

# Server Allocation with Delayed State Observation: Sufficient Conditions for the Optimality of an Index Policy

Navid Ehsan and Mingyan Liu

**Abstract**—In this paper we study an optimal server allocation problem, where a single server is shared among multiple queues based on the queue backlog information. Due to the physical nature of the system this information is delayed, in that when the allocation decision is made, the server only has the backlog information from an earlier time. Queues have different arrival processes as well as different buffering/holding costs. The objective is to minimize the expected total discounted holding cost over a finite or infinite horizon. We introduce an index policy where the index of a queue is a function of the state of the queue. Our primary interest is to characterize conditions under which this index policy is optimal. We present a fairly general method bounding the reward of serving one queue instead of another. Using this result, sufficient conditions on the optimality of the index policy can be derived for a variety of arrival processes and packet holding costs. These conditions are in general in the form of sufficient separation among indices, and they characterize the part of the state space where the index policy is optimal. We provide examples and derive the indices and illustrate the region where the index policy is optimal.

**Index Terms**—Optimal server allocation, resource allocation, optimization, index policy, delayed state observation, differentiated services, restless bandit, sufficient separation.

## I. INTRODUCTION

THE optimal use of available resources is a key element ensuring the efficiency of any system, wireless networks in particular, as resources (e.g., bandwidth) are shared and tend to be scarce. In this paper we study a class of bandwidth/resource allocation problems, where allocation decisions are based on partial and delayed information of the system state. Consider the problem of  $N$  users/queues competing for shares of a common channel to transmit packets. The channel consists of time frames of a fixed number of  $M$  time slots. Each slot is equivalent to one packet transmission time. A bandwidth allocation policy determines which slot to assign to which user within a frame, as shown in Fig. 1. The allocation decision is made once per frame based on backlog information, i.e., instantaneous queue occupancy, given by the

users/queues at the beginning of each frame. Due to non-negligible delay (e.g., propagation delay), such information reaches the server only in time for the allocation decision to be made for the *next* frame, by which time the queue occupancies may have changed due to packet arrivals within the current frame. In other words, the state information is delayed and partially obsolete. This results in possible over-allocation or under-allocation. Thus in this case the allocation needs to take into account unknown random arrivals that occur in between state information updates. Every queued packet incurs a cost at the beginning of each frame, known as the *buffering* or *holding* cost. This cost may depend on the queue size and may vary from one queue to another, allowing us to consider differentiated service classes. The objective is to minimize the total expected discounted cost over a finite or an infinite horizon.

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 TDMA frame that consists of a fixed number of slots. Users 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. Due to the long propagation delay of the satellite channel (250ms one way), the allocation decision for a particular frame is made based on the backlog information collected during the *previous* frame, which is partially obsolete by the time the allocation is used since by that time the backlog situation may have changed.

In this paper, we will largely focus on a class of policies known as *index policies*. Under an index policy, each user is assigned an *index*, which is a function of its own backlog and/or channel state, and does not depend on other users' states. An index policy is one that serves the user with the highest index at each instance of time. If the indices are chosen properly, the index policy can be shown to be optimal in certain scenarios or under certain conditions [1].

Manuscript received May 23, 2006; revised October 23, 2006; accepted February 1, 2008. The associate editor coordinating the review of this paper and approving it for publication was R. Negi.

This work was completed while N. Ehsan was at the University of Michigan, Ann Arbor. This work was partially supported by Hughes Network Systems, Germantown, MD.

N. Ehsan is currently with Nextwave Wireless Inc., San Diego.

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

Digital Object Identifier 10.1109/TWC.2009.060290

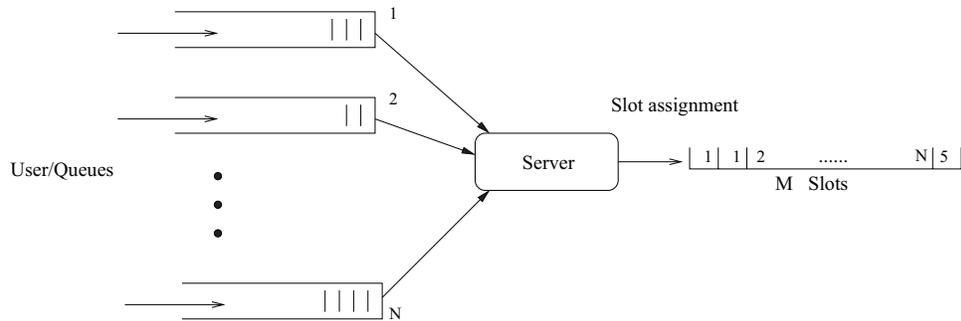


Fig. 1. The bandwidth allocation problem.

Optimal resource allocation problems have been extensively studied in the literature. Here we review those most relevant to the one under consideration in this paper. In [2]-[4] the problem of assigning a single server to parallel queues with different holding costs was considered, where packet transmissions are successful with a certain probability and that the state information on queue backlogs are perfectly observed. It was shown that the  $c\mu$  rule was optimal, where  $c$  is the unit holding cost and  $\mu$  is the probability of transmission success. This can be viewed as an index policy in that the server is always allocated to the non-empty queue with the highest  $c\mu$  value, the index. [5]-[7] considered the server allocation problem to multiple queues with varying connectivity probability but of the same service class. In each of these papers policies that maximize throughput over an infinite horizon were determined. [8] further considered a similar problem where different queues have different holding costs, with the objective being to minimize total discounted holding cost over a finite horizon. An interesting result is that the optimality of an index policy does not hold in general, but holds when the indices are sufficiently separated. The intuition, as pointed out in [8] is that due to different holding costs, allocation to shorter but more costly queues (which runs the risk of emptying the queue) is only justified (or compensated) if it is sufficiently more expensive than a longer but less costly queue. [9] studied the power allocation problem in a single cell with fading and showed that the optimal policy is to schedule the user with the best channel when considering a saturated system (i.e., queues always have packets to send). [10], [11] studied the stability of power allocation policies in a satellite channel.

One common feature of the above cited work is that the state of the system is known before making a decision (e.g., the queue size information in [2]-[4], or the channel gain in [9]). Therefore the policy does not need to take into account the uncertainty in the state observation (in our problem this is reflected as the unknown arrivals and thus the unknown queue size at each decision epoch). This is a major difference between the above cited work and the problem considered here.

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 [12]-[15], where a number of controlled Markov chains undergo state transitions even when they are not played or selected. [12] and [13] 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 Whittle's heuristic, which is sub-optimal in the finite (number of servers and projects) case and asymptotically optimal in the infinite case.

In [16] we studied a special case of the problem considered here, with single slot allocation ( $M = 1$ ), linear holding costs and Bernoulli arrivals. An index policy was introduced and shown to be optimal under sufficient separation of the indices. Given this prior work, a question naturally arises as to whether a method exists with which similar index policies along with their sufficient optimality conditions may be derived for cases of more general holding cost functions and arrival processes. Consequently our primary interest in this paper is to develop such a method. We will limit our attention to the case of allocating frames consisting of a single slot ( $M = 1$ ) in this paper. Discussion on batch assignment ( $M > 1$ ) is provided in Section VII.

In subsequent sections we present a fairly general method that achieves the above goal. We start by bounding the reward of serving one queue instead of another. We then introduce an index policy where the index of a queue is defined as the one step immediate reward of serving the queue, and depends only on its state and arrival statistics. The resulting index policy is greedy in nature.

Using those bounds, sufficient conditions on the optimality of the index policy are derived for a variety of arrival processes and packet holding costs (limited to convex functions). These sufficient conditions are in the form of sufficient separation among indices, and they characterize the part of the state space where the index policy is optimal. We emphasize that although we have considered a very specific problem scenario in this paper (delayed state information) the method itself is quite general and can be applied in a broader class of server allocation problems (for example it can be used to extend the results in [8] to the more general case of convex cost functions and arbitrary arrival and channel state processes).

The rest of the paper is organized as follows. In the next section we formulate the problem and state our assumptions. In Section III we derive the sufficient conditions on the optimality of serving one queue versus another, and apply these conditions to specific examples in Section IV. In Section V we introduce an index based on the one step reward and show by using the results from the previous sections that if

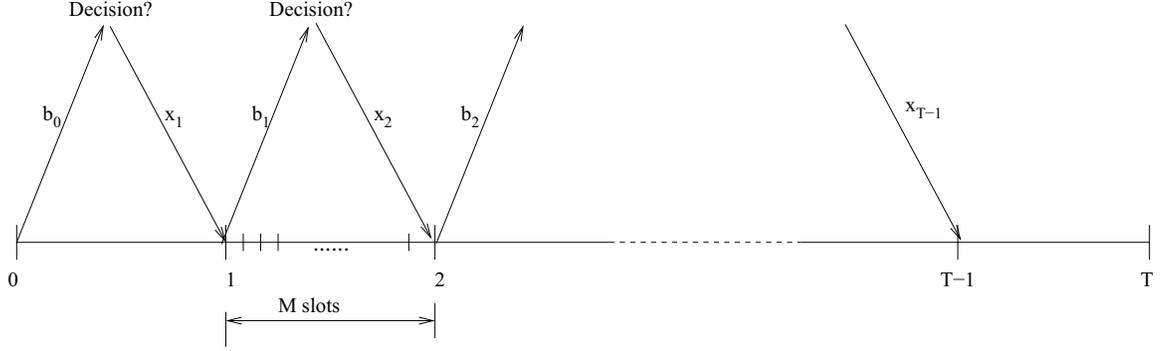


Fig. 2. The bandwidth allocation dynamics.

the indices are sufficiently separated, then the corresponding index policy is optimal. Section VI concludes the paper.

## II. PROBLEM FORMULATION AND ASSUMPTIONS

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

### A. Problem Formulation

Consider a set of  $\mathcal{N} = \{1, 2, \dots, N\}$  queues that need to transmit packets to a single server/receiver and compete for shares of a common channel. Time is slotted, and packets are of equal length and one packet transmission time equals one slot time. Transmissions are assumed to be always successful.

$M$  consecutive slots constitute a frame. The bandwidth allocation decision is based on the backlog information of each queue (number of packets in the queue) provided by the queues at the beginning of a frame. We will ignore the transmission time of such information. This is reasonable since one can always increase the frame length with dedicated fixed number of slots at the beginning for the transmission of such information, which does not affect our discussion of optimal allocation. Based on this information an allocation decision is made by the server and broadcast to all queues over a non-interfering channel. 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, as shown in Fig. 2.

Each user advertises to the server its queue size at the beginning of the  $t^{\text{th}}$  frame, denoted by the  $N$ -vector  $\mathbf{b}_t$ , with  $b_{i,t}$ ,  $i \in \mathcal{N}$  being the queue size of queue  $i$  at time  $t$ . The server allocates slots to be used for transmission in the next time frame, denoted by the  $N$ -vector  $\mathbf{x}_{t+1}$ , with  $x_{i,t+1}$  being the allocation to queue  $i$ , where  $0 \leq x_{i,t+1} \leq M$ ,  $i \in \mathcal{N}$  and  $\sum_{i=1}^N x_{i,t+1} = M$ . This procedure starts from  $t = 0$  and ends at  $t = T$  ( $T$  can be infinite). Note that in this scenario queues do not transmit during the first frame and only start transmitting in the second frame (starting  $t = 1$ ). Similarly, the state information update is not shown for the last frame (starting  $t = T - 1$ ) since the horizon ends at  $t = T$  for the finite horizon case.

For the rest of this paper we will only consider the special case where each frame consists of exactly one slot, i.e.,  $M = 1$ . Consequently,  $x_{i,t+1} \in \{0, 1\}$ ,  $i \in \mathcal{N}$  and  $\sum_{i=1}^N x_{i,t+1} = 1$ .

Let  $b_{i,t}$  denote the number of packets in queue  $i$  at time  $t$ , incurring a cost  $c_i(b_{i,t})$  for that time slot, with the total cost of the system being the summation over all queues, i.e.,  $c(\mathbf{b}_t) = \sum_{i=1}^N c_i(b_{i,t})$ . We will further assume that functions  $c_i(\cdot)$  are non-decreasing and convex<sup>1</sup>.

The objective is to find an allocation policy  $\pi$  that minimizes the following cost function:

$$J_T^\pi = E^\pi[C|\mathcal{F}_0],$$

$$C = \sum_{t=1}^T \beta^{t-1} \sum_{i=1}^N c_i(b_{i,t}),$$

where  $\mathcal{F}_0$  summarizes all the information available at time  $t = 0$ , and  $0 < \beta \leq 1$  is the discount factor<sup>2</sup>.

### B. Assumptions

Below we summarize key assumptions underlying our network model.

- 1) We assume that each user has an infinite buffer. Without this assumption we need to introduce penalty for packet dropping/blocking, which makes the problem drastically different.
- 2) We assume that if for some  $i$  and  $t$  we have  $x_{i,t} > b_{i,t}$  (which implies  $b_{i,t} = 0$  and  $x_{i,t} = 1$ ), then the one slot allocation cannot be used to transmit packets that may have arrived during the  $t^{\text{th}}$  frame, i.e., within  $[t, t+1)$ . This is because the exact arrival time of this packet is random, and unless it arrives right before  $t$  it cannot be transmitted during that slot.
- 3) We assume that the arrivals to each queue are mutually independent and they are also independent and identically distributed in each frame. The arrival process statistics are assumed known to the server.
- 4) The server recalls the latest allocation it has made.
- 5) We will also adopt the trivial assumption that  $\mathbf{x}_0 = \mathbf{0}$  for simplicity of discussion. It does not affect our results on optimal policy and can be easily relaxed in a straightforward way.

<sup>1</sup>Concave cost functions lead to unfair allocation, as the marginal gain for allocating a server to a queue decreases as the queue-size increases, resulting in more incentive to allocate the server to the smaller queues. This does not seem natural, thus the convex function assumption.

<sup>2</sup>The condition  $\beta < 1$  is required in the infinite horizon case to ensure that the cost of having an extra packet in a queue is bounded. For the finite horizon case this is not necessary, i.e., we can have  $\beta = 1$  in that case.

### C. Notations

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

As a general rule, boldface letters represent column vectors and normal letters represent scalars/random variables. Whenever we need to distinguish between two policies, we show the policy as a superscript. For example  $b_{i,t}^\pi$  means the buffer size of the  $i$ -th queue at time  $t$  under policy  $\pi$ . If  $w$  is a scalar,  $[w]^+$  takes value  $w$  or 0, whichever is greater. For a vector  $\mathbf{w}$ , we define  $[\mathbf{w}]^+$  in the same way for each component.

A list of notations is as follows.

$\mathbf{b}_t = [b_{1,t}, b_{2,t}, \dots, b_{N,t}]'$ : The column vector of all queue occupancies at time  $t$ .

$\mathbf{x}_t = [x_{1,t}, x_{2,t}, \dots, x_{N,t}]'$ : The number of slots (amount of bandwidth) allocated to users,  $x_i(t) \in \{0, 1\}$ ,  $i = 1, \dots, N$ ,  $t = 1, \dots, T - 1$ .

$\mathbf{d}_t = [\mathbf{b}_{t-1} - \mathbf{x}_{t-1}]^+$ . This value is completely determined from the buffer occupancy and allocation information of the  $(t-1)^{th}$  frame. We will call this amount the *existing backlog* since this is the amount carried over from the previous slot due to under-allocation (as opposed to new arrivals occurred during the previous slot). Alternatively we will also call this value the amount of *deterministic packets* to be distinguished from the random arrivals occurred during that frame.

$\mathbf{a}_t = [a_{1,t}, a_{2,t}, \dots, a_{N,t}]'$ : The number of packet arrivals during frame  $t$ . We sometimes use the notation  $a_i$  to denote the random arrivals in queue  $i$  during a time frame whenever it does not cause any confusion.

$p_i(k) = \mathbb{P}[a_{i,t} = k]$ ,  $\forall t$  (note that the arrivals are independent and identically distributed in each frame).

$\mathbf{d}_t^{i+} := \mathbf{d}_t + \mathbf{e}_i$  where  $\mathbf{e}_i$  is an  $N$ -dimensional vector with all entries zero except for a 1 in the  $i$ -th position.

$\mathbf{d}_t^{i-} := [\mathbf{d}_t - \mathbf{e}_i]^+$ .

$c_i(b_i)$ : The holding cost function for having  $b_i$  packets in queue  $i$ .

$C_t = \sum_{u=t}^T \beta^{u-t} \sum_{i=1}^N c_i(b_{i,u})$ : The cost to go, from time  $t$  on (note that  $C_1 = C$ ).

$\mathcal{F}_t$ : The  $\sigma$ -field of the information available up to time  $t$ .

**Remark 1:** The information available for making the allocation at time  $t$  is the queue occupancy of the previous frame  $\mathbf{b}_{t-1}$  and the allocation made earlier,  $\mathbf{x}_{t-1}$ . This will determine the number of deterministic packets in the buffer at time  $t$ ,  $\mathbf{d}_t$ . The total number of packets in the queue at time  $t$  is the sum of this deterministic part plus the random arrivals during slot  $t-1$ , i.e.,

$$\mathbf{b}_t = \mathbf{d}_t + \mathbf{a}_{t-1} . \quad (1)$$

Separating the queue size into the deterministic part and the random part will prove convenient in our analysis of the optimal policy.

Given that the server knows the arrival statistics and that the server recalls its last allocation decision, the state of the system at time  $t$  is completely described by the deterministic part of the queues,  $\mathbf{d}_t$ . Also note that for this problem it can be shown [17] that we can restrict ourselves to the class of

Markov policies (i.e. we can find a Markov policy that is optimal over the set of all general policies).

### III. SUFFICIENT CONDITIONS ON OPTIMALITY

In this section we derive a sufficient condition under which serving a particular queue  $i$  is optimal. We start by introducing an upper and lower bound on the cost difference in assigning the slot to different queues. We then show how these bounds may be calculated to produce the desired sufficient condition.

#### A. Sufficient Conditions

We are interested in sufficient conditions under which it is optimal to assign the slot to one queue instead of another. Therefore the key is to find bounds on the cost difference between the two allocations. The following definition characterizes these bounds.

**Definition 1:** Let  $\pi, \pi'$  be the optimal policies given the states  $\mathbf{d}_{t+1}, \mathbf{d}_{t+1}^{i-}$ , respectively. Suppose there exist functions  $S_i(\mathbf{d}, u)$  and  $R_i(\mathbf{d}, u)$  such that

$$\begin{aligned} \beta^t S_i(\mathbf{d}_{t+1}, T-t) & \leq E^\pi [C_{t+1} | \mathbf{d}_{t+1}, \mathcal{F}_{t+1}] - E^{\pi'} [C_{t+1} | \mathbf{d}_{t+1}^{i-}, \mathcal{F}_{t+1}] \\ & \leq \beta^t R_i(\mathbf{d}_{t+1}, T-t) \quad a.s., \end{aligned} \quad (2)$$

where  $T$  is the time horizon. We call function  $S_i$  the *lower bound on cost increase for having one more packet in queue  $i$*  or in short, the *lower bound on queue  $i$* . We call function  $R_i$  the *upper bound on cost increase for having one more packet in queue  $i$*  or in short, the *upper bound on queue  $i$* .

Functions  $R_i$  and  $S_i$  are not unique. We will focus on those that only depend on the state of queue  $i$ , i.e.,  $d_i$ . This turns out to be possible, as we will show later via Lemmas 2 and 3.

Suppose at time  $t$  we want to allocate the slot to one of the queues. The following lemma compares the allocation to two different queues.

**Lemma 1:** Let  $T$  be the time horizon and suppose  $\mathbf{d}_t$  is the state at time  $t$ . Let  $\pi$  be the policy that assigns the slot at time  $t$  to queue  $i$  and assigns optimally thereafter. Let  $\pi'$  be the policy that assigns the slot at  $t$  to queue  $j$  and assigns optimally thereafter. Suppose there exist functions  $R_k(\cdot, \cdot), S_k(\cdot, \cdot), k \in \mathcal{N}$  that satisfy (2). If the following condition holds:

$$E_{\mathbf{a}_{t-1}} [R_j(\mathbf{d}_t + \mathbf{a}_{t-1}, T-t) - S_i(\mathbf{d}_t + \mathbf{a}_{t-1}, T-t)] \leq 0, \quad (3)$$

then we have:

$$E^\pi [C_t | \mathbf{d}_t, \mathcal{F}_t] \leq E^{\pi'} [C_t | \mathbf{d}_t, \mathcal{F}_t] \quad a.s.$$

The proof of this lemma can be found in Appendix A. This lemma immediately leads to the following theorem.

**Theorem 1:** Suppose the state at time  $t$  is  $\mathbf{d}_t$  and that there exist functions  $R_k(\cdot, \cdot), S_k(\cdot, \cdot), k \in \mathcal{N}$  that satisfy (2). Then it is optimal to allocate the slot at  $t$  to queue  $i$  if the following condition holds:

$$\begin{aligned} E_{\mathbf{a}_{t-1}} [R_j(\mathbf{d}_t + \mathbf{a}_{t-1}, T-t) - S_i(\mathbf{d}_t + \mathbf{a}_{t-1}, T-t)] & \leq 0 \\ \forall j \neq i . \end{aligned} \quad (4)$$

*Proof:* This is a direct result of Lemma 1, by comparing allocation to queue  $i$  with allocation to all other queues. ■

### B. Calculating The Bounds

In this part we present a general method for finding functions  $S_i(\cdot, \cdot)$ ,  $R_i(\cdot, \cdot)$  for an arbitrary arrival process. Recall the following assumption on the cost function.

**Assumption 1:** The cost function  $c_i(b_i)$  is non-decreasing and convex in  $b_i$ .

**Definition 2:** For any user  $i$  and deterministic queue size  $d_i$ , define the cost  $\hat{c}_i(\cdot)$  to be

$$\hat{c}_i(d_i) = \sum_{a=0}^{\infty} p_i(a) c_i(d_i + a),$$

where  $\hat{c}_i$  is essentially the expected holding cost of queue  $i$  given that the deterministic part is  $d_i$ . This definition connects cost as a function of the deterministic queue and cost as a function of the actual buffer occupancy. Note that by Assumption 1, function  $\hat{c}_i$  is also non-decreasing and convex in  $d_i$ .

**Definition 3:** For any user  $i$  and  $b_i$ ,  $d_i > 0$  define  $\alpha_i(b_i)$  and  $\hat{\alpha}_i(d_i)$  to be:

$$\begin{aligned} \alpha_i(b_i) &= \inf\{\alpha \in \mathbb{R} | \alpha \geq 0, (c_i(b_i + 1) - c_i(b_i)) \\ &\leq (1 + \alpha)(c_i(b_i) - c_i(b_i - 1))\}, \\ \hat{\alpha}_i(d_i) &= \inf\{\hat{\alpha} \in \mathbb{R} | \hat{\alpha} \geq 0, (\hat{c}_i(d_i + 1) - \hat{c}_i(d_i)) \\ &\leq (1 + \hat{\alpha})(\hat{c}_i(d_i) - \hat{c}_i(d_i - 1))\}. \end{aligned} \quad (5)$$

The above definitions are introduced primarily for technical reasons.  $\alpha$  indicates how fast the function is growing compared to an exponential growth, thus it is a measure of how convex the cost function is. If the function grows exponentially, then  $\alpha$  is a constant. If the function grows slower than exponential, then  $\alpha$  decreases as the queue size increases. Note that if  $c_i(b_i)$  is linear in  $b_i$ , then  $\hat{c}_i(d_i)$  is linear in  $d_i$  (from Definition 2). In this case (linear cost functions) we have  $\alpha_i(b_i) = \hat{\alpha}_i(d_i) = 0$ ,  $\forall b_i, d_i \geq 1$ .

For simplicity of our discussion in the next few lemmas, we further introduce the following random processes.

Let  $\{X_{i,t}\}$  be a Markov chain taking values in the set  $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$  with transition probabilities

$$\mathbb{P}(X_{i,t+1} = l | X_{i,t} = k) = p_i(l - k), \quad (6)$$

where  $p_i(l - k) = 0$  for all  $l < k$ . Note that if  $X_{i,t_0} = d_{i,t_0}$ , then  $X_{i,t}$  represents the number of deterministic packets in queue  $i$  at time  $t \geq t_0$  if it is never served or allocated a slot.

Let  $\{Y_{i,t}\}$  be a Markov chain taking values in the set  $\mathbb{Z}^+$  with transition probabilities

$$\mathbb{P}(Y_{i,t+1} = l | Y_{i,t} = k) = p_i(l - k + 1), \quad \forall k \neq 0, \quad (7)$$

where  $p_i(l - k + 1) = 0$  for all  $l < k - 1$ , and  $\mathbb{P}(Y_{i,t+1} = 0 | Y_{i,t} = 0) = 1$  so that 0 is an absorbing state. Note that if  $Y_{i,t_0} = d_{i,t_0}$ , then  $Y_{i,t}$  represents the number of deterministic packets in queue  $i$  at time  $t \geq t_0$  if queue  $i$  is assigned every slot until its deterministic part becomes zero.

Note that the transition probabilities of both processes  $X_{i,t}$  and  $Y_{i,t}$  are functions of the  $i$ -th queue's arrival process.

When  $d_i = 0$ , then it is easy to see that  $R_i(\mathbf{d}, u) = S_i(\mathbf{d}, u) = 0$  satisfies (2). In the following lemmas we calculate these bounds when  $d_i > 0$ .

**Lemma 2:** Let  $\pi$  be the optimal policy given  $\mathbf{d}_{t+1}$  where  $d_{i,t+1} > 0$  and let  $\pi'$  be the optimal policy given  $\mathbf{d}_{t+1}^-$ . Then we have

$$\begin{aligned} &E^\pi[C_{t+1} | \mathbf{d}_{t+1}, \mathcal{F}_{t+1}] - E^{\pi'}[C_{t+1} | \mathbf{d}_{t+1}^-, \mathcal{F}_{t+1}] \\ &\geq \beta^t \sum_{k=1}^{\infty} (\hat{c}_i(k) - \hat{c}_i(k-1)) \\ &\cdot \sum_{u=0}^{T-t-1} \beta^u \mathbb{P}(Y_{i,u} = k | Y_{i,0} = d_{i,t+1}). \end{aligned} \quad (8)$$

Proof of this lemma can be found in Appendix B. This result shows that we can find function  $S_i$  that satisfies (2) and is a function of the state of queue  $i$ ,  $d_i$  rather than the whole vector  $\mathbf{d}$ .

**Remark 2:** Although Lemma 2 gives a good lower bound for the cost difference of starting from state  $\mathbf{d}$  rather than  $\mathbf{d}^{i-}$ , we will use the following bound instead. It is not as tight as that given in Lemma 2, but has a more explicit expression and will be more useful in the examples presented in the next section. Since  $d_{i,t+1} > 0$ , we have:

$$\begin{aligned} &\beta^t \sum_{k=1}^{\infty} (\hat{c}_i(k) - \hat{c}_i(k-1)) \sum_{u=0}^{T-t-1} \beta^u \mathbb{P}(Y_{i,u} = k | Y_{i,0} = d_{i,t+1}) \\ &\geq \beta^t \sum_{k=d_{i,t+1}}^{\infty} (\hat{c}_i(k) - \hat{c}_i(k-1)) \\ &\cdot \sum_{u=0}^{T-t-1} \beta^u \mathbb{P}(Y_{i,u} = k | Y_{i,0} = d_{i,t+1}) \\ &\geq \beta^t (\hat{c}_i(d_{i,t+1}) - \hat{c}_i(d_{i,t+1} - 1)) \\ &\cdot \sum_{u=0}^{T-t-1} \beta^u \sum_{k=d_{i,t+1}}^{\infty} \mathbb{P}(Y_{i,u} = k | Y_{i,0} = d_{i,t+1}) \\ &\geq \beta^t \Delta \hat{c}_i(d_{i,t+1}) \sum_{u=0}^{T-t-1} \beta^u \mathbb{P}(a_{i,t'} > 0, t \leq t' \leq t+u) \\ &= \beta^t \Delta \hat{c}_i(d_{i,t+1}) \frac{1 - (\beta(1 - p_i(0)))^{T-t}}{1 - \beta(1 - p_i(0))}, \end{aligned} \quad (9)$$

where  $\Delta \hat{c}_i(d_i) = \hat{c}_i(d_i) - \hat{c}_i([d_i - 1]^+)$ . Note that the first and the second inequalities result from the convexity of  $\hat{c}_i$  as a consequence of Assumption 1. The third inequality is due to the fact that  $\{a_{i,t'} > 0, t \leq t' \leq t+u\}$  is only one of the events resulting in  $Y_{i,u} \geq d_{i,t+1}$  given  $Y_{i,0} = d_{i,t+1}$ .

Next, we find an expression for  $R_i(\mathbf{d}, T)$ . We consider non-decreasing and convex cost functions that also satisfy the following assumption.

**Assumption 2:**  $\hat{\alpha}_i(d_i)$  is a non-increasing function of  $d_i$  for  $d_i > 0$  for all  $i \in \mathcal{N}$ .

A few comments are needed on this additional assumption before we proceed. Firstly, as long as  $\hat{c}_i$  does not grow faster than exponential in  $d_i$  then Assumption 2 is true. In fact under many commonly used cost functions  $\hat{\alpha}_i$  decreases with  $d_i$ , e.g., when the cost function is a polynomial of any degree and arrivals in each time frame are finite with probability one. For instance, consider  $c_i(b_i) = c_i b_i^n$  where  $n$  is a positive integer and assume that packets arrive in batches of  $K$  packets. During each frame a batch of  $K$  packets arrive with probability  $p_i$  and

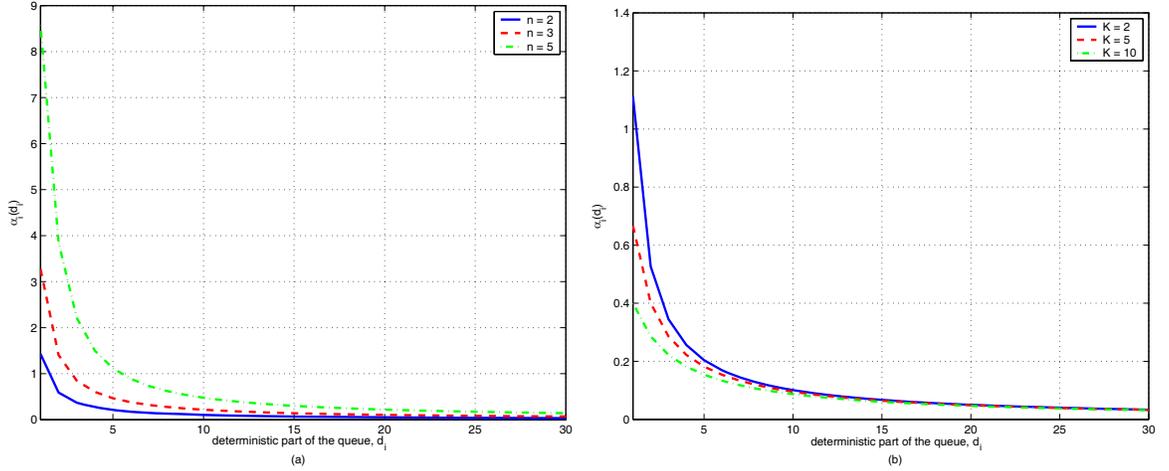


Fig. 3. The effect of cost function and arrival processes on  $\hat{\alpha}$  (the Y-axis in both figures is  $\hat{\alpha}_i$ ). (a)  $K = 1$  fixed,  $n$  variable; (b)  $n = 2$  fixed,  $K$  variable.

there will be no arrivals with probability  $1 - p_i$ . In this case we have  $\hat{c}_i(d_i) = c_i(1 - p_i)d_i^n + c_i p_i(d_i + K)^n$ . If we let  $c_i = 1$  and  $p_i = 0.2$ , Figure 3(a) plots  $\hat{\alpha}_i$  as a function of  $d_i$  for different values of  $n$  where  $K = 1$  is fixed. Figure 3(b) plots  $\hat{\alpha}_i$  as a function of  $d_i$  for different values of  $K$  where  $n = 2$  is fixed. It can be seen from these figures that  $\hat{\alpha}_i$  decays rapidly and approaches zero. Therefore we will also try to find simpler expressions from time to time for a class of convex cost functions where  $\hat{\alpha}_i(d_i) \rightarrow 0$  as  $d_i \rightarrow \infty$ .

**Lemma 3:** Let  $\pi$  be the optimal policy given  $\mathbf{d}_{t+1}$  where  $d_{i,t+1} > 0$  and let  $\pi'$  be the optimal policy given  $\mathbf{d}_{t+1}^-$ . If Assumptions 1 and 2 hold, then we have:

$$\begin{aligned} & E^\pi[C_{t+1}|\mathbf{d}_{t+1}, \mathcal{F}_{t+1}] - E^{\pi'}[C_{t+1}|\mathbf{d}_{t+1}^-, \mathcal{F}_{t+1}] \\ & \leq \beta^t(\hat{c}_i(d_{i,t+1}) - \hat{c}_i(d_{i,t+1} - 1)) \\ & \cdot \sum_{u=0}^{T-t-1} \beta^u \sum_{l=0}^{\infty} \mathbb{P}[X_{i,u} = l | X_{i,0} = 0] \cdot (1 + \hat{\alpha}_i(d_{i,t+1}))^l. \end{aligned} \quad (10)$$

Proof of this lemma can be found in Appendix C. Again we see from this lemma that there exists function  $R_i$  that satisfies (2) and is only a function of the state of queue  $i$ ,  $d_i$  rather than  $\mathbf{d}$ .

We can find an approximation to the above expression for the case where  $\hat{\alpha}$  is small. As  $\hat{\alpha} \rightarrow 0$  we have  $(1 + \hat{\alpha})^l \approx 1 + l\hat{\alpha}$ . Therefore when  $\alpha$  is close to zero we get (11).

To summarize, in this section we derived a sufficient condition under which serving a particular queue  $i$  is optimal. This sufficient condition is characterized by upper and lower bounds on the cost difference for having one extra packet in a particular queue. In the next section we derive these bounds in specific example scenarios.

#### IV. APPLICATIONS IN SPECIFIC SCENARIOS

In the previous section we proved that it is optimal to allocate the slot at time  $t$  to queue  $i$  if (4) holds. Using Lemmas 2 and 3 we can define functions  $S_i$  and  $R_i$  that satisfy (2) when  $d_i > 0$  as follows (note that we have used the bound

defined in Remark 2 to calculate  $S_i$ ):

$$\begin{aligned} R_i(\mathbf{d}, u) &= \Delta \hat{c}_i(d_i) \sum_{u'=0}^{u-1} \beta^{u'} \sum_{l=0}^{\infty} \mathbb{P}[X_{i,u'} = l | X_{i,0} = 0] \cdot (1 + \hat{\alpha}_i(d_i))^l, \end{aligned} \quad (12)$$

$$S_i(\mathbf{d}, u) = \Delta \hat{c}_i(d_i) \frac{1 - (\beta(1 - p_i(0)))^u}{1 - \beta(1 - p_i(0))}, \quad (13)$$

and when  $d_i = 0$  we have  $R_i(\mathbf{d}, u) = S_i(\mathbf{d}, u) = 0$ .

**Remark 3:** Note that in this case both functions depend on  $d_i$  rather than the whole vector  $\mathbf{d}$ , i.e. we have  $S_i(\mathbf{d}, u) = S_i(d_i, u)$  and  $R_j(\mathbf{d}, u) = R_j(d_j, u)$ . In this case condition (4) reduces to:

$$\sum_{k=0}^{\infty} p_i(k) S_i(d_{i,t} + k, T - t) \geq \sum_{l=0}^{\infty} p_j(l) R_j(d_{j,t} + l, T - t) \quad \forall j \neq i. \quad (14)$$

Therefore for the rest of this paper we will use functions  $R_i$  and  $S_i$  defined as above and use (14) as the sufficient condition for the optimality of allocating to queue  $i$  at time  $t$ .

For finite  $T$  the sufficient condition in (14) depends on  $T$ . In this section we apply these results to two scenarios to derive more specific sufficient conditions for the optimality of assigning the slot to a certain queue.

##### A. Batch and Bernoulli Arrivals

Suppose that the arrivals occur in batches of  $K$  packets. During each time frame, with probability  $p_i$  there are  $K$  arrivals in queue  $i$  and with probability  $1 - p_i$  there are no arrivals in queue  $i$ . Note that  $K = 1$  represents the Bernoulli arrival process.

We denote by  $S_i^B(d_i, u)$  and  $R_i^B(d_i, u)$  the lower and upper bounds satisfying (2). We have that  $p_i(0) = 1 - p_i$ , therefore we can calculate  $S_i^B(d_i, u)$  (for  $d_i > 0$ ) using (13) as follows:

$$S_i^B(d_i, u) = \Delta \hat{c}_i(d_i) \frac{1 - (\beta p_i)^u}{1 - \beta p_i}. \quad (15)$$

$R_i^B(d_i, u)$  can be calculated (for  $d_i > 0$ ) using (12). We have (16).

$$\begin{aligned}
 & \beta^t (\hat{c}_i(d_{i,t+1}) - \hat{c}_i(d_{i,t+1} - 1)) \sum_{u=0}^{T-t-1} \beta^u \sum_{l=0}^{\infty} \mathbb{P}[X_{i,u} = l | X_{i,0} = 0] \cdot (1 + \hat{\alpha}_i(d_{i,t+1}))^l \\
 \approx & \beta^t \Delta \hat{c}_i(d_{i,t+1}) \sum_{u=0}^{T-t-1} \beta^u \left\{ \sum_{l=0}^{\infty} \mathbb{P}[X_{i,u} = l | X_{i,0} = 0] + \hat{\alpha}_i(d_{i,t+1}) \sum_{l=0}^{\infty} l \mathbb{P}[X_{i,u} = l | X_{i,0} = 0] \right\} \\
 = & \beta^t \Delta \hat{c}_i(d_{i,t+1}) \left[ \frac{1 - \beta^{T-t}}{1 - \beta} + \hat{\alpha}_i(d_{i,t+1}) \sum_{u=0}^{T-t-1} \beta^u E(X_{i,u} | X_{i,0} = 0) \right]. \tag{11}
 \end{aligned}$$

$$\begin{aligned}
 R_i^B(d_i, u) &= \Delta \hat{c}_i(d_i) \sum_{u'=0}^{u-1} \beta^{u'} \sum_{l=0}^{\infty} \mathbb{P}[X_{i,u'} = l | X_{i,0} = 0] \cdot (1 + \hat{\alpha}_i(d_i))^l \\
 &= \Delta \hat{c}_i(d_i) \sum_{u'=0}^{u-1} \beta^{u'} \sum_{h=0}^{u'} \mathbb{P}[X_{i,u'} = h \cdot K | X_{i,0} = 0] \cdot (1 + \hat{\alpha}_i(d_i))^{h \cdot K} \\
 &= \Delta \hat{c}_i(d_i) \frac{1 - [\beta((1 - p_i) + p_i(1 + \hat{\alpha}_i(d_i))^K)]^u}{1 - \beta[(1 - p_i) + p_i(1 + \hat{\alpha}_i(d_i))^K]}. \tag{16}
 \end{aligned}$$

Using Theorem 1 for Batch arrivals as defined above we have the following result. It is optimal to allocate the packet at time slot  $t$  to queue  $i$  if

$$\begin{aligned}
 & p_j R_j^B(d_{j,t} + K, T - t) + (1 - p_j) R_j^B(d_{j,t}, T - t) \\
 & \leq p_i S_i^B(d_{i,t} + K, T - t) + (1 - p_i) S_i^B(d_{i,t}, T - t), \\
 & \forall j \in \mathcal{N}. \tag{17}
 \end{aligned}$$

By replacing the expressions for  $S_i^B$  and  $R_j^B$  for this special case we get the following result.

**Theorem 2:** Let  $T$  be the time horizon. Consider batch arrivals where during each time frame queue  $i$  has  $K$  arrivals with probability  $p_i$  and no arrivals with probability  $1 - p_i$ . Suppose the state at time  $t$  is  $\mathbf{d}_t$ , then it is optimal to allocate the slot at time  $t$  to queue  $i$  if the following two conditions hold.

i) For all  $j \neq i$  where  $d_{j,t} > 0$  we have,

$$\begin{aligned}
 & \frac{1 - (\beta p_i)^{T-t}}{1 - \beta p_i} \{p_i \Delta \hat{c}_i(d_{i,t} + K) + (1 - p_i) \Delta \hat{c}_i(d_{i,t})\} \\
 & \geq \frac{p_j (1 - \gamma_j(d_{j,t} + K))^{T-t} \Delta \hat{c}_i(d_{j,t} + K)}{1 - \gamma_j(d_{j,t} + K)} \\
 & + \frac{(1 - p_j) (1 - \gamma_j(d_{j,t}))^{T-t} \Delta \hat{c}_j(d_{j,t})}{1 - \gamma_j(d_{j,t})} \quad \text{if } d_{j,t} > 0.
 \end{aligned}$$

(18)

ii) For all  $j \neq i$  where  $d_{j,t} = 0$  we have,

$$\begin{aligned}
 & \frac{1 - (\beta p_i)^{T-t}}{1 - \beta p_i} p_i \Delta \hat{c}_i(d_{i,t} + K) \\
 & \geq \frac{1 - \gamma_j(d_{j,t} + K)^{T-t}}{1 - \gamma_j(d_{j,t} + K)} p_j \Delta \hat{c}_i(d_{j,t} + K),
 \end{aligned}$$

where  $\gamma_j(d_{j,t}) = \beta[(1 - p_j) + p_j(1 + \hat{\alpha}_j(d_{j,t}))^K]$ .

Theorem 2 holds true for all values of  $T$ . Therefore one can let  $T \rightarrow \infty$  to get the following result <sup>3</sup>.

<sup>3</sup>For the infinite horizon scenario, arrivals should be such that the right hand side of equation (14) remains bounded. Note that this condition holds for the batch and Poisson arrivals, as long as the cost function does not increase exponentially.

**Theorem 3:** Consider an infinite horizon and batch arrivals where during each time frame queue  $i$  has  $K$  arrivals with probability  $p_i$  and no arrivals with probability  $1 - p_i$ . Suppose the state at time  $t$  is  $\mathbf{d}_t$ . It is optimal to allocate the slot at time  $t$  to queue  $i$  if the following two conditions hold.

i) For all  $j \neq i$  such that  $d_{j,t} > 0$  we have  $\gamma_j(d_{j,t}) < 1$  and,

$$\begin{aligned}
 & \frac{1}{1 - \beta p_i} \{p_i \Delta \hat{c}_i(d_{i,t} + K) + (1 - p_i) \Delta \hat{c}_i(d_{i,t})\} \\
 & \geq \frac{p_j \Delta \hat{c}_i(d_{j,t} + K)}{1 - \beta[(1 - p_j) + p_j(1 + \hat{\alpha}_j(d_{j,t} + K))^K]} \\
 & + \frac{(1 - p_j) \Delta \hat{c}_j(d_{j,t})}{1 - \beta[(1 - p_j) + p_j(1 + \hat{\alpha}_j(d_{j,t}))^K]}, \tag{19}
 \end{aligned}$$

ii) For all  $j \neq i$  such that  $d_{j,t} = 0$  we have  $\gamma_j(K) < 1$  and,

$$\begin{aligned}
 & \frac{p_i}{1 - \beta p_i} \Delta \hat{c}_i(d_{i,t} + K) \\
 & \geq \frac{p_j}{1 - \beta[(1 - p_j) + p_j(1 + \hat{\alpha}_j(d_{j,t} + K))^K]} \Delta \hat{c}_i(d_{j,t} + K).
 \end{aligned}$$

### B. Poisson Arrivals

Suppose that arrivals occur according to a Poisson distribution with rate  $\lambda_i$  packets per frame, i.e. we have

$$p_i(k) = e^{-\lambda_i} \frac{\lambda_i^k}{k!}.$$

We denote by  $S_i^P(d_i)$  and  $R_i^P(d_i)$  the lower and upper bounds satisfying (2). We have that  $p_i(0) = 1 - e^{-\lambda_i}$ , therefore we can calculate  $S_i^P(d_i, u)$  using (13) as follows:

$$S_i^P(d_i, u) = \Delta \hat{c}_i(d_i) \frac{1 - (\beta(1 - e^{-\lambda_i}))^u}{1 - \beta(1 - e^{-\lambda_i})}. \tag{20}$$

A lower bound for the right hand side of (14) can be

calculated as follows. First assume that  $d_i > 0$ .

$$\begin{aligned} & \sum_{k=0}^{\infty} e^{-\lambda_i} \frac{\lambda_i^k}{k!} S_i^P(d_i + k, T - t) \\ &= \frac{1 - (\beta(1 - e^{-\lambda_i}))^{T-t}}{1 - \beta(1 - e^{-\lambda_i})} \sum_{k=0}^{\infty} \Delta \hat{c}_i(d_i + k) e^{-\lambda_i} \frac{\lambda_i^k}{k!} \\ &\geq \frac{1 - (\beta(1 - e^{-\lambda_i}))^{T-t}}{1 - \beta(1 - e^{-\lambda_i})} \Delta \hat{c}_i(d_i), \end{aligned}$$

where the inequality is due to the fact that  $\Delta \hat{c}_i$  is non-decreasing as  $\hat{c}_i$  is convex.

Similarly when  $d_i = 0$  it can be shown that:

$$\begin{aligned} & \sum_{k=0}^{\infty} e^{-\lambda_i} \frac{\lambda_i^k}{k!} S_i^P(d_i + k, T - t) \\ &\geq \frac{1 - (\beta(1 - e^{-\lambda_i}))^{T-t}}{1 - \beta(1 - e^{-\lambda_i})} (1 - e^{-\lambda_i}) \Delta \hat{c}_i(1). \end{aligned} \quad (21)$$

On the other hand,  $R_j^P(\cdot, \cdot)$  can be calculated using (12) as follows:

$$\begin{aligned} & R_j^P(d_j, u) \\ &= \Delta \hat{c}_j(d_j) \sum_{u'=0}^{u-1} \beta^{u'} \sum_{l=0}^{\infty} \mathbb{P}[X_{j,u'} = l | X_{j,0} = 0] \cdot (1 + \hat{\alpha}_j(d_j))^l \\ &= \Delta \hat{c}_j(d_j) \sum_{u'=0}^{u-1} \beta^{u'} \sum_{l=0}^{\infty} e^{-\lambda_j u'} \frac{(\lambda_j u')^l}{l!} (1 + \hat{\alpha}_j(d_j))^l \\ &= \Delta \hat{c}_j(d_j) \sum_{u'=0}^{u-1} \beta^{u'} e^{\lambda_j u' \hat{\alpha}_j(d_j)} = \Delta \hat{c}_j(d_j) \frac{1 - (\beta e^{\lambda_j \hat{\alpha}_j(d_j)})^u}{1 - \beta e^{\lambda_j \hat{\alpha}_j(d_j)}}. \end{aligned} \quad (22)$$

Now the left hand side of Equation (14) can be upper bounded as follows, again considering the two cases  $d_j > 0$  and  $d_j = 0$  separately.

In the case of  $d_j > 0$ :

$$\begin{aligned} & \sum_{l=0}^{\infty} p_j(l) R_j(d_j + l, T - t) \\ &= \sum_{l=0}^{\infty} e^{-\lambda_j} \frac{\lambda_j^l}{l!} \Delta \hat{c}_j(d_j + l) \frac{1 - (\beta e^{\lambda_j \hat{\alpha}_j(d_j+l)})^{T-t}}{1 - \beta e^{\lambda_j \hat{\alpha}_j(d_j+l)}} \\ &\leq \Delta \hat{c}_j(d_j) \frac{1 - (\beta e^{\lambda_j \hat{\alpha}_j(d_j)})^{T-t}}{1 - \beta e^{\lambda_j \hat{\alpha}_j(d_j)}} \sum_{l=0}^{\infty} e^{-\lambda_j} \frac{\lambda_j^l}{l!} (1 + \hat{\alpha}_j(d_j))^l \\ &= \Delta \hat{c}_j(d_j) \frac{1 - (\beta e^{\lambda_j \hat{\alpha}_j(d_j)})^{T-t}}{1 - \beta e^{\lambda_j \hat{\alpha}_j(d_j)}} e^{\lambda_j \hat{\alpha}_j(d_j)}, \end{aligned}$$

where the second equality is due to the calculation given in (22), and the inequality is due to the definition of  $\hat{\alpha}$  given in Definition 3, and the assumption that  $\hat{\alpha}$  is non-increasing.

In the case of  $d_j = 0$ :

$$\begin{aligned} & \sum_{l=0}^{\infty} p_j(l) R_j(d_j + l, T - t) \\ &= \sum_{l=1}^{\infty} e^{-\lambda_j} \frac{\lambda_j^l}{l!} \Delta \hat{c}_j(l) \frac{1 - (\beta e^{\lambda_j \hat{\alpha}_j(l)})^{T-t}}{1 - \beta e^{\lambda_j \hat{\alpha}_j(l)}} \\ &\leq \Delta \hat{c}_j(1) \frac{1 - (\beta e^{\lambda_j \hat{\alpha}_j(1)})^{T-t}}{1 - \beta e^{\lambda_j \hat{\alpha}_j(1)}} \sum_{l=1}^{\infty} e^{-\lambda_j} \frac{\lambda_j^l}{l!} (1 + \hat{\alpha}_j(1))^{l-1} \\ &= \Delta \hat{c}_j(1) \frac{(1 - (\beta e^{\lambda_j \hat{\alpha}_j(1)})^{T-t})(e^{\lambda_j \hat{\alpha}_j(1)} - e^{-\lambda_j})}{(1 - \beta e^{\lambda_j \hat{\alpha}_j(1)})(1 + \hat{\alpha}_j(1))}. \end{aligned}$$

The above derivation leads to the following result.

**Theorem 4:** Let  $T$  be the time horizon. Consider Poisson arrivals where during each time frame the number of arrivals to queue  $i$  follows a Poisson distribution with mean  $\lambda_i$ . Suppose the state at time  $t$  is  $\mathbf{d}_t$ , then it is optimal to allocate the slot at time  $t$  to queue  $i$  if the following two conditions hold.

i) For all  $j \neq i$  such that  $d_{j,t} > 0$  we have,

$$\begin{aligned} & \frac{1 - (\beta(1 - e^{-\lambda_i}))^{T-t}}{1 - \beta(1 - e^{-\lambda_i})} \Delta \hat{c}_i(d_i) \\ &\geq \frac{1 - (\beta e^{\lambda_j \hat{\alpha}_j(d_j)})^{T-t}}{1 - \beta e^{\lambda_j \hat{\alpha}_j(d_j)}} e^{\lambda_j \hat{\alpha}_j(d_j)} \Delta \hat{c}_j(d_j). \end{aligned}$$

ii) For all  $j \neq i$  such that  $d_{j,t} = 0$  we have,

$$\begin{aligned} & \frac{1 - (\beta(1 - e^{-\lambda_i}))^{T-t}}{1 - \beta(1 - e^{-\lambda_i})} (1 - e^{-\lambda_i}) \Delta \hat{c}_i(1) \\ &\geq \frac{(1 - (\beta e^{\lambda_j \hat{\alpha}_j(1)})^{T-t})(e^{\lambda_j \hat{\alpha}_j(1)} - e^{-\lambda_j})}{(1 - \beta e^{\lambda_j \hat{\alpha}_j(1)})(1 + \hat{\alpha}_j(1))} \Delta \hat{c}_j(1). \end{aligned}$$

Since Theorem 4 holds true for all  $T$  we can let  $T$  go to infinity and get the following result.

**Theorem 5:** Consider an infinite horizon and Poisson arrivals where during each time frame the number of arrivals to queue  $i$  follows a Poisson distribution with mean  $\lambda_i$ . Suppose the state at time  $t$  is  $\mathbf{d}_t$ , then it is optimal to allocate the slot at time  $t$  to queue  $i$  if the following two conditions hold:

i) For all  $j \neq i$  such that  $d_{j,t} > 0$  we have  $\beta e^{\lambda_j \hat{\alpha}_j(d_j)} < 1$  and,

$$\frac{\Delta \hat{c}_i(d_i)}{1 - \beta(1 - e^{-\lambda_i})} \geq \frac{e^{\lambda_j \hat{\alpha}_j(d_j)} \Delta \hat{c}_j(d_j)}{1 - \beta e^{\lambda_j \hat{\alpha}_j(d_j)}}.$$

ii) For all  $j \neq i$  such that  $d_{j,t} = 0$  we have  $\beta e^{\lambda_j \hat{\alpha}_j(1)} < 1$  and,

$$\frac{(1 - e^{-\lambda_i}) \Delta \hat{c}_i(1)}{1 - \beta(1 - e^{-\lambda_i})} \geq \frac{\Delta \hat{c}_j(1)(e^{\lambda_j \hat{\alpha}_j(1)} - e^{-\lambda_j})}{(1 - \beta e^{\lambda_j \hat{\alpha}_j(1)})(1 + \hat{\alpha}_j(1))}.$$

In this section we considered two specific examples and derived the upper and lower bounds on the gain for allocating to a queue. Using these bounds we were able to find sufficient conditions for the optimality of a policy. These conditions, while easy to verify, are not simple to interpret. In the next section we derive alternative sufficient conditions that appear as separation between the indices in an index policy.

## V. OPTIMALITY OF AN INDEX POLICY AND EXAMPLES

In this section we will use the sufficient conditions derived in previous sections to study the optimality of an index policy that is myopic/greedy in nature defined as follows.

**Definition 4:** The index of queue  $i$  at state  $d_i$  is defined as the immediate expected reward from assigning the slot to the queue:

$$\begin{aligned} I_i(d_i) &= \sum_{k=0}^{\infty} p_i(k) [\hat{c}_i(d_i + k) - \hat{c}_i([d_i + k - 1]^+)] \\ &= \sum_{k=0}^{\infty} p_i(k) \Delta \hat{c}_i(d_i + k). \end{aligned} \quad (23)$$

Note that  $\Delta \hat{c}_i(d_i) = 0$  when  $d_i = 0$ . The corresponding index policy is defined as one that assigns the slot to the queue with the highest index at each step.

This index policy is a natural one in that it compares the benefit in allocating the next slot to a user based on the expected reward gained in the next time slot. Results from previous sections can be utilized in the following way in the context of this index policy. Theorem 1 gives the sufficient condition under which it is optimal to assign a slot to queue  $i$ . By deriving appropriate functions  $R$  and  $S$  for given arrival process and cost functions, as shown in the previous section, we can obtain sufficient conditions under which the above index policy is optimal. We shall see that this sufficient condition appears as a separation condition in that the index policy is optimal when the highest index is sufficiently larger than the other indices.

**Remark 4:** The required separation exists only if the loss from not allocating to queue  $j$  is bounded for all  $j \neq i$ . Assumption 2 can be viewed as the condition required for this loss to be bounded.

In the remainder of this section we derive the sufficient conditions under which this index policy is optimal for the two special cases of batch and Poisson arrivals.

### A. Batch and Bernoulli Arrivals

Consider the model of batch arrivals. The index in this case is reduced to

$$I_i(d_i) = p_i \Delta \hat{c}_i(d_i + K) + (1 - p_i) \Delta \hat{c}_i(d_i). \quad (24)$$

Using Theorem 2 and the above index definition, we immediately obtain the following result.

**Theorem 6:** Let  $T$  be the time horizon. Consider batch arrivals where during each time frame queue  $i$  has  $K$  arrivals with probability  $p_i$  and no arrivals with probability  $1 - p_i$ . Suppose the state at time  $t$  is  $\mathbf{d}_t$ , then it is optimal to allocate the slot at time  $t$  to queue  $i$  if both of the following conditions hold.

i) For all  $j \neq i$  such that  $d_{j,t} > 0$  we have,

$$\begin{aligned} & \frac{1 - (\beta p_i)^{T-t}}{1 - \beta p_i} I_i(d_{i,t}) \\ & \geq \max \left\{ \frac{(1 - \gamma_j(d_{j,t} + K)^{T-t})}{1 - \gamma_j(d_{j,t} + K)}, \frac{(1 - \gamma_j(d_{j,t})^{T-t})}{1 - \gamma_j(d_{j,t})} \right\} I_j(d_{j,t}). \end{aligned}$$

ii) For all  $j \neq i$  such that  $d_{j,t} = 0$  we have,

$$\frac{1 - (\beta p_i)^{T-t}}{1 - \beta p_i} I_i(d_{i,t}) \geq \frac{(1 - \gamma_j(d_{j,t} + K)^{T-t})}{1 - \gamma_j(d_{j,t} + K)} I_j(d_{j,t}),$$

where  $\gamma_j(\cdot)$  is defined in the statement of Theorem 2.

**Remark 5:** Note that for  $\alpha_j = 0$  (e.g., when the costs are linear) and  $K = 1$  (Bernoulli arrival),  $\gamma(\cdot)$  reduces to  $\beta$ . Thus the condition in Theorem 6 reduces to the following. It is optimal to allocate to queue  $i$  at time  $t$  if for all  $j \neq i$  we have:

$$I_i(d_{i,t}) \left( \frac{1 - (p_i \beta)^{T-t}}{1 - p_i \beta} \right) \geq I_j(d_{j,t}) \left( \frac{1 - \beta^{T-t}}{1 - \beta} \right). \quad (25)$$

This is the same condition derived in [16] for linear cost and Bernoulli arrivals. In this case with linear cost function  $c_i(b_i) = c_i b_i$ , the index for queue  $i$  reduces to  $c_i$  when  $d_i \neq 0$ , and the index is  $p_i c_i$  when  $d_i = 0$ .

Using Theorem 3 and the index definition, we obtain the following result, noting that  $\hat{\alpha}_j(d_{j,t} + K) \leq \hat{\alpha}_j(d_{j,t})$ .

**Theorem 7:** Consider an infinite horizon and batch arrivals where during each time frame queue  $i$  has  $K$  arrivals with probability  $p_i$  and no arrivals with probability  $1 - p_i$ . Suppose the state at time  $t$  is  $\mathbf{d}_t$ . It is optimal to allocate the slot at time  $t$  to queue  $i$  if the following two conditions hold.

i) For all  $j \neq i$  such that  $d_{j,t} > 0$  we have  $\gamma_j(d_{j,t}) < 1$  and:

$$\begin{aligned} & \frac{1}{1 - \beta p_i} I_i(d_{i,t}) \\ & \geq \frac{1}{1 - \beta[(1 - p_j) + p_j(1 + \hat{\alpha}_j(d_{j,t}))^K]} I_j(d_{j,t}). \end{aligned} \quad (26)$$

ii) For all  $j \neq i$  such that  $d_{j,t} = 0$  we have  $\gamma_j(K) < 1$ :

$$\begin{aligned} & \frac{1}{1 - \beta p_i} I_i(d_{i,t}) \\ & \geq \frac{1}{1 - \beta[(1 - p_j) + p_j(1 + \hat{\alpha}_j(d_{j,t} + K))^K]} I_j(d_{j,t}). \end{aligned} \quad (27)$$

**Example 1:** Consider two queues  $i, j$  and let  $p_1 = 0.5, p_2 = 0.3, c_1 = 1, c_2 = 2$  and let  $c(\mathbf{b}) = c_1 b_1^2 + c_2 b_2^2$ . Figure 4 illustrates the separation discussed above with two values of  $\beta = 0.6$  and  $\beta = 0.8$ . Below the dot-dash line is the region where it is optimal to allocate the slot to queue 1 and to the left of the dash line is the region where it is optimal to allocate to the second queue. The solid line shows the boundary determined by the index policy (above the boundary allocate to queue 2 and below the boundary allocate to queue 1).

For the region where the sufficient condition is not met, the optimal policy can be found by solving the dynamic programming equations (e.g., using standard methods like value iteration or policy iteration). Finding a general, structured (as opposed to numerical) solution for this region remains an open problem. We are also not aware of any bounds on the cost function that apply here.

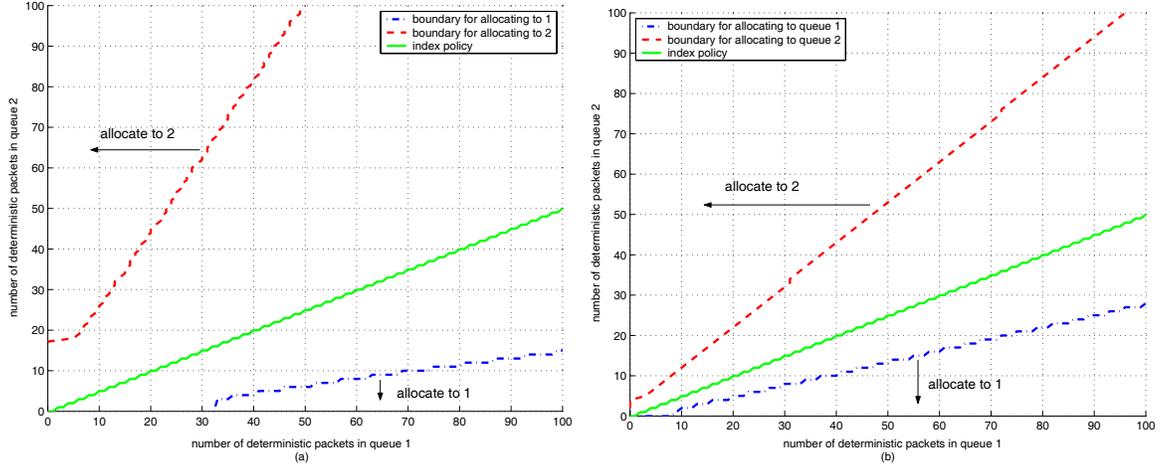


Fig. 4. Required separation between the indices of two queues (a)  $\beta = 0.8$ , (b)  $\beta = 0.6$ .

### B. Poisson Arrivals

Consider the case where queue  $i$  has a Poisson arrival process with parameter  $\lambda_i$ . The index in this case is

$$I_i(d_i) = \sum_{k=0}^{\infty} p_i(k) \Delta \hat{c}_i(d_i + k) = \sum_{k=0}^{\infty} e^{-\lambda_i} \frac{\lambda_i^k}{k!} \Delta \hat{c}_i(d_i + k). \quad (28)$$

**Theorem 8:** Let  $T$  be the time horizon. Consider Poisson arrivals where during each time frame, arrivals to queue  $i$  follow a Poisson distribution with rate  $\lambda_i$ . Suppose the state at time  $t$  is  $\mathbf{d}_t$ , then it is optimal to allocate the slot at time  $t$  to queue  $i$  if for all  $j \neq i$  we have

$$\begin{aligned} & \frac{1 - (\beta(1 - e^{-\lambda_i}))^{T-t}}{1 - \beta(1 - e^{-\lambda_i})} I_i(d_{i,t}) \\ & \geq \frac{1 - (\beta e^{\lambda_j \hat{\alpha}_j(d_{j,t} + \xi_{j,t})})^{T-t}}{1 - \beta e^{\lambda_j \hat{\alpha}_j(d_{j,t} + \xi_{j,t})}} I_j(d_{j,t}), \end{aligned} \quad (29)$$

where  $\xi_{j,t} = \begin{cases} 0, & \text{if } d_{j,t} > 0 \\ 1, & \text{if } d_{j,t} = 0 \end{cases}$ .

The proof of the following theorem can be found in [18]. Letting  $T$  go to infinity we immediately obtain the next result.

**Theorem 9:** Let  $T$  be the time horizon. Consider Poisson arrivals where during each time frame, arrivals to queue  $i$  follow a Poisson distribution with rate  $\lambda_i$ . Suppose the state at time  $t$  is  $\mathbf{d}_t$ , then it is optimal to allocate the slot at time  $t$  to queue  $i$  if for all  $j \neq i$  we have  $\beta e^{\lambda_j \hat{\alpha}_j(d_{j,t} + \xi_{j,t})} < 1$  and,

$$\frac{I_i(d_{i,t})}{1 - \beta(1 - e^{-\lambda_i})} \geq \frac{I_j(d_{j,t})}{1 - \beta e^{\lambda_j \hat{\alpha}_j(d_{j,t} + \xi_{j,t})}},$$

where  $\xi_{j,t}$  is the same as defined in Theorem 8.

The above theorems illustrate the required separation between the indices to ensure the optimality of the index policy. It's worth mentioning that even when the index cannot be explicitly derived, the results obtained in previous section (Theorems 4 and 5) can still be applied to find the region where it is optimal to allocate to any of the queues.

We next illustrate the separation between indices via an example in the case of Poisson arrivals. Finding a closed form expression for the index for all cost functions does not seem straightforward. We therefore consider a special cost function.

**Example 2:** Consider the cost function  $c_i(b_i) = c_i b_i^2$ . We have:

$$\hat{c}_i(d_i) = c_i \sum_{k=0}^{\infty} p_i(k) (d_i + k)^2 = c_i [d_i^2 + 2\lambda_i d_i + (\lambda^2 + \lambda)] \quad (30)$$

The index can be derived as follows. For  $d_i \geq 1$  we have

$$\begin{aligned} I_i(d_i) &= \sum_{k=0}^{\infty} p_i(k) [\hat{c}_i(d_i + k) - \hat{c}_i(d_i + k - 1)] \\ &= c_i \sum_{k=0}^{\infty} p_i(k) (2k + 2d_i + 2\lambda_i - 1) = c_i (4\lambda_i + 2d_i - 1). \end{aligned}$$

For  $d_i = 0$  we have

$$\begin{aligned} I_i(d_i) &= \sum_{k=1}^{\infty} p_i(k) [\hat{c}_i(d_i + k) - \hat{c}_i(d_i + k - 1)] \\ &= c_i \sum_{k=1}^{\infty} p_i(k) (2k + 2d_i + 2\lambda_i - 1) \\ &= c_i [2\lambda_i + (2d_i + 2\lambda_i - 1)(1 - e^{-\lambda_i})]. \end{aligned}$$

Therefore an index can be defined as  $I_i(d_i) = \begin{cases} c_i [4\lambda_i + 2d_i - 1], & \text{if } d_i \neq 0 \\ c_i [2\lambda_i + (2d_i + 2\lambda_i - 1)(1 - e^{-\lambda_i})], & \text{if } d_i = 0 \end{cases}$ .

Figure 5 illustrates the separation condition given by Theorem 9 in the infinite horizon case. We have assumed that  $c_1(b_1) = b_1^2$  and  $c_2(b_2) = 2b_2^2$ . Other parameters are  $\lambda_1 = 0.5$ ,  $\lambda_2 = 0.3$ , and results are shown for  $\beta = 0.6$  and  $\beta = 0.8$ , respectively. Below the dot-dash line is the region where it is optimal to allocate the slot to queue 1 and to the left of the dash line is the region where it is optimal to allocate to the second queue. The solid line shows the boundary determined by the index policy (above the boundary allocate to queue 2 and below the boundary allocate to queue 1).

Characterizing the tightness of these bounds can be very complex and depends on the convexity of the cost function as well as the arrival process. For linear cost functions the index policy can be shown to be optimal except when the deterministic part of all queues are zero [16]. In general the

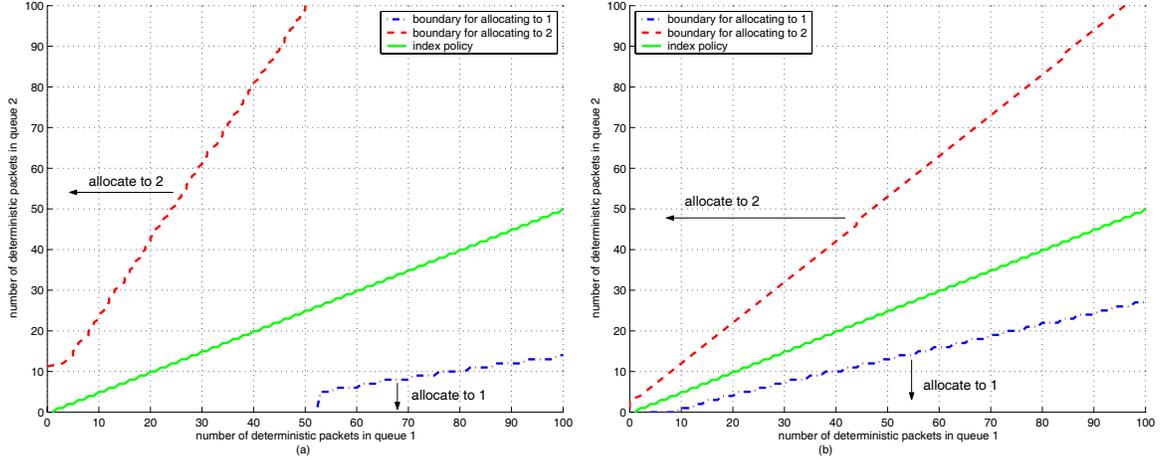


Fig. 5. Required separation between the indices of two queues for Poisson arrivals (a)  $\beta = 0.8$ , (b)  $\beta = 0.6$ .

closer the cost function gets to an exponential function, the looser the bounds will be.

The results presented in this section can be summarized as follows. If the one step reward for serving a queue is sufficiently larger than the one step reward of the other queues, then it is optimal to serve the former. Note that this separation is sufficient but not necessary. The main intuition behind this separation result lies in the need to balance the queue size and at the same time to empty high cost queues. That is, due to the uncertainty in queue sizes, allocating the server to a high cost queue with low probability of arrival runs the risk of wasting resources if there happen to be no arrivals. The separation ensures that the potential loss in doing so is sufficiently compensated by the high cost in case of an arrival. If the number of packets were exactly known at each instant of time, then it is not difficult to show that the index policy of allocating the server to the queue with the highest marginal cost is optimal, i.e., no separation is needed.

VI. DISCUSSION AND CONCLUSION

In this paper we considered the problem of sharing a single server among multiple queues when the queue backlog information is one step delayed. We introduced an index policy that is myopic in nature, and derived sufficient conditions under which it is optimal. This is done by bounding the difference in reward/cost between serving one or the other of any two queues. It is shown that the sufficient condition corresponds to having sufficient separation among the indices. This result is then applied to two specific cases where the arrivals are of the batch and Poisson types.

In this paper we have only considered the case of single slot allocation ( $M = 1$ ). When  $M > 1$  there are more than one slot/server to be allocated. The results in this paper can be extended to the case of  $M > 1$  by considering applying a policy for a single server repeatedly for  $M$  times as follows. We allocate the first slot (say it is allocated to queue  $i$ ). Then we reduce the deterministic part of queue  $i$  (if it is zero then we update the arrival process) and allocate the second slot in a similar manner, and so on. The sufficient condition derived in this paper applies to each slot allocation. If the policy used for a single slot is optimal (note the index policy studied in

this paper is not in general optimal for a single slot allocation, and the identification of an optimal policy remains open), then the optimality of this sequential procedure can be established under fairly general conditions [19]. This result only relies on the existence of an optimal policy, which we show in the current paper.

A main contribution of this paper is the methodology in deriving the sufficiency condition. The method developed in this paper is quite general and can be applied to more complicated arrival processes as well as more general forms of state information. For example it can be used to generalize the results obtained in [8] for server allocation in a simple on-off fading channel, to the case of non-linear convex cost functions [1].

As an extension, it would be interesting to derive similar results for the more general case of restless bandit problems. The results would then help characterize sufficient conditions for the existence of an index policy where the index of each queue is based solely on the property and state of that queue.

APPENDIX - A

*Proof of Lemma 1:* We can write the formula shown at the top of the next page.

Note that in deriving the inequality, we have used Definition 1. For the last equality we have used the fact that both  $\pi$  and  $\pi'$  are optimal from time  $t + 1$  on.

Therefore if (4) holds then we have,

$$E^\pi [C|\mathbf{d}_t, \mathcal{F}_t] \leq E^{\pi'} [C|\mathbf{d}_t, \mathcal{F}_t] \quad a.s.,$$

thus proving the lemma. ■

APPENDIX - B

*Proof of Lemma 2:* Let  $\pi$  be the optimal policy from time  $t + 1$  on given  $\mathbf{d}_{t+1}$ . Define policy  $\hat{\pi}$  for the starting condition  $\mathbf{d}_{t+1}^-$  as follows. Policy  $\hat{\pi}$  assigns slots to the same queue as policy  $\pi$  does under the starting condition  $\mathbf{d}_{t+1}$  for every slot. The “best case”, in the sense of minimizing the difference in cost between the two policies, is if both policies assign to queue  $i$  until (the deterministic part of the queue is) empty. If the slot is assigned to queue  $i$  in all the subsequent time

$$\begin{aligned}
& E^\pi[C_t|\mathbf{d}_t, \mathcal{F}_t] - E^{\pi'}[C_t|\mathbf{d}_t, \mathcal{F}_t] \\
= & E_{\mathbf{a}_{t-1}}\{E^\pi[C_{t+1}|\mathbf{d}_{t+1} = [\mathbf{d}_t + \mathbf{a}_{t-1} - \mathbf{e}_i]^+, \mathcal{F}_{t+1}] \\
& \quad - E^{\pi'}[C_{t+1}|\mathbf{d}_{t+1} = [\mathbf{d}_t + \mathbf{a}_{t-1} - \mathbf{e}_j]^+, \mathcal{F}_{t+1}]\} \\
\leq & E_{\mathbf{a}_{t-1}}\{E^\pi[C_{t+1}|\mathbf{d}_{t+1} = \mathbf{d}_t + \mathbf{a}_{t-1}, \mathcal{F}_{t+1}] - S_i(\mathbf{d}_t + \mathbf{a}_{t-1}, T-t) \\
& \quad + R_j(\mathbf{d}_t + \mathbf{a}_{t-1}, T-t) - E^{\pi'}[C_{t+1}|\mathbf{d}_{t+1} = \mathbf{d}_t + \mathbf{a}_{t-1}, \mathcal{F}_{t+1}]\} \\
= & E_{\mathbf{a}_{t-1}}[R_j(\mathbf{d}_t + \mathbf{a}_{t-1}, T-t) - S_i(\mathbf{d}_t + \mathbf{a}_{t-1}, T-t)] .
\end{aligned}$$

$$\begin{aligned}
& E^\pi[C_{t+1}|\mathbf{d}_{t+1}, \mathcal{F}_{t+1}] - E^{\pi'}[C_{t+1}|\mathbf{d}_{t+1}^{i-}, \mathcal{F}_{t+1}] \geq E^\pi[C_{t+1}|\mathbf{d}_{t+1}, \mathcal{F}_{t+1}] - E^{\hat{\pi}}[C_{t+1}|\mathbf{d}_{t+1}^{i-}, \mathcal{F}_{t+1}] \\
\geq & \sum_{t'=t+1}^T \beta^{t'-1} \sum_{k=1}^{\infty} \mathbb{P}(Y_{i,t'} = k | Y_{i,t+1} = d_{i,t+1} + 1) \cdot (\hat{c}_i(k) - \hat{c}_i(k-1)) \\
= & \sum_{k=1}^{\infty} (\hat{c}_i(k) - \hat{c}_i(k-1)) \sum_{t'=t+1}^T \beta^{t'-1} \mathbb{P}(Y_{i,t'} = k | Y_{i,t+1} = d_{i,t+1} + 1) \\
= & \sum_{k=1}^{\infty} (\hat{c}_i(k) - \hat{c}_i(k-1)) \sum_{u=0}^{T-t-1} \beta^{u+t} \mathbb{P}(Y_{i,t+u+1} = k | Y_{i,t+1} = d_{i,t+1} + 1) \\
= & \beta^t \sum_{k=1}^{\infty} (\hat{c}_i(k) - \hat{c}_i(k-1)) \sum_{u=0}^{T-t-1} \beta^u \mathbb{P}(Y_{i,u} = k | Y_{i,0} = d_{i,t+1} + 1) . \tag{B-1}
\end{aligned}$$

intervals (starting with  $d_{i,t+1} + 1$  deterministic packets in the queue), the number of deterministic packets in queue  $i$  at any time  $t' \geq t + 1$  will have the same distribution as the random process  $Y_{i,t'}$ , conditioned on  $Y_{i,t+1} = d_{i,t+1} + 1$ . This is true until the deterministic part of the queue hits zero. From that point on both policies  $\pi$  and  $\hat{\pi}$  will have the same performance and are both optimal (since  $\pi$  is optimal). Therefore we can write (B-1).

The first inequality is due to the fact that the policy  $\hat{\pi}$  is not necessarily optimal for the initial state  $\mathbf{d}_{t+1}^{i-}$  and the second inequality results from the evolution of the random process  $Y$  as the best case. The rest of the equalities are simple algebra, thus proving the lemma. ■

#### APPENDIX - C

*Proof of Lemma 3:* Let  $\pi'$  be the optimal policy given  $\mathbf{d}_{t+1}^{i-}$ . Define policy  $\hat{\pi}$  for the initial state  $\mathbf{d}_{t+1}$  as follows. Policy  $\hat{\pi}$  assigns the slot to the same queue as policy  $\pi'$  does under the starting condition  $\mathbf{d}_{t+1}^{i-}$  at time  $t + 1$ . The “worst case”, in the sense of maximizing the difference between the cost of two policies, is that queue  $i$  is never served again. Therefore the process for the number of packets in queue  $i$  at any time  $t' \geq t + 1$  (given the initial deterministic packets  $d_{i,t+1} + 1$ ) has the same distribution as  $X_{i,t'}$  given that  $X_{i,t+1} = d_{i,t+1}$ . Therefore we can write (C-1).

The first inequality is obtained since  $\hat{\pi}$  is not necessarily the optimal policy. The second inequality is due to the fact that we are considering the worst case when policy  $\hat{\pi}$  is being used. The next equality is just a change of variables  $l = k - d_i$ . The third inequality is obtained by using the definition of  $\hat{c}_i$ , by which we have:  $\hat{c}_i(d_i + l) - \hat{c}_i(d_i + l - 1) \leq (1 + \hat{\alpha}_i(d_i))^l (\hat{c}_i(d_i) - \hat{c}_i(d_i - 1))$  for  $d_{i,t+1} > 0$ . ■

#### REFERENCES

- [1] N. Ehsan and M. Liu, “Server allocation in wireless networks: the use of index policies,” in *Combinatorial Optimization in Communication Networks*, M. Cheng, Y. Li, and D.-Z. Du, (eds.), Kluwer Academic Publishers, 2006.
- [2] J. S. Baras, A. J. Dorsey, and A. M. Makowski, “Two competing queues with linear costs and geometric service requirements: the  $\mu c$  rule is often optimal,” *Advances in Applied Probability*, vol. 17, pp. 186–209, 1985.
- [3] C. Buyukkoc, P. Varaiya, and J. Warland, “The  $c\mu$ -rule revisited,” *Advances in Applied Probability*, vol. 17, pp. 237–238, 1985.
- [4] J. G. Shanthikumar and D. D. Yao, “Multi class queueing systems: polymatroid structure and optimal scheduling control,” *Oper. Res.*, vol. 40, pp. 293–299, 1992.
- [5] L. Tassiulas and A. Ephremides, “Dynamic server allocation to parallel queues with randomly varying connectivity,” *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 466–478, Mar. 1993.
- [6] L. Tassiulas, “Scheduling and performance limits of networks with constantly changing topology,” *IEEE Trans. Inform. Theory*, vol. 43, no. 3, pp. 1067–1073, May 1997.
- [7] N. Bambos and G. Michailidis, “On the stationary dynamics of parallel queues with random server connectivities,” in *Proc. 43th Conference on Decision and Control (CDC)*, pp. 3638–3643, 1995, New Orleans, LA.
- [8] 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.
- [9] R. Knopp and P. Humblet, “Information capacity and power control in single cell multiuser communications,” in *Proc. IEEE Int. Computer Conf (ICC' 95)*, 1995.
- [10] M. J. Neely, E. Modiano, and C. E. Rohrs, “Power allocation and routing in multibeam satellites with time-varying channels,” *IEEE/ACM Trans. Networking*, vol. 11, no. 1, pp. 138–152, 2003.
- [11] 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.
- [12] P. Whittle, “Restless bandits: activity allocation in a changing world,” *A Celebration of Applied Probability*, J. Gani, ed., *J. Applied Probability*, vol. 25A, pp. 287–298, 1988.
- [13] R. Weber and G. Weiss, “On an index policy for restless bandits,” *J. Applied Probability*, vol. 27, pp. 637–648, 1990.
- [14] J. Nino-Mora, “Restless bandits, partial conservation laws, and indexability,” *Advances in Applied Probability*, vol. 33, no. 1, pp. 76–98, 2001.

$$\begin{aligned}
& E^\pi [C_{t+1} | \mathbf{d}_{t+1}, \mathcal{F}_{t+1}] - E^{\pi'} [C_{t+1} | \mathbf{d}_{t+1}^i, \mathcal{F}_{t+1}] \leq E^{\hat{\pi}} [C_{t+1} | \mathbf{d}_{t+1}, \mathcal{F}_{t+1}] - E^{\pi'} [C_{t+1} | \mathbf{d}_{t+1}^i, \mathcal{F}_{t+1}] \\
\leq & \sum_{t'=t+1}^T \beta^{t'-1} \sum_{k=d_{i,t+1}}^{\infty} \mathbb{P}[X_{i,t'} = k | X_{i,t+1} = d_{i,t+1}] \cdot (\hat{c}(k) - \hat{c}(k-1)) \\
= & \sum_{t'=t+1}^T \beta^{t'-1} \sum_{l=0}^{\infty} \mathbb{P}[X_{i,t'} = l + d_{i,t+1} | X_{i,t+1} = d_{i,t+1}] \cdot (\hat{c}(l + d_{i,t+1}) - \hat{c}(l + d_{i,t+1} - 1)) \\
\leq & \sum_{t'=t+1}^T \beta^{t'-1} \sum_{l=0}^{\infty} \mathbb{P}[X_{i,t'} = l | X_{i,t+1} = 0] \cdot (1 + \hat{\alpha}_i(d_{i,t+1}))^l (\hat{c}_i(d_{i,t+1}) - \hat{c}_i(d_{i,t+1} - 1)) \\
\leq & \Delta \hat{c}_i(d_{i,t+1}) \sum_{t'=t+1}^T \beta^{t'-1} \sum_{l=0}^{\infty} \mathbb{P}[X_{i,t'} = l | X_{i,t+1} = 0] \cdot (1 + \hat{\alpha}_i(d_{i,t+1}))^l \\
= & \beta^t \Delta \hat{c}_i(d_{i,t+1}) \sum_{u=0}^{T-t-1} \beta^u \sum_{l=0}^{\infty} \mathbb{P}[X_{i,u} = l | X_{i,0} = 0] \cdot (1 + \hat{\alpha}_i(d_{i,t+1}))^l. \tag{C-1}
\end{aligned}$$

- [15] 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.
- [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*, Apr. 2004.
- [17] P. R. Kumar and P. Varaiya, *Stochastic Systems, Estimation, Identification and Adaptive Control*. Prentice Hall, 1986.
- [18] N. Ehsan and M. Liu, "Server allocation with delayed state observation: sufficient conditions for the optimality of an index policy," EECS Technical Report, University of Michigan, Ann Arbor, 2006.
- [19] N. Ehsan and M. Liu, "Optimal bandwidth allocation in a delay channel," *IEEE J. Select. Areas Commun.*, vol. 24, no. 8, pp. 1614–1626, Aug. 2006.



**Navid Ehsan** received his B.Sc. degree in 1998 from Sharif University of Technology, Tehran, Iran, and M.Sc. and PhD degrees in 2002 and 2005, respectively, from University of Michigan, Ann Arbor. He is currently with Nextwave Wireless Inc., San Diego. His research interests are in dynamic bandwidth allocation, optimal power allocation and admission control, medium access protocols for wireless systems, security, and stochastic optimization.



**Mingyan Liu** (M'00 / ACM'01) received her B.Sc. degree in electrical engineering in 1995 from the Nanjing University of Aero. and Astro., Nanjing, China, M.Sc. degree in systems engineering and Ph.D. Degree in electrical engineering from the University of Maryland, College Park, in 1997 and 2000, respectively. She joined the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, in September 2000, where she is currently an Associate Professor. Her research interests are in performance modeling, analysis, energy-efficiency and resource allocation issues in wireless, mobile ad hoc, and sensor networks. She is the recipient of the 2002 NSF CAREER Award, and the University of Michigan Elizabeth C. Crosby Research Award in 2003. She serves on the editorial board of IEEE/ACM TRANSACTIONS ON NETWORKING and IEEE TRANSACTIONS ON MOBILE COMPUTING.