RESEARCH   **Open Access**

CrossMark

Throughput maximization-based optimal power allocation for energy-harvesting cognitive radio networks with multiusers

Yuanyi Wang, Xiaohui Zhao[†] and Hui Liang[*†] (ID)

Abstract

An optimal power allocation (OPA) policy for orthogonal frequency division multiplexing (OFDM)-based cognitive radio networks (CRNs) using underlay spectrum access model is presented under multiple secondary users (SUs) with energy harvesting (EH). The proposed algorithm can allocate transmission power to each SU on each subcarrier with the objective of maximizing the average throughput of secondary network over a finite time interval. We consider both the interference power constraint limited by primary user (PU) and the minimum throughput constraint of each SU to improve the throughput of SUs while guaranteeing the communication quality of PU. To balance current throughput and expected future throughput, a dynamic programming (DP) problem is defined and solved by the backward induction method. Moreover, for each time slot, a convex immediate optimization is presented to obtain an optimal solution, which can be solved by the Lagrange dual method. Simulation results show that our policy can achieve better performance than some traditional policies and ensure good quality of service (QoS) of PU when SUs access the spectrum.

Keywords: Optimal power allocation, Energy harvesting, Cognitive radio network, Dynamic programming, Lagrange dual method

1 Introduction

Cognitive radio (CR) [1] is an optional solution to solve the spectrum shortage problem, which can both improve efficiency of spectrum utilization and ensure quality of service (QoS). In recent years, energy harvesting (EH) has obtained the widespread research attention on wireless communications, which has the ability to realize uninterrupted and self-sustainable operation [2]. In combination of EH technology and cognitive radio network (CRN), the secondary users (SUs) can collect energy from radio frequency (RF) signals such as primary channel and ambient sources such as solar by EH device [3]. Therefore, with energy harvesting device, SUs obtain renewable energy to improve energy efficiency and device flexibility.

Many researches have been conducted on optimal power allocation (OPA) policies with EH device under

wireless communication scenarios based on different considerations addressed in [4–7]. An OPA policy for EH wireless communications with limited channel feedback from receiver is investigated in [4] where the receiver periodically sends only 1-b feedback by comparing channel power gain with a predetermined threshold. In [5], a power allocation policy for an access-controlled transmitter with EH capability based on causal observations of channel fading state is considered. In [6], the authors study the OPA for an outage probability minimization problem in point-to-point fading channels with the constraints of the EH and the channel distribution at transmitter. In [7], a Markov decision process (MDP) model is proposed for the energy allocation problem over a finite horizon to maximize the throughput under point-to-point wireless communications, and both channel conditions and time varying energy sources are taken into account. According to the above research results, EH technology can obviously improve the energy efficiency; however, it also brings some new challenges to the power allocation strategy, such as the problem of uncertainty in EH process.

*Correspondence: lianghui@jlu.edu.cn
[†]Equal contributors
College of Communication Engineering, Jilin University, Nanhu Road, 130012 Changchun, China

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:7

Page 2 of 10

In fact, the former papers [4–7] have only considered point-to-point wireless communications without using CR technology. Recently, some researches have focused on the combination of EH technology and CRN, and different OPAs are discussed in [8–14]. The OPA for EH CRN is studied in [8], the authors consider the problem of system throughput maximization over a finite horizon rather than at a certain time slot, and they adopt a rate loss constraint to protect the transmission of primary user (PU). The access strategy for hybrid underlay-overlay CR networks with EH is analyzed in [9]. The partially observable Markov decision process (POMDP) framework is proposed to determine the action of SU; meanwhile, energy threshold is used to determine the transmission mode of SU. In [10], considering an EH-CR system, a power allocation policy with peak power constraints is proposed, and the target throughput maximization problem is solved by recursion machinery and geometric water-filling algorithm. A novel saving-sensing transmitting (SST) frame structure for EH CRNs is proposed in [11], where the authors aim to maximize the energy utilization efficiency of SU by jointly optimizing the save ratio and transmission power under both the energy causality constraint and the minimum throughput constraint. And the SST can make full use of residual battery energy as well as ensure enough time for spectrum sensing and data transmission. A generalized multislot spectrum sensing paradigm and two types of fusion rules (data fusion and decision fusion) are proposed in [12]. The authors focus on the trade-off of "harvesting-sensing-throughput" and joint optimization for save ratio, sensing duration, and sensing threshold as well as fusion rule to maximize the expected achievable throughput of SU while keeping the protection to PU. A POMDP is proposed in [13] to trade off energy consumption and throughput gain in hybrid CRN, where SU dynamically determines its operation mode for each time slot (e.g., to be idle or to transmit), sensing time, and access mode. In [14], for an overlay EH CRN, the authors aim to find an optimal sensing time to maximize throughput of SU and harvested RF energy. Most of the existing works only investigate OPA problem in CR system with single user, which is not pratical for real communication systems.

Based on the above discussions, we propose an OPA policy for the EH CRN considering multiusers with the underlay spectrum access model. And this paper only considers that SUs harvest energy from ambient environment such as solar [8, 9]. The major contributions of this paper are as follows:

- We consider the system model with multiple SUs where each SU transmitter is equipped with an EH device. In our study, we focus on the maximization of the average throughput of the secondary system

within K time slots under the maximum transmission power constraint, the minimum throughput constraint of SUs, and the interference power constraint defined by PU. The proposed algorithm can maximize average throughput and satisfy all constrains simultaneously to achieve better performance for SUs and the QoS of PU.

- Orthogonal frequency division multiplexing access (OFDMA) scheme is used in the process of spectrum sharing, where the available spectrum is divided into a set of subcarriers. The optimal transmission power of each SU can be obtain by using dynamic programming (DP) algorithm and immediate optimization solved by the backward induction method and the Lagrange dual method, respectively.

The rest of the paper is organized as follows. In Section 2, a system model of EH CRNs, a subcarrier allocation model of SUs, and a subcarrier occupation state are described. We introduce EH process in Section 3 and present the throughput optimization problem and the immediate OPA algorithm. Then, the system state and the DP-based scheme are formulated, and the backward induction method is given in Section 4. In Section 5, we present our simulation results and performance analysis through the comparison between our proposed policy and other policies. Finally, the conclusion of the whole paper is provided in Section 6.

2 System model

We consider an EH-CR network with a PU and M SUs as shown in Fig. 1a. Each SU transmitter is equipped with an EH device. We let the set $\mathcal{M} = \{1, 2, \ldots, m, \ldots, M\}$ denote the number of SU, and $\forall m \in \mathcal{M}$. The PU bands are divided into N subcarriers, and each of the subcarriers with the same bandwidth. We let the set $\mathcal{N} = \{1, 2, \ldots, n, \ldots, N\}$ denote the number of subcarrier, and $\forall n \in \mathcal{N}$. The number of the subcarriers occupied by PU is l. The SU can obtain the channel state information (CSI) and the working state of PU by spectrum sensing algorithm [15], such as matched filter detection, energy detection, and multiple identification spectrum detection. We let PU-Tx and PU-Rx denote the primary transmitter and receiver, respectively. Similarly, SU-Tx and SU-Rx denote the secondary transmitter and receiver, respectively.

We consider Rayleigh fading channels modeled as a two-state Markov chain [16, 17] as shown in Fig. 1b. The channel gains from PU-Tx to PU-Rx, the mth SU-Tx to the mth SU-Rx, the mth SU-Tx to PU-Rx, and PU-Tx to the mth SU-Rx over the nth subcarrier are denoted by w^n, h_m^n, a_m^n, and b_m^n, respectively. Moreover, we select the underlay spectrum access model in which SU has the opportunity to coexist with PU, while SU must control the interference to PU under a certain threshold.
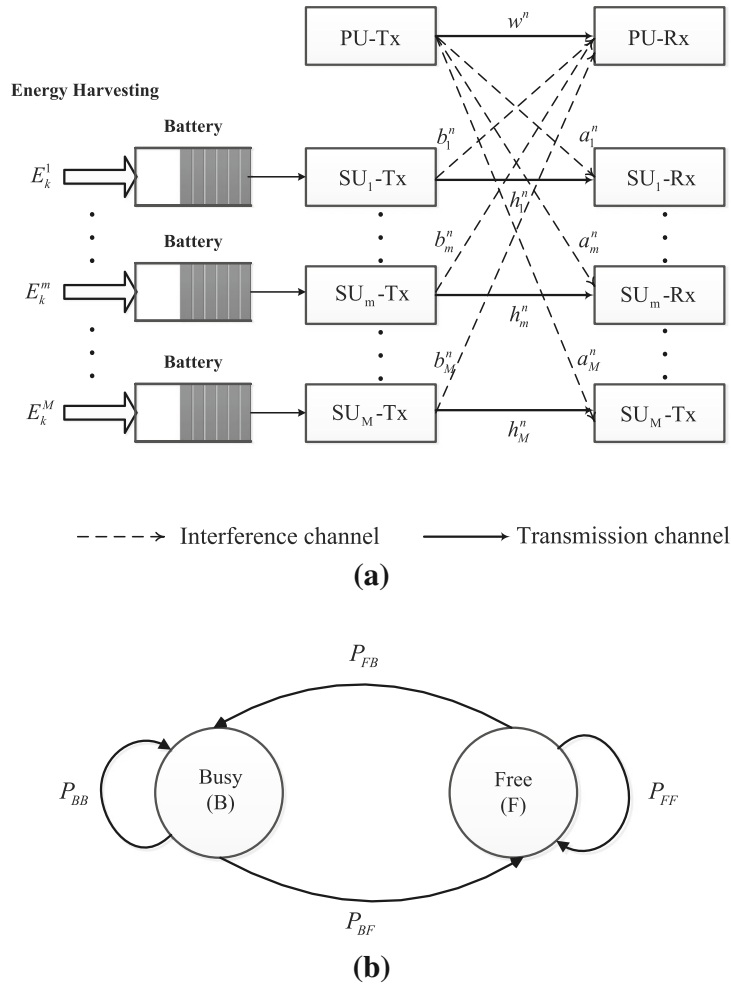
Fig. 1 System model and Markov chain model. **a** Time-slotted EH multiuser CR communication (SU-TX with finite battery capacity). **b** Markov chain model

2.1 Subcarrier allocation model

According to the above system model, we determine a policy to allocate N subcarriers to M SUs. Both PU and SU systems use the OFDMA sheme, and we assume that one subcarrier can only be used by one SU at each time slot, which means the interference between SUs is not considered. We consider that each SU can use multiple subcarriers at each time slot, the subcarriers occupied by the mth SU can be denoted as the set $\mathcal{K}_m$, $\mathcal{K}_m \in \mathcal{N}$. We define the minimum rate requirement [18] of each SU denoted as $R_{\min}^m, \forall m \in \mathcal{M}$. The allocation policy can be described as follows.

First, we determine the priorities of each SU according to $R_{\min}^m$, where the larger value of $R_{\min}^m$ reaches the higher priority of SU has. Then, the subcarriers that have a good channel condition and are not occupied by PU can be allocated to SU with high priority. According to this allocation policy, we can allocate the subcarriers to SUs.

2.2 Subcarrier occupation state

Since the subcarrier occupation state is constantly changing over time slot, SU should perform spectrum sensing to determine the behavior of PU at the beginning of each time slot. From Fig. 1b, the channel may be in one of the two states: busy (B) or free (F). The state B denotes that PU is active; conversely, the state F denotes the inactiveness of PU. We consider K time slots in our study; the time slot set can be defined as $\mathcal{K} = \{0, 1, \ldots, k, \ldots, K-1\}$ and $\forall k \in \mathcal{K}$.

We define x_k^n to indicate the state of nth channel at the time slot k, which has two possible values

$$x_k^n = \begin{cases} 0, & \text{the } n^{th} \text{channel is in state } F \\ 1, & \text{the } n^{th} \text{channel is in state } B \end{cases} \quad (1)$$

In addition, we define O_k to indicate the state set of all channels at the time slot k. Let N denotes the total subcarriers and l denotes the number of subcarriers occupied

by PU, then the number of random subcarrier occupation state can be defined as $L = \begin{pmatrix} N \\ L \end{pmatrix}$. Therefore, the elements of O_k can be given by

$$y_i = \left[x_k^1, \ldots, x_k^N\right]^T, i \in \{1, 2, \ldots, L\} \tag{2}$$

The transition matrix of PU occupation state is defined as $\mathbf{P}^o$. At first, the state transition probability of the nth channel can be given by

$$P_{BF}^n = \Pr\left\{x_{k+1}^n = 0 | x_k^n = 1\right\} \tag{3}$$

$$P_{FB}^n = \Pr\left\{x_{k+1}^n = 1 | x_k^n = 0\right\} \tag{4}$$

The transition probability of $\mathbf{P}^o$ is defined as

$$p_{ij}^o = \Pr\left\{O_{k+1} = y_j | O_k = y_i\right\} \tag{5}$$

and it further can be expressed as

$$p_{ij}^o = \prod_{n=1}^{N} \Pr\left\{x_{k+1}^n | x_k^n\right\} \tag{6}$$

According to P_{BF}^n and P_{FB}^n, we can get the value of p_{ij}^o.

3 Optimal power allocation policy
3.1 Energy-harvesting process
Here, we assume that the finite capacity of energy harvest battery (i.e., energy queue) is attached to each SU transmitter and can be used for signal transmission. We also consider that the energy harvesters only harvest energy from ambient environment such as solar. The energy harvested packets can be denoted as $E_k^m \in \mathbf{e} = \{e_1, e_2, \ldots, e_H\}$ following a Poisson process [9] with mean e_λ, $\forall k \in \mathcal{K}$, $\forall m \in \mathcal{M}$. Thus, its probability distribution can be described as follows

$$\Pr\left(E_k^m = e_j\right) = e^{-e_\lambda} \frac{(e_\lambda)^j}{j!}, j = 1, 2, \ldots, H \tag{7}$$

Besides, we use B_k^m to represent the remaining battery energy of the mth SU at the time slot k; therefore, the battery energy update value at the next time slot $k+1$ can be given by

$$B_{k+1}^m = \min\left\{B_k^m - p_k^m T + E_k^m, B_{\max}\right\}, \forall k \in \mathcal{K} \tag{8}$$

where $p_k^{m,n}$ denotes the transmission power allocated to the mth SU in the nth channel at the time slot k, T denotes the duration of one time slot, and $B_{\max}$ is the maximum battery capacity.

3.2 Problem formulation
In our study, the optimization objective is to maximize the average throughput of the secondary system within K time slots. We define $SINR_k^{m,n}$ as signal-interference-noise

radio (SINR) of the mth SU in the nth channel at the time slot k, and

$$SINR_k^{m,n} = \frac{h_k^{m,n} p_k^{m,n}}{a_k^{m,n} p_k^{p,n} + \sigma^2}, \forall m, \forall n \tag{9}$$

then defined

$$g_k^{m,n} = \frac{h_k^{m,n}}{a_k^{m,n} p_k^{p,n} + \sigma^2}, \forall m, \forall n \tag{10}$$

where $p_k^{p,n}$ denotes the transmission power of PU in the nth channel at the time slot k, and σ^2 is the noise power. Therefore, the throughput of the mth SU at the time slot k can be defined as

$$R_k^m = \sum_{n=1}^{N} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right), \forall m \tag{11}$$

And the optimization problem can be formulated as **OP1**

$$\max_{p_k^{m,n}} \quad E\left\{\frac{1}{K} \sum_{k=0}^{K-1} \sum_{n=1}^{N} \sum_{m=1}^{M} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right)\right\}$$
$$\text{s.t.} \quad C1: \sum_{n=1}^{N} p_k^{m,n} \leq \frac{B_k^m}{T}, \forall m, \forall k$$
$$C2: \sum_{n=1}^{N} \sum_{m=1}^{M} b_k^{m,n} p_k^{m,n} \leq I^{th}, \forall k$$
$$C3: R_k^m \geq R_{\min}^m, \forall k$$
$$C4: p_k^{m,n} \geq 0, \forall m, \forall n, \forall k \tag{12}$$

where $E\{\cdot\}$ denotes the expectation of the channel gain distribution and the subcarrier-occupied state at each time slot. $C1$ denotes the maximum transmission power constraint, and B_k^m/T is the total transmission power budget for the mth SU at the time slot k. This constraint can ensure the transmission power of each SU at each time slot not to exceed the energy budget. $C2$ denotes the interference power constraint to guarantee the interference to PU remains under I^{th}, where I^{th} is the interference threshold prescribed by the PU receiver. $C3$ represents the minimum throughput constraint which can keep the throughput above the minimum throughput requirement in the network. The constrain $C4$ can make the transmission power of SU conform to the actual situation.

3.3 Immediate optimization problem solution
To solve the optimal problem **OP1**, we consider both the throughput at the current time slot and the future throughput. In order to simplify the problem, we set $K = 1$, that means we only consider the optimal problem at one time slot. Therefore, we can formulate the immediate optimization problem as **OP2**

Wang et al. EURASIP Journal on Wireless Communications and Networking (2018) 2018:7

Page 5 of 10

$$\max_{p_k^{m,n}} \sum_{n=1}^{N} \sum_{m=1}^{M} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right) \quad (13)$$
$$\text{s.t.} \quad C1 \sim C4$$

OP2 is a convex problem which can be solved by the Lagrange dual method [19]. First, we define the Lagrange function

$$L\left(\{p_k^{m,n}\}, \{\lambda_m\}, \mu, \{\xi_m\}\right)$$
$$= \sum_{n=1}^{N} \sum_{m=1}^{M} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right) - \sum_{m=1}^{M} \lambda_m\left(\sum_{n=1}^{N} p_k^{m,n} - B_k^m/T\right)$$
$$- \mu\left(\sum_{n=1}^{N} \sum_{m=1}^{M} b_k^{m,n} p_k^{m,n} - I^{th}\right) - \sum_{m=1}^{M} \xi_m\left(R_{\min}^m - \sum_{n=1}^{N} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right)\right)$$

$$(14)$$

where λ_m, μ, $\xi_m \geq 0$ are the Lagrange multipliers. The dual variable λ_m relates to the maximum transmission power constraint, the dual variable μ connects with the interference power constraint, and the dual variable ξ_m contacts with the minimum throughput constraint.

Moreover, the dual problem of the Lagrange function is

$$D\left(\{\lambda_m\}, \mu, \{\xi_m\}\right)$$
$$= \max_{\{p_k^{m,n}\}} L\left(\{p_k^{m,n}\}, \{\lambda_m\}, \mu, \{\xi_m\}\right)$$
$$= \sum_{m=1}^{M} \max_{\{p_k^{m,n}\}} L_m\left(\{p_k^{m,n}\}, \lambda_m, \mu, \xi_m\right) + \sum_{m=1}^{M} \lambda_m\left(B_k^m/T\right)$$
$$+ \mu I^{th} - \sum_{m=1}^{M} \xi_m R_{\min}^m$$

$$(15)$$

where

$$L_m\left(p_k^{m,n}, \lambda_m, \mu, \xi_m\right)$$
$$= \sum_{n=1}^{N} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right) - \lambda_m\left(\sum_{n=1}^{N} p_k^{m,n}\right)$$
$$- \mu\left(\sum_{n=1}^{N} b_k^{m,n} p_k^{m,n}\right) + \xi_m\left(\sum_{n=1}^{N} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right)\right)$$

$$(16)$$

The dual optimization problem d^* of (13) can be formulated as

$$\min_{\lambda_m, \mu, \xi_m \geq 0} D\left(\lambda_m, \mu, \xi_m\right), \forall m \quad (17)$$

Since L_m is a convex function, according to the Karush-Kuhn-Tucker (KKT) conditions [20], the optimal transmission power $p_k^{m,n}$ at the mth SU transmitter can be calculated by $\partial L_m / \partial p_k^{m,n} = 0$. Thus, the optimal solution is

$$p_k^{m,n*} = \left[\frac{(1 + \xi_m)}{\ln 2\left(\lambda_m + \mu b_k^{m,n}\right)} - \frac{1}{g_k^{m,n}}\right]^+ \quad (18)$$

where $[\cdot]^+ = \max(0, \cdot)$.

The Lagrange multipliers λ_m, μ, and ξ_m should ensure a fast convergence rate. We can use the sub-gradient methods to update these multipliers, and their recursive forms are

$$\lambda_m(i+1) = \left[\lambda_m(i) + \alpha_1\left(\sum_{n=1}^{N} p_k^{m,n} - B_k^m/T\right)\right]^+$$

$$(19)$$

$$\mu(i+1) = \left[\mu(i) + \alpha_2\left(\sum_{n=1}^{N} \sum_{m=1}^{M} b_k^{m,n} p_k^{m,n} - I^{th}\right)\right]^+$$

$$(20)$$

$$\xi_m(i+1) = \left[\xi_m(i) + \alpha_3\left(R_{min}^m - \sum_{n=1}^{N} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right)\right)\right]^+$$

$$(21)$$

where i denotes the iteration number. α_1, α_2, and $\alpha_3 \geq 0$ are small step sizes. The proper selection of the step size can ensure the stability and convergence of this dual algorithm [21]. Finally, we can get the optimal solution $p_k^{m,n*}$. Then, by taking this solution into (11), the optimal throughput of each SU can be calculated.

Thus, our proposed immediate OPA algorithm can be summarized as Table 1.

However, this solution does not directly apply to the multiple time slots, namely, $K > 1$, in which we should balance the current throughput and the expected future throughput. According to our system model, the subcarrier occupation state, the harvested energy state, and the total energy budget are all time-varying state. In order to further develop this problem for $K > 1$, we use DP algorithm, which will be introduced in the next section.

4 Dynamic programming formulation and backward induction method
4.1 System state
Considering the future expected throughput, the transition probability of the system state should be determined.

Table 1 Immediate OPA algorithm introduction

Algorithm 1 Immediate OPA algorithm
1: Initialization: set $i = 0, \lambda_m(0) > 0, \mu(0) > 0, \xi_m(0) > 0$; $\alpha_1(0) > 0, \alpha_2(0) > 0, \alpha_3(0) > 0; I^{th} > 0, R_{\min}^m > 0$;
2: Solve the optimization problem (13) to obtain $p_k^{m,n}, \forall k$, thus we can get the throughput of the secondary system;
3: Update the Lagrange multiplier by (19) $\sim$ (21);
4: Update the transmission power $p_k^{m,n}$ by (18);
5: Go to 2 until $\left\|\hat{p}_k^{m,n}(i+1) - \hat{p}_k^{m,n}(i)\right\| \leq \varepsilon$, where ε represents iteration precision usually a very small positive constant;
6: End: The optimal transmission power $p_k^{m,n*}$ can be calculated by (18), and take $p_k^{m,n*}$ into (11), the optimal throughput can be obtained.

Wang *et al. EURASIP Journal on Wireless Communications and Networking* (2018) 2018:7

Page 6 of 10

Through the introduction of the previous section, we know that the system states include the subcarrier occupation and the harvested energy which are independent of each other. Since the process of EH for each SU is independent, the states of harvested energy for each SU are also independent.

We define the system state as $S_k, \forall k \in \mathcal{K}$, and the number of system state is $J = L \times H^M$. The elements of S_k can be described as follows

$$
\begin{aligned}
&s_i = \left(y_g, e_{h_1}, e_{h_2}, \ldots, e_{h_M}\right), \\
&\forall i \in \left\{1, 2, \ldots\ldots, L \times H^M\right\} \\
&\forall g \in \left\{1, 2, \ldots\ldots, L\right\} \\
&e_{h_1}, e_{h_2}, \ldots, e_{h_M} \in \mathbf{e} = \left\{e_1, e_2, \ldots\ldots, e_H\right\}
\end{aligned} \tag{22}
$$

We use $\mathbf{P}^s$ to denote the transition matrix of the system state, and its dimension is $D = J^2 = \left(L \times H^M\right)^2$. We assume $S_{k+1} = s_j$ at the time slot $k + 1$, where

$$
\begin{aligned}
&s_j = \left(y_f, e_{H_1}, e_{H_2}, \ldots, e_{H_M}\right), \\
&\forall j \in \left\{1, 2, \ldots\ldots, L \times H^M\right\} \\
&\forall f \in \left\{1, 2, \ldots\ldots, L\right\} \\
&e_{H_1}, e_{H_2}, \ldots, e_{H_M} \in \mathbf{e} = \left\{e_1, e_2, \ldots\ldots, e_H\right\}
\end{aligned} \tag{23}
$$

Therefore, the transition probability of $\mathbf{P}^s$ can be given as

$$
\begin{aligned}
p_{ij}^s &= \Pr\left\{S_{k+1} = s_j | S_k = s_i\right\} \\
&= \Pr\left\{O_{k+1} = y_f | O_k = y_g\right\} \prod_{m=1}^{M} \Pr\left(E_{k+1}^m = e_{H_m}\right)
\end{aligned} \tag{24}
$$

According to (5), (6), and (7), the transition probability can be calculated.

4.2 Dynamic programming formulation

In this section, we define a DP formulation [22] to solve the optimization problem **OP1**. We define a reward function which can be understood as a maximum of the sum of the throughput at the current time slot and the expected cumulative throughput at the future time slot from the current system state. We set $V_k\left(B_k^1, B_k^2, \ldots\ldots, B_k^M, S_k\right)$ to denote the reward function at the time slot k, which is a function of the current energy budget B_k^m of each SU and the current system state S_k, and it can be expressed as

$$
\begin{aligned}
&V_k\left(B_k^1, B_k^2, \ldots\ldots, B_k^M, S_k\right) \\
&= \max_{p_k^{m,n}} \mathrm{E}\left\{\sum_{t=k}^{K-1} \sum_{n=1}^{N} \sum_{m=1}^{M} \log_2\left(1 + g_k^{m,n} p_k^{m,n}\right)\right\}, \forall k \in \mathcal{K}
\end{aligned} \tag{25}
$$

4.3 Backward induction method

The backward induction method [23] can be used to solve (25). The immediate reward function at the current time slot can be defined as $R_k\left(B_k^1, B_k^2, \ldots\ldots, B_k^M, S_k\right)$, while we set $V_{k+1}\left(B_{k+1}^1, B_{k+1}^2, \ldots\ldots, B_{k+1}^M, S_{k+1}\right)$ to denote the future reward function at the next time slot $k + 1$.

Since we consider the time slot set as $\mathcal{K} = \{0, 1, \ldots, K - 1\}$, the reward function at time slot K is

$$
V_K\left(B_K^1, B_K^2, \ldots\ldots, B_K^M, S_K\right) = 0 \tag{26}
$$

For $k = K - 1, K - 2, \ldots, 0$, the reward functions can be expressed as

$$
\begin{aligned}
&V_{K-1}\left(B_{K-1}^1, B_{K-1}^2, \ldots\ldots, B_{K-1}^M, S_{K-1}\right) \\
&= \max_{p_k^{m,n}}\left\{R_{K-1}\left(B_{K-1}^1, B_{K-1}^2, \ldots\ldots, B_{K-1}^M, S_{K-1}\right)\right\}
\end{aligned} \tag{27}
$$

$$
\begin{aligned}
&V_k\left(B_k^1, B_k^2, \ldots\ldots, B_k^M, S_k\right) \\
&= \max_{p_k^{m,n}}\left\{R_k\left(B_k^1, B_k^2, \ldots\ldots, B_k^M, S_k\right) + \right. \\
&\quad \left. \mathrm{E}\left[V_{k+1}\left(B_{k+1}^1, B_{k+1}^2, \ldots\ldots, B_{k+1}^M, S_{k+1}\right)\right]\right\}
\end{aligned} \tag{28}
$$

$$
\begin{aligned}
&V_0\left(B_0^1, B_0^2, \ldots\ldots, B_0^M, S_0\right) = \\
&\max_{p_k^{m,n}}\left\{\mathrm{E}\left[V_1\left(B_1^1, B_1^2, \ldots\ldots, B_1^M, S_1\right)\right]\right\}
\end{aligned} \tag{29}
$$

Thus, the reward function can be further calculated in the time-reversal order.

1) Time slot $k = K - 1$

$$
\begin{aligned}
&V_{K-1}\left(B_{K-1}^1, B_{K-1}^2, \ldots\ldots, B_{K-1}^M, S_{K-1}\right) \\
&= \max_{0 \le p_{K-1}^m \le B_{K-1}^m}\left\{R_{K-1}\left(B_{K-1}^1, B_{K-1}^2, \ldots\ldots, B_{K-1}^M, S_{K-1}\right)\right\}
\end{aligned} \tag{30}
$$

where $p_{K-1}^m = \sum_{n=1}^{N} p_{K-1}^{m,n}, \forall m \in \mathcal{M}$. In this case, we only need to consider the immediate reward function, which can be achieved when the transmission power of each SU on each subcarrier satisfies the optimal solution (18).

2) Time slots from $k = K - 2$ to $k = 1$

$$
\begin{aligned}
&V_k\left(B_k^1, B_k^2, \ldots\ldots, B_k^M, S_k = i\right) \\
&= \max_{0 \le p_k^m \le B_k^m}\left\{R_k\left(B_k^1, B_k^2, \ldots\ldots, B_k^M, S_k = i\right) + \right. \\
&\quad \left. \mathrm{E}\left[V_{k+1}\left(B_{k+1}^1, B_{k+1}^2, \ldots\ldots, B_{k+1}^M, S_{k+1} = j\right)\right]\right\} \\
&= \max_{0 \le p_k^m \le B_k^m}\left\{R_k\left(B_k^1, B_k^2, \ldots\ldots, B_k^M, S_k = i\right) + \right. \\
&\quad \left. \sum_{j=1}^{J} p_{ij}^s V_{k+1}\left(B_{k+1}^1, B_{k+1}^2, \ldots\ldots, B_{k+1}^M, S_{k+1} = j\right)\right\}
\end{aligned} \tag{31}
$$

where B_{k+1}^m can be updated by (8), $\forall m \in \mathcal{M}$. Through the state transition probability p_{ij}^s, we can get the expected reward function. By considering the trade-off between the current reward and the potential reward at next time slot, we can get the optimal transmission power.

3) Time slot $k = 0$

$$
\begin{aligned}
&V_0\left(B_0^1, B_0^2, \ldots\ldots, B_0^M, S_0 = i\right) \\
&= \max_{0 \le p_0^m \le B_0^m}\left\{\mathrm{E}\left[V_1\left(B_1^1, B_1^2, \ldots\ldots, B_1^M, S_1 = j\right)\right]\right\} \\
&= \max_{0 \le p_0^m \le B_0^m}\left\{\sum_{j=1}^{J} p_{ij}^s V_1\left(B_1^1, B_1^2, \ldots\ldots, B_1^M, S_1 = j\right)\right\}
\end{aligned} \tag{32}
$$

where B_0^m denote the energy budget for SUs at the beginning of transmission, $\forall m \in \mathcal{M}$, and S_0 is the system initial state. At the initial time slot, we only need to satisfy the maximum of the expected reward at time slot $k = 1$.

Therefore, our DP power allocation algorithm can be summarized as Table 2.

As we can see, the results can be stored by a table with the time slot index; according to this table, SUs can determine the optimal transmission power.

4.4 Performance analysis

We first analyze the performance of immediate power allocation algorithm. This algorithm can guarantee the interference of PU below a certain threshold and the throughput of each SU above a proper threshold at each time slot.

From (19) to (21), we can see that the Lagrange multipliers can be updated only by local information, which can effectively improve the calculate speed and reduce the algorithm complexity. If the transmission power $p_k^{m,n}$ is relatively high, it will result in $\sum_{n=1}^{N} p_k^{m,n} > B_k^m/T$ and $\sum_{n=1}^{N} \sum_{m=1}^{M} b_k^{m,n} p_k^{m,n} > I^{th}$, which violates the constraint conditions, so that λ_m and μ will increase and ξ_m will decrease. Following (18), we can find $p_k^{m,n}$ will reduce. As a result, the transmission power can be adjusted to satisfy the constraint conditions. However, this power does not infinitely decrease, if $p_k^{m,n}$ becomes relatively small, λ_m and μ will decrease and ξ_m will increase. Therefore, $p_k^{m,n}$ will increase and goes back to the appropriate range. This adaptive iterative process can ensure good QoS for both PU and SU.

Based on the Lipschitz continuity [24] of the dual function and the proper step parameters of the Lagrange multipliers, this algorithm can converge quickly. According to the Lipschitz continuity, there exist a Lipschitz constant δ which can make the function $d^*(\lambda_m, \mu, \xi_m)$ satisfy the following condition:

$\|d^*(\lambda_1, \mu_1, \xi_1) - d^*(\lambda_2, \mu_2, \xi_2)\|_2 \leq \delta \left\| [\lambda_1, \mu_1, \xi_1]^T - [\lambda_2, \mu_2, \xi_2]^T \right\|_2$, where $\lambda_1, \lambda_2 \in \lambda_m, \mu_1, \mu_2 \in \mu, \xi_1, \xi_2 \in \xi_m$, and $\|\cdot\|_2$ denotes the norm of vector. Thus, we can determine the dual function d^* of (13) is uniformly continuous. When $p_k^{m,n}$ satisfies all constraints C1 to C4 with λ_m, μ and $\xi_m \geq 0$, $(p_k^{m,n*}, \lambda_m^*, \mu^*, \xi_m^*)$ can converge to a feasible region. Owing to the duality property between the dual problem and the original problem, the immediate power allocation algorithm can converge to the optimal solution.

Using the immediate power allocation algorithm, we can realize DP power allocation algorithm with consideration of the throughput optimization problem for the whole K time slots. We store the system state S_k, the energy budget level B_k^m, the immediate reward function R_k, and the reward function V_k in a look-up table indexing with the time slot, which contains all the possible situations. Therefore, each SU can determine his optimal power policy from this table which greatly reduces the computational complexity.

5 Simulation results

In this section, we present some simulation results to evaluate our proposed algorithm by comparing with two policies. The first policy is the conservative power policy. We use this scheme to allocate half of available energy to each time slot for power allocation, i.e., $p_k^m = B_k^m/(2T)$, $\forall k \in \mathcal{K}$. The second policy is the greedy power policy which uses whole available energy for power allocation at each time slot, i.e., $p_k^m = B_k^m/T$, $\forall k \in \mathcal{K}$.

In our simulations, we assume that there are four SUs, i.e., $M = 4$, one PU, and eight subcarriers, i.e., $N = 8$. In addition, this PU occupies two subcarriers in each time slot. As the result, the number of subcarrier occupation state is twenty-eight, i.e., $L = 28$. We assume that the total transmission power of the primary user is constant, and average allocation in each occupied subcarriers. The throughput performance is compared over the total bandwidth of $B = 1$ MHz. The channel suffers with the frequency-selective fading by a six-ray Rayleigh model with exponential profile and maximal $5\mu s$ multipath delay [8]. Moreover, we define the maximum battery capacity of each SU is 5 J, i.e., $B_{\max} = 5$J. Thus, the total energy budget of the secondary system is 20 J. We set $T = 1$ s, and the energy budget and the transmission power with resolutions of 0.5 J and 0.5 W, respectively. The minimum throughput of each SU, i.e., $R_{\min}^m$, is a positive constant depending on the energy budget at current time slot. We set $R_{\min}^1 > R_{\min}^2 > R_{\min}^3 > R_{\min}^4$, which means the priority of SUs is SU_1, SU_2, SU_3, and SU_4. The interference temperature (IT) at the PU receiver is $I^{th} = 0.1$W. The simulation results are presented in Figs. 2, 3, 4, 5, 6, and 7.

Table 2 DP power allocation algorithm

Algorithm 2	DP power allocation algorithm
1:	Input all statistics of the channel gain information, subcarriers occupation state and EH process;
2:	For $k = 0, 1, \ldots, K-1$:
	Calculate and store R_k for all system states according to Algorithm 1;
3:	For $k = K-1, K-2, \ldots, 0$:
	Calculate and store V_k for all system states based on (30), (31) and (32);
	Then, calculate and store $p_k^{m,n*} = \arg\max_{p_k^{m,n}} V_k$.

Fig. 2 Convergence of average throughput of each SU under one time slot

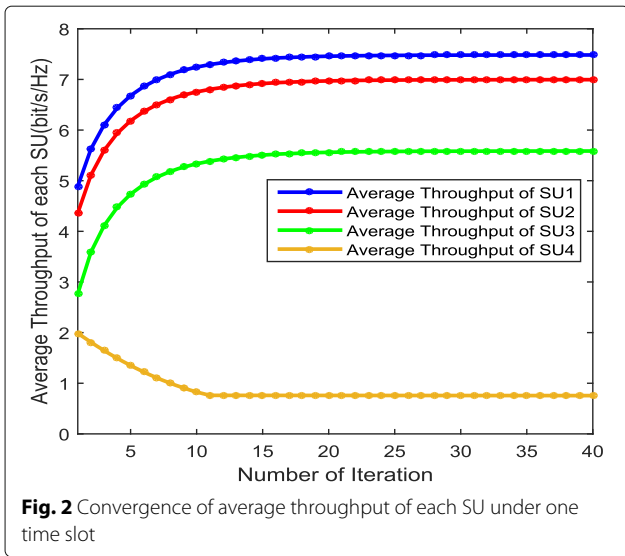**Fig. 4** Performance comparison among the policies with different number of time slot

Figure 2 shows the convergence of the average throughput of four SUs with the proposed algorithm, respectively. In this simulation, we set the energy budget of each SU as 5 J. From Fig. 2, we can clearly see that this scheme can quickly converge to the equilibrium point. Obviously, Fig. 2 show that the average throughput of SU_1 is the best of all SUs, and the average throughput of SU_4 is less than all SUs. The reason is that according to the priority of SUs and the subcarriers allocation model, the SU with higher priority can transmit data under good channel condition. However, SU_4 must deal with the interference power constraint to ensure the QoS of PU.

In Fig. 3, we consider the average throughput comparisons among our proposed power allocation policy, the conservative power policy, and the greedy power policy. We compare the average throughput of the secondary system on the variation of total energy budget. From this

Fig. 3 Performance comparison among the policies with different total energy budget under $K = 5$
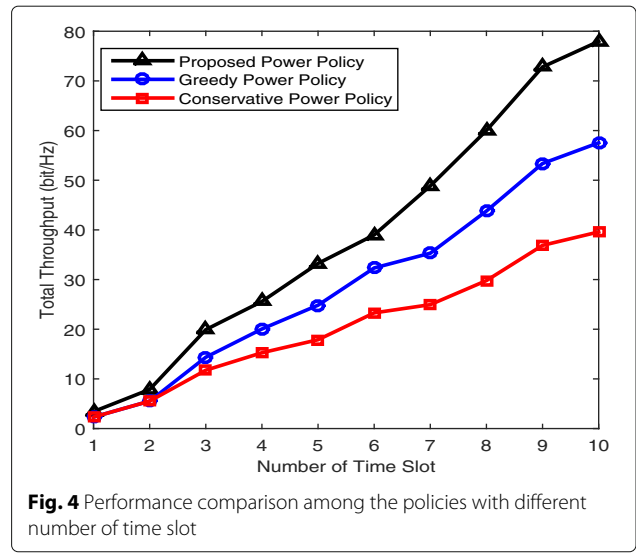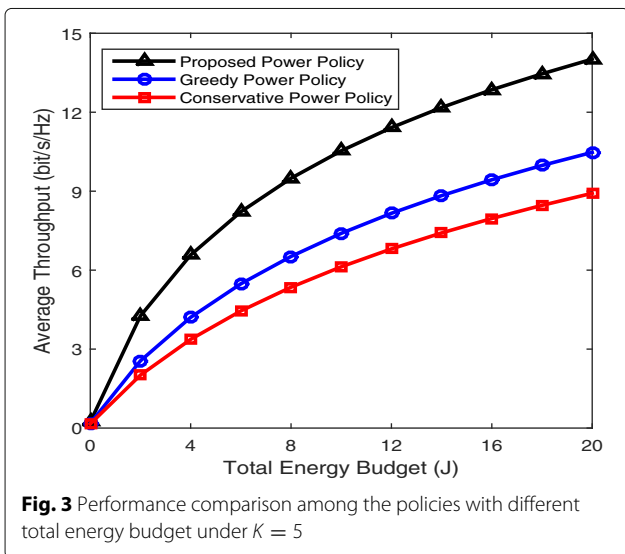
figure, we know that the average throughput of our propose policy is much higher than those of the other two policies over the examined range of total energy budget, since our policy considers not only the immediate OPA but also the whole time slot transmit performance. In addition, the average throughput increases for all three policies with the increase of the total energy budget, since the increase of the energy budget means we have more energy for power allocation for the increase of the throughput. But this throughput increases much more rapidly through our policy.

Figure 4 shows the comparisons of total throughput among three policies on the variation of time slot. In this case, we also set the energy budget of each SU is 5 J. Obviously, the throughput of our policy increases significantly with the increase number of time slot. Moveover, we can find that our policy has the best performance among the three policies over the whole range of number of time slot. From the simulation result, we can conclude that our policy can guarantee optimal performance for long-run operations.

Figure 5 illustrates the comparison of interference power at PU receiver for three policies. We set the IT level as $I^{th} = 0.1$ W. Figure 5a provides the convergence of the interference power under arbitrary time slot. It is clear that the interferences from three algorithms can quickly converge to their stable points, and these three policies can guarantee the interference power at the PU receiver below the IT level. Figure 5b presents the average interference comparisons among three policies on the variation of total energy budget under $K = 5$. We can find that with the increase of the total energy budget, three average interferences increase gradually and are always less than the IT level even if the energy budget reaches the battery capacity. From Fig. 5a, b, we can see that the interference
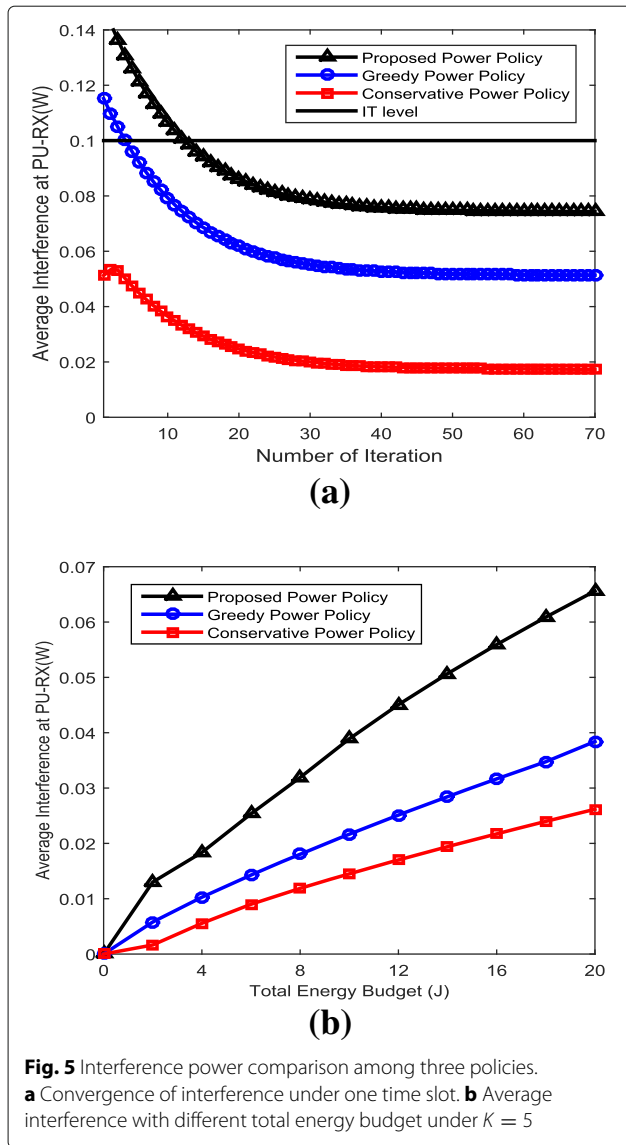
(a)

(b)

Fig. 5 Interference power comparison among three policies. **a** Convergence of interference under one time slot. **b** Average interference with different total energy budget under $K = 5$

Fig. 6 Average throughput with different IT level and energy budget under one time slot

transmission power will eventually tend to a constant value. In addition, the interference power constraint represents the distance, with the increasing distance between SU and PU; more transmit power is allocated to achieve higher throughput.

Finally, in Fig. 7, we present the average throughput on the variation of the number of time slot of our proposed algorithm for different IT levels. In this simulation, we set the energy budget of each SU is 0.5 J. From Fig. 7, we can find that the higher IT level can obtain higher total throughput. Since that if PU can tolerate higher interference power, SUs can be allocated with higher transmission power. From other perspective, with the increase number of time slot, the overall trend of throughput goes up. However, the throughputs at $K = 6, 9$ slightly decrease, since the channel gain, the subcarrier occupation state, and the harvested energy state are random at each time slot. According to Figs. 6 and 7, we can get the conclusion that

from our policy is slightly higher than those of other two policies under the premise of the interference power constraint, i.e., 0.01 W $\sim$ 0.02 W. Moreover, from Figs. 3, 4, and 5, we conclude that the proposed policy can provide better throughput performance of the secondary system at the cost of little more interference at the PU receiver within the tolerance of PU.

To show the impact of different IT level on the system performance, Fig. 6 presents the average throughput on the variation of the IT level with different energy budget. From Fig. 6, we find that the average throughput increases first then tends to a constant value with the increasing of the IT level. The reason for this phenomenon is that the higher IT level attains, the higher interference power can be tolerated by PU. Thus, SU can get more transmission power for more throughput. Meanwhile, due to the constraint of the energy budget in our policy, the
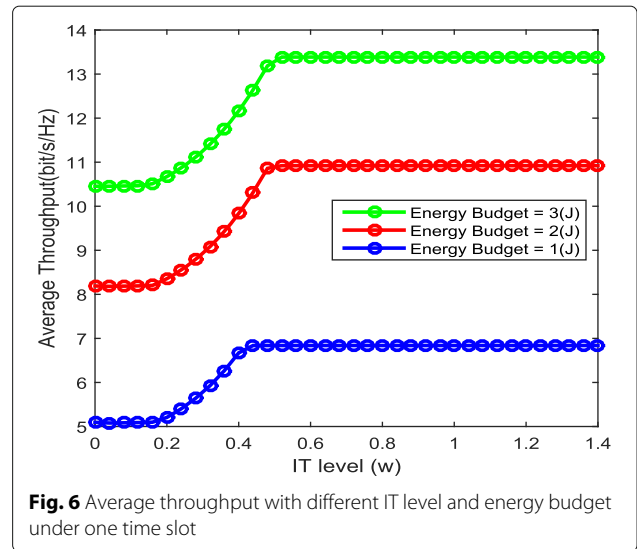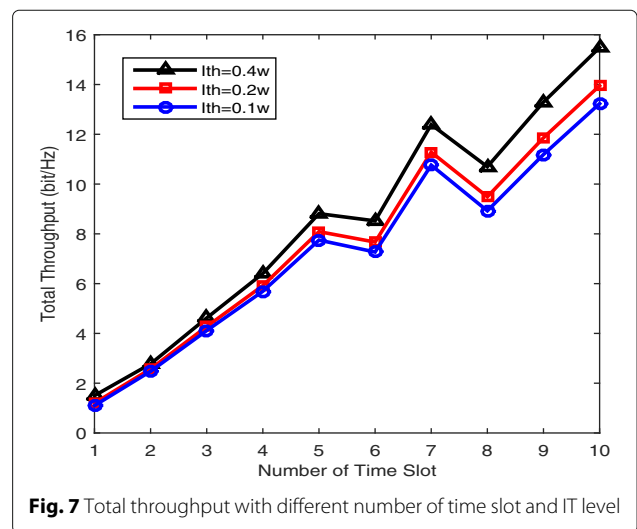

Fig. 7 Total throughput with different number of time slot and IT level

it is necessary to choose an appropriate IT level to balance the PU protection and the secondary system performance.

6 Conclusions

In this paper, we study the OPA problem in the EH CRN and propose an OPA policy to maximize the average throughput of the secondary system within K time slots, where the maximum transmission power constraint, the interference power constraint, and the minimum throughput constraint are considered. The optimal transmission power in each time slot can be obtained by joint utilization of the immediate OPA algorithm and the DP method. The simulation results show that, compared with the conservative power policy and the greedy power policy without DP approach, our policy can achieve better throughput performance on the variation of total energy budget and the number of time slot. Meanwhile, our algorithm can provide well protection for the basic communication of PU through introduction of interference power constraint. However, this policy improves the performance of the secondary system at the expense of little more interference compared with the other two policies.

Authors' contributions
YW contributed in the conception of the study and design of the study and wrote the manuscript. Furthermore, YW carried out the simulation and revised the manuscript. XZ and HL helped to perform the analysis with constructive discussions and helped to draft the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. J Mitola, GQ Maguire, Cognitive radio: making software radios more personal. IEEE Pers. Commun. **6**(4), 13–18 (1999)
2. D Gunduz, K Stamatiou, N Michelusi, et al, Designing intelligent energy harvesting communication systems. IEEE Commun. Mag. **52**(1), 210–216 (2014)
3. OH Hayoung, RAN Rong, Stochastic policy-based wireless energy harvesting in green cognitive radio network. Eurasip J. Wirel. Commun. Netw. **2015**(177), 2–10 (2015)
4. R Ma, W Zhang, Adaptive MQAM for energy harvesting wireless communications with 1-bit channel feedback. IEEE Trans. Wirel. Commun. **14**(11), 6459–6470 (2015)
5. Z Wang, V Aggarwal, X Wang, Power allocation for energy harvesting transmitter with causal information [J]. IEEE Trans. Commun. **62**(11), 4080–4093 (2014)
6. C Huang, R Zhang, S Cui, Optimal power allocation for outage probability minimization in fading channels with energy harvesting constraints. IEEE Trans. Wirel. Commun. **13**(2), 1074–1087 (2014)
7. CK Ho, R Zhang, Optimal energy allocation for wireless communications with energy harvesting constraints. IEEE Trans. Signal Process. **60**(9), 4808–4818 (2012)
8. H Liang, X Zhao, in *International Conference on Computing, Networking and Communications (ICNC)*. Optimal power allocation for energy harvesting cognitive radio networks with primary rate protection (IEEE, Kauai, 2016), pp. 1–6. doi:10.1109/ICCNC.2016.7440658
9. M Usman, I Koo, Access strategy for hybrid underlay-overlay cognitive radios with energy harvesting. IEEE Sensors J. **14**(9), 3164–3173 (2014)
10. P He, L Zhao, in *IEEE International Conference on Communications (ICC)*. Optimal power control for energy harvesting cognitive radio networks (IEEE, London, 2015), pp. 92–97. doi:10.1109/ICC.2015.7248304
11. C Wu, Q Shi, C He, et al, Energy utilization efficient frame structure for energy harvesting cognitive radio networks. IEEE Wirel. Commun. Lett. **5**(5), 488–491 (2016)
12. S Yin, Z Qu, S Li, Achievable throughput optimization in energy harvesting cognitive radio systems. IEEE J. Sel. Areas Commun. **33**(3), 407–422 (2015)
13. Y Zhang, Q Zhang, Y Wang, et al, in *IEEE/CIC International Conference on Communications in China (ICCC)*. Energy and throughput tradeoff in hybrid cognitive radio networks based on POMDP (IEEE, Xi'an, 2013), pp. 668–673. doi:10.1109/ICCChina.2013.6671196
14. A Bhowmick, SD Roy, S Kundu, Throughput of a cognitive radio network with energy-harvesting based on primary user signal. IEEE Wirel. Commun. Lett. **5**(2), 136–139 (2016)
15. Letaief KB, Zhang W, Cooperative communications for cognitive radio networks. Proc. IEEE. **97**(5), 878–893 (2009)
16. HS Wang, N Moayeri, Finite-state Markov channel—a useful model for radio communication channels. IEEE Trans. Veh. Technol. **44**(1), 163–171 (1995)
17. Q Zhang, SA Kassam, Finite-state Markov model for Rayleigh fading channels. IEEE Trans. Commun. **47**(11), 1688–1692 (1999)
18. C Huang, R Zhang, S Cui, Throughput maximization for the Gaussian relay channel with energy harvesting constraints. IEEE J. Sel. Areas Commun. **31**(8), 1469–1479 (2013)
19. S Boyd, L Vandenberghe, *Convex optimization*. (Cambridge University Press, Cambridge, 2004)
20. N Rahbari-Asr, MY Chow, Cooperative distributed demand management for community charging of PHEV/PEVs based on KKT conditions and consensus networks. IEEE Trans. Ind. Inform. **10**(3), 1907–1916 (2014)
21. DP Palomar, M Chiang, A tutorial on decomposition methods for network utility maximization. IEEE J. Sel. Areas Commun. **24**(8), 1439–1451 (2016)
22. DP Bertsekas, *Dynamic Programming and Optimal Control, vol 1*. (Athena Scientific, Belmont, 1995)
23. ML Puterman, *Markov Decision Process, Discrete Stochastic Dynamic Programming*. (Wiley, New York, 1994)
24. K Eriksson, D Estep, C Johnson, *Lipschitz continuity*. (Springer-Verlag, Berlin Heidelberg, 2004)