Game Theoretic Analysis for Spectrum Sharing with Multi-hop Relaying

Yong Xiao, Guoan Bi, *Senior Member, IEEE,* and Dusit Niyato, *Member, IEEE*

***Abstract*—This paper studies spatial spectrum sharing (SSS) based multi-user cognitive radio (CR) networks that allow secondary users (SU) to access the licensed spectrum as long as the interference powers of primary users (PU) to be lower than a certain threshold. Although recent results have shown that multi-hop relaying has a great potential on improving the performance of CR networks, finding effective methods to control and manage SUs to achieve the optimal performance is still a challenging problem. In this paper, we model CR networks as a non-cooperative game in which each SU obtains benefits through both spectrum sharing by paying prices to PUs and multi-hop relaying by paying price to nearby SUs. Optimal power allocation methods for SUs are investigated under different assumptions and pricing functions. The conditions under which the optimal Nash Equilibrium (NE) is obtained when all SUs use multi-hop relaying are discussed. Our results are extended into large multi-user CR networks with K source-to-destination pairs. Two distributed algorithms are proposed. The first one is a sub-gradient based power allocation algorithm in which SUs can iteratively adjust their transmit powers to approach the payoff of a NE. The other one is a Q-learning based relay selection algorithm which enables each SU to iteratively search for a NE-achieving relaying scheme.**

I. Introduction

A report published by FCC observes that most of the spectrum allocated to current mobile service subscribers was either unused or rarely used for most of the time [1]. This motivates the study of using the cognitive radio (CR) to solve the spectrum under-utilization problem. In CR networks, unlicensed users, known as secondary users (SU), learn from the environment and intelligently decide what and how can the licensed spectrum owned by the licensed users, known as primary users (PU), can be further utilized.

In this paper, we consider a special type of CR networks, called spatial spectrum sharing (SSS), in which PUs can maintain a certain level of QoS as long as their interference powers are lower than a tolerable threshold, called interference-temperature limit [2]. The channel capacity and information theoretic bounds for SSS based CR networks were first studied in [3]. The result was extended into fading channels in [4] in which optimal power allocation methods for SUs were derived under different fading channel models. Recently, it was shown that the performance of CR networks can be further improved if SUs could help each other during the transmission [5] [6]. However, it was also observed that multi-hop relaying introduces both flexibility and complexity to the network operation, and if parameters are not properly selected, it cannot provide any improvement over the direct transmission. Currently, finding effective ways to control and manage SSS based CR networks with multi-hop relaying, especially in networks with multiple secondary source-to-destination pairs, is still an open problem.

This paper focuses on optimizing parameters and operations of multi-user CR networks. This issue becomes more complex when multi-hop relaying is allowed because if SUs are selfish, they tend to keep forcing others to serve as relays without returning the favor. In this paper, we try to find a balanced point for the above network by using a game theoretic model in which each SU obtains benefits through both spectrum sharing by paying prices to PUs and multi-hop relaying by paying prices to SUs who serve as its relays. One advantage of our work over the previously reported results [3] [4] is that we not only consider the performance of one specific source-to-destination pair, but also investigate the interaction among SUs and that between SUs and PUs. More specifically, we define two types of pricing functions: the first one is charged by SUs for selling their relaying service, and the other is charged by PUs for selling the licensed spectrum. The objective for each SU is to choose proper parameters and strategies to improve its revenue and simultaneously minimize the prices paid to PUs and relays. The optimal power allocation methods for SUs are derived under three different assumptions. Except for extending two traditional assumptions [4] into our network model, i.e., transmitters know full channel state information (CSI) or only know average channel gains, we also introduce a new power allocation method which could optimize the performance of SUs even if transmitters cannot do any predictions but only know the current channel gains. The Nash equilibrium (NE) of the game is discussed, and the conditions under which all SUs choosing multi-hop relaying achieves the optimal NE are studied.

To simplify our illustration, we first consider a simple network model with two source-to-destination pairs and then extend our results to the K source-to-destination pair case. In addition, we investigate two new questions for large multi-user CR networks: how to distributively control and optimize the transmit powers of SUs and how to choose the "right" relaying schemes for each SU in terms of the number and quality of relays. To answer these questions, two distributed optimization algorithms are proposed. The first one is a sub-gradient based power allocation algorithm which enables SUs to iteratively adjust their powers to approach the performance of a NE. The other one is a Q-learning based relay selection algorithm for

Y. Xiao and G. Bi are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (email: xiao0021@e.ntu.edu.sg, egbi@ntu.edu.sg).

D. Niyato is with the School of Computer Engineering, Nanyang Technological University, Singapore (email: dniyato@ntu.edu.sg).

each SU to search for a NE-achieving relaying scheme. We discuss possible extensions of our work and present numerical results to demonstrate the performance improvement brought by using the multi-hop relaying in CR networks.

The rest of this paper is organized as follows. The network model and game formulation are presented in Section II. The game theoretic analysis for CR networks with two secondary source-to-destination pairs are presented in Section III. The results are extended into large multi-user CR networks in Section IV. The numerical results are presented in Section V and the paper is concluded in Section VI.

II. Network Model and Game Theoretic Setup

A. Network Model

Consider a CR network model in which K secondary source-to-destination pairs, S_1 to D_1, S_2 to D_2, $\ldots$, S_K to D_K, share the licensed spectrum with M PUs, $P_1, P_2, \ldots, P_M$. To simplify our illustration, in this paper, we only focus on the transmitter cooperation in which S_i can only serve as the relay for other secondary sources, i.e., S_j for $j \neq i, j \in [1, K]$. The results of the receiver cooperation can be similarly obtained. Let the transmit power of S_i used to send/forward signals for S_j be $w_{i,j}$ for $i, j \in [1, K]$. Let the channel gain between user l and m be $h_{l,m}$ where $l, m \in \{S_1, S_2, ..., S_K\} \cup \{D_1, D_2, ..., D_K\} \cup \{P_1, P_2, ..., P_K\}$

In our model, SUs can either directly send their signals or cooperate with each other by using the decode-and-forward multi-hop relaying. Assume that SUs use time division half duplex transmission mode to void self interference and each secondary source has been allocated the same amount of time for its own transmission. For example, consider a U time duration of transmission. If S_i chooses the direct transmission, it will directly send its own information to D_i in $\frac{U}{K}$ amount of time. However, if S_i chooses to use L_i-hop relaying to send its signals for $L_i \geq 2$, it only spends $\frac{U}{KL_i}$ in sending its source information to the closest relay and leaves all the rest $\frac{U(L_i-1)}{KL_i}$ amount of time for the relaying transmission (each relay will spend $\frac{U}{KL_i}$ in forwarding the signal of S_i). Note that, in our settings, the relaying scheme chosen by each secondary source does not affect the transmission of other source-to-destination pairs. To simplify our discussion, we define a time scheduler (TS) for the multi-user CR network as follows.

Assumption 1: We assume that there is a TS in the SU network. Two main objectives for the TS are: 1) to ensure that there is no collisions (more than one SU uses the licensed spectrum at the same time); 2) to arrange the relays and time for the transmission of each secondary source-to-destination pair. More specifically, if S_i chooses L_i-hop relaying, it will first send the request to the TS to schedule $(L_i - 1)$ intermediate SUs (relays) and the respective time segments for the relaying transmission. Note that the TS only provides the time scheduling information and each SU cannot obtain any other information from the TS. We assume the delay for each SU to wait for the availability of relays and the time scheduling information is negligible.

Note that a TS may not necessarily be a central controller. For example, SUs can follow a pre-defined protocol to schedule their transmissions, or they can compete/cooperate with each other for the limited transmission opportunities [7].

To simplify our problem, in this paper, we assume there is no pricing competition [8] and all PUs treat each SU equally. Hence, we can combine the effects of M PUs as one, labeled as PU 0. The interference power increment of PU 0 caused by S_i is given by $I_{S_i} = w_{i,j} h_{S_i,0}$ where $h_{S_i,0} = \sum_{k=1}^{M} h_{S_i,P_k}$.

B. Game Theoretic Setup

Define the following elements in our model: 1) *Players* are SUs who share the licensed spectrum, 2) *Strategies* of S_i, denoted by $s_i \in \mathcal{S}$ where $\mathcal{S}$ is the set of all strategies for SUs, are the options of actions. For example, in relay selection games, s_i can be either direct transmission or multi-hop relaying. In power allocation games, s_i can be different transmit powers of S_i, 3) *Revenue* of S_i, denoted by $r_{i,j}$, is the benefit earned by S_i for using PU 0's spectrum ($r_{i,i}$) and/or selling relaying services to S_j ($r_{i,j}$ for $j \neq i$ and $i, j \in \{1, 2, ..., K\}$), 4) *Pricing function* paid by S_i is the price charged by PU 0, denoted by $c_{0,i}$, and/or S_j, denoted by $c_{i,j}$, for selling the spectrum and/or providing relaying services, 5) *Payoff* of S_i, denoted by π_i, is the difference between its revenue and price, i.e., $\pi_i = \sum_{j=1}^{K} (r_{i,j} - c_{i,j})$.

Our model can be regarded as a special case of spectrum leasing in which PUs are regarded as the spectral resource owners and could lease a part of the resource to SUs in exchange for remuneration [9] [10] [11]. The price charged by PUs and SUs can either be used to exchange services with each other [9] or be the profits of the network operators [10].

The main objective for each SU is to choose proper parameters and strategies to maximize its "long term" payoff. Let us consider a repeated game with T time slots of transmission and use subscript $[t]$ to denote parameters during the tth time slot. In every time slot, each source-to-destination pair can send a certain amount of information. We assume the transmissions in different time slots are independent and each player cares for the performance of the future as same as that of the present i.e., the time discount factor [12, Definition 6.1.2] is 1. In this paper, transmit powers of SUs are limited in every time slot, and hence the power constraint is defined as follows:

$$E\left(\sum\nolimits_{j\in[1,K]} \sum\nolimits_{i\in[1,K]} h_{S_i,0[t]} w_{i,j[t]}\right) \leq Q_{0[t]}, \qquad (1)$$

where $w_{j,i[t]} = 0$ if S_i does not require S_j to serve as its relay for $i \neq j, i, j \in [1, K]$, and $Q_{0[t]}$ is the interference temperature limit of PU 0 in time slot t.

Since the main effect of allowing spectrum sharing is to increase the interference level of PU 0, it is reasonable to define the price charged by PU 0 from S_i to be a function of its resulting interference power increment caused by S_i, i.e., we have

$$\hat{c}_{0,i} = \sum\nolimits_{t=1}^{T} c_{0,i[t]} = \sum\nolimits_{t=1}^{T} \sum\nolimits_{j=1}^{K} \tilde{\alpha}_{0,j[t]} h_{S_j,0[t]} w_{j,i[t]}, \quad (2)$$

where $\tilde{\alpha}_{0,i[t]}$ is the pricing coefficient of PU 0 in time slot t defined as,

$$\tilde{\alpha}_{0,i[t]} = \left(b_{0,i[t]} w_{i,j[t]}\right)^{\eta} + a_{0,i[t]}. \qquad (3)$$

and $\eta, a_{0,i[t]}$ and $b_{0,i[t]}$ are constants in time slot t. Note that $\tilde{\alpha}_{0,i[t]}$ characterizes the interaction between SUs and PU 0. More specifically, a small $\tilde{\alpha}_{0,i[t]}$ indicates that a slight increment of interference power will not cause much QoS degrading for PU 0. However, if $\tilde{\alpha}_{0,i[t]}$ is large, PU 0 is less likely to share the spectrum with SUs and each SU needs to pay a relatively high price for accessing the licensed spectrum. In this paper, we mainly focus on two special types of $\tilde{\alpha}_{0,i[t]}$: $\tilde{\alpha}_{0,i[t]}$ is a constant ($c_{0,i[t]}$ is a linear pricing function of $w_{i,j[t]}$), i.e., $\eta = 0$, and $\tilde{\alpha}_{0,i[t]}$ is a linear function of the transmit power of S_i ($c_{0,i[t]}$ is a quadratic pricing function of $w_{i,j[t]}$), i.e., $\eta = 1$, which are considered in Subsections III-B and III-C, respectively. Assuming the decoding and re-encoding delay is negligible and the transmit power is the main resource used by each relay, we define the price charged by S_i for serving as the relay for the jth source-to-destination pair to be proportional to $w_{i,j}$, i.e., $\hat{c}_{i,j} = \sum_{t=1}^{T} c_{i,j[t]} = \sum_{t=1}^{T} \beta_{i,j} w_{i,j[t]}$, where $\beta_{i,j}$ is a constant. Assume each SU tries to maximize its data rate, and therefore the revenue of S_i should be proportional to its achievable rate $\mathbb{R}_i$, i.e., $\hat{r}_{i,i} = \sum_{t=1}^{T} r_{i,i[t]} = \sum_{t=1}^{T} \gamma_i \mathbb{R}_{i[t]}$, where $i, j \in \{1, 2, ..., K\}$, and γ_i is a constant which should be large if spectrum sharing is the only way to maintain a certain level of achievable rate, or small if there exist some alternative ways/resources for S_i to transmit important data. Revenues obtained by S_i to serve as the relay for S_j is defined as $\hat{r}_{i,j} = \sum_{t=1}^{T} r_{i,j[t]} = \sum_{t=1}^{T} \tilde{\beta}_{i,j} w_{i,j[t]}$, for $i \neq j, i, j \in \{1, 2, ..., K\}$ where $\tilde{\beta}_{i,j}$ is a constant.

Assumption 2: We assume only transmitters, i.e., S_1 and S_2, can decide which operation should be chosen (direct transmission or multi-hop relaying). Therefore, no SUs can obtain revenues by forcing others to relay information. Except from that, the revenue earned by S_i by serving as the relay of S_j can only be used for exchanging services with other SUs, i.e., S_k for $k \neq i, k, i, j \in [1, K]$. We assume, during T time slots of transmission, each SU maintains a balance between the overall price paid to buy relaying services from others and the revenue obtained for serving as relays, i.e., we have

$$\hat{r}_{i,-i} = \sum_{t\in[1,T]} \sum_{\substack{j\in[1,K] \\ j\neq i}} r_{i,j[t]} = \hat{c}_{-i,i} = \sum_{t\in[1,T]} \sum_{\substack{j\in[1,K] \\ j\neq i}} c_{j,i[t]}. \quad (4)$$

The above assumption can be satisfied for a relatively long period of transmission in which all SUs are selfish and always tend to spend all relaying revenues on improving their payoffs. In addition, we also assume no SUs can refuse the relaying request of others, i.e., a TS is built to ensure the transmission and relaying requests of SUs to be satisfied. These assumptions avoid the case that selfish SUs in the networks keep using multi-hop relaying and refuse to help others (free-riders) without being punished, or some SUs always forward signals for others without getting any benefits.

In the network with multiple secondary source-to-destination pairs, one of the main objectives is to find the NE. In other words, if the network operates at a NE, each SU cannot achieve a higher payoff by choosing a different strategy, given the strategies of other SUs. The formal definition of NE is given below.

Definition 1: [13, Definition 23.1] A strategy profile s^* is a NE if, for every player S_i and every strategy s_i for $s^*, s_i \in \mathcal{S}$ and $i \in [1, K]$, s^* is at least as good as the strategy profile (s_i, s_j^*) in which player S_i chooses s_i while every other player chooses s^*, i.e., for every player i, $\pi_i(s_i^*, s_{-i}^*) \geq \pi_i(s_i, s_{-i}^*)$, where subscript $-i$ denotes all players except S_i.

III. Game Theoretic Analysis for Two Source-to-Destination Pair Case

Let us consider the CR network model with two secondary source-to-destination pairs, i.e., S_1 to D_1 and S_2 to D_2.

A. Game Theoretic Analysis

Following the game theoretic model established in Subsection II-B, in the tth time slot, four strategy pairs $(s_{1[t]}, s_{2[t]})$ can be selected by S_1 and S_2 as discussed below.

1) Both S_1 and S_2 Using Direct Transmissions: In this case, the price paid by S_i for $i \in \{1, 2\}$ to PU 0 is given by $c_{0,i[t]}^{DT} = \tilde{\alpha}_{0,i[t]} h_{S_i,0[t]} w_{i,i[t]}$. S_i obtains the following revenue which is proportional to its achievable rate: $r_{i,i[t]}^{DT} = \gamma_i \log\left(1 + h_{S_i,D_i[t]} w_{i,i[t]}\right)$. Therefore, the payoff of S_i is given by

$$\pi_{i[t]}^{DT} = r_{i,i[t]}^{DT} - c_{0,i[t]}^{DT}. \quad (5)$$

2) Both S_1 and S_2 Using Multi-hop Relaying: In this case, the cost of S_i for $i \in \{1, 2\}$ contains two parts: one is charged by PU 0 from S_i (for sending the source information) and S_{-i} (for forwarding signals of S_i), denoted by $c_{0,i[t]}^{CT} = \tilde{c}_{0,i[t]}^{CT} + c_{0,-i[t]}^{\prime CT} = \tilde{\alpha}_{0,i[t]} h_{S_i,0[t]} w_{i,i[t]} + \tilde{\alpha}_{0,-i[t]} h_{S_{-i},0[t]} w_{-i,i[t]}$, and the other is charged by the relay (i.e., S_{-i}), denoted by $c_{i,-i[t]}^{CT} = \beta_{-i,i} w_{-i,i[t]}$. The revenue of S_i also contains two parts: one is proportional to its achievable rate $\mathbb{R}_{i[t]}$, denoted as

$$r_{i,i[t]}^{CT} = \frac{\gamma_i}{2} \min\left\{E \log\left(1 + h_{S_i,S_{-i}[t]} w_{i,i[t]}\right), E \log\left(1 + h_{S_{-i},D_i[t]} w_{-i,i[t]}\right)\right\}, \quad (6)$$

and the other is earned by serving as the relay for S_{-i}, denoted as $r_{i,-i[t]}^{CT} = \tilde{\beta}_{i,-i} w_{i,-i[t]}$. The payoff of S_i using multi-hop relaying is given by

$$\pi_{i[t]}^{CT} = r_{i,i[t]}^{CT} + r_{i,-i[t]}^{CT} - c_{0,i[t]}^{CT} - c_{i,-i[t]}^{CT}. \quad (7)$$

3) One SU Using Multi-hop Relaying and the Other SU Using Direct Transmission: Assume only S_i for $i \in \{1, 2\}$ uses multi-hop relaying. In this case, S_{-i} not only transmits its own information to D_{-i} but also serves as the relay for S_i. Following the same line as the previous cases, the payoff functions for S_i and S_{-i} are given by

$$\pi_{i[t]}^{CTD} = r_{i,i[t]}^{CT} - c_{0,i[t]}^{CT} - c_{i,-i[t]}^{CT}, \quad (8)$$

$$\pi_{-i[t]}^{DTC} = r_{-i,-i[t]}^{DT} + r_{-i,i[t]}^{CT} - c_{0,-i[t]}^{DT}. \quad (9)$$

From (8) and (9), we have the following results,

$$\pi_{i[t]}^{CTD} = \pi_{i[t]}^{CT} - r_{i,-i[t]}^{CT}, \qquad \pi_{-i[t]}^{DTC} = \pi_{-i[t]}^{DT} + r_{-i,i[t]}^{CT}. \quad (10)$$

The revenue earned by each SU for serving as the relay should have different meanings for different systems. In a large network, for example, each SU can use this revenue to buy the relaying service from others [9]. In some other systems, the price charged by relays can be simply used for avoiding free-riders[14].

B. Optimal Power Allocation Methods for CR Networks with Linear Pricing Function

As observed in the previous section, the revenue of each SU is closely related to its transmit power. Hence, finding the optimal transmit power for each SU is very important for the interference limited SSS systems. Let us introduce the following assumption.

Assumption 3: It is observed that the achievable rate of the multi-hop relay channel is always limited by the hop with the lowest capacity and hence, if the transmit powers of the source and relays are equal, there is always a part of the powers being wasted. This problem becomes more serious in our proposed model because a high transmit power also means a high price charged by relays and PU 0. One way to solve this problem is to adapt the forwarding powers of the relay to the CSIs of its connected channels. More specifically, if S_i chooses the multi-hop relaying, $w_{-i,i}$ should be given by

$$w_{-i,i[t]} = f_i w_{i,i[t]}, \tag{11}$$

where $f_i = h_{S_i,S_{-i}[t]}/h_{S_{-i},D_i[t]}$ for $i \in \{1,2\}$.

In this way, the overall achievable rate of multi-hop relaying channels should always be equal to that of the first hop and no powers will be wasted during the relaying transmission, i.e., $\mathbb{R}^{CT}_{i[t]} = E\log\left(1 + h_{S_i,S_{-i}[t]} w_{i,i[t]}\right)$.

Following Assumption 2, each SU (i.e., S_i) should balance the accumulated price paid to relaying SUs and the accumulated revenue obtained by serving as relays for others. In other words, we can actually neglect the effect of $c_{-i,i[t]}$ and $r_{i,-i[t]}$ in the payoff optimization problem because these two values will be eventually canceled during T time slots of transmission. Let us consider the linear pricing function case and assume $\tilde{\alpha}_{0,j[t]}$ of PU 0 in (3) to be a constant in time slot t, i.e., $\tilde{\alpha}_{0,j[t]} = \tilde{\alpha}_{0[t]}$. Since the transmissions in different time slots are independent with each other, in the rest of this section, we focus on the transmission in the tth time slot and drop the subscript $[t]$ to simplify the notation.

1) Only Average Channel Gains are Known by Transmitters: In this case, $w_{i,i}$ for $i \in \{1,2\}$ should be a constant. If both S_1 and S_2 directly transmit their signals to the corresponding destinations, $w_{i,i}$ should satisfy,

$$\begin{aligned} &E\left(h_{S_1,0}w_{1,1} + h_{S_2,0}w_{2,2}\right) \\ &\quad = w_{1,1}E\left(h_{S_1,0}\right) + w_{2,2}E\left(h_{S_2,0}\right) \le Q_0, \end{aligned} \tag{12}$$

where if we assume the statistics of $h_{S_1,0}$ and $h_{S_2,0}$ are fixed during each time slot, $E(h_{S_1,0})$ and $E(h_{S_2,0})$ are two constants which can be obtained through feedbacks of PU 0.

If the SNR is low, the optimal transmit power of S_i for $i \in \{1,2\}$ is given by

$$w_{i,i} = \left(\frac{\gamma_i}{\lambda E(h_{S_i,0})} - \frac{1}{E(h_{S_i,D_i})}\right)^+, \tag{13}$$

where λ needs to satisfy $\lambda \ge \frac{\gamma_1+\gamma_2}{Q+E(h_{S_1,0})/E(h_{S_1,D_1})+E(h_{S_2,0})/E(h_{S_2,D_2})}$ if $w_{1,1}, w_{2,2} > 0$.

Consider the case that both S_1 and S_2 use multi-hop relaying. Following Assumption 3, if the SNR is low, the optimal transmit power of S_i for $i \in \{1,2\}$ is given by

$$w_{i,i} = \left(\frac{\gamma_i}{\lambda\left[E(h_{S_i,0}) + E\left(h_{S_{-i},0}f_i\right)\right]} - \frac{1}{E(h_{S_i,S_{-i}})}\right)^+, \tag{14}$$

where λ needs to satisfy

$$\lambda \ge \frac{\gamma_1+\gamma_2}{2Q + \frac{E(h_{S_1,0})+E\left(h_{S_2,0}f_1\right)}{E\left(h_{S_1,S_2}\right)} + \frac{E(h_{S_2,0})+E\left(h_{S_1,0}f_2\right)}{E\left(h_{S_2,S_1}\right)}} \tag{15}$$

if both $w_{1,1}$ and $w_{2,2}$ are positive. The detailed proof is given in Appendix A. By using (10) and the same methods in Appendix A, we can directly obtain the optimal power allocation method for the case with one SU choosing direct transmission and the other one choosing multi-hop relaying.

Note that our results in (13) and (14) are different from the single source-to-destination pair case in [4] in the sense that $w_{i,i}$ is limited by both SU-to-PU and SU-to-SU channels.

2) Full CSIs are Known by Transmitters: In this case, $w_{i,i}$ can be adapted to instantaneous value of the channel gains to further increase the revenue of S_i. If both S_1 and S_2 use direct transmission, we can obtain the following optimal power allocation methods by maximizing $\pi_1^{DT} + \pi_2^{DT}$ over $w_{1,1}$ and $w_{2,2}$, respectively. The optimal transmit power of S_i for $i \in \{1,2\}$ is given by

$$w_{i,i} = \left(\frac{\gamma_i}{\lambda h_{S_i,0}} - \frac{1}{h_{S_i,D_i}}\right)^+ \tag{16}$$

where λ needs to satisfy $\lambda \ge \frac{\gamma_1+\gamma_2}{Q+h_{S_1,0}/h_{S_1,D_1}+h_{S_2,0}/h_{S_2,D_2}}$ if $w_{1,1}, w_{2,2} > 0$. Note that in (16), $\frac{\gamma_1}{\lambda h_{S_1,0}}$ and $\frac{\gamma_2}{\lambda h_{S_2,0}}$ are called "water-levels" of $w_{1,1}$ and $w_{2,2}$, respectively. In other words, S_i should not transmit any signals if h_{S_i,D_i} is lower than $\frac{\lambda h_{S_i,0}}{\gamma_i}$.

Similarly, if both S_1 and S_2 use multi-hop relaying, the optimal transmit power of S_i can be obtained by maximizing $\pi^{CT}_{1[t]} + \pi^{CT}_{2[t]}$ over $w_{i,i}$, which is given by

$$w_{i,i} = \left(\frac{\gamma_i}{\lambda\left(h_{S_i,0} + h_{S_{-i},0}f_i\right)} - \frac{1}{h_{S_1,S_2}}\right)^+, \tag{17}$$

where λ needs to satisfy $\lambda \ge \frac{\gamma_1+\gamma_2}{2Q+\left(h_{S_1,0}+h_{S_2,0}f_1\right)/h_{S_1,S_2}+\left(h_{S_2,0}+h_{S_1,0}f_2\right)/h_{S_2,S_1}}$ if both $w_{1,1}$ and $w_{2,2}$ are positive. The detailed proofs are given in Appendix B.

3) Only Current Channel Gains are Known by Transmitters: In previous subsections, we study two extreme cases: 1) transmitters do not know any CSIs except for the average channel gains, 2) full CSIs can be accurately predicted by transmitters. These assumptions may not always be applied. For example, in many practical systems, transmitters cannot predict any information, i.e., future CSIs or average channel gains. In this subsection, we propose a new power allocation method to solve the above problem. In our method, transmitters only observe current channel gains.

Following the network model in Section II, we assume that each time slot can be further divided into B short segments during each of which the channel fading coefficients can be regarded as constants. We use l to denote the index of segment for $l \in [1, B]$. Since SUs cannot do any predictions about future channel gains, the power constraint in (1) should be applied in every segment, i.e., we have the following power constraint,

$$\sum\nolimits_{j=1}^{2} \sum\nolimits_{i=1}^{2} h_{S_i,0}(l) w_{i,j}(l) \le Q_0. \tag{18}$$

Due to the space limit, in the rest of this subsection, we only present results for the case that both S_1 and S_2 use multi-hop relaying to transmit signals. The direct transmission case can be obtained by using the similar method. For the lth segment of transmission, $\pi_{i,i}^{CT}(l)$ can be re-written as

$$\begin{aligned} \pi_i^{CT}(l) &= \gamma_i \log\left(1 + h_{S_i,S_{-i}}(l) w_{i,i}(l)\right) \\ &\quad - \tilde{\alpha}_0 h_{S_i,0}(l) w_{i,i}(l) - \tilde{\alpha}_0 h_{S_{-i},0}(l) w_{-i,i}(l). \end{aligned} \tag{19}$$

Since $\pi_i^{CT}(l)$ is a quasi-concave function of $w_{i,i}$, the optimal transmit power of S_i can be easily obtained by maximizing (19) over $w_{i,i}$ if both $w_{1,1}$ and $w_{2,2}$ satisfy (18). However, if (18) cannot be satisfied, we need to find the optimal values of $w_{i,i}$ and $w_{-i,-i}$ in a linear function of $\sum_{i=1}^{2} \left(h_{S_i,0}(l) + h_{S_{-i},0}(l) f_i(l)\right) w_{i,i} = Q_0$. Therefore, the optimal transmit power for S_i in segment l is given by

$$\begin{aligned} w_{i,i}(l) &= \left(\frac{\gamma_i}{\lambda^*(l)\left(h_{S_i,0}(l) + h_{S_{-i},0}(l) f_i(l)\right)} \right. \\ &\qquad \left. - \frac{1}{h_{S_i,S_{-i}}(l)}\right)^+ \end{aligned} \tag{20}$$

where λ^* is given by,

$$\lambda^*(l) = \begin{cases} \frac{\gamma_i}{2Q_0 + \nu_{S_i}(l)}, & \text{if (18) is not satisfied and } w_{i,i} > 0, w_{-i,-i} = 0, \\ \frac{\gamma_1 + \gamma_2}{2Q_0 + \sum_{j=1}^{2} \nu_{S_j}(l)}, & \text{if (18) is not satisfied and } w_{1,1}, w_{2,2} > 0, \\ \tilde{\alpha}_0, & \text{Otherwise,} \end{cases} \tag{21}$$

and $\nu_{S_i}(l) = \left[h_{S_i,0}(l) + h_{S_{-i},0}(l) f_i(l)\right] / h_{S_i,S_{-i}}(l)$. From (20), it is observed that knowing $\tilde{\alpha}_0$ from PU 0 only improves the payoffs of SUs when the received power constraint defined in (1) is not tight, i.e., Q_0 being large enough.

C. Optimal Power Allocation Methods for CR Networks with Quadratic Pricing Function

In Subsection III-B, we studied a CR network model in which the price charged by PU 0 is a linear function of its interference power, i.e., $\tilde{\alpha}_{0,i[t]}$ is a constant. However, the pricing function for a practical system may be more complex. For example, in an interference sensitive network, even a slight increment of the interference power could cause a huge QoS degrading for PU 0. In this case, setting $\tilde{\alpha}_{0,i[t]}$ to be a high order polynomial function of the transmit power of S_i, i.e., $\eta > 1$ in (2), could accelerate the power decreasing process of SUs to avoid a long period of high interference to PU 0. In this subsection, we consider the case that the pricing function of PU 0 is a quadratic function of the transmit power of S_i, i.e., $\tilde{\alpha}_{0,i[t]}$ is given by

$$\tilde{\alpha}_{0,i[t]} = b_{0[t]} w_{i,j[t]} + a_{0[t]}. \tag{22}$$

Following the same line as Subsection III-B, we assume $\hat{c}_{i,-i} = \hat{r}_{-i,i}$ and focus on one time slot of transmission to neglect the subscript $[t]$. Let us first consider the case that both S_1 and S_2 directly transmit signals to D_1 and D_2, respectively. Following the same setting as Section III, the payoff function of S_i can be written as follows:

$$\pi_i^{DT} = \gamma_i E \log\left(1 + h_{S_i,D_i} w_{i,i}\right) - \left(b_0 w_{i,i} + a_0\right) h_{S_i,0} w_{i,i}.$$

By maximizing π_i^{DT} over $w_{i,i}$ under the constraint in (1), we can derive the optimal transmit power of S_i for $i \in \{1, 2\}$ as follows:

$$\begin{aligned} w_{i,i} &= \left(\sqrt{\frac{1}{4h_{S_i,D_i}^2} - \frac{a_0 + \lambda}{4b_0 h_{S_i,D_i}} + \frac{(a_0 + \lambda)^2}{16 b_0^2} + \frac{\gamma_i}{2b_0 h_{S_i,0}}} \right. \\ &\qquad \left. - \frac{1}{2h_{S_i,D_i}} - \frac{a_0 + \lambda}{4b_0}\right)^+, \end{aligned} \tag{23}$$

where λ is a constant to ensure the power constraint in (1) being satisfied.

Let us consider the case that both S_1 and S_2 use multi-hop relaying to transmit signals. In this setting, π_i^{CT} for $i \in \{1, 2\}$ can be re-written as follows:

$$\begin{aligned} \pi_i^{CT} &= \gamma_i E \log\left(1 + h_{S_i,S_{-i}} w_{i,i}\right) - \left(b_0 w_{i,i} + a_0\right) h_{S_i,0} w_{i,i} \\ &\quad - \left(b_0 w_{-i,i} + a_0\right) h_{S_{-i},0} w_{-i,i}. \end{aligned} \tag{24}$$

By maximizing (24) over $w_{i,i}$ and using Assumption 3, the optimal transmit power for S_i with full CSIs known by transmitters is given by

$$\begin{aligned} w_{i,i} &= \left(\sqrt{\frac{1}{4h_{S_i,S_{-i}}^2} - \frac{D_i}{4A_i h_{S_i,S_{-i}}} + \frac{D_i^2}{16A_i^2} + \frac{\gamma_i}{2A_i}} \right. \\ &\qquad \left. - \frac{1}{2h_{S_i,S_{-i}}} - \frac{D_i}{4A_i}\right)^+, \end{aligned} \tag{25}$$

where A_i and D_i are given by

$$A_i = b_0 \left(h_{S_i,0} + h_{S_{-i},0} \frac{h_{S_i,S_{-i}}^2}{h_{S_i,D_i}^2}\right), \tag{26}$$

$$D_i = (a_0 + \lambda) \left(h_{S_i,0} + h_{S_{-i},0} \frac{h_{S_i,S_{-i}}}{h_{S_i,D_i}}\right), \tag{27}$$

and λ is a constant to ensure that $w_{i,j}$ for $i, j \in \{1, 2\}$ satisfies the received power constraint in (1). Note that $w_{i,i}$ in (25) is different from the result in (17) and is always larger than zero if $\gamma_i > \frac{D_i}{2h_{S_i,S_{-i}}}$. By using the same method discussed in Section III-B, we can calculate results for the cases that only average or current CSIs are known by transmitters.

D. Cooperative Nash Equilibrium

Based on the above discussion, there are four strategy pairs in each time slot of the considered game: $\mathcal{S} = \{(\pi_{1[t]}^{DT}, \pi_{2[t]}^{DT}), (\pi_{1[t]}^{CTD}, \pi_{2[t]}^{DTC}), (\pi_{1[t]}^{DTC}, \pi_{2[t]}^{CTD}), (\pi_{1[t]}^{CT}, \pi_{2[t]}^{CT})\}$. Assume that each SU fixes its transmission scheme during T time slots of communication. To simplify our discussion, we only focus on the symmetric network and assume $\hat{\pi}_1^{CT} = \hat{\pi}_2^{CT}$, $\hat{\pi}_1^{DT} = \hat{\pi}_2^{DT}$, $\hat{\pi}_1^{CTD} = \hat{\pi}_2^{CTD}$, and $\hat{\pi}_1^{DTC} = \hat{\pi}_2^{DTC}$ in this subsection. From Definition 1, we have the following results about the NE supported by the 2 source-to-destination pair game,

1) If $\hat{\pi}_i^{DT} > \hat{\pi}_i^{CT}$ for $i \in \{1, 2\}$, then $(\hat{\pi}_1^{DT}, \hat{\pi}_2^{DT})$ is the only NE of the game,
2) If $\hat{\pi}_i^{CT} \geq \hat{\pi}_i^{DT}$ for $i \in \{1, 2\}$, then $(\hat{\pi}_1^{CT}, \hat{\pi}_2^{CT})$ is a NE. Furthermore, if we assume $\hat{\pi}_i^{CT} \geq \hat{\pi}_i^{DTC}$, then $(\hat{\pi}_1^{CT}, \hat{\pi}_2^{CT})$ is the only NE of the game.

Result 1) defines the conditions under which both S_1 and S_2 directly transmitting signals to the corresponding destinations achieve the unique NE of the game. More specifically, in our game theoretic model, S_1 and S_2 should not use multi-hop relaying to transmit signals if the extra revenue brought by using multi-hop relaying is smaller than the increased price charged by PU 0 from the relay, i.e., we have

$$\gamma_i \Delta \hat{\mathbb{R}}_i^{CT-DT} = \sum_{t=1}^{T} \gamma_i \mathbb{R}_{i[t]}^{CT} - \sum_{t=1}^{T} \gamma_i \mathbb{R}_{i[t]}^{DT} \leq \sum_{t=1}^{T} c_{0,-i[t]}^{\prime CT}. \quad (28)$$

This setting makes each SU to cautiously decide on whether or not to use multi-hop relaying and hence avoid the scenarios that some SUs impulsively request the relaying service of others simply because they want to spend the revenues earned by selling their relaying services during previous time slots.

As observed in [13], a NE is generally not the optimal solution. However, it can be shown that, if the condition in result 2) is satisfied, applying the multi-hop relaying to both S_1 and S_2 will be the unique NE as well as the payoff maximization solution of our game. Let us summarize this observation in the following Lemma.

Lemma 1: Assume each SU fixes its transmission scheme during T time slots of transmission. Using multi-hop relaying for both S_1 and S_2 is the optimal NE of symmetric SSS based CR networks if the improvement of the weighed achievable rate $\gamma_i \Delta \hat{\mathbb{R}}_{i[t]}^{CT-DT}$ for S_i is larger than the sum of the price paid to S_{-i} and the price charged by PU 0 from the relay (S_{-i}) for $i \in [1, 2]$, i.e.,

$$\gamma_i \Delta \hat{\mathbb{R}}_i^{CT-DT} > \hat{c}_{X,i}, \quad (29)$$

where $\hat{c}_{X,i} = \sum_{t=1}^{T} \left(c_{0,-i[t]}^{\prime CT} + c_{-i,i[t]}^{CT} \right)$.

IV. Extending to Large CR Networks

In this section, we extend the results in Section III into a general CR network model in which K sources transmit signals to the corresponding destinations over the licensed spectrum. To simplify our illustration, we mainly focus on the operations of one source-to-destination pair, labeled by S_1 to D_1, and assume there are $L-1$ intermediate SUs, labeled as $S_2, S_3, \ldots, S_L$ for $L \leq K$, available to decode-and-forward signals for S_1. Note that these intermediate SUs are also secondary sources which have their own information to transmit. Therefore, if S_1 wants to exploit intermediate SUs to relay its signals, it needs to pay prices to each of these SUs and returns the favor when necessary. Let the set of all SUs who require the relaying service of S_1 in time slot t be $\tilde{\mathcal{L}}_{1[t]}$ for $|\tilde{\mathcal{L}}_{1[t]}| \leq K-1$. Following the network model and game theoretic notations introduced in Section II, we define the revenue of S_1 to be $\hat{r}_{1,1}^{CT} = \sum_{t=1}^{T} r_{1,1[t]}^{CT} = \sum_{t=1}^{T} \gamma_1 \mathbb{R}_{1[t]}$ where γ_1 is a constant. Due to space limitation, let us only provide the results for the linear pricing function case as discussed in Subsection III-B. Define the revenue of S_1, earned by being relays for other SUs, as $\hat{r}_{1,\tilde{\mathcal{L}}_1}^{CT} = \sum_{t=1}^{T} r_{1,\tilde{\mathcal{L}}_{1[t]}}^{CT} = \sum_{t=1}^{T} \sum_{i \in \tilde{\mathcal{L}}_{1[t]}} \tilde{\beta}_{1,i} w_{1,i[t]}$, where $\tilde{\beta}_{1,i}$ is a positive constant. The prices charged by intermediate SUs from S_1 are $\hat{c}_{\bar{L},1}^{CT} = \sum_{t=1}^{T} c_{\bar{L},1[t]}^{CT} = \sum_{t=1}^{T} \sum_{i=2}^{L} \beta_{i,1} w_{i,1[t]}$, where $\bar{L} = [2, 3, ..., L]$. If S_1 uses multi-hop relaying, it also needs to pay price to PU 0 for the relaying transmission of $S_2, S_3, \ldots, S_L$. Define the price charged by PU 0 to be $\hat{c}_{0,1}^{CT} = \sum_{t=1}^{T} c_{0,1[t]}^{CT} = \sum_{t=1}^{T} \sum_{i=2}^{L} \tilde{\alpha}_{0,i} h_{S_i,0[t]} w_{i,1[t]}$. Following Assumption 2, we have $\hat{r}_{1,\tilde{\mathcal{L}}_1}^{CT} = \hat{c}_{\bar{L},1}^{CT}$. Using this setting, the payoff function of S_1 can be written as,

$$\pi_1^{CT} = \sum\nolimits_{t=1}^{T} \Big[\gamma_1 \mathbb{R}_{1[t]} + \sum\nolimits_{j \in \tilde{\mathcal{L}}_{1[t]}} \tilde{\beta}_{1,j} w_{1,j[t]} \quad (30)$$
$$- \sum\nolimits_{k=2}^{L} \beta_{k,1} w_{k,1[t]} - \sum\nolimits_{l=1}^{L} \tilde{\alpha}_{0,l} h_{S_l,0[t]} w_{l,1[t]} \Big].$$

In a CR network with K source-to-destination pair, the main objective is to solve the following optimization problem:

$$\max_{w_{i,j[t]}} \sum\nolimits_{t=1}^{T} \sum\nolimits_{i=1}^{K} \pi_{i,i[t]} \quad (31)$$
$$\text{s.t.} \sum\nolimits_{j=1}^{K} \sum\nolimits_{i=1}^{K} E\left(h_{S_i,0[t]} w_{i,j[t]} \right) \leq Q_{0[t]}.$$

All results presented in the previous section can be directly applied into the above model. For example, let us consider the transmission from S_1 to D_1 with $(L-1)$-hop relaying. By assuming the transmission of SUs in each time slot to be independent and transmit powers of SUs to satisfy $w_{i,1[t]} = f_{i-1[t]} w_{i-1,1[t]}$ and $w_{L,1[t]} = \frac{h_{S_{L-1},S_L[t]}}{h_{S_L,D_1[t]}} w_{L-1,1[t]}$ where $f_{i-1[t]} = \frac{h_{S_{i-1},S_i[t]}}{h_{S_i,S_{i+1}[t]}}$ for $i \in [2, L-1]$, the optimal transmit power for S_1 with full CSIs known by transmitters can be obtained as follows:

$$w_{1,1[t]} = \left(\frac{\gamma_1}{\lambda (h_{S_1,0[t]} + \sum_{i=2}^{L-1} h_{S_i,0[t]} f_{i-1})} - \frac{1}{h_{S_1,S_2[t]}} \right)^{+}, \quad (32)$$

where λ is a constant to ensure transmit powers of all SUs satisfying the subjective function in (31). Similarly, the optimal power allocation methods under other operations and assumptions discussed in Subsections III-B and III-C can be directly obtained. We omit the detailed results due to the limit of space.

It is observed that in large multi-user networks, the network structures and user interactions greatly increase the complexity of the optimization problem. More specifically, two new questions that are naturally raised for this network are:

Q1) How to manage and optimize the transmit powers for CR networks in a distributed fashion?

Consider the transmission in time slot t with B time segments,

1) *Initialization Step*:
   S_i chooses the initial powers $w_{i,i[t]}(1)$ and $w_{i,i[t]}(2)$ to transmit signals $x_{i,i[t]}(1)$ and $x_{i,i[t]}(2)$ during the first two time segments. After receiving signals $x_{i,i[t]}(1)$ and $x_{i,i[t]}(2)$, D_i feedbacks the revenue $r_{i,i[t]}(1)$ and $r_{i,i[t]}(2)$ to S_i. PU 0 and relaying SUs also send the prices $c_{0,i[t]}(1)$, $c_{0,i[t]}(2)$, $c_{k,i[t]}(1)$and $c_{k,i[t]}(2)$ to S_i for $k \neq i, i, k \in [1, K]$. At the end of the second segment, S_i calculates payoffs $\pi_{i[t]}(1)$ and $\pi_{i[t]}(2)$.
2) *Iteration Step*: For $l = 3: B$,
   - If the power constraint satisfies (1), S_i sends signal $x_{i,i[t]}(l)$ using power $w_{i,i[t]}(l) = \left(w_{i,i[t]}(l-1) + v(l-1)g_{i,i[t]}(l-1)\right)^+$ where $(a)^+ = \max\{0, a\}$, $g_{i,i[t]}(l)$ is the sub-gradient of $\pi_{i[t]}$ which can be calculated by $\pi_{i[t]}\left(w_{i,i[t]}(l)\right) \geq \pi_{i[t]}\left(w_{i,i[t]}(l-1)\right) + g_{i,i[t]}(l)(w_{i,i[t]}(l) - w_{i,i[t]}(l-1))$, and $v(l) = \frac{u}{\sqrt{l}}$ where u is the iteration step size.
   - If PU 0 detects higher than tolerable interference power (Power constraint in (1) is not satisfied), it sends the necessary information for each SU to calculate the projection of its transmit power $w_{i,i[t]}(l-1)$ on the power constraint in (1), i.e., $w_{i,i[t]}(l)$ is given by
   $$w_{i,i[t]}(l) = \mathcal{P}(\boldsymbol{w}_{[t]}(l-1))^+$$
   where $\boldsymbol{w}_{[t]}(l-1) = \left[w_{1,1[t]}(l-1), w_{2,2[t]}(l-1), \ldots, w_{K,K[t]}(l-1)\right]^T$ and $\mathcal{P}(\boldsymbol{a})$ denotes the projection function of $\boldsymbol{a}$ on (1) defined by
   $$\mathcal{P}(\boldsymbol{a}) = \boldsymbol{a} - \boldsymbol{h}^T(\boldsymbol{h}\boldsymbol{h}^T)^{-1}(\boldsymbol{h}\boldsymbol{a} - Q_0)$$
   where $h = [h_{S_1,0}, h_{S_2,0}, \ldots, h_{S_K,0}]$.

   End
3) *Termination Step*: The above process continues until $l = B$.

Fig. 1. Illustration of Algorithm 1.

Q2) How to choose "proper" relays (in terms of the number and quality) for each secondary source-to-destination pair to maximize the payoff?

In the following subsections, we propose two distributed algorithms to solve the above problems. The first one is a distributed power allocation method and the other is a distributed relay selection algorithm.

A. A Distributed Power Allocation Algorithm

As discussed in Section III-B, achieving the optimal payoff without allowing the transmitters to know the CSIs is a very challenging task. In this subsection, we assume the channel gains can be regarded as constants during each time slot and the transmission scheme selected by each SU is fixed. A distributed algorithm, denoted as Algorithm 1, is presented in Figure 1. The proposed algorithm does not require the CSIs to be known by transmitters or SUs to communicate with each other. Following the same methods in [15], it can be proved that, by using Algorithm 1, the payoffs of SU networks approaches to a NE if the number of iterations is large enough. In Figure 2, we show the time average payoff of S_1 under different numbers of iterations. It is observed that the convergence performance of Algorithm 1 is closely related to the iteration step size u. Generally speaking, the larger the value of u, the faster rate of convergence to the payoff of a NE. However, if u is too large, the average payoff will be fluctuated during the first few iterations.

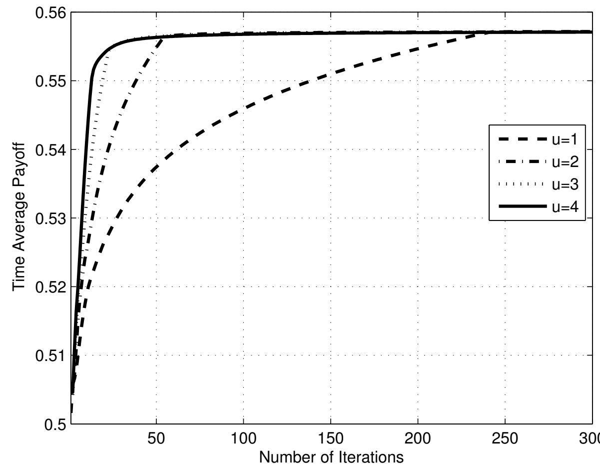

Fig. 2. Convergence rate of Algorithm 1 with different step sizes.

Consider the transmission in time slot t with B time segments,

1) *Initialization Step*: During the first segment, each SU (i.e., S_i) chooses a strategy (transmission scheme) $s_{i[t]}(1) = j$ to send signals. D_i feedbacks a successful-decoding message to S_i after the decoding process finished. Assume this feedback message can be eavesdropped by the TS to make sure all signals sent by secondary sources to be successfully decoded by their destinations in each time slot. S_i, after receiving the successful-decoding message from D_i and the pricing functions from relaying SUs and PU 0, calculates $\pi_{i[t]}(s_{i[t]}(1) = j)$ for its strategy j. We use $\mathcal{S}$ to denote the set of all the possible strategies that can be chosen by S_i. Assume S_i chooses the initial value of **Q**-function as $\boldsymbol{Q}_{S_i[t]}\left(s_{i[t]}(1) = j\right) = 0$.
2) *Iteration Step*: For $l = 2: B$,
   - S_i updates the value of $\boldsymbol{Q}$-function as follows,
   $$\boldsymbol{Q}_{S_i[t]}\left(s_{i[t]}(l) = j\right) = \boldsymbol{Q}_{S_i}\left(s_{i[t]}(l-1) = k\right) + \Delta_{[t]}(l)\left(\pi_{i[t]}\left(s_{i[t]}(l-1) = k\right)\Pr\left(s_{i[t]}(l-1) = k\right)\right)$$
   where $\Delta_{[t]}(l)$ is the iteration step size defined by $\Delta_{[t]}(l) = \frac{\Delta_{0[t]}}{l}$ and $\Delta_{0[t]}$is a constant in time slot t.
   - S_i randomly chooses its strategies $j \in \mathcal{S}$ according to the Boltzmann distribution,
   $$\Pr\left(s_{i[t]}(l) = j\right) = \frac{\exp\left(\boldsymbol{Q}_{S_i}(s_{i[t]}(l) = j)/\rho\right)}{\sum_{k \in \mathcal{S}} \exp\left(\boldsymbol{Q}_{S_i}(s_{i[t]}(l) = k)/\rho\right)}$$
   where $\Pr\left(s_{i[t]}(l) = j\right)$ is the probability of S_i choosing strategy j in segment l.
3) *Termination Step*: The above process stops if $l = B$.

Fig. 3. Illustration of Algorithm 2.

B. A Distributed Relay Selection Algorithm

Following the same setting as Algorithm 1, channel gains are assumed to be constants and the number of SUs who have positive powers is fixed in each time slot. Let us divide the time duration of each time slot into B small segments. Define the **Q**-function for S_1 in segment l of time slot t as $\mathbf{Q}_{S_1[t]}(s_1(l)) = E[\pi_{1[t]}(s_1(l))]$, where s_1 denotes the transmission scheme selected by S_1, i.e., which and how many nearby SUs should be chosen as relays of S_1. Assume there

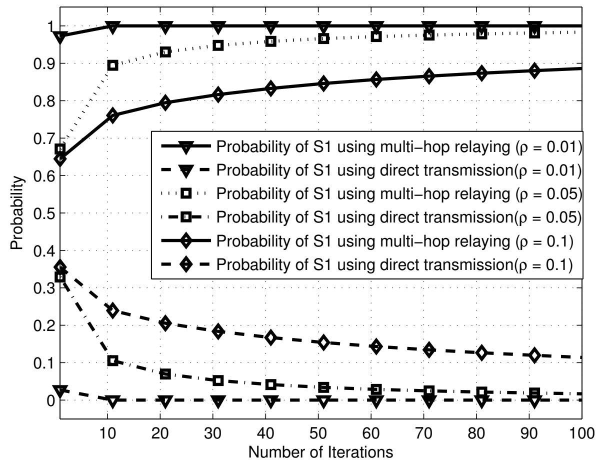

Fig. 4. Convergence rate of Algorithm 2 with different values of ρ.

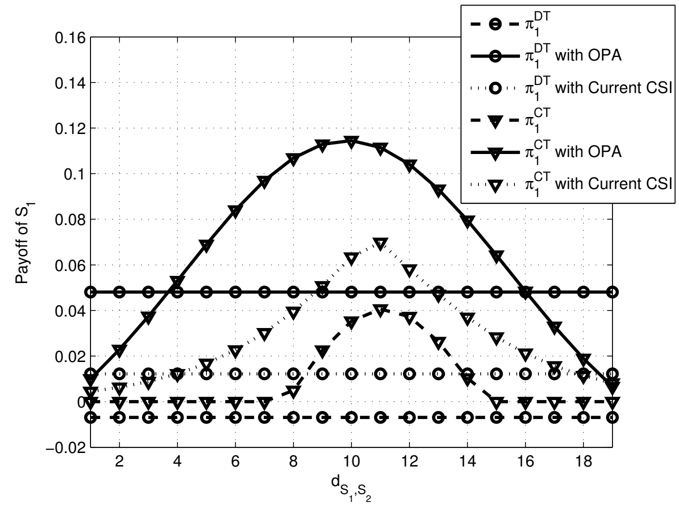

Fig. 5. Numerical results for the payoff of S_1 in SSS based CR networks with multi-hop relaying

are L intermediate SUs located between S_1 and D_1, and hence there are $2^{(L-1)}-1$ strategies can be chosen. The operations of Algorithm 2 are described in Figure 3. Note that this algorithm is similar to Robbins-Monro algorithm [16], and hence $\mathbf{Q}_{S_1[t]}$ will converge to the average payoff of strategy s_1. Each SU will be more likely to choose the scheme with the highest $\mathbf{Q}$ value which will eventually approach to a NE-achieving relaying scheme [17]. The convergence proof is similar to that of Robbins-Monro algorithm and hence is omitted here. The convergence rate of Algorithm 2 is closely related to the value of ρ. In Figure 4, we consider the 2 source-to-destination pair case and present the convergence performance of the probability of S_1 choosing different schemes (either direct transmission or multi-hop relaying). Note that in the proposed Algorithm, the value of ρ controls the frequency of the exploration, i.e., the smaller value of ρ, the higher probability for S_i to choose the high $\mathbf{Q}$ value strategy. This explains the observation in Figure 4 that the convergence rate of Algorithm 2 is decreased with ρ.

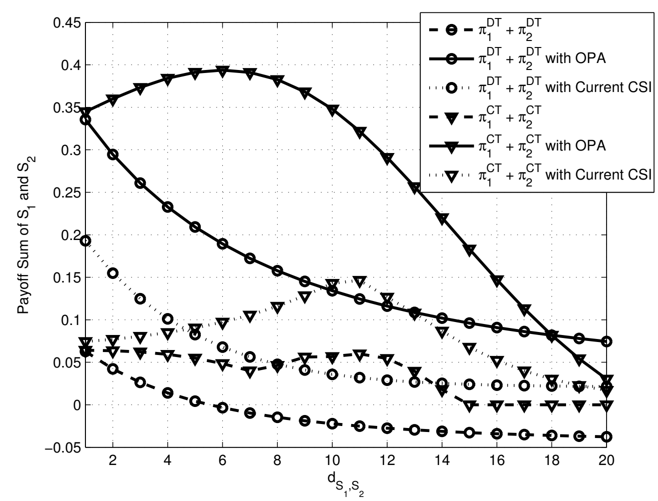

Fig. 6. Numerical results for the payoff sum of S_1 and S_2 in SSS based CR networks with multi-hop relaying

V. Numerical Results

As discussed in Section IV, a large multi-user CR network can be regarded as an extended version of the network with 2 source-to-destination pair. Therefore, in this section, we will first present numerical results for the CR network model discussed in Section III and leave the results of K source-to-destination pair case to the end of this section.

Denote the distance between two SUs as $d_{i,j}$ for $i,j \in \{S_1, S_2, D_1, D_2, 0\}$. Let us consider the following channel models,

$$h_{i,j} = \tilde{h}_{i,j}/d_{i,j}^{\delta}, \tag{33}$$

where $\tilde{h}_{i,j}$ is a Rayleigh distribution random variable unrelated to the distance of transmission and δ is the channel attenuation exponent. To simplify our illustration, we focus on two types of networks: 1) *symmetric networks* in which the locations of two secondary source-to-destination pairs are geometrically symmetric, i.e., $d_{S_1,D_1} = d_{S_2,D_2} = d_{S_1,S_2} + d_{S_2,D_1} =$

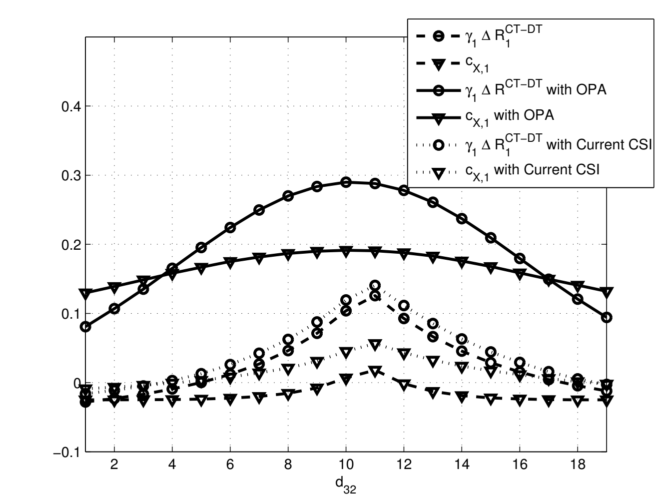

Fig. 7. Numerical results for the payoff of S_1 in SSS based CR networks with multi-hop relaying

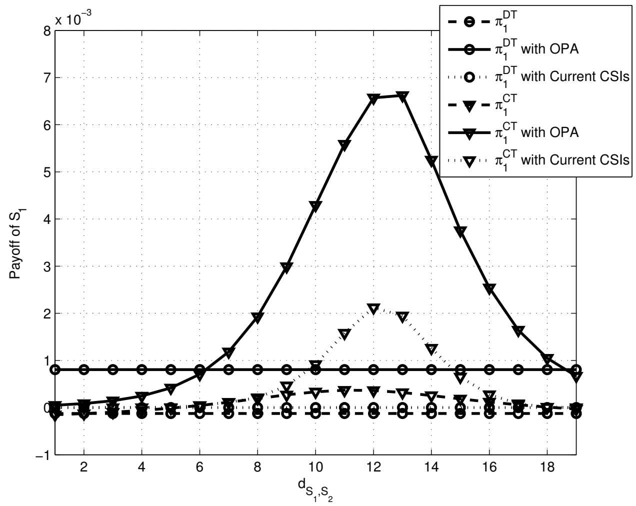

Fig. 8. Numerical results for SSS based CR networks with multi-hop relaying

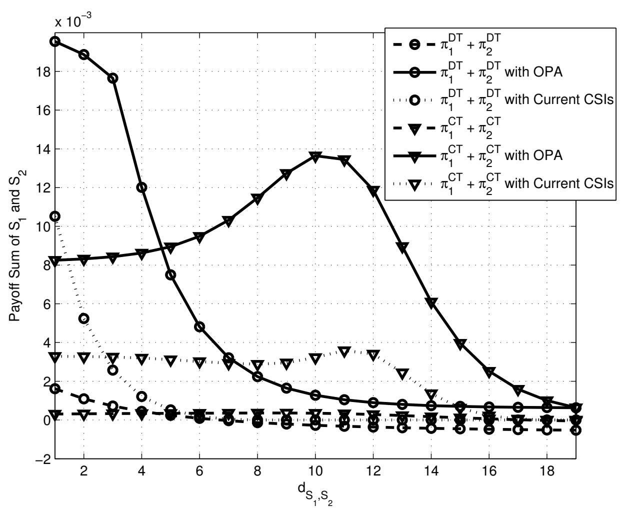

Fig. 9. Numerical results for the payoff sum of S_1 and S_2 in SSS based CR networks with multi-hop relaying

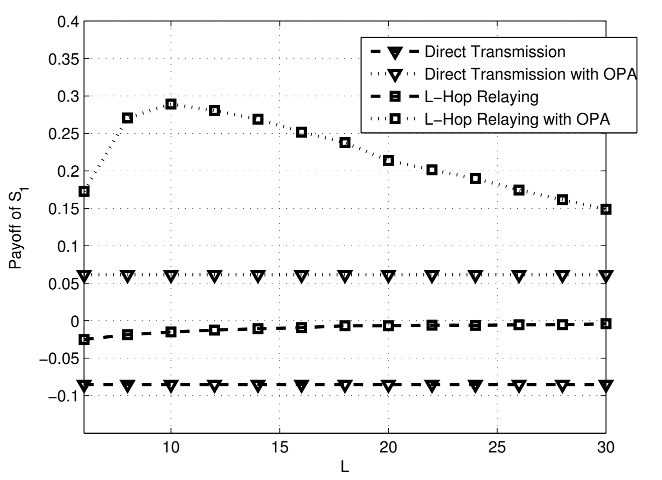

Fig. 10. Numerical results for SSS based CR networks with L-hop relaying

$d_{S_2,S_1}+d_{S_1,D_2}$, and hence we fix d_{S_1,D_1} to study the payoff of S_1 under different d_{S_1,S_2}, 2) *asymmetric networks* in which d_{S_1,D_2} and d_{S_1,D_1} are fixed (i.e., $d_{S_1,D_2}=\frac{d_{S_1,D_1}}{2}$) and the payoff sum of S_1 and S_2 under different d_{S_1,S_2} are considered. Based on this setting, the difference between $E(h_{S_1,S_2})$ and $E(h_{S_2,D_1})$ (or between $E(h_{S_1,S_2})$ and $E(h_{S_1,D_2})$) will become small if d_{S_1,S_2} approaches to $\frac{d_{S_1,D_1}}{2}$, or large if d_{S_1,S_2} approaches to 0 or d_{S_1,D_1}.

Consider symmetric networks with a linear pricing function first. It is observed in Figure 5 that the payoff of S_1 using multi-hop relaying is closely related to the location of the relay. More specifically, if S_2 is located near the middle point between S_1 and D_1, multi-hop relaying could greatly improve the performance of CR networks. However, the improvement decreases when d_{S_1,S_2} approaches to 0 or d_{S_1,D_1}. This is because, by following Assumption 3, if $E(h_{S_1,S_2})$ is much larger than $E(h_{S_2,D_1})$, S_2 will raise its forwarding power to ensure all received signals to be successfully decoded by D_1, which also increases the price charged by PU 0. Note that the optimal location of S_2 which maximizes the payoff of S_1 can be obtained by substituting (33) into π_1^{CT} and maximizing the π_1^{CT} over d_{S_1,S_2}. The detailed derivation is omitted due to space limit. Another interesting result in Figure 5 is that if S_2 is located in the middle point between S_1 and D_1, multi-hop relaying with fixed transmit powers can provide a higher payoff than the direct transmission with optimal power allocation methods. In other words, multi-hop relaying can be a more useful tool than optimal power allocation methods to improve the performance of SSS based CR networks under certain conditions. Numerical results for asymmetric networks with linear pricing function are presented in Figure 6. It is observed that, the payoff sum of S_1 and S_2 is decreased with d_{S_1,S_2} if both secondary sources use direct transmission. However, when using multi-hop relaying, the payoff sum can be greatly improved, especially when d_{S_1,S_2} approaches to $\frac{d_{S_1,D_1}}{2}$.

Note that the value of payoff in our setting can be negative. For some systems, this could mean the price is too high for SUs to use the licensed spectrum or some specific strategies. In this case, SUs should either use alternative strategies and optimization methods, such as multi-hop relaying, optimal power allocation methods, etc., to increase their payoffs, or stop sending signals over the licensed spectrum.

As discussed in Lemma 1, using multi-hop relaying for both secondary sources achieves the performance of a NE if the weighed achievable rate gain $\gamma_i\Delta\mathbb{R}_i^{CT-DT}$ for S_i is larger than the price $\hat{c}_{X,i}$ charged by PU 0 and relays for $i\in[1,2]$. Therefore, we present numerical results for the symmetric CR network in Figure 7 in which $\gamma_1\Delta\mathbb{R}_1^{CT-DT}$ defined in the left-hand-side of (29) and $\hat{c}_{X,1}$defined in the right-hand-side of (29) are compared. It is observed that multi-hop relaying can always be a NE-achieving scheme if the location of relay is close to the middle point between source and destination, i.e., $d_{S_1,S_2}\approx\frac{d_{S_1,D_1}}{2}$. However, $\gamma_1\Delta\mathbb{R}_1^{CT-DT}$ will be exceeded by $\hat{c}_{X,1}$ if the relay is located close to either source or destination.

Let us consider CR networks with quadratic pricing functions. Figure 8 shows the payoff of S_1 when the network is symmetric. It is observed that with the increasing of d_{S_1,S_2}, the payoff of S_1 with multi-hop relaying changes in a much faster speed than that in the linear pricing function case. This is because a part of the price charged by PU 0 from the relay is a second degree polynomial function of the forwarding power which, as mentioned in Assumption 2, should be proportional to $\frac{h_{S_1,S_2[t]}}{h_{S_2,D_1[t]}}$. In addition, similar to the linear pricing function case, multi-hop relaying cannot provide a higher payoff than that of direct transmission when the relay is close to either the source or destination. For the asymmetric network, it is observed in Figure 9 that the payoff sum is decreased with d_{S_1,S_2} if both S_1 and S_2 use direct transmission because the revenue of S_2 is decreased with $d_{S_2,D_2} = d_{S_1,D_2} + d_{S_1,S_2}$. However, multi-hop relaying is an effective way to alleviate the payoff decreasing process caused by the increasing of d_{S_1,D_1} especially when the relay approaches to the middle point between the source and destination.

In Figure 10, we consider CR networks with multiple source-to-destination pairs and assume there are $L-1$ intermediate SUs located between S_1 and D_1. Similar to the 2 source-to-destination pair case, it is shown that $(L-1)$-hop relaying provides a great performance improvement for CR networks. It is also observed that there is a tradeoff between the number of relays and payoff of S_1. More specifically, with the increasing of L, the distance of each hop is reduced, which increases the revenue (or achievable rate). On the other hand, the time duration used for S_1 to transmit the source information will be shorten, which decreases the revenue.

VI. Conclusion

This paper has exploited the game theoretic model to study the optimization problem for SSS based multi-user CR networks with multi-hop relaying. In the considered system, each secondary source can obtain benefits through both spectrum sharing by paying prices to PUs and multi-hop relaying by paying prices to the relaying SUs. The payoff of secondary users under different settings and different types of pricing functions of primary users have been studied. Furthermore, we have proposed three power allocation methods corresponding to the assumptions that transmitters have average knowledge, full knowledge and current knowledge of channel gains. The NE for the proposed game under different settings have been discussed and the conditions for which applying multi-hop relaying to all secondary users achieves the optimal NE have been considered. Our results are extended into the multi-user CR network model with K secondary source-to-destination pair. Two distributed algorithms have been proposed for each secondary source to search for the NE-achieving power allocation method and relaying scheme, respectively. Numerical results are presented to show the performance of CR networks with multi-hop relaying in both symmetric and non-symmetric network cases.

Acknowledgment

The authors would like to thank Professor Husheng Li for helpful comments on the initial version of the manuscript.

Appendix A
Optimal Power Allocation Methods for Direct Transmission

Consider the network model in Section II. Assuming the transmissions for each secondary source during different time slots are independent with each other. Our objective is to maximize the payoff sum of SUs during the tth time slot. More specifically, we need to find the solution of the following optimization problem:

$$\begin{aligned} \max_{w_{1,1[t]},w_{2,2[t]}} \quad & \pi^{DT}_{1[t]} + \pi^{DT}_{2[t]} \\ \text{s.t.} \quad & w_{1,1[t]}E\left(h_{S_1,0[t]}\right) + w_{2,2[t]}E\left(h_{S_2,0[t]}\right) \le Q_{0[t]}. \end{aligned} \tag{34}$$

If the SNR is low, we have $E(\log(1+SNR)) \approx \log(1+E(SNR))$. Therefore, the Lagrange multiplier of (34) is given by

$$\begin{aligned} LM^{DT} = {} & \gamma_1 \log\left(1 + E(h_{S_1,D_1[t]})w_{1,1[t]}\right) \\ & +\gamma_2 \log\left(1 + E(h_{S_2,D_2[t]})w_{2,2[t]}\right) - \tilde{\alpha}_{0[t]}E(h_{S_1,0[t]})w_{1,1[t]} \\ & -\tilde{\alpha}_{0[t]}E(h_{S_2,0[t]})w_{2,2[t]} + \lambda'_t\left(Q_{0[t]} - w_{1,1[t]}E(h_{S_1,0[t]})\right. \\ & \left. +w_{2,2[t]}E(h_{S_2,0[t]})\right). \end{aligned} \tag{35}$$

We have KKT conditions of the above equation as follows,

$$\frac{\partial LM^{DT}}{\partial w_{1,1[t]}} = 0 \Rightarrow w_{1,1[t]} = \left(\frac{\gamma_1}{\lambda E(h_{S_1,0[t]})} - \frac{1}{E(h_{S_1,D_1[t]})}\right)^+, \tag{36}$$

$$\frac{\partial LM^{DT}}{\partial w_{2,2[t]}} = 0 \Rightarrow w_{2,2[t]} = \left(\frac{\gamma_2}{\lambda E(h_{S_2,0[t]})} - \frac{1}{E(h_{S_2,D_2[t]})}\right)^+. \tag{37}$$

where $\lambda_{[t]} = \alpha_{0[t]} + \lambda'_{[t]}$.

By substituting (36) and (37) into the subjective function in (34), we can show that λ needs to satisfy $\lambda_{[t]} \ge \frac{\gamma_1+\gamma_2}{Q_{0[t]}+E(h_{S_1,0[t]})/E(h_{S_1,D_1[t]})+E(h_{S_2,0[t]})/E(h_{S_2,D_2[t]})}$ if the transmit powers of both S_1 and S_2 are positive.

The optimal power allocation methods for the case that transmitters have full CSIs can be obtained by replacing the average values with instantaneous value of CSIs.

Appendix B
Optimal Power Allocation Method for Both S_1 and S_2 Using Multi-hop Relaying

Following the same line as Appendix A, let us consider the case that both S_1 and S_2 use multi-hop relaying. Using Assumption 3, we can write the optimization problem as follows:

$$\begin{aligned} \max_{w_{1,1[t]},w_{2,2[t]}} \quad & \pi^{CT}_{1[t]} + \pi^{CT}_{2[t]} \\ \text{s.t.} \quad & \frac{1}{2}\left[E\left(h_{S_1,0[t]}w_{1,1[t]}\right) + E\left(h_{S_2,0[t]}w_{2,2[t]}\right)\right. \\ & \left. +E\left(h_{S_1,0[t]}f_{2[t]}w_{2,2[t]}\right) + E\left(h_{S_2,0[t]}f_{1[t]}w_{1,1[t]}\right)\right] \le Q_{0[t]} \end{aligned} \tag{38}$$

Similar to Appendix A, we only consider the low SNR case and re-write the objective function of (38) as follows:

$$\begin{aligned} \pi^{CT}_{1[t]} + \pi^{CT}_{2[t]} = {} & \gamma_1 \log\left(1 + E(h_{S_1,S_2[t]})w_{1,1[t]}\right) \\ & +\gamma_2 \log\left(1 + E(h_{S_2,S_1[t]})w_{2,2[t]}\right) - \tilde{\alpha}_{0[t]}E(h_{S_1,0[t]})w_{1,1[t]} \\ & -\tilde{\alpha}_{0[t]}E(h_{S_2,0[t]})w_{2,1[t]} - \tilde{\alpha}_{0[t]}E(h_{S_2,0[t]})w_{2,2[t]} \\ & -\tilde{\alpha}_{0[t]}E(h_{S_1,0[t]})w_{1,2[t]}. \end{aligned} \tag{39}$$

Solving the Lagrange multiplier, we have the following KKT conditions:

$$w_{1,1[t]} = \left(\frac{\gamma_1}{\lambda \left[E(h_{S_1,0[t]}) + E\left(h_{S_2,0[t]} f_{1[t]}\right) \right]} - \frac{1}{E(h_{S_1,S_2[t]})} \right)^+,$$

$$w_{2,2[t]} = \left(\frac{\gamma_2}{\lambda \left[E(h_{S_2,0[t]}) + E\left(h_{S_1,0[t]} f_{2[t]}\right) \right]} - \frac{1}{E(h_{S_2,S_1[t]})} \right)^+,$$

Substituting the above equations into the subjective function in (38), we can have the condition of λ as $\lambda \geq \frac{\gamma_1+\gamma_2}{2Q+\frac{E(h_{S_1,0[t]})+E(h_{S_2,0[t]}f_{1[t]})}{E(h_{S_1,S_2[t]})}+\frac{E(h_{S_2,0[t]})+E(h_{S_1,0[t]}f_{2[t]})}{E(h_{S_2,S_1[t]})}}$ if both $w_{1,1[t]}$ and $w_{2,2[t]}$ are positive.

The optimal power allocation methods for the case that transmitters have full CSIs can be similarly obtained by replacing the average values with instantaneous value of CSIs.

References

[1] FCC, "Spectrum policy task force report," *Rep. ET Docket*, no. 02-135, Nov. 2002.

[2] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 201–220, 2005.

[3] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471–487, 2007.

[4] G. Amir and S. S. Elvino, "Fundamental limits of spectrum-sharing in fading environments," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 649–658, 2007.

[5] Y. T. Hou, S. Yi, and H. D. Sherali, "Spectrum sharing for multi-hop networking with cognitive radios," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 146–155, 2008.

[6] A. Pandharipande, H. Yang, and W. Ying, "Sensing and communication protocols in cognitive sensor relay networks," in *2008 IEEE Sensors*, pp. 617–620, 2008.

[7] M. van der Schaar and F. Fangwen, "Spectrum access games and strategic learning in cognitive radio networks for delay-critical applications," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 720–740, 2009.

[8] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 192–202, 2008.

[9] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 203–213, 2008.

[10] D. Lingjie, H. Jianwei, and S. Biying, "Competition with dynamic spectrum leasing," in *2010 IEEE Symposium on New Frontiers in Dynamic Spectrum*, pp. 1–11, 2010.

[11] G. Vazquez-Vilar, C. Mosquera, and S. Jayaweera, "Primary user enters the game: Performance of dynamic spectrum leasing in cognitive radio networks," *Accepted in IEEE Trans. Wireless Commun.*, 2010.

[12] Y. Shoham and K. Leyton-Brown, *Multiagent systems: algorithmic, game-theoretic, and logical foundations*. Cambridge Univ Press, 2008.

[13] M. Osborne, *An introduction to game theory*. Oxford University Press, New York, NY, 2004.

[14] D. W. R. Mahajan, M. Rodrig and J. Zahorjan, "Sustaining cooperation in multi-hop wireless networks," in *Proc. Second USENIX Symp. on Networked System Design and Implementation (NSDI 05), Boston, MA*, pp. 1–5, 2005.

[15] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[16] H. Kushner and G. Yin, *Stochastic approximation and recursive algorithms and applications*. Springer Verlag, 2003.

[17] H. Li, "Multiagent q-learning for aloha-like spectrum access in cognitive radio systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, 2010.

Yong Xiao received the M.Sc. degree from Hong Kong University of Science and Technology in 2006 and the B.S. degree from China University of Geosciences, Wuhan, China in 2002 both in electrical engineering. He is currently working towards the Ph.D. degree in Nanyang Technological University, Singapore. His current research interests include the applications of game theory and network coding to wireless networks.

Guoan Bi (SM'89) received the B.Sc degree in Radio communications, Dalian University of Technology, P. R. China, 1982, M.Sc degree in Telecommunication Systems and Ph.D degree in Electronics Systems, Essex University, UK, 1985 and 1988, respectively. Between 1988 and 1990,he was a research fellow at the University of Surrey, U.K.. Since 1991, he has been with the school of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.

Dusit Niyato (M'09) is currently an Assistant Professor in the Division of Computer Communications, School of Computer Engineering, Nanyang Technological University, Singapore. He received the B.E. degree from King Mongkuts Institute of Technology Ladkrabang, Bangkok, Thailand, in 1999 and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Winnipeg, MB, Canada, in 2008. His current research interests include design, analysis, and optimization of wireless communication, smart grid systems, green radio communications, and mobile cloud computing.