Properties of Optimal Resource Sharing in a Delay Channel

Navid Ehsan and Mingyan Liu

Abstract— **In this paper we consider the problem of allocating bandwidth/server to two user transmitters/queues with arbitrary arrival processes, to minimize the total expected holding cost of backlogged packets in the system over a finite horizon. However, the queue backlog information is delayed due to communication delay in the channel. In addition, the bandwidth allocation is done in batches, so that a queue can be assigned any number of slots not exceeding the total number in a batch. This problem is motivated by channel allocation in a communication system involving large propagation delay, e.g., a typical satellite data communication scenario. Our principal interest in this paper is to investigate whether the optimal assignment of a batch of slots can be achieved by sequentially using a strategy that is optimal in assigning a single slot, which is typically much easier to find. In this paper we show that if the cost $c(\mathbf{x})$, as a function of the packet backlog $\mathbf{x}$ in the system, is non-decreasing, supermodular and superconvex, then (1) the value function 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.**

Index Terms— **Resource allocation, Delay channel, Stochastic systems, Optimal control.**

I. Introduction

In this paper we study a problem of optimally allocating bandwidth (or servers) to parallel queues when the state information is delayed. Special features of this problem are that servers can be assigned *in batches*, i.e., multiple servers can be allocated to the same queue at a time so that multiple packets may be removed from the queue, and that there is a significant delay between when the allocation decision is made and when the queues are being served, i.e., allocation decision is based on obsolete or delayed state observations.

This bandwidth allocation problem is primarily motivated by communication systems that have large propagation delay, e.g., a typical data communication scenario in a satellite network. Users/terminals transmit packets to the Network Operating Center (NOC) via the satellite. The data communication link from users to the satellite, also known as the *return channel*, follows a dynamic TDMA schedule. Each user is assigned/allocated a certain number of *slots* within a *frame* that consists of a fixed number of slots. A user can only transmit within its assigned slots during every frame. A user also informs the NOC of its current queuing situation (e.g., number of backlogged packets) carried either in packet headers or in a special packet at the beginning of its transmission. The assignment/allocation could be determined either by the satellite or by the NOC, and is broadcast to the users over a *forward channel*, which is separate from (non-interfering with) the return 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 (250 ms from ground/user to satellite and back, or 500 ms from ground/user to ground/NOC via 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 since by that time the backlog situation may have changed. 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.

Similar resource allocation problems also arise in systems where resource allocation is done relatively infrequently compared to packet transmission time, due to cost or design constraints.

In this paper we limit our attention to two users. The extension to more than two users are briefly discussed in the last section. We assume that backlogged packets incur a cost, and formulate an optimal bandwidth allocation problem with the objective of minimizing the expected total packet holding cost over a finite time horizon. We identify three conditions that characterize a class of cost functions, namely monotonicity (non-decreasing), supermodularity, and superconvexity, and explore the properties of an optimal policy under these three conditions. Our contribution is summarized as follows.

1) We show that if the cost function is non-decreasing, supermodular and superconvex, then the value function at each time frame will also satisfy those properties. Furthermore, the optimal policy for assigning a single-slot-frame is of the threshold type.

2) If the cost function is non-decreasing, supermodular and superconvex, then the problem of optimally allocating M slots within each time frame reduces to sequentially allocating a single slot optimally. In other words, a policy that assigns each slot optimally given the previous allocations, is optimal in assigning the entire batch of M slots.

Bandwidth allocation problems 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 them determined policies that maximize

N. Ehsan and M. Liu are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor {nehsan,mingyan}@eecs.umich.edu

throughput over an infinite horizon. In particular, [3] derived the sufficient condition for stability and has shown that serving the Longest Connected Queue (LCQ) policy stabilizes the system if system is stabilizable. The same policy minimizes the delay in the special case of symmetric queues. [6] further considered a similar problem but with differentiated service classes where different queues have different (although linear) holding costs, with the objective being to minimize total discounted holding cost over a finite horizon. It was shown that the optimality of the $c\mu$ rule holds when the indices are sufficiently separated. In all the work mentioned above, the state of the system, i.e., connectivity and the number of packets in each queue, is always precisely known before server allocation is made. In addition, it is also commonly assumed that no more than one server can be allocated to a queue at a time. These constitute the main differences between this paper and the above cited work.

In [7], [8] we have studied a similar problem but with linear cost assumptions. [7] derived the optimal policy when users have the same unit holding cost and identical arrival processes, while [8] investigated optimal policies for differentiated holding costs in the case of a single slot allocation and Bernoulli arrivals. [9] studies the problem of routing to two parallel queues with delayed state observation and shows that when the information is one step delayed it is optimal to join the queue with smaller expected length. [10] studied the problem of optimally routing to two queues with imperfect and noisy information. [11] and [12] present some general results about the propagation of specific properties in some queueing models. The idea in these papers is to show that certain properties of the value function propagate with time. Similar ideas are adopted and explored in this paper.

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 Section III we investigate the optimal policy in the case of two queues. Section IV concludes the paper.

II. PROBLEM FORMULATION

In this section we describe the network model we adopted 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. The Network Model and Notation

Consider N queues that transmit packets to a single server/receiver and in doing so compete for shares of a common channel that consists of *time slots*. Packets arrive at each queue according to some arbitrary random processes. Packets are assumed to be of equal length and one packet transmission time occupies one slot time. M consecutive slots constitute a *frame*. The allocation of the channel is done once for 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 an allocation of a frame, 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 will be served.

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 the 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/receiver and broadcast to all queues over a non-interfering channel. Due to extensive propagation delay in both directions (feedback and the broadcast), 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. Thus there is a one-step delay in state observation by the server. 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 time frame $[t + 1, t + 2)$, denoted by $\mathbf{w}_{t+1}$. However, the server does not know for sure the queue backlog at time $t + 1$ due to random arrivals that occurred during $[t, t + 1)$. The decision is thus made based on partially obsolete information. This procedure begins from $t = 0$ and ends at $t = T$, the 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 $t = 1$).

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.

$\mathbf{x}_t = [\mathbf{b}_{t-1} - \mathbf{w}_{t-1}]^+$: 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 allocation for the period $[t-1, t)$, $\mathbf{w}_{t-1}$, this quantity $\mathbf{x}_t$ is the number 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.

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

$\mathbf{e}_i$: 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 $x^+ = x$ if $x \geq 0$ and is equal to zero otherwise. For a vector $\mathbf{x}$, we define $\mathbf{x}^+$ the same way component-wise. Also for two vectors $\mathbf{x}$ and $\mathbf{y}$ by $\mathbf{x} \leq \mathbf{y}$ we mean that 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 with $\hat{f}$, $\hat{g}$, etc.

The objective is to find an allocation policy π that minimizes the following cost function.

$$J = E^{\pi}[C|\mathbf{b_0}, \mathbf{w_0}] \ , \quad (1)$$
$$\text{where} \quad C = \sum_{t=1}^{T} c(\mathbf{b}_t) \ .$$

For now the packet holding cost $c(\mathbf{b})$ is an arbitrary function. Later, we will introduce specific properties that the cost function should possess so that those properties propagate.

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 is not straightforward and requires stronger assumptions on the cost function. Therefore it will be considered in a separate study.
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 allocated bandwidth for a user is greater than its buffer occupancy in the beginning of a frame, the newly-arrived packets during that frame cannot be transmitted using the extra slots during that frame. This is because the exact arrival times of the packets in a frame is random (e.g., could be toward the beginning or the end of that frame). 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. Consequently we will have to take into account different orderings of the allocation (i.e, not only that user i is assigned j slots, but also which j slots within the frame). This makes the problem very different and much more complicated. In this paper we will limit our attention to the simpler scenario prescribed by this assumption.
5) The server recalls the latest allocation it has made. Note that the expected cost conditioned on the latest allocation, $\mathbf{w}_t$ and buffer occupancy $\mathbf{b}_t$ is independent of arrivals that occurred before frame t. (Note that $\mathbf{b}_t$ 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 $\mathbf{w}_t$).

C. Problem Formulation

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 [13]. 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 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 before. 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 [14]. Define $c'_t(\mathbf{x}) = E_{\mathbf{a}_{t-1}}[c(\mathbf{x} + \mathbf{a}_{t-1})]$, then the dynamic program of the problem is as follows.

$$\begin{aligned} V_T(\mathbf{x}) &= c'_T(\mathbf{x}) \\ V_t(\mathbf{x}) &= c'_t(\mathbf{x}) + \qquad (2) \\ &\min_{\sum_{i=1}^{N} w_{i,t} = M} \{E_{\mathbf{a}_{t-1}}[V_{t+1}([\mathbf{x} + \mathbf{a}_{t-1} - \mathbf{w}_t]^+)]\} \end{aligned}$$

where

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

Remark 1: For the rest of the paper, we make the following additional assumption. We assume that the joint probability mass function of the arrival processes does not change with time, i.e. $p_t(\mathbf{a}_t) = p(\mathbf{a})$; $\forall t$. This assumption is only for convenience in notation and as will be discussed in section IV does not have any effect on our results. Note that by this assumption, we have $c'_t(\mathbf{x}) = c'(\mathbf{x})$ for all t.

Definition 1: Define two operators T_1 and T_M to be

$$T_1 f(\mathbf{x}) = \min_{i\in\{1,2\}} \{f(\mathbf{x}-\mathbf{e}_i)\} \quad (4)$$

$$T_M f(\mathbf{x}) = \min_{\mathbf{w}:w_1+w_2=M} \{f(\mathbf{x}-\mathbf{w})\}\ . \quad (5)$$

Loosely speaking, if $f(.)$ represents the value function, then $T_1 f(.)$ can be viewed as 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.

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

$$\hat{S}_t(\mathbf{x}) = \sum_{\mathbf{a}} p(\mathbf{a}) V_t([\mathbf{x}+\mathbf{a}]^+). \quad (6)$$

We have the following lemma as a result of the definitions above.

Lemma 1: For all values $0 < t < T$, $V_t(\mathbf{x})$ is equal to $\hat{c}'(\mathbf{x}) + T_M \hat{S}_{t+1}(\mathbf{x})$ restricted to $\mathbf{x} \in \mathbb{Z}^2_+$.

III. Optimal Policy For Two Users

We first study the case when each frame consists of only a single slot ($M = 1$), i.e., single slot allocation, and show that the optimal policy is of the threshold type. We then consider $M > 1$ and prove that finding the optimal policy in this case reduces to finding the optimal policy for the single slot assignment.

A. Single Slot Allocation

In this case we have for $\mathbf{x} \in \mathbb{Z}^2_+$,

$$V_t(\mathbf{x}) = c'(\mathbf{x}) + T_1 \hat{S}_{t+1}(\mathbf{x})\ ,$$

$$V_T(\mathbf{x}) = c'(\mathbf{x})\ .$$

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

In what follows we will show that certain properties possessed by $V_t(\mathbf{x})$ propagate backward with t.

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 [11].

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**.

In the remainder of this subsection 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 used in Theorem 1 to show the properties in $\mathcal{F}$ propagate backwards in time for the value functions defined in our problem. A few lemmas are needed for its proof. The proof of the following lemmas can be found in [15].

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 $\mathbf{a} \in \mathbb{Z}^2_+$.

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 are constants.

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 are constants.

The following two lemmas have also been stated in [16].

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

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 now state the first main result of this paper in the following theorem, assuming two users and frames that consist of only one slot for assignment.

Theorem 1: Suppose there are two users and one slot in each frame to be allocated. If the cost function $c(\mathbf{b}) \in \mathcal{F}$, 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.

Proof:

(a) We prove the result by induction. First note that if $c(\mathbf{b}) \in \mathcal{F}$, then $c'(\mathbf{x}) \in \mathcal{F}$ by Lemma 3, therefore $V_T(\mathbf{x}) = c'(\mathbf{x})$ is in $\mathcal{F}$.

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 $T_1 \hat{S}_{t+1}(\mathbf{x})$ restricted to non-negative values is in $\mathcal{F}$. Since $c'(\mathbf{x}) \in \mathcal{F}$, therefore $c'(\mathbf{x}) + T_1 \hat{S}_{t+1}(\mathbf{x})$ 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.

(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

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

By replacing $\mathbf{x}$ with $\mathbf{x}-\mathbf{e}_1-\mathbf{e}_2$ we have

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

Rearranging, we get

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

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 $\mathbf{x}+\mathbf{e}_2$. The threshold may be expressed by the following function.

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

$h_t(x_1) = \infty$ 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, proving the optimality of a threshold policy. ∎

B. Multiple Slot Allocation

In this part we consider the problem of allocating $M > 1$ slots for each time frame and show that this problem reduces to the single slot allocation.

Definition 5: Define recursively the operator T_1^k as follows:

$$T_1^k f(\mathbf{x}) = T_1(T_1^{k-1} f(\mathbf{x})) \ .$$

Theorem 2: If $\hat{f}(\mathbf{x}) \in \hat{\mathcal{F}}$, then we have

$$T_M \hat{f}(\mathbf{x}) = T_1^M \hat{f}(\mathbf{x}).$$

Remark 3: 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 will use Theorem 2 to show that the same policy for a single slot allocation can be repeatedly used (sequentially) M times, and it results in the optimal policy for allocating the batch of M slot.

Proof of the theorem: We use induction on M, interpreted as "the number of allocation slots" in the following. Note that the induction basis for $M=1$ is trivially true. Suppose that the theorem holds for M, we want to show that it holds for $M+1$.

Suppose we have;

$$\mathbf{w}^1 = \arg \min_{\mathbf{w}: w_1+w_2=M} \{\hat{f}(\mathbf{x}-\mathbf{w})\} \tag{7}$$

Suppose we increase the number of slots to $M+1$. We have:

$$T_{M+1}\hat{f}(\mathbf{x}) = \min_{\mathbf{w}: w_1+w_2=M+1} \{\hat{f}(\mathbf{x}-\mathbf{w})\} \tag{8}$$

In this case we show that the vector $\mathbf{w}^1+\mathbf{e}_i$, $i \in \{1,2\}$ is at least "as good as" all vectors 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 the sense of minimizing the right hand side of (8). Note that as j and k vary, this form covers all the possible allocations. We need to show the following:

$$\hat{f}(\mathbf{x}-(w_i^1+k+1)\mathbf{e}_i-(w_j-k)\mathbf{e}_j) \geq \hat{f}(\mathbf{x}-(w_i^1+1)\mathbf{e}_i-w_j^1\mathbf{e}_j) \ . \tag{9}$$

If this holds, then it means that if $\mathbf{w}^1+(k+1)\mathbf{e}_i-k\mathbf{e}_j$ minimizes the right hand side of (8), then $\mathbf{w}^1+\mathbf{e}_i$ will also minimize that equation. Therefore, assume that $\mathbf{w}^1 + (k+1)\mathbf{e}_i - k\mathbf{e}_j$ minimizes the right hand side of (8) and let $\mathbf{w}^2 = \mathbf{w}^1+\mathbf{e}_i$.

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

$$\begin{aligned} & \hat{f}(\mathbf{x}-(w_i^1+k)\mathbf{e}_i-(w_j^1-k)\mathbf{e}_j) \\ & \qquad - \hat{f}(\mathbf{x}-w_i^1\mathbf{e}_i-w_j^1\mathbf{e}_j) \\ \leq \ & \hat{f}(\mathbf{x}-(w_i^1+k+1)\mathbf{e}_i-(w_j^1-k)\mathbf{e}_j) \\ & \qquad - \hat{f}(\mathbf{x}-(w_i^1+1)\mathbf{e}_i-w_j^1\mathbf{e}_j) \ . \end{aligned} \tag{10}$$

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

$$\begin{aligned} & \hat{f}(\mathbf{x}-(w_i^1+1)\mathbf{e}_i-(w_j^1-1)\mathbf{e}_j) \\ & \qquad - \hat{f}(\mathbf{x}-w_i^1\mathbf{e}_i-w_j^1\mathbf{e}_j) \\ \leq \ & \hat{f}(\mathbf{x}-(w_i^1+2)\mathbf{e}_i-(w_j-1)\mathbf{e}_j) \\ & \qquad - \hat{f}(\mathbf{x}-(w_i^1+1)\mathbf{e}_i-w_j^1\mathbf{e}_j) \ . \end{aligned} \tag{11}$$

(11) 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** (if $i=1$ then we use property **C.3.a** and if $i=2$ we use **C.3.b**). Now assume (10) is true for k, $1 \leq k \leq w_j^1$, we want to show that is also true for $k+1$.

In property **C.3** (if $i=1$ then we use property **C.3.a** and if $i=2$ we use **C.3.b**), substituting $\mathbf{x}$ for $\mathbf{x}-(w_i^1+k+2)\mathbf{e}_i-(w_j^1-k)\mathbf{e}_j$ gives

$$\begin{aligned} & \hat{f}(\mathbf{x}-(w_i^1+k+1)\mathbf{e}_i-(w_j^1-k-1)\mathbf{e}_j) \\ & \qquad - \hat{f}(\mathbf{x}-(w_i^1+k)\mathbf{e}_i-(w_j^1-k)\mathbf{e}_j) \\ \leq \ & \hat{f}(\mathbf{x}-(w_i^1+k+2)\mathbf{e}_i-(w_j-k-1)\mathbf{e}_j) \\ & \qquad - \hat{f}(\mathbf{x}-(w_i^1+k+1)\mathbf{e}_i-(w_j^1-k)\mathbf{e}_j) \ . \end{aligned} \tag{12}$$

Combining inequalities (10) and (12) gives the result for case of $k+1$ 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.

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

Therefore the left hand side of (10) is always greater than or equal to zero. Thus the right hand side will also be greater than or equal to zero, i.e.,

$$\hat{f}(\mathbf{x}-(w_i^1+k+1)\mathbf{e}_i-(w_j-k)\mathbf{e}_j) \geq \hat{f}(\mathbf{x}-(w_i^1+1)\mathbf{e}_i-w_j^1\mathbf{e}_j) \ . \tag{13}$$

This means that $\mathbf{w}^2$ minimizes the right hand side of Equation (8).

The above result shows that the minimizer on the right hand side of (8) 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 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. Therefore we have $T_{M+1}\hat{f}(\mathbf{x}) = T_1^{M+1}\hat{f}(\mathbf{x})$, which completes the proof. ■

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

Proof: We use backward induction on t. Since $c(\mathbf{b}) \in \mathcal{F}$, 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 $V_t(\mathbf{x}) \in \mathcal{F}$, we have $\hat{S}_t(\mathbf{x}) \in \hat{\mathcal{F}}$. For $\mathbf{x} \in \mathbb{Z}_+^2$ we have,

$$\begin{aligned} V_{t-1}(\mathbf{x}) &= c'(\mathbf{x}) + T_M\hat{S}_t(\mathbf{x}) \\ &= c'(\mathbf{x}) + T_1^M(\hat{S}_t(\mathbf{x})) \end{aligned}$$

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 $c'(\mathbf{x}) \in \mathcal{F}$ 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 the problem reduces to optimally allocating a single slot. Fix any time t and consider M slot allocation. Suppose $\mathbf{w}$ is the optimal allocation for time frame t given $\mathbf{x}_t$ and $m < M$ slot allocations (i.e. $w_1 + w_2 = m$). Since $V_{t+1}(\mathbf{x}) \in \mathcal{F}$, then by Theorem 2, it can be seen that the $(m+1)$-th slot should be assigned to user i such that;

$$i = \arg\min\{E_{\mathbf{a}}V_{t+1}([\mathbf{x}+\mathbf{a}-\mathbf{w}-\mathbf{e}_i]^+)\}$$

This simply shows that it is optimal to allocate the $(m+1)$-th slot given that the state of the system is $\mathbf{x}-\mathbf{w}$, i.e. the problem can be solved as follows. Allocate the slots sequentially by assigning the $m+1$-st slot optimally given the state of the system and the previous allocations ■

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

IV. Conclusion

In this paper we studied the problem of optimal bandwidth allocation to two users with delayed information about the queue occupancy. We proved the following results:

1) If the cost function $c(\mathbf{b})$ satisfies the three conditions **C.1-C.3**, then the value function at each time frame will also satisfy those conditions. Furthermore, the optimal policy for assigning a single slot at each time frame is of the threshold type.

2) If the cost function $c(\mathbf{b})$ satisfies the three conditions **C.1-C.3**, then the problem of optimally allocating M slots for each time frame reduces to the single slot allocation. In other words a policy that assigns each slot optimally given the previous allocations, is optimal.

Also note that the assumption that the arrival process does not change with time did not appear in any of the proofs. Basically, we used induction at each step and showed that the properties in $\mathcal{F}$ propagate under any arrival process for the previous time interval. Therefore, the results of section III can be generalized to the case where the arrival process changes with time.

The results above are proved for the two user scenario. Extension of this result to more than two users is not straightforward and requires stronger assumptions on the cost function. This is part of our ongoing research.

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] 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.

[8] 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.

[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 domain and thresholds," *Stochastics and Stochastic Reports, Vol. 26*, pp. 81–100, 1989.

[11] G. M. Koole, "Structural results for the control of queueing systems using event-based dynamic programming," *Queueing Systems,Vol. 30*, pp. 323–339, 1998.

[12] E. Altman and G. M. Koole, "On submodular value functions of dynamic programming," *Technical Report 2658, INRIA Sophia Antipolis*, 1995.

[13] E. Altman and P. Nain, "Closed-loop control with delayed information," *Performance Evaluation Review, vol. 20, no. 1*, pp. 193–204, 1992.

[14] P. R. Kumar and P. Varaiya, *Stochastic Systems, Estimation, Identification and Adaptive Control,* Prentice Hall, 1986.

[15] N. Ehsan and M. Liu, "Optimal server allocation in batches," *EECS Technical Report CGR 04-10, University of Michigan, Ann Arbor,* August 2004.

[16] B. Hajek, "Optimal control of two interacting service stations," *IEEE Trans. Auto. Control. AC-29*, pp. 491–499, 1984.