Hierarchic Power Allocation for Spectrum Sharing in OFDM-Based Cognitive Radio Networks

Tian Zhang, Wei Chen, *Senior Member, IEEE*, Zhu Han, *Fellow, IEEE*, and Zhigang Cao, *Senior Member, IEEE*

***Abstract*—In this paper, a Stackelberg game is built to model the joint power allocation of the primary user (PU) network and the secondary user (SU) network hierarchically in orthogonal frequency division multiplexing (OFDM)-based cognitive radio (CR) networks. We formulate the PU and SUs as the leader and the followers, respectively. We consider two constraints: the total power constraint and the interference-to-signal ratio (ISR) constraint, in which the ratio between the accumulated interference and the received signal power at each PU should not exceed a certain threshold. First, we focus on the single-PU–multi-SU scenario. Based on the analysis of the Stackelberg equilibrium (SE) for the proposed Stackelberg game, an analytical hierarchic power-allocation method is proposed when the PU can acquire the additional information to anticipate SUs' reactions. The analytical algorithm has two steps. First, the PU optimizes its power allocation by considering the SUs' reactions to its action. In the power optimization of the PU, there is a subgame for power allocation of SUs given the fixed transmit power of the PU. The existence and uniqueness for the Nash equilibrium (NE) of the subgame are investigated. We also propose an iterative algorithm to obtain the NE and derive the closed-form solutions of the NE for the perfectly symmetric channel. Second, the SUs allocate the power according to the NE of the subgame given the PU's optimal power allocation. Furthermore, we design two distributed iterative algorithms for the general channel even when private information of the SUs is unavailable at the PU. The first iterative algorithm has a guaranteed convergence performance and the second iterative algorithm employs asynchronous power update to improve time efficiency. Finally, we extend to the multi-PU–multi-SU scenario, and a distributed iterative algorithm is presented.**

***Index Terms*—Cognitive radio (CR), distributed iterative algorithm, game theory, hierarchic power allocation, joint power allocation.**

Manuscript received March 11, 2013; revised September 6, 2013 and January 2, 2014; accepted January 28, 2014. Date of publication February 4, 2014; date of current version October 14, 2014. This work was supported in part by the National Basic Research Program of China (973 Program) under Grant 2013CB336600 and Grant 2012CB316000; by the National Natural Science Foundation of China through the Young Investigator Program under Grant 61322111; by the Chinese Ministry of Education through the New Century Talent Program under Grant NCET-12-0302; by the Beijing Nova Program under Grant Z121101002512051; by the National Science and Technology Key Project under Grant 2013ZX03003006-005 and Grant 2013ZX03003004-002; by the U.S. National Science Foundation under Grant CNS-1265268, Grant CNS-1117560, Grant ECCS-1028782, and Grant CNS-0953377; and by the Electric Power Analytics Consortium. The review of this paper was coordinated by Dr. C. Ibars.

T. Zhang is with the School of Information Science and Engineering, Shandong University, Jinan 250100, China (e-mail: tianzhang.ee@gmail.com).

W. Chen and Z. Cao are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: wchen@tsinghua.edu.cn; czg-dee@tsinghua.edu.cn).

Z. Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005 USA (e-mail: zhan2@uh.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2014.2304618

I. Introduction

COGNITIVE radio (CR) technology has gained much attention because of its ability to improve spectrum utilization efficiency [1]. In CR networks, the CRs transmit in an opportunistic way or coexist with the primary systems simultaneously under the constraints that the primary systems will not be harmed. Due to the scarcity of power and hostile characteristics of wireless channels, efficient power-allocation schemes are necessary for the design of high-performance CR networks. Meanwhile, as the game theory is suitable for analyzing conflict and cooperation among rational decision-makers, it has emerged as a very powerful tool for power allocation in CR networks [2].

In game-theory-based power-allocation frameworks, the nodes are modeled as self-interested or group-rational players, and compete or cooperate with each other to maximize their utilities by viewing the power as the strategies [3]–[10]. The Stackelberg game, which is also referred to as the leader–follower game, is a game in which the leader moves first and then the followers move sequentially. The problem is then transformed to find an optimal strategy for the leader, assuming that the followers react in such a rational way that they optimize their objective functions given the leader's actions [11]. In [12]–[14],[1] the Stackelberg game was applied for the multiuser power control problem in interference channels. Utilization of the Stackelberg game in wireless communications can be also found in [15]–[18].

As the Stackelberg game is defined for the cases in which a hierarchy of actions exists between players, it is a natural fit for the CR scenario. The Stackelberg game was employed in CR networks in [19] and [20]. A Stackelberg game model was proposed for frequency bands in which a licensed user has priority over opportunistic CRs. In [21], the Stackelberg game was applied for the utility-based cooperative CR networks. In [22], the resource allocation in CR networks was studied by using the Stackelberg game to characterize the asymmetry of PUs and SUs. Allocation of underutilized spectrum resources from PUs to multiple SUs was modeled as the seller–buyer game. Similar work can also be found in [23], although the authors did not claim the use of Stackelberg game explicitly. A decentralized Stackelberg game formulation for power allocation was developed in [24]. Distributed optimization for CR networks using the Stackelberg game was considered in [25]. Distributed power control method for SUs and optimal pricing for the PU were obtained, and the algorithm for finding

[1] In [14], the Stackelberg equilibrium is a special case of the CE.

the optimal price was proposed. In [26], the focus is on how an SU chooses its power level to obtain maximal cognitive network capacity and guarantee the performance of the PU. Power allocation in the downlink of the secondary system was considered using the Stackelberg game in [27]. Constraints such as protecting PUs and the maximum power limitations of base stations (BSs) were considered. Distributed power control for spectrum-sharing femtocell networks was investigated by using the Stackelberg game in [28]. The Stackelberg equilibrium (SE) was studied, and an effective distributed interference price bargaining algorithm with guaranteed convergence was presented to achieve the equilibrium.

Recently, orthogonal frequency-division multiplexing (OFDM) has been recognized as an attractive modulation candidate for CR systems. In practice, the efficient algorithm of allocating power to subcarriers in both the OFDM-based PU network and SU network is important. However, most of the aforementioned works focus on the power control of the SU network only; the hierarchic joint power allocation for the OFDM-based PU network and SU network by using the Stackelberg game has not been extensively studied yet. When power control for the PU network and that for the SU network are jointly considered, we should consider not only the interference among SUs but also the interference among PUs and the mutual interference between the PU and SU networks. Meanwhile, to meet quality-of-service requirement of the PU precisely, the interference-to-signal ratio (ISR),[2] which is defined as the ratio between the accumulated interference and the received signal power, should be less than a certain constant at the PU. Then, the power allocation of the PU network and that of the SU network are tightly coupled. In addition, the transmission from the primary transmitter to its receiver needs to be analyzed. Thus, the transmission merit, such as rate, should be taken into consideration in the utility function of the PU. Due to the above reasons, the hierarchic power allocation is challenging, particularly when the PU network cannot acquire private information of the SU network. Even when the private information is available, it is difficult to design the time-efficient algorithm because of complexity of the game.

To overcome these challenges, the main contributions of this paper are summarized as follows.

- A Stackelberg game is formulated to describe the priority for the PU network in the joint power allocation with the SU network. We analyze the mutual effect between power allocation for the PU network and that of the SU network in two aspects: ISR constraint and mutual interference between the PUs and SUs. The former impacts the feasible power-allocation set and the latter influences the utility.
- When there is only one PU, the Stackelberg game can be written as an optimization problem that contains a noncooperative subgame. The subgame can be viewed as the power game of the SU network given the PU's power. We analyze existence for the NE of the subgame and give a sufficient condition of uniqueness. Moreover, an iterative algorithm, which converges to the NE, is presented for the general channel condition, and the closed-form solutions for the NE are derived in a perfectly symmetric channel.
- Based on the Stackelberg game analysis, the hierarchic joint power-allocation algorithms for the PU and SU networks are proposed. Considering availability of the private information for the SUs at the PU, two scenarios are investigated. When the private information is available and the perfectly symmetric channel conditions can be satisfied, the PU can allocate power by solving a specific optimization problem, and the SU can analytically allocate power in the perfectly symmetric channel. Otherwise, the iterative distributed power-allocation algorithms are presented. We also investigate the convergence and effectiveness of the proposed iterative algorithms.
- The extension to the multi-PU–multi-SU scenario is discussed, and we present an iterative distributed algorithm for the hierarchic power allocation.

[2]We use ISR instead of signal-to-interference ratio (SIR) to emphasize the interference controlling in the paper. ISR is the inverse of SIR, and it has also been widely used in literature.

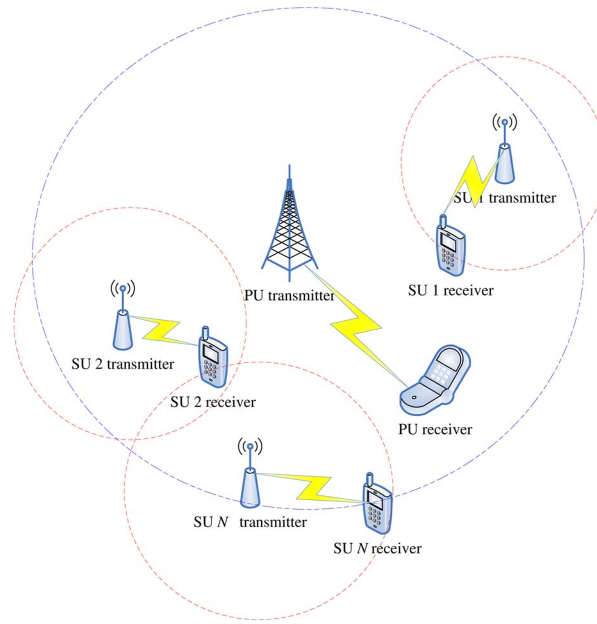


Fig. 1. PU system coexists with the SU system.

The remainder of this paper is structured as follows. In Section II, we introduce the system model under consideration and formulate the Stackelberg game. In Section III, the game analysis is performed. In Section IV, the hierarchic power-allocation methods for the PU and SUs are proposed. Then, the numerical results are presented in Section V. Finally, we conclude this paper in Section VI.

II. System Model and Problem Formulation

A. System Model

We consider a spectrum-sharing scenario, as shown in Fig. 1, in which a PU system coexists with an SU system. There is one PU and several SUs, each formed by a single transmitter–receiver pair using OFDM. The PU is denoted as user 1 and the SUs are denoted as user 2, $\ldots$, user L, respectively, i.e., the PU set $\mathbb{P} = \{1\}$, and the SU set $\mathbb{S} = \{2, \ldots, L\}$.

It is assumed that the total number of OFDM subchannels is N, and each subchannel experiences flat fading. The sampled signal on the fth subchannel at the receiver of user j is

$$y_j^f = \sqrt{P_j^f} h_{j,j}^f x_j^f + \sum_{i \neq j \in \mathbb{P} \cup \mathbb{S}} \sqrt{P_i^f} h_{i,j}^f x_i^f + w_i^f \tag{1}$$

where P_j^f and $h_{i,j}^f$ denote the transmitted power of user j's transmitter and the channel coefficient between the transmitter of user i and the receiver of user j on the fth subchannel, respectively. x_j^f is the transmitted symbol of user j at subchannel f and is assumed to have unit energy. w_i^f is the additive white Gaussian noise with $w_i^f \sim \mathcal{CN}(0, N_i^f)$. Each user's transmitter has a limited power budget, i.e., $\sum_{f=1}^N P_j^f \leq P_j^{\max}, \forall j \in \mathbb{P} \cup \mathbb{S}$. Treating the interference as noise and assuming Gaussian signaling, the maximum rate that user j can obtain on the fth subchannel can be expressed as

$$R_j^f = \log\left(1 + \frac{P_j^f \left|h_{j,j}^f\right|^2}{\sum_{i \neq j \in \mathbb{P} \cup \mathbb{S}} P_i^f \left|h_{i,j}^f\right|^2 + N_j^f}\right) \text{ (nats/s/Hz).} \tag{2}$$

B. Stackelberg Game Formulation

We formulate the PU and the SUs as the leader and the followers, respectively. The PU first selects its transmission power by maximization of its utility, in which it tries to anticipate the SUs' reactions to its action. Then, based on the PU's power, the SUs compete with each other to maximize their own rates by adjusting transmit power. The ISR constraint, i.e., $(\sum_{i \in \Omega} P_i^f |h_{i,1}^f|^2)/(P_1^f |h_{1,1}^f|^2) \leq \rho$ with ρ being the ISR threshold, needs to be satisfied to guarantee primary service.[3]

The formulation of the proposed Stackelberg game can be decomposed into two levels: lower level of the SUs and upper level for the PU.

1) Lower Level: Given the PU's transmit power, the SUs' noncooperative subgame can be mathematically formulated as

$$\mathcal{G} = \{\Omega, \{\mathcal{S}_i\}_{i \in \Omega}, \{u_i\}_{i \in \Omega}\} \tag{3}$$

where $\Omega = \mathbb{S}$ is the set of active players. The set of admissible power-allocation strategies for user i is given by

$$\mathcal{S}_i = \left\{ \mathbf{P}_i = \left(P_i^1, P_i^2, \ldots, P_i^N\right) : \sum_{f=1}^N P_i^f \leq P_i^{\max}; \right. \\ \left. \forall f \in \{1, 2, \ldots, N\},\ P_i^f \geq 0 \right\}. \tag{4}$$

The utility function of user i is defined as $u_i(\mathbf{P}_i, \mathbf{P}_{-i}) = \sum_{f=1}^N R_i^f$,[4] where $\mathbf{P}_{-i} := \{\mathbf{P}_k\}_{k \in \Omega/\{i\}}$.

2) Upper Level: For the PU, if it can anticipate the SUs' reactions to its action, we have the following problem:

$$\max_{\mathbf{P}_1} u_1 = \sum_{f=1}^N \log\left(1 + \frac{P_1^f \left|h_{1,1}^f\right|^2}{\sum_{i \in \Omega} P_i^{f^*} \left|h_{i,1}^f\right|^2 + N_1^f}\right)$$

$$\text{s.t.} \quad \sum_{f=1}^N P_1^f \leq P_1^{\max},\ P_1^f \geq 0,\ \frac{\sum_{i \in \Omega} P_i^{f^*} \left|h_{i,1}^f\right|^2}{P_1^f \left|h_{1,1}^f\right|^2} \leq \rho \tag{5}$$

where $\mathbf{P}_1 = (P_1^1, P_1^2, \ldots, P_1^N)$, $\mathbf{P}_i^* = (P_i^{1*}, P_i^{2*}, \ldots, P_i^{N*})$ with $i \in \Omega$, and $(\mathbf{P}_i^*, \mathbf{P}_{-i}^*)$ is the NE of $\mathcal{G}$ when $\mathbf{P}_1$ is given.[5]

When the leader (PU) chooses a power strategy $\mathbf{P}_1$, the followers (SUs) get their power strategies $\{\mathbf{P}_i^*\}_{i \in \mathbb{S}}$ from $\mathcal{G}$ accordingly. Then, there is an u_1. Once the PU could obtain the private information of the SU network (e.g., c, h_{1i}^f, h_{ii}^f, h_{i1}^f, and $P_i^{\max}$ $(i \in \Omega, f = 1, \ldots, N)$), the PU could know the SUs' reactions by solving the subgame $\mathcal{G}$. Thus, the PU could exactly know u_1 corresponding to a $\mathbf{P}_1$. Then, it can choose its optimal power allocation.

III. Game Analysis

In Section II-B, the Stackelberg game is formulated in two levels: the SUs' subgame for given PU's power and the PU's optimization problem when it could anticipate the SUs' reactions. Then, naturally, we can transform the analysis of the whole problem into the successive analysis of the lower problem and the upper problem. We first analyze the subgame (which is a noncooperative game) and focus on the NE. Second, we analyze the upper problem. Since only one PU is considered, once we obtain the solution of the SUs' problem (through the analysis of the lower level problem), the upper level problem is a common optimization problem. Then, we focus on the analysis of subgame $\mathcal{G}$.

The most important thing is that this successive analysis is enough for proposing the power-allocation algorithm in Section IV. The proposed analytical algorithm can obtain the SE and we will prove this. The Stackelberg game can be theoretically analyzed as a whole, e.g., directly analyzing the SE. The analysis of SE will be considered in future work.

Here, the existence, uniqueness, and solution for the NE of subgame $\mathcal{G}$ are analyzed. An iterative algorithm to obtain the NE of the subgame is given. We also investigate the convergence of the iterative algorithm. Furthermore, the closed-form solutions for the NE are derived for the perfectly symmetric channel.

First, for subgame $\mathcal{G}$, its NE is defined as follows:

Definition 1: $(\mathbf{P}_i^*, \mathbf{P}_{-i}^*)$ is the NE if $u_i(\mathbf{P}_i^*, \mathbf{P}_{-i}^*) \geq u_i(\mathbf{P}_i, \mathbf{P}_{-i}^*)$ for all $\mathbf{P}_i \in \mathcal{S}_i$ and $i \in \Omega$.

With respect to the existence of the NE for $\mathcal{G}$, we have the following proposition.

[3]We only need to guarantee that the power allocation in the stable state, i.e., the SE (its definition will be given in the following) or the convergent outcomes of the iterative algorithm, should satisfy the ISR constraint.

[4]The utility function can be defined in other forms, i.e., the proposed framework is general enough to allow different definitions of the utility function. Concerning the obtained conclusions, some are independent of the utility function definition, and others can be adapted easily for new definitions of the utility function.

[5]The definition of NE will be given in Section III. Equation (5) is the formulated Stackelberg game, where it contains subgame $\mathcal{G}$. We should observe that the ISR constraint is not considered in $\mathcal{G}$. However, as the ISR constraint is considered in (5), the solutions of the Stackelberg game comply with the ISR constraint.

Proposition 1: The subgame $\mathcal{G}$ has at least one pure NE.

Proof: See Appendix A. ■

The uniqueness of the NE can be given by the following.

Proposition 2: Define

$$\mathbf{M}_{i,j} = \begin{cases} -\max\limits_{f\in[1,N]} \left\{ \frac{|h_{ji}^f|^2}{|h_{jj}^f|^2} \frac{N_j^f + P_1^f |h_{1j}^f|^2 + \sum_{l\in\Omega} |h_{lj}^f|^2 P_l^{\max}}{N_i^f + P_1^f |h_{1i}^f|^2} \right\}, & i\neq j \\ 1, & i=j. \end{cases} \tag{6}$$

If $\mathbf{M}$ is a positive definite matrix, $\mathcal{G}$ has a unique NE.

Proof: See Appendix B. ■

The condition in Proposition 2 can be viewed as the weak interference condition. Specifically, $\mathbf{M}$ is positive definite if both of the following conditions hold[6]

$$\frac{1}{w_i}\sum_{j\neq i} w_j \times \max_{f\in[1,N]} \left\{ \frac{\left|h_{ji}^f\right|^2}{\left|h_{jj}^f\right|^2} \frac{N_j^f + P_1^f \left|h_{1j}^f\right|^2 + \sum_{l\in\Omega} \left|h_{lj}^f\right|^2 P_l^{\max}}{N_i^f + P_1^f \left|h_{1i}^f\right|^2} \right\} \leq 1 \tag{7}$$

$$\frac{1}{w_j}\sum_{i\neq j} w_i \times \max_{f\in[1,N]} \left\{ \frac{\left|h_{ji}^f\right|^2}{\left|h_{jj}^f\right|^2} \frac{N_j^f + P_1^f \left|h_{1j}^f\right|^2 + \sum_{l\in\Omega} \left|h_{lj}^f\right|^2 P_l^{\max}}{N_i^f + P_1^f \left|h_{1i}^f\right|^2} \right\} \leq 1 \tag{8}$$

where $(w_i)_{i\in\mathbb{S}}$ is a certain positive vector. Equation (7) can be viewed as an upper bound on the amount of interference that each receiver can tolerate, and (8) introduces a constraint on the level of interference that each transmitter can generate.

In the following, we give an iterative algorithm to obtain the NE. The best response for user i $(i\in\Omega)$ can be expressed as

$$P_i^f = BR_i\left(P_1^f, P_{-i}^f\right) = \left(\frac{1}{\mu_i} - \frac{P_1^f \left|h_{1,i}^f\right|^2 + \sum_{j\in\Omega/i} P_j^f \left|h_{j,i}^f\right|^2 + N_i^f}{|h_{ii}|^2}\right)^+ \tag{9}$$

where $P_{-i}^f(k) = \{P_j^f(k)\}_{j\in\Omega/i}$, $(\cdot)^+ = \max(\cdot, 0)$, and μ_i is a constant satisfying $\sum_{f=1}^N P_i^f \leq P_i^{\max}$. Based on (9), an iterative distributed algorithm (Algorithm 1), which can converge to the NE, can be given. In the algorithm, SU i only has to obtain its own channel state h_{ii}, and measure the aggregated interference it receives, i.e., $P_1^f|h_{1,i}^f|^2 + \sum_{j\in\Omega/i} P_j^f(k)|h_{j,i}^f|^2$. Hence, Algorithm 1 can be implemented distributively.

[6]See [8, Coroll. 4].

Algorithm 1: Iterative Distributed Algorithm for Obtaining NE

Step 1: $k=0$,
initialize feasible $\{\mathbf{P}_i(0) = (P_i^1(0), \ldots, P_i^N(0))\}_{i\in\Omega}$.
Step 2: $P_i^f(k+1) = BR_i(P_1^f, P_{-i}^f(k))$
for every $i\in\Omega$ and $f=1,\ldots,N$.
Step 3: $k=k+1$, go to Step 2 until convergence.

Following the existing literature (such as [10] and [29]), a sufficient condition for the convergence of Algorithm 1 is given by the following proposition.

Proposition 3: Let $\mathbf{C}$ be a partitioned matrix with zero diagonal blocks; the $(i-1, j-1)$th block is an $N\times N$ matrix whose (f,f) entry is $c_{i,j}^f = |h_{i,j}^f|^2/|h_{j,j}^f|^2$, for $i, j \in \{2,3,\ldots,L\}$ and $f\in\{1,2,\ldots,N\}$. If $\|\mathbf{C}\| < 1$, where $\|\cdot\|$ is any induced matrix norm with its corresponding vector norm being monotone, Algorithm 1 converges.

Proof: See Appendix C. ■

Under a special circumstance, i.e., a perfectly symmetric channel, we derive the closed-form solutions of NE.

Proposition 4: When $|h_{i,j}^f|/|h_{j,j}^f| = |h_{j,i}^{f'}|/|h_{i,i}^{f'}| = \sqrt{c} < 1$, $N_i^f/|h_{ii}^f|^2 = N_j^f/|h_{jj}^f|^2$ and $|h_{1i}^f|/|h_{ii}^f| = |h_{1j}^f|/|h_{jj}^f|$ for f, $f' = 1,\ldots,N$ and $i\neq j\in\Omega$, the perfectly symmetric channel conditions hold. Then, for $L=3$, the NE of $\mathcal{G}$ has the following closed-form solutions[7]:

$$P_2^{f^*} = \begin{cases} t_1^* - \frac{ct_2^*+\sigma_f}{1+c}, & f\in[1,k_2] \\ t_1^* - \sigma_f, & f\in[k_2+1,k_1] \\ 0, & f\in[k_1+1,N] \end{cases} \tag{10}$$

$$P_3^{f^*} = \begin{cases} \frac{t_2^*-\sigma_f}{1+c}, & f\in[1,k_2] \\ 0, & f\in[k_2+1,N] \end{cases} \tag{11}$$

where $\sigma_f = (N_i^f + P_1^f|h_{1i}^f|^2)|h_{ii}^f|^{-2}$, $t_2^* = k_2^{-1}[(1+c)P_3^{\max} + \sum_{f=1}^{k_2}\sigma_f]$, and k_2 can be found from $\varphi_{k_2}^2 < P_3^{\max} \leq \varphi_{k_2+1}^2$ with

$$\varphi_k^2 = \begin{cases} \frac{1}{1+c}\sum_{f=1}^{k}(\sigma_k - \sigma_f), & 1\leq k\leq N \\ \infty, & k=N+1. \end{cases} \tag{12}$$

$t_1^* = (P_2^{\max} + \sum_{f=k_2+1}^{k_1}\sigma_f + (1/1+c)\sum_{f=1}^{k_2}(ct_2^*+\sigma_f))/(k_1)$, where $k_1 = k_2$ when $P_2^{\max} \leq \varphi_{k_2+1}^1$; otherwise, k_1 is the solution of $\varphi_{k_1}^1 < P_2^{\max} \leq \varphi_{k_1+1}^1$, and φ_k^1 is defined as

$$\varphi_k^1 = \begin{cases} \sum_{f=k_2+1}^{k}(\sigma_k - \sigma_f) + \frac{1}{1+c} \\ \sum_{f=1}^{k_2}((1+c)\sigma_k - \sigma_f - ct_2^*), & k\in[k_2+1,N] \\ \infty, & k=N+1. \end{cases} \tag{13}$$

Proof: See Appendix D. ■

The above proposition is for the two-SU scenario. However, following the proof of this proposition, the closed-form solutions for the multi-SU scenario can be obtained similarly. Using Proposition 4, the power for SUs in the perfectly symmetric

[7]Without loss of generality, we assume $P_2^{\max} > P_3^{\max}$. σ_f is only distinguished by the number of subchannels in the perfectly symmetric channel; the subchannels can be renumbered according to the strength of noise plus received interference from the PU. Thus, it is also assumed that $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_N$. Subcarriers should be renumbered at the beginning and we need to recover the number of subcarriers in the end.

channel can be allocated analytically with simple computation. Moreover, if we suppose that $\{P_i^{\max}\}_{i\in\Omega}$ is known at user $i(i \in \Omega)$, user $i(i \in \Omega)$ only needs to obtain c (i.e., $h_{j,i}^f$ and $h_{i,i}^f$) and measure the received interference from PU, i.e., $P_1^f|h_{1i}|^2$. Thus, Proposition 4 can be distributively applied.

Equations (10) and (11) and Algorithm 1 can be used to obtain the NE of $\mathcal{G}$ in the two-SU scenario. When the perfectly symmetric channel conditions hold, the analytical solutions are given in (10) and (11); otherwise, Algorithm 1 can find the solution for the general case.

IV. Power Allocation Algorithm

Here, we consider the hierarchic joint power allocation for the PU and SUs. If the PU can acquire the additional information about the SUs to anticipate SUs' reactions to its action, we propose an analytical power-allocation algorithm in Section IV-A. Otherwise, the iterative power-allocation algorithms are developed in Section IV-B. Furthermore, we consider the extension to the multi-PU–multi-SU scenario in Section IV-C.

A. Analytical Power-Allocation Algorithm

The definition of the SE is given by the following.

Definition 2: $(\mathbf{P}_1^*, \hat{\mathbf{P}}_i^*, \hat{\mathbf{P}}_{-i}^*)$ is an SE for the proposed Stackelberg game when it satisfies the following:

1) $u_i(\mathbf{P}_1^*, \hat{\mathbf{P}}_i^*, \hat{\mathbf{P}}_{-i}^*) \geq u_i(\mathbf{P}_1^*, \mathbf{P}_i, \hat{\mathbf{P}}_{-i}^*), \forall\, i \in \Omega, \mathbf{P}_i \in \mathcal{S}_i$;
2) $u_1(\mathbf{P}_1^*, \hat{\mathbf{P}}_i^*, \hat{\mathbf{P}}_{-i}^*) \geq u_1(\mathbf{P}_1, \mathbf{P}_i^*, \mathbf{P}_{-i}^*)$ for any feasible $\mathbf{P}_1$.

In the following, we discuss on how to find the SE solution of the proposed Stackelberg game. First, we have the following lemma.

Lemma 1: Denote the optimal solution of (5) as $\mathbf{P}_1^*$; then, $(\mathbf{P}_1^*, \mathrm{Ne}(\mathbf{P}_1^*))$ is an SE.

Proof: See Appendix E. ■

According to Lemma 1, to get the SE, we should solve (5) first. Then, the focus becomes solving (5). In Proposition 4, we have obtained the closed-form NE of the SUs' subgame for perfectly symmetric channels given PU power. By substituting the closed-form NE (which is the function of PU's power level) into the PU's problem (5), the PU's problem becomes a conventional optimization problem. If the PU knows c, h_{1i}^f, h_{ii}^f, h_{i1}^f, $P_i^{\max}(i \in \Omega, f = 1, \ldots, N)$ and its own channel state h_{11}^f, it can solve the optimization problem (at least numerically).

Based on the earlier discussions, we get the analytical power-allocation method. First, the PU obtains the optimal power allocation $\mathbf{P}_1^*$ by substituting the expressions of subgame's NE in to (5) and solving (5) thereafter. Then, the SUs allocate the power according to the NE of $\mathcal{G}$, $(\hat{\mathbf{P}}_i^*, \hat{\mathbf{P}}_{-i}^*)$, given $\mathbf{P}_1 = \mathbf{P}_1^*$. In particular, for the two-SU scenario with perfectly symmetric channels, by substituting (10) and (11) into (5), the PU problem becomes a conventional optimization problem. By solving the problem, we obtain the optimal power-allocation strategy of the PU, i.e., $\mathbf{P}_1^*$. Replacing $\mathbf{P}_1$ by $\mathbf{P}_1^*$ in (10) and (11), we get the NE of $\mathcal{G}$ given the optimal power allocation of the PU, which is denoted $(\hat{\mathbf{P}}_2^*, \hat{\mathbf{P}}_3^*)$. Then, the SUs allocate the power according to $\hat{\mathbf{P}}_2^*$ and $\hat{\mathbf{P}}_3^*$, respectively. Observe that $(\mathbf{P}_1^*, \hat{\mathbf{P}}_2^*, \hat{\mathbf{P}}_3^*)$ is the SE of the Stackelberg game according to Lemma 1.

In the analytical power-allocation method, we should know the expressions of NE to transform the PU's problem (which includes a subgame) to be a conventional optimization problem with respect to the PU's power. Meanwhile, the private information of the SU network is needed to solve the transformed optimization problem. The analytical power-allocation method can be utilized in the perfectly symmetric channels[8] when the private information of the SU network is available.

B. Iterative Power Allocation Algorithm

If the private information of the SUs is unknown to the PU, the PU cannot set an optimal power level by solving the optimization problem even under the perfectly symmetric channel conditions. Alternatively, the iterative algorithms are needed to identify the power level.

To play the SE, the PU must have the ability to anticipate the SUs' reactions to its action. However, it is impossible to exactly anticipate the SUs' reactions to the PU's action when the PU cannot obtain the private information about the SUs. The PU should know the SUs' private information such as the strategy set (exactly, c, h_{1i}^f, h_{ii}^f, h_{i1}^f, and $P_i^{\max}(i \in \Omega, f = 1, \ldots, N)$ are needed) to anticipate the SUs' reactions to its action. Although the SE can be viewed as a special case of the conjectural equilibrium (CE) [14], the CE assumes that the foresighted user knows its stationary interference and the first derivatives with respect to the allocated power. (ISR constraint is not considered in [14].) Hence, no algorithms can derive the SE solution in the case that the PU cannot obtain the private information about the SUs, particularly when the ISR constraint is considered. The outcomes of the iterative algorithms are not the SE solution.

1) Proposed Iterative Algorithm: The PU sets an initial power level in Step 1. In each iteration, based on PU's power allocation in the former iteration, the SUs allocate their power levels $\{\mathbf{P}_i(n) = (P_i^1(n), \ldots, P_i^N(n))\}_{i\in\mathbb{S}}$ according to the NE of the SUs' subgame by using Proposition 1 or Algorithm 1. Given the updated power levels of the SUs, the PU updates its power by maximizing its utility under total power and interference constraints,[9] i.e., $P_1^f(n+1)$ is the solution of the following convex optimization problem:

$$\max_{\mathbf{P}_1} u_1 = \sum_{f=1}^{N} \log\left(1 + \frac{P_1^f\left|h_{1,1}^f\right|^2}{I^f(n) + N_1^f}\right)$$
$$\text{s.t.} \sum_{f=1}^{N} P_1^f \leq P_1^{\max},\ P_1^f \geq 0,\ \frac{I^f(n)}{P_1^f\left|h_{1,1}^f\right|^2} \leq \rho \tag{14}$$

where $I^f(n) = \sum_{i\in\mathbb{S}} P_i^f(n)|h_{i,1}^f|^2$ is the received interference at the PU. The ISR constraint $(I^f(n))/(P_1^f|h_{1,1}^f|^2) \leq \rho$

[8]For general channels, using Algorithm 1, we can get values of the NE for $\mathcal{G}$ given a PU power level but not analytical expressions of the NE regarding the arbitrary PU power level.

[9]See (5). To some extent, the ISR constraint is imposed on the PU network in the iterative algorithm. In [28], the interference constraint has been imposed on PU to decrease the complexity of the power-allocation algorithms. Here, we impose the ISR constraint on the PU network for the same reason.

in (14) is equivalent to a minimal power constraint $(I^f(n))/(\rho|h^f_{1,1}|^2) \leq P^f_1$. Consequently, it can be solved by a two-step algorithm. The minimal power to meet the ISR constraint is first allocated to each subchannel, i.e., we allocate $(I^f(n))/(\rho|h^f_{1,1}|^2)$ for subchannel f; then, we subtract the allocated power from $P^{\max}_1$ and allocate the remaining power to the subchannels by using the water-filling method [30]. The iteration continues until convergence. We observe that the PU only needs to know its own channel information h^f_{11} and received interference $I^f(n)$. The specific distributed power-allocation algorithm is described in Algorithm 2.

Algorithm 2: Joint Iterative Distributed Power-Allocation Algorithm for PU and SUs (single-PU–multi-SU)

Step 1: $n = 0$, initialize $\mathbf{P}_1(0) = (P^1_1(0), \ldots, P^N_1(0))$.

Step 2: Given $\mathbf{P}_1(n)$, the SUs allocate the NE power according to (10) and (11) when the perfectly symmetric conditions can be satisfied in the two-SU scenario. Otherwise, the SUs apply Algorithm 1 in the general scenario. Denote the allocated power for SUs as $\{\mathbf{P}_i(n) = (P^1_i(n), \ldots, P^N_i(n))\}_{i \in \mathbb{S}}$.

Step 3: Update PU's power by using $P^f_1(n+1) = (1-\eta)P^f_1(n) + \eta[(I^f(n))/(\rho|h^f_{1,1}|^2) + (\lambda - (I^f(n) + N^f_1)/(|h_{11}|^2))^+]$, where λ is a constant to meet $\sum_{f=1}^N [(I^f(n))/(\rho|h^f_{1,1}|^2) + (\lambda - (I^f(n) + N^f_1)/(|h_{11}|^2))^+] \leq P^{\max}_1$, i.e., $\sum_{f=1}^N (\lambda - (I^f(n) + N^f_1)/(|h^f_{11}|^2))^+ \leq P^{\max}_1 - \sum_{f=1}^N (I^f(n))/(\rho|h^f_{1,1}|^2)$ ($P^{\max}_1 \geq \sum_{f=1}^N (I^f(n))/(\rho|h^f_{1,1}|^2)$ is assumed in this paper), and $\eta \in (0,1)$ is a fixed step size.

Step 4: $n = n+1$, go to Step 2 until convergence.

When the private information of the SUs (followers) cannot be acquired by the PU (leader), the PU has no information at the beginning, and it cannot anticipate the interference from the SUs with respect to its own power allocation; the only thing it can do is randomly set an initial feasible power allocation. Then, according to the PU's power allocation, the SUs play their subgame to obtain the power allocations. Then, define the nth ($n = 1, 2 \ldots$) round as "the PU allocates its power $\mathbf{P}_1(n)$, and the SUs allocate the power $\{\mathbf{P}_i(n)|i \in \Omega\}$ subsequently." In the nth round, the PU can only know the interference of the SUs with respect to the PU's former power allocation (power allocation in the former round), i.e., $I^f(n-1)$ (history information of the interference can be obtained by measuring the total interference it received), and it cannot exactly anticipate the interference of the SUs with respect to the PU's allocation in the same round, i.e., $I^f(n)$ (future information of the interference); therefore, it can only allocate the power by utilizing the history information $I^f(n-1)$. Then, based on the PU's power allocation, the SUs play their subgame to obtain the power allocations in the same round. In addition, the ISR constraint should be considered in the power allocation.

Due to the condition that the PU cannot obtain the private information about the SUs, the PU cannot exactly anticipate the future information of the interference,[10] and it can only utilize the history information of the interference. In conclusion, the unavailability of the private information and the ISR constraint leads to Algorithm 2. There are many methods to utilize the history information; we choose a simple one in our algorithm.

2) Convergence: Regarding the convergence of Algorithm 2, we have the following lemma.

Lemma 2: When $P^{\max}_i$, h_{ij}, and N_i ($i, j \in \mathbb{S} \cup \mathbb{P}$) are fixed, there exists a constant $\xi > 0$, and when $\eta < \xi$, Algorithm 2 converges.

Proof: See Appendix F. ■

Remark: If the algorithm does not converge with a certain step size, we can choose smaller step size to make the algorithm converge. Lemma 2 guarantees the existence of such convergent step size.

3) Comparison With Previous Algorithms: In [14], the conjecture-based rate maximization (CRM) algorithms are developed even if the foresighted user has no *a priori* knowledge of its competitors' private information. The CRM algorithm can achieve better performance than those of the NE.[11] However, there are shortcomings of CRM algorithm: 1) it is not guaranteed to converge to a CE; 2) it cannot be utilized for the scenarios in which multiple foresighted users coexist; and 3) the number of frequency bins should be sufficiently large. In contrast, there are no constraints on the number of frequency bins in our proposed algorithm, and it has the guaranteed convergence performance. Moreover, our proposed algorithm can be extended to the multileader case (see Section IV-C). Finally, no ISR constraints are considered in [14].[12] As explained in this paper, the ISR constraint will greatly couple the power allocations of the PU (leader) with the power allocations of the SUs (followers). That is, the CRM algorithm cannot be applied under our system model, where the ISR constraint should be considered, and there is no constraint on the number of subcarriers. In other word, we deal with a more complicated problem in this paper.

4) Asynchronous Algorithm: In Algorithm 2, as the PU waits for the convergence of the power profiles of the SUs (Step 2), it then updates its power. It will be time-consuming particularly when the number of SUs is large. For the purpose of further improving time efficiency, we propose the asynchronous algorithm in Algorithm 3.

Algorithm 3: Asynchronous Joint Iterative Distributed Power Allocation Algorithm for PU and SUs (single-PU multi-SU)

Step 1: $n=0$, $k=1$, initialize $\mathbf{P}_1(0) = (P^1_1(0), \ldots, P^N_1(0))$ and $\{\mathbf{P}_i(0) = (P^1_i(0), \ldots, P^N_i(0))\}_{i \in \mathbb{S}}$, $\mathbf{P}_1(0)$ and $\{\mathbf{P}_i(0)\}_{i \in \mathbb{S}}$ satisfy their respective total power constraints and the ISR constraint.

[10]Based on the history information of the interference, the PU may predict the future information of the interference by using prediction methods, but it is not an exact prediction.

[11]Observe that the CRM algorithm cannot derive the SE.

[12]The system model considered in [14] is the interference channel.

Step 2: Given $\mathbf{P}_1(n)$, the SUs update power allocation $\{\mathbf{P}_i(n+1) = (P_i^1(n+1), \ldots, P_i^N(n+1))\}_{i\in\mathbb{S}}$ according to (10) and (11) in the two-SU scenario when the perfectly symmetric conditions can be satisfied. Otherwise
$P_i^f(n+1) = BR_i(P_1^f(n), P_{-i}^f(n))$
for every $i \in \mathbb{S}$ and $f = 1, \ldots, N$.
Step 3: Let $\{\tau_k\}_{k=1}^{\infty}$ be a subsequence of $\{n\}_{n=0}^{\infty}$ with $\tau_{k+1} - \tau_k < \infty$ for finite k.
The PU updates its power asynchronously by

$$P_1^f(n+1) = \begin{cases} (1-\delta)P_1^f(n) \\ \quad +\delta\left[\frac{I^f(n+1)}{\rho\left|h_{1,1}^f\right|^2} + \left(\lambda - \frac{I^f(n+1)+N_1^f}{|h_{11}|^2}\right)^+\right], & n = \tau_k \\ P_1^f(n), & \text{otherwise.} \end{cases}$$

$f = 1, 2, \ldots, N$, where $\delta \in (0,1)$ is the fixed step size.
If $n = \tau_k$, $k = k+1$.
Step 4: $n = n+1$, go to Step 2 until convergence or $n = N_{\max}$.

Remark: The PU asynchronously updates its power allocation in Algorithm 3. It does not need to wait for the convergence of SUs' power allocation. Consequently, it is more time efficient.

C. Extension to the Multi-PU–Multi-SU Scenario

When considering the multi-PU scenario, there are multiple leaders in the Stackelberg game; they compete with each other to maximize their individual utility. Each PU considers not only the power allocation of other PUs but also the rational reaction of the SU network to the power allocation of the PU network. In addition, we need to guarantee all PUs' ISR constraints. By minor adjustments, the proposed algorithms can be applied in the multi-PU–multi-SU scenario. In Algorithm 1, SU i still measure the aggregated received interference, but the interference is generated by all PUs and other SUs in this scenario. In Algorithms 2 and 3, the update of each PU's power can still utilize the former method. However, the received interference should take other PUs' power allocation into consideration. In Algorithm 2, the convergence of PUs' power allocation should be achieved before the next iteration in the multi-PU case. A renewed algorithm of Algorithm 2 for multi-PU is outlined as Algorithm 4.

Algorithm 4: Joint Iterative Distributed Power Allocation Algorithm for PUs and SUs (multi-PU multi-SU)

Step 1: $n=0$, initialize $\mathbf{P}_i(0) = (P_1^1(0), \ldots, P_1^N(0))$, $i \in \mathbb{P}$.
Step 2: Given $\{\mathbf{P}_i(n)\}_{i\in\mathbb{P}}$, the SUs allocate the NE power according to (10) and (11) when the perfectly symmetric conditions can be satisfied in the two-SU scenario. Otherwise, the SUs apply Algorithm 1 in the general scenario (Observe that $P_1^f|h_{1,i}^f|^2$ should be replaced by $\sum_{l\in\mathbb{P}} P_l^f(n)|h_{l,i}^f|^2$).
Denote the allocated power for SUs as
$\{\mathbb{P}_i(n) = (P_i^1(n), \ldots, P_i^N(n))\}_{i\in\mathbb{S}}$.
Step 3:
Substep 3.1: $k = 0$, $\mathbf{P}_i(k) = \mathbf{P}_i(n)$ for all $i \in \mathbb{P}$.
Substep 3.2: For every $i \in \mathbb{P}$, PU i updates its power by using

$$P_i^f(k+1) = (1-\eta_i)P_i^f(k) + \eta_i\left[\frac{I_i^f(k)}{\rho\left|h_{i,i}^f\right|^2} + \left(\lambda_i - \frac{I_i^f(k)+N_i^f}{\left|h_{ii}^f\right|^2}\right)^+\right]$$

where $I_i^f(k) = \sum_{l\neq i\in\mathbb{P}} P_l^f(k)|h_{l,i}^f|^2 + \sum_{j\in\mathbb{S}} P_j^f(n)|h_{j,i}^f|^2$ is the total received interference, λ_i is a constant to meet $\sum_{f=1}^N[(I_i^f(k))/(\rho|h_{i,i}^f|^2)+(\lambda_i - (I_i^f(k)+N_i^f)/(|h_{ii}|^2)^+] \le P_i^{\max}$, i.e., $\sum_{f=1}^N(\lambda_i - (I_i^f(k)+N_i^f)/(|h_{ii}^f|^2))^+ \le P_i^{\max} - \sum_{f=1}^N(I_i^f(k))/(\rho|h_{i,i}^f|^2)$, and $\eta_i \in (0,1)$ is a fixed step size.
Substep 3.3: $k = k+1$, go to Substep 3.2 until convergence.
Substep 3.4: $\mathbf{P}_i(n+1) = \mathbf{P}_i(k)$ for $i \in \mathbb{P}$.
Step 4: $n = n+1$, go to Step 2 until convergence or $n = N_{\max}$.

In terms of the rate performance, the analytical method, which can get the SE, is the first choice. When the analytical method cannot be applied, e.g., private information of SU network is unavailable, the iterative method is utilized. (The iterative method is proposed for the scenarios that the analytical method cannot be applied, and there is rate performance degeneration; see Section IV-B). We have performed comparisons in Fig. 8. From the simulation results, we can find that, for the iterative method, the rate of the PU has approximately 2% degeneration and the rates of SUs are almost the same as those of the analytical method.

Regarding the complexity and overhead, in the analytical method, after substituting the expressions of NE into PU's problem, the transformed conventional optimization problem maybe nonconvex. When the interior method [36] is utilized for solving the problem, the iteration complexity is $O(3LN + (L-1)(L-2)N + L)$. In addition, the PU should know the private information of the SU network, and each SU should know information of other Sus. The exchanging information has $[2 + (L-1)(L-2)N]b_1 + [(L-1) + (L-2/2)(L-1)]b_2$ bits, where b_1 and b_2 are the numbers of bits to denote the channel state and the maximal power budget, respectively. These information exchanges produce extra overhead. With respect to the iterative method, for Algorithm 1 step 2, the update of μ_i can be done by a bisection method. The computational complexity of Algorithm 1 is $O((L-1)N)$. For Algorithm 2, the computational complexity for each iteration is $O(m(L-1)N + N)$, where m is the iteration number for Algorithm 1. Additionally, there is no information exchange between the PU and SU networks.

The proposed hierarchic joint power allocation can be employed to improve the rate performance of a practical network, in which a licensed BS and a licensed mobile phone correspond to the PU transmitter and the PU receiver (in the considered system model), respectively. Secondary access points and

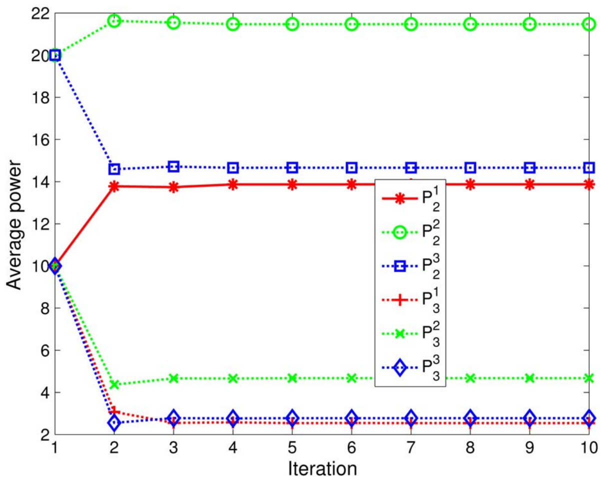


Fig. 2. Power allocation of the SUs by using Algorithm 1 in the perfectly symmetric channel case; there are three subcarriers, and PU's power is $P_1 =$ [70 10 30].

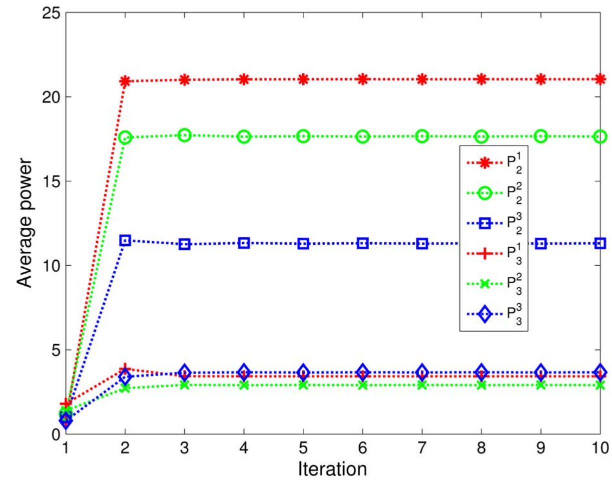


Fig. 3. Convergence performance of Algorithm 1 in the general case; there are three subcarriers, and PU's power is $P_1 =$ [10 30 70].

cognitive mobile phones correspond to the SU transmitters and the SU receivers, respectively. When the licensed BS–mobile-phone network can obtain the private information of the cognitive access-point–mobile-phone network, the analytical method can be applied for perfectly symmetric channels. Otherwise, if the information exchange between the licensed network and cognitive network is expensive or impossible, the iterative algorithms can be utilized.

V. Numerical Results

Here, we perform simulations to verify our analysis. The convergence performance of the iterative algorithm, as well as the rate performance for analytical and iterative algorithms, is given numerically. The PU and SUs are uniformly located in a square area of 100×100. Unless specified otherwise, the channel coefficients are generated as $h_{i,j} = d_{i,j}^{-\alpha}\tilde{h}_{i,j}$, where $d_{i,j}$ represents the distance between the transmitter of user i and the receiver of user j, and $\alpha = 3$ is the path loss. Slow fading gain $\tilde{h}_{i,j}$ is modeled as the independently circular symmetric Gaussian distributed random vector. "Average" (e.g., average power and average rate) is taken over 10^4 channel realizations. Unless specified otherwise, the units of power and rate are in megawatts and nats/s/Hz, respectively.

A. NE for the subGame of SUs

First, we compare the analytical solutions in (10) and (11) with Algorithm 1 in the perfectly symmetric channel case. In the simulations, we set $N = 3$, $P_1 = [70\ 10\ 30]$(mW), $P_2^{\max} = 50$(mW), $P_3^{\max} = 10$(mW), $N_2 = N_3 = [10^{-10} 10^{-10} 10^{-10}]$(mW), $h_{22} = h_{33}$, $\tilde{h}_{22} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$, $h_{12} = h_{13}$, $\tilde{h}_{12} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.2\ \ 0.3\ \ 0.4])$, and $h_{23} = h_{32} = 0.5 \times h_{22}$ (i.e., $c = 0.25$). Using (10) and (11), we obtain the average NE power levels $\mathbf{P}_2^* = [13.6871\ 21.6799\ 14.6329]$ and $\mathbf{P}_3^* = [2.4389\ 4.8819\ 2.6792]$. Fig. 2 shows the results of Algorithm 1. Observe that Algorithm 1 converges to the same results as the analytical solutions since the fifth iteration.

Fig. 3 shows the convergence performance of Algorithm 1 when perfectly symmetric channel conditions do not hold. In the simulations, we set $N = 3$, $P_1 = [10\ 30\ 70]$(mW), $P_2^{\max} = 50$(mW), $P_3^{\max} = 10$(mW), $N_2 = N_3 = [10^{-10}\ 10^{-10}\ 10^{-10}]$(mW), $\tilde{h}_{22} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$, $\tilde{h}_{33} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$, $\tilde{h}_{12} \sim \mathcal{CN}(0,\ 0.5 * \mathrm{diag}[0.5\ \ 0.3\ \ 0.4])$, $\tilde{h}_{13} \sim \mathcal{CN}(0,\ 0.5 * \mathrm{diag}[0.36\ 0.3\ 0.0625])$, $\tilde{h}_{23} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.16\ 0.49])$, and $\tilde{h}_{32} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.04\ \ 0.09\ \ 0.25])$. We can notice that Algorithm 1 converges to $\mathbf{P}_2 = [21.0453\ 17.6421\ 11.3126]$ and $\mathbf{P}_3 = [3.4257\ 2.9075\ 3.6668]$ since the fourth iteration.

B. Convergence Performance of Iterative Hierarchic Power Allocation Algorithm for the PU and SUs

The number of inner iterations for Algorithm 2 (i.e., the number of iterations for Algorithm 1) is set to be 10. In Algorithm 3, we let $\tau_k = 3 \times k$.

1) Convergence Performance in Different Channel States With the Same ISR Constraint, the Same Total Power Constraints, and the Same Step Size: We set $N = 3$, $N_1 = N_2 = N_3 = [10^{-10}\ 10^{-10}\ 10^{-10}]$(mW), $\rho = 0.2$, $P_1^{\max} = 100$(mW), $P_2^{\max} = 50$(mW), $P_3^{\max} = 60$(mW), and step size $\delta = \eta = 0.1$.

Figs. 4 and 5 plot the convergence performance of Algorithms 2 and 3 with different channel parameters. It is observed that Algorithms 2 and 3 converge to the same results in Figs. 4 and 5, respectively. Algorithm 2 converges since about the 50th iteration, and Algorithm 3 converges since the 100th iteration. We notice that there are ten inner iterations in each iteration of Algorithm 2; therefore, Algorithm 3 is more time efficient. Comparing Figs. 4 and 5, we can see that the rate performance is better in Fig. 5. This can be explained as follows: Comparing the channel parameters used in Figs. 4 and 5, there is stronger interference in Fig. 4. As a result, the performance is better in Fig. 5.

2) Convergence Performance in the Same Channel State With Different Step Sizes: Parameters are chosen as follows: $N = 3$, $N_1 = N_2 = N_3 = [10^{-10}\ 10^{-10}\ 10^{-10}]$(mW), $\rho = 0.1$, $P_1^{\max} = 100$(mW), $P_2^{\max} = 60$(mW), $P_3^{\max} = 40$(mW), $h_{12} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.25\ 0.36]$, $h_{13} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.09])$,

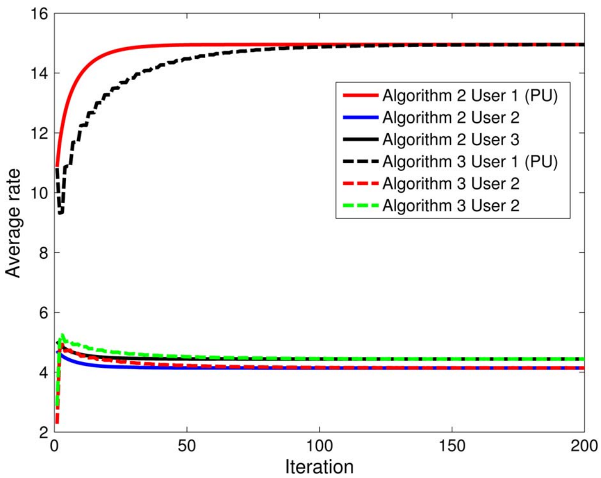


Fig. 4. Convergence performance of Algorithms 2 and 3 with $\tilde{h}_{12} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.49\ 0.25\ 0.36])$, $\tilde{h}_{13} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.49])$, $\tilde{h}_{21} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.25\ 0.36])$, $\tilde{h}_{31} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.16])$, $\tilde{h}_{23} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.25])$, $\tilde{h}_{32} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.25])$, $\tilde{h}_{11} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$, $\tilde{h}_{22} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$, and $\tilde{h}_{33} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$.

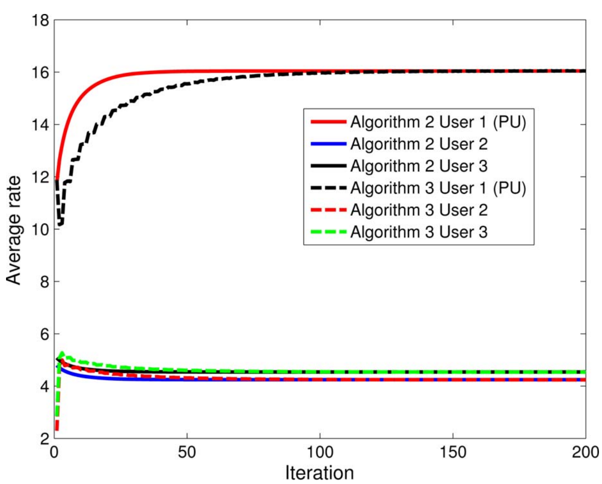


Fig. 5. Convergence performance of Algorithms 2 and 3 with $\tilde{h}_{12} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.25\ 0.36])$, $\tilde{h}_{13} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.09])$, $\tilde{h}_{21} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.36\ 0.25\ 0.36])$, $\tilde{h}_{31} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.49\ 0.25\ 0.16])$, $\tilde{h}_{23} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.09\ 0.81])$, $\tilde{h}_{32} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.25\ 0.36])$, $\tilde{h}_{11} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[4\ 4\ 4])$, $\tilde{h}_{22} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$, and $\tilde{h}_{33} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$.

$h_{21} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.36\ \ 0.25\ \ 0.36])$, $h_{31} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.49\ 0.25\ 0.16])$, $h_{23} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.25])$, $h_{32} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ \ 0.25\ \ 0.25])$, $h_{11} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ \ 1\ \ 1])$, $h_{22} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ \ 1\ \ 1])$, and $h_{33} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$. Fig. 6 demonstrates the convergence performance of Algorithm 2 with different step sizes. We can observe that the algorithm converges with step sizes $\eta = 0.1$ and $\eta = 0.2$. However, when $\eta = 0.8$ and $\eta = 0.9$, the algorithm oscillates, i.e., does not converge. It can be interpreted by using Lemma 2. The upper bound for the convergent step size for all channel realizations lies between 0.2 and 0.8, i.e., $0.2 < \min \xi < 0.8$. Then, the condition in Lemma 2 can be satisfied when $\eta = 0.1$ and $\eta = 0.2$. Consequently, the algorithm

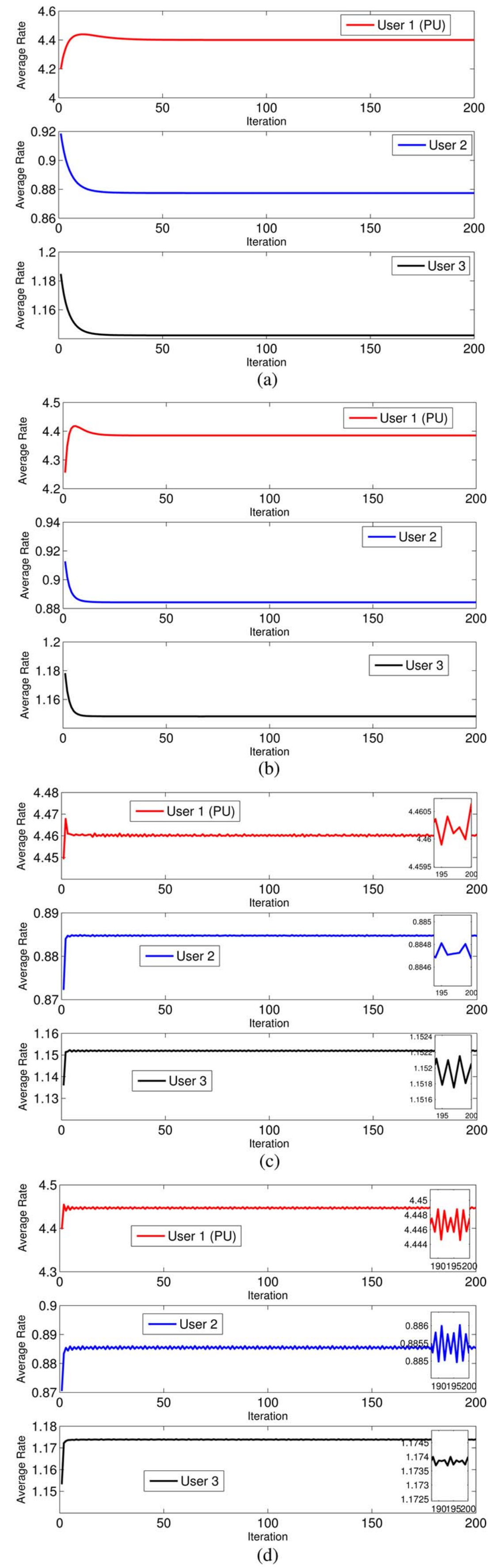


Fig. 6. Convergence performance of Algorithm 2 with different step sizes. (a) $\eta = 0.1$. (b) $\eta = 0.2$. (c) $\eta = 0.8$. (d) $\eta = 0.9$.

converges for all channel realizations, and the average rate converges. When $\eta = 0.8$ and $\eta = 0.9$, $\eta < \xi$ does not hold. Thus, the convergence cannot be guaranteed.[13] It can also be noticed that the step size does not affect the final results of the algorithm when it converges. Fig. 7 shows the convergence performance of Algorithm 3 with different step sizes. Similarly, we observe that the algorithm converges when the step size is set to be 0.1 and 0.2, and it oscillates when the step size equals to 0.9 and 0.99. The algorithm converges to the same results for $\delta = 0.1$ and $\delta = 0.2$.

C. Rate Performance Comparison for Iterative Power Allocation Algorithm and Analytical Algorithm in the Perfectly Symmetric Channel

In the perfectly symmetric channel, both analytical and iterative power allocation for the PU and SUs can be applied.[14] We compare the rate performances here.

Fig. 8 shows the rate performance of the analytical hierarchic power allocation and iterative power allocation for the PU and SUs with different power constraints for the PU, i.e., $P_1^{\max}$. We can observe that the rate performance of the PU decreases slightly in the iterative power allocation because of the unavailability of SUs' private information, but the rate performance of the SUs is almost the same as the analytical algorithm. This verifies the effectiveness of the iterative power allocation.

D. Multi-PU and Multi-SU Scenario

Fig. 9(a) plots the rate performance versus the power constraint of PU when there are 2 PUs and 2 SUs. In the simulation, the parameters are chosen as follows: $N = 3$, $\rho = 0.2$, $P_1^{\max} = P_2^{\max} = P_{\max}$, $P_3^{\max} = 100$, $P_4^{\max} = 150$, and $N_1 = N_2 = N_3 = N_4 = [1\ 1\ 1]$. The channel settings are listed in Table I. Fig. 9(b) shows the two-PU–three-SU scenario. Additionally, $P_5^{\max} = 120$, $h_{15} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.09\ 0.25])$, $h_{25} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.49\ 0.36\ 0.49])$, $h_{35} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.09\ 0.25\ 0.36])$, $h_{45} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.49\ 0.49])$, $h_{55} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.25])$, and $N_5 = [1\ 1\ 1]$. In Fig. 9(a) and (b), we can find that with the increase in $P_{\max}$, the rate performance of PUs increases and that of SUs decreases. Comparing Fig. 9(a) and (b), we can notice that the entry of a new SU (user 5) will degrade the rate performance of the existing users since additional interference will be incurred.

VI. Conclusion

We consider the power allocation for the PU network and the SU network jointly using the Stackelberg game to describe the hierarchy. The PU network is considered the leader and the SU network acts as the follower. We consider the ISR constraint to guarantee the primary service in the Stackelberg game. Based on the analysis of the Stackelberg game, the hierarchic joint power-allocation algorithms are given. The analytical method

[13] Only the convergence cannot be guaranteed but not definitely divergent.

[14] The analytical method is applied when private information is available and the iterative method is used otherwise.

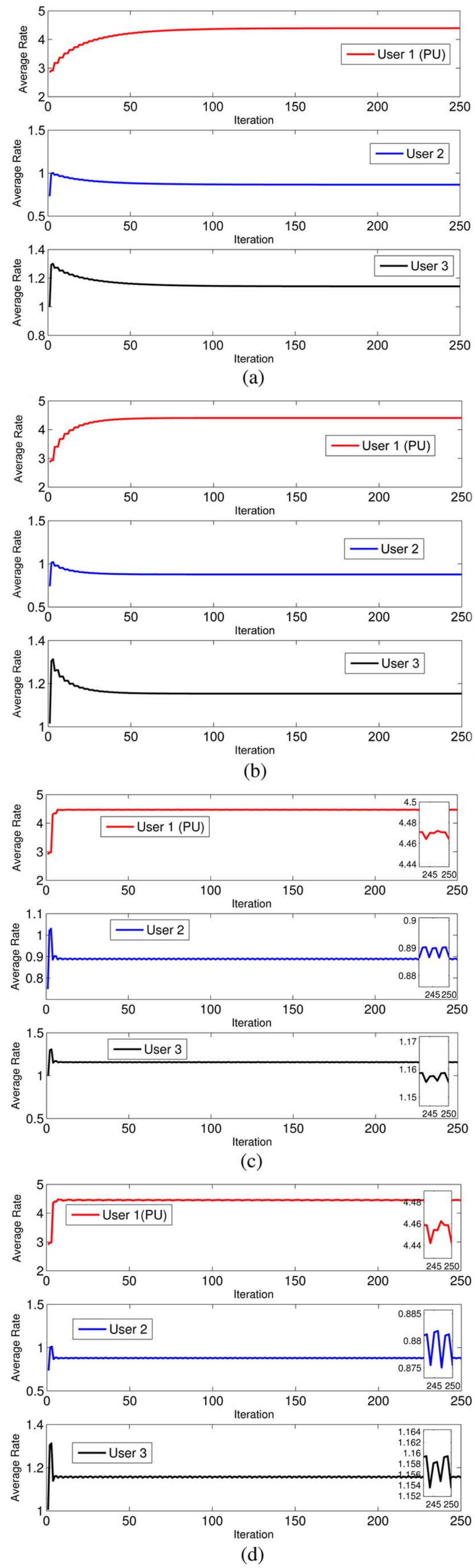


Fig. 7. Convergence performance of Algorithm 3 with different step sizes. (a) $\delta = 0.1$. (b) $\delta = 0.2$. (c) $\delta = 0.9$. (d) $\delta = 0.99$.

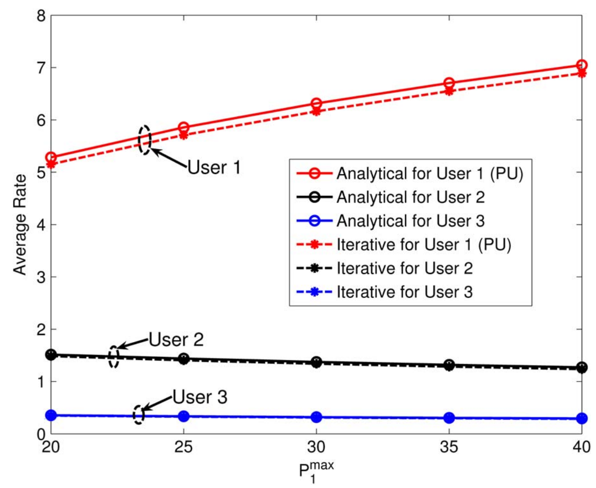
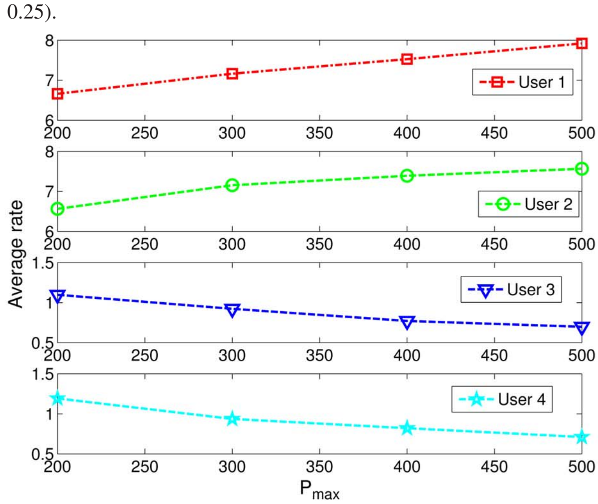
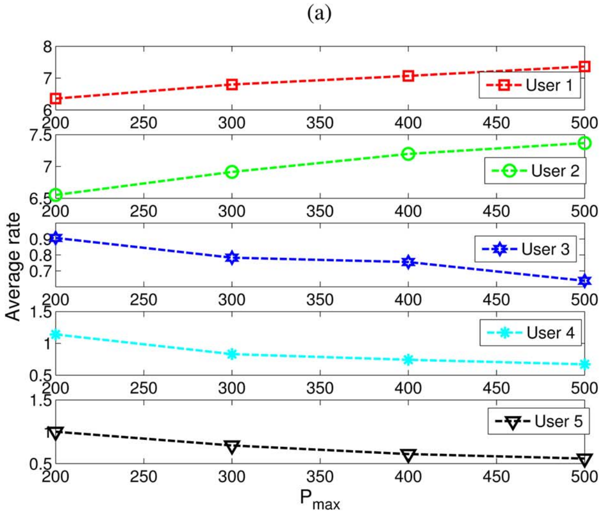


Fig. 8. Rate performance of the analytical power allocation and iterative power allocation in the perfectly symmetric channel with different $P_1^{\max}$. The other parameters are $N = 3$, $N_1 = N_2 = N_3 = [0.5 \quad 0.5 \quad 0.5]$, $\rho = 0.1$, $P_2^{\max} = 5$, $P_3^{\max} = 1$, $h_{11} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1 \quad 1 \quad 1])$, $h_{22} = h_{33} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[1 \quad 1 \quad 1])$, $h_{12} = h_{13} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.2 \quad 0.3 \quad 0.4])$, $h_{21} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.09 \quad 0.36 \quad 0.25])$, $h_{31} \sim \mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.25\ 0.16])$, and $h_{23} = h_{32} = 0.5 \times h_{22}$ (i.e., $c = 0.25$).



(a)



(b)

Fig. 9. Rate performance versus power constraint of PU in the multi-PU–multi-SU scenario applying the distributed iterative algorithm (Algorithm 4), $\eta = 0.001$, and 100 iterations. (a) Two PUs (users 1 and 2) and two SUs (users 3 and 4). (b) Two PUs (users 1 and 2) and three SUs (users 3, 4, and 5).

TABLE I
CHANNEL SETTINGS FOR FIG. 9(a)

h_{12}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.04\ 0.01])$
h_{31}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.09\ 0.49\ 0.04])$
h_{13}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.09])$
h_{32}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.04\ 0.01\ 0.25])$
h_{14}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.25\ 0.36])$
h_{34}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.36\ 0.64\ 0.36])$
h_{21}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.01\ 0.36\ 0.01])$
h_{41}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.16\ 0.09\ 0.04])$
h_{23}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.36\ 0.09])$
h_{42}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.04\ 0.09\ 0.09])$
h_{24}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.36\ 0.36\ 0.25])$
h_{43}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.36\ 0.49])$
h_{11}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$
h_{22}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[1\ 1\ 1])$
h_{33}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.25])$
h_{44}	$\mathcal{CN}(0, 0.5 * \mathrm{diag}[0.25\ 0.25\ 0.25])$

is presented when the PU can obtain the information for the SUs. Once the PU cannot obtain the information for the SUs, distributed iterative methods are proposed. Simulation results demonstrate the effectiveness (in rate and convergence) of the proposed hierarchic joint power-allocation algorithms.

APPENDIX A
PROOF OF PROPOSITION 1

First, $\forall \mathbf{P},\ \mathbf{P}' \in \mathcal{S}_i$, we have $\alpha\mathbf{P} + (1-\alpha)\mathbf{P}' \in \mathcal{S}_i$ ($\alpha \in [0, 1]$), i.e., $\mathcal{S}_i$ is a convex set. Meanwhile, as $P_i^{\max} < \infty$, $\mathcal{S}_i \subseteq \mathbb{E}^N$ is closed and bounded; therefore, it is compact. Next, $u_i(\mathbf{P}_i,\ \mathbf{P}_{-i})$ is continuous in $\mathbf{P}_{-i}$. Third, $\forall \tau \in \mathbb{R}$, we can prove that the upper contour set $\mathcal{U}_\tau = \{\mathbf{P}_i \in \mathcal{S}_i,\ u_i(\mathbf{P}_i,\ \mathbf{P}_{-i}) \geq \tau\}$ is convex. That is, $\forall \mathbf{P},\ \mathbf{P}' \in \mathcal{U}_\tau$, $\alpha\mathbf{P} + (1-\alpha)\mathbf{P}' \in \mathcal{U}_\tau$ ($\alpha \in [0,\ 1]$). Consequently, $u_i(\mathbf{P}_i, \mathbf{P}_{-i})$ is quasi-concave in $\mathbf{P}_i$. Using the Debreu–Fan–Glicksberg theorem [31], the lemma can be proven.

APPENDIX B
PROOF OF PROPOSITION 2

Define $\Lambda(\mathbf{P}) := (\Lambda_2(\mathbf{P})^T, \ldots, \Lambda_{|\Omega|+1}(\mathbf{P})^T)^T$, where

$$\Lambda_i(\mathbf{P}) = -\nabla_{\mathbf{P}_i} u_i(\mathbf{P}_i,\ \mathbf{P}_{-i}) = \left(-\frac{1}{\hat{N}_i^f + \sum_{j\in\mathbb{S}} P_j^f \left|\hat{h}_{j,i}^f\right|^2}\right)_{f=1}^{N} \tag{15}$$

with $\nabla_{\mathbf{P}_i}(\cdot)$ being the gradient vector with respect to $\mathbf{P}_i$ [8], $\hat{N}_i^f = (N_i^f + P_1^f|h_{1,i}^f|^2)/(|h_{i,i}^f|^2)$, and $\hat{h}_{j,i}^f = (h_{j,i}^f)/(h_{i,i}^f)$. Denote $\mathcal{S} = \mathcal{S}_2 \times \cdots \times \mathcal{S}_{|\Omega|+1}$ with a Cartesian structure. For $\mathbf{P} = (\mathbf{P}_2, \ldots, \mathbf{P}_{|\Omega|+1})$, $\mathbf{P}' = (\mathbf{P}'_2, \ldots, \mathbf{P}'_{|\Omega|+1}) \in \mathcal{S}$, define

$$\Theta_i(f) = \sqrt{\hat{N}_i^f + \sum_{j\in\mathbb{S}} P_j^f |\hat{h}_{j,i}^f|^2}\sqrt{\hat{N}_i^f + \sum_{j\in\mathbb{S}} {P'}_j^f |\hat{h}_{j,i}^f|^2} \tag{16}$$

and $e_i(f) = (P_i^f - P_i'^f)/(\Theta_i(f)), i \in \mathbb{S}$. Note that $\hat{N}_i^f \leq \Theta_i(f) \leq \Theta_i^{\max}(f) := \hat{N}_i^f + \sum_{j\in\mathbb{S}} P_j^{\max}|\hat{h}_{j,i}^f|^2$.[15] We have

$$\begin{aligned}
&(\mathbf{P}_i - \mathbf{P}_i')^T \left[\Lambda_i(\mathbf{P}) - \Lambda_i(\mathbf{P}')\right] \\
&= \sum_{f=1}^{N} \frac{P_i^f - P_i'^f}{\Theta_i(f)} \frac{\sum_{j\in\mathbb{S}} \left|\hat{h}_{j,i}^f\right|^2 \left(P_j^f - P_j'^f\right)}{\Theta_i(f)} \\
&\geq \sum_{f=1}^{N} e_i(f)^2 - \sum_{j\neq i}\sum_{f=1}^{N} e_i(f) \frac{\left|\hat{h}_{j,i}^f\right|^2 \Theta_j(f)}{\Theta_i(f)} e_j(f) \\
&\overset{(a)}{\geq} \sum_{f=1}^{N} e_i(f)^2 - \sum_{j\neq i} \left(\sum_{f=1}^{N} e_i(f)^2\right)^{\frac{1}{2}} \\
&\quad \times \max_{1\leq f\leq N} \left\{\frac{\left|\hat{h}_{j,i}^f\right|^2 \Theta_j(f)}{\Theta_i(f)}\right\} \left(\sum_{f=1}^{N} e_j(f)^2\right)^{\frac{1}{2}} \\
&\geq \hat{e}_i \sum_{j\in\mathbb{S}} \mathbf{M}_{i,j} \hat{e}_j
\end{aligned} \tag{17}$$

where (a) holds since the Cauchy–Schwarz inequality and $\hat{e}_j = (\sum_{f=1}^{N} e_j(f)^2)^{1/2}$. When $\mathbf{M}$ is a positive definite matrix, it is a P-matrix.[16] Then, based on [33, Th. 3.3.4(b)], $\alpha(\mathbf{M}) = \min_{\|\mathbf{x}\|_2=1}\{\max_{i\in\mathbb{S}} x_j(\mathbf{M}\mathbf{x})_i\}$ is positive. That is, $\forall \mathbf{x} \in \mathbb{R}^{L-1}$

$$\max_{i\in\mathbb{S}} x_i(\mathbf{M}\mathbf{x})_i \geq \alpha(\mathbf{M})\|\mathbf{x}\|_2^2 \tag{18}$$

where $\|\cdot\|_2$ is the spectral norm. By combining (17) and (18), we have

$$\begin{aligned}
&\max_{i\in\Omega} \left\{(\mathbf{P}_i - \mathbf{P}_i')^T \left[\Lambda_i(\mathbf{P}) - \Lambda_i(\mathbf{P}')\right]\right\} \\
&\geq \alpha(\mathbf{M}) \sum_{j\in\mathbb{S}}\sum_{f=1}^{N} e_j(f)^2 \\
&\geq \frac{\alpha(\mathbf{M})}{\max_{j\in\mathbb{S}} \max_{1\leq f\leq N} \left(\Theta_j^{\max}(f)\right)^2} \sum_{j\in\mathbb{S}}\sum_{f=1}^{N} \left(p_j^f - p_j'^f\right)^2.
\end{aligned} \tag{19}$$

That is, $\exists\gamma > 0$, $\max_{i\in\Omega}\{(\mathbf{P}_i - \mathbf{P}_i')^T[\Lambda_i(\mathbf{P}) - \Lambda_i(\mathbf{P}')]\} \geq \gamma\|\mathbf{P} - \mathbf{P}'\|_2^2$. Consequently, $\mathcal{G}$ has a unique NE according to [34, Prop. 3.5.10(a)].

APPENDIX C
PROOF OF PROPOSITION 3

Denote $(\mathbf{P}_i(k+1))_{i=2}^{L} = \mathfrak{F}((\mathbf{P}_i(k))_{i=2}^{L})$ in Algorithm 1. First, we can derive that $\mathfrak{F}$ is piecewise affine.[17] The domain of function $\mathfrak{F}$ can be partitioned into finitely many polyhedral regions, and in each region, $\mathfrak{F}(\mathbf{P})$ is equal to an affine function $\mathbf{ACP} + \mathbf{b}$, where $\mathbf{A} = \mathrm{diag}(\mathbf{W}(\phi_2), \mathbf{W}(\phi_3), \ldots, \mathbf{W}(\phi_L))$ with

$$[\mathbf{W}(\phi_i)]_{kl} = \begin{cases} 0, & k \notin \phi_i \text{ or } l \notin \phi_i \\ 1/|\phi_i|, & k, l \in \phi_i \text{ and } k \neq l \\ -1 + 1/|\phi_i|, & k, l \in \phi_i \text{ and } k = l \end{cases} \tag{20}$$

being an $(L-1)N \times (L-1)N$ block diagonal matrix, for some choice of $\phi_i \subseteq \{1, 2, \ldots, N\}$, $i = 2, 3, \ldots, L$ [29].[18] Next, it is sufficient to show that each block of A_σ has spectrum norm equal to 1. For this purpose, we can prove that

$$W_m^T W_m = W_m^2 = 1 \tag{21}$$

where $W_m = (1/m)\mathbf{J} - \mathbf{I}$, with $\mathbf{J}$ and $\mathbf{I}$ being an $m \times m$ all-one matrix and an identity matrix, respectively, is a square matrix; it is obtained by removing the zero columns and zero rows in A_σ. From (21), we can derive that the eigenvalues of W_m are 0 and -1.

APPENDIX D
PROOF OF PROPOSITION 4

Since $u_i(\mathbf{P}_i, \mathbf{P}_{-i})$ is concave on $\mathbf{P}_i$, using the Karush–Kuhn–Tucker conditions,[19] $(\mathbf{P}_2, \ldots, \mathbf{P}_{|\Omega|+1})$ is the NE if and only if there are nonnegative $\{\lambda_i\}$ satisfying

$$\begin{aligned}
&\frac{\partial u_i(\mathbf{P}_i, \mathbf{P}_{-i})}{\partial P_i^f} \\
&= \left[P_i^f + \frac{N_i^f + P_1^f \left|h_{1i}^f\right|^2}{\left|h_{ii}^f\right|^2} + \frac{\sum_{j\neq i\in\Omega} P_j^f \left|h_{ji}^f\right|^2}{\left|h_{ii}^f\right|^2}\right]^{-1} \qquad (22) \\
&= \left[P_i^f + \sigma^f + c\sum_{j\neq i\in\Omega} P_j^f\right]^{-1} \begin{cases} = \lambda_i, & P_i^f > 0 \\ \leq \lambda_i, & P_i^f = 0. \end{cases} \qquad (23)
\end{aligned}$$

Consequently, let $\tau_r^k = (1/1-c)((1+(|\Omega|-1-r+k)c)/(\lambda_k) - c\sum_{j=1}^{|\Omega|-r+k}(1/\lambda_j))$ with $\lambda_1 \leq \cdots \leq \lambda_{|\Omega|}$; each NE is of the following form:

$$P_{k+1}^f = \begin{cases} \frac{1}{1+(|\Omega|-1)c}\left(\tau_k^k - \sigma_f\right), & \sigma_f < \tau_{|\Omega|}^{|\Omega|} \\ \frac{1}{1+(|\Omega|-1-r+k)c}\left(\tau_r^k - \sigma_f\right), & \tau_{|\Omega|}^{|\Omega|+k+1-r} \\ \leq \sigma_f < \tau_{|\Omega|}^{|\Omega|+k-r}, & r \in [k+1, |\Omega|] \\ 0, & \tau_{|\Omega|}^k \leq \sigma_f. \end{cases} \tag{24}$$

[15] $0 \leq P_j^f$ and $\sum_{f=1}^{N} P_j^f \leq P_j^{\max}$; thus, $P_j^f \leq P_j^{\max}$. Similarly, $P_j'^f \leq P_j^{\max}$.

[16] A matrix $\mathbf{M}$ is called P-matrix if every principal minor of $\mathbf{M}$ is positive. Any positive definite matrix is P-matrix, but the reverse does not hold [32], [33].

[17] A function g mapping from a domain in $\mathcal{V}$ to vector space $\mathcal{W}$ is called piecewise affine if the domain of g can be partitioned into finitely many polyhedra, e.g., $\mathcal{R}_1, \ldots, \mathcal{R}_S$, such that, for $\sigma = 1, \ldots, S$, the function g restricted to the region $\mathcal{R}_\sigma$ is equal to an affine function, i.e., $g(x) = A_\sigma x + c_\sigma$, for all $x \in \mathcal{R}_\sigma$, for some suitable choice of matrix A_σ and vector c_σ.

[18] $|X|$ is the cardinality of set X.

[19] See [35, Ch. 5.5.3].

For user $(|\Omega|+1)$, we have

$$\begin{aligned}\sum_{f=1}^{N} P_{|\Omega|+1}^{f} &= \sum_{\sigma_f<\tau_{|\Omega|}^{|\Omega|}} P_{|\Omega|+1}^{f}\\ &= \frac{1}{1+(|\Omega|-1)\,c} \sum_{\sigma_f<\tau_{|\Omega|}^{|\Omega|}} \left(\tau_{|\Omega|}^{|\Omega|}-\sigma_f\right)\\ &\le P_{|\Omega|+1}^{\max}. \end{aligned} \tag{25}$$

When the equality holds, we have $\tau_{|\Omega|}^{|\Omega|*} = ((1+(|\Omega|-1)c)P_{|\Omega|+1}^{\max} + \sum_{f=1}^{k_{|\Omega|}} \sigma_f)/(k_{|\Omega|})$, where $k_{|\Omega|}$ is given by $\phi_{k_{|\Omega|}}^{|\Omega|} < P_{|\Omega|+1}^{\max} \le \phi_{k_{|\Omega|}+1}^{|\Omega|}$, and $\phi_k^{|\Omega|} = (1)/(1+(|\Omega|-1)c)\sum_{f=1}^{k}(\sigma_k-\sigma_f)$. Consequently, the equilibrium power allocation for user $(|\Omega|+1)$ is given by

$$P_{|\Omega|+1}^{f*} = \begin{cases} \frac{\tau_{|\Omega|}^{|\Omega|*}-\sigma_f}{1+(|\Omega|-1)c}, & f\in\left[1,k_{|\Omega|}\right]\\ 0, & f\in\left[k_{|\Omega|}+1,N\right]. \end{cases} \tag{26}$$

$\tau_{|\Omega|-1}^{|\Omega|-1} = (1+(|\Omega|-1)c)/(1+(|\Omega|-2)c)\tau_{|\Omega|}^{|\Omega|-1} - (c)/(1+(|\Omega|-2)c)\tau_{|\Omega|}^{|\Omega|}$, then regarding user $|\Omega|$

$$\begin{aligned}\sum_{f=1}^{N} P_{|\Omega|}^{f} = &\frac{1}{1+(|\Omega|-1)\,c} \sum_{\sigma_f<\tau_{|\Omega|}^{|\Omega|}} \left(\tau_{|\Omega|-1}^{|\Omega|-1}-\sigma_f\right)\\ &+ \frac{1}{1+(|\Omega|-2)\,c} \sum_{\tau_{|\Omega|}^{|\Omega|}\le\sigma_f<\tau_{|\Omega|}^{|\Omega|-1}} \left(\tau_{|\Omega|}^{|\Omega|-1}-\sigma_f\right)\\ \le &P_{|\Omega|}^{\max}. \end{aligned} \tag{27}$$

Utilizing the equality, we get $\tau_{|\Omega|}^{(|\Omega|-1)*} = (P_{|\Omega|}^{\max} + (\sum_{f=k_{|\Omega|}}^{k_{|\Omega|-1}} \sigma_f)/(1+(|\Omega|-2)c) + (\sum_{f=1}^{k_{|\Omega|}}((c\tau_{|\Omega|}^{|\Omega|*})/(1+(|\Omega|-2)c+\sigma_f))/(1+(|\Omega|-1)c))((1+(|\Omega|-2)c)/(k_{|\Omega|-1}))$, where $k_{|\Omega|-1}$ is derived by

$$\begin{cases} k_{|\Omega|-1} = k_{|\Omega|}, & P_{|\Omega|}^{\max} \le \phi_{k_{|\Omega|}+1}^{|\Omega|-1}\\ \phi_{k_{|\Omega|-1}}^{|\Omega|-1} < P_{|\Omega|}^{\max} \le \phi_{k_{|\Omega|-1}+1}^{|\Omega|-1}, & \text{otherwise} \end{cases} \tag{28}$$

with

$$\begin{aligned}\phi_k^{|\Omega|-1} = &\sum_{f=k_{|\Omega|}+1}^{k} \frac{\sigma_k-\sigma_f}{1+(|\Omega|-2)\,c} + \sum_{f=1}^{k_{|\Omega|}} \frac{1}{1+(|\Omega|-2)\,c}\\ &\times\left(\frac{1+(|\Omega|-1)\,c}{1+(|\Omega|-2)\,c}\sigma_k - \sigma_f + \frac{c}{1+(|\Omega|-2)\,c}\tau_{|\Omega|}^{|\Omega|*}\right). \end{aligned} \tag{29}$$

Then

$$P_{|\Omega|}^{f*} = \begin{cases} \frac{\tau_{|\Omega|}^{(|\Omega|-1)*}}{1+(|\Omega|-2)c} - \frac{\frac{c\tau_{|\Omega|}^{|\Omega|*}}{1+(|\Omega|-2)c}+\sigma_f}{1+(|\Omega|-1)c}, & f\in\left[1,k_{|\Omega|}\right]\\ \frac{\tau_{|\Omega|}^{(|\Omega|-1)*}-\sigma_f}{1+(|\Omega|-2)c}, & f\in\left[k_{|\Omega|}+1,k_{|\Omega|-1}\right]\\ 0, & f\in\left[k_{|\Omega|-1}+1,N\right]. \end{cases} \tag{30}$$

As $|\Omega|=2$, we arrive at the proposition, which completes the proof.

Appendix E
Proof of Lemma 1

In Definition 2, Inequality 1 implies that $(\hat{\mathbf{P}}_i^*, \hat{\mathbf{P}}_{-i}^*)$ is the NE of $\mathcal{G}$ given $\mathbf{P}_1^*$. As $(\mathbf{P}_i^*, \mathbf{P}_{-i}^*)$ denotes the NE of $\mathcal{G}$ given $\mathbf{P}_1$, we have an equivalent definition, i.e., $(\mathbf{P}_1^*, \mathrm{Ne}(\mathbf{P}_1^*))$ is an SE if $u_1(\mathbf{P}_1^*, \mathrm{Ne}(\mathbf{P}_1^*)) \ge u_1(\mathbf{P}_1, \mathrm{Ne}(\mathbf{P}_1))$ for any feasible $\mathbf{P}_1$, where $\mathrm{Ne}(x)$ denotes the NE of $\mathcal{G}$ given $\mathbf{P}_1 = x$. Since $\mathbf{P}_1^*$ is the optimal solution of (5), $u_1(\mathbf{P}_1^*, \mathrm{Ne}(\mathbf{P}_1^*)) \ge u_1(\mathbf{P}_1, \mathrm{Ne}(\mathbf{P}_1))$ for any feasible $\mathbf{P}_1$. According to the equivalent definition of Definition 2, we claim that $(\mathbf{P}_1, \mathrm{Ne}(\mathbf{P}_1))$ is an SE.

Appendix F
Proof of Lemma 2

Denote $\chi(\mathbf{P}_1(n)) = [(I^f(n))/(\rho|h_{1,1}^f|^2) + (\lambda - (I^f(n)+N_1^f)/(|h_{11}|^2))^+]_{f=1}^N$, where $[x_i]_{i=1}^n = (x_1,\ldots,x_n)$. Then

$$\begin{aligned}\mathbf{P}_1(n+1) &= (1-\eta)\mathbf{P}_1(n) + \eta\chi\left(\mathbf{P}_1(n)\right)\\ &:= F\left(\mathbf{P}_1(n)\right). \end{aligned} \tag{31}$$

First, $\forall \mathbf{P}_1^{(1)} \ne \mathbf{P}_1^{(2)}$ in PU's feasible power set, as $\sum_{f=1}^N P_i^f \le P_i^{\max}$ for $i\in\mathbb{P}\cup\mathbb{S}$, $\exists\beta>0$ satisfies

$$\begin{aligned}&\left(\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\right)\left[\chi\left(\mathbf{P}_1^{(1)}\right)-\chi\left(\mathbf{P}_1^{(2)}\right)\right]^T\\ &\qquad\ge -\beta\left\|\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\right\|_2^2. \end{aligned} \tag{32}$$

Next, from (31), we get

$$\begin{aligned}&\left(\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\right)\left[F\left(\mathbf{P}_1^{(1)}\right)-F\left(\mathbf{P}_1^{(2)}\right)\right]^T\\ &= (1-\eta)\left(\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\right)\left(\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\right)^T\\ &\quad+\eta\left(\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\right)\left[\chi\left(\mathbf{P}_1^{(1)}\right)-\chi\left(\mathbf{P}_1^{(2)}\right)\right]^T\\ &\overset{(b)}{\ge} [1-(1+\beta)\eta]\left\|\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\right\|_2^2 \end{aligned} \tag{33}$$

where (b) holds since (32). On the other hand, $\exists\theta>0$, $\|\chi(\mathbf{P}_1^{(1)})-\chi(\mathbf{P}_1^{(2)})\|_2 \le \theta\|\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\|_2$. Consequently, we derive

$$\begin{aligned}&\left(\mathbf{P}_1^{(1)}-\mathbf{P}_1^{(2)}\right)\left[F\left(\mathbf{P}_1^{(1)}\right)-F\left(\mathbf{P}_1^{(2)}\right)\right]^T\\ &\qquad\ge (1-(1+\beta)\eta)\,\theta^{-2}\left\|\chi\left(\mathbf{P}_1^{(1)}\right)-\chi\left(\mathbf{P}_1^{(2)}\right)\right\|_2^2. \end{aligned} \tag{34}$$

When $\eta<(1+\beta)^{-1}$, $F(\cdot)$ is cocoercive with constant $[1-(1+\beta)\eta]\theta^{-2}$. Then, applying [34, Th. Th. 12.1.8], if $\eta<2[1-(1+\beta)\eta]\theta^{-2}$, i.e., $\eta<2[2(1+\beta)-\theta^2]^{-1}$, the iterative algorithm converges. In conclusion, if $\eta<\min\{(1+\beta)^{-1}, 2[2(1+\beta)-\theta^2]^{-1}\} = (1+\beta)^{-1} := \xi$, the iterative algorithm converges, which completes the proof.

Acknowledgment

The authors would like to thank the anonymous reviewers and the editor for their constructive comments.

References

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[2] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjrungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[3] C.-G. Yang, J.-D. Li, and Z. Tian, "Optimal power control for cognitive radio networks under coupled interference constraints: A cooperative game-theoretic perspective," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1696–1706, May 2010.

[4] L. Cao and H. Zheng, "Distributed spectrum allocation via local bargaining," in *Proc. IEEE SECON*, Santa Clara, CA, USA, 2005, pp. 475–486.

[5] J. Suris, L. A. DaSilva, Z. Han, A. B. MacKenzie, and R. S. Komali, "Asymptotic optimality for distributed spectrum sharing using bargaining solutions," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5225–5237, Oct. 2009.

[6] O. N. Gharehshiran, A. Attar, and V. Krishnamurthy, "Dynamic coalition formation for resource allocation in cognitive radio networks," in *Proc. IEEE ICC*, Cape town, South Africa, May 2010, pp. 1–6.

[7] Y.-E. Lin, K.-H. Liu, and H.-Y. Hsieh, "Design of power control protocols for spectrum sharing in cognitive radio networks: A game-theoretic perspective," in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010, pp. 1–6.

[8] J. S. Pang, A. Scutari, D. Palomar, and F. Facchinei, "Design of cognitive radio systems under temperature-interference constraints: A variational inequality approach," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3251–3271, Jun. 2010.

[9] P. Lin, J. Jia, Q. Zhang, and M. Hamdi, "Dynamic spectrum sharing with multiple primary and secondary users," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1756–1765, May 2011.

[10] M. Hong and A. Garcia, "Equilibrium pricing of interference in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6058–6072, Dec. 2011.

[11] R. Ilies, F. P. Morgeson, and J. D. Nahrgang, "Authentic leadership and eudaemonic well-being: Understanding leader-follower outcomes," *Lead. Quart.*, vol. 16, no. 3, pp. 373–394, Jun. 2005.

[12] M. Bennis, M. Le Treust, S. Lasaulce, M. Debbah, and J. Lilleberg, "Spectrum sharing games on the interference channel," in *Proc. GameNets*, Istanbul, Turkey, May 2009, pp. 515–522.

[13] Y. Su and M. van der Schaar, "A new perspective on multi-user power control games in interference channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2910–2919, Jun. 2009.

[14] Y. Su and M. van der Schaar, "Conjectural equilibrium in multi-user power control games," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3638–3650, Sep. 2009.

[15] S. Lasaulce, Y. Hayel, R. El Azouzi, and M. Debbah, "Introducing hierarchy in energy games," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3833–3843, Jul. 2009.

[16] G. He, S. Lasaulce, and Y. Hayel, "Stackelberg games for energy-efficient power control in wireless networks," in *Proc. IEEE INFOCOM*, Shanghai, China, Apr. 2011, pp. 591–595.

[17] B. Wang, Z. Han, and K. J. R. Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using Stackelberg game," *IEEE Trans. Mobile Comput.*, vol. 8, no. 7, pp. 975–990, Jul. 2009.

[18] S. Guruacharya, D. Niyato, E. Hossain, and D. I. Kim, "Hierarchical competition in femtocell-based cellular networks," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 2010, pp. 1–5.

[19] M. Bloem, T. Alpcan, and T. Basar, "A Stackelberg game for power control and channel allocation in cognitive radio networks," presented at the Proc. GameComm, Nantes, France, Oct. 2007, Paper 4.

[20] A. A. Daoud, T. Alpcan, S. Agarwal, and M. Alanyali, "A Stackelberg game for pricing uplink power in wide-band cognitive radio networks," in *Proc. IEEE CDC*, Cancun, Mexico, Dec. 2008, pp. 1422–1427.

[21] J. Zhang and Q. Zhang, "Stackelberg game for utility-based cooperative cognitive radio networks," in *Proc. ACM MobiHoc*, New Orleans, LA, USA, May 2009, pp. 23–31.

[22] Y. Li, X. Wang, and M. Guizani, "Resource pricing with primary service guarantees in cognitive radio networks: A Stackelberg game approach," in *Proc. IEEE GLOBECOM*, Honolulu, HI, USA, Dec. 2009, pp. 1–5.

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

[24] M. Razaviyayn, Y. Morin, and Z.-Q. Luo, "A Stackelberg game approach to distributed spectrum management," in *Proc. IEEE ICASSP*, Dallas, TX, USA, Mar. 2010, pp. 3006–3009.

[25] Y. Xiao, G. Bi, and D. Niyato, "Distributed optimization for cognitive radio networks using Stackelberg game," in *Proc. IEEE ICCS*, San Francisco, CA, USA, Oct. 2010, pp. 77–81.

[26] C. Yang and J. Li, "Capacity maximization in cognitive networks: A Stackelberg game-theoretic perspective," in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010, pp. 1–5.

[27] N. Omidvar and B. H. Khalaj, "A game theoretic approach for power allocation in the downlink of cognitive radio networks," in *Proc. IEEE CAMAD*, Kyoto, Japan, Jun. 2011, pp. 158–162.

[28] X. Kang, Y.-C. Liang, and H. K. Garg, "Distributed power control for spectrum-sharing femtocell networks using Stackelberg game," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 2011, pp. 1–5.

[29] K. W. Shum, K.-K. Leung, and C. Sung, "Convergence of iterative waterfilling algorithm for Gaussian interference channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 6, pp. 1091–1100, Aug. 2007.

[30] J. G. Proakis and M. Salehi, *Digital Communications,* 5th ed. New York, NY, USA: McGraw Hill, 2007.

[31] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991.

[32] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA, USA: SIAM, 1987.

[33] R. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*. Cambridge, U.K.: Cambridge Univ. Press, 1992.

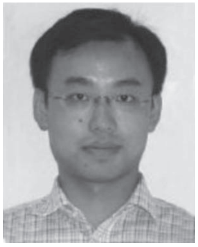
[34] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problem*. New York, NY, USA: Springer-Verlag, 2003.

[35] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[36] H. Y. Benson, D. F. Shanno, and R. J. Vanderbei, "Interior-point methods for nonconvex nonlinear programming: Filter methods and merit functions," Dept. Oper. Res. Financial Eng., Princeton Univ., Princeton, NJ, USA, Tech. Rep. ORFE-00-06, 2000.

Tian Zhang received the B.S. and M.S. degrees from Shandong Normal University, Jinan, China, in 2006 and 2009, respectively. He is currently working toward the Ph.D. degree with the School of Information Science and Engineering, Shandong University, Jinan, China.

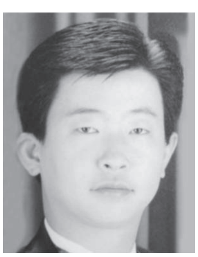
From March 2010 to June 2013, he was also a Visiting Research Staff with Tsinghua University, Beijing, China. His research interests include wireless communications and smart grids.



Wei Chen (S'05–M'07–SM'13) received the B.S. and Ph.D. degrees in electronic engineering (both with the highest honors) from Tsinghua University, Beijing, China, in 2002 and 2007, respectively.

From 2005 to 2007, he was a Visiting Research Staff with the Hong Kong University of Science and Technology, Kowloon, Hong Kong. Since July 2007, he has been with the Department of Electronic Engineering, Tsinghua University, where he is currently a Full Professor and the Vice Director of the Institute of Communications. From June 2010 to September 2010, he was a Visiting Researcher with the University of Southampton, Southampton, U.K. His research interests include wireless communications, information theory, and applied optimizations.

Dr. Chen is a 973 Youth Project Chief Scientist and is supported by the New Century Talent Program of the Chinese Ministry of Education, by the Beijing Nova Program, and by the 100 Fundamental Research Talents Program of Tsinghua University (also known as 221 Talents Program). He is the Vice Director of the Youth Committee of the China Institute of Communications. He served as a Student Travel Grant Chair for the 2008 IEEE International Conference on Communications (IEEE ICC), as a Cochair for the Technical Program Committee (TPC) of the Wireless Communication Symposium of the 2010 IEEE ICC and for the TPC of the 2011 Spring IEEE Vehicular Technology Conference, as the Publication Chair for the 2012 IEEE International Conference on Communications in China, as a Track Cochair of the wireless track for the 2013 IEEE Consumer Communications and Networking Conference, and as a Tutorial Cochair for the 2013 IEEE ICC. He served as an Editor for the IEEE WIRELESS COMMUNICATIONS LETTERS. He received the Best Paper Awards at the 2006 IEEE ICC, the 2007 IEEE International Workshop on Cross Layer Design, and the 2012 IEEE International Conference on Smart Grid Communications; the 2009 IEEE Marconi Prize Paper Award; the 2010 IEEE Communication Society Asia–Pacific Board Best Young Researcher Award; the 2011 Tsinghua Raising Academic Star Award; the 2012 Tsinghua Young Faculty Teaching Excellence Award; the First Prize in the First National Young Faculty Teaching Competition, the First Prize in the Seventh Beijing Young Faculty Teaching Competition; the First Prize in the Fifth Tsinghua University Young Faculty Teaching Competition; and the National May 1st Medal. He also holds the honorary title of Outstanding Teacher in Beijing.



Zhu Han (S'01–M'04–SM'09–F'14) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1997 and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, MD, USA, in 1999 and 2003, respectively.

From 2000 to 2002, he was a Research and Development Engineer with JDS Uniphase Corporation, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, Boise, ID, USA. He is currently an Associate Professor of the Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, wireless multimedia, security, and smart grid communications.

Dr. Han has been an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since 2010. He received the National Science Foundation CAREER Award in 2010 and the IEEE Fred W. Ellersick Prize in 2011.

Zhigang Cao (M'84–SM'85) graduated from Tsinghua University, Beijing, China, in 1962.

Since then, he has been with Tsinghua University, where he is currently a Professor with the Department of Electronic Engineering. From 1984 to 1986, he was a Visiting Scholar with Stanford University, Stanford, CA, USA. In 1997, he was a Visiting Professor with Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong. He is the author of six books and more than 500 papers on communications and signal processing fields, and is the holder of over 20 patents. His current research interests include mobile communications and satellite communications.

Mr. Cao is a Fellow of the Chinese Institute of Communication, a Senior Member of the Chinese Institute of Electronics, and a Member of the Institute of Electronics, Information and Communication Engineers. He currently serves as an Associate Editor-In-Chief of *ACTA Electronics Sinica*. He also serves as the Editor for *China Communications*, *Journal of Astronautics*, and *Frontiers of Electrical and Electronic Engineering*. He has received 11 research awards and special grants from the Chinese Government for his outstanding contributions to education and research. He has also coreceived several best paper awards, including the 2009 IEEE Marconi Prize Paper Award.