Optimal Server Allocation in Batches

Navid Ehsan, Mingyan Liu
Electrical Engineering and Computer Science Department
University of Michigan, Ann Arbor
{nehsan,mingyan}@eecs.umich.edu

Abstract

In this paper we consider the problem of allocating bandwidth to two transmitters/queues with arbitrary arrival processes, so as to minimize the total expected (discounted) holding cost of backlogged packets in the system over a finite (infinite) horizon. The bandwidth is in the form of time slots in a TDMA schedule. Allocation decisions are made based on the queue backlog information, which is delayed. 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. In this paper we show that if the packet holding cost as a function of the packet backlog in the system is non-decreasing, supermodular and superconvex, then (1) the value function (or the cost to go) at each time 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 typically much easier to obtain. We provide sufficient conditions for the same results to hold in the case of discounted cost over an infinite horizon and in the case of average cost criterion. The above results are then applied to a special case where the holding costs are linear and equal for all queues, in which minimizing holding cost is equivalent to maximizing the system throughput.

Index Terms

Optimal resource allocation, delayed state observation, stochastic processes, batch allocation

I. Introduction

In this paper we study the problem of optimally allocating bandwidth (in the form of time slots in a slotted system, or equivalently *servers*) 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).

In the case of a satellite network, users/terminals transmitting to the satellite are assumed to follow a dynamic TDMA schedule, each assigned a certain number of slots within a *frame* that consists of a fixed number of *slots*. Users also inform the satellite their current backlog, carried in packet headers. The assignment is made based on the backlog information and broadcast to the users over a non-interfering channel. An allocation specifies which slot in the upcoming frame is reserved for/to be used by which user. In such a scenario, 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 delayed and partially obsolete by the time the allocation is used. This results in possible over-allocation or under-allocation. Therefore in this case the allocation needs to take into account unknown random arrivals that occur in between observations/state information updates. In the case of low-power devices similar resource allocation problems arise where users share the channel for transmitting to the common server. In these systems time is divided into active and inactive periods. During the inactive period users turn their transmitters and receivers off in order to conserve energy and turn them back on at the beginning of each active period. At the beginning of an active period, the server sends a beacon containing information about the slot allocation for the current active period (each active period can contain multiple time slots). The users then transmit according to this allocation. The users also send their current backlog information to the server. The server will use this information to make decision for the next active frame. Due to the long inactive period, the backlog information has most likely changed at the time when the allocation is used. Therefore the server has to consider this uncertainty in the backlog due to random arrivals during the inactive period. Although dynamics of both of the above systems are the same and the specific type of the system does not affect our discussion on the optimal policy, in this paper we consider the satellite scenario to model our system.

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 (discounted) packet holding cost or average cost over a finite (infinite) time horizon. While in general reducing holding cost has the effect of reducing packet delay, different forms of the cost function lead to different performance criteria. For example, under a linear cost function equal to all queues (i.e., each packet incurs a constant unit cost) minimizing the cost is equivalent to maximizing system throughput. Different cost functions also lead to different optimal strategies, to be further explored in the 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], [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], [4], [5] considered the server allocation problem to multiple queues with varying connectivity but of the same service class. Each of these studies determined policies that maximize throughput over an infinite horizon. In particular, [3] derived the sufficient condition for stability and showed that serving the Longest Connected Queue (LCQ) policy stabilizes the system if the system is stabilizable. [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 server 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.

The problem studied in this paper (in the case of an infinite horizon) can also be cast as a special case of the *restless bandit* problem [11], [12], [13], [14], where projects undergo state transitions even when they are not played or selected. This is because in our case the backlog of each queue continues to change as packets arrive. [11] and [12] studied the asymptotic behavior of this class of problems when the number of projects (queues in this case) and servers (slots in a frame in this case) go to infinity with a fixed ratio. A general optimal solution is not known for this class of problems. However, an index policy can be defined based on the Whittle's heuristic, which is sub-optimal in the finite (number of servers and projects) case and asymptotically optimal in the infinite case.

In [15], [16] we have studied problems similar to the one presented in this paper, but with simpler, linear cost assumptions. In [15] we derived the optimal policy when users have the same unit holding cost and identical arrival processes, while in [16] 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 ideas similar to that used in [17] and [18], 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* (non-decreasing), *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 non-decreasing, 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 *threshold* type.
2) If the cost function is non-decreasing, 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 first represents a fundamental result on the nature of this problem, and may also help us derive the optimal policy. We will provide examples further illustrating the threshold property. 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 typically 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.

The rest of the paper is organized as follows. In the next section 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 the average cost criterion case. In Section VI we use these results to find the optimal policy for the special case of linear and equal holding cost. In section VII we present some properties of the threshold policy through numerical examples. Section VIII concludes the paper.

II. Problem Formulation

In this section we describe the network model we adopt as an abstraction of the bandwidth allocation problem described in the previous section, and formally present the optimization problem along with a summary of assumptions and notations.

A. Network Model and Notation

Consider N queues that transmit packets to a single receiver and in doing so compete for shares of a common channel that consists of time slots. Packets arrive at queues according to arbitrary random processes. Packets are assumed to be of equal length and 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 for all M slots in a frame (M may or may not be greater than N). In other words, the channel assignment is done in batches of M slots. Under a particular allocation, 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. Different from most of the prior work, here multiple servers can be assigned to a single queue. When this happens, multiple packets are served. For the rest of our discussion, we will adopt the slot allocation model and use the term *server* to mean the controller that makes the allocation decisions.

We consider time evolution in discrete time steps indexed by $t = 0, 1, \cdots T$, 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 based on the backlog information of each queue (number of packets waiting/existing in the queue) provided by queues at the beginning of a frame. We will ignore the transmission time of such information. This 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. Based on this information an allocation decision is made by the server and broadcast to all queues over a non-interfering channel. Due to extensive propagation delay in the system, this broadcast is received by the queues at the end of that frame, in time to be used for the next frame. The same procedure then repeats, resulting in a one-step delay in state observation by the server as shown in Figure 1. Specifically, at time t, each user advertises its buffer size (denoted by $\mathbf{b}_t$) to the server. The server allocates slots to be used for transmission in the next frame $[t+1, t+2)$, denoted by $\mathbf{w}_{t+1}$. However, the server does not know the queue backlog at time $t+1$ due to random arrivals that occurred during $[t, t+1)$. This procedure begins from $t = 0$ and ends at $t = T$ (in the case of finite time horizon). Note that in this scenario during the first frame queues do not have allocated slots and only start transmitting in the second frame (starting $t = 1$).

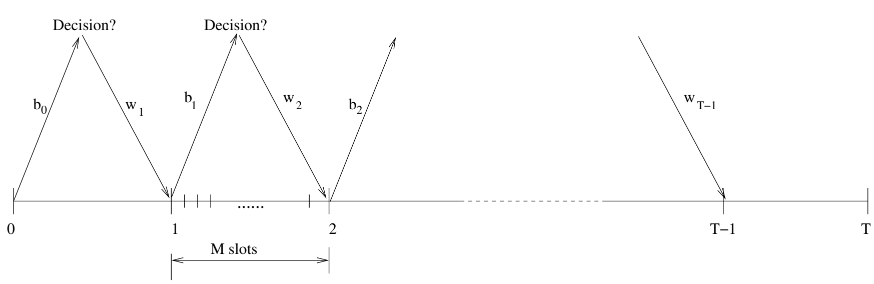

Fig. 1. The bandwidth allocation dynamics

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

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 $(b_{1,t}, b_{2,t}, \cdots, b_{N,t})$. We use the same convention for other quantities as defined below.

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

$\mathbf{a}_t = (a_{1,t}, \cdots, a_{N,t})$: 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 $[t, t+1)$.

$\mathbf{x}_t = [\mathbf{b}_{t-1} - \mathbf{w}_{t-1}]^+$: This is the part of the queue backlog at time t that is precisely known to the server 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 for sure in the queue, excluding the random arrivals that occurred during $[t-1, t)$. It's 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, i.e., a vector with all elements being zero except a one in the i-th position.

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

For a function f defined on $\mathbb{Z}_+^2$, let $\hat{f}$, defined on $\mathbb{Z}^2$, be $\hat{f}(\mathbf{x}) = f(\mathbf{x}^+)$. In general if the domain of a function is $\mathbb{Z}_+^2$ we use f, g, etc., and if the domain is $\mathbb{Z}^2$ we denote the functions by $\hat{f}$, $\hat{g}$, etc. The above 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,

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

where $\mathbf{w}_0 = \mathbf{0}$. For now the packet holding cost $c(\mathbf{b})$ is an arbitrary function. Later, we will restrict c to belong to a certain class of functions.

B. Assumptions

Below 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 the results to more than two users remains an open problem and is out of the scope of this paper. Limited 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 for that frame. This is because the exact arrival times of the packets in a frame is 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.

5) The server recalls the latest allocation it has made. Note that the 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. ($\mathbf{b}_t$ is a Markov chain with state space $\{(b_1, b_2) : b_1, b_2 \in \mathbb{Z}_+\}$ where the transition probabilities depend on the control action $\mathbf{w}_t$ and arrival statistics).

C. Problem Formulation and Preliminaries

Although the state of the system is not perfectly observed, we can extend the state space to convert a Markov chain with imperfect state observation into a Markov chain with perfect state observation [19]. In our problem we could consider $(\mathbf{b}_{t-1}, \mathbf{w}_{t-1})$ to be the state at time t. However, one can see that in our specific problem the states and their transitions only depend on $\mathbf{x}_t = [\mathbf{b}_{t-1} - \mathbf{w}_{t-1}]^+$, which is the deterministic portion of the queue at time t as defined earlier. The actual queue size at time t is $\mathbf{x}_t + \mathbf{a}_{t-1}$.

Using $\mathbf{x}_t$ as the state, this problem can be solved via dynamic programming [20]. Define

$$\bar{c}_t(\mathbf{x}) = E_{\mathbf{a}_{t-1}}[c(\mathbf{x} + \mathbf{a}_{t-1})], \quad (2)$$

$$\text{where} \quad E_{\mathbf{a}_t}[f(\mathbf{a}_t)] = \sum_{\mathbf{a}_t} p_t(\mathbf{a}_t) f(\mathbf{a}_t) \quad (3)$$

for some function f. Then the dynamic program of the problem is as follows.

$$V_T(\mathbf{x}) = \bar{c}_T(\mathbf{x}),$$
$$V_t(\mathbf{x}) = \bar{c}_t(\mathbf{x}) + \min_{\sum_{i=1}^{N} w_{i,t} = M} \{E_{\mathbf{a}_{t-1}}[V_{t+1}([\mathbf{x} + \mathbf{a}_{t-1} - \mathbf{w}_t]^+)]\}, \quad (4)$$

where V_t is the *value function* or the cost to go at time t.

Remark 1: For the rest of the paper, we make the following additional assumption. The joint probability mass function of the arrival processes does not change with time. Thus we have $p_t(\mathbf{a}_t) = p(\mathbf{a}), \ \forall t$. This assumption is only for the simplicity in notation and as will be discussed in Section VIII can be easily relaxed. Note that by this assumption, we have $\bar{c}_t(\mathbf{x}) = \bar{c}(\mathbf{x})$ for all t.

Definition 1: Define $\hat{S}_t(\mathbf{x}) : \mathbb{Z}^2 \to \mathbb{R}$ as follows:

$$\hat{S}_t(\mathbf{x}) = \sum_{\mathbf{a}} p(\mathbf{a}) V_t([\mathbf{x} + \mathbf{a}]^+). \quad (5)$$

Definition 2: For some function $f : \mathbb{Z}^2 \to \mathbb{R}$ or $f : \mathbb{Z}_+^2 \to \mathbb{R}$, define two operators T_1 and T_M to be

$$T_1 f(\mathbf{x}) = \min_{i \in \{1,2\}} \{f(\mathbf{x} - \mathbf{e}_i)\}, \quad (6)$$

$$T_M f(\mathbf{x}) = \min_{\mathbf{w}: w_1 + w_2 = M} \{f(\mathbf{x} - \mathbf{w})\} \ . \quad (7)$$

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. One of the key results to be shown is the conditions under which T_M may be obtained by repeatedly using T_1.

The following lemma immediately follows as a result of the definitions above.

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

In the next two sections we will first study the case of $M = 1$, and then consider $M > 1$.

III. Optimal Policy For a Single Slot Allocation

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

$$V_T(\mathbf{x}) = \bar{c}(\mathbf{x}),$$
$$V_t(\mathbf{x}) = \bar{c}(\mathbf{x}) + T_1\hat{S}_{t+1}(\mathbf{x}), \quad 1 \le t \le T-1, \tag{8}$$

where $\hat{S}_t(\mathbf{x})$ is defined in the previous section.

Definition 3: A function $f : \mathbb{Z}_+^2 \to \mathbb{R}$ belongs to the set $\mathcal{F}$ if $f(\mathbf{x})$ satisfies the following conditions:

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

C.2 $f(\mathbf{x}+\mathbf{e}_1) + f(\mathbf{x}+\mathbf{e}_2) \le f(\mathbf{x}) + f(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2)$;

C.3.a $f(\mathbf{x}+\mathbf{e}_1) + f(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \le f(\mathbf{x}+\mathbf{e}_2) + f(\mathbf{x}+2\mathbf{e}_1)$;

C.3.b $f(\mathbf{x}+\mathbf{e}_2) + f(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \le f(\mathbf{x}+\mathbf{e}_1) + f(\mathbf{x}+2\mathbf{e}_2)$.

C.1 is the *monotonicity* condition and requires the function $f(\mathbf{x})$ to be non-decreasing in both its elements, **C.2** is the *supermodularity* condition, and **C.3** is the *superconvexity* condition following the terminology used in [17]. Note that these are rather benign conditions, and they specify a very large class of cost functions of practical interest.

Remark 2: Note that conditions **C.2** and **C.3.a** result in the convexity of f in x_1. Similarly, **C.2** and **C.3.b** imply the convexity of f in x_2.

Definition 4: Define $\hat{\mathcal{F}}$ to be the set of all functions $\hat{f} : \mathbb{Z}^2 \to \mathbb{R}$ that satisfy conditions **C.1** - **C.3**.

It should be immediately clear that $f \in \mathcal{F} \Rightarrow \hat{f} \in \hat{\mathcal{F}}$.

The main result of this section is the following theorem.

Theorem 1: Suppose there are two users and one slot in each frame to be allocated. If the cost function $c(\cdot) \in \mathcal{F}$, then
(a) for all time t we have $V_t(\mathbf{x}) \in \mathcal{F}$; and
(b) the optimal policy in assigning one slot is of the threshold type.

In the remainder of this section we 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 number of lemmas.

Lemma 2: If $f \in \mathcal{F}$, then the function $\hat{g} : \mathbb{Z}^2 \to \mathbb{R}$ 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 $i = 1, 2$,

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

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

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

$$\hat{g}(\mathbf{x}+\mathbf{e}_1) + \hat{g}(\mathbf{x}+\mathbf{e}_2) \le \hat{g}(\mathbf{x}) + \hat{g}(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \ . \tag{9}$$

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 (9) becomes

$$\begin{aligned} & f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_1) + f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_2) \\ \leq\; & f([\mathbf{x}+\mathbf{a}]^+) + f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_1 + \mathbf{e}_2), \end{aligned}$$

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 (9) 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 (9) becomes

$$f([\mathbf{x}+\mathbf{a}]^+) + f([\mathbf{x}+\mathbf{a}]^+) \leq f([\mathbf{x}+\mathbf{a}]^+) + f([\mathbf{x}+\mathbf{a}]^+),$$

which is trivially true.

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

$$\hat{g}(\mathbf{x}+\mathbf{e}_1) + \hat{g}(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \leq \hat{g}(\mathbf{x}+\mathbf{e}_2) + \hat{g}(\mathbf{x}+2\mathbf{e}_1). \tag{10}$$

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

1) If $y_1, y_2 \geq 0$, then (10) becomes

$$\begin{aligned} & f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_1) + f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_1 + \mathbf{e}_2) \\ \leq\; & f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_2) + f([\mathbf{x}+\mathbf{a}]^+ + 2\mathbf{e}_1), \end{aligned}$$

which is true since f satisfies **C.3**, by replacing $\mathbf{x}$ with $[\mathbf{x}+\mathbf{a}]^+$ in **C.3**.

2) If $y_1 < 0, y_2 \geq 0$, then (10) becomes

$$\begin{aligned} & f([\mathbf{x}+\mathbf{a}]^+) + f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_2) \\ \leq\; & f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_2) + f([\mathbf{x}+\mathbf{a}+2\mathbf{e}_1]^+), \end{aligned}$$

which is true by the monotonicity of f.

3) If $y_2 < 0, y_1 \geq 0$, then (10) becomes

$$\begin{aligned} & f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_1) + f([\mathbf{x}+\mathbf{a}]^+ + \mathbf{e}_1) \\ \leq\; & f([\mathbf{x}+\mathbf{a}]^+) + f([\mathbf{x}+\mathbf{a}]^+ + 2\mathbf{e}_1), \end{aligned}$$

which is true by the convexity of f (combining **C.2** and **C.3.a**).

4) If $y_1, y_2 < 0$, then (10) 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.3.b can be proven in the same way and is thus omitted for brevity.

Therefore we conclude $\hat{g}(\mathbf{x}) \in \hat{\mathcal{F}}$. ■

Lemma 3: If $f_1, f_2, \cdots$ 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

$$\begin{aligned}
& g(\mathbf{x}+\mathbf{e}_1)+g(\mathbf{x}+\mathbf{e}_2) \\
= & \sum_l p_l \cdot (f_l(\mathbf{x}+\mathbf{e}_1)+f_l(\mathbf{x}+\mathbf{e}_2)) \\
\leq & \sum_l p_l \cdot (f_l(\mathbf{x})+f_l(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2)) \\
= & g(\mathbf{x})+g(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \ , \qquad (11)
\end{aligned}$$

where the inequality is due to the supermodularity of f_l.

(iii) *Superconvexity*: This holds because

$$\begin{aligned}
& g(\mathbf{x}+\mathbf{e}_1)+g(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \\
= & \sum_l p_l \cdot (f_l(\mathbf{x}+\mathbf{e}_1)+f_l(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2)) \\
\leq & \sum_l p_l \cdot (f_l(\mathbf{x}+\mathbf{e}_2)+f_l(\mathbf{x}+2\mathbf{e}_1)) \\
= & g(\mathbf{x}+\mathbf{e}_2)+g(\mathbf{x}+2\mathbf{e}_1) \ , \qquad (12)
\end{aligned}$$

where the inequality is due to the superconvexity of f_l.

$\mathbf{C.3.b}$ can be shown in the same way and is thus omitted for brevity. ■

Lemma 4: If $\hat{f}_1, \hat{f}_2, \cdots$ 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 of this lemma is the same as that of 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

$$\hat{g}(\mathbf{x}) = T_1\hat{f}(\mathbf{x}) = \min\{\hat{f}(\mathbf{x}-\mathbf{e}_1), \hat{f}(\mathbf{x}-\mathbf{e}_2)\}. \qquad (13)$$

(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

$$\hat{g}(\mathbf{x}+\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_2) \leq \hat{g}(\mathbf{x})+\hat{g}(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2). \qquad (14)$$

We will consider different cases depending on the minimizers of $\hat{g}(\mathbf{x})$ and $\hat{g}(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2)$ on the right hand side of (13), 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

$$\hat{g}(\mathbf{x}+\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_2) \leq \hat{f}(\mathbf{x}-\mathbf{e}_1)+\hat{f}(\mathbf{x}+\mathbf{e}_2). \qquad (15)$$

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 $\hat{g}(\mathbf{x}+\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_2) \le \hat{f}(\mathbf{x})+\hat{f}(\mathbf{x}+\mathbf{e}_2-\mathbf{e}_1)$.
Letting $\mathbf{y}=\mathbf{x}-\mathbf{e}_1$, the above becomes

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

where the second inequality is true by the supermodularity of $\hat{f}$, thus proving (15).

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

$$\hat{g}(\mathbf{x}+\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_2) \le \hat{f}(\mathbf{x}-\mathbf{e}_1)+\hat{f}(\mathbf{x}+\mathbf{e}_1)\ . \tag{16}$$

To show this, consider

$$\begin{aligned} & \hat{g}(\mathbf{x}+\mathbf{e}_1) = \min\{\hat{f}(\mathbf{x}), \hat{f}(\mathbf{x}+\mathbf{e}_1-\mathbf{e}_2)\} \le \hat{f}(\mathbf{x}), \\ & \hat{g}(\mathbf{x}+\mathbf{e}_2) = \min\{\hat{f}(\mathbf{x}+\mathbf{e}_2-\mathbf{e}_1), \hat{f}(\mathbf{x})\} \le \hat{f}(\mathbf{x}), \\ \Rightarrow\ & \hat{g}(\mathbf{x}+\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_2) \le 2\hat{f}(\mathbf{x}) \\ \le\ & \hat{f}(\mathbf{x}-\mathbf{e}_1)+\hat{f}(\mathbf{x}+\mathbf{e}_1), \end{aligned} \tag{17}$$

where the last inequality is due to the convexity of $\hat{f}$, thus proving (16).

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

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

$$\hat{g}(\mathbf{x}+\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \le \hat{g}(\mathbf{x}+2\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_2)\ . \tag{18}$$

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

1) $m_1=m_2=1$: In this case (18) becomes

$$\hat{g}(\mathbf{x}+\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \le \hat{f}(\mathbf{x}+\mathbf{e}_1)+\hat{f}(\mathbf{x}+\mathbf{e}_2-\mathbf{e}_1)$$

To show this we have

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

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

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

where the second inequality is due to the superconvexity of $\hat{f}$, thus proving (18).

2) $m_1=1, m_2=2$: In this case (18) becomes

$$\hat{g}(\mathbf{x}+\mathbf{e}_1)+\hat{g}(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \le \hat{f}(\mathbf{x}+\mathbf{e}_1)+\hat{f}(\mathbf{x})\ . \tag{19}$$

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 (19).

3) $m_1 = 2, m_2 = 1$: By superconvexity of 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.3.b** and the other two inequalities are a consequence of **C.3.a**. Combining (adding) these inequalities we get

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

However, note that whenever $m_1 = 2$, the right hand side of the above equation is non-positive, thus the left hand side is also non-positive. This implies that $m_2 = 2$ (i.e., $m_1 = 2, m_2 = 1 \Rightarrow m_1 = 2, m_2 = 2$, 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 (included in the case of) $m_1 = 2, m_2 = 2$, which is dealt with next.

4) $m_1 = 2, m_2 = 2$: In this case (18) 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

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

thus proving (18). That $\hat{g}$ also satisfies **C.3.b** can be shown in a similar way and is thus not repeated here.

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. ∎

The following lemma is also stated in [21].

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:

(a) 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.

(b) By part (a) of this theorem, $V_{t+1} \in \mathcal{F}$ for all t. Therefore $\hat{S}_{t+1} \in \mathcal{F}$. Thus by property **C.3.a** we have

$$\begin{aligned} & \hat{S}_{t+1}(\mathbf{x}+\mathbf{e}_1) + \hat{S}_{t+1}(\mathbf{x}+\mathbf{e}_1+\mathbf{e}_2) \\ \leq \quad & \hat{S}_{t+1}(\mathbf{x}+2\mathbf{e}_1) + \hat{S}_{t+1}(\mathbf{x}+\mathbf{e}_2) \ . \end{aligned}$$

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 $\mathbf{C.3.b}$ 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 define the threshold as following.

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

$h_t(x_1) = \infty$ 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 the optimal decision rule is to assign the slot to queue 1 (if the set is empty then the threshold is infinity), proving the optimality of a threshold policy. ∎

IV. Multiple Slot Batch Allocation

In this part 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 $c(\mathbf{b}_t) = b_{1,t}^2 \cdot b_{2,t}$. All cost quantities are in some unspecified unit. 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 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 $x_{1,1}^* = 0, x_{2,1}^* = 2$, resulting in a cost of zero at $t = 2$.

Now consider the sequential allocation, which proceeds as follows. Suppose we only have 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 $d_{1,t} = d_{2,t} = 2$. For the allocation of the second slot, again suppose we only have 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. These two sequential steps result in both slots being allocated to queue 1 and none to queue 2, with a cost of 3 at $t = 2$. Obviously this policy is not optimal.

In this section we show that under certain conditions on the cost function the optimal policy can be achieved by sequentially allocating the slots according to the optimal policy for a single slot allocation. It turns out that the required conditions for this property are the same conditions as we defined for the functions to belong to $\mathcal{F}$.

Definition 5: 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 number of slots. Note that 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 $\mathbf{w}^1$

$$\mathbf{w}^1 = \text{argmin}_{\mathbf{w}: w_1 + w_2 = m} \{\hat{f}(\mathbf{x} - \mathbf{w})\} \ . \tag{20}$$

Suppose we have $m + 1$ slots to assign. By definition we have

$$T_{m+1} \hat{f}(\mathbf{x}) = \min_{\mathbf{w}: w_1 + w_2 = m+1} \{\hat{f}(\mathbf{x} - \mathbf{w})\} \ . \tag{21}$$

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 (21), i.e. we want to show the following for $i \neq j$:

$$\begin{aligned} & \hat{f}(\mathbf{x} - (w_i^1 + k + 1)\mathbf{e}_i - (w_j^1 - k)\mathbf{e}_j) \\ \geq \ & \hat{f}(\mathbf{x} - (w_i^1 + 1)\mathbf{e}_i - w_j^1 \mathbf{e}_j) \ . \end{aligned} \tag{22}$$

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 (22) then we will have established that $\mathbf{w}^1 + \mathbf{e}_i$ minimizes the right hand side of (21).

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 (21), then $\mathbf{w}^1 + \mathbf{e}_i$ will also minimize the right hand side of (21). Therefore, assume that $\mathbf{w}^1 + (k + 1)\mathbf{e}_i - k\mathbf{e}_j$ minimizes the right hand side of (21) and let $\mathbf{w}^2 = \mathbf{w}^1 + \mathbf{e}_i$.

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

$$\begin{aligned} & \hat{f}(\mathbf{x} - (w_i^1 + k)\mathbf{e}_i - (w_j^1 - k)\mathbf{e}_j) \\ & \quad - \hat{f}(\mathbf{x} - w_i^1 \mathbf{e}_i - w_j^1 \mathbf{e}_j) \\ \leq \ & \hat{f}(\mathbf{x} - (w_i^1 + k + 1)\mathbf{e}_i - (w_j^1 - k)\mathbf{e}_j) \\ & \quad - \hat{f}(\mathbf{x} - (w_i^1 + 1)\mathbf{e}_i - w_j^1 \mathbf{e}_j) \ . \end{aligned} \tag{23}$$

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

$$\begin{aligned} & \hat{f}(\mathbf{x}-(w_i^1+1)\mathbf{e}_i-(w_j^1-1)\mathbf{e}_j) \\ & \quad -\hat{f}(\mathbf{x}-w_i^1\mathbf{e}_i-w_j^1\mathbf{e}_j) \\ \leq\ & \hat{f}(\mathbf{x}-(w_i^1+2)\mathbf{e}_i-(w_j-1)\mathbf{e}_j) \\ & \quad -\hat{f}(\mathbf{x}-(w_i^1+1)\mathbf{e}_i-w_j^1\mathbf{e}_j)\ . \end{aligned} \tag{24}$$

(24) 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.3.a** if $i=1$ and use **C.3.b** if $i=2$). Thus the induction basis is established.

Now assume (23) 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.3.a** if $i=1$ and use **C.3.b** 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

$$\begin{aligned} & \hat{f}(\mathbf{x}-(w_i^1+l+1)\mathbf{e}_i-(w_j^1-l-1)\mathbf{e}_j) \\ & \quad -\hat{f}(\mathbf{x}-(w_i^1+l)\mathbf{e}_i-(w_j^1-l)\mathbf{e}_j) \\ \leq\ & \hat{f}(\mathbf{x}-(w_i^1+l+2)\mathbf{e}_i-(w_j-l-1)\mathbf{e}_j) \\ & \quad -\hat{f}(\mathbf{x}-(w_i^1+l+1)\mathbf{e}_i-(w_j^1-l)\mathbf{e}_j)\ . \end{aligned} \tag{25}$$

Combining the induction hypothesis and (25) gives the result for case of $k=l+1$ and the induction is complete.

Next note that the following inequality holds due to the optimality of $\mathbf{w}^1$ when there are m slots to allocate, for $1\leq k\leq w_j^1$.

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

Therefore the left hand side of (23) is always greater than or equal to zero. Thus the right hand side will also be greater than or equal to zero, i.e.,

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

This means that $\mathbf{w}^2$ minimizes the right hand side of equation (21).

The above result shows that the minimizer on the right hand side of (21) 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 allocation slots, 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 $T_{m+1}\hat{f}(\mathbf{x})=T_1T^m\hat{f}(\mathbf{x})$, Using the induction hypothesis. Thus we have $T_{m+1}\hat{f}(\mathbf{x})=T_1^{m+1}\hat{f}(\mathbf{x})$, which completes the proof. ■

Consider two users and M allocation slots in each time frame. Also 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 two users and 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 $\mathbf{x} \in \mathbb{Z}_+^2$

$$\begin{aligned} V_{t-1}(\mathbf{x}) &= \bar{c}(\mathbf{x}) + T_M \hat{S}_t(\mathbf{x}) \\ &= \bar{c}(\mathbf{x}) + T_1^M(\hat{S}_t(\mathbf{x}))\ . \end{aligned} \tag{26}$$

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 (26) 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 $\mathbf{w}$. That is, the problem can be solved as follows: allocate slots sequentially by assigning the $(m+1)$-th slot optimally given the state of the system and previous allocation in the same frame. ■

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 $T \to \infty$. Note that the cost defined in (1) is infinite as $T \to \infty$, 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 β ($0 < \beta < 1$), and define the t step minimum cost function

$$W_t(\mathbf{x}) = \min_\pi E^\pi\{\sum_{u=1}^{t} \beta^{u-1} \bar{c}(\mathbf{x}_u) | \mathbf{x}_1 = \mathbf{x}\}\ . \tag{27}$$

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:

$$\begin{aligned} W_0(\mathbf{x}) &= 0; \\ W_t(\mathbf{x}) &= \bar{c}(\mathbf{x}) + \beta \min_{\mathbf{w}: w_1+w_2=M} E_{\mathbf{a}}[W_{t-1}([\mathbf{x} + \mathbf{a} - \mathbf{w}]^+)]\ . \end{aligned} \tag{28}$$

Definition 6: Define $\hat{R}(\mathbf{x}) : \mathbb{Z}^2 \to \mathbb{R} \cup \{\infty\}$ as follows:

$$\hat{R}_t(\mathbf{x}) = \sum_{\mathbf{a}} p(\mathbf{a}) W_t([\mathbf{x}+\mathbf{a}]^+). \tag{29}$$

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 theorem 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:

$$W_\infty(\mathbf{x}) = \min_\pi E^\pi \{ \lim_{t\to\infty} \sum_{u=1}^{t} \beta^{u-1} \bar{c}(\mathbf{x}_u) | \mathbf{x}_1 = \mathbf{x} \} \ . \tag{30}$$

Note that $c(\mathbf{x})$ is not necessarily bounded. However, if we have $c(\mathbf{x}) \geq 0$ for all $x \geq 0$, then $W_\infty(\mathbf{x})$ satisfies the following (for more details and proof see [22], Chapter 5.4).

$$\begin{aligned} W_\infty(\mathbf{x}) &= \bar{c}(\mathbf{x}) + \beta \min_{\mathbf{w}: w_1+w_2=M} E_{\mathbf{a}}[W_\infty([\mathbf{x}+\mathbf{a}-\mathbf{w}]^+)] \\ W_\infty(\mathbf{x}) &= \lim_{t\to\infty} W_t(\mathbf{x}) \ . \end{aligned} \tag{31}$$

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 point-wise limit of functions, i.e. if $f_1, f_2, \cdots$ is a sequence of functions and $f_i \in \mathcal{F}$, $\forall i$, and if $f = \lim_{n\to\infty} f_n$, then $f \in \mathcal{F}$. Therefore by using Lemma 8 and Equation (31) 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:

$$\bar{J}^\pi = E^\pi[\bar{C}|\mathbf{b}_0, \mathbf{w}_0], \quad \bar{C} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} c(\mathbf{b}_t), \tag{32}$$

Recall the infinite horizon discounted cost defined before:

$$W_\beta(\mathbf{x}) = \min_\pi E^\pi \{ \lim_{t\to\infty} \sum_{u=1}^{t} \beta^{u-1} \bar{c}(\mathbf{x}_u) | \mathbf{x}_1 = \mathbf{x} \} \ . \tag{33}$$

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 section 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 (31).

$$W_\beta(\mathbf{x}) = \bar{c}(\mathbf{x}) + \beta \min_{\mathbf{w}: w_1+w_2=M} E_{\mathbf{a}}[W_\beta([\mathbf{x}+\mathbf{a}-\mathbf{w}]^+)] \ . \tag{34}$$

Consider the following assumption:

Assumption 1: For any state $\mathbf{x} > 0$ 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 (non-discounted) cost for this transition be denoted by $U(\mathbf{x})$.

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

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

If $\beta_n \to 1^-$, then it is shown in Lemma A-3 that under Assumption 1 one can find a subsequence α_n such that $\lim_{n\to\infty} h_{\alpha_n}(\mathbf{x})$ 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 1 holds. Then,

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

$$J^* + h(\mathbf{x}) \geq \bar{c}(\mathbf{x}) + \min_{\mathbf{w}: w_1 + w_2 = M} E_{\mathbf{a}}[h(\mathbf{x}+\mathbf{a})] \ . \tag{35}$$

(b) Let π^* be a policy that minimizes the right hand side of (35). 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 [23] (Chapter 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 $h(\mathbf{x}) = \lim_{\beta\to 1^-} W_\beta(\mathbf{x}) - W_\beta(\mathbf{0})$. 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, and 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. It can be shown (see for example [15]) that a slot should always be allocated to a queue with non-zero deterministic packets. However, when both queues have zero deterministic parts, the allocation depends on the arrival processes. In this section we will use results from 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 $c_1 = c_2$. Then for all $\mathbf{x} \in \mathbb{Z}_+^2$ we have

$$W_t(\mathbf{x}+\mathbf{e}_1) = W_t(\mathbf{x}+\mathbf{e}_2).$$

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 non-empty (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 $t = 0$. Now, suppose the statement is true for $t-1$, i.e $W_{t-1}(\mathbf{x}+\mathbf{e}_1) = W_{t-1}(\mathbf{x}+\mathbf{e}_2)$, $\forall \mathbf{x} \in \mathbb{Z}_+^2$. We want to show $W_t(\mathbf{x}+\mathbf{e}_1) = W_t(\mathbf{x}+\mathbf{e}_2)$.

We first show that when the state is $\mathbf{x}+\mathbf{e}_1$, then it is optimal to allocate the slot to the first queue (similarly, if the state is $\mathbf{x}+\mathbf{e}_2$, then it is optimal to allocate the slot to the second queue).

Suppose the state is $\mathbf{x}+\mathbf{e}_1$ for some $\mathbf{x}\geq 0$. The dynamic equation for the problem is given in (28). The slot is allocated to the first queue if

$$\sum_{a_1,a_2=0}^{\infty} p_1(a_1)p_2(a_2)W_{t-1}(\mathbf{x}+a_1\mathbf{e}_1+a_2\mathbf{e}_2) \leq \sum_{a_1,a_2=0}^{\infty} p_1(a_1)p_2(a_2)W_{t-1}([\mathbf{x}+(a_1+1)\mathbf{e}_1+(a_2-1)\mathbf{e}_2]^+) \ . \quad (36)$$

Using the non-decreasing property of $W_{t-1}(.)$ and the induction hypothesis, we have that for any value of $a_1, a_2 \geq 0$,

$$\begin{aligned} & W_{t-1}(\mathbf{x}+a_1\mathbf{e}_1+a_2\mathbf{e}_2) \\ \leq \ & W_{t-1}([\mathbf{x}+a_1\mathbf{e}_1+(a_2-1)\mathbf{e}_2]^+ +\mathbf{e}_2) \\ = \ & W_{t-1}([\mathbf{x}+(a_1+1)\mathbf{e}_1+(a_2-1)\mathbf{e}_2]^+). \end{aligned} \quad (37)$$

Thus (36) holds and it is optimal to allocate the slot to the first queue. Similar arguments can be used to show the same for the second queue if the state is $\mathbf{x}+\mathbf{e}_2$.

Now we can write

$$\begin{aligned} & W_t(\mathbf{x}+\mathbf{e}_1) = \bar{c}(\mathbf{x}+\mathbf{e}_1) \\ + \ & \beta \sum_{a_1=0}^{\infty}\sum_{a_2=0}^{\infty} p_1(a_1)p_2(a_2)W_{t-1}(\mathbf{x}+a_1\mathbf{e}_1+a_2\mathbf{e}_2) \\ & W_t(\mathbf{x}+\mathbf{e}_2) = \bar{c}(\mathbf{x}+\mathbf{e}_2) \\ + \ & \beta \sum_{a_1=0}^{\infty}\sum_{a_2=0}^{\infty} p_1(a_1)p_2(a_2)W_{t-1}(\mathbf{x}+a_1\mathbf{e}_1+a_2\mathbf{e}_2) \end{aligned}$$

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

It is also easy to see in this case that if one of the queues is empty and the other is non-empty, then it is optimal to allocate the slot to the non-empty queue. Due to space limit the formal proof is not provided here but may be found in [15]. Next we examine the optimal allocation when both queues are empty. Note that Lemma 9 holds true for $t\to\infty$.

Definition 7: 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 $\mathbb{Z}_+$,

$$q_{p_1}(x) \geq q_{p_2}(x),$$

where

$$q_{p_i}(x) = \sum_{y\geq x} p_i(y) \ .$$

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

Theorem 7: Consider time horizon t (t can be ∞) and 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

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

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

$$\begin{aligned}
&\sum_{a_1=0}^{\infty}\sum_{a_2=0}^{\infty} p_1(a_1)p_2(a_2)[W_{t-1}(a_1\mathbf{e}_1+[a_2-1]^+\mathbf{e}_2) - W_{t-1}([a_1-1]^+\mathbf{e}_1+a_2\mathbf{e}_2)] \\
=\;& p_2(0)\sum_{a_1=1}^{\infty} p_1(a_1)[W_{t-1}(a_1\mathbf{e}_1) - W_{t-1}([a_1-1]^+\mathbf{e}_1)] - p_1(0)\sum_{a_2=1}^{\infty} p_2(a_2)[W_{t-1}(a_2\mathbf{e}_2) - W_{t-1}([a_2-1]^+\mathbf{e}_2)] \\
=\;& \sum_{a=1}^{\infty}(p_2(0)p_1(a) - p_1(0)p_2(a))\cdot[W_{t-1}(a\mathbf{e}_1) - W_{t-1}((a-1)\mathbf{e}_1),
\end{aligned} \tag{38}$$

where the first equality is due to Lemma 9 and the second equality uses the relation $W_{t-1}(a\mathbf{e}_1) = W_{t-1}(a\mathbf{e}_2)$, which can be shown using Lemma 9 and a simple induction.

By the monotonicity and convexity of W_{t-1} the expression in (38) is greater than zero if for any $a' > 0$ we have:

$$\sum_{a=a'}^{\infty}(p_2(0)p_1(a) - p_1(0)p_2(a)) \geq 0 \Longleftrightarrow p_2(0)\sum_{a=a'}^{\infty} p_1(a) \geq p_1(0)\sum_{a=a'}^{\infty} p_2(a),$$

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

Let $p_1(a)$, $p_2(a)$ denote the arrival processes for queue 1 and 2, respectively. 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* $(*)$
Stop

Putting the above results together, we see that an optimal policy for this linear equal cost scenario allocates every slot to an non-empty 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 [15]. Interestingly, it was also shown in [15] that in this special case (equal cost, identical arrival) the optimality of 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. Some Numerical Examples

In this section we illustrate some features of the threshold property of the optimal policy in allocating a single slot using numerical examples. Here "time" refers to the actual step or time in the optimization and not "time to go". We will also denote by π^* the greedy policy defined as follows. Policy π^* allocates the next slot to the queue that minimizes the cost only for the next step ahead. This policy is optimal for step $T-1$, but it is not necessarily optimal in general.

We first show the effect of time on the threshold via the following example. Consider $T = 30,\ p_1(0) = 0.1,\ p_1(1) = 0.1,\ p_1(2) = 0.8,\ p_2(0) = 0.8,\ p_2(1) = 0.1$ and $p_2(2) = 0.1$. We want to compare the optimal policy at different time instants. As we proved earlier, at each time instant the optimum policy is of the threshold type. The threshold however may vary over time. Figure 2 illustrates the difference between policies in different time instants when $c(\mathbf{x}) = x_1^2 + x_2^2$. For example, the threshold line for $t = 20$ indicates that at time 20, for all queue sizes on or below this line it is optimal to allocate to the first queue and for all points above this line it is optimal to allocate to the second queue. Note that we do not discount the cost in this case. As t increases (with fewer steps to go), the optimal threshold converges to the greedy policy π^* (i.e., t=29).

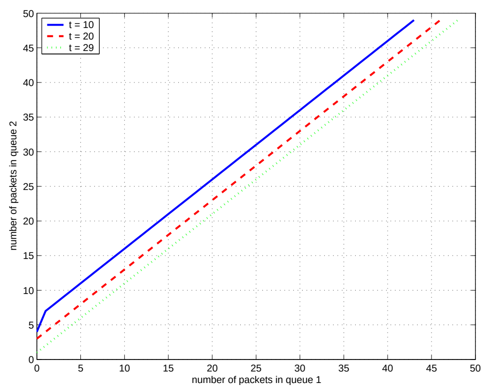

Fig. 2. The effect of time on the optimal threshold level

Consider now the same parameters and cost function as before, but this time with cost discounted by β. Figure 3 shows the thresholds at $t = 10$. It can be seen that as β decreases (heavier and heavier discount, i.e., future becomes less and less important), the optimal policy converges to the greedy policy that optimizes only the next step.

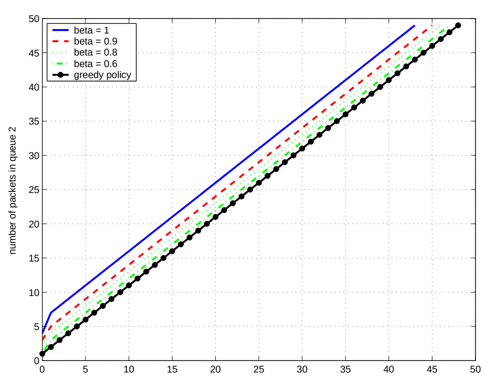

Fig. 3. The effect of the discount factor on the optimal threshold level

Finally, Figure 4 shows the effect of cost function on the optimal threshold. Same parameters are used with $\beta = 1$, and $c(\mathbf{x}) = x_1^n + x_2^n$ where n is a variable. Figure 4 compares the optimal threshold at $t = 10$ for $n = 2, 3$ and 5. As can be seen, as n increases, the threshold moves in favor of the user with more aggressive packet arrivals.

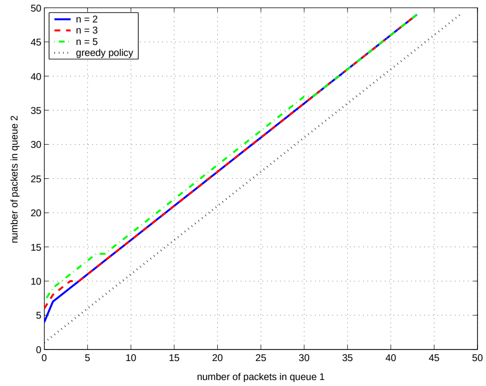

Fig. 4. The effect of the cost function on the optimal threshold level

VIII. Conclusion

In this paper we studied the problem of optimal bandwidth allocation to two users with delayed information about the queue occupancy and 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.

Note that the assumption that the arrival process does not change with time did not appear in any of the proofs in Sections III and IV. We essentially used induction at each step and showed that the properties of $\mathcal{F}$ propagate under any arrival process for the previous time interval. Therefore, the results of Sections III and IV can be generalized to the case where the arrival process changes with time. One key generalization of this work is to the case of more than two queues. This extension is not straightforward and is part of our ongoing research.

Appendix

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

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

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

Proof: In order to show $W_\beta(\mathbf{x})$ is non-decreasing we need to show $W_\beta(\mathbf{0}) \leq W_\beta(\mathbf{x})$. Fix β. We use induction on t to show that $W_t(\mathbf{x})$ (as defined in (28) is non-decreasing for all t. First note that this is true for $t = 1$, since $c(\mathbf{x})$ is non-decreasing. 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 non-decreasing property of $c(\cdot)$ and $W_t(\cdot)$, using the induction hypothesis:

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

Taking the limit on both sides of (A-2) and using (31) we get $W_\beta(\mathbf{x}) \leq W_\beta(\mathbf{x}+\mathbf{e}_i)$, thus $W_\beta(\mathbf{x})$ is non-decreasing 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

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

thus proving the lemma. ■

Lemma A-2: Suppose $c(\mathbf{x}) \geq 0$ for all $\mathbf{x} \geq \mathbf{0}$. Then under Assumption 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 1 implies that $E^{\pi_{\mathbf{0}}}[c(\mathbf{x}_t)|\mathbf{x}_1 = \mathbf{0}] \leq U(\mathbf{0})$. This can be argued as follows. Under policy $\pi_{\mathbf{0}}$, 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

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

where the first inequality is due to the fact that $\pi_{\mathbf{0}}$ 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 non-decreasing) and the last inequality holds by Assumption 1.

■

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

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

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

Proof: Note that $h_{\beta_n} = W_{\beta_n}(\mathbf{x}) - W_{\beta_n}(0) \leq U(\mathbf{x})$ 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 Tychnoff theorem [24]. 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 $0 \leq h(\mathbf{x}) \leq U(\mathbf{x})$.
■

Proof of Theorem 5: Take Equation (34), subtract $\beta W_\beta(\mathbf{0})$ from both sides, and add and subtract $W_\beta(\mathbf{0})$ from the left hand side. We get

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

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

Replace β with γ_n in Equation (A-3) and take the limit infimum on both sides. Using Fatou's Lemma [23] we obtain:

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

Now assume that policy π^* minimizes the right hand side of (35). First we show that $\bar{J}^{\pi^*} \leq J^*$. Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{t+1}$ be the (random) states that are visited at times $1, 2, \cdots, t+1$, then using (A-4) we have (note that $E_{\mathbf{a}}[[\mathbf{x}_t+\mathbf{a}-\mathbf{w}]^+]$ is nothing but $E[\mathbf{x}_{t+1}|\mathbf{x}_t]$),

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

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

$$\frac{1}{t}\sum_{u=1}^{t} E[c(\mathbf{x}_u)] \leq J^* + \frac{E[(h(\mathbf{x}_1) - h(\mathbf{x}_{t+1}))]}{t} \leq J^* + \frac{E[h(\mathbf{x}_1)]}{t}, \tag{A-5}$$

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 $t \to \infty$ 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 [25]),

$$\bar{J}^{\pi^*} \leq J^* \leq \limsup_{\beta\to 1^-}(1-\beta)W_\beta(\mathbf{x}) \leq \limsup_{\beta\to 1^-}(1-\beta)W_\beta^{\pi'}(\mathbf{x}) \leq \bar{J}^{\pi'}\ . \tag{A-6}$$

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 ([26]) or bounded cost functions [20]. For countably infinite state space and unbounded cost functions, [21] has approached the average cost problem for linear cost functions through a limit of finite horizon problems. Other methods can be found in [27], [28] 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 [23]. The assumptions used in [23] are different than Assumption 1 here. However we use the lemmas to show that if Assumption 1 holds, then the three assumptions in [23] 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," *Advances in Applied Probability*, 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. Prob.*, vol. 17, pp. 186–209, 1985.

[3] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 466–478, March 1993.

[4] L. Tassiulas, “Scheduling and performance limits of networks with constantly changing topology,” *IEEE Transactions on Information Theory*, vol. 43, no. 3, pp. 1067–73, May 1997.
[5] N. Bambos and G. Michailidis, “On the stationary dynamics of parallel queues with random server connectivities,” *Proc. 43th Conference on Decision and Control (CDC)*, pp. 3638–43, 1995, New Orleans, LA.
[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,” *Probability in the Engineering and Informational Sciences*, vol. 14, no. 3, pp. 259–297, July 2000.
[7] M. J. Neely, E. Modiano, and C. E. Rohrs, “Power allocation and routing in multibeam satellites with time-varying channels,” *IEEE/ACM Transactions on Networking, Vol. 11, N0. 1*, pp. 138–152, 2003.
[8] M. J. Neely, E. Modiano, and C. E. Rohrs, “Power and server allocation in a multibeam satellite with time-varying channels,” *in Proc. IEEE INFOCOM*, vol. 3, pp. 1451–1460, 2002.
[9] J. Kuri and A. Kumar, “Optimal control of arrivals to queues with delayed queue length information,” *IEEE Transactions on Automatic Control*, vol. 40, no. 8, pp. 1444–1450, August 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] P. Whittle, “Restless bandits: Activity allocation in a changing world,” *A Celebration of Applied Probability, ed. J. Gani, Journal of applied probability*, vol. 25A, pp. 287–298, 1988.
[12] R. Weber and G. Weiss, “On an index policy for restless bandits,” *Journal of Applied Probability*, vol. 27, pp. 637–648, 1990.
[13] J. Nino-Mora, “Restless bandits, patial conservation laws, and indexability,” *Advances in Applied Probability, Vol. 33, no. 1*, pp. 76–98, 2001.
[14] C. H. Papadimitriou and J. N. Tsitsiklis, “The complexity of optimal queueing network control,” *Mathematics of Operations Research, Vol. 24, No. 2*, pp. 293–305, May 1999.
[15] N. Ehsan and M. Liu, “Optimal bandwidth allocation with delayed state observation and batch assignment,” *EECS Technical Report CGR 03-11, University of Michigan, Ann Arbor*, 2003.
[16] N. Ehsan and M. Liu, “On the optimality of an index policy for bandwidth allocation with delayed state observation and differentiated services,” *in Proc. IEEE INFOCOM, Hong Kong*, March 2004.
[17] G. M. Koole, “Structural results for the control of queueing systems using event-based dynamic programming,” *Queueing Systems,Vol. 30*, pp. 323–339, 1998.
[18] E. Altman and G. M. Koole, “On submodular value functions of dynamic programming,” *Technical Report 2658, INRIA Sophia Antipolis*, 1995.
[19] E. Altman and P. Nain, “Closed-loop control with delayed information,” *Performance Evaluation Review, vol. 20, no. 1*, pp. 193–204, 1992.
[20] P. R. Kumar and P. Varaiya, *Stochastic Systems, Estimation, Identification and Adaptive Control*, Prentice Hall, 1986.
[21] B. Hajek, “Optimal control of two interacting service stations,” *IEEE Trans. Auto. Control. AC-29*, pp. 491–499, 1984.
[22] D. P. Bertsekas, *Dynamic Programming, Deterministic And Stochastic Models*, Prentice Hall, 1987.
[23] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*, Wiley Series in Probability and Statistics, 1999.
[24] J. R. Munkres, *Topology*, Prentice Hall, second edition, 2000.
[25] L. Sennott, “A new condition for the existence of optimum stationary policies in average cost morkov decision processes-unbounded cost case,” *Proc. 25th Conf. Decision and Control, Athens, Greece*, pp. 1719–1721, 1986.
[26] S. Ross, *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, 1970.
[27] F. Lu and R. F. Serfozo, “M/m/1 queueing decision processes with monotone hysteritic optimal policies,” *Operat. Res.*, vol. 32, pp. 1116–1132, 1984.
[28] R. R. Weber and S. Stidham, “Optimal control of service rates in networks of queues,” *Adv. Appl. Prob.*, pp. 202–218, 1987.