

*Section V*

**Flow Control**



## TIME-AVERAGE AND ASYMPTOTICALLY OPTIMAL FLOW CONTROL POLICIES IN NETWORKS WITH MULTIPLE TRANSMITTERS

Redha M. BOURNAS, Frederick J. BEUTLER and Demosthenis TENEKETZIS  
*Department of Electrical Engineering and Computer Science, The University of Michigan,  
Ann Arbor, MI 48109, USA*

### Abstract

We consider  $M$  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. Our objective is to characterize policies that minimize the long-term average of the total number of messages awaiting service at the  $M$  transmitters. We establish necessary and sufficient conditions on the arrival processes at the transmitters for the existence of finite cost time-average policies; it is not enough that the average arrival rate is strictly less than the slot capacity. We construct a pure strategy that attains a finite average cost under these conditions. This in turn leads to the existence of an optimal time-average pure policy for each phase length  $T$ , and to upper and lower bounds on the cost this policy achieves. Furthermore, we show that such an optimal time-average policy has the same properties as those of optimal discounted policies investigated by the authors in a previous paper. Finally, we prove that in the absence of costs accrued by messages within the phase, there exists a policy such that the time-average cost tends toward zero as the phase length  $T \rightarrow \infty$ .

### 1. Introduction

We consider a flow control problem that arises in the performance modelling 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 [7, 8, 17]. The hop-by-hop scheme studied in this paper is the same as the one in [2–5], its purpose being to maintain a smooth flow of traffic between  $M$  transmitting stations attempting to send messages through a single communication channel to an adjacent receiving station. The time axis 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  $j$  has an independent generally distributed arrival process of packets per slot with finite first moment  $\lambda^{(j)}$  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.

$T$  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 during which each transmitter is allowed to transmit. In making a decision on the assigned window size, the receiver uses the knowledge of arrival statistics, the number of queued packets for each transmitter at the beginning of the preceding phase, and the window size assigned for the preceding phase. The number of queued packets at a phase is sent by each transmitter to the receiver with negligible overhead some time before the beginning of the next phase. Due to the arrival of new packets, the number of queued packets changes by the time the receiver is able to use the information for the next window assignment. The window allocations by the receiver thus constitute a discrete-time Markov decision process with partial information.

Optimal flow control allocations were first analyzed by Rosberg and Gopal [16], who considered a single transmitter, and a cost function reflecting the number of queued packets together with the number of unutilized (i.e. wasted) transmission slots. Subsequently, Cansever and Milito [3] investigated the problem for two transmitters ( $M = 2$ ) with identical arrival statistics, with later generalizations to heterogeneous arrivals [4]. Cansever and Milito also conjectured results for  $M > 2$  transmitters. Later, they extended their work to more complex networks with multiple states in a layered tree-like network [5].

Our model in [2] is similar to that of [3,4]. As in these references, the cost per phase in [2] is the expectation of the sum of the number of untransmitted packets at the respective stations. Our objective then was to dynamically allocate a fixed number  $T$  of slots among the  $M \geq 2$  transmitters to minimize the total discounted cost. Our results in [2] include a partial characterization of a set of optimal allocation policies. These structural properties enable us to prove that, for the set of all discount factors  $\beta < 1$ , a finite number of dynamic optimal allocations suffice to completely describe an optimal allocation policy. For  $M = 2$ , we prove in addition that the optimal policy is a monotone function of a state, and that the total cost is convex. When the process of message generation at one transmitter is stochastically larger than the message generation process at the other transmitter, we further characterize an optimal allocation. Finally, if the message-generating processes at the  $M \geq 2$  transmitter are i.i.d., we find an explicit form of the optimal allocation policy (compare [3]) that does not depend on the discount factor  $\beta$ .

Here, we turn our attention to time-average policies. We believe that a time-average cost criterion is a more natural setting for flow control problems, since it represents the long-term performance measure of the flow control algorithm. Moreover, in the time scale under which a flow control system generally operates, any discount attached to past data ought to be minimal; hence, discounting appears to us to be an artifice that facilitates solutions, at the cost of detracting from the validity of the model.

It is not surprising that the existence of time-average policies of finite cost requires that the average arrival rate is strictly less than the slot capacity, i.e.  $\rho < 1$

in terms of a traffic intensity. We show that more is required: if  $\rho < 1$ , a necessary and sufficient condition for the existence of finite cost strategies is the finiteness of the second moment of the number of arrivals during a phase.

We exhibit a pure strategy that attains a finite average cost under the condition of the preceding paragraph. This in turn leads to four further results:

- (1) For each phase length  $T$ , there exists an optimal time-average policy.
- (2) The time-average optimal policy can be obtained as a limit of infinite horizon optimal discounted policies as the discounting factor  $\beta \rightarrow 1$ .
- (3) The properties of the time-average optimal policy are the same as those derived in [2] for optimal discounted policies.
- (4) Upper and lower performance bounds are obtained for the cost attained by the optimal policy.

For each phase length  $T$ , the existence of time-average optimal policies and the derivation of their structural properties are based on: (1) the work of Bournas et al. [2] on the infinite horizon optimal discounted cost and structural properties of the optimal discounted stationary policies, and (2) the work of Sennott [18].

Finally, we prove that in the absence of costs accrued by messages within the phase, there exists a policy such that the time-average cost tends toward zero as the phase length  $T \rightarrow \infty$ . This result is motivated by our interest in investigating the long-term average cost as a function of the phase length  $T$ , which in turn leads to an optimal phase length size. This problem turns out to be a very difficult one and we have not solved it in this paper. However, we have been able to solve a closely related problem, as we shall explain. For each phase, the cost has two additive components: (1) the number of packets awaiting transmission at the beginning of the phase, and (2) the waiting times accumulated by packets arriving during the phase, and not being available for transmission until the beginning of the next phase. Using the strong law of large numbers and the theory of convergence of probability measures as in Billingsley [1], we have been able to show that the cost component (1) asymptotically goes to zero at  $T \rightarrow \infty$ . In addition, the corresponding asymptotically optimal policies are state independent and proportional to the arrival processes rates. The asymptotical behaviour of component (1) at  $T \rightarrow \infty$  combined with the monotonicity in  $T$  of the cost component (2) shed more light on the issues that should be addressed to determine the optimal phase length. Some of these issues are discussed in section 5.

The paper is organized as follows. In section 2, we formalize the model and formulate the problem. In section 3, we first derive necessary conditions on the statistics of the arrival processes that ensure system-stable behaviour. Under these conditions, we then construct a pure strategy possessed of finite long-term average cost. In section 4, we demonstrate the existence of optimal time-average stationary policies, and show that these policies have the same properties as those derived in [2] for optimal discounted stationary strategies. In section 5, we exhibit the

existence of a stationary nonrandomized policy under which the long-term average number of packets awaiting transmission at the beginning of each phase converges to zero as  $T \rightarrow \infty$ . The flow control problem with priorities is briefly discussed in section 6. Finally, conclusions are presented in section 7.

## 2. Model formulation

### 2.1. DEFINITIONS AND PROBLEM STATEMENT

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 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 for the delay within the phase. These assumptions may be considered as a restriction of the model. Relaxing them results in formulating a 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 not be addressed here.

The processes of message generation at each transmitter are stochastic processes with known statistics. The number of packets generated at transmitter  $j$  during slot  $i$   $\{\xi_i^{(j)}\}_{i=1}^{\infty}$ , are i.i.d. random variables with finite first moment  $\lambda^{(j)}$ . We assume that the arrival processes to distinct stations are mutually independent.

Let  $Y_k^{(j)}$  be the number of packets generated at transmitter  $j$  ( $j = 0, 1, \dots, M$ ) during phase  $k$  ( $k = 0, 1, \dots$ ). For each  $j$ , the  $Y_k^{(j)}$  are i.i.d. in  $k$ . Indeed, each  $Y_k^{(j)}$  is the sum of  $T$  i.i.d. random variables representing packets generated at the respective slots of the phase. For simplicity, we often use the notation  $Y_k$  to denote the vector whose  $M$  components are the  $Y_k^{(j)}$ .

We now write the evolution equations for the flow control system. The number of packets in the buffer at the beginning of phase  $k$  is called  $N_k$ ; the same vector convention holds for  $Y_k$ . In a similar vein, define  $w_k$  as the allocation vector for phase  $k$ ,  $w_k^{(j)}$  being the number of slots assigned to transmitter  $j$  for phase  $k$ . Finally, we let  $X_k$  be the number of packets "left over" at the beginning of phase  $k$ , in the sense that they had been buffered at the beginning of phase  $k-1$ , but not transmitted during the course of that phase. More precisely, we shall define  $X_k$  by the relation

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

where we adopt the vector notation

$$x^+ \triangleq \sum_{j=1}^M \max(0, x^{(j)}).$$

Inherent in (2.1) is the assumption that arrivals during phase  $k - 1$  cannot be transmitted during that phase, but are available for transmission in phase  $k$ . It follows that the number of buffered packets for the respective transmitters at the beginning of phase  $k$  is described by

$$N_k = X_k + Y_{k-1}. \tag{2.2}$$

Observe that (2.2) holds only for  $k \geq 1$ ; to complete the set of dynamic equations, we assume that  $N_0$  and  $w_0$  are given. From (2.1) and (2.2), we thus arrive at the dynamical equations of evolution

$$X_k = \begin{cases} (X_{k-1} + Y_{k-2} - w_{k-1})^+ & \text{if } k \geq 2, \\ x_1 & \text{if } k = 1. \end{cases} \tag{2.3}$$

It will be seen that  $\{X_k\}$  turns out to be the more natural state variable, not only for the evolution equations, but also for the cost expressions and the allocation rules.

It is convenient as well as reasonable to define for the cost function during a single phase

$$r(N) = \sum_{j=1}^M \sum_{k=0}^{T-1} \left[ N^{(j)} + \sum_{m=1}^k \xi_{i,m}^{(j)} \right], \tag{2.4}$$

where we have used  $\xi_{i,m}^{(j)}$  to denote the number of packet arrivals to transmitter  $j$  during slot  $m$  of any phase  $i$ . The cost (2.4) may be interpreted as a total waiting time via Little's formula, except that we have not found it feasible to take account of the particular slot during which a packet is transmitted. From (2.3), it follows that the expected cost per phase is furnished by

$$E[r(N)] = T \sum_{j=1}^M E[X^{(j)}] + \frac{T(T+1)}{2} \sum_{j=1}^M \lambda^{(j)}. \tag{2.5}$$

For the allocation  $w$  determined by the receiver, we first define the action space  $A$  to describe the possible allocations of slots within a phase, namely

$$A \triangleq \left\{ w = (w^{(1)}, w^{(2)}, \dots, w^{(M)}) \in \mathbf{Z}_+^M : \sum_{j=1}^M w^{(j)} \leq T \right\}. \tag{2.6}$$

During the course of phase  $k - 1$ , each transmitter is able (with negligible overhead) to apprise the receiver of his value of  $N_{k-1}^{(j)}$ . Since the receiver also knows  $w_{k-1}^{(j)}$ , he can deduce  $X_k^{(j)}$  by the relation (2.1). It follows from the evolution equation (2.3) that  $\{X_k\}$  is a Markov decision process, whose optimal control requires only the most recently available state (cf. [13], section 6.7). In short, we need only consider the set of admissible control policies  $\pi = \prod_{k=1}^{\infty} \pi_k$ , where

$$\pi_k : \mathbf{Z}_+^M \rightarrow \mathbf{A}. \quad (2.7)$$

More succinctly, we write  $w_k(X_k)$  to indicate that the allocation of slots for phase  $k$  is based on the current state  $X_k$ . We emphasize once again that our allocation is based on imperfect information; at phase  $k$ ,  $X_k$  represents data from the beginning of phase  $k - 1$ , and does not take into account the arrivals  $Y_{k-1}$  that contribute to the current buffer content  $N_k$ . For future reference, the set of admissible controls as described above will be called  $\mathbf{P}_T$ . Finally, when we consider only stationary policies, we shall omit the subscript from  $w_k$ .

When a policy  $\pi$  is employed, we define the long-term average cost for phase length  $T$  by taking the time average of the expected cost (2.5), and conditioning on the beginning state  $x$ . We thus obtain

$$W_T(\pi, x) = \limsup_{n \rightarrow \infty} n^{-1} \sum_{k=1}^n \sum_{j=1}^M E_x^\pi [X_k^{(j)}] + \frac{T+1}{2} \sum_{j=1}^M \lambda^{(j)}. \quad (2.8)$$

Attention is called to the second term, which varies only with  $T$ , and does not depend on the allocation algorithm for the slots. Because this term plays no role in the choice of optimal allocation, we shall focus our attention primarily on the first term.

If we call the total number of unsent packets at the beginning of phase  $n$

$$S_n \triangleq \sum_{j=1}^M X_n^{(j)}, \quad (2.9)$$

then for each fixed  $T$  there is no loss of generality (with respect to either system stability or optimality of flow control) in restricting attention to the cost function

$$\bar{V}_T(\pi, x) \triangleq \limsup_{n \rightarrow \infty} \frac{1}{n} E_x^\pi \left[ \sum_{k=1}^n S_k \right]. \quad (2.10)$$

We also define the value function

$$\bar{V}_T(x) \triangleq \inf_{\pi \in \mathbf{P}_T} \bar{V}_T(\pi, x). \quad (2.11)$$

We are now ready to state the problems we address in this paper:

- (1) For each fixed  $T$ , find necessary and sufficient conditions on the statistics of the arrival processes at the transmitters that guarantee stable system behaviour. That is, under these conditions there will exist flow control policies possessed of finite long-term average cost  $\bar{V}_T(\cdot, \cdot)$ . This problem is decomposed into two subproblems: (a) establish the necessary stability conditions first, and (b) show that these conditions are sufficient for system stability by constructing a pure policy possessed of finite long-term average cost for each phase length  $T$ .
- (2) For each fixed  $T$ , establish the existence of an optimal flow control policy for the long-term average cost criterion as given by (2.10) and derive qualitative properties of this optimal control policy.
- (3) Explore the asymptotic behaviour of  $\bar{V}_T(x)$  as  $T \rightarrow \infty$ .

It is worthwhile at this point to make the following observation. In (3) above, we study the behaviour of the minimal asymptotic long-term average cost consisting only of the number of packets awaiting transmission at the beginning of each phase as  $T \rightarrow \infty$ . We do not incorporate the additional cost due to holding packets arriving within each phase. The long-term average of the sum of these two costs is as given by (2.9). By inspection, the second term on the right-hand side of (2.8) tends toward infinity as  $T \rightarrow \infty$  while, as we shall show in section 5, the first term can be made to tend toward zero under some specified allocation strategy. Therefore, there exists at least one optimal phase length, say  $T^*$ , that minimizes the optimal value of  $W_T$  over  $T$ . We will elaborate on this optimal design problem in section 5.

We now move to the next subsection to discover a fundamental relation between the state evolution of this queueing network and the waiting time process of some  $G/G/1$  queue, that we shall shortly define.

## 2.2. RELATION TO THE $G/G/1$ QUEUE WAITING TIME PROCESS

In this subsection, we relate the state evolution of this queueing network to the waiting time process of a  $G/G/1$  queue. The key to this connection is the following simple inequality: for any real numbers  $\{x_i^{(k)}, 1 \leq i \leq n, 1 \leq k \leq m\}$ , we have

$$\max \left\{ \sum_{k=1}^m x_i^{(k)} : 1 \leq i \leq n \right\} \leq \sum_{k=1}^m \max \{ x_i^{(k)} : 1 \leq i \leq n \}. \tag{2.12}$$

Applying (2.12) to the sum of the components of (2.3), we obtain

$$\sum_{j=1}^M X_n^{(j)} \geq \left( \sum_{j=1}^M X_{n-1}^{(j)} + Y_{n-2}^{(j)} - w_{n-1}^{(j)} \right)^+. \tag{2.13}$$

We first define

$$Y_n \triangleq \sum_{j=1}^M Y_n^{(j)}. \quad (2.14)$$

Since  $\sum_{j=1}^M w_k^{(j)} \leq T$  for all  $k$ , we then infer from (2.9) and (2.13) that

$$S_n \geq (S_{n-1} + Y_{n-2} - T)^+. \quad (2.15)$$

We next define the random process

$$W_n \triangleq \begin{cases} S_n & \text{if } n = 1, \\ (W_{n-1} + Y_{n-1} - T)^+ & \text{if } n > 1. \end{cases} \quad (2.16)$$

The random process  $\{W_n\}$  behaves exactly like the waiting time for a  $D/G/1$  queue with interarrival times equal to the phase length  $T$ , and i.i.d. service times  $\{Y_n\}$ , see Lindley [14]. Combining (2.15) and (2.16) produces the relation

$$S_n \geq W_n, \quad (2.17)$$

which is to say that under any control policy, the total cost at epoch  $n$  is at least as high as the waiting time of the  $n$ th customer of a  $D/G/1$  system. The relation (2.17) is of particular importance in the sequel, for it will enable us to draw the following conclusions:

- (i) any unstable behaviour of  $\{W_n\}$  will imply unstable behaviour of  $\{S_n\}$ ;
- (ii) the minimal long-term average cost  $\bar{V}_T(x)$  is always higher than the time average of  $\{W_n\}$ , whenever the latter exists.

In the next section, we will then take advantage of the established theoretical results of the  $G/G/1$  queue [14,10,11] and the theory of random walks [6,9] to study the stability behaviour of the queueing network.

### 3. Existence of finite average cost policies

In this section, we derive necessary and sufficient conditions on the statistics of the arrival processes at the transmitters that will guarantee the existence of finite long-term average cost flow control policies. In this regard, we shall first define system stability.

#### DEFINITION

The system is stable under control policy  $\pi$  if the long-term average cost (2.10) is finite.

In the remainder of the paper, we shall set

$$\rho \triangleq \sum_{j=1}^M \lambda^{(j)}. \tag{3.1}$$

We shall establish the following results.

- (1) If  $\rho > 1$ , or  $\rho = 1$  and the arrival processes at the transmitters are not all deterministic, the system is unstable; that is, under any control policy, the expected number of unsent messages grows without bound.
- (2) If  $\rho < 1$  and the second moment of the arrival process at some transmitter is infinite, the system is unstable under any control policy.
- (3) If  $\rho < 1$  and the second moment of the arrival process at each transmitter is finite, we construct a nonrandomized stationary strategy that leads to a stable system for each phase length  $T$ .

### 3.1. UNSTABLE BEHAVIOUR

We begin by proving assertion (1), and then describe a policy with finite long-term average cost when  $\rho = 1$  and the arrivals are deterministic.

#### THEOREM 3.1

If  $\rho > 1$ , or  $\rho = 1$  and the arrival processes at the transmitters are not all deterministic, then the system is unstable under any allocation policy in the sense that for any control policy  $\pi$  and any initial state  $X_1 = x$ ,

- (a) If  $\rho > 1$ , then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n S_k(\pi) = \lim_{n \rightarrow \infty} S_n(\pi) = +\infty \quad \text{w.p.1}; \tag{3.2}$$

- (b) If  $\rho = 1$ , then

$$\limsup_{n \rightarrow \infty} S_n(\pi) = +\infty \quad \text{w.p.1}. \tag{3.3}$$

Before we prove the theorem, we need the following lemma. It will be used to prove that if  $\rho = 1$  and the arrival processes at the transmitters are not all deterministic, then the system is unstable under any control policy.

#### LEMMA 3.2

Let  $X_1, \dots, X_M$  be mutually independent random variables. Then  $\sum_{j=1}^M X_j \equiv T$  if and only if for each  $j$ ,  $X_j \equiv n_j$  for some non-negative integers  $n_1, \dots, n_M$  such that  $\sum_{j=1}^M n_j = T$ .

*Proof*

If for some non-negative integers  $n_1, \dots, n_M$ ,  $X_j \equiv n_j$  for each  $j$  and  $\sum_{j=1}^M n_j = T$ , then trivially  $\sum_{j=1}^M X_j \equiv T$ . Conversely, suppose  $\sum_{j=1}^M X_j \equiv T$ . Let

$$\phi^{(j)}(z) = \sum_{i=0}^{\infty} q_i^{(j)} z^i$$

be the generating function of  $X_j$ , where  $\{q_i^{(j)}\}_{i=0}^{\infty}$  is the probability mass function of  $X_j$ . Since the  $X_j$ 's are independent, the generating function  $\phi(z)$  of  $\sum_{j=1}^M X_j$  must satisfy

$$\phi(z) = \prod_{j=1}^M \phi^{(j)}(z) = z^T. \quad (3.4)$$

Since the coefficients of each  $\phi^{(j)}(z)$  are non-negative, then by the above equation each  $\phi^{(j)}$  must be of finite order. It then follows that for each  $j$ , the polynomial  $\phi^{(j)}(z)$  must divide  $z^T$ . This is possible if and only if  $\phi^{(j)}(z) = q_{n_j}^{(j)} z^{n_j}$  for some integer  $n_j$ ,  $0 \leq n_j \leq T$ . This then implies that  $q_{n_j}^{(j)} = 1$  and  $\sum_{j=1}^M n_j = T$ .  $\square$

*Proof of theorem 3.1*

Let us remind the reader of the relation (2.17), which we state here again for convenience:

$$S_n \geq W_n. \quad (3.5)$$

Recall that  $W_n$ , as defined by (2.16), is the waiting time process of a  $D/G/1$  queue. The inequality (3.5) is the key to the following proof.

If  $\rho > 1$ ,  $W_n$  does not possess a limiting distribution, see [14]. In addition,  $\lim_{n \rightarrow \infty} W_n = \infty$  almost surely. This result is not explicitly stated in [14], but it can be seen as follows. Define the sequence of i.i.d. random variables

$$U_k \triangleq Y_k - T, \quad (3.6)$$

and the random walk

$$V_n \triangleq \begin{cases} 0 & \text{if } n = 0, \\ \sum_{k=0}^{n-1} U_k & \text{if } n \geq 1. \end{cases} \quad (3.7)$$

We first prove by induction that

$$W_n \geq V_{n-1}. \quad (3.8)$$

Equation (3.8) is trivial for  $n = 1$ . We now suppose it holds for  $n$  and prove it for  $n + 1$ . Since  $V_n = V_n + Y_{n-1} - T$ , then

$$W_{n+1} = (W_n + Y_{n-1} - T)^+ = (W_n - V_{n-1} + V_n)^+ \geq V_n, \tag{3.9}$$

the last inequality following from the induction hypothesis. If  $\rho > 1$ , then the drift of  $V_n, E(U_0) = (\rho - 1)T$ , is positive so that  $\lim_{n \rightarrow \infty} V_n = \infty$  almost surely (see Feller [6] or Gut [9]). Thus, (3.2) follows immediately from (3.5) and (3.8). Assume next that  $\rho = 1$ , and the arrival processes at the transmitters are not all deterministic, i.e. for some  $i$  and all  $n \in \mathbf{Z}_+$ ,  $P[Y_0^{(i)} = n] < 1$ . This implies by lemma 3.2 that  $P[U_0 = 0] < 1$ , so that  $\limsup_{n \rightarrow \infty} V_n = +\infty$  w.p.1, by [5] or [6]. Expression (3.3) is now immediate upon invoking (3.5) and (3.8).  $\square$

*Remark*

We note that if  $\rho = 1$  and the arrival processes at all transmitters are deterministic, a control policy with finite long-term average cost exists. Suppose that for each  $j, Y_0^{(j)} \equiv n_j$  for some non-negative integers  $n_1, \dots, n_M$  such that  $\sum_{j=1}^M n_j = T$ . Consider the stationary control policy  $(w_i^{(1)}, \dots, w_i^{(M)}) = (n_1, \dots, n_M)$  for all  $i \geq 1$ . By (2.3),  $X_n^{(j)} = x^{(j)}$ , for all  $n \geq 1$ , and this strategy possesses a long-term average cost equal to  $\sum_{j=1}^M x^{(j)}$ .

We next show that in addition to  $\rho < 1$ , finiteness of the second moment of the arrival process at each transmitter is a necessary condition for the existence of control policies that lead to a stable system. Let  $\{X_n^{(j)}(\pi), n \geq 2\}$  denote the state of transmitter  $j$  at epoch  $n$  when policy  $\pi = \prod_{i=1}^\infty \pi_i$  is employed, starting with the initial state  $X_1^{(j)} = x^{(j)}$ .

**THEOREM 3.3**

Assume  $\rho < 1$ . If for some  $i, E[(\xi_1^{(i)})^2] = \infty$ , then the system is unstable under any control policy, i.e. for any control policy  $\pi$  and any initial state  $X_1 = x$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n S_k(\pi) = \lim_{n \rightarrow \infty} E_x[S_n(\pi)] = \infty \quad \text{w.p.1.} \tag{3.10}$$

*Proof*

When  $\rho < 1$ , the waiting time process  $\{W_n\}$ , as defined by (2.16), tends in distribution to a finite random variable, say  $W$  (see Lindley [14] or Kiefer and Wolfowitz [10]). Since  $E[(\xi_1^{(i)})^2] = \infty$  implies that  $E[U_0^2] = \infty$ , then by theorem 3 of Kiefer and Wolfowitz [11],  $E(W) = \infty$ . In addition, by [11], theorem 1, we have

$$P \left[ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n W_k = E(W) \right] = 1. \quad (3.11)$$

This result together with (3.5) and (3.8) imply (3.10).  $\square$

Having established the necessary stability conditions through theorem 3.1 and theorem 3.3, we shall show next that these conditions are sufficient for stable system behaviour.

### 3.2. STABLE BEHAVIOUR: CONSTRUCTION OF A FINITE AVERAGE COST PURE POLICY

In this subsection, we exhibit the existence of a nonrandomized stationary strategy that leads to a stable system for each phase length  $T$ . Before proceeding with the construction of this pure strategy, we shall remind the reader of a key lemma that will enable us to interchange limits in distribution (or probability, or with probability 1) and expectations of a sequence of random variables. The key to this interchange is the *uniform integrability* of the sequence as indicated in [1], theorem 5.4. In this regard, we shall often make use of the following facts.

#### LEMMA 3.4 (UNIFORM INTEGRABILITY)

- (a) Let  $Z, \{Z_n\}$  be a sequence of non-negative integrable random variables, and suppose that  $Z_n$  converges in distribution (or with probability 1, or in probability) to  $Z$ . Then  $\{Z_n\}$  is uniformly integrable if and only if  $E(Z_n) \rightarrow E(Z)$ .
- (b) Any sequence of i.i.d. random variables with finite mean is uniformly integrable.
- (c) Suppose  $\{Z_n\}$  is a sequence such that  $|Z_n| \leq Z$  w.p.1 and  $E(Z) < \infty$ . Then  $\{Z_n\}$  is uniformly integrable.

In the remainder of the paper, we shall consistently suppose that

$$\rho < 1, \quad (3.12)$$

$$E[(\xi_1^{(j)})^2] < \infty, \quad 1 \leq j \leq M. \quad (3.13)$$

To avoid trivialities, also assume

$$0 < P[\xi_1^{(j)} = 0] < 1, \quad 1 \leq j \leq M; \quad (3.14)$$

the inequality on the right precludes channels without message inputs, while the left-hand side inequality is implied by the stability requirement (3.12). We recall that (cf. (2.14))

$$Y_n \triangleq \sum_{j=1}^M Y_n^{(j)}. \tag{3.15}$$

For future reference, we call

$$p_k \triangleq P[Y_0 = k], \tag{3.16}$$

and note that (3.14) implies  $p_0 > 0$ .

We also recall that (cf. [29])

$$S_n = \sum_{j=1}^M X_n^{(j)}; \tag{3.17}$$

$\{S_n\}$  need not be a Markov chain under an arbitrary policy, but it is one under the policy we shall describe. Let  $w_n^{(j)}$  be the number of slots allocated to transmitter  $j$  at epoch  $n$ . Apply the following nonrandomized stationary policy to  $\{X_n\}$ :

$$\text{if } S_n \leq T, \text{ take } w_n^{(j)} = x_n^{(j)}, \tag{3.18}$$

$$\text{if } S_n > T, \text{ take } w_n^{(j)} \leq x_n^{(j)} \text{ and } \sum_{j=1}^M w_n^{(j)} = T. \tag{3.19}$$

This is actually a description of a class of policies, but any such policy will be adequate for our purpose. The motivation of policies of this class is provided by the properties of optimal discounted policies, as given in theorem 3.8 of [2]. Such a policy, applied to  $\{X_n\}$ , induces an  $\{S_n\}$  that will meet our needs. In fact,  $\{S_n\}$  is simply the total cost at epoch  $n$ , so that the time average of  $\{S_n\}$  becomes the time average of the total cost. The state evolution of  $\{S_n\}$  then follows the recursive equations

$$S_{n+1} = \begin{cases} Y_{n-1} & S_n \leq T, \\ S_n + Y_{n-1} - T & S_n > T. \end{cases} \tag{3.20}$$

The remainder of this section is devoted to proving that the stated pure policy indeed leads to a finite average cost, i.e.

$$\limsup_{n \rightarrow \infty} \sum_{k=1}^n E[S_k(x)] < \infty, \tag{3.21}$$

where  $S_n(x)$  is as in (3.17), except that the parameter  $x$  now indicates the initial number of packets stored at the respective transmitters. For (3.21) to be valid, it suffices to demonstrate that

$$\sup_n E[S_n(x)] < \infty, \tag{3.22}$$

from which (3.21) follows immediately.

We shall again obtain the desired result by comparing the average cost associated with the allocation policy (3.18) and (3.19), and the average waiting time cost for a  $D/G/1$  queue. For this purpose, we introduce the notation  $W_n(x)$  to denote the waiting time for the  $D/G/1$  queue, under the supposition that the zeroth customer undergoes waiting time  $x$ . We shall prove that  $S_n(x)$  and  $W_n(x)$  are related by

$$W_n(x) \leq S_n(x) \leq W_n(x) + T \quad (3.23)$$

for all  $n$  when both receive the same inputs  $\{Y_n\}$ . In addition, we verify that  $E[W_n(x)]$  remains bounded in  $n$  by virtue of uniform integrability.

We use induction to prove (3.23). It is clearly true that (3.23) holds for  $n = 0$ , where  $S_n(x)$  and  $W_n(x)$  are equal. Now assume (3.23) holds for  $n$ . In the case  $S_n(x) \leq T$ ,  $S_{n+1}(x) = Y_{n-1}$  by (3.20). On the other hand, we have (see (3.9))

$$W_{n+1}(x) = (W_n(x) + Y_{n-1} - T)^+, \quad (3.24)$$

whence

$$Y_{n-1} - T \leq W_{n+1}(x) \leq (S_n(x) + Y_{n-1} - T)^+ \leq Y_{n-1}.$$

The complementary case  $S_n(x) > t$  yields

$$S_{n+1}(x) - S_n(x) = W_{n+1}(x) - W_n(x)$$

by (3.20) and (3.24), so the result (3.23) again follows.

We now show that  $\{W_n(x)\}$  is uniformly integrable; this implies that

$$\sup_n E[W_n(x)] < \infty \quad (3.25)$$

for each  $x$ . Under the condition  $\rho < 1$ , it was established in [10] and [11] (see also [14]) that there is the limit in distribution

$$W_n(x) \xrightarrow{d} W, \quad (3.26)$$

where  $W$  does not depend on  $x$ . In addition, it is shown in [11] that

$$\lim_{n \rightarrow \infty} E[W_n(0)] = E[W] < \infty, \quad (3.27)$$

with the convergence and the finiteness of  $E[W]$  following from the existence of a finite second moment for  $Y_n$ . Then (3.26) (with  $x = 0$ ), taken together with (3.27), implies that  $\{W_n(0)\}$  is uniformly integrable; this follows from lemma 3.4(a).

To compare  $W_n(0)$  and  $W_n(x)$ , we note that (3.24) can be extended by an easy calculation to

$$W_n(x) = \max \left( 0, x \delta_{0,k} + \sum_{k=r}^{n-1} Y_k - (n-r)T : k = 0, \dots, n-1 \right). \quad (3.28)$$

where  $\delta$  is the Kronecker delta. We therefore obtain

$$W_n(x) \leq W_n(0) + x. \quad (3.29)$$

For each  $x$ , the uniform integrability of  $\{W_n(0)\}$  conveys the same property to  $\{W_n(x)\}$ . Thus,  $\{E[W_n(x)]\}$  satisfies (3.25), and by (3.23) the same must be true for  $\{E[S_n(x)]\}$ . Then (3.21) holds also, and our argument is complete.

*Remark 1*

Before leaving this section, we make the following observation that will be used in section 4. From the state transition matrix of  $\{S_n\}$ , one checks that state zero is reachable from any other state; this follows because  $p_0 > 0$ , as we have already mentioned (see (3.14) and (3.16)). Moreover, the finiteness of the long-term average cost of  $\{S_n\}$ , starting from any initial state, implies that the total expected cost to reach state zero from state  $x$  is finite.

*Remark 2*

Through the use of the inequality (3.22), one also proves that

$$E[S_n(x)] \leq E[W_n(x)] + \rho T. \quad (3.30)$$

For this purpose, consider  $E[S_{n+1}(x) | S_n(x)]$ . In the event  $\{S_n(x) \leq T\}$ , we have from (3.20)

$$E[S_{n+1}(x) | S_n(x)] = E(Y_{n-1}) (= \rho T).$$

On the other hand, in the event  $\{S_n(x) > T\}$ , an application of (3.20) yields

$$E[S_{n+1}(x) | S_n(x)] = S_n(x) + E(Y_{n-1}) - T,$$

so that by the right-hand side inequality of (3.22),

$$E[S_{n+1}(x) | S_n(x)] \leq W_n(x) + \rho T. \quad (3.31)$$

Thus, taking expectations of the conditional expectation leads to (3.30).

*Remark 3*

Using more subtle arguments, one shows that  $E[S_n(x)] \rightarrow E(W) + \rho T$  for all  $x$ .

### 3.3. AN UPPER AND A LOWER BOUND FOR $\bar{V}_T(x)$

We shall derive an upper and lower bound for the minimal achievable long-term average cost  $\bar{V}_T(x)$ . The optimal time-average policy certainly achieves a cost at most as high as the one of the pure policy of section 3.2 as defined by (3.18)–(3.19). Hence, by remark 2 of the preceding section,

$$\bar{V}_T(x) \leq E(W) + \rho T. \quad (3.32)$$

On the other hand, since under any policy  $S_n \geq W_n$  (cf. (2.17)), we also obtain

$$\bar{V}_T(x) \geq E(W). \quad (3.33)$$

For completeness, we need an expression for  $E(W)$  in terms of the arrival processes statistics. In general, there is no such explicit formula. However, from the bounds of [12] on the waiting time process of the  $G/G/1$  queue, we obtain

$$\frac{E[(\sum_{j=1}^M Y_0^{(j)} - T)^+]^2}{2(1-\rho)T} \leq E(W) \leq \frac{\sum_{j=1}^M (\sigma^{(j)})^2}{2(1-\rho)}, \quad (3.34)$$

where  $(\sigma^{(j)})^2 = \text{Var}(\xi^{(j)})$ . Combining (3.32), (3.33) and (3.34) leads to

$$\frac{E[(\sum_{j=1}^M Y_0^{(j)} - T)^+]^2}{2(1-\rho)T} \leq \bar{V}_T(x) \leq \rho T + \frac{\sum_{j=1}^M (\sigma^{(j)})^2}{2(1-\rho)}. \quad (3.35)$$

Observe that the magnitude of the difference between the bounds of (3.32) and (3.33) tends to infinity as  $T \rightarrow \infty$ . These bounds therefore do not give us any insight on the asymptotic behaviour of  $\bar{V}_T(x)$  as  $T \rightarrow \infty$ . We are, however, able to solve this problem in section 5 using a different approach.

In summary, the pure policy of this section enables us to conclude that the minimal achievable long-term average cost  $\bar{V}_T(x)$  is finite for each phase length  $T$ . We now move to the next section to show that  $\bar{V}_T(x)$  is achieved by a Markov policy. Additionally, we derive qualitative properties of this optimal control strategy.

## 4. Existence and properties of time-average optimal policies

In this section, we establish the existence of an optimal nonrandomized stationary strategy for the long-term average cost criterion and derive qualitative properties of this optimal control strategy. Throughout this section,  $T$  is arbitrary but fixed. To be precise, we seek a pure policy, say  $\pi^*$ , such that

$$\bar{V}_T(\pi^*, x) = \inf_{\pi \in \mathcal{P}_T} \bar{V}_T(\pi, x), \quad (4.1)$$

where (cf. (2.10))

$$\bar{V}_T(\pi, x) \triangleq \limsup_{n \rightarrow \infty} \frac{1}{n} E_x^\pi \left[ \sum_{k=1}^n S_k \right], \tag{4.2}$$

$x = (x^{(1)}, \dots, x^{(M)})$  is the initial system state,  $S_n = \sum_{j=1}^M X_n^{(j)}$  and  $X_k^{(j)}$  is the number of packets awaiting transmission at the beginning of phase  $k$  at transmitter  $j$ . Recall that for each  $j$  (cf. (2.3))

$$X_{k+1}^{(j)} = (Y_{k-1}^{(j)} + X_k^{(j)} - w_k^{(j)})^+, \tag{4.3}$$

where  $w_k^{(j)}$  is the number of packets allocated to transmitter  $j$  at phase  $k$ . We remind the reader that the cost (4.2) represents the long-term average number of packets awaiting transmission at the beginning of each phase. Since the long-term average cost of holding packets arriving in each phase is constant (cf. (2.9)), there is no loss of optimality in restricting attention to the cost (4.2).

We shall prove the existence of time-average optimal policies and investigate their qualitative properties based on the following results: (1) the properties of the total expected discounted infinite horizon cost of [2], (2) the properties of the Markov chain induced by the pure policy of the previous section, and (3) the work of Sennott [18] on average cost optimal stationary policies.

We first summarize the properties of the minimal achievable total expected discounted cost and the properties of the optimal discounted policies as given in [2]. For  $0 \leq \beta < 1$ , the total expected  $\beta$ -discounted cost incurred by a policy  $\pi$  is given by

$$V^\beta(\pi, x) \triangleq E_x^\pi \left[ \sum_{n=1}^{\infty} \beta^{n-1} S_{n+1} \right]. \tag{4.4}$$

Let

$$V^\beta(x) \triangleq \inf_{\pi \in \mathcal{P}_T} V^\beta(\pi, x) \tag{4.5}$$

be the minimal achievable total expected  $\beta$ -discounted cost when the initial system state is  $x$ . In [2], the authors study the properties of the discounted value function  $V^\beta(x)$  and of the optimal policies that attain the infimum (4.5). It is shown in [2], lemma 2.2, lemma 3.1, eqs. (3.8) and (3.16), respectively, that  $V^\beta(x)$  has the following properties:

- (P1) For every state  $x$  and discount factor  $\beta$ ,  $V^\beta(x)$  is finite.
- (P2)  $V^\beta(x)$  is non-decreasing in  $x$ ; that is, for each  $i$ ,  $V^\beta(x + e_i) \geq V^\beta(x)$ , where  $e_i$  is an  $M$ -component row vector with 1 in the  $i$ th entry and zeroes in all the other entries.

(P3)  $V^\beta(x)$  satisfies the optimality equation of dynamic programming

$$V^\beta(x) = \min_{w \in \bar{A}} \{L(x, w) + \beta E[V^\beta([Y + x - w]^+)]\} \triangleq \min_{w \in \bar{A}} G^\beta(x, w), \quad (4.6)$$

where  $\bar{A} \triangleq \{(w^{(1)}, \dots, w^{(M)}) \in \mathbf{Z}_+^M : \sum_{j=1}^M w^{(j)} = T\}$ ,  $Y \triangleq (Y^{(1)}, \dots, Y^{(M)})$ ,  $Y^{(j)}$  denotes the random sequence  $\{Y_n^{(j)}\}_{n=0}^\infty$ , and  $L(x, w)$  is the expected cost per phase, i.e.  $L(x, w) = \sum_{j=1}^M E[(Y^{(j)} + x^{(j)} - w^{(j)})^+]$ .

The properties of the  $\beta$ -discounted optimal policies are as follows (see [2], theorem 3.8, lemma 3.6). Let  $x$  be the initial system state and  $\alpha \triangleq \sum_{j=1}^M x^{(j)}$ ; then

(P4) if  $\alpha \geq T$ , any decision rule  $w(x) \in \bar{A}$  such that  $w^{(l)}(x) \leq x^{(l)}$  for each  $l$  is optimal, i.e.  $V^\beta(x) = G^\beta(x, w(x))$ ;

(P5) if  $\alpha \leq T$ , there exists an optimal decision rule  $w(x) \in \bar{A}$  such that  $w^{(l)}(x) \geq x^{(l)}$  for each  $l$ , i.e.  $V^\beta(x) = G^\beta(x, w(x))$ .

Property (P4) assures that for a large number of messages, the slots are allocated so that each one will carry a packet, none of them being "empty" and hence possibly wasted. By similar reasoning, property (P5) assures that the slots are allocated so that all the queued messages that are known to the receiver are transmitted and hence the number of wasted slots is minimized.

We now verify that assumptions 1–3 of [18] which ensure the existence of an average cost optimal stationary policy are satisfied. Assumption 1 is exactly property (P1). From property (P2),  $V^\beta(x) - V^\beta(0) \geq 0$ , so that assumption 2 is met. To verify assumption 3 without irreducibility conditions, we need to show that for every  $x = (x^{(1)}, \dots, x^{(M)})$  there exists non-negative  $M(x) < \infty$  such that

$$V^\beta(x) - V^\beta(0) \leq M(x), \quad (4.7)$$

and that there exists an allocation rule  $w(x)$  such that

$$\sum_y p_{xy}(w(x))M(y) < \infty, \quad (4.8)$$

where  $p_{xy}(w(x))$  is the probability of a transition from  $x$  to  $y$  under the allocation scheme  $w(x)$ . For every  $x$ , let  $w(x) = (w^{(1)}(x), \dots, w^{(M)}(x))$  be the allocation of slots under the stationary policy of section 3.2, as defined by (3.18)–(3.19). Let  $c_{x0}$  be the expected cost of a first passage from  $x$  to zero under this policy. While  $c_{x0}$  depends on  $\beta$ , by remark 1 at the end of section 3.2,  $c_{x0} < \infty$  for every  $x$  even in the worst case, which is  $\beta$  equals 1. Starting from state  $x$ , suppose now that we apply policy  $w(x)$  until we reach state zero and then we continue according to an optimal policy afterwards. We then incur a cost of no more than  $c_{x0} + V^\beta(0)$ . Since any

optimal policy is at least as good as the one employed above, it must be that  $V^\beta(x) \leq c_{x0} + V^\beta(0)$ . Letting  $M(x) = c_{x0}$  for every  $x$ , (4.7) is then satisfied. In addition, under policy  $w(x)$ , we have

$$c_{x0} = \left[ \max \left\{ \sum_{j=1}^M x^{(j)}, T \right\} - T + E(Y_0) \right] + \left[ \sum_{y \neq 0} p_{xy} (w(x)) c_{y0} \right], \tag{4.9}$$

where the first term in [ . . . ] on the RHS of (4.9) is the instantaneous cost incurred when in state  $x$ . Since  $c_{x0} < \infty$ , then the second term in [ . . . ] on the RHS of (4.9) is finite. As  $M(y) = c_{y0}$  for every  $y$ , (4.8) is then satisfied.

We next construct a stationary allocation policy  $f$  that is a limit point of a sequence of optimal allocation policies  $\{f_{\beta_n}\}$  associated with a sequence of discount factors  $\{\beta_n\} \rightarrow 1$ ; indeed, starting with a sequence  $\beta \rightarrow 1$ , we shall be able to choose a subsequence such that  $f = f_{\beta_n}$  for all  $n$ . By the lemma on p. 628 of [18], we already know that a convergent subsequence of  $\{f_{\beta_n}\}$  exists, and from the theorem on the same page, it follows from assumptions 1–3 in [18] that the limit of  $f$  is a time-average optimal allocation.

To prove that all the allocation policies  $f_{\beta_n}$  can be chosen to be the same, it suffices to demonstrate that there exists a finite set of allocation strategies among which the optimal strategy may be chosen for all  $\beta < 1$ . Since the action space  $\mathbf{A}$  can consist of no more than  $M^T$  elements when all  $T$  slots are allocated, the set of all possible allocations on any finite subspace of the state space  $\mathbf{Z}_+^M$  is necessarily finite. Thus, the restriction of optimal policies to the subspace  $\{x : x \in \mathbf{Z}_+^M, \alpha < MT\}$  is finite. On the other hand, for any  $x \in \mathbf{Z}_+^M$  such that  $\alpha \geq MT$ , there exists a first index  $i$  such that its  $i$ th component  $x_i$  satisfies  $x_i \geq T$ . For this  $x$ , (P4) indicates that an optimal allocation is  $w(x) = T e_i$ , where  $e_i$  is the unit vector along component  $i$ ; moreover, the same allocation is optimal for any  $\beta < 1$ .

With the application of the quoted results from [18], together with the finiteness of the set of optimal allocation strategies, we obtain:

**THEOREM 4.1**

Every sequence of discount factors  $\beta$  converging to unity has a subsequence  $\{\beta_n\}$  such that the corresponding optimal stationary allocation policies  $\{f_{\beta_n}\}$  satisfy  $f = f_{\beta_n}$  for all  $n$ . This stationary allocation policy  $f$  is average cost optimal with average cost

$$g = \lim_{\beta \rightarrow 1} (1 - \beta) V^\beta(x), \tag{4.10}$$

where the limit does not depend on  $x$ . Furthermore,  $f$  satisfies properties (P4) and (P5) of the optimal discounted policies, as well as the properties found in sections 4 and 5 of [2] for  $M = 2$ . □

### 5. An asymptotically optimal stationary policy

In this section, we study the asymptotic behaviour of the queueing system as a function of the phase length  $T$ . We remind the reader that we do not incorporate the waiting cost of packets arriving within a phase in the cost function. A cost is incurred only when packets awaiting service at the beginning of a phase are not transmitted. We exhibit the existence of a stationary nonrandomized strategy under which the long-term average number of queued packets at the beginning of each phase converges to zero as  $T$  tends to infinity. This strategy is defined only for large values of  $T$ , and depends only on  $T$  and the average arrival processes rates. It is described in words as follows. There is a  $T_0$  such that for each  $T \geq T_0$  : allocate to each transmitter at each phase of the decision process some fixed number of slots that is higher than the average number of arrivals per phase.

Briefly recall that  $\{\xi_i^{(j)}, i \geq 0\}$ , the number of arrivals per slot at transmitter  $j$ , is an i.i.d. sequence with finite mean  $\lambda^{(j)}$ . To avoid unstable behaviour, we require

$$\sum_{j=1}^M \lambda^{(j)} < 1, \tag{5.1}$$

$$E[(\xi_1^{(j)})^2] < \infty, \quad 1 \leq j \leq M. \tag{5.2}$$

Choose  $\lambda_1^{(j)} > \lambda^{(j)}$  for each  $j$ , such that  $\sum_{j=1}^M \lambda_1^{(j)} < 1$ , and let

$$T_0 \triangleq \left\lceil \frac{M}{1 - \sum_{j=1}^M \lambda_1^{(j)}} \right\rceil, \tag{5.3}$$

where  $\lceil x \rceil$  is the smallest integer not less than  $x$ . Consider the allocation scheme

$$w^{(j)}(T) \triangleq \lceil \lambda_1^{(j)} T \rceil, \quad 1 \leq j \leq M. \tag{5.4}$$

Under the specific strategy considered here, the assignment of slots does *not* depend on the information state  $x$ . The assignment does vary with the phase length  $T$ , which is the parameter we are studying in this section. With this in mind, we shall refer to  $w^{(j)}(T)$ , rather than to  $w^{(j)}(x)$  we used in referring to fixed  $T$  and an arbitrary policy. The allocation scheme (5.4) is well defined for  $T \geq T_0$  and satisfies

$$w^{(j)}(T) \geq \lambda_1^{(j)} T > \lambda^{(j)} T, \quad 1 \leq j \leq M, \quad \forall T \geq T_0. \tag{5.5}$$

Indeed, (5.5) is implied by (5.4), and as  $\lceil \lambda_1^{(j)} T \rceil < \lambda_1^{(j)} T + 1$ , we then obtain using (5.3)

$$\sum_{j=1}^M w^{(j)}(T) < T, \quad \text{for all } T \geq T_0. \tag{5.6}$$

By the strong law of large numbers, the number of arrivals at transmitter  $j$  is about  $\lambda^{(j)}T$  for sufficiently large  $T$ . In addition, since strategy (5.4) allocates invariably more than  $\lambda^{(j)}T$  slots to transmitter  $j$ , it is then intuitive that the mean queue size of transmitter  $j$  tends to be empty at the beginning of each phase whenever  $T$  is sufficiently large. To be precise, under the allocation scheme (5.4) with the phase length fixed at  $T$ , let  $X_n^{(j)}(T, x^{(j)})$  be the state of transmitter  $j$  at epoch  $n$ , where  $x^{(j)} \triangleq X_1^{(j)}(T)$  is the initial state. The number of packets arriving during phase  $n$  at transmitter  $j$  will be denoted by  $Y_n^{(j)}(T)$ . Our goal is to prove the following theorem.

**THEOREM 5.1**

For each initial system state  $\{X_1^{(j)}(T) = x^{(j)}, 1 \leq j \leq M\}$ , we have

$$\forall j, 1 \leq j \leq M, \quad \lim_{T \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n E[X_k^{(j)}(T, x^{(j)})] = 0. \tag{5.7}$$

*Remark*

Since we prove (5.7) separately for each transmitter, we will omit the superscript  $j$  throughout the proof of theorem 5.1 for simplicity. We will reference each transmitter individually only when making a formal statement such as in a theorem, lemma, or corollary.

The main idea behind the proof of theorem 5.1 is that as  $T \rightarrow \infty$ ,  $E[X_n(T, x)]$  converges to zero uniformly in  $n$  for each initial state  $x$ . To establish the latter, we introduce an auxiliary random walk  $W_n(T, x)$  that is equal in distribution to  $X_n(T, x)$  and show that:

- (i)  $\sup_{n \geq 2} \{W_n(T, x)\} \rightarrow 0$  almost surely as  $T \rightarrow \infty$ .
- (ii)  $\{\sup_{n \geq 2} \{W_n(T, x)\}, T \geq T_0\}$  is uniformly integrable.

Statements (i) and (ii) then ensure that  $E[\sup_{n \geq 2} \{W_n(T, x)\}]$  converges to zero as  $T \rightarrow \infty$ , and this also entails that  $\lim_{T \rightarrow \infty} \sup_{n \geq 2} \{E[W_n(T, x)]\} = 0$ . Since  $X_n(T, x) \stackrel{d}{=} W_n(T, x)$ , the last assertion then implies that the expected number of packets awaiting transmission at the beginning of each phase converges uniformly to zero at  $T \rightarrow \infty$ , that is,  $\lim_{T \rightarrow \infty} \sup_{n \geq 2} \{E[X_n(T, x)]\} = 0$ . This result then immediately entails the assertion of theorem 5.1. We shall next proceed to prove (i) and (ii).

Under the allocation scheme (5.4), the system state equations (2.3) are expressible, following an induction step, as

$$X_{n+1}(T, x) = \max \left\{ \sum_{k=1}^i Y_{n-k}(T) - iw(T) + x\delta(n-i) : 0 \leq i \leq n \right\}, \tag{5.8}$$

where by convention the sum on the RHS of (5.8) is zero when  $i = 0$ . Introduce the auxiliary random process

$$W_{n+1}(T, x) = \max \left\{ \sum_{k=1}^i Y_{k-1}(T) - iw(T) + x\delta(n-i) : 0 \leq i \leq n \right\}, \quad (5.9)$$

and note that since  $\{Y_k(T), k \geq 0\}$  is an i.i.d. sequence,  $X_n(T, x) \triangleq W_n(T, x)$ . This technique of substituting a random walk is a well-known approach to  $G/G/1$  queues, as in Lindley [14]. Our first result is that  $W_n(T, x)$  converges uniformly to zero w.p.1 as  $T \rightarrow \infty$ , and this entails convergence to zero in probability of  $X_n(T, x)$  for each  $n \geq 2$  as  $T \rightarrow \infty$ . Remark that for the proof of this result, we only require that assumption (5.1) be met.

#### THEOREM 5.2

For each initial system state  $\{X_1^{(j)}(T) = x^{(j)}, 1 \leq j \leq M\}$ , we have

$$\forall j, 1 \leq j \leq M, \sup_{n \geq 2} \{W_n^{(j)}(T, x^{(j)})\} \rightarrow 0 \text{ w.p.1 as } T \rightarrow \infty. \quad (5.10)$$

#### Proof

Write  $Y_n(T) = \sum_{k=m}^{(m+1)T} \xi_k$ , where  $\{\xi_k\}$ , the number of arrivals per slot, is an i.i.d. sequence with  $E(\xi_1) = \lambda$ , and define the i.i.d. zero mean sequence

$$Z_k \triangleq \xi_k - \lambda. \quad (5.11)$$

Expression (5.9) is then equivalent to

$$W_{n+1}(T, x) = \max \left\{ \sum_{k=1}^{iT} Z_k + i[\lambda T - w(T)] + x\delta(n-i) : 0 \leq i \leq n \right\}. \quad (5.12)$$

Let

$$\mu \triangleq \frac{\lambda_1 - \lambda}{2} > 0, \quad (5.13)$$

$$A_m^{(T)}(v) \triangleq \left\{ \omega : \frac{1}{mT} \sum_{k=1}^{mT} Z_k(\omega) > v \right\}, \quad (5.14)$$

$$T_x \triangleq \min \left\{ i \in \mathbf{N} : i > \max \left\{ T_0, \frac{x}{\mu} \right\} \right\}. \quad (5.15)$$

By (5.12) and straightforward algebra, we obtain

$$\forall T \geq T_x, \{ \omega : W_{n+1}(T, x, \omega) > 0 \} \subset \left[ \bigcup_{i=1}^{n-1} A_i^{(T)}(2\mu) \right] \cup A_n^{(T)}(\mu) \subset \bigcup_{i=1}^n A_i^{(T)}(\mu). \tag{5.16}$$

From (5.16), we obtain

$$P \left[ \sup_{n \geq 2} \{ W_n(T, x) \} > 0 \right] \leq P \left[ \bigcup_{i=1}^{\infty} A_i^T(\mu) \right] \tag{5.17}$$

$$\leq P \left[ \sup_{m \geq T} \left\{ m^{-1} \sum_{k=1}^m Z_k > \mu \right\} \right], \tag{5.18}$$

and the right-hand side goes to zero as  $T \rightarrow \infty$  by the strong law of large numbers. However, since  $W_n(T, x) \geq 0$ , (5.10) is then established.  $\square$

An immediate consequence of (5.10) is that the number of packets awaiting transmission at the beginning of each phase converges to zero in probability as the phase length increases indefinitely.

**COROLLARY 5.3**

For each initial system state  $\{X_1^{(j)}(T) = x^{(j)}, 1 \leq j \leq M\}$ , we have

$$\forall j, 1 \leq j \leq M, X_n^{(j)}(T, x^{(j)}) \xrightarrow{P} 0, \quad \forall n \geq 2, \text{ as } T \rightarrow \infty. \tag{5.19}$$

*Proof*

By an earlier remark,  $X_n(T, x) \triangleq W_n(T, x)$ , so that for any Borel set  $B$ ,  $P[X_n(T, x) \in B] = P[W_n(T, x) \in B]$ . In addition, since for each  $n$ , almost sure convergence of  $\{W_n(T, x)\}$  entails convergence in probability of  $\{W_n(T, x)\}$  as  $T \rightarrow \infty$ , (5.19) immediately follows from (5.10).  $\square$

We next show that the random variables  $\{\sup_{n \geq 2} \{W_n(T, x)\}, T \geq T_0\}$  are dominated by an integrable random variable, which then implies (cf. lemma 3.4(c)) that they are uniformly integrable. Recalling (5.11), we first write  $W_{n+1}(T, x)$ , as given by (5.12), under the equivalent form

$$W_{n+1}(T, x) = \max \left\{ \sum_{k=1}^{iT} \xi_k - iw(T) + x\delta(n-i) : 0 \leq i \leq n \right\}. \tag{5.20}$$

We next introduce the auxiliary random process

$$\bar{W}_{n+1}(T, x) \triangleq \max \left\{ \sum_{k=1}^{iT} \xi_k - i\lambda_1 T + x\delta(n-i) : 0 \leq i \leq n \right\}, \tag{5.21}$$

and note that as  $w(T) \geq \lambda_1 T$  (cf. (5.5)), we have the simple relationship

$$W_{n+1}(T, x) \leq \bar{W}_{n+1}(T, x). \quad (5.22)$$

Our aim is to prove the following.

LEMMA 5.4

For each initial system state  $\{X_1^{(j)}(T) = x^{(j)}, 1 \leq j \leq M\}$ , we have

$$\begin{aligned} \forall j, 1 \leq j \leq M, \forall T \geq T_0, \sup_{n \geq 2} \{W_n^{(j)}(T, x^{(j)})\} \\ \leq \sup_{n \geq 2} \{\bar{W}_n^{(j)}(1, 0)\} + x^{(j)}. \end{aligned} \quad (5.23)$$

Furthermore,  $E[\sup_{n \geq 2} \{\bar{W}_n^{(j)}(1, 0)\}] < \infty$ .

*Proof*

Let us first define

$$W(T, x) \triangleq \sup_{n \geq 2} \{W_n(T, x)\}, \quad (5.24)$$

$$\bar{W}(T, x) \triangleq \sup_{n \geq 2} \{\bar{W}_n(T, x)\}. \quad (5.25)$$

By (5.21), we obtain

$$\begin{aligned} \bar{W}_{nT+1}(1, x) &= \max \left\{ \sum_{k=1}^i \xi_k - i\lambda_1 + x\delta(n-i) : 0 \leq i \leq nT \right\} \\ &\geq \max \left\{ \sum_{k=1}^i \xi_k - i\lambda_1 + x\delta(n-i) : i \in \{0, T, 2T, \dots, nT\} \right\} \\ &= \max \left\{ \sum_{k=1}^{iT} \xi_k - i\lambda_1 T + x\delta(n-i) : 0 \leq i \leq n \right\} \\ &= \bar{W}_{n+1}(T, x). \end{aligned} \quad (5.26)$$

Combining (5.22) and (5.26), we obtain  $W_{n+1}(T, x) \leq \bar{W}_{nT+1}(1, x)$ . Recalling (5.24)–(5.25), the latter inequality then entails

$$W(T, x) \leq \bar{W}(1, x). \tag{5.27}$$

Since  $x \geq 0$ , we obtain from (5.21)

$$\begin{aligned} \forall T, \forall x, \forall n, \quad \bar{W}_n(T, x) &\leq \max \left\{ \sum_{k=1}^{iT} \xi_k - i\lambda_1 T + x : 0 \leq i \leq n \right\} \\ &= \bar{W}_n(T, 0) + x, \end{aligned} \tag{5.28}$$

so that invoking (5.25), we obtain  $\bar{W}(1, x) \leq \bar{W}(1, 0) + x$ . This result together with (5.27) yields (5.23). Furthermore, by [11], theorem 5, the negative drift of the random walk underlying  $\{\bar{W}_n\}$  and (5.2) ensure that  $\bar{W}(1, 0)$  is integrable.  $\square$

From lemma 3.4(c) and lemma 5.4, we then deduce that  $\{W(T, x), T \geq T_0\}$  is uniformly integrable. This result and the almost sure convergence to zero of  $W(T, x)$  as  $T \rightarrow \infty$ , imply the following (cf. lemma 3.4(a)).

**THEOREM 5.5**

For each initial system state  $\{X\}^{(j)}(T) = x^{(j)}, 1 \leq j \leq m\}$ , we have

$$\forall j, 1 \leq j \leq M, \quad \lim_{T \rightarrow \infty} E \left[ \sup_{n \geq 2} \{W_n^{(j)}(T, x^{(j)})\} \right] = 0. \tag{5.29}$$

An immediate consequence of theorem 5.5 is that the expected number of untransmitted packets at the beginning of each phase converges uniformly to zero as the phase length tends to infinity.

**COROLLARY 5.6**

For each initial system state  $\{X\}^{(j)}(T) = x^{(j)}, 1 \leq j \leq m\}$ , we have

$$\forall j, 1 \leq j \leq M, \quad \lim_{T \rightarrow \infty} \sup_{n \geq 2} \{E[X_n^{(j)}(T, x^{(j)})]\} = 0. \tag{5.30}$$

*Proof*

The proof is immediate by (5.29) and the inequality

$$\sup_{n \geq 2} \{E[X_n(T, x)]\} = \sup_{n \geq 2} \{E[W_n(T, x)]\} \leq E \left[ \sup_{n \geq 2} \{W_n(T, x)\} \right]. \tag{5.31}$$

$\square$

The proof of theorem 5.1 (cf. (5.7)) is now a direct consequence of (5.30).

**Remark**

Suppose that we generalize assumption (5.2) to

$$E[(\xi_1^{(j)})^{m+1}] < \infty, \text{ for some integer } m \geq 1, 1 \leq j \leq M. \quad (5.32)$$

Applying theorem 5 of [9] to  $\bar{W}(1, 0)$ , we obtain that  $E[\bar{W}^m(1, 0)] < \infty$ . This observation and the majorization of  $W^k(T, x)$  by  $(\bar{W}(1, 0) + x)^k$  for positive values of  $k$  (cf. (5.23)), then imply that the family  $\{W^k(T, x), T \geq T_0\}$  is uniformly integrable for each  $k$ ,  $0 < k \leq m$ . This result and the almost sure convergence to zero of  $W(T, x)$  as  $T \rightarrow \infty$  then entail that  $\lim_{T \rightarrow \infty} E[W^k(T, x)] = 0$  for each  $k$ ,  $0 < k \leq m$ . By the inequality for positive  $k$

$$\sup_{n \geq 2} \{E[X_n^k(T, x)]\} = \sup_{n \geq 2} \{E[W_n^k(T, x)]\} \leq E \left[ \sup_{n \geq 2} \{W_n(T, x)\} \right]^k, \quad (5.33)$$

we deduce that for each  $k$ ,  $0 < k \leq m$ , the  $k$ th moment of the number of unsent packets at the beginning of each phase converges uniformly to zero as the phase length tends to infinity, that is  $\lim_{T \rightarrow \infty} \sup_{n \geq 2} \{E[X_n^k(T, x)]\} = 0$ . Suppose now that we generalize the cost function, say  $f(x_n)$ , to a polynomial of the number of untransmitted packets at the beginning of each phase  $x_n$  satisfying: (1)  $f(0) = 0$ , and (2) the degree of  $f$  is at most  $m$ . Then using (5.33), we have the following generalization of theorem 5.1.

**THEOREM 5.7**

Let the cost function at transmitter  $j$ ,  $f^{(j)}$ , be a polynomial of the number of untransmitted packets at the beginning of each phase. If the degree of  $f^{(j)}$  is at most  $m$ , and  $f^{(j)}(0) = 0$ , then for each initial system state  $\{X_1^{(j)}(T) = x^{(j)}, 1 \leq j \leq M\}$ , we have under assumption (5.32)

$$\forall j, 1 \leq j \leq M, \quad \lim_{T \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n E\{f^{(j)}[X_k^{(j)}(T, x^{(j)})]\} = 0. \quad (5.34)$$

We conclude this section by a discussion of the following open problem. In the derivation of the zero asymptotic cost as  $T \rightarrow \infty$ , we neglected the holding costs of arrivals within the phase. When these costs are incorporated in the cost function per phase, one would like to study the behaviour of the minimal long-term average cost as a function of the phase length  $T$ . This will then lead to an optimal design of the phase length, a problem of practical importance.

Theorem 5.7 states that, by an appropriate choice of  $\pi$ , the first term  $\bar{V}_T(\pi, x)$  of the total cost  $W_T(\pi, x)$  in (2.8) can be made to go to zero at  $T \rightarrow \infty$ . However, the second term grows with  $T$ , independently of any allocation policy choice  $\pi$ . It follows that there exists at least one optimal  $T$ , that is, one phase length that

minimizes the total cost. Our results on the behaviour of  $\bar{V}_T(\pi, x)$  do not reveal the form of this function, so we cannot guarantee that the optimality on  $T$  is unique, nor can we suggest how such an optimal  $T$  is to be calculated. For example, if we could prove (as we conjecture) that the optimal cost  $\bar{V}_T(x)$  is monotone in  $T$ , the uniqueness of the optimal  $T$  is obvious, and some trial-and-error scheme might be suitable for its determination.

### 6. Flow control with priorities

Priorities among the messages at the respective  $M$  transmitters are often discussed in the context of different types of transmissions, such as voice and data. We shall discuss optimal flow control allocation policies with priorities in a future paper. At this juncture, we shall limit ourselves to some simple generalizations involving priorities.

In our model, priorities appear in terms of weighting factors  $c^{(j)}$  for the various transmitters. Thus, as a direct extension of (2.8), we write

$$W_T^c(\pi, x) = \limsup_{n \rightarrow \infty} n^{-1} \sum_{k=1}^n \sum_{j=1}^M c^{(j)} E_x^\pi [X_k^{(j)}] + \frac{T+1}{2} \sum_{j=1}^M c^{(j)} \lambda^{(j)}. \tag{6.1}$$

If we define

$$\underline{c} \triangleq \min c^{(j)} \quad \text{and} \quad \bar{c} \triangleq \max c^{(j)}$$

we obtain

$$\underline{c} W_T(\pi, x) \leq W_T^c(\pi, x) \leq \bar{c} W_T(\pi, x). \tag{6.2}$$

It follows that none of the results on system stable behaviour are changed. The conditions of theorem 3.3 on instability, and the stability of the allocation policy of section 3.2 remain unaffected. In section 3.3, the upper and lower bounds are modified in an obvious manner that we shall not detail here. Moreover, the asymptotic values of the two types of costs as  $T \rightarrow \infty$  are as indicated in section 5 without any change in the applicable arguments. Indeed, the only modification comes in section 4, since properties (P4) and (P5) may no longer hold. We then obtain a weaker version of theorem 4.1. There is still a limiting allocation policy  $f$  as a pointwise limit of a subsequence of the  $f_{\beta_n}$ . This policy continues to satisfy (4.10), but one must be content with lesser properties than (P4) and (P5).

### 7. Conclusions

In this paper, we have shown that for each phase length  $T$ , optimal pure policies exist for the average cost criterion under the conditions: (1) the traffic

intensity is less than unity, and (2) the intensity of the arrival stream has finite second moment. We also proved that these time-average optimal strategies have the same properties as those derived in [2] for optimal discounted strategies. This result is of practical importance since: (1) the time-average cost criterion is a more natural setting for flow control problems, and (2) these qualitative properties are very useful in the search for optimal time-average policies. Finally, we proved that in the absence of costs accrued by messages within the phase, there exists a policy such that the time-average cost tends toward zero as the phase length  $T \rightarrow \infty$ . We believe that this result is a first step in understanding the system behaviour as a function of the phase length  $T$  when the holding costs of messages arriving within the phase are incorporated in the cost function. Our ultimate goal in this direction is to determine the optimal value of the phase length  $T$ . This is a difficult problem that we have reserved for possible future publication.

### Acknowledgements

The work of R. Bournas was supported by IBM Endicott, NY 13760 and the work of D. Teneketzis was supported by ONR Grant No. N00014-87-K-0540.

### References

- [1] P. Billingsley, *Convergence of Probability Measures* (Wiley, New York, 1968).
- [2] 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. Auto. Control* AC-36(1991)1450–1463.
- [3] D.H. Cansever and R.A. Milito, Optimal hop-by-hop flow control policies with multiple transmitters, *Proc. 26th Conf. on Decision and Control*, Los Angeles, CA (December 1987), pp. 1858–1862.
- [4] D.H. Cansever and R.A. Milito, Optimal hop-by-hop flow control policies with multiple transmitters, *Proc. 27th Conf. on Decision and Control*, Austin, TX (December 1988), pp. 1291–1296.
- [5] D.H. Cansever and R.A. Milito, Optimal multistage hop-by-hop flow control policies, *Proc. 28th Conf. on Decision and Control*, Tampa, FL (December 1989), pp. 2530–2535.
- [6] W. Feller, *An Introduction to Probability Theory and Its Application*, Vol. 2, 2nd ed. (Wiley, New York, 1971).
- [7] M. Gerla and L. Kleinrock, Flow control: A comparative survey, *IEEE Trans. Commun.* COM-28(1980)533–574.
- [8] M. Gerla and L. Kleinrock, Flow control protocols, in: *Computer Network Architectures and Protocols*, ed. P.E. Green, Jr. (Plenum, New York, 1982), pp. 361–412.
- [9] A. Gut, *Stopped Random Walks, Limit Theorems and Application* (Springer, New York, 1988).
- [10] J. Kiefer and J. Wolfowitz, On the theory of queues with many servers, *Trans. Amer. Math. Soc.* 78(1955)1–18.
- [11] J. Kiefer and J. Wolfowitz, On the characteristics of the general queueing process with applications to random walk, *Ann. Math. Stat.* 27(1956)147–161.
- [12] J.F.C. Kingman, Inequalities in the theory of queues, *J. Roy. Statist. Soc. Ser. B* 32(1970)102–110.
- [13] P. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control* (Prentice-Hall, Englewood Cliffs, NJ, 1986).
- [14] D.V. Lindley, The theory of queues with a single server, *Proc. Cambridge Phil. Soc.* 48(1952) 277–289.

- [15] S.A. Lippman, On dynamic programming with unbounded rewards, *Manag. Sci.* 21(1975) 1225–1233.
- [16] Z. Rosberg and I.S. Gopal, Optimal hop-by-hop flow control in computer networks, *IEEE Trans. Auto. Control* AC-31(1986)813–822.
- [17] M. Schwartz, *Computer Communication Networks, Design and Analysis* (Prentice–Hall, Englewood Cliffs, NJ, 1977).
- [18] L.I. Sennott, Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs, *Oper. Res.* 37(1989)626–633.