Minimizing Backlog for Downlink of Energy Harvesting Networks

Venkhat V, Prasanna Chaporkar, Abhay Karandikar
Department of Electrical Engineering
IIT Bombay, Mumbai, India
Email:{venkhat, chaporkar, karandi}@ee.iitb.ac.in

***Abstract*—A transmitter powered by a renewable energy source becomes self sustainable. In this paper, we consider the broadcast channel with a transmitter and N receivers. The transmitter is powered by a renewable energy source and has finite battery capacity. Transmitter requires power P_i to transmit a packet to i^{th} user. In this setting, our objective is to minimize the expected backlog at the transmitter while accounting for randomness in the arrival and the recharge processes. We formulate the problem as an infinite horizon Markov Decision Process (MDP) problem and obtain the structural properties of an optimal policy. These structural properties provide valuable insights for designing close to optimal policies that are computationally efficient for real life implementations. In special cases, we provide complete description of an optimal policy.**

I. Introduction

World is in power crisis as most conventional energy sources are draining rapidly. Future communication devices are aiming at becoming self sustainable with the use of green energy sources. However, for self sustainability one has to effectively deal with the randomness in energy harvesting process. This introduces new challenges e.g. the existing optimal scheduling algorithms with respect to the conventional sources may not be optimal in energy harvesting networks. So, it's necessary to investigate and obtain optimal algorithms for energy harvesting scenarios. In this paper, we investigate the broadcast channel with multiple receivers and a single transmitter having a finite capacity battery that is powered by a green energy source. Our aim is to minimize the expected backlog at the transmitter. Note that minimizing backlog is equivalent to minimizing queuing delay.

Data communication in energy harvesting systems has been explored in different scenarios, e.g. see [1]–[7]. Optimal offline policy for minimization of transmission completion time in a broadcast channel with finite battery setup is computed assuming the knowledge of future energy values in [1]. Optimal online and offline policies for maximizing throughput and minimizing transmission time is obtained for a wireless fading channel in [2]. The sum throughput is maximized using techniques from calculus of variations in [3]. Here the battery is modeled using the storage dam model. There has also been work that looked into information theoretic capacity of channels in energy harvesting scenarios e.g. see [4]–[6], but this body of work does not consider delay minimization. Optimal policy for maximizing total amount of data transmitted in a given finite duration is obtained in [7]. Transmission completion time under a deterministic setting, i.e. when the arriving energy values are known apriori, is minimized in [8]. The problem of throughput maximization in a point to point link is framed as a Markov Decision Process (MDP) problem and monotone property of an optimal policy is obtained in [9]. Note that most of this work aims at throughput maximization and do not consider delay. In this work, we focus on delay minimization.

Optimal policy that minimizes delay in every slot for any sample path of packet arrivals for tandem and parallel queuing systems is obtained in [10]. Optimal policy for throughput maximization and delay minimization in every slot in a parallel queuing system is obtained in [11]. Trade-offs between average power and average delay has been analyzed for a fading wireless channel in [12]. Average packet transmission delay is minimized for a single user and multiuser uplink fading channel respectively controlling the power and rate dynamically under conventional energy setup in [13], [14]. Average waiting time for a head of line packet is minimized using dynamic programming in loss tolerant MAC layer multicast in [15]. Under average delay constraint, average power is minimized for a single user fading channel and online implementation using stochastic approximation is obtained in [16]. Order optimal delay result in a one hop wireless network with N users and ON-OFF channels is shown in [17]. In renewable energy paradigm, an online algorithm for minimizing delay is proposed and its competitive ratio is analyzed for a arbitrary wireless channel in finite time and Gaussian single user, multi user channels respectively in [18] and [19]. Above branch of work considers some variant of delay optimization in wireless networks. However they do not consider energy harvesting networks. Here our aim is to minimize delay for energy harvesting networks.

In this paper, we investigate the problem of minimizing the backlog at the renewable energy empowered transmitter in a broadcast channel. The transmitter is

978-3-9018-8273-9/15/ ©2015 IFIP

assumed to have a finite battery that is recharged as per some stationary stochastic process. We first demonstrate that, on account of randomness in recharge process, no policy can minimize the backlog on every sample path. This is in contrast with the sample path wise optimality obtained for wireless system with conventional energy sources in [11]. This motivates design of policies that minimize the expected backlog. To this end, we formulate the problem as an MDP, and obtain structural properties of an optimal policy. These structural properties not only aid in reducing computations for obtaining optimal policy, but also provide insight that prove useful for designing near optimal heuristics. In some special cases, we provide the complete description of an optimal policy.

The remaining paper is organized as follows. Section II describes the system model and challenges involved in designing optimal policy. The problem is formulated as MDP in Section IV. In Section V, structural properties of the optimal policy are proved. Simulation results are presented inVI. We conclude the paper in Section VII.

II. System Model

Consider a single server and N parallel queues. Packets are of the constant length l. The time is divided into intervals of fixed length τ called time slots. Server can serve one packet in a slot. This is similar to downlink network with N separate queues for N users and a single base station which decides which user is scheduled. We only assume slow fading, i.e., the channel gains do not vary over time. Let the channel gains be $h_1, \ldots, h_N$. Without loss of generality, $|h_i| > |h_{i+1}|, \forall i \in \{1, \ldots, N-1\}$.The transmission rate is given by the Shannon's capacity formula $B \log_2 \left(1 + \frac{P|h|^2}{BN_0}\right)$, where N_0 is Noise power spectral density. Without loss of generality, assume bandwidth B to be 1. The power required to transmit a packet from the base station to user within a slot is given by $P_i = \frac{N_0}{|h_i|^2}\left(2^{\frac{l}{\tau}} - 1\right)$ for $i = 1, \ldots, N$. So, $P_i < P_{i+1}$ $\forall i$.

Let $A_i(t)$ be the number of packets arriving in queue i at beginning of slot t, for $t \geq 1$. Let $R(t)$ be the recharge energy arrivals which are added to the energy buffer/battery at the beginning of slot t, for $t \geq 1$. The system model with the arrivals are shown in Fig. 1. Action or decision is taken after the arrivals. Queues are considered to be of infinite capacity. So the packet loss never happens. Battery is of finite capacity with ξ_m being the maximum value. Let $Q_i(t)$ indicate the number of packets in the queue i at the beginning of slot t. Let $E(t)$ indicate the amount of energy in the battery at the beginning of slot t. A queue i is said to be *connected*, if $Q_i(t) > 0$ and $E(t) \geq P_i$. Thus in any slot t, a packet can be transmitted only from the set of connected queues, not otherwise. We define a few terms which will be used in this paper hereafter.

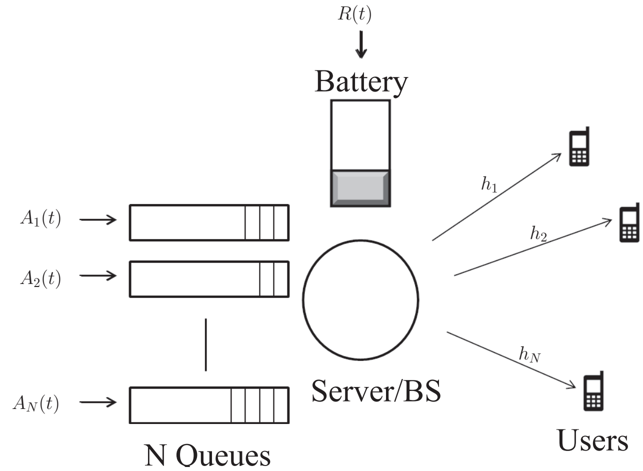

Fig. 1. System Model

Definition 1 (Scheduling Policy)**.** *Scheduling policy is a sequence of decision rules at each slot t, that chooses a connected queue from which a packet will be transmitted in the slot or decides to stay idle.*

We assume scheduling policy to be causal i.e., the action taken is a function of the past actions, energy arrivals and packet arrivals.

Definition 2 (Stationary Policy)**.** *Stationary policy is a map $\pi : \mathcal{S} \to \{0, 1, \ldots, N\}$, i.e. the policy maps the system state s to an action $\{0\} \cup \mathcal{C}_s$, where $\mathcal{C}_s$ is the set of connected queues in state s.*

The stationary policy does not depend on time. Also, given the current state, the decision solely depends on the state and not on the past.

Definition 3 (Non-idling Policy)**.** *A policy π is non-idling,*

1) *if it is stationary*
2) *if $\mathcal{C}_s \neq \emptyset$, then the policy schedules a queue from $\mathcal{C}_s$ for every s*

Any stationary policy which is not non-idling is referred to as idling policy.

Queue length and the battery energy level depend on the scheduling policy. This dependence is made clear by mentioning the policy in the superscript. Let us define a indicator variable, $I_i^\pi(t)$ which is 1, if a packet is scheduled from queue i by policy π in slot t and 0 otherwise. Also, $P^\pi(t)$ denotes power spent in slot t and ξ_m denotes battery capacity. $(\boldsymbol{Q(0)}, E(0))$ denotes the initial values in slot 0. Queue length under policy π evolves as follows. For every $t \geq 1$ and $i \in \{1, \ldots, N\}$,

$$Q_i^\pi(t) = Q_i^\pi(t-1) + A_i(t) - I_i^\pi(t-1). \quad (1)$$

The queue state at t is $\boldsymbol{Q^\pi(t)} = [Q_1^\pi(t), \ldots, Q_N^\pi(t)]^T$, where $\boldsymbol{x}^T$ denotes the transpose of vector $\boldsymbol{x}$. Thus in vector notation queue state evolves as follows:

$$\boldsymbol{Q}^{\pi}(t) = \boldsymbol{Q}^{\pi}(t-1) + \boldsymbol{A}(t) - \boldsymbol{I}^{\pi}(t-1).$$

The battery energy level under policy π in slot t is given as follows,

$$E^{\pi}(t) = \min\left\{E^{\pi}(t-1) + R(t) - P^{\pi}(t-1), \xi_m\right\}. \tag{2}$$

Recall that policy π can schedule only from a connected queue, which implies $E^{\pi}(t-1) + R(t) \geq P^{\pi}(t-1)$ always. On the next section we discuss challenges involved in designing delay optimal policy for energy harvesting network.

III. Challenges in Designing Optimal Policy

Let us first define notion of delay optimality akin to the one considered in [11].

Definition 4 (Backlog optimality everywhere)**.** *Scheduling policy π is backlog optimal everywhere if it satisfies $\sum_{i=1}^{N} Q_i^{\pi}(t) \leq \sum_{i=1}^{N} Q_i^{\pi'}(t) \quad \forall \pi', t = \{0, 1, 2, \ldots\}$, under any packet and recharge energy arrivals and any initial state $(\boldsymbol{Q}(\boldsymbol{0}), E(0))$.*

Note that the backlog optimality everywhere is the sample path wise optimality. Different systems have been studied in a similar way in [10], [11]. In [10], authors have described the policy which achieves delay optimality everywhere for tandem queuing and parallel queuing systems with adjacency constraints on servers. In the parallel queuing system, the policy which schedules such that it serves most number of queues, achieves backlog optimality everywhere. Whereas in a tandem queuing system with a single destination, the following policy is shown to be delay optimal everywhere: Select a non-empty queue (say i) that is closest to the destination, then choose a non-empty queue that is closest to the queue i and does not interfere with i's transmission and repeat this until no further queue can be selected. The policy schedules a packet from all the chosen queues simultaneously in a slot. For multi user downlink with random connectivity, when all users' arrival and channel connectivity processes are identical. The authors show that the longest connected queue (LCQ) policy minimizes the backlog in the system at every time slot in [11]. From these papers, it is observed that for some systems, there exists a policy which has optimality at every time slot.

Note that the aforementioned work does not consider energy harvesting scenario. In our model, if all queues remain connected in all slots, then any non-idling policy is backlog optimal everywhere. The queues remain connected if for example, energy arrival $R(t) > P_N, \ \forall\ t$. Next we address the existence of backlog optimality everywhere when *queues do not remain connected in all the slots* . Specifically we show that backlog optimality everywhere does not exist in this scenario. We state this formally in the following theorem.

Theorem 1. *There does not exist an optimal stationary policy π that achieves backlog optimality everywhere.*

This theorem is proved in the following four subparts. In the first part we show that if there exists an optimal policy, then there exists a stationary policy which is optimal. In the second part, it is shown that if an optimal stationary policy exists, then it belongs to the class of non-idling policies. Next we show that the optimal non-idling policy must schedule the lowest index connected queue (LICQ). In the last part, it is shown that LICQ policy is not an opitmal policy with the help of an example. These four parts are proved in the following four lemmas.

Lemma 1. *If there exists a policy that is backlog optimal everywhere, then there exists a stationary policy that is optimal.*

Proof: Let π be an optimal policy. Let the initial system state be s. Let the sample path be $\{\boldsymbol{A}(t), R(t)\}, \ t \geq 1$. Let us denote $u^{\pi}(t)$ as the action chosen by policy π in slot t. Let us shift the optimal policy to the left by one slot and denote it as π'. So, π' is such that $u^{\pi'}(t) = u^{\pi}(t+1), \ \forall t \geq 0$. Next, let us shift the packet and energy arrivals to the left by one slot and let them be $\{\boldsymbol{A}'(t), R'(t)\}, \ t \geq 1$. So $\boldsymbol{A}'(t) = \boldsymbol{A}'(t+1)$ & $R'(t) = R(t+1), \forall t \geq 1$. At slot 1, let the system state under policy π be $s' = (\boldsymbol{Q}^{\pi}(1), E^{\pi}(1))$. If the system starts at state s', with arrivals $\{\boldsymbol{A}'(t), R'(t)\}, \ t \geq 1$ and under policy π', then

$$\sum_{i=1}^{N} Q_i^{\pi'}(t) = \sum_{i=1}^{N} Q_i^{\pi}(t+1), \quad \forall t \geq 0. \tag{3}$$

Since π is optimal for every sample path, π is also optimal for the the shifted packet and energy arrivals $\{\boldsymbol{A}'(t), R'(t)\}, \ t \geq 1$. Hence from Eq. 3, it follows that π' is also optimal. Thus if optimal policy exists, then there exists a stationary policy which is optimal. ■

Lemma 2. *Optimal stationary policy belongs to the class of non-idling policies.*

Proof: Suppose a policy π_1 which does not belong to non-idling policies is optimal. π_1 has atleast one system state s such that $\mathcal{C}_s \neq \emptyset$, where it idles without choosing any of the connected queues. Let π_2 be a non-idling policy same as policy π_1 except at state s, where it chooses any one connected queue. Then if the system starts at state s, then $\sum_{i=1}^{N} Q_i^{\pi_2}(1) = \sum_{i=1}^{N} Q_i^{\pi_1}(1) - 1$. So $\sum_{i=1}^{N} Q_i^{\pi_1}(1) > \sum_{i=1}^{N} Q_i^{\pi_2}(1)$. Hence a contradiction. So, any policy which idles cannot be an optimal policy. ■

Now let us the define the notion of Lowest index connected queue policy, which is used often hereafter.

Definition 5 (Lowest Index Connected Queue (LICQ) Policy)**.** *The non-idling policy which chooses the con-*

nected queue with the lowest power requirement is referred as Lowest Index Connected Queue policy. $u^*(s) = \min \mathcal{C}_s, \ \forall s$ *such that* $|\mathcal{C}_s| > 0$.

Now we state and prove the third of the four lemmas.

Lemma 3. *Among class of non-idling policies any policy other than the Lowest Index Connected Queue (LICQ) policy is not optimal*

Proof: Let π_1 be an optimal non-idling policy that is different from LICQ policy. Then there exists a state s, such that $\mathcal{C}_s \neq \emptyset$ and $|\mathcal{C}_s| > 1$, in which π_1 does not choose the lowest index connected queue. Lets assume that the system starts in state s. Let $\min \mathcal{C}_s = i$. Let π_{LICQ} be the LICQ policy. Then π_{LICQ} chooses i whereas the other policy π_1 chooses another connected queue, say j, such that $j > i$. Then, energy remaining in policy π_2 is $E(0) - P_i$, which can be written as $kP_1 \leq E(0) - P_i < (k+1)P_1$ for some integer k. From slot 1 till slot k, assume packet arrivals to be $[1 \ 0 \ \dots \ 0]^T$ and zero energy arrivals. In slots 1 to k, π_{LICQ} transmit a packet from queue 1 whereas π_1 may transmit from any of the connected queues. At the end of k^{th} slot, energy remaining in the battery under π_1 is strictly smaller than π_{LICQ}. Moreover, the energy remaining in π_{LICQ} is smaller than P_1. At slot k+1, assume packer arrival be $[1 \ 0 \ \dots \ 0]^T$ and energy arrival be $R(k+1) = P_1 - (E(0) - P_i - kP_1)$. So $E^{\pi_{LICQ}}(k+1) = P_1$ and $E^{\pi_1}(k+1) < E^{\pi_{LICQ}}(k+1)$. So, LICQ policy π_{LICQ} chooses queue 1, whereas policy π_1 stays idle in slot $k+1$. So $\sum_{i=1}^{N} Q_i^{\pi_{LICQ}}(k+1) \leq \sum_{i=1}^{N} Q_i^{\pi_1}(k+1) - 1$. So $\sum_{i=1}^{N} Q_i^{\pi_1}(k+1) > \sum_{i=1}^{N} Q_i^{\pi_{LICQ}}(k+1)$. Hence a contradiction. Thus we have shown an example of packet and energy arrivals where every non-idling policy other than LICQ policy fails to attain backlog optimality everywhere. ∎

Intuitively, choosing LICQ i.e., queue with the lowest power requirement seems optimal as it retains the most energy in the battery for future transmissions.

Lemma 4. *Lowest Index Connected Queue (LICQ) policy is not an optimal policy.*

Proof: Let us consider a system with $N = 2$. Let the policy π_{LICQ} be optimal. Let us consider π_2 as an idling policy, which transmits packets only from queue 1 and stays idle if queue 1 is not connected. We show an example where policy π_2 achieves lesser backlog than the LICQ policy. Initial state is $(\boldsymbol{Q(0)}, E(0)) = ([0,1]^T, P_2)$. From slot 1 till slot $k-1$, energy arrival, $R(t) = P_2$ and packet arrivals, $A(t) = [0,1]^T$. For all slots greater than k-1, the energy arrival, $R(t) = 0$ and packet arrivals, $A[t] = [1,0]^T$.

TABLE I. ACTION AT EACH SLOT (QUEUE FROM WHICH PACKET IS SCHEDULED)

t	0	1	–	k-1	k	k+1	–	$k + \frac{kP_2}{P_1} - 1$
$\pi_{LICQ}(t)$	2	2	–	2	0	0	–	0
$\pi_2(t)$	0	0	–	0	1	1	–	1

$$\sum_{i=1}^{N} Q_i^{\pi_{LICQ}} \left(k \left(1 + \frac{P_2}{P_1} \right) \right) - \sum_{i=1}^{N} Q_i^{\pi_2} \left(k \left(1 + \frac{P_2}{P_1} \right) \right) = k \left(\frac{P_2}{P_1} - 1 \right)$$

We have shown that there exists packet arrivals and energy arrivals under which LICQ policy does not minimize backlog everywhere and hence it is non-optimal in this sense. ∎

Remark 1. *The example can be generalized to any number of queues.*

Remark 2. *Difference in the backlog under LICQ policy and that under policy* π_2 *is* $k\left(\frac{P_2}{P_1} - 1\right)$ *and can become unbounded as* k *increases. However, battery capacity needs to be at least* kP_2*. Hence under the given example, the difference between the backlog increases if the battery is scaled appropriately.*

Remark 3. *Even within the class of non-idling policies, it can be shown that LICQ policy is not backlog optimal everywhere. In fact the backlog under* π_{LICQ} *can grow arbitrarily larger than a non-idling policy along some sample path. Kindly refer to the technical report [20] for the example, where backlog under* π_{LICQ} *grows arbitrarily large. Thus,* π_{LICQ} *policy is not even bounded distance away from optimal.*

Proof of Theorem 1: From Lemma 1, we know that if an optimal policy exists, then there exists a stationary policy that is backlog optimal everywhere. In Lemma 2 and Lemma 3, we have shown that if an optimal stationary policy exists, then the optimal policy must be the LICQ policy. Finally, we show that LICQ policy is not backlog optimal everywhere. Hence as a consequence of the four lemmas, it is proven that there does not exist an optimal policy that achieves backlog optimality everywhere. ∎

This motivates us to construct policies that are backlog optimal in the expected sense. In the following section, we present our approach in detail.

IV. MDP Formulation

In this section, the problem of minimizing backlog is formulated as a discounted infinite horizon Markov decision process (MDP) problem. For a user i, the arrival process $\{A_i(t)\}_{t \geq 1}$ is assumed to be independent and identically distributed (i.i.d). The arrival processes for different users are assumed to be independent. Also for simplicity we assume that $A_i(t) \in \{0,1\}$ for every i and t. The recharge process is $\{R(t)\}_{t \geq 0}$ is i.i.d.

Let e_{max} be the maximum value of recharge arrival. Assume $e_{max} < P_1$. We define the system state to be $s = (\boldsymbol{q}, \xi)$, where $\boldsymbol{q}$ denotes the number of packets present in the queues and ξ denotes the energy present in the battery. Note that the state space $\mathcal{S}$ is $N+1$ dimensional. An action chosen by a policy in any state could be either to remain idle or to schedule from a connected queue. Thus, in a state $(\boldsymbol{q}, \xi)$, possible actions are $U(\boldsymbol{q}, \xi) = \{0\} \cup \{i : q_i > 0 \text{ and } \xi \geq P_i; i \in \{1, \ldots, N\}\}$. Action $u = 0$ implies that no queue is scheduled and it is possible in every state. Union of all action spaces are $U = \{0, \ldots, N\}$. We assume that the queue buffer capacity to be large, but finite. The reward function $r : \mathcal{S} \times U \to \Re_+$ is

$$r(\boldsymbol{q}, \xi, u) = \sum_{i=1}^{N} q_i$$

Let us consider $\lambda \in (0,1)$ as a discount factor, $u^\pi(t)$ is the queue scheduled by policy π in slot t. We refer to $u^\pi(t)$ as the action taken by policy π in slot t. Let us define the cost function of policy π, $J^\pi : \mathcal{S} \to R_0^+$ for the state $(\boldsymbol{q}, \xi)$ that we start with.

$$J^\pi(\boldsymbol{q}, \xi) = \lim_{T\to\infty} E\left[\sum_{t=0}^{T} \lambda^t r(\boldsymbol{Q^\pi(t)}, E^\pi(t), u^\pi(t)) \middle| S_0 = (\boldsymbol{q}, \xi)\right]$$
$$= E\left[\sum_{t=0}^{\infty} \lambda^t \sum_{i=1}^{N} Q_i^\pi(t)\right]$$

Note that since queue is finite, reward is finite and hence limit and expectation can be interchanged. Now, let us define the notion of expected backlog optimality.

Definition 6 (Expected backlog optimality)**.** *A scheduling policy π is expected backlog optimal if it satisfies the following relation,*

$$J^\pi(\boldsymbol{q}, \xi) \leq J^{\pi'}(\boldsymbol{q}, \xi) \quad \forall \boldsymbol{q}, \xi \tag{4}$$

So, here our objective is to minimize the queue backlog at the transmitter. Let us define p_e as the probability of energy arrival value being e, with e assumed to be discrete valued. Let $\alpha_{(\boldsymbol{q},\boldsymbol{q}')}$ be the transition probability from queue state $\boldsymbol{q}$ to $\boldsymbol{q}'$. $p_{\boldsymbol{a}}$ is the probability of packet arrival being $\boldsymbol{a}$ with $\boldsymbol{a} = [a_1, \ldots, a_N]^T$; $a_i \in \{0,1\}$. $\mathbb{I}_u$- $[0, \ldots, 0, 1, 0 \ldots, 0]^T$ is a $N \times 1$ vector with 1 in the u^{th} position, zeros elsewhere; $\mathbb{I}_0$ =zero vector. The optimal reward function satisfies the Bellman's equation of dynamic programming, given by

$$\begin{aligned} J^*(\boldsymbol{q}, \xi) = & \min_{u \in U(\boldsymbol{q},\xi)} \Big\{ r(\boldsymbol{q}, \xi, u) \\ & + \lambda \sum_{e=0}^{\xi_m - \xi + P_u - 1} p_e \sum_{\boldsymbol{q}'} \alpha_{(\boldsymbol{q},\boldsymbol{q}')} J^*(\boldsymbol{q}', \xi - P_u + e) \\ & + \lambda P(e \geq \xi_m - \xi + P_u) \sum_{\boldsymbol{q}'} \alpha_{(\boldsymbol{q},\boldsymbol{q}')} J^*(\boldsymbol{q}', \xi_m) \Big\} \\ = & \min_{u \in U(\boldsymbol{q},\xi)} \Big\{ r(\boldsymbol{q}, \xi, u) \\ & + \lambda \sum_{e=0}^{\xi_m - \xi + P_u - 1} p_e \sum_{\boldsymbol{a}} p_{\boldsymbol{a}} J^*(\boldsymbol{q} - \mathbb{I}_u + \boldsymbol{a}, \xi - P_u + e) \\ & + \lambda P(e \geq \xi_m - \xi + P_u) \sum_{\boldsymbol{a}} p_{\boldsymbol{a}} J^*(\boldsymbol{q} - \mathbb{I}_u + \boldsymbol{a}, \xi_m) \Big\} \end{aligned}$$

At each epoch, the policy maps the state to its optimal action. Since, it is a infinite horizon problem with discounted rewards and state space is finite, we know from [21] that, there exists a stationary deterministic policy which attains optimality. Let $\pi^* = \{u^*, u^*, \ldots\}$ represent the optimal stationary deterministic policy. The optimal action at each state is given by,

$$\begin{aligned} u^*(\boldsymbol{q}, \xi) = & \arg \min_{u \in U(\boldsymbol{q},\xi)} \Big\{ r(\boldsymbol{q}, \xi, u) \\ & + \lambda \sum_{e=0}^{\xi_m - \xi + P_u - 1} p_e \sum_{\boldsymbol{q}'} \alpha_{(\boldsymbol{q},\boldsymbol{q}')} J^*(\boldsymbol{q}', \xi - P_u + e) \\ & + \lambda P(e \geq \xi_m - \xi + P_u) \sum_{\boldsymbol{a}} p_{\boldsymbol{q}'} J^*(\boldsymbol{q}', \xi_m) \Big\} \end{aligned}$$

In the next section we obtain structural properties of the optimal policy.

V. Structural properties of optimal policy

In our first result, we show that when available energy level is high, the optimal policy is non-idling.

Theorem 2 (Work Conservation of Optimal Policy)**.** *There exists an energy threshold ξ_{th} such that for every state $s = (q, \xi)$ such that $\xi > \xi_{th}$ and $\mathcal{C}_s > 0$, then the optimal action $u^*(s) \neq 0$.*

Proof: Let us assume a state s such that $\xi > \xi_{th}$ and suppose $u^*(s) = 0$. Let the system start with state s. Let π_1 be non stationary policy which chooses action 0 in slot 0. Let π_2 be an optimal policy. Let us compare between actions 0 and N. There are two possible

TABLE II. Action at each slot

slot	0	1	2	-	m	m+1	m+2	-	$m + \frac{P_N}{P_1}$
π_1	0	N	N	–	N	1	1	–	1
π_2	N	N	N	–	N	0	0	–	0

explanations, according to the nature of recharge values. They are as follows:

a)Let us consider recharge values to be 0 from slot 0 till slot $m + \frac{P_N}{P_1}$. Here m would be $\frac{\xi}{P_N} - 1$ as shown in Table II. At slot $m+1$, Energy in policy π_2 becomes zero. At the end of slot $m + \frac{P_N}{P_1}$, Energy in policy π_1 is also zero. Policy π_2 transmits m packets from queue N until battery gets drained. Since, policy π_1 has transmited only $m-1$ packets in the same number of slots, it has P_N energy more than π_1. Now if π_1 transmits from first queue, it can transmit many packets and reduce the backlog. So this is the only way policy π_1 can minimize backlog better than policy π_2. So when this happens, we show that if energy is greater than some threshold, policy π_1 can never be better.

$$J^{\pi_1}(\boldsymbol{q},\xi) - J^{\pi_2}(\boldsymbol{q},\xi) \geq \overbrace{\lambda + \lambda^2 + \ldots + \lambda^{\frac{\xi}{P_N}}}^{\text{term1}} - \underbrace{\frac{P_N - P_1}{P_1}\lambda^{\frac{\xi}{P_N}+2}(\sum_{k=0}^{\infty}\lambda^k)}_{\text{term2}}$$

For (term1 $-$ term2) $> 0, \xi > P_N\left(\frac{\rho(N)}{\log\lambda} - 1\right) \geq \xi_{th}$ (5)

where,

$$\rho(j) \triangleq \log\left(\frac{\lambda}{1+\lambda\left(\frac{P_j}{P_1} - 1\right)}\right).$$

b) When recharge values from slot 0 are non zero, term 1 in the above equation may increase, term 2 may decrease, so eventually value difference increases.

Note that, when we compare with action $j < N$, then the threshold value obtained will be less than that of action N.

$$\text{For } \xi > P_N\left(\frac{\rho(N)}{\log\lambda} - 1\right) \geq \xi_{th}, \quad u^*(s) \neq 0 \qquad (6)$$

So, if energy is greater than this ξ_{th}, then optimal action at this state s, $u^*(s) \neq 0$. ■

Above some energy threshold, it is never optimal to stay idle. Only reason for which a policy may want to stay idle is to wait for packets to arrive in a lower indexed queue rather than transmitting a packet now from a higher indexed queue which may require a lot more energy. However, when enough energy is available, it becomes more prudent to transmit a packet to reduce cost now rather than conserving energy for future potential cost reduction. The value of ξ_{th} depends on discount factor λ.

Let us assume that the battery capacity $\xi_m > \xi_{th} + P_N$. In the next result, we show that the optimal action is to either remain idle or it follows π_{LICQ} i.e. it transmits from the lowest index connected queue. Formally we show the following.

Theorem 3. *At a state $s = (\boldsymbol{q},\xi)$ such that $|\mathcal{C}_s| > 1$ and if $u^*(s) \neq 0$, then the optimal action is to choose the LICQ. $u^*(s) = \min\mathcal{C}_s = i$.*

Proof: Let π_1 be an optimal policy. Suppose there exists a state s such that $u^*(s) = j$, even when $\min\mathcal{C}_s = i$. Let the system start with state s. Let π_2 be a non-stationary policy, which chooses action i at slot 0. As a consequence of Theorem 2, whenever battery level in policy π_1 crosses ξ_{th}, it transmits and since $e_{max} < P_1$, battery level in π_1 nevers reaches ξ_m in any sample path. Since we know that $\xi_m > \xi_{th} + P_N \geq \xi_{th} + (P_j - P_i) + e_{max}$, energy level under policy π_2 as well does not reach ξ_m. If optimal policy π_1 chooses action i in some slot, say t' as shown in Table III, then in slots 1 to $t'-1$, policy π_2 chooses same actions as optimal policy π_1. In slot t', policy π_1 chooses action i and $E^{\pi_2}(t') = E^{\pi_1}(t') + P_j - P - i$, π_2 chooses action j. Hence from slot $t'+1$, the queue state and energy state are same for both policies π_1 and π_2 and their rewards become equal. It is possible that, optimal policy π_1 may never choose action i. In that case, from slot 1 as energy under policy π_2 is higher, it can do better or as good as policy π_1. By choosing action i in slot 0, there exists a policy which is better or atleast as good as policy π_1. Hence, the optimal action at state s, $u^*(s) = \min\mathcal{C}_s = i$.

TABLE III.

Slot	0	1	-	-	-	t'
π_1	j	u_2	u_3	$-$	$-$	i
π_2	i	u_2	u_3	$-$	$-$	j

■

Spending lower power saves more energy in battery, so more packets can be transmitted in future and hence backlog is lesser when compared to transmitting from any other connected queue.

In the next result we show that if queue 1 is connected then the optimal action is to choose 1.

Theorem 4. *At a state $s = (\boldsymbol{q},\xi)$ such that $1 \in \mathcal{C}_s$, then the optimal action at this state, $u^*(s) = 1$.*

Proof: Let π_1 be an optimal policy. Suppose there exists a state s such that $u^*(s) = 0$, even when $\min\mathcal{C}_s = 1$. Let the system start with state s. Let π_2 be a non-stationary policy, which chooses action 1 at slot 0. As a consequence of Theorem 2, battery level in policies π_1 and π_2 nevers reaches ξ_m in any sample path. Let $\tilde{t} \geq 1$ be the first instance when policy π_1 chooses to transmit a packet from a connected queue, say x. Note that if optimal policy decides never to transmit a packet in any slot, in that case, policy π_2 has lesser reward than π_1 and hence π_2 is better than π_1. So, when $\tilde{t}$ exists, policy π_2 stays idle in slots 1 to $\tilde{t}$. If optimal policy π_1 chooses action 1 in some slot, say t' as shown in Table IV, then in slots $\tilde{t}+1$ to $t'-1$, policy π_2 chooses same actions as optimal policy π_1. In slot t', policy π_1 chooses action 1 and π_2 chooses action x. Hence from slot $t'+1$, the queue state and energy state are same for both policies π_1 and π_2 and their rewards become equal. It is possible that, optimal policy π_1 may never choose action 1. In that case, from slot 1 as energy under policy π_2 is higher, it can do better or atleast as good as policy π_1. At slot

TABLE IV.

slot	0	1	-	-	t	-	-	-	t'
π_1	0	0	–	0	x	u_1	u_2	–	1
π_2	1	0	–	0	0	u_1	u_2	–	x

$\tilde{t}$, $J^{\pi_1} = J^{\pi_2} + (1 + \lambda + \lambda^2 + \cdots + \lambda^{\tilde{t}-1})$. Between slots $\tilde{t}$ and t', rewards are same. After t', the state is the same in both π_1 and π_2 and the rewards will be equal. So, $J^{\pi_2} < J^{\pi_1}$. Hence, a contradiction. Optimal action $u^*(s)$ is not 0. So, when $u^*(s) \neq 0$, we know from Theorem 3, the optimal action $u^*(s) = \min \mathcal{C}_s = 1$. ■

This follows the Theorem 3 in which there is a necessity for knowing the optimal action to be non zero, whereas here in this theorem we characterize it completely, without any apriori knowledge about the optimal action, that the optimal action is 1, whenever queue 1 is connected.

As a consequence of Theorem 3 and 4, we show that if only non idling policies are allowed, then π_{LICQ} is optimal.

Corollary 1. *Among the class of non-idling policies, the policy that chooses the connected queue with the lowest index i.e., LICQ policy is expected backlog optimal.*

Proof: Under a non idling policy, whenever there is a connected queue, the action is not 0. So, based on the proof of Theorem. 3, it is observed that, transmitting a packet from a connected queue with the lowest power requirement i.e., Lowest Index Connected Queue (LICQ) is better than transmitting from any other connected queue. So, LICQ policy is an expected backlog optimal policy among this class of non-idling policies. ■

In this special class of policies, we have completely characterized an optimal policy that minimizes the expected backlog.

VI. Simulation Results

In this section, the results from the simulation have been described. The scenario is simulated and performance of LICQ policy is shown with respect to different metrics. The simulation parameters are as follows. The number of users are $N = 3$. The power required to transmit a packet from the queues are $[P_1\ P_2\ P_3] = [4\ 6\ 9]$ respectively. Number of slots are 100,000 over which the simulations are carried out. The packet arrivals are Bernoulli process with values 0 and 1 with mean arrival rate $\alpha = [0.1\ 0.1\ 0.1]$. The recharge energy arrivals are of Poisson distribution with mean $\bar{E}$. The battery capacity ξ_m is assumed to be 50 units. Note that the simulations are carried out without the assumptions that $e_{max} < P_1$ and finite queue buffer, which were required for analytical guarantees. Also notice that on account of infinite state space computation of optimal policy through methods like policy iteration and value iteration is not possible. Hence, we simulate the performance of LICQ policy,

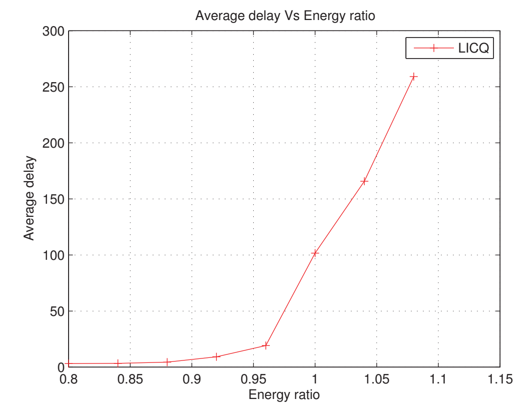

Fig. 2. Average Delay vs Energy Ratio

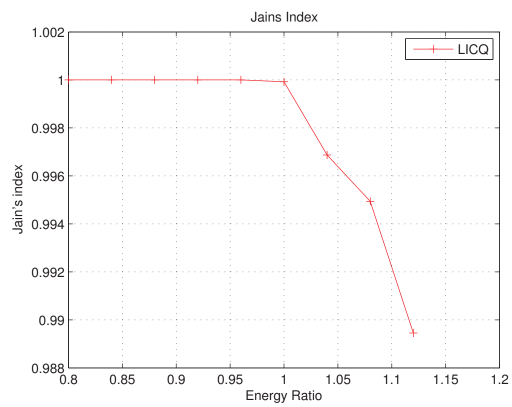

Fig. 3. Jain Index vs Energy ratio

which was shown to be optimal in the class of non idling policies.

Let us define energy ratio to be $\frac{\alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3}{\bar{E}}$. Note that energy ratio is equivalent to Erlang load on the energy queue. In Fig 2 we plot average delay in network as function of energy ratio. As expected, the average delay increases with energy ratio. In Figure 3, using Jain's index, we investigate fairness of the LICQ policy, in terms of delay for various users. It can be shown that, as energy ratio increases the fairness goes down as most of the times LICQ transmits from lowest index queue.

VII. Conclusion

Under finite battery setup in minimizing expected backlog, structural properties of an optimal policy has been proved. Importantly, we have shown that above some threshold in battery energy, it is optimal to transmit, rather than staying idle. Among class of non-idling policies, the policy that schedules the connected queue with the lowest index (LICQ), i.e lowest power requirement is optimal. Hence under this special class of policies, optimal policy is completely characterized. But, the same LICQ policy is not backlog optimal everywhere and is justified via a counter example. From the analysis of backlog optimality at every slot, it can be inferred that with energy being a random value, an optimal policy does not exist.

ACKNOWLEDGEMENT

This work is supported by the India-UK Advanced Technology Centre of Excellence in Next Generation Networks, Systems and Services (IU-ATC) and funded by the Department of Science and Technology (DST), Government of India.

REFERENCES

[1] O. Ozel, J. Yang, and S. Ulukus, "Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 6, pp. 2193–2203, June 2012.

[2] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 8, pp. 1732–1743, Sept 2011.

[3] M. Khuzani and P. Mitran, "On online energy harvesting in multiple access communication systems," *Information Theory, IEEE Transactions on*, vol. 60, no. 3, pp. 1883–1898, March 2014.

[4] O. Ozel and S. Ulukus, "Information-theoretic analysis of an energy harvesting communication system," in *Personal, Indoor and Mobile Radio Communications Workshops (PIMRC Workshops), IEEE 21st International Symposium on*, Sept 2010.

[5] ——, "Achieving awgn capacity under stochastic energy harvesting," *Information Theory, IEEE Transactions on*, vol. 58, no. 10, pp. 6471–6483, Oct 2012.

[6] W. Mao and B. Hassibi, "On the capacity of a communication system with energy harvesting and a limited battery," in *Information Theory Proceedings (ISIT), IEEE International Symposium on*, July 2013.

[7] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 3, pp. 1180–1189, March 2012.

[8] J. Yang and S. Ulukus, "Transmission completion time minimization in an energy harvesting system," in *Information Sciences and Systems (CISS), 44th Annual Conference on*, March 2010.

[9] R. Serfozo, "Monotone optimal policies for markov decision processes," in *Stochastic Systems: Modeling, Identification and Optimization, II*, ser. Mathematical Programming Studies, R.-B. Wets, Ed. Springer Berlin Heidelberg, 1976, vol. 6, pp. 202–215.

[10] L. Tassiulas and A. Ephremides, "Dynamic scheduling for minimum delay in tandem and parallel constrained queueing models," *Annals of Operations Research*, vol. 48, no. 4, pp. 333–355, 1994.

[11] ——, "Dynamic server allocation to parallel queues with randomly varying connectivity," *Information Theory, IEEE Transactions on*, vol. 39, no. 2, pp. 466–478, Mar 1993.

[12] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," *Information Theory, IEEE Transactions on*, vol. 48, no. 5, pp. 1135–1149, May 2002.

[13] M. Goyal, A. Kumar, and V. Sharma, "Power constrained and delay optimal policies for scheduling transmission over a fading channel," in *INFOCOM 2003, IEEE*, vol. 1, March 2003, pp. 311–320.

[14] ——, "Optimal cross-layer scheduling of transmissions over a fading multiaccess channel," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3518–3537, Aug 2008.

[15] P. Chaporkar and S. Sarkar, "Minimizing delay in loss-tolerant mac layer multicast," *Information Theory, IEEE Transactions on*, vol. 52, no. 10, pp. 4701–4713, Oct 2006.

[16] N. Salodkar, A. Bhorkar, A. Karandikar, and V. Borkar, "An online learning algorithm for energy efficient delay constrained scheduling over a fading channel," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 4, pp. 732–742, May 2008.

[17] M. Neely, "Order optimal delay for opportunistic scheduling in multi-user wireless uplinks and downlinks," *Networking, IEEE/ACM Transactions on*, vol. 16, no. 5, pp. 1188–1199, Oct 2008.

[18] R. Vaze, R. Garg, and N. Pathak, "Dynamic power allocation for maximizing throughput in energy-harvesting communication system," *Networking, IEEE/ACM Transactions on*, vol. PP, no. 99, pp. 1–1, 2013.

[19] R. Vaze, "Competitive ratio analysis of online algorithms to minimize packet transmission time in energy harvesting communication system," in *INFOCOM, Proceedings IEEE*, April 2013.

[20] V. V, P. Chaporkar, and A. Karandikar, "Minimizing backlog for downlink of energy harvesting networks," Department of Electrical Engineering, IIT Bombay, Mumbai, India, Tech. Rep., 2015. [Online]. Available: http://ticet.iitb.ac.in/documents/venkhat_report.php

[21] M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 1st ed. John Wiley & Sons, Inc., 1994.