

Optimal Bandwidth Allocation in a Delay Channel

Navid Ehsan, *Member, IEEE*, and Mingyan Liu, *Member, IEEE*

Abstract—In this paper, we consider the problem of allocating bandwidth to two queues with arbitrary arrival processes, so as to minimize the total expected packet holding cost over a finite or infinite horizon. Bandwidth is in the form of time slots in a time-division multiple-access schedule. Allocation decisions are made based on one-step delayed queue backlog information. In addition, the allocation is done in batches, in that a queue can be assigned any number of slots not exceeding the total number in a batch. We show for a two queue system that if the holding cost as a function of the packet backlog in the system is nondecreasing, supermodular, and superconvex, then: 1) the value function at each slot will also satisfy these properties; 2) the optimal policy for assigning a single slot is of the threshold type; and 3) optimally allocating M slots at a time can be achieved by repeatedly using a policy that assigns each slot optimally given the previous allocations. Thus, the problem of finding the optimal allocation strategy for a batch of slots reduces to that of optimally allocating a single slot, which is conceptually much easier to obtain. These results are applied to the case of linear and equal holding costs, and we also present a special case where the above results extend to more than two queues.

Index Terms—Convexity, delayed state observation, optimal resource/bandwidth allocation, satellite communication, stochastic optimal control, superconvexity, supermodularity.

I. INTRODUCTION

IN THIS PAPER, we study the problem of optimally allocating bandwidth (in the form of time slots in a slotted system) to parallel queues when the channel introduces significant feedback delay. Special features of this problem include that 1) servers/slots are assigned *in batches*, i.e., multiple servers/slots may be allocated to the same queue at a time so that multiple packets may be served from the queue and 2) the allocation decision is based on partially obsolete state observations (queue backlogs) due to the significant delay in the system.

This optimal bandwidth allocation problem is primarily motivated by wireless communication systems that either have large propagation delay (e.g., in satellite data communication), or where resource allocation is done relatively infrequently compared to packet transmission time, due to cost or design constraint such as energy (e.g., under the IEEE 802.15.4 standard for low-power indoor wireless networks).

Manuscript received September 1, 2005; revised April 22, 2006. This work was supported in part by the National Science Foundation under Grant ANI-0238035 and in part by Hughes Network Systems, Germantown, MD. This paper was presented in part at the IEEE Conference on Decision and Control (CDC), Paradise Island, Bahamas, 2004.

N. Ehsan is with the University of California, San Diego, CA 92093 USA (e-mail: nehsan@ucsd.edu).

M. Liu is with the University of Michigan, Ann Arbor, MI 48109-2122 USA (e-mail: mingyan@eecs.umich.edu).

Digital Object Identifier 10.1109/JSAC.2006.879406

In the case of a satellite network, users/terminals transmitting to the satellite are assumed to follow a dynamic time-division multiple-access (TDMA) schedule, each assigned a certain number of slots within a *frame* that consists of a fixed number of *slots*. Users inform the satellite their current backlog, carried in packet headers, and the assignment is made based on this information and broadcast to the users over a noninterfering channel. An allocation specifies which slot in the upcoming frame is reserved for/to be used by which user. Due to the long propagation delay of the satellite channel (approximately 250 ms from ground/user to satellite and back), the allocation decision for a particular frame is made based on the backlog information collected during the *previous* frame, which is partially obsolete by the time the allocation is used. This results in possible overallocation or underallocation. Therefore, in this case, the allocation policy needs to take into account unknown random arrivals that occur in between observations/state information updates. In the case of low-power devices sharing a common channel, similar resource allocation problems may arise when time is divided into active and inactive periods (with transceivers turned off to conserve energy), and the bandwidth allocation decision for a given active period is made based on backlog information provided during the previous active period. While the dynamics of the above systems are similar, in this paper we will focus on the satellite scenario to formulate our problem.

Our primary interest is in deriving allocation strategies that allow the system to perform in the most efficient way. Specifically, we assume that backlogged packets incur a cost, and consider an optimal bandwidth allocation problem with the objective of minimizing the expected total packet holding cost over a finite or infinite time horizon.

While, in general, reducing the holding cost has the effect of reducing packet delay, different forms of the cost function lead to different performance criteria. Different cost functions also lead to different optimal strategies, to be further explored in this paper.

Resource allocation problems of similar types have been extensively studied in the literature under various scenarios. Here we review studies most relevant to the one investigated in this paper. In [1] and [2], the problem of parallel queues with different holding costs and a single server was considered, and the simple $c\mu$ rule was shown to be optimal. [3]–[5] considered the server allocation problem to multiple queues with varying connectivity but of the same service class (identical cost functions). [6] further considered a similar problem but with differentiated service classes where different queues have different holding costs. [7], [8] studied the stability of power allocation policies. In all of the above work, the state of the system, i.e., connectivity and the number of packets in each queue, is precisely known before the allocation is made. This is a major difference between the above cited work and the problem considered here.

[9] studied the problem of routing to two parallel queues with delayed state observation and showed that when the information is one-step delayed the policy to join the queue with smaller expected length minimizes the total discounted sum of the number of packets in both queues. [10] studied the problem of optimally routing to two queues with imperfect and noisy information. [11] studied the problem of optimal subcarrier allocation in an orthogonal frequency-division multiplexing (OFDM) channel, and used a model very similar to ours. However, in the case of [11], the bandwidth is shared in the frequency-domain and the state of the system is assumed to be perfectly observed, i.e., no delayed information as we consider in this paper.

In [12] and [13], we have studied problems similar to the one presented in this paper, but with simpler, linear cost assumptions. In [12], we derived the optimal policy when users have the same unit holding cost and identical arrival processes, while in [13], we investigated optimal policies for differentiated linear holding costs in the case of a single-slot allocation and Bernoulli arrivals. By contrast, in this paper we consider general cost functions and arrival processes, and the problem of assigning a batch of slots at a time. We will adopt and explore similar ideas to that used in [14] and [15], where certain properties of the value function were shown to propagate in time for specific queueing models. In particular, we identify three conditions that characterize a class of cost functions, namely, *monotonicity* (nondecreasing), *supermodularity*, and *superconvexity* (to be defined precisely later), and show the following main results by limiting our attention to two queues/users.

- 1) When allocating one slot at a time (single server scenario), if the cost function is nondecreasing, supermodular, and superconvex, then the value function (or cost to go) at each time slot will also satisfy these properties. Furthermore, the optimal policy for assigning a single slot is of the *threshold* type.
- 2) If the cost function is nondecreasing, supermodular, and superconvex, then the problem of optimally allocating M slots at a time reduces to sequentially allocating a single-slot optimally. In other words, a policy that assigns each slot optimally given the previous allocations in the batch, is optimal in assigning the entire batch of M slots.

The second is an important result, as it indicates that if the cost function satisfies those properties, then we may limit our attention to finding the optimal allocation strategy for a single slot instead of for the whole batch. The former is conceptually much easier to obtain. We will also apply the above results to the special case of linear and equal holding cost and show an example where the above results also extend to more than two queues. General extension to more than two queues remains a challenging and open problem.

The rest of this paper is organized as follows. In Section II, we describe the general network model and formulate the corresponding optimization problem. In Sections III and IV, we investigate the optimal policy of allocating a single slot and multiple slots to two queues, respectively. In Section V, we extend our results to the infinite horizon case and examine the discounted cost and average cost criteria. In Section VI, we use these results to find the optimal policy for the special case of linear and equal holding costs. Section VII concludes this paper.

II. PROBLEM FORMULATION

A. Network Model and Notation

Consider N queues that transmit packets to a single receiver, competing for shares of a common channel that consists of time slots. Packets arrive at queues according to arbitrary random processes, and are assumed to be of equal length. One packet transmission time occupies one time slot (i.e., transmissions are assumed to be successful). M consecutive slots constitute a *frame*. The allocation of the channel is done once per frame. M may or may not be greater than N , and a queue may be assigned any number of slots not exceeding M . Alternatively, the above model can be viewed as one, where N queues are being served by M servers, and multiple packets from the same queue are served if it is allocated multiple servers.

We consider time evolution in discrete time steps indexed by $t = 0, 1, \dots, T$, with each increment representing a frame length. Frame t refers to the frame defined by the interval $[t, t + 1)$. In subsequent discussions, we will use terms *frames*, *steps*, and *stages* interchangeably. We will also use the terms *bandwidth* and *slots* interchangeably.

The allocation decision is made (either by the satellite or a ground control center) based on the backlog information (i.e., buffer occupancy denoted by \mathbf{b}_t) provided by queues at the beginning of frame t . We will ignore the transmission time of such information.¹ The decision (denoted by \mathbf{w}_{t+1}) is then broadcast to all queues over a noninterfering channel, and received by the queues at the end of frame t , due to propagation delay, in time to be used for the next frame $t + 1$. The same procedure then repeats till $t = T$, the time horizon, resulting in a one-step delay in state observation, as shown in Fig. 1. Note that in this scenario during the first frame queues do not have allocated slots and only start transmitting in the second frame (starting $t = 1$).

Below, we summarize key notations used in subsequent sections. In general, bold face letters and normal letters represent vectors and scalars, respectively.

Let $b_{i,t}$ be the backlog of queue/user i at the beginning of frame t (more precisely, this is the backlog of queue i at time instant t^-). Denote by \mathbf{b}_t the vector $(b_{1,t}, b_{2,t}, \dots, b_{N,t})$.

$\mathbf{w}_t = (w_{1,t}, \dots, w_{N,t})$: Allocation (in number of slots) for each queue to be used for packet transmission during the t th frame (in the interval $[t, t + 1)$).

$\mathbf{a}_t = (a_{1,t}, \dots, a_{N,t})$: Random arrivals during $[t, t + 1)$ to each queue.

$p_t(\mathbf{a}_t)$: The joint probability mass function for having \mathbf{a}_t arrivals between $[t, t + 1)$.

For any scalar x define $x^+ = [x]^+ = x$ if $x \geq 0$ and 0 otherwise. For a vector \mathbf{x} , we define $\mathbf{x}^+ = [\mathbf{x}]^+$ the same way componentwise. For two vectors \mathbf{x} and \mathbf{y} , by $\mathbf{x} \leq \mathbf{y}$, we mean that the inequality holds component by component.

$\mathbf{x}_t = [\mathbf{b}_{t-1} - \mathbf{w}_{t-1}]^+$: The part of the queue backlog at time t that is precisely known to the controller at time t^- . Given the backlog at $t - 1$, \mathbf{b}_{t-1} , and the past allocation for the period $[t - 1, t)$, \mathbf{w}_{t-1} , this quantity is the amount of packets that are

¹This does not affect our analysis since one can always increase the frame length with dedicated fixed number of slots at the beginning for the transmission of such information.

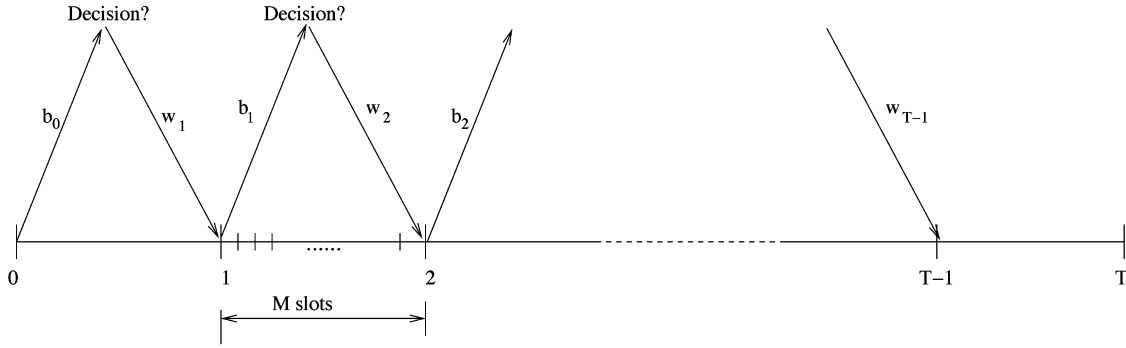


Fig. 1. Bandwidth allocation dynamics.

for sure in the queue, not including the random arrivals that occurred during $[t-1, t)$. It is either zero (when the previous allocation is sufficient or more) or positive (when the previous allocation is not sufficient). We will also refer to this quantity as the *deterministic* part of the queue.

\mathbf{e}_i : The i th N -dimensional unit vector.

For a function f defined on \mathbb{Z}_+^2 , let \hat{f} , defined on \mathbb{Z}^2 , be $\hat{f}(\mathbf{x}) = f(\mathbf{x}^+)$. This definition will prove to be helpful since we do not need to be concerned with boundary conditions for \mathbf{x} when using \hat{f} .

Our objective is to find an allocation policy π that minimizes the following cost function:

$$J = E^\pi[C|\mathbf{b}_0, \mathbf{w}_0], \quad C = \sum_{t=1}^T c_t(\mathbf{b}_t) \quad (1)$$

where $\mathbf{w}_0 = \mathbf{0}$ and $c_t(\mathbf{b})$ is the (potentially time-dependent) cost function.

We summarize important assumptions adopted by this paper.

- 1) We will consider a system with only two users, i.e., $N = 2$. The extension of our results to more than two users remains an open problem and is out of the scope of this paper. Such results exist with stronger assumptions on the cost function, and we will present an example in Section VI.
- 2) We assume that each user has an infinite buffer size. Without this assumption we need to introduce penalty for packet dropping/blocking, which makes the problem drastically different.
- 3) We assume that the arrivals are independent of the queue size and the allocation policy.
- 4) We assume that if the number of allocated slots for a user is greater than its buffer occupancy at the beginning of a frame, the newly arrived packets during that frame cannot be transmitted using the extra slots.²
- 5) Finally, we assume that the controller recalls at least the latest allocation it has made. Note, however, that due to the Markovian nature of the problem, memory more than the

²This is because the exact arrival times of the packets in a frame are random, thus whether an extra slot could be used for a new arrival or not depends on the position of the allocated slot (e.g., the first slot or the last slot of the M slots in the frame) and the arrival time of the packet.

latest allocation will not make any difference to the optimal allocation strategy.³

B. Problem Formulation and Preliminaries

We will use $\mathbf{x}_t = [\mathbf{b}_{t-1} - \mathbf{w}_{t-1}]^+$, the deterministic portion of the queue, as the system state at time t , as it completely determines the system transition. Subsequently, when we say a queue is empty (nonempty), we are referring to this quantity. Note that the actual queue size at time t is $\mathbf{x}_t + \mathbf{a}_{t-1}$.

Define $\bar{c}_t(\mathbf{x}) := E_{\mathbf{a}_{t-1}}[c_t(\mathbf{x} + \mathbf{a}_{t-1})]$, where $E_{\mathbf{a}_t}[f(\mathbf{a}_t)] = \sum_{\mathbf{a}_t} p_t(\mathbf{a}_t)f(\mathbf{a}_t)$ for some function f .

Define the cost-to-go at time t as follows:

$$V_t(\mathbf{x}) = \min_{\pi} E^\pi \left\{ \sum_{s=t}^T \bar{c}_s(\mathbf{x}_s) | \mathbf{x}_t = \mathbf{x} \right\}.$$

$V_t(\mathbf{x})$ can be obtained via the following dynamic program [16], [17]:

$$\begin{aligned} V_T(\mathbf{x}) &= \bar{c}_T(\mathbf{x}), \\ V_t(\mathbf{x}) &= \bar{c}_t(\mathbf{x}) \\ &+ \min_{\sum_{i=1}^N w_{i,t}=M} \left\{ E_{\mathbf{a}_{t-1}} [V_{t+1}([\mathbf{x} + \mathbf{a}_{t-1} - \mathbf{w}_t]^+)] \right\}. \end{aligned} \quad (2)$$

For the rest of this paper, we further assume that the joint pmf of the arrival processes does not change with time, and that the cost function is also time invariant. Thus, we have $p_t(\mathbf{a}_t) = p(\mathbf{a}_t)$, $c_t(\mathbf{b}_t) = c(\mathbf{b}_t)$, $\forall t$. These two assumptions are for the simplicity of notation and as will be discussed at the end of Section III can be easily relaxed. Note that by these two assumptions, we have $\bar{c}_t(\mathbf{x}) = \bar{c}(\mathbf{x})$ for all t .

We will also use the notation $\hat{S}_t(\mathbf{x}) : \mathbb{Z}^2 \rightarrow \mathbb{R}$ defined as

$$\hat{S}_t(\mathbf{x}) := \sum_{\mathbf{a}} p(\mathbf{a}) V_t([\mathbf{x} + \mathbf{a}]^+). \quad (3)$$

Definition 1: For some function $f : \mathbb{Z}^2 \rightarrow \mathbb{R}$, define the operator T_m to be $T_m f(\mathbf{x}) := \min_{\mathbf{w}: w_1+w_2=m} \{f(\mathbf{x} - \mathbf{w})\}$.

³The expected cost occurred after time t conditioned on the latest allocation, \mathbf{w}_t and buffer occupancy \mathbf{b}_t is independent of arrivals that occurred before frame t , and independent of the allocations made before t .

If $f(\mathbf{x})$ represents the value function at state \mathbf{x} , then T_1 represents the minimum between assigning one slot to user 1 and user 2, whereas T_M is the minimum among all possible ways of dividing M slots between two users. The following lemma immediately follows as a result of the definitions above.

Lemma 1: $\forall t, 0 < t < T, V_t(\mathbf{x})$ is equal to $\bar{c}(\mathbf{x}) + T_M \hat{S}_{t+1}(\mathbf{x})$ restricted to $\mathbf{x} \in \mathbb{Z}_+^2$.

The next two sections study the two cases $M = 1$ and $M > 1$, respectively. We are particularly interested in conditions under which T_M may be obtained by repeatedly using T_1 .

III. OPTIMAL POLICY FOR A SINGLE-SLOT ALLOCATION

We first study the case when each frame consists of only a single-slot ($M = 1$). In this case, we have for $\mathbf{x} \in \mathbb{Z}_+^2$

$$\begin{aligned} V_T(\mathbf{x}) &= \bar{c}(\mathbf{x}) \\ V_t(\mathbf{x}) &= \bar{c}(\mathbf{x}) + T_1 \hat{S}_{t+1}(\mathbf{x}), \quad 1 \leq t \leq T-1. \end{aligned} \quad (4)$$

Definition 2: A function $f : \mathbb{Z}_+^2 \rightarrow \mathbb{R}$ belongs to the set \mathcal{F} if $f(\mathbf{x})$ satisfies the following conditions.

C.1) (Monotonicity or Nondecreasing Condition)

$$f(\mathbf{x}) \leq f(\mathbf{x} + \mathbf{e}_i), \quad i \in \{1, 2\}.$$

C.2) (Supermodularity Condition)

$$f(\mathbf{x} + \mathbf{e}_1) + f(\mathbf{x} + \mathbf{e}_2) \leq f(\mathbf{x}) + f(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2).$$

C.3a) (Superconvexity Condition)

$$f(\mathbf{x} + \mathbf{e}_1) + f(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq f(\mathbf{x} + \mathbf{e}_2) + f(\mathbf{x} + 2\mathbf{e}_1).$$

C.3b) (Superconvexity Condition)

$$f(\mathbf{x} + \mathbf{e}_2) + f(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq f(\mathbf{x} + \mathbf{e}_1) + f(\mathbf{x} + 2\mathbf{e}_2).$$

Here, the terminologies follow that used in [14]. Note that these are rather benign conditions, and they specify a very large class of cost functions of practical interest. For example, all functions of the form $c_1 x_1^{n_1} + c_2 x_2^{n_2}$ (for any $c_1, c_2, n_1, n_2 > 0$) satisfy these conditions. An example of a function that does not satisfy the above conditions is $f(x_1, x_2) = x_1 x_2$ which fails conditions **C.3a** and **C.3b**. Also, note that conditions **C.2** and **C.3a** result in the convexity of f in x_1 . Similarly, **C.2** and **C.3b** imply the convexity of f in x_2 .

Definition 3: $\hat{\mathcal{F}}$ is the set of all functions $\hat{f} : \mathbb{Z}^2 \rightarrow \mathbb{R}$ satisfying **C.1–C.3**.

It immediately follows that $f \in \mathcal{F} \Rightarrow \hat{f} \in \hat{\mathcal{F}}$. The main result of this section is the following theorem.

Theorem 1: For the single-slot allocation problem, if the cost function $c(\cdot) \in \mathcal{F}$, then:

- 1) for any time t , we have $V_t(\mathbf{x}) \in \mathcal{F}$;
- 2) the optimal policy in assigning one slot is of the threshold type.

In the remainder of this section, we first show that if $V_{t+1}(\mathbf{x}) \in \mathcal{F}$, then $T_1 \hat{S}_{t+1}(\mathbf{x})$ restricted to $\mathbf{x} \in \mathbb{Z}_+^2$ is in \mathcal{F} . This is then used to prove Theorem 1. We proceed with a few lemmas.

Lemma 2: If $f \in \mathcal{F}$, then $\hat{g} : \mathbb{Z}^2 \rightarrow \mathbb{R}$ defined as $\hat{g}(\mathbf{x}) = f([\mathbf{x} + \mathbf{a}]^+)$ is in $\hat{\mathcal{F}}$ for all $\mathbf{a} \in \mathbb{Z}_+^2$.

Proof: We need to show that $\hat{g}(\mathbf{x}) = f([\mathbf{x} + \mathbf{a}]^+)$ satisfies conditions **C.1–C.3**.

- i) *Monotonicity:* $\hat{g}(\mathbf{x})$ obviously satisfies monotonicity since for $i = 1, 2$

$$\begin{aligned} \hat{g}(\mathbf{x} + \mathbf{e}_i) &= \begin{cases} f([\mathbf{x} + \mathbf{a}]^+), & (\mathbf{x} + \mathbf{a})_i < 0 \\ f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_i), & \text{else} \end{cases} \\ &\geq f([\mathbf{x} + \mathbf{a}]^+) = \hat{g}(\mathbf{x}) \end{aligned}$$

where the inequality is a result of the monotonicity of f .

- ii) *Supermodularity:* To prove this, we need to show

$$\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_2) \leq \hat{g}(\mathbf{x}) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2). \quad (5)$$

Letting $\mathbf{y} = (y_1, y_2) = \mathbf{x} + \mathbf{a}$, we consider the following four cases.

- 1) If $y_1, y_2 \geq 0$, then (5) becomes $f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_1) + f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_2) \leq f([\mathbf{x} + \mathbf{a}]^+) + f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_1 + \mathbf{e}_2)$, which is true since f satisfies **C.2**, by replacing \mathbf{x} with $[\mathbf{x} + \mathbf{a}]^+$ in **C.2**.
- 2) If $y_1 \geq 0, y_2 < 0$, then (5) becomes $f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_1) + f([\mathbf{x} + \mathbf{a}]^+) \leq f([\mathbf{x} + \mathbf{a}]^+) + f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_1)$, which is trivially true.
- 3) If $y_2 \geq 0, y_1 < 0$, the proof is the same as in case 2).
- 4) If $y_1, y_2 < 0$, then (5) becomes $f([\mathbf{x} + \mathbf{a}]^+) + f([\mathbf{x} + \mathbf{a}]^+) \leq f([\mathbf{x} + \mathbf{a}]^+) + f([\mathbf{x} + \mathbf{a}]^+)$, which again is trivially true.

- iii) *Superconvexity:* To prove **C.3a**, we need to show

$$\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq \hat{g}(\mathbf{x} + \mathbf{e}_2) + \hat{g}(\mathbf{x} + 2\mathbf{e}_1). \quad (6)$$

Again, let $\mathbf{y} = \mathbf{x} + \mathbf{a}$ and consider the same four cases.

- 1) If $y_1, y_2 \geq 0$, then (6) becomes $f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_1) + f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_1 + \mathbf{e}_2) \leq f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_2) + f([\mathbf{x} + \mathbf{a}]^+ + 2\mathbf{e}_1)$, which is true since f satisfies **C.3** (replacing \mathbf{x} with $[\mathbf{x} + \mathbf{a}]^+$ in **C.3**).
- 2) If $y_1 < 0, y_2 \geq 0$, then (6) becomes $f([\mathbf{x} + \mathbf{a}]^+) + f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_2) \leq f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_2) + f([\mathbf{x} + \mathbf{a}]^+ + 2\mathbf{e}_1)$, which is true by the monotonicity of f .
- 3) If $y_2 < 0, y_1 \geq 0$, then (6) becomes $f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_1) + f([\mathbf{x} + \mathbf{a}]^+ + \mathbf{e}_1) \leq f([\mathbf{x} + \mathbf{a}]^+) + f([\mathbf{x} + \mathbf{a}]^+ + 2\mathbf{e}_1)$, which is true by the convexity of f (combining **C.2** and **C.3a**).
- 4) If $y_1, y_2 < 0$, then (6) becomes $f([\mathbf{x} + \mathbf{a}]^+) + f([\mathbf{x} + \mathbf{a}]^+) \leq f([\mathbf{x} + \mathbf{a}]^+) + f([\mathbf{x} + \mathbf{a}]^+ + 2\mathbf{e}_1)$, which is true by the monotonicity of f .

C.3b can be proven in the same way and is not repeated. Therefore, we conclude $\hat{g}(\mathbf{x}) \in \hat{\mathcal{F}}$. ■

Lemma 3: If f_1, f_2, \dots are a sequence of functions that belong to \mathcal{F} , then $g(\mathbf{x}) = \sum_l p_l f_l(\mathbf{x})$ also belongs to \mathcal{F} , where p_l 's are non-negative constants.

Proof: We need to show that $g(\mathbf{x})$ satisfies **C.1–C.3**.

- i) *Monotonicity:* By the monotonicity of f_l , we have $g(\mathbf{x}) = \sum_l p_l f_l(\mathbf{x}) \leq \sum_l p_l f_l(\mathbf{x} + \mathbf{e}_1) = g(\mathbf{x} + \mathbf{e}_1)$, proving g 's monotonicity.

ii) *Supermodularity*: This holds because

$$\begin{aligned} g(\mathbf{x} + \mathbf{e}_1) + g(\mathbf{x} + \mathbf{e}_2) &= \sum_l p_l \cdot (f_l(\mathbf{x} + \mathbf{e}_1) + f_l(\mathbf{x} + \mathbf{e}_2)) \\ &\leq \sum_l p_l \cdot (f_l(\mathbf{x}) + f_l(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2)) \\ &= g(\mathbf{x}) + g(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \end{aligned}$$

where the inequality is due to the supermodularity of f_l .

iii) *Superconvexity*: This holds because

$$\begin{aligned} g(\mathbf{x} + \mathbf{e}_1) + g(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) &= \sum_l p_l \cdot (f_l(\mathbf{x} + \mathbf{e}_1) + f_l(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2)) \\ &\leq \sum_l p_l \cdot (f_l(\mathbf{x} + \mathbf{e}_2) + f_l(\mathbf{x} + 2\mathbf{e}_1)) \\ &= g(\mathbf{x} + \mathbf{e}_2) + g(\mathbf{x} + 2\mathbf{e}_1) \end{aligned}$$

where the inequality is due to the superconvexity of f_l .

C.3b can be shown in the same way and is thus omitted for brevity. ■

Lemma 4: If $\hat{f}_1, \hat{f}_2, \dots$ are a sequence of functions that belong to $\hat{\mathcal{F}}$, then $\hat{g}(\mathbf{x}) = \sum_l p_l \hat{f}_l(\mathbf{x})$ also belongs to $\hat{\mathcal{F}}$, where p_l 's are non-negative constants.

The proof is the same as in Lemma 3, and is thus not presented for brevity.

Lemma 5: If $\hat{f} \in \hat{\mathcal{F}}$, then $T_1 \hat{f} \in \hat{\mathcal{F}}$.

Proof: Let $\hat{g}(\mathbf{x}) = T_1 \hat{f}(\mathbf{x}) = \min\{\hat{f}(\mathbf{x} - \mathbf{e}_1), \hat{f}(\mathbf{x} - \mathbf{e}_2)\}$.

- i) *Monotonicity*: $\hat{g}(\mathbf{x}) \leq \hat{g}(\mathbf{x} + \mathbf{e}_1)$ holds, since the monotonicity of f results in an increment in both elements.
- ii) *Supermodularity*: We need to show that $\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_2) \leq \hat{g}(\mathbf{x}) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2)$. We will consider different cases depending on the minimizers of $\hat{g}(\mathbf{x})$ and $\hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2)$, denoted by m_1 and m_2 , respectively. For example, $m_1 = i, m_2 = j, i, j = 1, 2$ means $\hat{g}(\mathbf{x}) = \hat{f}(\mathbf{x} - \mathbf{e}_i)$, and $\hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) = \hat{f}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_j)$.

- 1) $m_1 = m_2 = 1$: In this case, the supermodularity condition we need to show becomes

$$\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_2) \leq \hat{f}(\mathbf{x} - \mathbf{e}_1) + \hat{f}(\mathbf{x} + \mathbf{e}_2). \quad (7)$$

To show this, consider

$$\begin{aligned} \hat{g}(\mathbf{x} + \mathbf{e}_1) &= \min\{\hat{f}(\mathbf{x}), \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2)\} \leq \hat{f}(\mathbf{x}) \\ \hat{g}(\mathbf{x} + \mathbf{e}_2) &= \min\{\hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1), \hat{f}(\mathbf{x})\} \leq \hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1) \end{aligned}$$

which yields $\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_2) \leq \hat{f}(\mathbf{x}) + \hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1)$. Letting $\mathbf{y} = \mathbf{x} - \mathbf{e}_1$, this becomes

$$\begin{aligned} \hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_2) &\leq \hat{f}(\mathbf{y} + \mathbf{e}_1) + \hat{f}(\mathbf{y} + \mathbf{e}_2) \\ &\leq \hat{f}(\mathbf{y}) + \hat{f}(\mathbf{y} + \mathbf{e}_1 + \mathbf{e}_2) \\ &= \hat{f}(\mathbf{x} - \mathbf{e}_1) + \hat{f}(\mathbf{x} + \mathbf{e}_2) \end{aligned}$$

where the second inequality is true by the supermodularity of \hat{f} , thus proving (7).

- 2) $m_1 = 1, m_2 = 2$: In this case, the supermodularity condition we need to show is

$$\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_2) \leq \hat{f}(\mathbf{x} - \mathbf{e}_1) + \hat{f}(\mathbf{x} + \mathbf{e}_1). \quad (8)$$

To show this, consider

$$\begin{aligned} \hat{g}(\mathbf{x} + \mathbf{e}_1) &= \min\{\hat{f}(\mathbf{x}), \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2)\} \leq \hat{f}(\mathbf{x}) \\ \hat{g}(\mathbf{x} + \mathbf{e}_2) &= \min\{\hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1), \hat{f}(\mathbf{x})\} \leq \hat{f}(\mathbf{x}) \\ &\Rightarrow \hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_2) \leq 2\hat{f}(\mathbf{x}) \\ &\leq \hat{f}(\mathbf{x} - \mathbf{e}_1) + \hat{f}(\mathbf{x} + \mathbf{e}_1) \end{aligned}$$

where the last inequality is due to the convexity of \hat{f} , thus proving (8).

The two remaining cases where $m_1 = m_2 = 2$ and $m_1 = 2, m_2 = 1$ can be shown similarly, and are not repeated here.

- iii) *Superconvexity*: First, we show that \hat{g} satisfies **C.3a**, i.e.,

$$\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq \hat{g}(\mathbf{x} + 2\mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_2). \quad (9)$$

We again consider different cases depending on the minimizers for the two terms on the right, respectively, denoted by m_1 and m_2 , as in the case of supermodularity.

- 1) $m_1 = m_2 = 1$: In this case, (9) becomes $\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq \hat{f}(\mathbf{x} + \mathbf{e}_1) + \hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1)$. To show this, we have

$$\begin{aligned} \hat{g}(\mathbf{x} + \mathbf{e}_1) &= \min\{\hat{f}(\mathbf{x}), \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2)\} \leq \hat{f}(\mathbf{x}) \\ \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) &= \min\{\hat{f}(\mathbf{x} + \mathbf{e}_2), \hat{f}(\mathbf{x} + \mathbf{e}_1)\} \leq \hat{f}(\mathbf{x} + \mathbf{e}_2). \end{aligned}$$

Therefore, by letting $\mathbf{y} = \mathbf{x} - \mathbf{e}_1$, we have

$$\begin{aligned} \hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) &\leq \hat{f}(\mathbf{x}) + \hat{f}(\mathbf{x} + \mathbf{e}_2) \\ &= \hat{f}(\mathbf{y} + \mathbf{e}_1) + \hat{f}(\mathbf{y} + \mathbf{e}_1 + \mathbf{e}_2) \\ &\leq \hat{f}(\mathbf{y} + 2\mathbf{e}_1) + \hat{f}(\mathbf{y} + \mathbf{e}_2) \\ &= \hat{f}(\mathbf{x} + \mathbf{e}_1) + \hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1) \end{aligned}$$

where the second inequality is due to the superconvexity of \hat{f} , thus proving (9).

- 2) $m_1 = 1, m_2 = 2$: In this case, (9) becomes $\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq \hat{f}(\mathbf{x} + \mathbf{e}_1) + \hat{f}(\mathbf{x})$. In order to show this consider

$$\begin{aligned} \hat{g}(\mathbf{x} + \mathbf{e}_1) &= \min\{\hat{f}(\mathbf{x}), \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2)\} \leq \hat{f}(\mathbf{x}) \\ \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) &= \min\{\hat{f}(\mathbf{x} + \mathbf{e}_2), \hat{f}(\mathbf{x} + \mathbf{e}_1)\} \leq \hat{f}(\mathbf{x} + \mathbf{e}_1) \\ &\Rightarrow \hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq \hat{f}(\mathbf{x}) + \hat{f}(\mathbf{x} + \mathbf{e}_1) \end{aligned}$$

proving (9).

- 3) $m_1 = 2, m_2 = 1$: By superconvexity of f (and thus \hat{f}), we have

$$\begin{aligned}\hat{f}(\mathbf{x}) - \hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1) &\leq \hat{f}(\mathbf{x} - \mathbf{e}_2) - \hat{f}(\mathbf{x} - \mathbf{e}_1) \\ \hat{f}(\mathbf{x} - \mathbf{e}_2) - \hat{f}(\mathbf{x} - \mathbf{e}_1) &\leq \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2) - \hat{f}(\mathbf{x}) \\ \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2) - \hat{f}(\mathbf{x}) &\leq \hat{f}(\mathbf{x} + 2\mathbf{e}_1 - \mathbf{e}_2) - \hat{f}(\mathbf{x} + \mathbf{e}_1)\end{aligned}$$

where the first inequality results from **C.3b** and the other two inequalities are a consequence of **C.3a**. Combining (adding) these inequalities, we get

$$\hat{f}(\mathbf{x}) - \hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1) \leq \hat{f}(\mathbf{x} + 2\mathbf{e}_1 - \mathbf{e}_2) - \hat{f}(\mathbf{x} + \mathbf{e}_1). \quad (10)$$

However, note that whenever $m_1 = 2$ (i.e., $\hat{f}(\mathbf{x} + 2\mathbf{e}_1 - \mathbf{e}_2) \leq \hat{f}(\mathbf{x} + \mathbf{e}_1)$), the right-hand side of the above equation is non-positive, thus the left-hand side is also non-positive. This implies that $m_2 = 2$. In other words, $m_1 = 2, m_2 = 1$ implies $m_1 = 2, m_2 = 2$, meaning $\hat{g}(\mathbf{x} + \mathbf{e}_2) = \hat{f}(\mathbf{x} + \mathbf{e}_2 - \mathbf{e}_1) = \hat{f}(\mathbf{x})$. Therefore, the case of $m_1 = 2, m_2 = 1$ is a special case of $m_1 = 2, m_2 = 2$, which is dealt with next.

- 4) $m_1 = 2, m_2 = 2$: In this case, (9) becomes $\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq \hat{f}(\mathbf{x} + 2\mathbf{e}_1 - \mathbf{e}_2) + \hat{f}(\mathbf{x})$. To show this consider

$$\begin{aligned}\hat{g}(\mathbf{x} + \mathbf{e}_1) &= \min\{\hat{f}(\mathbf{x}), \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2)\} \leq \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2) \\ \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) &= \min\{\hat{f}(\mathbf{x} + \mathbf{e}_2), \hat{f}(\mathbf{x} + \mathbf{e}_1)\} \leq \hat{f}(\mathbf{x} + \mathbf{e}_1).\end{aligned}$$

Letting $\mathbf{y} = \mathbf{x} - \mathbf{e}_2$, we have

$$\begin{aligned}\hat{g}(\mathbf{x} + \mathbf{e}_1) + \hat{g}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) &\leq \hat{f}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2) + \hat{f}(\mathbf{x} + \mathbf{e}_1) \\ &= \hat{f}(\mathbf{y} + \mathbf{e}_1) + \hat{f}(\mathbf{y} + \mathbf{e}_1 + \mathbf{e}_2) \\ &\leq \hat{f}(\mathbf{y} + 2\mathbf{e}_1) + \hat{f}(\mathbf{y} + \mathbf{e}_2) \\ &= \hat{f}(\mathbf{x} + 2\mathbf{e}_1 - \mathbf{e}_2) + \hat{f}(\mathbf{x})\end{aligned} \quad (11)$$

thus proving (9). That \hat{g} also satisfies **C.3b** can be shown in a similar way and is not repeated. Therefore, we conclude that if $\hat{f} \in \hat{\mathcal{F}}$, then $\hat{g} = T_1\hat{f} \in \hat{\mathcal{F}}$, proving the lemma. \blacksquare

The following lemma is also stated in [18].

Lemma 6: If $\hat{f}(\mathbf{x}) \in \hat{\mathcal{F}}$, then the restriction of $\hat{f}(\mathbf{x})$ to non-negative values is in \mathcal{F} .

We are now ready to prove Theorem 1, assuming two users and single-slot frames.

Proof of Theorem 1:

- 1) We prove the result by induction. First, note that if $c(\cdot) \in \mathcal{F}$, then $\bar{c}(\mathbf{x}) \in \mathcal{F}$ by Lemma 3, therefore $V_T(\mathbf{x}) = \bar{c}(\mathbf{x})$ is in \mathcal{F} . This completes the induction basis. Next, we show that if $V_{t+1}(\mathbf{x}) \in \mathcal{F}$, then $V_t(\mathbf{x}) \in \mathcal{F}$. By Lemmas 2 and 4, we have that if $V_{t+1}(\mathbf{x}) \in \mathcal{F}$, then $\hat{S}_{t+1}(\mathbf{x}) \in \hat{\mathcal{F}}$. Therefore, by Lemma 5, $T_1\hat{S}_{t+1}(\mathbf{x}) \in \hat{\mathcal{F}}$.

Using Lemma 6, we have that $T_1\hat{S}_{t+1}(\mathbf{x})$ restricted to non-negative values is in \mathcal{F} . Since $\bar{c}(\mathbf{x}) \in \mathcal{F}$, $\bar{c}(\mathbf{x}) + T_1\hat{S}_{t+1}(\mathbf{x})$ restricted to non-negative values is in \mathcal{F} by Lemma 3, and by Lemma 1 this value is equal to $V_t(\mathbf{x})$. Thus, $V_t(\mathbf{x}) \in \mathcal{F}$, completing the induction step.

- 2) By part 1) of this theorem, $V_{t+1} \in \mathcal{F}$ for all t . Therefore, $\hat{S}_{t+1} \in \hat{\mathcal{F}}$. Thus, by property **C.3a**, we have $\hat{S}_{t+1}(\mathbf{x} + \mathbf{e}_1) + \hat{S}_{t+1}(\mathbf{x} + \mathbf{e}_1 + \mathbf{e}_2) \leq \hat{S}_{t+1}(\mathbf{x} + 2\mathbf{e}_1) + \hat{S}_{t+1}(\mathbf{x} + \mathbf{e}_2)$.

By replacing \mathbf{x} with $\mathbf{x} - \mathbf{e}_1 - \mathbf{e}_2$, we have $\hat{S}_{t+1}(\mathbf{x} - \mathbf{e}_2) + \hat{S}_{t+1}(\mathbf{x}) \leq \hat{S}_{t+1}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2) + \hat{S}_{t+1}(\mathbf{x} - \mathbf{e}_1)$. Rearranging, we get

$$\hat{S}_{t+1}(\mathbf{x} - \mathbf{e}_2) - \hat{S}_{t+1}(\mathbf{x} - \mathbf{e}_1) \leq \hat{S}_{t+1}(\mathbf{x} + \mathbf{e}_1 - \mathbf{e}_2) - \hat{S}_{t+1}(\mathbf{x}).$$

The last inequality suggests that if the left-hand side is non-negative, then the right-hand side is also non-negative. Therefore, if the optimal decision is to allocate to the first queue when the state is \mathbf{x} for some \mathbf{x} , then it is optimal to allocate the slot to the first queue when the state is $\mathbf{x} + \mathbf{e}_1$. Similarly using **C.3b**, we can show that if the optimal decision is to allocate to the second queue when the state is \mathbf{x} , then it is optimal to allocate the slot to the second queue when the state is $\mathbf{x} + \mathbf{e}_2$. We can then define a threshold as follows:

$$h_t(x_1) = \min\{x_2 | \hat{S}_{t+1}(\mathbf{x} - \mathbf{e}_2) \leq \hat{S}_{t+1}(\mathbf{x} - \mathbf{e}_1)\} \quad (12)$$

and $h_t(x_1) = \infty$ when the above set is empty. If we have $x_{2,t} \geq h_t(x_{1,t})$, then the optimal policy is to assign the slot at time t to queue 2, otherwise, to queue 1 (if the set is empty then the threshold is infinity), proving the optimality of a threshold policy. \blacksquare

While Theorem 1 shows that the optimal scheduler is of the threshold type, it is worth pointing out that it is, in general, difficult to obtain the quantitative value of the threshold. The threshold is given by (12), where the current cost-to-go function needs to be calculated. This can be computationally expensive.

Note that throughout our discussion, none of the results obtained depends on the arrival process and the cost function being time invariant. The proof of Theorem 1 is based on induction, i.e., as long as the induction basis holds, the induction step is established by using the properties of the value function from the induction hypothesis and using the fact that $c(\cdot) \in \mathcal{F}$. In other words, all results developed here are equally applicable to a time-varying cost function $c_t(\cdot)$ and time-varying arrival process $p_t(\cdot)$. In particular, Lemma 1 will hold with $\bar{c}(\mathbf{x})$ replaced with $\bar{c}_t(\mathbf{x})$, and Theorem 1 will hold by requiring that $c_t(\cdot) \in \mathcal{F}, \forall t$, rather than requiring that $c(\cdot) \in \mathcal{F}$. Therefore, as we mentioned earlier, the time-invariant assumptions are merely for the simplicity of notation and can be easily relaxed. The same argument also holds for results obtained in the next two sections.

IV. MULTIPLE SLOT BATCH ALLOCATION

In this section, we consider the problem of allocating $M > 1$ slots for each time frame. The following example shows that,

in general, a sequential allocation of slots does not necessarily lead to the optimal policy for allocating M slots.

Example 1: Suppose $T = 2$ and let $p_1(0) = p_2(0) = 1$, i.e., there are no arrivals. Let $b_{1,0} = 3$, $b_{2,0} = 2$, and $c(\mathbf{b}_t) = b_{1,t}^2 \cdot b_{2,t}$. Finally, let $M = 2$. Since $T = 2$, the queues only get to transmit during the second frame. The queue occupancy thus remains the same for $t = 0$ and $t = 1$ no matter what strategy is used. Therefore, to minimize the total cost, we need only to focus on $t = T = 2$ and minimize the cost at time $t = 2$. It can be easily verified that the optimal allocation at $t = 1$ is $w_{1,1}^* = 0$, $w_{2,1}^* = 2$, resulting in a cost of zero at $t = 2$.

Now, consider the sequential allocation, which proceeds as follows. We suppose there is only one slot in the frame to allocate and it needs to be allocated in such a way to minimize the cost at $t = 2$. If the slot is allocated to queue 1, the cost at $t = 2$ will be 8 and if the slot is allocated to queue 2, the cost at $t = 2$ will be 9. Thus, the optimal allocation of the first slot is to queue 1. The updated state at $t = 1$ given the first allocation (to queue 1) is $x_{1,t} = x_{2,t} = 2$. For the allocation of the second slot, again we suppose there is only one slot in the frame to allocate to minimize the cost at $t = 2$. It can be seen that the second slot should also be allocated to the first queue. The two sequential steps result in both slots allocated to queue 1 and none to queue 2, with a cost of 3 at $t = 2$. Obviously, this policy is not optimal.

In the rest of this section, we show that under the conditions introduced in the previous section, an optimal policy for allocating M slots can be obtained by sequentially allocating the slots according to the optimal policy for a single-slot allocation.

Definition 4: Define recursively the operator T_1^k as $T_1^k f(\mathbf{x}) := T_1(T_1^{k-1} f(\mathbf{x}))$.

Theorem 2: If $\hat{f}(\mathbf{x}) \in \hat{\mathcal{F}}$, then we have $T_M \hat{f}(\mathbf{x}) = T_1^M \hat{f}(\mathbf{x})$.

Proof: We use induction on M . The induction basis for $M = 1$ is trivially true. Suppose that the theorem holds for $M = m$, i.e., $T_m \hat{f}(\mathbf{x}) = T_1^m \hat{f}(\mathbf{x})$, we want to show that it holds for $M = m + 1$.

Denote by $\mathbf{w}^1 := \arg \min_{\mathbf{w}: w_1 + w_2 = m} \{\hat{f}(\mathbf{x} - \mathbf{w})\}$. Suppose we have $m + 1$ slots to assign: By definition, we have

$$T_{m+1} \hat{f}(\mathbf{x}) = \min_{\mathbf{w}: w_1 + w_2 = m+1} \left\{ \hat{f}(\mathbf{x} - \mathbf{w}) \right\}. \quad (13)$$

Below, we show that the allocation $\mathbf{w}^1 + \mathbf{e}_i$, $i \in \{1, 2\}$, is at least “as good as” all allocations of the form $\mathbf{w}^1 + (k+1)\mathbf{e}_i - k\mathbf{e}_j$, for all $1 \leq k \leq w_j^1$, in minimizing the right-hand side of (13), i.e., we want to show the following for $i \neq j$:

$$\begin{aligned} \hat{f}(\mathbf{x} - (w_i^1 + k + 1)\mathbf{e}_i - (w_j^1 - k)\mathbf{e}_j) \\ \geq \hat{f}(\mathbf{x} - (w_i^1 + 1)\mathbf{e}_i - w_j^1 \mathbf{e}_j). \end{aligned} \quad (14)$$

Since $\mathbf{w}^1 + (k+1)\mathbf{e}_i - k\mathbf{e}_j$, $1 \leq k \leq w_j^1$, denotes all possible allocations between the two users other than the allocation denoted by $\mathbf{w}^1 + \mathbf{e}_i$, if we can show (14), then we will have established that $\mathbf{w}^1 + \mathbf{e}_i$ minimizes the right-hand side of (13). It is thus sufficient to show that if $\mathbf{w}^1 + (k+1)\mathbf{e}_i - k\mathbf{e}_j$ minimizes the right-hand side of (13), then $\mathbf{w}^1 + \mathbf{e}_i$ will also minimize the right-hand side of (13). Assume now that $\mathbf{w}^1 + (k+1)\mathbf{e}_i - k\mathbf{e}_j$ minimizes the right-hand side of (13) and let $\mathbf{w}^2 = \mathbf{w}^1 + \mathbf{e}_i$. We

proceed by first showing that the following holds for all values $1 \leq k \leq w_j^1$:

$$\begin{aligned} \hat{f}(\mathbf{x} - (w_i^1 + k)\mathbf{e}_i - (w_j^1 - k)\mathbf{e}_j) - \hat{f}(\mathbf{x} - w_i^1 \mathbf{e}_i - w_j^1 \mathbf{e}_j) \\ \leq \hat{f}(\mathbf{x} - (w_i^1 + k + 1)\mathbf{e}_i - (w_j^1 - k)\mathbf{e}_j) \\ - \hat{f}(\mathbf{x} - (w_i^1 + 1)\mathbf{e}_i - w_j^1 \mathbf{e}_j). \end{aligned} \quad (15)$$

We show this by using induction on k . First, consider $k = 1$, i.e., we need to show

$$\begin{aligned} \hat{f}(\mathbf{x} - (w_i^1 + 1)\mathbf{e}_i - (w_j^1 - 1)\mathbf{e}_j) - \hat{f}(\mathbf{x} - w_i^1 \mathbf{e}_i - w_j^1 \mathbf{e}_j) \\ \leq \hat{f}(\mathbf{x} - (w_i^1 + 2)\mathbf{e}_i - (w_j^1 - 1)\mathbf{e}_j) \\ - \hat{f}(\mathbf{x} - (w_i^1 + 1)\mathbf{e}_i - w_j^1 \mathbf{e}_j). \end{aligned} \quad (16)$$

Equation (16) can be obtained by replacing \mathbf{x} with $\mathbf{x} - (w_i^1 + 2)\mathbf{e}_i - w_j^1 \mathbf{e}_j$ in property **C.3** (use **C.3a** if $i = 1$ and use **C.3b** if $i = 2$). Thus, the induction basis is established.

Now, assume (15) is true for $k = l$, $1 \leq l < w_j^1$, we want to show that is also true for $k = l + 1$. In property **C.3** (use **C.3a** if $i = 1$ and use **C.3b** if $i = 2$), substituting \mathbf{x} for $\mathbf{x} - (w_i^1 + l + 2)\mathbf{e}_i - (w_j^1 - l)\mathbf{e}_j$ gives

$$\begin{aligned} \hat{f}(\mathbf{x} - (w_i^1 + l + 1)\mathbf{e}_i - (w_j^1 - l - 1)\mathbf{e}_j) \\ - \hat{f}(\mathbf{x} - (w_i^1 + l)\mathbf{e}_i - (w_j^1 - l)\mathbf{e}_j) \\ \leq \hat{f}(\mathbf{x} - (w_i^1 + l + 2)\mathbf{e}_i - (w_j^1 - l - 1)\mathbf{e}_j) \\ - \hat{f}(\mathbf{x} - (w_i^1 + l + 1)\mathbf{e}_i - (w_j^1 - l)\mathbf{e}_j). \end{aligned} \quad (17)$$

Combining the induction hypothesis and (17) gives the result for the case of $k = l + 1$ and the induction is complete.

Next, note that we have $\hat{f}(\mathbf{x} - (w_i^1 + k)\mathbf{e}_i - (w_j^1 - k)\mathbf{e}_j) \geq \hat{f}(\mathbf{x} - w_i^1 \mathbf{e}_i - w_j^1 \mathbf{e}_j)$, $1 \leq k \leq w_j^1$, due to the optimality of \mathbf{w}^1 when there are m slots to allocate. Therefore, the left-hand side of (15) is always non-negative, thus so is the right-hand side, i.e.,

$$\begin{aligned} \hat{f}(\mathbf{x} - (w_i^1 + k + 1)\mathbf{e}_i - (w_j^1 - k)\mathbf{e}_j) \\ \geq \hat{f}(\mathbf{x} - (w_i^1 + 1)\mathbf{e}_i - w_j^1 \mathbf{e}_j). \end{aligned}$$

This means that \mathbf{w}^2 minimizes the right-hand side of (13).

The above result shows that the minimizer on the right-hand side of (13) can be found by taking the minimum between $\mathbf{w}^1 + \mathbf{e}_1$ and $\mathbf{w}^1 + \mathbf{e}_2$. Following this result, for the $(m + 1)$ th slot, we have $T_{m+1} \hat{f}(\mathbf{x}) = \min_{i \in \{1, 2\}} \{\hat{f}(\mathbf{x} - \mathbf{w}^1 - \mathbf{e}_i)\}$, where \mathbf{w}^1 is the minimizer for m slots, i.e., $\hat{f}(\mathbf{x} - \mathbf{w}^1) = T_m \hat{f}(\mathbf{x})$. Thus, we have $T_{m+1} \hat{f}(\mathbf{x}) = T_1 T_m \hat{f}(\mathbf{x})$ using the induction hypothesis. By the recursive definition of the operator, we have $T_{m+1} \hat{f}(\mathbf{x}) = T_1^{m+1} \hat{f}(\mathbf{x})$, completing the proof. ■

Consider two users and M slots in each time frame, and assume that the optimal policy is known for the single-slot allocation. We next use Theorem 2 to show that the same policy for a single-slot allocation can be repeatedly/sequentially used M times, and it results in the optimal policy for allocating the batch of M slots.

Theorem 3: Consider M slots to allocate. If $c(\cdot) \in \mathcal{F}$, then $V_t(\mathbf{x}) \in \mathcal{F}$ for all $t \leq T$. Furthermore, the policy that sequentially assigns each slot optimally given the state and the previous allocations, is optimal.

Proof: We use backward induction on t . Since $c(\cdot) \in \mathcal{F}$, we have $V_T(\mathbf{x}) \in \mathcal{F}$, which establishes the induction basis. Next suppose that $V_t(\mathbf{x}) \in \mathcal{F}$. We want to show that $V_{t-1} \in \mathcal{F}$. Since $V_t(\mathbf{x}) \in \mathcal{F}$, $\hat{S}_t(\mathbf{x}) \in \hat{\mathcal{F}}$, using Theorem 2, we have for $\mathbf{x} \in \mathbb{Z}_+^2$

$$V_{t-1}(\mathbf{x}) = \bar{c}(\mathbf{x}) + T_M \hat{S}_t(\mathbf{x}) = \bar{c}(\mathbf{x}) + T_1^M (\hat{S}_t(\mathbf{x})). \quad (18)$$

By Lemma 5, we have $T_1^M (\hat{S}_t(\mathbf{x})) \in \hat{\mathcal{F}}$, therefore its restriction to \mathbb{Z}_+^2 is in \mathcal{F} by Lemma 6. Also, we have $\bar{c}(\mathbf{x}) \in \mathcal{F}$ since $c(\mathbf{b}) \in \mathcal{F}$. Therefore, the right-hand side of the above equation is in \mathcal{F} by Lemma 3, thus $V_{t-1}(\mathbf{x}) \in \mathcal{F}$, completing the induction.

Next, we show that this allocation problem reduces to optimally allocating a single slot. It should be evident from (18) that finding the allocation vector $\mathbf{w} : w_1 + w_2 = M$ by solving $T_M \hat{S}_t(\mathbf{x})$ is equivalent to solving $T_1^M \hat{S}_t(\mathbf{x})$, which implies allocating one slot at a time. More specifically, consider allocating M slots within frame t . Having already allocated m slots ($m < M$) within the frame with allocation $\mathbf{w} : w_1 + w_2 = m$, the optimal allocation of the next slot, by definition of $T_1 T^m$, is $\arg \min_{i=1,2} \{E_{\mathbf{a}} V_{t-1}([\mathbf{x} + \mathbf{a} - \mathbf{w} - \mathbf{e}_i]^+)\}$, which simply shows that it is optimal to allocate the $(m+1)$ th slot given the system state \mathbf{x} and prior allocation in the same frame \mathbf{w} . ■

The above result shows that the M slot allocation problem reduces to the single-slot allocation problem.

V. INFINITE HORIZON DISCOUNTED COST AND AVERAGE COST

In this section, we study the properties of the optimal policy when $T \rightarrow \infty$. Note that the cost defined in (1) is infinite as $T \rightarrow \infty$, except for certain special cases. In this section, we consider two alternatives for defining the cost over an infinite horizon, the discounted cost and the average cost.

A. Discounted Cost

Consider the discount factor $\beta(0 < \beta < 1)$, and define the t step minimum cost function

$$W_t(\mathbf{x}) = \min_{\pi} E^{\pi} \left\{ \sum_{u=1}^t \beta^{u-1} \bar{c}(\mathbf{x}_u) | \mathbf{x}_1 = \mathbf{x} \right\}. \quad (19)$$

Note here t denotes the number of frames to go (or the horizon), rather than the actual time as in previous sections. It can be shown that $W_t(\mathbf{x})$ satisfies the following recursion:

$$\begin{aligned} W_0(\mathbf{x}) &= 0; \\ W_t(\mathbf{x}) &= \bar{c}(\mathbf{x}) \\ &+ \beta \min_{\mathbf{w}: w_1+w_2=M} E_{\mathbf{a}} [W_{t-1}([\mathbf{x} + \mathbf{a} - \mathbf{w}]^+)]. \end{aligned} \quad (20)$$

Definition 5: Define $\hat{R}(\mathbf{x}) : \mathbb{Z}^2 \rightarrow \mathbb{R} \cup \{\infty\}$ as follows:

$$\hat{R}_t(\mathbf{x}) = \sum_{\mathbf{a}} p(\mathbf{a}) W_t([\mathbf{x} + \mathbf{a}]^+). \quad (21)$$

The following lemma then follows directly.

Lemma 7: For all values $0 < t < T$, $W_t(\mathbf{x})$ is equal to $\bar{c}(\mathbf{x}) + \beta T_M \hat{R}_{t-1}(\mathbf{x})$ restricted to $\mathbf{x} \in \mathbb{Z}_+^2$.

Lemma 8: Consider two users and M slots to allocate. If $c(\cdot) \in \mathcal{F}$, then $W_t(\mathbf{x}) \in \mathcal{F}$ for all $t \geq 0$.

The proof of this lemma is similar to that of the same result for V_t in the previous section, except that instead of backward induction, we need to use forward induction for W_t , noting that $W_0(\mathbf{x}) = 0$, and thus $W_0(\mathbf{x}) \in \mathcal{F}$. The complete proof is not presented for brevity.

Define the infinite horizon cost as follows:

$$W_{\infty}(\mathbf{x}) = \min_{\pi} E^{\pi} \left\{ \lim_{t \rightarrow \infty} \sum_{u=1}^t \beta^{u-1} \bar{c}(\mathbf{x}_u) | \mathbf{x}_1 = \mathbf{x} \right\}. \quad (22)$$

Note that $c(\mathbf{x})$ is not necessarily bounded. However, if we have $c(\mathbf{x}) \geq 0$ for all $\mathbf{x} \geq 0$, then $W_{\infty}(\mathbf{x})$ satisfies the following (for more details and proof see [17, Ch. 5.4]):

$$\begin{aligned} W_{\infty}(\mathbf{x}) &= \bar{c}(\mathbf{x}) + \beta \min_{\mathbf{w}: w_1+w_2=M} E_{\mathbf{a}} [W_{\infty}([\mathbf{x} + \mathbf{a} - \mathbf{w}]^+)] \\ W_{\infty}(\mathbf{x}) &= \lim_{t \rightarrow \infty} W_t(\mathbf{x}). \end{aligned} \quad (23)$$

Theorem 4: Consider two users and M slots to allocate. If $c(\cdot) \in \mathcal{F}$ and is non-negative, then $W_{\infty}(\mathbf{x}) \in \mathcal{F}$ and the optimal policy for a single-slot allocation is of the threshold type. Furthermore, the policy that assigns each slot optimally given the state and the previous allocation in the same frame, is optimal.

Proof: Note that $W_t(\mathbf{x}) \in \mathcal{F}$ for all t and that the set \mathcal{F} is closed under pointwise limit of functions, i.e., if f_1, f_2, \dots is a sequence of functions and $f_i \in \mathcal{F}, \forall i$, and if $f = \lim_{n \rightarrow \infty} f_n$, then $f \in \mathcal{F}$. Therefore, by using Lemma 8 and (23), we have $W_{\infty}(\mathbf{x}) \in \mathcal{F}$. The rest of the theorem follows from the same arguments used in the proofs of Theorems 1 and 3. ■

B. Average Cost

One may also choose to minimize the average cost over time, rather than discounted cost. Consider the following cost function:

$$\bar{J}^{\pi} = E^{\pi} [\bar{C} | \mathbf{b}_0, \mathbf{w}_0], \quad \bar{C} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T c(\mathbf{b}_t). \quad (24)$$

Recall the infinite horizon discounted cost defined before

$$W_{\beta}(\mathbf{x}) = \min_{\pi} E^{\pi} \left\{ \lim_{t \rightarrow \infty} \sum_{u=1}^t \beta^{u-1} \bar{c}(\mathbf{x}_u) | \mathbf{x}_1 = \mathbf{x} \right\}.$$

Here, we have used $W_{\beta}(\mathbf{x})$ to denote this cost rather than $W_{\infty}(\mathbf{x})$ as used before. This is because in this subsection we will focus on this cost as a function of the value β , while always taking the horizon to be infinite.

Recall we have shown that the following holds in (23):

$$W_{\beta}(\mathbf{x}) = \bar{c}(\mathbf{x}) + \beta \min_{\mathbf{w}: w_1+w_2=M} E_{\mathbf{a}} [W_{\beta}([\mathbf{x} + \mathbf{a} - \mathbf{w}]^+)]. \quad (25)$$

Consider the following assumption.

Assumption (AVG-1): For any state $\mathbf{x} > \mathbf{0}$ there exists a policy $\pi_{\mathbf{x}}$ such that starting from state \mathbf{x} , it takes the queue-size back to state $\mathbf{0}$ with finite expected number of steps and finite expected cost. Let the expected (nondiscounted) cost for this transition be denoted by $U(\mathbf{x})$.

Define $h_{\beta}(\mathbf{x})$ as follows:

$$h_{\beta}(\mathbf{x}) = W_{\beta}(\mathbf{x}) - W_{\beta}(\mathbf{0}).$$

Let $\{\beta_n\}$ be a sequence of real numbers such that $\beta_n \rightarrow 1^-$ as $n \rightarrow \infty$. Then, it is shown in Lemma A-3 that under Assumption AVG-1 one can find a subsequence α_n such that $\lim_{n \rightarrow \infty} h_{\alpha_n}(\mathbf{x})$ exists. We call this limit function $h(\mathbf{x})$. We then have the following theorem.

Theorem 5: Suppose $c(\mathbf{x}) \geq 0$ for all $\mathbf{x} \geq \mathbf{0}$ and that Assumption AVG-1 holds. Then we have the following.

- a) There exists a finite constant J^* that satisfies the following inequality:

$$J^* + h(\mathbf{x}) \geq \bar{c}(\mathbf{x}) + \min_{w_1+w_2=M} E_{\mathbf{a}} [h(\mathbf{x} + \mathbf{a})]. \quad (26)$$

- b) Let π^* be a policy that minimizes the right-hand side of (26). Then, π^* is the optimal average cost policy.
c) J^* is the optimum average cost.

The proof of this theorem follows closely the argument used in [19, Ch. 7]. However, for self-sufficiency, we have included the proof in the appendix.

Theorem 6: Consider two users and M slots to allocate. If $c(\cdot) \in \mathcal{F}$ and is non-negative, then $h(\mathbf{x}) \in \mathcal{F}$ and the optimal average cost policy for a single-slot allocation is of the threshold type. Furthermore, the policy that assigns each slot optimally given the state and the previous allocation in the same frame, is optimal.

Proof: Note that $h(\mathbf{x}) = \lim_{\beta \rightarrow 1^-} W_{\beta}(\mathbf{x}) - W_{\beta}(\mathbf{0})$. Since we have $W_{\beta}(\mathbf{x}) \in \mathcal{F}$ by Theorem 4, we conclude that $h(\mathbf{x}) \in \mathcal{F}$. The rest of the proof is very similar to the proofs of Theorems 1 and 3, and is not repeated for brevity. ■

VI. LINEAR, EQUAL HOLDING COST

In this section, we consider the special case when the cost function is linear and equal for both queues. Let c be the cost of having a packet in queue, then the cost of queue i at time t would be $cb_{i,t}$. We also assume that the arrivals to different queues are independent, i.e., $p(\mathbf{a}) = p_1(a_1)p_2(a_2)$, where $p_i(a)$ is the probability of having a arrivals in queue i during a time frame. We will focus on the finite horizon problem and use the results derived in the previous sections to characterize the optimal allocation in this case. From Section IV, it suffices to concentrate on allocating a single slot.

Lemma 9: Suppose for two queues we have $c(\mathbf{b}_t) = cb_{1,t} + cb_{2,t}$. Then, for all $\mathbf{x} \in \mathbb{Z}_+^2$, we have $V_t(\mathbf{x} + \mathbf{e}_1) = V_t(\mathbf{x} + \mathbf{e}_2)$.

This lemma essentially says that because the two queues are symmetric, the future cost to go remains the same as long as the total number of packets in the system is the same, regardless of which queue they are in. This in turn suggests that when both queues are nonempty (the deterministic part), it is equally optimal to allocate the slot to either queue.

Proof: We use induction on t to prove the lemma. The statement is obviously true for $t = T$. Now, suppose the statement is true for $t + 1$, i.e., $V_{t+1}(\mathbf{x} + \mathbf{e}_1) = V_{t+1}(\mathbf{x} + \mathbf{e}_2)$, $\forall \mathbf{x} \in \mathbb{Z}_+^2$. We want to show that $V_t(\mathbf{x} + \mathbf{e}_1) = V_t(\mathbf{x} + \mathbf{e}_2)$.

Suppose the state is $\mathbf{x} + \mathbf{e}_1$ at t for some $\mathbf{x} \geq \mathbf{0}$. By the dynamic equation of the problem given in (2), the slot is allocated to the first queue if

$$\begin{aligned} & \sum_{a_1, a_2=0}^{\infty} p_1(a_1)p_2(a_2)V_{t+1}(\mathbf{x} + a_1\mathbf{e}_1 + a_2\mathbf{e}_2) \\ & \leq \sum_{a_1, a_2=0}^{\infty} p_1(a_1)p_2(a_2) \\ & \quad \times V_{t+1}([\mathbf{x} + (a_1 + 1)\mathbf{e}_1 + (a_2 - 1)\mathbf{e}_2]^+). \end{aligned} \quad (27)$$

Using the nondecreasing property of $V_{t+1}(\cdot)$ and the induction hypothesis, we have that for any value of $a_1, a_2 \geq 0$

$$\begin{aligned} & V_{t+1}(\mathbf{x} + a_1\mathbf{e}_1 + a_2\mathbf{e}_2) \\ & \leq V_{t+1}([\mathbf{x} + a_1\mathbf{e}_1 + (a_2 - 1)\mathbf{e}_2]^+ + \mathbf{e}_2) \\ & = V_{t+1}([\mathbf{x} + (a_1 + 1)\mathbf{e}_1 + (a_2 - 1)\mathbf{e}_2]^+). \end{aligned}$$

Thus, (27) holds, and we conclude that it is optimal to allocate the slot to the first queue. Similar arguments can be used to show that it is optimal to allocate the slot to the second queue at state $\mathbf{x} + \mathbf{e}_2$. Now, we can write

$$\begin{aligned} V_t(\mathbf{x} + \mathbf{e}_1) &= \bar{c}(\mathbf{x} + \mathbf{e}_1) + \sum_{a_1=0}^{\infty} \sum_{a_2=0}^{\infty} p_1(a_1)p_2(a_2) \\ & \quad \times V_{t+1}(\mathbf{x} + a_1\mathbf{e}_1 + a_2\mathbf{e}_2) \\ V_t(\mathbf{x} + \mathbf{e}_2) &= \bar{c}(\mathbf{x} + \mathbf{e}_2) + \sum_{a_1=0}^{\infty} \sum_{a_2=0}^{\infty} p_1(a_1)p_2(a_2) \\ & \quad \times V_{t+1}(\mathbf{x} + a_1\mathbf{e}_1 + a_2\mathbf{e}_2). \end{aligned}$$

Since $\bar{c}(\mathbf{x} + \mathbf{e}_1) = \bar{c}(\mathbf{x} + \mathbf{e}_2)$ due to the equal cost assumption, we have $V_t(\mathbf{x} + \mathbf{e}_1) = V_t(\mathbf{x} + \mathbf{e}_2)$, completing the induction. ■

It is also easy to see in this case that if one of the queues is empty and the other is nonempty, then it is optimal to allocate the slot to the nonempty queue. Next, we examine the optimal allocation when both queues are empty.

Definition 6: Let p_1, p_2 denote two probability measures on \mathbb{Z}_+ (We denote by \mathcal{P} the set of all probability measures on \mathbb{Z}_+). We say p_1 is stochastically greater than p_2 (in symbols $p_1 \succ p_2$) if for all elements in \mathbb{Z}_+ , if $q_{p_1}(x) \geq q_{p_2}(x)$, where $q_{p_i}(x) = \sum_{y \geq x} p_i(y)$.

The next theorem shows that whenever both queues have zero deterministic part, it is optimal to allocate the slot to the user whose arrival process is stochastically dominant.

Theorem 7: Suppose the initial state is $\mathbf{x}_1 = \mathbf{0}$. Let $p_i(a)$ denote the probability that there will be a arrivals in queue i , $i = 1, 2$ during a time frame. If $p_i \succ p_j$, then it is optimal to allocate the slot to user i .

Proof: Suppose $p_1 \succ p_2$. We show that it is optimal to allocate the packet to queue 1. Note that it is optimal to allocate the slot at time $t = 1$ to the first queue if

$$\sum_{a_1=0}^{\infty} \sum_{a_2=0}^{\infty} p_1(a_1)p_2(a_2) [V_{t+1}(a_1\mathbf{e}_1 + [a_2 - 1]^+\mathbf{e}_2) - V_{t+1}([a_1 - 1]^+\mathbf{e}_1 + a_2\mathbf{e}_2)] \geq 0. \quad (28)$$

By separating the sums conditioning on a_1, a_2 and using Lemma 9, we get

$$\begin{aligned} & \sum_{a_1=0}^{\infty} \sum_{a_2=0}^{\infty} p_1(a_1)p_2(a_2) \\ & \times [V_{t+1}(a_1\mathbf{e}_1 + [a_2 - 1]^+\mathbf{e}_2) - V_{t+1}([a_1 - 1]^+\mathbf{e}_1 + a_2\mathbf{e}_2)] \\ & = p_2(0) \sum_{a_1=1}^{\infty} p_1(a_1) [V_{t+1}(a_1\mathbf{e}_1) - V_{t+1}([a_1 - 1]^+\mathbf{e}_1)] \\ & \quad - p_1(0) \sum_{a_2=1}^{\infty} p_2(a_2) [V_{t+1}(a_2\mathbf{e}_2) - V_{t+1}([a_2 - 1]^+\mathbf{e}_2)] \\ & \quad + \sum_{a_1=1}^{\infty} \sum_{a_2=1}^{\infty} p_1(a_1)p_2(a_2) \\ & \quad \times [V_{t+1}(a_1\mathbf{e}_1 + (a_2 - 1)\mathbf{e}_2) - V_{t+1}((a_1 - 1)\mathbf{e}_1 + a_2\mathbf{e}_2)] \\ & = \sum_{a=1}^{\infty} (p_2(0)p_1(a) - p_1(0)p_2(a)) V_{t+1}(a\mathbf{e}_1) \\ & \quad - \sum_{a=1}^{\infty} (p_2(0)p_1(a) - p_1(0)p_2(a)) \\ & \quad \times V_{t+1}((a - 1)\mathbf{e}_1) + 0 \end{aligned} \quad (29)$$

where the second equality is due to Lemma 9 and uses the relation $V_{t+1}(a\mathbf{e}_1) = V_{t+1}(a\mathbf{e}_2)$, which can be shown using Lemma 9 and a simple induction. By the monotonicity and convexity of V_{t+1} , the expression in (29) is greater than zero if for any $a' > 0$, we have

$$\begin{aligned} & \sum_{a=a'}^{\infty} (p_2(0)p_1(a) - p_1(0)p_2(a)) \geq 0 \\ & \iff p_2(0) \sum_{a=a'}^{\infty} p_1(a) \geq p_1(0) \sum_{a=a'}^{\infty} p_2(a) \end{aligned}$$

which is satisfied whenever $p_1 \succ p_2$. ■

The above result says that when both queues are empty (the deterministic part), we should allocate a slot to the queue with a stochastically dominating arrival process. However, it is not always possible to compare arrival processes using stochastic dominance. It is, therefore, tempting to see if one could find policies that does not require such comparison, particularly for the case of costs being linear and equal. For instance, would it be optimal if we allocate the slot to the user with a smaller chance of being empty when the deterministic part of both queues are zero? The following example shows that this is not necessarily the case.

Example 2: Suppose we have two queues that are empty (the deterministic part), and we want to allocate one slot to one of them. Suppose the first queue has one arrival with probability 0.6 and no arrivals with probability 0.4; the second queue has $K > 1$ arrivals with probability 0.1 and no arrivals with probability 0.9. If we allocate to the queue with a smaller chance of being empty, we should allocate the time slot to queue 1 since it has a smaller probability of having no arrivals in the next time slot. Let π_i be the policy that allocates the next slot t to queue i . Let $V_t^{\pi_i}(\mathbf{x})$ be the expected cost after one step given that the number of packets in the queues is \mathbf{x} . For π_1 , we have

$$\begin{aligned} V_t^{\pi_1}(\mathbf{0}) & = 0.1 \times 0.6V_{t+1}(0, K) + 0.1 \times 0.4V_{t+1}(0, K) \\ & \quad + 0.9 \times 0.6V_{t+1}(0, 0) + 0.9 \times 0.4V_{t+1}(0, 0). \end{aligned}$$

For π_2 , we have

$$\begin{aligned} V_t^{\pi_2}(\mathbf{0}) & = 0.1 \times 0.6V_{t+1}(1, K - 1) + 0.1 \times 0.4V_{t+1}(0, K - 1) \\ & \quad + 0.9 \times 0.6V_{t+1}(1, 0) + 0.9 \times 0.4V_{t+1}(0, 0). \end{aligned}$$

Note that $V_{t+1}(1, K - 1) = V_{t+1}(0, K)$ and that $V_{t+1}(0, K) - V_{t+1}(0, K - 1)$ can be arbitrarily larger than $V_{t+1}(1, 0) - V_{t+1}(0, 0)$ by making K sufficiently large ($V_{t+1}(\mathbf{x})$ is a convex function of \mathbf{x}). Therefore, it is possible to have $V_t^{\pi_2} < V_t^{\pi_1}$, i.e., the optimal policy may be to allocate the next time slot to queue 2 (which has a higher chance of having no arrivals). This example clearly shows that it is not always optimal to allocate to the queue with a smaller probability of being empty.

Similarly, it can be shown that a policy that allocates the time slot to the queue with higher expected number of arrivals is also not necessarily optimal.

The above discussion illustrates some of the difficulty in trying to obtain conditions weaker than the one given in Theorem 7.

Using the result from Section IV, it can be seen that for the case of multiple slot allocation (when the deterministic part of both queues is zero), the following algorithm finds the optimal policy if the sufficient condition of Theorem 7 is satisfied in each step.

$m = 0$

(*) If $p_i \prec p_j$ allocate the $(m + 1)$ -th slot to queue j .

$w_j = w_j + 1$; For $i = 1, 2$, let $p_i(a) \rightarrow p_i(a + w_i)$; $m = m + 1$.

If $m < M$ go to (*); otherwise stop.

Putting the above results together, we see that an optimal policy for this linear equal cost scenario allocates every slot to a nonempty queue if it exists, and otherwise, allocates it to a queue with stochastically dominant arrival process (updated as shown above). This policy further reduces to, in the case of identical arrival processes, one that allocates slots in a max-min fair fashion among queues when they are all empty [12]. Interestingly, it was also shown in [12] that in this special case (equal cost, identical arrival) the optimality of this policy holds for any

number of queues ($N \geq 2$). Thus, this special case is an example where the main results derived in this paper extend to more than two queues.

VII. CONCLUSION

In this paper, we studied the problem of optimal bandwidth allocation to two users with delayed information on queue backlog and derived fundamental properties of the optimal policy. We proved that when the cost function satisfies certain conditions the optimal single-slot assignment is of the threshold type, and that optimal multiple slot assignment can be obtained by repeatedly using optimal single-slot assignment. We also provided sufficient conditions under which the same properties hold over an infinite horizon, for both the discounted cost and the average cost. We then applied the results to the case of linear and equal holding cost and proved that when both queues have zero deterministic parts, it is optimal to serve the queue with stochastically dominating arrival process. We also presented an example where these results extend to more than two queues.

Unfortunately, other than the example we presented, we have not been able to obtain general results on extending this work to more than two queues. This generalization appears highly nontrivial and higher dimensional versions of the conditions given in Section III may be required. Note that the conditions we used in this paper (monotonicity, supermodularity, and superconvexity) are in essence the equivalent of the continuous convexity conditions in two dimensions for the discrete-time case. In order to extend our results to more than two queues, we would require similar conditions for higher dimensions. Although there has been some work done in this regard (see, for example, [20]), we have not been able to identify a set of convexity conditions that can be shown to propagate over time.

APPENDIX

In this appendix, we present the proof of Theorem 5. A few lemmas are needed to prove the Theorem.

Lemma A-1: $W_\beta(\mathbf{x})$ is nondecreasing in \mathbf{x} . Moreover, under Assumption AVG-1, we have

$$W_\beta(\mathbf{x}) - W_\beta(\mathbf{0}) \leq U(\mathbf{x}). \quad (\text{A-1})$$

Proof: Fix β . We use induction on t to show that $W_t(\mathbf{x})$ as defined in (20) is nondecreasing for all t . First, note that this is true for $t = 1$, since $c(\mathbf{x})$ is nondecreasing. Assuming it holds for t , we want to show that it holds for $t + 1$. Note that we have

$$\begin{aligned} W_{t+1}(\mathbf{x}) &= \bar{c}(\mathbf{x}) \\ &+ \beta \min_{\mathbf{w}: w_1 + w_2 = M} E_{\mathbf{a}} [W_t([\mathbf{x} + \mathbf{a} - \mathbf{w}]^+)], \\ W_{t+1}(\mathbf{x} + \mathbf{e}_i) &= \bar{c}(\mathbf{x} + \mathbf{e}_i) \\ &+ \beta \min_{\mathbf{w}: w_1 + w_2 = M} E_{\mathbf{a}} [W_t([\mathbf{x} + \mathbf{e}_i + \mathbf{a} - \mathbf{w}]^+)]. \end{aligned}$$

The result for $t + 1$ follows from the nondecreasing property of $c(\cdot)$ and $W_t(\cdot)$, using the induction hypothesis

$$W_t(\mathbf{x}) \leq W_t(\mathbf{x} + \mathbf{e}_i). \quad (\text{A-2})$$

Taking the limit on both sides of (A-2) and using (23), we get $W_\beta(\mathbf{x}) \leq W_\beta(\mathbf{x} + \mathbf{e}_i)$, thus $W_\beta(\mathbf{x})$ is nondecreasing in \mathbf{x} .

To show that (A-1) holds, consider the policy π^* that follows policy $\pi_{\mathbf{x}}$ until the first time state $\mathbf{0}$ is reached, and then follows the optimal policy. Therefore, we have

$$W_\beta(\mathbf{x}) \leq W_\beta^{\pi^*}(\mathbf{x}) = U(\mathbf{x}) + W_\beta(\mathbf{0})$$

thus proving the lemma. \blacksquare

Lemma A-2: Suppose $c(\mathbf{x}) \geq 0$ for all $\mathbf{x} \geq \mathbf{0}$. Then, under Assumption AVG-1 the quantity $(1 - \beta)W_\beta(\mathbf{0})$ is bounded for $\beta \in (0, 1)$.

Proof: Note that when $c(\mathbf{x}) \geq 0$, Assumption AVG-1 implies that $E^{\pi_0}[c(\mathbf{x}_t)|\mathbf{x}_1 = \mathbf{0}] \leq U(\mathbf{0})$ for all $t \geq 0$. This can be argued as follows. Under policy π_0 , state $\mathbf{0}$ is a recurrent state, and thus any state at time t lies in between two consecutive occurrences of state $\mathbf{0}$. Since the expected sum of all costs in between those two occurrences is less than or equal to $U(\mathbf{0})$ and all costs are non-negative, the cost at each time step has to be less than or equal to $U(\mathbf{0})$. Thus, we have

$$\begin{aligned} (1 - \beta)W_\beta(\mathbf{0}) &\leq (1 - \beta)W_\beta^{\pi_0}(\mathbf{0}) \\ &= (1 - \beta)E^{\pi_0} \left[\lim_{t \rightarrow \infty} \sum_{u=1}^t \beta^{u-1} c(\mathbf{x}_u) | \mathbf{x}_1 = \mathbf{0} \right] \\ &= (1 - \beta) \lim_{t \rightarrow \infty} \sum_{u=1}^t \beta^{u-1} E^{\pi_0} [c(\mathbf{x}_u) | \mathbf{x}_1 = \mathbf{0}] \\ &\leq (1 - \beta) \lim_{t \rightarrow \infty} \sum_{u=1}^t \beta^{u-1} \cdot U(\mathbf{0}) = U(\mathbf{0}) \end{aligned}$$

where the first inequality is due to the fact that π_0 is not necessarily the optimal policy. The exchange of the limit and expectation is a result of the assumption that $c(\mathbf{x}) \geq 0$ (and consequently, the fact that the sum inside the expectation is nondecreasing) and the last inequality holds by Assumption AVG-1. \blacksquare

Lemma A-3: Let β_n be a sequence of real numbers such that $\beta_n \rightarrow 1^-$ as $n \rightarrow \infty$. If Assumption AVG-1 holds, then there exists a subsequence α_n such that

$$\lim_{n \rightarrow \infty} (W_{\alpha_n}(\mathbf{x}) - W_{\alpha_n}(\mathbf{0})) = h(\mathbf{x})$$

where $0 \leq h(\mathbf{x}) \leq U(\mathbf{x})$ for all $\mathbf{x} > \mathbf{0}$.

Proof: Note that $h_{\beta_n} = W_{\beta_n}(\mathbf{x}) - W_{\beta_n}(\mathbf{0}) \leq U(\mathbf{x})$ by Lemma A-1. The sequence h_{β_n} can be considered as a point in the product topology $\prod_{n=1}^{\infty} [0, U(\mathbf{x})]$, which is a compact space by Tychonoff theorem [21]. Therefore, there exists a subsequence α_n for which $h_{\alpha_n}(\mathbf{x})$ converges. Let $h(\mathbf{x})$ be the limit point of $h_{\alpha_n}(\mathbf{x})$. Since $0 \leq h_{\alpha_n}(\mathbf{x}) \leq U(\mathbf{x})$ for all n , we have $0 \leq h(\mathbf{x}) \leq U(\mathbf{x})$. \blacksquare

Proof of Theorem 5: Take (25), subtract $\beta W_\beta(\mathbf{0})$ from both sides, and add and subtract $W_\beta(\mathbf{0})$ from the left-hand side. We get

$$\begin{aligned} (1 - \beta)W_\beta(\mathbf{0}) + (W_\beta(\mathbf{x}) - W_\beta(\mathbf{0})) &= \bar{c}(\mathbf{x}) \\ &+ \beta \min_{\mathbf{w}: w_1 + w_2 = M} E_{\mathbf{a}} [W_\beta([\mathbf{x} + \mathbf{a} - \mathbf{w}]^+) - W_\beta(\mathbf{0})]. \quad (\text{A-3}) \end{aligned}$$

Let β_n be a sequence of real numbers such that $\beta_n \rightarrow 1^-$ as $n \rightarrow \infty$ and let α_n be a subsequence, as defined in Lemma A-3. We have $\alpha_n \rightarrow 1^-$. Since the quantity $(1 - \alpha_n)W_{\alpha_n}(\mathbf{0})$ is bounded by Lemma A-2, there exists a subsequence γ_n such that $\lim_{n \rightarrow \infty} (1 - \gamma_n)W_{\gamma_n}(\mathbf{0})$ exists and is finite. Let this value be J^* .

Replace β with γ_n in (A-3) and take the limit infimum on both sides. Using Fatou's Lemma [19], we obtain

$$J^* + h(\mathbf{x}) \geq \bar{c}(\mathbf{x}) + \min_{\mathbf{w}: w_1 + w_2 = M} E_{\mathbf{a}} [h(\mathbf{x} + \mathbf{a} - \mathbf{w}^+)]. \quad (\text{A-4})$$

Now, assume that policy π^* minimizes the right-hand side of (26). First, we show that $\bar{J}^{\pi^*} \leq J^*$. Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{t+1}$ be the (random) states that are visited at times 1, 2, $\dots, t+1$, then using (A-4), we have (note that $E_{\mathbf{a}}[\mathbf{x}_t + \mathbf{a} - \mathbf{w}^+]$ is nothing but $E[\mathbf{x}_{t+1}|\mathbf{x}_t]$)

$$\begin{aligned} J^* + h(\mathbf{x}_1) &\geq \bar{c}(\mathbf{x}_1) + E[h(\mathbf{x}_2)|\mathbf{x}_1] \\ J^* + h(\mathbf{x}_2) &\geq \bar{c}(\mathbf{x}_2) + E[h(\mathbf{x}_3)|\mathbf{x}_2] \\ &\dots \\ J^* + h(\mathbf{x}_t) &\geq \bar{c}(\mathbf{x}_t) + E[h(\mathbf{x}_{t+1})|\mathbf{x}_t]. \end{aligned}$$

Taking the expected value on both sides, adding the equations and dividing by t , we get

$$\begin{aligned} \frac{1}{t} \sum_{u=1}^t E[c(\mathbf{x}_u)] &\leq J^* + \frac{E[h(\mathbf{x}_1) - h(\mathbf{x}_{t+1})]}{t} \\ &\leq J^* + \frac{E[h(\mathbf{x}_1)]}{t} \end{aligned} \quad (\text{A-5})$$

where the second inequality is due to the fact that $E[h(\mathbf{x}_{t+1})] \geq 0$. Taking the limit on both sides of (A-5) as $t \rightarrow \infty$ and using the fact that $h(\mathbf{x}) \leq U(\mathbf{x})$, we have $\bar{J}^{\pi^*} \leq J^*$.

Now, consider any other policy π' . We have (see [22])

$$\begin{aligned} \bar{J}^{\pi^*} \leq J^* &\leq \limsup_{\beta \rightarrow 1^-} (1 - \beta)W_{\beta}(\mathbf{x}) \\ &\leq \limsup_{\beta \rightarrow 1^-} (1 - \beta)W_{\beta}^{\pi'}(\mathbf{x}) \leq \bar{J}^{\pi'}. \end{aligned}$$

Therefore, π is the optimal average cost policy. On the other hand, if we let $\pi' = \pi^*$, then we can see that J^* is the optimal average cost, thus proving Theorem 5.

Note 1: The major step in extending the results from the discounted infinite horizon case to the average cost problem is Theorem 5. This step has been justified in the literature in many scenarios. For example, for the case of finite-state space ([23]) or bounded cost functions [16]. For countably infinite-state space and unbounded cost functions, [18] has approached the average cost problem for linear cost functions through a limit of finite horizon problems. Other methods can be found in [24] and [25] that have approached the problem via the limit of discounted cost problems. The method used here is essentially the same as the one used in [19]. The assumptions used in [19] are different than Assumption AVG-1 here. However, we use the lemmas to

show that if Assumption AVG-1 holds, then the three assumptions in [19] will hold, and then use the same argument used there to prove Theorem 5.

REFERENCES

- [1] C. Buyukkoc, P. Varaiya, and J. Warland, "The $c\mu$ -rule revisited," *Adv. Appl. Probab.*, vol. 17, pp. 237–238, 1985.
- [2] J. S. Baras, A. J. Dorsey, and A. M. Makowski, "Two competing queues with linear costs and geometric service requirements: The μc rule is often optimal," *Adv. Appl. Probab.*, vol. 17, pp. 186–209, 1985.
- [3] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 466–478, Mar. 1993.
- [4] L. Tassiulas, "Scheduling and performance limits of networks with constantly changing topology," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 1067–1073, May 1997.
- [5] N. Bambos and G. Michailidis, "On the stationary dynamics of parallel queues with random server connectivities," in *Proc. 43th Conf. Decision Control*, New Orleans, LA, 1995, pp. 3638–3643.
- [6] C. Lott and D. Teneketzis, "On the optimality of an index rule in multi-channel allocation for single-hop mobile networks with multiple service classes," *Probab. Eng. Inf. Sci.*, vol. 14, no. 3, pp. 259–297, Jul. 2000.
- [7] M. J. Neely, E. Modiano, and C. E. Rohrs, "Power allocation and routing in multibeam satellites with time-varying channels," *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 138–152, 2003.
- [8] —, "Dynamic power allocation and routing for time-varying wireless networks," *IEEE J. Sel. Areas Commun. (Special Issue on Wireless Ad Hoc Networks)*, vol. 23, no. 1, pp. 89–103, 2005.
- [9] J. Kuri and A. Kumar, "Optimal control of arrivals to queues with delayed queue length information," *IEEE Trans. Autom. Control*, vol. 40, no. 8, pp. 1444–1450, Aug. 1995.
- [10] F. J. Beutler and D. Teneketzis, "Routing in queueing networks under imperfect information: Stochastic dominance and thresholds," *Stochastics and Stochastic Reports*, vol. 26, pp. 81–100, 1989.
- [11] S. Kittipiyakul and T. Javidi, "A fresh look at optimal subcarrier allocation in OFDMA systems," in *Proc. IEEE Conf. Decision Control*, Dec. 2004.
- [12] N. Ehsan and M. Liu, Optimal bandwidth allocation with delayed state observation and batch assignment, Univ. Michigan, Ann Arbor, EECS, Tech. Rep. CGR 03-11, 2003.
- [13] —, "On the optimality of an index policy for bandwidth allocation with delayed state observation and differentiated services," in *Proc. IEEE INFOCOM*, Hong Kong, Mar. 2004, pp. 1974–1983.
- [14] G. M. Koole, "Structural results for the control of queueing systems using event-based dynamic programming," *Queueing Syst.*, vol. 30, pp. 323–339, 1998.
- [15] E. Altman and G. M. Koole, On submodular value functions of dynamic programming, INRIA Sophia Antipolis, Tech. Rep. 2658, 1995.
- [16] P. R. Kumar and P. Varsaiya, *Stochastic Systems, Estimation, Identification and Adaptive Control*. Englewood Cliffs, NJ: Prentice-Hall, 1986.
- [17] D. P. Bertsekas, *Dynamic Programming, Deterministic and Stochastic Models*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [18] B. Hajek, "Optimal control of two interacting service stations," *IEEE Trans. Autom. Control*, no. AC-29, pp. 491–499, 1984.
- [19] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*, ser. Probability and Statistics. New York: Wiley, 1999.
- [20] K. Murota, *Discrete Convex Analysis*. Philadelphia, PA: SIAM, 2003.
- [21] J. R. Munkres, *Topology*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2000.
- [22] L. Sennott, "A new condition for the existence of optimum stationary policies in average cost markov decision processes-unbounded cost case," in *Proc. 25th Conf. Decision Control*, Athens, Greece, 1986, pp. 1719–1721.
- [23] S. Ross, *Applied Probability Models With Optimization Applications*. San Francisco, CA: Holden-Day, 1970.
- [24] F. Lu and R. F. Serfozo, "M/M/1 queueing decision processes with monotone hysteretic optimal policies," *Oper. Res.*, vol. 32, pp. 1116–1132, 1984.
- [25] R. R. Weber and S. Stidham, "Optimal control of service rates in networks of queues," *Adv. Appl. Probab.*, pp. 202–218, 1987.



Navid Ehsan (S'00–M'05) received the B.Sc. degree from the Sharif University of Technology, Tehran, Iran, in 1998, and the M.Sc. and Ph.D. degrees from the University of Michigan, Ann Arbor, in 2002 and 2005, respectively.

He is currently a Postdoctoral Researcher at the University of California, San Diego. His research interests are in dynamic bandwidth allocation, optimal power allocation and admission control in wireless systems, medium access protocols for ad hoc networks, energy efficient protocols in sensor networks, and stochastic optimization.



Mingyan Liu (M'00) received the B.Sc. degree in electrical engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1995, the M.Sc. degree in systems engineering and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 1997 and 2000, respectively.

She joined the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, in September 2000, where she is currently an Assistant Professor. Her research interests are in performance modeling, analysis, energy-efficiency and resource allocation issues in wireless mobile ad hoc networks, wireless sensor networks, and terrestrial satellite hybrid networks.

Dr. Liu is the recipient of the 2002 National Science Foundation (NSF) CAREER Award, and the University of Michigan Elizabeth C. Crosby Research Award in 2003.