+) \end{aligned}$$

for some distinct indexes j_1, j_2, \dots, j_M . Since $\{Y^{(j_i)}, i = 1, 2, \dots, M\}$ are i.i.d., then $E[V_k([Y - \delta_*]^+)] = E[V_k([Y - \sigma_*]^+)]$. Hence $G_{k+1}(0, \delta_*) = G_{k+1}(0, \sigma_*) = V_{k+1}(x)$. The proof of a) is now complete.

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

Case 1: $\alpha \geq T$. Let $V_{k+1}(x) = G_{k+1}(x, w_*)$ for some $w_* \in B(T)$. By Theorem 3.8, there exists an optimal w_* such that $w_*^{(l)} \leq x^{(l)}$ for each l . Then

$$V_{k+1}(x) = \alpha - T + ME[Y] + \beta E[V_k(Y + x - w_*)].$$

Applying Theorem 3.8 again, i.e., w_* is also optimal when the system state is $x + e_i$ or $x + e_j$ or $x + e_i + e_j$, we obtain

$$\begin{aligned} V_{k+1}(x + e_i) &= \alpha + 1 - T + ME[Y^{(1)}] \\ &\quad + \beta E[V_k(Y + x + e_i - w_*)], \end{aligned}$$

$$\begin{aligned} V_{k+1}(x + e_j) &= \alpha + 1 - T + ME[Y^{(1)}] \\ &\quad + \beta E[V_k(Y + x + e_j - w_*)], \end{aligned}$$

$$\begin{aligned} V_{k+1}(x + e_i + e_j) &= \alpha + 2 - T + ME[Y^{(1)}] \\ &\quad + \beta E[V_k(Y + x + e_i + e_j - w_*)]. \end{aligned}$$

Thus by the induction hypothesis, i.e., the convexity of $V_k(x)$, we see that $V_{k+1}(x)$ is convex.

Case 2: $\alpha \leq T - 2$. By (3.26)

$$V_{k+1}(x + e_i + e_j) = \overline{V_{k+1}}(0, \gamma - 2) = G_{k+1}(0, \sigma_*)$$

for some $\sigma_* \in B(\gamma - 2)$ satisfying, by a), $|\sigma_*^{(m)} - \sigma_*^{(n)}| \leq 1$ for all m and n . Without loss of generality, assume $\sigma_*^{(1)} \leq \sigma_*^{(2)} \leq \dots \leq \sigma_*^{(M)}$. Note that such an ordering is justified by a). Clearly, $\sigma_* + e_1 \in B(\gamma - 1)$ satisfies the hypothesis of a). Thus

$$\begin{aligned} V_{k+1}(x + e_i) &= V_{k+1}(x + e_j) \\ &= \overline{V_{k+1}}(0, \gamma - 1) = G_{k+1}(0, \sigma_* + e_1). \end{aligned}$$

Moreover by the optimality of $\sigma_* + e_1$, we have

$$\forall l, \quad G_{k+1}(0, \sigma_* + e_l) \geq G_{k+1}(0, \sigma_* + e_1).$$

or equivalently, $\forall l$

$$\begin{aligned} &E[(Y^{(1)} - \sigma_*^{(l)} - 1)^+] - E[(Y^{(1)} - \sigma_*^{(l)})^+] \\ &\quad + \beta E[V_k([Y - \sigma_* - e_l]^+)] \\ &\geq E[(Y^{(1)} - \sigma_*^{(1)} - 1)^+] - E[(Y^{(1)} - \sigma_*^{(1)})^+] \\ &\quad + \beta E[V_k([Y - \sigma_* - e_1]^+)]. \end{aligned} \quad (5.8)$$

Note also that $\sigma_* + e_1 + e_2 \in B(\gamma)$ satisfies the hypothesis of part a). Thus

$$V_{k+1}(x) = \overline{V_{k+1}}(0, \gamma) = G_{k+1}(0, \sigma_* + e_1 + e_2).$$

We then obtain

$$\begin{aligned} &V_{k+1}(x + e_i + e_j) - V_{k+1}(x + e_i) \\ &= E[(Y^{(1)} - \sigma_*^{(1)})^+] - E[(Y^{(1)} - \sigma_*^{(1)} - 1)^+] \\ &\quad + \beta \{E[V_k([Y - \sigma_*]^+)] \\ &\quad - E[V_k([Y - \sigma_* - e_1]^+)]\}, \\ &V_k(x + e_j) - V_k(x) \\ &= E[(Y^{(1)} - \sigma_*^{(2)})^+] - E[(Y^{(1)} - \sigma_*^{(2)} - 1)^+] \\ &\quad + \beta \{E[V_k([Y - \sigma_* - e_1]^+)] \\ &\quad - E[V_k([Y - \sigma_* - e_1 - e_2]^+)]\}, \end{aligned}$$

so that

$$\begin{aligned} &V_{k+1}(x + e_i + e_j) - V_{k+1}(x + e_i) \\ &\quad - V_{k+1}(x + e_j) + V_{k+1}(x) \\ &= \left\{ \left(E[(Y^{(1)} - \sigma_*^{(2)} - 1)^+] - E[(Y^{(1)} - \sigma_*^{(2)})^+] \right) \right. \\ &\quad + \beta E[V_k([Y - \sigma_* - e_2]^+)] \\ &\quad - \left(E[(Y^{(1)} - \sigma_*^{(1)} - 1)^+] - E[(Y^{(1)} - \sigma_*^{(1)})^+] \right) \\ &\quad + \beta E[V_k([Y - \sigma_* - e_1]^+)] \left. \right\} \\ &\quad + \beta \{ (E[V_k([Y - \sigma_*]^+)] \\ &\quad - E[V_k([Y - \sigma_* - e_1]^+)] \\ &\quad - (E[V_k([Y - \sigma_* - e_2]^+)] \\ &\quad - E[V_k([Y - \sigma_* - e_1 - e_2]^+)])) \}. \end{aligned}$$

By (5.8), the first expression in $\{ \}$ on the RHS of the above identity is nonnegative. By the convexity of $V_k(x^+)$, the second expression in $\{ \}$ on the RHS of the above identity is nonnegative. Thus $V_{k+1}(x)$ is convex.

Case 3: $\alpha = T - 1$. Applying Theorem 3.8, we obtain

$$\begin{aligned} V_{k+1}(x + e_i) &= V_{k+1}(x + e_j) = G_{k+1}(x + e_i, x + e_j) \\ &= ME[Y^{(1)}] + \beta E[V_k(Y)], \\ V_{k+1}(x + e_i + e_j) &= G_{k+1}(x + e_i + e_j, x + e_j) \\ &= 1 + ME[Y^{(1)}] + \beta E[V_k(Y + e_i)]. \end{aligned}$$

Moreover by (3.26)

$$V_{k+1}(x) = \overline{V_{k+1}}(0, 1) = G_{k+1}(0, e_i)$$

the last equality being justified by a). Then

$$\begin{aligned} &V_{k+1}(x + e_i + e_j) - V_{k+1}(x + e_i) \\ &\quad - V_{k+1}(x + e_j) + V_{k+1}(x) \\ &= \{1 - E[Y^{(1)}] + E[(Y^{(1)} - 1)^+]\} \\ &\quad + \beta \{ (E[V_k([Y + e_i]^+)] - E[V_k(Y)]) \\ &\quad - (E[V_k(Y)] - E[V_k([Y - e_i]^+)])) \}. \end{aligned}$$

Clearly, the first expression in $\{ \}$ on the RHS of the above identity is nonnegative. By the convexity of $V_k(x^+)$, the second expression in $\{ \}$ on the RHS of the above identity is nonnegative. Therefore $V_{k+1}(x)$ is convex. This completes the proof of b) and the theorem. ■

CONCLUSION

In this paper, we have generalized the model of [4], [5] to $M \geq 2$. By deriving several qualitative properties of an optimal allocation policy, we have shown that finding dynamic optimal allocations for only T states suffices to completely describe an optimal allocation policy. Furthermore, when $M = 2$, for all but one of the T states there are only two allocations that are candidates for optimality. It is desirable to extend this property to $M > 2$. However, it appears that the technique used for $M = 2$ does not generalize to $M > 2$. New methods to prove such properties are a subject of further investigation. The case when the holding costs at the transmitters are distinct will appear in a future publication [14].

REFERENCES

- [1] M. Gerla and L. Kleinrock, "Flow control: A comparative study," *IEEE Trans. Commun.*, vol. COMM-28, pp. 533-574, 1980.
- [2] —, "Flow control protocols," in *Computer Network Architectures and Protocols*, P. E. Green, Jr., Ed. New York: Plenum, 1982, pp. 361-412.
- [3] M. Schwartz, *Computer Communication Networks, Design and Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1977.
- [4] 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.
- [5] —, "Optimal hop-by-hop flow control with multiple heterogeneous transmitters," in *Proc. 27th IEEE Conf. Decision Contr.*, Austin, TX, Dec. 1988, pp. 1291-1296.
- [6] Z. Rosberg and I. S. Gopal, "Optimal hop-by-hop flow control in computer networks," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 813-822, 1986.
- [7] H. Kushner, *Introduction to Stochastic Control*. New York: Holt, Rinehart, and Winston, 1971.
- [8] S. A. Lippman, "On dynamic programming with unbounded rewards," *Management Sci.*, vol. 21, pp. 1225-1233, 1975.

- [9] L. Kleinrock, *Queueing Systems, Vol. 1, Theory*. New York: Wiley, 1975.
- [10] B. Hajek, "Extremal splittings of point processes," *Math. Operat. Res.*, vol. 10, no. 4, pp. 543-556, Nov. 1985.
- [11] S. M. Ross, *Stochastic Processes* (Wiley Series in Probability and Mathematical Statistics). New York: Wiley, 1983.
- [12] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Englewood Cliffs, NJ: Prentice-Hall, 1986.
- [13] F. J. Beutler and D. Teneketzis, "Routing in queueing networks under imperfect information: Stochastic dominance and thresholds," *Stochastics and Stochastics Reports*, vol. 26, pp. 81-100, 1989.
- [14] R. M. Bournas, F. J. Beutler, and D. Teneketzis, "Optimal flow control allocation policies in communication networks with priorities," Tech. Rep., Dep. Electr. Eng. Comp. Sci., Univ. Michigan, Ann Arbor, Jan. 1991.
- [15] J. E. Wieselthier and A. Ephremides, "New classes of protocols for multiple access in satellite networks," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 5, pp. 865-879, 1980.
- [16] —, "Channel-access control for integrated voice/data networks," in *Proc. 29th Conf. Decision Contr.*, 1990, pp. 921-924.



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

His research interests include multidimensional queueing systems and networking systems (transmission protocols, performance analysis).

Dr. Bournas was awarded a three-year IBM graduate fellowship from 1987 to 1990. He received the outstanding graduate student 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) 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 editorial 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, Patras, Greece, in 1974, and the M.S., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, 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. His current research interests include stochastic systems and control, decentralized systems, queueing networks, communication networks, stochastic scheduling, and resource allocation problems.