Online Transmission Policies for Cognitive Radio Networks with Energy Harvesting Secondary Users

Burak Varan Aylin Yener
Wireless Communications and Networking Laboratory
Electrical Engineering Department
School of Electrical Engineering and Computer Science
The Pennsylvania State University, University Park, PA 16802
varan@psu.edu yener@ee.psu.edu

***Abstract*—In this paper, we consider a cognitive radio network with energy harvesting secondary users and orthogonal channels for the primary users. Secondary users have finite capacity batteries to store harvested energy for future use. Secondary users can use the primary users' channel and transmit their data if the channel is not occupied by the primary user. We define an instantaneous reward for the secondary users which is equal to the throughput if the channel is available. Otherwise, we assign a negative reward to the secondary user as a penalty. We formulate the long term average reward maximization problem for each secondary user. We extend the fixed fraction policy in [1] to our setup and show that it yields near optimal performance. We derive an upper bound on the long term average reward, as well as additive and multiplicative lower bounds. We observe that, for energy arrivals that are frequent and when secondary users have several channel options to choose from, the secondary users are better off spending all available energy immediately, i.e., a greedy policy is a better choice whereas the fixed fraction policy is better when the energy arrivals are scarce.**

I. Introduction

Cognitive radio networks are composed of wireless nodes that can detect available channels and access them to send their data. The legitimate owners of these channels are called the primary users and cognitive radio nodes are often secondary users. The secondary users sense the primary users' channels and use the ones that they believe to be unoccupied by the primary user. If a secondary user attempts to transmit over an occupied channel, there is a collision and the primary user may not be able to use its channel either. Naturally, collisions are undesirable. We consider such a network where the secondary users are energy harvesting nodes, i.e., they acquire the energy needed for their operation from nature in an intermittent fashion [2]. Further, we consider the case where the nodes have only causal knowledge of the energy amounts they harvest, i.e., the online case.

Cognitive radio networks with energy harvesting have been studied in the recent wireless communications literature. Reference [3] has studied secondary users that can harvest energy from the primary users' transmission and use the harvested energy to send their own data via the unoccupied channels. Reference [4] has proposed an optimal spectrum sensing policy for energy harvesting secondary users with infinite batteries. Reference [5] has studied the tradeoff between false alarms and misdetections in the secondary users' sensing decisions. Reference [6] has studied secondary users that can harvest energy from the primary users' transmission while not accessing the primary users' spectrum. Other related references include [7]–[11].

This work was supported in part by the National Science Foundation Grants CCF 1422347 and CNS 1526165.

A reference that is closely related to our work is [1]. In this paper, an energy harvesting transmitter is studied and the long term throughput maximization problem is formulated. A near optimal policy, named the fixed fraction policy, is proposed and shown to be within additive and multiplicative gaps. In this work, we extend this policy to cognitive radio networks with energy harvesting secondary users. We propose to use the near optimal power values found by the fixed fraction policy for the single user case as *sum power constraints* for each secondary user. We show that each secondary user can partition the sum power between several available channels and obtain a reward that is within additive and multiplicative gaps to optimality. We employ a reward function that penalizes incorrect estimates of availability of the primary users' channels. We compare the fixed fraction policy with the greedy policy which spends all available energy in each time slot, and find that the greedy policy can outperform the fixed fraction policy in cases where frequency and energy resources are not scarce.

II. System Model

Consider a cognitive radio network with N primary users (PU_1, PU_2, $\dots$, PU_N) and M secondary users (SU_1, SU_2, $\dots$, SU_M). The primary users have orthogonal links to their respective receivers (PR_1, PR_2, $\dots$, PR_N), i.e., they can communicate with their receivers without interference if they have any data to transmit. When a primary user does not have any data to transmit, a secondary user can utilize the available spectrum to communicate with its receiver. Secondary users can sense a primary user's channel to determine whether it is being used. They can then contend for a channel which they believe to be available, i.e., not occupied by the primary user. If several secondary users contend for a primary user's channel, each is equally likely to acquire the channel and transmit, while the rest back off from that channel. If a secondary user incorrectly estimates a primary user's channel to be available and attempts to transmit, it will incur a penalty. The penalty

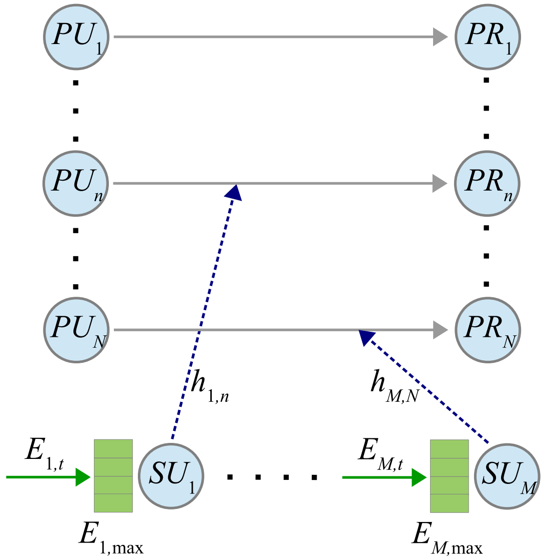

Fig. 1. Cognitive radio network, dashed lines represent channel access by secondary users.

is in the form of a negative instantaneous reward and will be elaborated on in the sequel.

The secondary users are energy harvesting nodes. We consider a time slotted communication scenario where the time slots are of duration 1 second, without loss of generality. Secondary user m receives $E_{m,t}$ units of energy in time slot t and has an energy storage device of capacity $E_{m,\max}$. We consider an independent and identically distributed (i.i.d.) non-negative random process for the energy arrivals at all secondary users. That is, $E_{m,t} \geq 0$, for all m and t, is the realization of random variable $\mathbf{E}_{m,t}$, and $\mathbf{E}_{m,t}$ and $\mathbf{E}_{m,t'}$ are i.i.d. for all m and $t \neq t'$. Let $p_{m,t}$ denote the energy consumption of secondary user m in time slot t and $B_{m,t}$ the energy in its battery at the end of time slot t. We have that

$$B_{m,t} = \min\{B_{m,t-1} + E_{m,t} - p_{m,t}, E_{m,\max}\} \tag{1}$$

which describes the evolution of the battery state $B_{m,t}$. In the sequel, we will use $B_{m,t}$ to constrain the energy expenditure of secondary user m as

$$p_{m,t} \leq B_{m,t-1} + E_{m,t}. \tag{2}$$

That is, secondary user m has $B_{m,t-1}$ units of energy in its battery at the beginning of time slot t. It then receives $E_{m,t}$ units and has $B_{m,t-1} + E_{m,t}$ units to spend. The remaining $B_{m,t-1} + E_{m,t} - p_{m,t}$ units will be stored in the battery up to $E_{m,\max}$, see (1).

We denote by $h_{m,n}$ the channel power gain that secondary user m experiences it if uses primary user n's channel. Each secondary user has three decisions to make for each time slot. The first one is $a_{m,t} = 0, 1, 2, \dots, N$ which determines which channel to sense for secondary user m. If $a_{m,t} = 0$, secondary user m will not sense at all. The second decision is $c_{m,t} = 0, 1, 2, \dots, N$ that determines which channel secondary user m will contend for. If $c_{m,t} = 0$, secondary user m will not contend. The last decision is $p_{m,t} \geq 0$ which determines the transmit power for the time slot for secondary user m. Let $\psi_{m,t} = 1$ denote that the channel is available and let $\psi_{m,t} = 0$ denote that channel is occupied. Similarly, let $\phi_{m,t} = 1$ denote that the secondary user has won the contention and $\phi_{m,t} = 0$ denote that it has lost the contention. We consider an instant reward function that is based on the throughput for each time slot instead of the transmission duration to emphasize the impact of energy scheduling. We denote all action variables as $u_{m,t} \triangleq (a_{m,t}, c_{m,t}, p_{m,t})$ and all state variables as $x_{m,t} \triangleq (\psi_{m,t}, \phi_{m,t}, h_{m,n}, B_{m,t})$. For $c_{m,t} = n$, the instant reward is

$$R(x_{m,t}, u_{m,t}) = \begin{cases} 0 & \text{if } \phi_{m,t} = 0 \\ -\frac{\alpha}{2}\log(1 + h_{m,n}p_{m,t}) & \text{if } \phi_{m,t} = 1, \psi_{m,t} = 0 \\ \frac{1}{2}\log(1 + h_{m,n}p_{m,t}) & \text{if } \phi_{m,t} = 1, \psi_{m,t} = 1 \end{cases} \tag{3}$$

or more compactly

$$R(x_{m,t}, u_{m,t}) = \frac{\phi_{m,t}[(1+\alpha)\psi_{m,t} - \alpha]}{2}\log(1 + h_{m,n}p_{m,t}) \tag{4}$$

which we will use in the sequel. Here, $\alpha \geq 0$ denotes the penalty factor for the secondary users, i.e., the secondary user expends energy, but obtains a negative reward.

In the next section, we formulate the long term reward maximization problem for a secondary user in a distributed fashion and derive bounds on the optimal value of this problem using the fixed fraction power policy proposed in [1].

III. Bounds on the Long Term Reward

Let us consider a distributed scenario where each secondary user wishes to solve the long term average reward maximization problem. Let $u_m = (u_{m,1}, u_{m,2}, \dots)$ for all m. We formulate the long term average throughput maximization problem for secondary user m as

$$\max_{u_m} \ \Theta(u_m) \triangleq \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[R(x_{m,t}, u_{m,t})] \tag{5a}$$

$$\text{s.t.} \quad 0 \leq p_{m,t} \leq B_{m,t-1} + E_{m,t}, \forall t, \tag{5b}$$

$$a_{m,t}, c_{m,t} \in \{0, 1, 2, \dots, N\}, \forall t, \tag{5c}$$

$$B_{m,t} = \min\{B_{m,t-1} + E_{m,t} - p_{m,t}, E_{m,\max}\}, \forall t. \tag{5d}$$

The objective of (5) can be simplified as follows.

$$\Theta(u_m) = \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[R(x_{m,t}, u_{m,t})], \tag{6a}$$

$$= \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathbb{P}(c_{m,t} = i) \times \mathbb{E}[R(x_{m,t}, u_{m,t}) \mid c_{m,t} = i], \tag{6b}$$

$$= \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N} \mathbb{P}(c_{m,t} = i)$$

$$\times \mathbb{E}\left[\frac{\phi_{m,t}[(1+\alpha)\psi_{m,t}-\alpha]}{2}\log(1+h_{m,i}p_{m,t}) \mid c_{m,t}=i\right], \tag{6c}$$

$$= \liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\mathbb{P}(c_{m,t}=i)\mathbb{E}[\phi_{m,t} \mid c_{m,t}=i] \times [(1+\alpha)\mathbb{E}[\psi_{m,t} \mid c_{m,t}=i]-\alpha] \times \mathbb{E}\left[\frac{1}{2}\log(1+h_{m,i}p_{m,t}) \mid c_{m,t}=i\right], \tag{6d}$$

$$= \liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\mathbb{P}(c_{m,t}=i)\pi_\phi(m,t,i) \times [(1+\alpha)\pi_\psi(t,i)-\alpha]\mathbb{E}\left[\frac{1}{2}\log(1+h_{m,i}p_{m,t}) \mid c_{m,t}=i\right], \tag{6e}$$

where $\pi_\phi(m,t,i) \triangleq \mathbb{E}[\phi_{m,t} \mid c_{m,t}=i]$ is the probability that secondary user m wins the contention for channel i in time slot t. Assuming a fair contention scheme, $\pi_\phi(m,t,i)$ is 1 divided by the number of secondary users contending for channel i in time slot t for all contenders. $\pi_\psi(t,i) \triangleq \mathbb{E}[\psi_{m,t} \mid c_{m,t}=i]$ is the probability that channel i is available in time slot t. Each secondary user can estimate this probability using its current observation and its observation and action history. Equation (6d) follows from the independence of $\phi_{m,t}$, $\psi_{m,t}$, and $p_{m,t}$ given that secondary user m has chosen channel i to contend for. That is, $\phi_{m,t}$ depends only on the number of secondary users contending for channel i, $\psi_{m,t}$ depends only on whether or not primary user i is using its channel, and $p_{m,t}$ depends on the energy harvests at secondary user m. In (6e), we use the fact that $\phi_{m,t}$ and $\psi_{m,t}$ are indicators and we define the corresponding probabilities. The instant reward is zero if $c_{m,t}=0$.

Equation (6e) reveals that once the secondary user has chosen a channel to contend for, the remaining power allocation problem is the same as the long term throughput maximization problem in [1]. However, there are several channels in this problem unlike [1]. We aim to maximize the weighted sum of the long term throughputs for each channel, which is equal to the long term average reward for the secondary users. We express the long term average reward as

$$\Theta(u_m) = \liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\mathcal{T}(m,t,i) \tag{7}$$

where

$$\sigma_{i,t} \triangleq \mathbb{P}(c_{m,t}=i) \tag{8}$$

are to be designed, and

$$\gamma_{i,t} \triangleq \pi_\phi(m,t,i)[(1+\alpha)\pi_\psi(t,i)-\alpha] \tag{9}$$

are known multipliers which act as weights on the individual throughputs yielded by the available channels, which can be stated as

$$\mathcal{T}(m,t,i) \triangleq \mathbb{E}\left[\frac{1}{2}\log(1+h_{m,i}p_{m,t}) \mid c_{m,t}=i\right]. \tag{10}$$

We determine the contention probabilities $\mathbb{P}(c_{m,t}=i)$. Suppose the secondary user has an estimate for $\pi_\psi(t,i)$ and for the number of contenders. We know how to calculate the near-optimal fixed fraction policy [1] for each channel. We define $\sigma_{i,t}$ and $\gamma_{i,t}$ as in (7). From this point forward, we omit the secondary user index m since we have a single user problem. Suppose $(p_1^*, p_2^*, \dots)$ is the fixed fraction policy in [1] that is generated according to the energy arrivals at the secondary user. We can adopt this policy by limiting the average power consumption for each time slot by p_t^*. Let $\bar{p}_{i,t}$ denote the power scheduled for channel i in time slot t. We reformulate the problem as

$$\max_{\bar{p}_{i,t},\sigma_{i,t}} \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t}) \tag{11a}$$

$$\text{s.t.} \quad \bar{p}_{i,t} \geq 0,\ \forall t=1,2,\dots,T,\ i=1,2,\dots,N, \tag{11b}$$

$$0 \leq \sigma_{i,t} \leq 1,\ \forall t=1,2,\dots,T,\ i=1,2,\dots,N, \tag{11c}$$

$$\sum_{i=1}^{N}\sigma_{i,t}=1,\ \forall t=1,2,\dots,T, \tag{11d}$$

$$\sum_{i=1}^{N}\sigma_{i,t}\bar{p}_{i,t} \leq p_t^*,\ \forall t=1,2,\dots,T. \tag{11e}$$

Here, we can assume that all $\gamma_{i,t} \geq 0$ since we would have to schedule zero power to any channel with a negative reward. Equivalently, we can redefine $\gamma_{i,t} = \max\{\pi_\phi(m,t,i)[(1+\alpha)\pi_\psi(t,i)-\alpha],0\}$.

This is essentially the same problem as [12, Equation (10)] which gives the maximum achievable rate for the hybrid relaying scheme in [12]. We solve (8) in a similar way to [12]. Let $r_t(p)$ be the concavification of the following curve:

$$\tilde{r}_t(p) = \max_{i=1,2,\dots,N}\gamma_{i,t}\frac{1}{2}\log(1+h_ip). \tag{12}$$

That is, we take the maximum of all possible instantaneous rewards and concavify the resulting curve since we can improve the reward further by time sharing between several channels. We can use $r_t(p)$ with any power value p_t^* to find a convex combination of power values that satisfy the constraints of (8) and thus are feasible, yielding the multi link extension of the fixed fraction policy in [1].

To find an upper bound on the maximum of (11), let us first note that $\gamma_{i,t} \leq 1$ for all i and t since

$$\gamma_{i,t} = \max\{\pi_\phi(m,t,i)[(1+\alpha)\pi_\psi(t,i)-\alpha],0\}, \tag{13a}$$

$$= \max\{\underbrace{\pi_\phi(m,t,i)}_{\leq 1}[\underbrace{\pi_\psi(t,i)}_{\leq 1}-\underbrace{\alpha(1-\pi_\psi(t,i))}_{\geq 0}],0\}. \tag{13b}$$

We have

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t})$$

$$\leq \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t}), \tag{14a}$$

$$\leq \frac{1}{T}\sum_{t=1}^{T}\frac{1}{2}\log(1+\max_{i=1,2,\ldots,N}\{h_i\}p_t^*), \tag{14b}$$

where we choose the best channel and allocate all the power to it. Since the individual rate curves in (12) are not scaled by different $\gamma_{i,t}$ values, there is no need for concavification and

$$r_t(p)=\tilde{r}_t(p)=\frac{1}{2}\log(1+\max_{i=1,2,\ldots,N}\{h_i\}p). \tag{15}$$

Taking $\liminf_{T\to\infty}$ of both sides of (14b) and applying the upper bound in [1, Proposition 2], we obtain

$$\liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t}) \leq \frac{1}{2}\log\left(1+\left(\max_{i=1,2,\ldots,N}h_i\right)\pi_m E_m\right). \tag{16}$$

To find a lower bound on the maximum of (11), we note

$$r_t(p_t^*)\geq\tilde{r}_t(p_t^*)\geq\gamma_{i,t}\frac{1}{2}\log(1+h_ip_t^*) \tag{17}$$

for all $i=1,2,\ldots,N$ due to (12). This means that sticking to only one channel and applying the fixed fraction policy in [1] cannot improve the long term reward as compared to that achieved by the solution of (11). Consequently, we have that

$$\liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t}) \geq \liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\gamma_{i,t}\frac{1}{2}\log(1+h_ip_t^*) \tag{18}$$

for all $i=1,2,\ldots,N$. Consider a Markov structure for the availability of the primary users' channels and the contention decisions of the secondary users. Consider the following state transition matrix for the availability of channel i:

$$Z_i=\begin{bmatrix}\zeta_i' & 1-\zeta_i'\\ 1-\zeta_i'' & \zeta_i''\end{bmatrix} \tag{19}$$

where primary user i uses its channel in state 1 and does not in state 2. The steady state distribution for Z_i is found as

$$[\zeta_i \quad 1-\zeta_i]=\left[\frac{1-\zeta_i''}{2-\zeta_i'-\zeta_i''} \quad \frac{1-\zeta_i'}{2-\zeta_i'-\zeta_i''}\right]. \tag{20}$$

Similarly, we can consider a Markov model for each secondary user with $N+1$ states where the first state corresponds to the decision that the secondary user does not contend, and state n represents contention for channel $n-1$, $n=2,\ldots,N+1$. Let the steady state for this Markov model be given as

$$[\xi_{m,0} \quad \xi_{m,1} \quad \ldots \quad \xi_{m,N}] \tag{21}$$

for secondary user m, $m=1,2,\ldots,M$. After the steady states for these two Markov models have been achieved, the time dependence of the $\gamma_{i,t}$ multiplier in the throughput vanishes. Since we are looking at the long term throughput, we can express $\gamma_{i,t}$ in terms of the steady state probabilities as

$$\gamma_{i,t}\triangleq\gamma_i=\pi_\phi(m,t,i)[(1+\alpha)\pi_\psi(t,i)-\alpha], \tag{22a}$$

$$=\left(\sum_{m=1}^{M}\xi_{m,i}\right)^{-1}\left[(1+\alpha)\frac{1-\zeta_i'}{2-\zeta_i'-\zeta_i''}-\alpha\right] \tag{22b}$$

where the winning contender is chosen with equal probability for all contenders and the expected number of contenders is $\sum_{m=1}^{M}\xi_{m,i}$, and the probability for channel availability is the second probability in the steady state distribution of Z_i.

Continuing from (18), we have that

$$\liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t})$$

$$\geq \liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\gamma_i\frac{1}{2}\log(1+h_ip_t^*), \tag{23a}$$

$$=\gamma_i\liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\frac{1}{2}\log(1+h_ip_t^*), \tag{23b}$$

$$=\frac{\gamma_i}{2}\log(1+h_i\pi_mE_m)-0.72\gamma_i \tag{23c}$$

where we have used the first bound in [1, Theorem 2], and

$$\liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t}) \geq \frac{\gamma_i}{4}\log(1+h_i\pi_mE_m) \tag{24}$$

for all $i=1,\ldots,N$, where we have used the second bound in [1, Theorem 2]. Finally, we have

$$\liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t}) \geq \max_{i=1,\ldots,N}\left[\frac{\gamma_i}{2}\log(1+h_i\pi_mE_m)-0.72\gamma_i\right] \tag{25}$$

and

$$\liminf_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sigma_{i,t}\gamma_{i,t}\frac{1}{2}\log(1+h_i\bar{p}_{i,t}) \geq \max_{i=1,\ldots,N}\frac{\gamma_i}{4}\log(1+h_i\pi_mE_m) \tag{26}$$

for all $i=1,\ldots,N$.

IV. Numerical Results

In this section, we provide numerical results for the cognitive radio setup that we have studied in this work. We consider a cognitive radio network of 100 primary users and 100 secondary users. The nodes are uniformly placed on a 100 m $\times$ 100 m square. We consider a 1 MHz band for each orthogonal link, carrier frequency 900 MHz, noise density 10^{-19} W/Hz, and Rayleigh fading. We calculate the mean fading level over a distance of d meters as -40 dB$/d^2$ [13], [14]. We consider a communication session of 1000 time slots

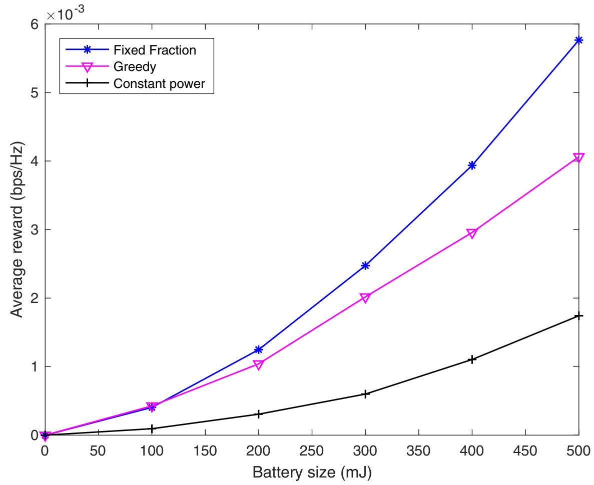

Fig. 2. Average reward per secondary user with Bernoulli(0.01) energy arrivals.

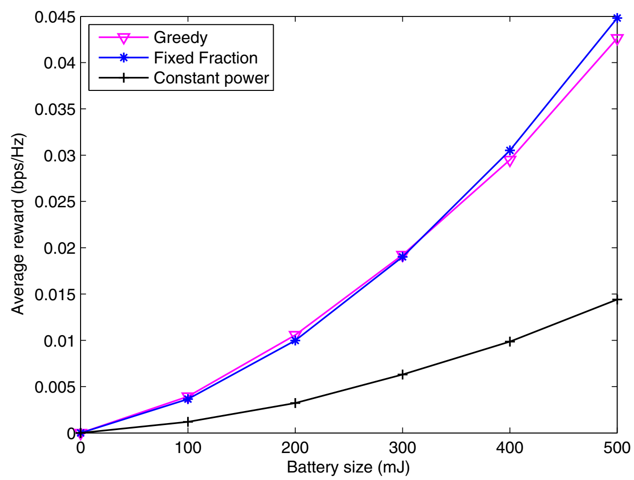

Fig. 3. Average reward per secondary user with Bernoulli(0.1) energy arrivals.

of duration 1 second each, averaging the results over 1000 runs. We compare three online power policies for this setup. The first one is the fixed fraction policy [1] which spends the $\mathbb{E}[E_{m,t}]/E_{m,\max}$ fraction of the available energy in the battery in every time slot. The other two policies are the greedy policy which spends all the energy in the battery in each time slot instead of a fixed fraction of it, and the constant power policy which transmits at the mean harvest rate only if there is enough energy in the battery.

The average reward per secondary user is shown in Figs. 2–5 for Bernoulli arrivals with positive harvest probability 0.01, 0.1, 0.5, and 0.9, respectively. When the Bernoulli parameter is low, the harvested energy becomes more intermittent and the fixed fraction policy outperforms the greedy policy if the battery is large enough, i.e., if energy can be stored without incurring too much loss. The difference grows as the battery size increases. This confirms the advantage of saving

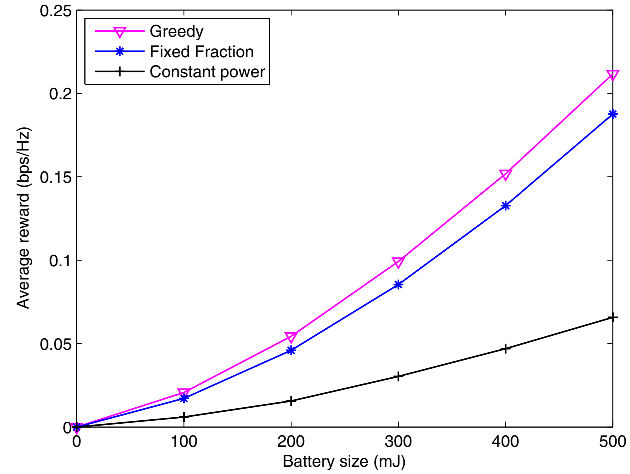

Fig. 4. Average reward per secondary user with Bernoulli(0.5) energy arrivals.

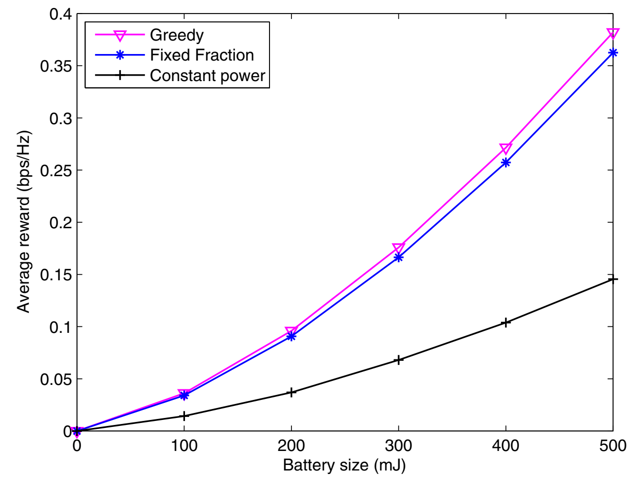

Fig. 5. Average reward per secondary user with Bernoulli(0.9) energy arrivals.

energy for future time slots with scarce energy arrivals. As the Bernoulli parameter is increased, energy becomes more abundant and the greedy policy is observed to perform better as in [1].

The simulation is repeated for uniform arrivals in Fig. 6 and for exponential arrivals in Fig. 7. These plots offer the same insights as those in [1] in that for exponential arrivals, all three policies yield similar throughput/reward values, and for uniform arrivals, the fixed fraction and greedy policies outperform the constant power policy. This is because the constant power policy is optimal for constant energy arrivals, and the uniform distribution over $[0, E_{m,\max}]$ maximizes the uncertainty in the energy arrival process.

For most simulation scenarios in consideration, we observe that the greedy policy is slightly better than the fixed fraction policy. The fixed fraction policy is suitable for conditions with sparse energy arrivals as it promotes saving the harvested

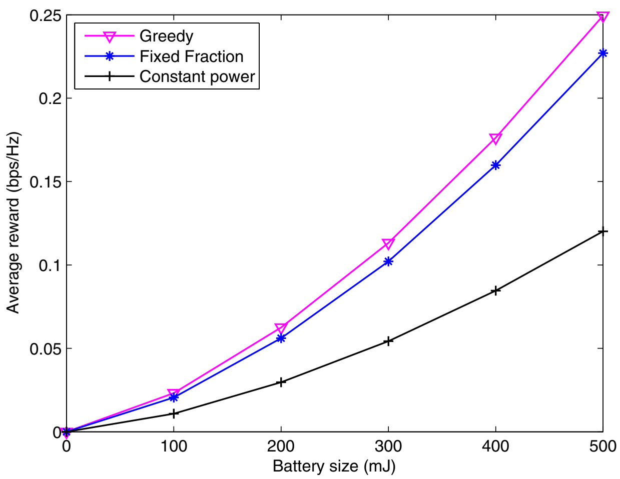

Fig. 6. Average reward per secondary user with uniform energy arrivals.

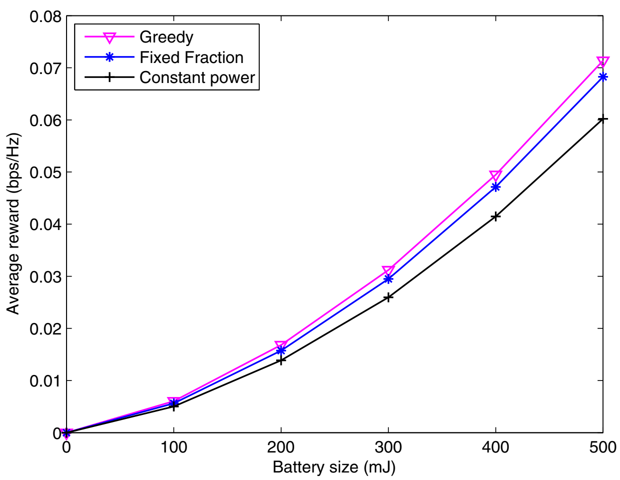

Fig. 7. Average reward per secondary user with $\exp(\frac{1}{0.1E_{m,\max}})$ energy arrivals.

energy for future. By contrast, the greedy policy is better for cases where resources are not scarce. In our cognitive radio setup, each secondary user may have several options that offer the possibility of a positive instantaneous reward compared to the single link [1]. The secondary user can choose the channel with the largest expected reward in every time slot and contend for it. To test this reasoning, we simulate a setup with fewer channels ($N = 20$) in Fig. 8 and observe the fixed fraction policy to outperform the greedy policy when fewer options are present for secondary users.

V. Conclusion

In this paper, we have studied a cognitive radio network with energy harvesting secondary users. We have extended the online fixed fraction policy that was shown to be near optimal for a single user system in [1] to cognitive radio networks. We have identified upper and lower bounds on the long term reward for the secondary users using the bounds derived in [1].

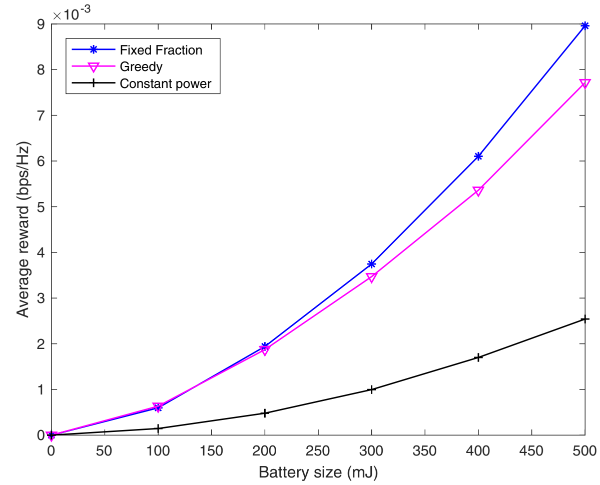

Fig. 8. Average reward per secondary user with Bernoulli(0.1) energy arrivals and $N = 20$ and $M = 100$.

We have observed that the greedy policy can outperform the fixed fraction policy under simulation scenarios without scarce energy arrivals, and when the secondary users have several channels they can access and possibly obtain a positive reward.

References

[1] D. Shaviv and A. Ozgur, "Universally near-optimal online power control for energy harvesting nodes," *arXiv preprint: 1511.00353*, Nov. 2015.
[2] O. Ozel, K. Tutuncuoglu, S. Ulukus, and A. Yener, "Fundamental limits of energy harvesting communications," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 126–132, Apr. 2015.
[3] S. Lee, R. Zhang, and K. Huang, "Opportunistic wireless energy harvesting in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4788–4799, Sep. 2013.
[4] S. Park, H. Kim, and D. Hong, "Cognitive radio networks with energy harvesting," *IEEE Transactions on Wireless Communications*, vol. 12, no. 3, pp. 1386–1397, Mar. 2013.
[5] S. Park and D. Hong, "Optimal spectrum access for energy harvesting cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 6166–6179, Dec. 2013.
[6] S. Park, J. Heo, B. Kim, W. Chung, H. Wang, and D. Hong, "Optimal mode selection for cognitive radio sensor networks with rf energy harvesting," in *IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2012, pp. 2155–2159.
[7] A. Sultan, "Sensing and transmit energy optimization for an energy harvesting cognitive radio," *IEEE Wireless Communications Letters*, vol. 1, no. 5, pp. 500–503, Oct. 2012.
[8] X. Lu, P. Wang, D. Niyato, and E. Hossain, "Dynamic spectrum access in cognitive radio networks with RF energy harvesting," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 102–110, Jun. 2014.
[9] J. P. Jeyaraj, S. S. Kalamkar, and A. Banerjee, "Energy harvesting cognitive radio with channel-aware sensing strategy," *IEEE Communications Letters*, vol. 18, no. 7, pp. 1171–1174, Jul. 2014.
[10] G. Zheng, Z. Ho, E. A. Jorswieck, and B. Ottersten, "Information and energy cooperation in cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2290–2303, May 2014.
[11] S. Yin, E. Zhang, Z. Qu, L. Yin, and S. Li, "Optimal cooperation strategy in cognitive radio systems with energy harvesting," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4693–4707, Sep. 2014.
[12] K. Tutuncuoglu, B. Varan, and A. Yener, "Throughput maximization for two-way relay channels with energy harvesting nodes: The impact of relaying strategies," *IEEE Transactions on Communications*, vol. 63, no. 6, pp. 2081–2093, Jun. 2015.
[13] A. Goldsmith, *Wireless communications*. Cambridge Univ. press, 2005.
[14] T. S. Rappaport, *Wireless communications: Principles and practice*. Prentice Hall PTR New Jersey, 1996.