Received March 21, 2018, accepted April 14, 2018, date of publication April 17, 2018, date of current version May 9, 2018.

*Digital Object Identifier 10.1109/ACCESS.2018.2828033*

# A One-Leader Multi-Follower Bayesian-Stackelberg Game for Anti-Jamming Transmission in UAV Communication Networks

**YIFAN XU[1,2], (Student Member, IEEE), GUOCHUN REN[1], JIN CHEN[1], YUNPENG LUO[1], LULIANG JIA[1], (Student Member, IEEE), XIN LIU[3], YANG YANG[1], AND YUHUA XU[1,2], (Member, IEEE)**

[1]College of Communications Engineering, Army Engineering University of PLA, Nanjing 210000, China
[2]Science and Technology on Communication Networks Laboratory, Shijiazhuang 050002, China
[3]College of Information Science and Engineering, Guilin University of Technology, Guilin 541000, China

Corresponding authors: Guochun Ren (guochunren@yeah.net) and Yuhua Xu (yuhuaenator@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771488, Grant 61671473, and Grant 61631020, in part by the Natural Science Foundation for Distinguished Young Scholars of Jiangsu Province under Grant BK20160034, and in part by the Open Research Foundation of Science and Technology on Communication Networks Laboratory.

**ABSTRACT** This paper focuses on the anti-jamming transmission problem in unmanned aerial vehicle (UAV) communication networks. Considering the incomplete information constraint and the co-channel mutual interference, a Bayesian Stackelberg game is proposed to formulate the competitive relations between UAVs (users) and the jammer. Specifically, the jammer acts as the leader, whereas users act as followers of the proposed game. Based on their utility functions, the jammer and users select their optimal power control strategies respectively. In addition, the observation error is also investigated in this paper. Due to the incomplete information constraint and the existence of co-channel mutual interference, it is challenging to obtain Stackelberg equilibrium (SE) of the proposed game. Thus, a sub-gradient based Bayesian Stackelberg iterative algorithm is proposed to obtain SE. Finally, simulations are conducted to demonstrate the performance of the proposed algorithm, and the impact of incomplete information and observation error on users' utilities are also presented.

**INDEX TERMS** Anti-jamming, unmanned aerial vehicle, Bayesian Stackelberg game, power control, incomplete information.

## I. INTRODUCTION

Unmanned aerial vehicles (UAV) communication networks have been a hot topic recently [1]. In this regard, anti-jamming transmission is playing an increasingly important role in building UAV communication networks. When UAVs suffer serious jamming, they can no longer build connections to other UAVs as well as the control site, which is a crucial problem to be solved. In addition, when several UAVs share one common channel, mutual interference would be triggered, which is also a factor that deteriorates the system performance.

Since jamming attacks pose significant threat to the wireless communication security, more attention has been paid to anti-jamming studies in recent years [2]. A large number of anti-jamming countermeasures were proposed for the purpose of reducing or avoiding the influences from malicious jammer, e.g., [3]–[7]. In the anti-jamming transmission filed, there are still several problems remaining unsolved. Firstly, it should be noticed that most existing studies rarely considered multi-user case in anti-jamming transmission field, and the co-channel mutual interference, as an indispensable factor, has hardly been investigated neither. Secondly, most existing work made an assumption that user-side and jammer-side can obtain complete information of their opponents, which is not realistic in some cases.

In this paper, we mainly concentrate on the anti-jamming transmission problem in UAV communication networks, and a Bayesian Stackelberg game approach is proposed. The reason for using this game model is twofold. Firstly, considering the hierarchical interactions between UAV-side and jammer-side in the anti-jamming transmission field, and that each decision-maker makes decisions spontaneously and independently, Stackelberg game is a befitting game model to formulate. Secondly, due to the mobility of UAVs, the acquisition of information is more difficult in UAV communication networks than in traditional wireless communication networks, which is the main cause of incomplete information constraint.

In addition, we focus a UAV power transmission system where jamming attack and mutual interference exist simultaneously. For jamming attack, the smart jammer, who can adjust its strategy adaptively, is mainly investigated. Whereas for mutual interference, the co-channel interference is mainly considered.

While analyzing all these factors mentioned above, we formulate a Bayesian Stackelberg game in UAV communication networks to cope with the anti-jamming transmission problem among UAVs. Besides, we propose a sub-gradient [8] based Bayesian Stackelberg iterative algorithm (SBBSIA), with which UAVs and the jammer can adjust their strategies to be optimal. The main contributions of this paper can be summarized as follows:

- The anti-jamming transmission issue in UAV networks is investigated using a Bayesian Stackelberg game, with multi-users case taken into consideration. Besides, the co-channel mutual interference among users is also investigated in anti-jamming transmission field.
- The Stackelberg Equilibrium (SE) of the proposed game is obtained. For the purpose of obtaining SE, the SBBSIA algorithm is designed. In addition, the existence as well as the uniqueness of SE have been proved.
- The process of the scheme is presented, and masses of simulations have been conducted to support the theoretical analysis. Simulation results show that the proposed SBBSIA algorithm is effective.

Note that an anti-jamming Stackelberg game approach in UAV communication networks was presented in our previous work [9], and the main differences and new contributions are: (i) Incomplete information and observation error are introduced in this work. (ii) By introducing incomplete information and observation error, it brings difficulty in analyzing the game model, and the model proposed in [9] can not be employed directly. Moreover, a Bayesian Stackelberg approach can be found in our work [37]. The main differences are: (i) Multi-user case is mainly considered in our paper, while one-user case was investigated in [37]. (ii) By introducing the multi-user case, co-channel mutual interference is also considered.

The rest part of this paper is organized as follows. In Section II, we review the related work. In Section III, the system model and problem formulation are presented. In Section IV, the anti-jamming Bayesian Stackelberg game with incomplete information is presented, and the operation process of the proposed game is shown. In Section V, plentiful simulations are conducted, and the convergence as well as the influence of incomplete information and observation error are demonstrated. In the end, we make conclusion in Section VI.

## II. RELATED WORK

Under the background of wireless communication networks, abundant studies with respect to the anti-jamming transmission problem have been proposed. For example, a network performance maximization problem under the existence of jammer was modeled as a joint power control and user scheduling optimization in [6]. While considering the competition between the user-side and jammer-side, game theory [10]–[17] is a fantastic theoretical tool in modeling the jamming and anti-jamming problems. Xiao [18] applied game theory approaches to study the competitive interactions between secondary user and jammer, and game theoretic solutions were also provided in anti-jamming CRNs. Xiao *et al.* [19] proposed a novelty scheme which uses unmanned aerial vehicles (UAVs) to relay the message of an OBU (onboard units) and to improve the communication performance of VANETs against smart jammers. Moreover, an anti-jamming UAV relay game was formulated. In [20], a non-zero-sum power control game was modeled for the anti-jamming issue, with transmission cost taken into consideration. In [21], a zero-sum game which provided several defense schemes for user-side was modeled in cognitive radio networks. Considering the time-varying spectrum environment, an anti-jamming stochastic game [22] was also proposed. In cognitive radio communications, the interaction between secondary users and attackers was formulated as a dogfight game [23]. In [24], a stochastic game was proposed to fight against sweep attack in cognitive ad-hoc networks. In [25] and [26], Markov game was proposed to formulate as well as to obtain the optimal strategy for the legitimate transmitter. In [27], a three-player game consisting of two secondary users and a jammer was formulated with the awareness of "rendezvous" channel.

In further analysis, considering the hierarchical interactions between user-side and jammer-side, Stackelberg game [28], [29] was formulated in anti-jamming transmission field. Yang *et al.* [30], [31] formulated the anti-jamming problem using a Stackelberg game approach. The user was assumed to be the leader who took actions firstly, while the smart jammer acted secondarily as the follower. The strategy action of the jammer was based on the strategy action of the user. Li *et al.* [32] proposed an anti-jamming Stackelberg game in cooperative wireless networks, viewing the relay node as the vice leader of the game. In [33], a hierarchical learning solution on discrete power strategies was proposed for anti-jamming Stackelberg game.

There are also some literature taking incomplete information into consideration. In [34], it was discussed that users did not have complete information about the traffic dynamics, channel characteristics and other important parameters of other users in wireless networks, based on which, a jamming game was modeled. Altman *et al.* [35] solved the problem from game-theoretic perspective under the assumption that the user knew little about the distribution of jammer and certain information of fading channel. Moreover, In [36], Xiao *et al.* viewed the observation error as a influence factor while proposing a Stackelberg game model containing one jammer and one user, and the power constraint of each side was investigated. Jia *et al.* [37] applied an anti-jamming Bayesian Stackelberg game approach to investigate the best transmission schemes for both players. Theoretically, the demonstration of Bayesian Stackelberg game was also given smoothly.

Different from most of the existing work, in this paper, we formulate the anti-jamming transmission problem as a one-leader multi-follower Bayesian Stackelberg game in UAV communication networks. Besides, the existence of malicious jammer and co-channel mutual interference is considered simultaneously.

## III. SYSTEM MODEL AND PROBLEM FORMULATION

### A. SYSTEM MODEL

The system model is shown in Fig.1. It is assumed that there are one jammer and one UAV group which includes $n$ users. A transmitter-receiver pair is treated as one user in the system. The UAV group is assigned with one common channel while flying towards its destination. Simultaneously, the transmitters in the group are transmitting to the receivers. In general case, when users are transmitting, the malicious jammer will disturb the normal transmission among users. Both the smart jammer and users can adjust their transmission power adaptively. In the system model, the smart jammer has the ability to learn and to sense when and where the user would start transmitting, and it has the intention to jam the channel assigned to the UAV group.

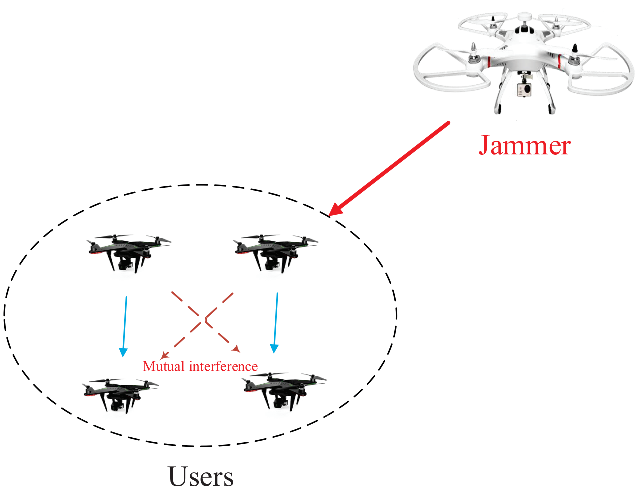



**FIGURE 1. The existence of malicious jammer and mutual interference in UAV communication networks.**

According to the model mentioned above, the users' set is denoted as $\mathcal{N} = [1, 2, \ldots, N]$, $i \in \mathcal{N}$ denotes the $i$th user in the UAV group. It is assumed that the channel state keeps unchanged when users are transmitting. The power vector of the UAVs is denoted as $\mathbf{P} = [P_1, \ldots, P_i, \ldots, P_N]$. Moreover, the smart jammer's transmission power is denoted as $J$. The background noise is denoted as $N_0$ for the purpose of simplification. For user $i$, it suffers from the co-channel mutual interference caused by other users, and the malicious jamming signal caused by the smart jammer. The SINR level of the user $i$ is determined by:

$$\gamma_i\left(P_i, \mathbf{P_{-i}}, J\right) = \frac{P_i \alpha_i}{N_0 + \beta_i J + I_i\left(\mathbf{P_{-i}}\right)}, \tag{1}$$

where $\alpha_i$ denotes the channel gain of user $i$'s transmission link, and $\beta_i$ represents the channel gain of the smart jammer's jamming link to user $i$. $\mathbf{P_{-i}} = [P_0, P_1, \ldots, P_{i-1}, P_{i+1}, \ldots, P_N]$ is a vector of all users' transmission power except user $i$. $I_i\left(\mathbf{P_{-i}}\right) = \sum_{m \neq i, m \in N} P_m \theta_{m,i}$ represents the co-channel mutual interference, and $\theta_{m,i}$ is the mutual interference gain from user $m$ to user $i$. Based on the simplified path-loss model [38] which has been widely used in jamming and anti-jamming games [39]–[41], it is assumed that the signals are influenced by propagation losses and power decay. Specifically, the channel gain of user $i$'s transmission link is denoted as $\alpha_i = d_i^{-\delta}$, the channel gain of the smart jammer's jamming link to user $i$ is denoted as $\beta_i = d_{J,i}{}^{-\delta}$, and the co-channel mutual interference gain from user $m$ to user $i$ is represented as $\theta_{m,i} = d_{m,i}{}^{-\delta}$, where $\delta$ is the path-loss exponent, $d_i$ denotes the distance between the user $i$'s transmitter-receiver pair, $d_{J,i}$ represents the jamming distance to user $i$, and $d_{m,i}$ is the distance from user $m$ to user $i$.

According to the definition shown above, the user $i$'s throughput can be denoted as:

$$Tr_i(B) = B\log_2\left(1 + \gamma_i\left(P_i, \mathbf{P_{-i}}, J\right)\right). \tag{2}$$

After making a normalization on the channel bandwidth $B$, the throughput of user $i$ is presented as:

$$Tr_i = \log_2\left(1 + \gamma_i\left(P_i, \mathbf{P_{-i}}, J\right)\right). \tag{3}$$

When users or smart jammer starts transmitting, it can not be ignored that the existence of transmitting cost will have a great impact on the decision of both user-side and jammer-side. Motivated by [20], we introduce $C_u$ and $C_j$ for users and the jammer, which denotes the transmission cost per unit power. When transmitting with power $P_i$, the transmission cost of user $i$ can be denoted as:

$$Tc_i = C_u P_i. \tag{4}$$

Similarly, the transmission cost of the smart jammer can be expressed as:

$$Tc_j = C_j J. \tag{5}$$

In the system model, the existence of both mutual interference and malicious jammer has been considered. However, in a realistic scenario, there exist uncertainties for both users and the smart jammer while sensing the precise information of the opponent. Motivated by [37], such uncertainties of the channel state information (CSI) have also been taken into consideration in our paper.

For users who can share internal information within a UAV group, it is assumed that each user knows the information of other users perfectly. However, users usually could not obtain precise information about the smart jammer, they could only get the joint probability distribution of the opponent. Thus, the description of the incomplete information for user $i$ is shown in Assumption 1.

*Assumption 1:* As for user $i$, the jamming gain $\beta_i$ is assumed to be $G$ positive states, which could be denoted as $\beta_i(1), \ldots, \beta_i(g), \ldots, \beta_i(G)$. The probability of the $g$th state is $\sigma_{\beta_i}(g)$, and $\sum_{g=1}^{G} \sigma_{\beta_i}(g) = 1$.

For the smart jammer, when it starts jamming the legitimate communication of users, it has the intention to know the channel gain and the mutual interference state as precisely as possible. In Assumption 2, the description of the incomplete information for the smart jammer is expressed as:

*Assumption 2:* As for the jammer, the user $i$'s transmission gain $\alpha_i$ has $K$ states which are $\alpha_i(1), \ldots, \alpha_i(k), \ldots, \alpha_i(K)$, and the mutual interference gain $\theta_{m,i}$ has $W$ states which are $\theta_{m,i}(1), \ldots, \theta_{m,i}(w), \ldots \theta_{m,i}(W)$. The joint probability is $\rho_J\left(\alpha_i(k), \boldsymbol{\theta}_{\mathbf{m,i}}(\mathbf{w})\right)$, where $\boldsymbol{\theta}_{\mathbf{m,i}}(\mathbf{w}) = (\theta_{1,i}(w), \ldots, \theta_{i-1,i}(w), \theta_{i+1,i}(w), \ldots, \theta_{N,i}(w))$ denotes the $w$th combination of the mutual interference, and $\sum_{k=1}^{K} \sum_{w=1}^{W} (\rho_J(\alpha_i(k), \boldsymbol{\theta}_{\mathbf{m,i}}(\mathbf{w}))) = 1$.

### B. PROBLEM FORMULATION

In UAV communication networks which consist of users as well as a malicious jammer, the competitive interaction and the incomplete information constraint between two sides make the decision more complicated. Usually, the jammer's signal power is bigger than users', and it has the intention to jam the users proactively. Motivated by these, we use a Bayesian Stackelberg game to formulate the anti-jamming system model. Based on the Bayesian Stackelberg game, we propose a one-leader, multi-follower scheme in which the smart jammer acts firstly as leader of the game whereas users act secondarily after obtaining the strategy of the leader. In addition, the incomplete information constraint is also formulated. Mathematically, the Bayesian Stackelberg game is denoted as:

$$\mathcal{G} = \left\{\mathcal{N}, \mathbf{J}, \{\mathcal{P}_i\}_{i\in\mathcal{N}}, \{\mathcal{J}\}, \{\mathcal{U}_i\}_{i\in\mathcal{N}}, \left\{\mathcal{V}_j\right\}\right\}, \tag{6}$$

where $\{\mathcal{J}\}$ and $\{\mathcal{P}_i\}$ denote the jammer's and user $i$'s strategy space set. Specifically, the strategy set of the smart jammer can be denoted as:

$$\mathcal{J} = \{J : 0 \leq J \leq J_{\max}\}. \tag{7}$$

Similarly, the strategy set of user $i$ is given by:

$$\mathcal{P}_i = \{P_i : 0 \leq P_i \leq P_{\max}\}. \tag{8}$$

In equation (7) and equation (8), $J$ and $P_i$ represent transmission power of the smart jammer and user $i$ respectively. $J_{max}$ and $P_{max}$ denote the maximum value of jammer's and user's transmission power. Moreover, $\left\{\mathcal{V}_j\right\}$ and $\{\mathcal{U}_i\}_{i\in\mathcal{N}}$ are utility functions of the jammer or users respectively.

Inspired by [30], considering the transmission payoff which includes the throughput and transmission cost, the user $i$'s utility function can be defined as:

$$\begin{aligned}
&\mathcal{U}_i\left(P_i, \mathbf{P_{-i}}, \tilde{J}\right) \\
&\quad = \sum_{g=1}^{G} Tr_i(\beta_i(g)) - Tc_i \\
&\quad = \sum_{g=1}^{G} \sigma_{\beta_i}(g) \log_2\left(1 + \frac{\alpha_i P_i}{N_0 + \beta_i(g)\tilde{J} + \sum_{m\neq i} P_m\theta_{m,i}}\right) \\
&\qquad - C_u P_i.
\end{aligned} \tag{9}$$

As shown in equation (9), incomplete information and observation error have been considered in users' side, and $\tilde{J}$ denotes the observation value of $J$. When transmission cost is considerably high for UAVs to afford, it is not wise for UAVs to increase the transmission power as much as possible. Making a tradeoff between the throughput and transmission cost is more reasonable.

From the opponent's perspective, the purpose of the jammer is to minimize the communication throughput of UAV group. When it starts jamming, the transmission cost is also a non-ignorable factor. Thus, the smart jammer's utility function is denoted as follows:

$$\begin{aligned}
V_j &= -\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{w=1}^{W} Tr_i\left(\alpha_i(k), \boldsymbol{\theta}_{\mathbf{m,i}}(\mathbf{w})\right) - Tc_j \\
&= -\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{w=1}^{W} \rho_J\left(\alpha_i(k), \boldsymbol{\theta}_{\mathbf{m,i}}(\mathbf{w})\right) \\
&\quad \cdot\log_2(1 + \frac{\alpha_i(k)P_i}{N_0 + \beta_i J + \sum_{m\neq i} P_m\theta_{m,i}(w)}) - C_j J.
\end{aligned} \tag{10}$$

Assuming that random variables $\alpha_i$ and $\boldsymbol{\theta}_{\mathbf{m,i}}$ are independent for the smart jammer. The channel gain of user $i$ has $K$ states $\alpha_i(1), \ldots, \alpha_i(k), \ldots, \alpha_i(K)$, the probabilities of these states are $\sigma_{\alpha_i}(1), \ldots, \sigma_{\alpha_i}(k), \ldots, \sigma_{\alpha_i}(K)$, and $\sum_{k=1}^{K} \sigma_{\alpha_i}(k) = 1$.

Similarly, the mutual interference gain $\boldsymbol{\theta}_{\mathbf{m,i}}$ has $W$ states which are $\boldsymbol{\theta}_{\mathbf{m,i}}(1), \ldots, \boldsymbol{\theta}_{\mathbf{m,i}}(w), \ldots, \boldsymbol{\theta}_{\mathbf{m,i}}(W)$, the probabilities are $\sigma_{\boldsymbol{\theta}_{\mathbf{m,i}}}(1), \ldots, \sigma_{\boldsymbol{\theta}_{\mathbf{m,i}}}(w), \ldots, \sigma_{\boldsymbol{\theta}_{\mathbf{m,i}}}(W)$, and $\sum_{w=1}^{W} \sigma_{\boldsymbol{\theta}_{\mathbf{m,i}}}(w) = 1$.

Considering the independent relationship between $\alpha_i$ and $\boldsymbol{\theta}_{\mathbf{m,i}}$, the smart jammer's utility function is redefined as:

$$\begin{aligned}
&\mathcal{V}_j(J, P_1, \ldots, P_N) \\
&\quad = -\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{w=1}^{W} \sigma_{\alpha_i}(k)\, \sigma_{\boldsymbol{\theta}_{\mathbf{m,i}}}(w) \\
&\qquad \cdot\log_2(1 + \frac{\alpha_i(k)P_i}{N_0 + \beta_i J + \sum_{m\neq i} P_m\theta_{m,i}(w)}) - C_j J.
\end{aligned} \tag{11}$$

The parameters mentioned above have been defined in Section III-A. As for the jammer, it aims to maximize its utility to increase the jamming payoff, which can be formulated as:

$$\max_{J\geq 0} \quad \mathcal{V}_j\left(J, P_1, \ldots, P_N\right). \tag{12}$$

As for the user $i$, the optimization problem is presented as:

$$\max_{P_i\geq 0} \mathcal{U}_i\left(P_i, \mathbf{P_{-i}}, \tilde{J}\right). \tag{13}$$

While introducing the co-channel mutual interference and incomplete information, the process of solving the optimization problem for users and the smart jammer is challenging, which will be shown in the next section.

## IV. ANTI-JAMMING BAYESIAN STACKELBERG GAME WITH INCOMPLETE INFORMATION

In this section, we make the definition of Stackelberg Equilibrium (SE) firstly. Subsequently, the process of the solution will be shown. Moreover, the SBBSIA algorithm is also proposed.

As for the formulated anti-jamming Stackelberg game, the Stackelberg Equilibrium is the best strategy combination for users and the smart jammer. A strategy pair $({P_1}^*, {P_2}^*, \ldots, {P_N}^*, J^*)$ is called SE if it is satisfied that $J$ maximizes the utility of the smart jammer, and for UAVs, $({P_1}^*, {P_2}^*, \ldots, {P_N}^*)$ is the best response to the jammer's strategy. For any $P_i > 0, i \in N, J > 0$, the following conditions should be satisfied:

$$\begin{cases} \mathcal{U}_i\left({P_i}^*, \mathbf{P_{-i}}^*, J^*\right) \geq \mathcal{U}_i\left(P_i, \mathbf{P_{-i}}^*, J^*\right), \\ \mathcal{V}_j\left(J^*, {P_1}^*, \ldots, {P_N}^*\right) \geq \mathcal{V}_j\left(J, {P_1}^*, \ldots, {P_N}^*\right), \end{cases} \tag{14}$$

where $[\cdot]^*$ denotes the best strategy of each decision-maker, $\mathbf{P_{-i}}^*$ denotes that all users choose the best strategies except the $i$th user.

Stackelberg game is a kind of non-cooperative game. While in normal non-cooperative game, Nash equilibrium (NE) is a stable point where no player can improve its utility unilaterally. Moreover, SE can be decomposed into finding the NEs for both users and the jammer [28].

### A. FOLLOWERS SUB-GAME

In this paper, we apply the backward induction method to find the SE, which is drawn from [9] and [37]. Given the jammer's transmission power, each user is going to take the best response independently for the purpose of maximizing its utility. Hence, the users' strategies are firstly studied, and then a non-cooperative game is formulated to describe the competition among users, which is expressed as follows:

$$\mathcal{G}_f = \left\{\mathcal{N}, \{\mathcal{P}_i\}_{i\in\mathcal{N}}, \{\mathcal{U}_i\}_{i\in\mathcal{N}}\right\}. \tag{15}$$

When given the transmission power of the smart jammer and other users except user $i$, the user $i$'s optimization problem is represented as:

$$\begin{aligned} P_i &= \arg\max \mathcal{U}_i\left(P_i, \mathbf{P_{-i}}, \tilde{J}\right), \\ &\quad s.t.\ 0 \leq P_i \leq P_{\max}. \end{aligned} \tag{16}$$

*Theorem 1:* For user $i$, the best strategy is obtained through solving the following equation:

$$\sum_{g=1}^{G} \frac{\sigma_{\beta_i}(g)\,\alpha_i}{\ln 2\left(N_0 + \beta_i(g)\tilde{J} + \sum\limits_{m\neq i} P_m\theta_{m,i} + \alpha_i P_i\right)} - (C_u + \lambda_i) = 0. \tag{17}$$

*Proof:* For user $i$, the utility function is concave while taking the second partial derivative of $P_i$, since:

$$\frac{\partial^2 \mathcal{U}_i}{\partial {P_i}^2} = -\sum_{g=1}^{G} \frac{\sigma_{\beta_i}(g)\,{\alpha_i}^2}{\ln 2\left(N_0 + \beta_i(g)\tilde{J} + \sum\limits_{m\neq i} P_m\theta_{m,i} + \alpha_i P_i\right)^2}. \tag{18}$$

It is obvious that $\partial^2\mathcal{U}_i/\partial {P_i}^2 < 0$ constantly established, and the utility optimization problem of user $i$ can be viewed as a convex problem. Moreover, the user $i$'s utility function is concave so that it can get the only one maximal value when differentiating $\mathcal{U}_i$ in regard to $P_i$ and setting the result equal to 0. Taking nonnegative dual variable $\lambda_i$ into consideration, the Lagrange function of user $i$ is represented as:

$$\begin{aligned} L_i&\left(P_i, \mathbf{P_{-i}}, \tilde{J}\right) \\ &= \sum_{g=1}^{G} \sigma_{\beta_i}(g) \log_2\Big(1 + \frac{\alpha_i P_i}{N_0 + \beta_i(g)\tilde{J} + \sum\limits_{m\neq i} P_m\theta_{m,i}}\Big) \\ &\quad - C_u P_i + \lambda_i\left(P_{\max} - P_i\right). \end{aligned} \tag{19}$$

The Lagrange dual function is shown as follows:

$$D_i(\lambda_i) = \max_{P_i\geq 0} L_i\left(P_i, \mathbf{P_{-i}}, \tilde{J}, \lambda_i\right). \tag{20}$$

Moreover, the dual problem is:

$$d^* = \min_{\lambda_i \geq 0} D_i\left(\lambda_i\right). \tag{21}$$

On the basis of Karush-Kuhn-Tucker (KKT) conditions [8], [42], by setting the resulting derivative equal to 0, which is shown as:

$$\begin{aligned} &\frac{\partial L_i(P_i, \mathbf{P_{-i}}, \tilde{J})}{\partial P_i} \\ &\quad = \sum_{g=1}^{G} \frac{\sigma_{\beta_i}(g)\alpha_i}{\ln 2\left(N_0 + \beta_i(g)\tilde{J} + \sum\limits_{m\neq i} P_m\theta_{m,i} + \alpha_i P_i\right)} \\ &\qquad - C_u - \lambda_i = 0. \end{aligned} \tag{22}$$

When transmission power $\mathbf{P_{-i}}$ and $\tilde{J}$ are known by user $i$, $P_i$ can be obtained from equation (22). ■

The optimization problem for any user is a convex optimization problem, and the utility function holds a strong duality. Moreover, the duality gap is zero. Thus, the optimal solution for the dual problem is the same as initial problem proposed. It is found that if there are more than three users in the UAV group, the optimization problem do not have a analytic solution, and the numerical solution can been obtained instead.

### B. LEADER SUB-GAME

In our work, the smart jammer acts as the leader. Similar to the followers, the leader's game can be formulated as:

$$\mathcal{G}_J = \left\{\mathbf{J}, \mathcal{J}, \mathcal{V}_j\right\}. \tag{23}$$

Usually, the jammer's optimal transmission power is able to be obtained after the following optimization problem being solved:

$$\begin{aligned} J &= \arg\max \mathcal{V}_j\left(P_1(J), \ldots, P_N(J), J\right) \\ &s.t.\ 0 \le J \le J_{\max} \end{aligned} \tag{24}$$

For the smart jammer, a discussion on the users' actions which are learned by the smart jammer is firstly made. As for user $i$, the best response is shown in equation (17). However, as for the smart jammer, the information of itself is completely known, and the observation error can be eliminated as well. Thus, the estimation of the user $i$'s action from the smart jammer's perspective can be denoted as:

$$\frac{\alpha_i}{\ln 2\left(N_0 + \beta_i J + \sum\limits_{m \ne i} P_m \theta_{m,i} + \alpha_i P_i\right)} - (C_u + \lambda_i) = 0. \tag{25}$$

Then an analytic solution of the user $i$ which is obtained in the jammer-side can be expressed as:

$$P_i(J) = \left(\frac{1}{\ln 2 (C_u + \lambda_i)} - \frac{N_0 + \beta_i J + \sum\limits_{m \ne i} P_m \theta_{m,i}}{\alpha_i}\right)^+, \tag{26}$$

where $(\cdot)^+ \triangleq \max(\cdot, 0)$. According to (26), if the transmission power of the jammer is too large, which is shown as follows:

$$J \ge \frac{1}{\beta_i}\left[\frac{\alpha_i}{\ln 2 (C_u + \lambda_i)} - \left(N_0 + \sum_{m \ne i} P_m \theta_{m,i}\right)\right], \tag{27}$$

then from the jammer's perspective, user $i$ will stop transmitting which means $P_i(J) = 0$. Generally, let $\Lambda_i, i \in \mathcal{N}$ denote the $i$th threshold that influence the decision of the smart jammer, which can be expressed as:

$$\begin{aligned} \Lambda_i &= \frac{1}{\beta_i}\left[\frac{\alpha_i}{\ln 2 (C_u + \lambda_i)} - \left(N_0 + \sum_{m \ne i} P_m \theta_{m,i}\right)\right], \quad i \in \mathcal{N}, \\ \Lambda_{\min} &= \min\left(\Lambda_1, \ldots, \Lambda_N\right), \\ \Lambda_{\max} &= \max\left(\Lambda_1, \ldots, \Lambda_N\right). \end{aligned} \tag{28}$$

Moreover, assuming that $\Lambda_{\min} = \Lambda_1 \le \Lambda_2 \le \ldots \le \Lambda_{\max}$, the jammer's utility function can be expressed as follows after substituting equation (26) into equation (11):

$$\begin{aligned} &\mathcal{V}_j\left(J, P_1(J), \ldots, P_N(J)\right) \\ &= \begin{cases} -\sum\limits_{i=1}^{N}\sum\limits_{k=1}^{K}\sum\limits_{w=1}^{W} \sigma_{\alpha_i}(k)\, \sigma_{\theta_{\mathbf{m,i}}}(w) \log_2 \cdot \left(\frac{\alpha_i(k)}{\ln 2\left(N_0 + \beta_i J + \sum\limits_{m \ne i} P_m \theta_{m,i}(w)\right)(C_u + \lambda_i)}\right) \\ \quad -C_j J, \; J \le \Lambda_{\min}; \\ -\sum\limits_{i=i}^{N}\sum\limits_{k=1}^{K}\sum\limits_{w=1}^{W} \sigma_{\alpha_i}(k)\, \sigma_{\theta_{\mathbf{m,i}}}(w) \log_2 \cdot \left(\frac{\alpha_i(k)}{\ln 2\left(N_0 + \beta_i J + \sum\limits_{m \ne i} P_m \theta_{m,i}(w)\right)(C_u + \lambda_i)}\right) \\ \quad -C_j J, \; \Lambda_{i-1} \le J \le \Lambda_i, i \in \mathcal{N}; \\ -C_j J, \; J \ge \Lambda_{\max}. \end{cases} \end{aligned} \tag{29}$$

*Theorem 2:* The optimal transmission power $J^*$ can be expressed as follows:

$$J^* = \begin{cases} 0, & J_{opt1} \le 0, \quad J \le \Lambda_{\min}; \\ J_{opt1}, & 0 \le J_{opt1} \le \Lambda_{\min}, \quad J \le \Lambda_{\min}; \\ \Lambda_{\min}, & J_{opt1} \ge \Lambda_{\min}, \quad J \le \Lambda_{\min}; \\ \Lambda_{i-1}, & J_{opti} \le \Lambda_{i-1}, \quad \Lambda_{i-1} \le J \le \Lambda_i, \; i \in \mathcal{N}; \\ J_{opti}, & \Lambda_{i-1} \le J_{opti} \le \Lambda_i, \; \Lambda_{i-1} \le J \le \Lambda_i, \quad i \in \mathcal{N}; \\ \Lambda_i, & J_{opt1} \ge \Lambda_i, \quad \Lambda_{i-1} \le J \le \Lambda_i, \; i \in \mathcal{N}; \\ \Lambda_{\max}, & J \ge \Lambda_{\max}. \end{cases} \tag{30}$$

*Proof:* As shown in equation (29), the jammer's utility is a linear function in regard to $J$ when $J \ge \Lambda_{\max}$, and it is a concave function with respect to $J$ in interval $J \le \Lambda_{\min}$ and $\Lambda_{i-1} \le J \le \Lambda_i, i \in \mathcal{N}$, since

$$\frac{\partial^2 V_j}{\partial J^2} = \begin{cases} -\sum\limits_{i=1}^{N}\sum\limits_{k=1}^{K}\sum\limits_{w=1}^{W} \frac{\sigma_{\alpha_i}(k)\, \sigma_{\theta_{\mathbf{m,i}}}(w)\, \beta_i^2}{\ln 2\left(N_0 + \beta_i J + \sum\limits_{m \ne i} P_m \theta_{m,i}(w)\right)^2}, \\ \qquad J \le \Lambda_{\min} \\ -\sum\limits_{i=i}^{N}\sum\limits_{k=1}^{K}\sum\limits_{w=1}^{W} \frac{\sigma_{\alpha_i}(k)\, \sigma_{\theta_{\mathbf{m,i}}}(w)\, \beta_i^2}{\ln 2\left(N_0 + \beta_i J + \sum\limits_{m \ne i} P_m \theta_{m,i}(w)\right)^2}, \\ \qquad \Lambda_{i-1} \le J \le \Lambda_i, i \in \mathcal{N}. \end{cases} \tag{31}$$

By introducing a non-negative dual variable $\mu$, the Lagrange function of the smart jammer is:

$$\begin{aligned}&L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)\\&= -\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{w=1}^{W}\sigma_{\alpha_i}\left(k\right)\sigma_{\theta_{\mathbf{m},\mathbf{i}}}\left(w\right)\log_2\\&\cdot\left(\frac{\alpha_i\left(k\right)}{\ln 2\left(N_0+\beta_i J+\sum\limits_{m\neq i}P_m\theta_{m,i}\left(w\right)\right)\left(C_u+\lambda_i\right)}\right)\\&-C_jJ+\mu\left(J_{\max}-J\right).\end{aligned}\tag{32}$$

Then, we differentiate $L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)$ with respect to $J$ and set the result equal to zero, which can be shown as follows:

$$\begin{aligned}&\frac{\partial L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)}{\partial J}\\&=\sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{w=1}^{W}\frac{\sigma_{\alpha_i}\left(k\right)\sigma_{\theta_{\mathbf{m},\mathbf{i}}}\left(w\right)\beta_i}{\ln 2\left(N_0+\beta_i J+\sum\limits_{m\neq i}P_m\theta_{m,i}\left(w\right)\right)}\\&\quad-C_j-\mu=0.\end{aligned}\tag{33}$$

According to equation (33), we can get the numerical solution of $J$ and find the best solution $J_{opt1}$. Similarly, under the condition $\Lambda_{i-1}\leq J\leq\Lambda_i, i\in\mathcal{N}$, the best solution $J_{opti}$ can be obtained as well. Thus, the best strategy $J^*$ can be obtained for different cases as equation (30) shows. ■

### C. OPTIMAL STRATEGIES FOR USERS AND THE JAMMER

From the users' perspective, for different $J^*$, they obtain different $\tilde{J}^*$ respectively. Substituting $\tilde{J}^*$ into (22), and the best solution $P_{opti}$ of utility function $L_i(P_i, \mathbf{P_{-i}}, \tilde{J})$ can be obtained as well. Thus, the best strategies of users and the smart jammer $(P_1{}^*, \ldots, P_i{}^*, \ldots, P_N{}^*, J^*)$ are shown with following cases.

1) $J_{opt1}\leq 0, J\leq\Lambda_{\min}$: In this case, the jammer's utility function $L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)$ can obtain its maximum value at $J^*=0$, under this condition, the optimal value of users can be obtained at $P_i{}^*=P_{opti}, i=1,\ldots,N$.
2) $0\leq J_{opt1}\leq\Lambda_{\min}, J\leq\Lambda_{\min}$: In this case, the jammer's utility function $L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)$ can derive its maximum value at $J^*=J_{opt1}$, under this condition, the optimal value of users can be obtained at $P_i{}^*=P_{opti}, i=1,\ldots,N$.
3) $J_{opt1}\geq\Lambda_{\min}, J\leq\Lambda_{\min}$: In this case, the jammer's utility function $L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)$ can achieve its maximum value at $J^*=\Lambda_{\min}$, under this condition, the optimal value of users can be obtained at $P_i{}^*=P_{opti}, i=1,\ldots,N$.
4) $J_{opti}\leq\Lambda_{i-1}, \Lambda_{i-1}\leq J\leq\Lambda_i, i\in\mathcal{N}$: In this case, the jammer's utility function $L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)$ can obtain its maximum value at $J^*=\Lambda_{i-1}$, under this condition, the optimal value of users can be obtained at $P_x{}^*=P_{optx}, x=1,\ldots,i; P_x{}^*=0, x=i+1,\ldots,N$.
5) $\Lambda_{i-1}\leq J_{opti}\leq\Lambda_i, \Lambda_{i-1}\leq J\leq\Lambda_i, i\in\mathcal{N}$: In this case, the jammer's utility function $L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)$ can obtain its maximum value at $J^*=J_{opti}$, under this condition, the optimal value of users can be derived at $P_x{}^*=P_{optx}, x=1,\ldots,i; P_x{}^*=0, x=i+1,\ldots,N$.
6) $J_{opt1}\geq\Lambda_i, \Lambda_{i-1}\leq J\leq\Lambda_i, i\in\mathcal{N}$: In this case, the jammer's utility function $L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)$ can obtain its maximum value at $J^*=\Lambda_i$, under this condition, the optimal value of users can be obtained at $P_x{}^*=P_{optx}, x=1,\ldots,i; P_x{}^*=0, x=i+1,\ldots,N$.
7) $J\geq\Lambda_{\max}$: In this case, the jammer's utility function $L_j\left(J, P_1\left(J\right), \ldots, P_N\left(J\right)\right)$ can achieve its maximum value at $\Lambda_{\max}$, under this condition, the optimal value of users can be obtained at $P_i{}^*=0, i=1,\ldots,N$.

### D. EXISTENCE AND UNIQUENESS OF STACKELBERG EQUILIBRIUM

The definition of Stackelberg Equilibrium has been given in Section IV. In this part, the demonstration of existence and uniqueness of SE are presented.

*Theorem 3:* The Stackelberg Equilibrium could always exist in the proposed game.

*Proof:* For users, when the jammer's power $J$ is given, the followers sub-game is a non-cooperative game for each user. Moreover, the strategy space of users is a non-empty, compact and convex subset of some Euclidean space. Furthermore, each user's utility function is continuous and concave with respect to its transmission power. According to [43], there exists no less than one NE for each follower when given a value $J$. Let NE $(J)$ denote the best response combination of users when given $J$, and then the SE is expanded as:

$$\mathcal{V}_j\left(J^*, NE\left(J^*\right)\right)\geq\mathcal{V}_j\left(J, NE\left(J\right)\right).\tag{34}$$

Thus, it is easy to prove that there exists $J^*$ which satisfies the following condition:

$$\mathcal{V}_j\left(J^*, NE\left(J^*\right)\right)=\sup_{J\geq 0}\mathcal{V}_j\left(J, NE\left(J\right)\right).\tag{35}$$

The existence of SE has been proved as shown in equation (34) and (35). ■

*Theorem 4:* The Stackelberg Equilibrium is unique in the anti-jamming hierarchical power control game.

*Proof:* Based on equation (18), the second-order derivative of every user, it shows that $\partial^2\mathcal{U}_i/\partial P_i{}^2<0$. Thus, the utility function $\mathcal{U}_i$ of user $i$ is a concave function of $P_i$ when given the observational transmission power $\tilde{J}$. Therefore, based on duality optimization theory [8], [42], for every user in UAV group, the existence of the best response NE $(J^*)$ is unique. In addition, for the smart jammer, an exhaustive analysis has been given in Section IV-B, which shows that the jammer has a unique optimal $J^*$.

**Algorithm 1** Sub-gradient based Bayesian Stackelberg iterative

**1. Initialization**
(1) Initialization of followers:
Transmission power $P_i(t)$, maximum power $P_{max}$, channel gain $\alpha_i$, dual variable $\lambda_i(t)$, transmission cost $C_u$, possible state $\beta_i(g)$, and the probability $\sigma_{\beta_i}(g)$. $i = 1, \ldots, N$; $g = 1, \ldots, G$.
(2) Initialization of leader:
Transmission power $J(t)$, maximum power $J_{max}$, channel gain $\beta_i$, dual variable $\mu_i(t)$, transmission cost $C_j$, possible state $\alpha_i(k)$, and the probability $\sigma_{\alpha_i}(k)$.
Possible mutual interference state $\boldsymbol{\theta}_{\mathbf{m,i}}(\mathbf{w})$, and the probability $\sigma_{\theta_{\mathbf{m,i}}}(W)$,
$i = 1, \ldots, N$; $k = 1, \ldots, K$; $w = 1, \ldots, W$.
(3) Set iteration count $t$=0 and the maximum iteration count $t_{\max}$.
**2. Repeat iterations**
(1) $t = t + 1$.
(2) **for** $i$=1:$N$
(3) Users obtain the observation value $\tilde{J}(t)$ of the jammer.
(4) $P_i{}^*(t+1) = \arg\max L_i\left(P_i, \mathbf{P_{-i}}(t), \tilde{J}(t), \lambda_i(t)\right)$.
(5) The optimal strategies of followers in leader's sight:
$$P_i(J, t+1) = \left(\frac{1}{\ln 2(C_u + \lambda_i(t))} - \frac{N_0 + \beta_i J + \sum\limits_{m \neq i} P_m(t)\theta_{m,i}}{\alpha_i}\right)^+.$$
(6) **end for**
(7) $J^*(t+1) = \arg\max L_j(J, P_i(J, t+1), \ldots, P_N(J, t+1), \mu(t))$.
(8) **for** $i$=1:$N$
(9) $\lambda_i(t+1) = \left[\lambda_i(t) - \Delta_{\lambda_i}^t (P_{\max} - P_i{}^*(t+1))\right]^+$, where $\Delta_{\lambda_i}^t$ is the iteration step of $\lambda_i$.
(10) **end for**
(11) $\mu(t+1) = \left[\mu(t) - \Delta_\mu^t (J_{\max} - J^*(t+1))\right]^+$, where $\Delta_\mu^t$ is the iteration step of $\mu$.
(12) **Until** $t \geq t_{\max}$.
**End iterations**
**3. Output**
(1) **for** $i$=1:$N$
(2) Obtain $P_i{}^*(t_{\max})$.
(3) **end for**
(4) Obtain $J^*(t_{\max})$.
(5) Obtain $\mathcal{U}_i\left(P_i{}^*(t_{\max}), \mathbf{P_{-i}}{}^*(t_{\max}), \tilde{J}^*(t_{\max})\right)$.
(6) Obtain $\mathcal{V}_j(J^*(t_{\max}), P_1{}^*(t_{\max}), \ldots, P_N{}^*(t_{\max}))$.

Thus, the Stackelberg Equilibrium is unique in the proposed game. ∎

### E. OPERATION PROCESS OF THE GAME AND ALGORITHM

In this part, the operation process of the Bayesian Stackelberg game and the SBBSIA algorithm are presented in detail. The operation process is shown in Fig. 2. Moreover, the details of the proposed algorithm are shown in Algorithm 1. The initialization process is implemented when users and the jammer start collecting information which is incomplete but relevant to the opponent. Specifically, for users, they are going to acquire the information about the possible state and corresponding probability of the jammer's channel gain. While for the jammer, it is going to acquire the users' channel information as well as the mutual interference gain. Firstly, the jammer selects its strategy and determines the jamming power. Then, for users, the transmission power is updated until the followers sub-game converges to equilibrium, and the transmission power combination of users composes an NE which maximizes the utility of each user. Finally, both the leader and followers sub-game converge to equilibrium after several iterations, which means the proposed game converges to the SE.

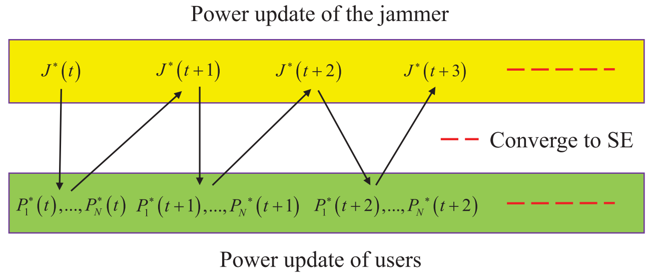



**FIGURE 2.** **Implementation process of the proposed game.**

In addition, the convergence proof of the proposed algorithm is also presented in this section.

*Theorem 5:* The proposed algorithm can converge to the SE.

*Proof:* In this paper, we have given the demonstration of the existence and uniqueness of SE according to Theorem 3 and Theorem 4 in Section IV-D. Moreover, the process of iterative Algorithm 1 is also presented in Section IV-E. In brief, the convergence of Algorithm 1 is supported by the demonstration of the existence and uniqueness of SE, and the operation process of Algorithm 1 is the reflection of obtaining SE for the proposed Bayesian Stackelberg game. In addition, the convergence of the sub-gradient update method is analyzed, which is shown as follows:

$$\lambda_i(t+1) = \left[\lambda_i(t) - \Delta_{\lambda_i}^t\left(P_{\max} - P_i{}^*(t+1)\right)\right]^+, \quad (36)$$

$$\mu(t+1) = \left[\mu(t) - \Delta_\mu^t\left(J_{\max} - J^*(t+1)\right)\right]^+. \quad (37)$$

On the basis of the theoretical analysis in [8], the convergence of the sub-gradient iterative can be guaranteed on condition that the iteration step $\Delta_{\lambda_i}^t$ and $\Delta_\mu^t$ are chosen appropriately.

Combine the analysis on the existence and uniqueness of SE, the operation process of iterative Algorithm 1, and the convergence proof of the sub-gradient update method, we can make the conclusion that the proposed algorithm converges to the SE, and it has also been verified that the Algorithm 1 converges quickly in the simulation parts. ∎

**TABLE 1. Complexity analysis of the proposed algorithm.**

| Computation | Complexity |
|---|---|
| $\tilde{J}(t)$ | $O(N_u C_1)$ |
| $P_i^*(t+1) = \arg\max L_i\left(P_i, \mathbf{P_{-i}}(t), \tilde{J}(t), \lambda_i(t)\right)$ | $O(N_u C_2)$ |
| $P_i(J,t+1) = \left(\frac{1}{\ln 2(C_u+\lambda_i(t))} - \frac{N_0+\beta_i J+\sum_{m\neq i} P_m(t)\theta_{m,i}}{\alpha_i}\right)^+$ | $O(N_u C_3)$ |
| $J^*(t+1) = \arg\max L_j(J, P_i(J,t+1), ..., P_N(J,t+1), \mu(t))$ | $O(C_4)$ |
| $\lambda_i(t+1) = \left[\lambda_i(t) - \Delta_{\lambda_i}^t(P_{\max} - P_i^*(t+1))\right]^+$ <br> $\mu(t+1) = \left[\mu(t) - \Delta_{\mu}^t(J_{\max} - J^*(t+1))\right]^+$ | $O(N_u C_5 + C_5)$ |

### F. COMPLEXITY ANALYSIS OF THE PROPOSED ALGORITHM

In this part, we analyze the computation complexity of the SBBSIA algorithm. Motivated by [44], denote the number of convergence iterations as $N_{it}$, and the number of users as $N$, the computational complexity is shown in Table 1. Moreover, the details are shown as follows:

1) The computation complexity of obtaining the observation value $\tilde{J}(t)$ for every user $i$ is $O(C_1)$, where $C_1$ is a constant determined by the observation process. Then the computation for all users to obtain $\tilde{J}(t)$ is $O(NC_1)$. This part is shown in the Repeat iterations step (3).
2) The computation complexity of computing $P_i^*(t+1) = \arg\max L_i\left(P_i, \mathbf{P_{-i}}(t), \tilde{J}(t), \lambda_i(t)\right)$ for every user $i$ is $O(C_2)$, where $C_2$ is a constant determined by the computing process shown in equation (22). Then the computation for all users to obtain $P_i^*(t+1)$ is $O(NC_2)$. This part is shown in the Repeat iterations step (4).
3) The process of computing the optimal strategies of followers in leader's sight:

$$P_i(J,t+1) = \left(\frac{1}{\ln 2(C_u+\lambda_i(t))} - \frac{N_0+\beta_i J+\sum_{m\neq i} P_m(t)\theta_{m,i}}{\alpha_i}\right)^+. \quad (38)$$

As is shown in the Repeat iterations step (5), for every user $i$, the computation complexity of computing $P_i(J,t+1)$ is $O(C_3)$, where $C_3$ is a small constant. The demonstration is shown in equation (26). Thus, the computation for all users to obtain $P_i(J,t+1)$ is $O(NC_3)$.

4) The computation complexity of computing the jammers best strategy:

$$J^*(t+1) = \arg\max L_j(J, P_i(J,t+1), \ldots, P_N(J,t+1), \mu(t)). \quad (39)$$

As is shown in the Repeat iterations step (7), the computation complexity of computing $J^*(t+1)$ is $O(C_4)$, where $C_4$ is a constant determined by equation (33) in our paper.

5) The part of sub-gradient update:

$$\lambda_i(t+1) = \left[\lambda_i(t) - \Delta_{\lambda_i}^t\left(P_{\max} - P_i^*(t+1)\right)\right]^+, \quad (40)$$
$$\mu(t+1) = \left[\mu(t) - \Delta_{\mu}^t\left(J_{\max} - J^*(t+1)\right)\right]^+. \quad (41)$$

As is shown in the Repeat iterations step (9) and step (11), the computation complexity for every user $i$ to compute $\lambda_i(t+1)$, and for the smart jammer to compute $\mu(t+1)$ is $C_5$, where $C_5$ is a small constant. Thus, the computational complexity for all players to update sub-gradient is $O(NC_5 + C_5)$.

In a word, the total computation complexity of the proposed algorithm can be expressed as:

$$C_{alg} = N_{it}(O(NC_1) + O(NC_2) + O(NC_3) + O(C_4) + O(NC_5 + C_5)) \quad (42)$$

It is shown that the computation complexity is related to the iterations $N_{it}$ and the user number $N$. In our paper, the algorithm can converge quickly with a small $N_{it}$, and the user number $N$ is assumed to be 2. Thus, the proposed Algorithm 1 has relatively low complexity.

## V. SIMULATION RESULTS

In this part, simulations results are presented. The location setting is shown in Fig. 3. Assuming there are four nodes in the UAV group, two of them constitute user 1, whereas the rest constitute user 2. The coordinates of the transmitter

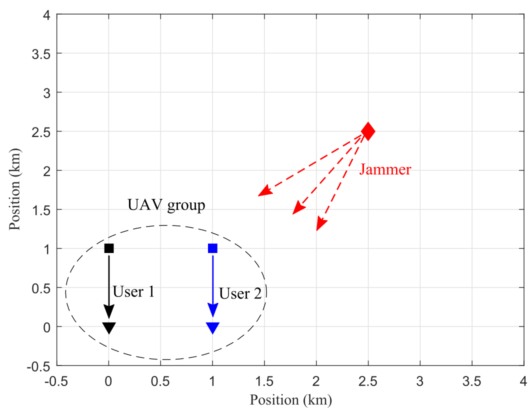



**FIGURE 3.** Locations of the jammer and the UAV group.

and receiver in user 1 are (0km, 1km) and (0km, 0km). The coordinates of the transmitter and receiver in user 2 are (1km, 1km) and (1km, 0km). The jammer is located in (2.5km, 2.5km).

The parameters are given as follow: $\delta_i = \delta_{J,i} = \delta_{m,i} = 2, i = 1, \ldots, N; m = 1, \ldots, i-1, i+1, \ldots, N$. $N_0 = -174$ dBm, $C_u = 0.12$, $C_j = 0.05$, $d_i$, $d_{J,i}$ and $d_{m,i}$ can be calculated through the locations given in Fig. 3. Thus, the values of channel gain, mutual interference gain and jamming gain can be obtained. The maximal transmission power of users is $P_{max} = 10$W, and the maximal transmission power of the jammer is $J_{max} = 100$W. The dual variable $\lambda_i = 5, i = 1, \ldots, N$, $\mu = 10$. The iteration steps are set as 0.2 in all simulations.

Moreover, to describe the incomplete information and observation error more specifically, the fluctuation coefficient $f_u, f_j$ and the observation error coefficient $e_r$ are introduced. From the jammer's perspective, assuming that the channel gain of user $i$ has two states $[\alpha_i, \alpha_i + f_u\alpha_i]$ with probability [0.5, 0.5], and the mutual interference gain of user $i$ also has two combined states $\left[\boldsymbol{\theta}_{\mathbf{m},\mathbf{i}}, \boldsymbol{\theta}_{\mathbf{m},\mathbf{i}} + f_u\boldsymbol{\theta}_{\mathbf{m},\mathbf{i}}\right]$ with the same probability. $f_u$ represents the fluctuation coefficient of users. Whereas from the user $i$'s perspective, the channel gain of the jammer has two states $\left[\beta_i, \beta_i + f_j\beta_i\right]$ with probability [0.5, 0.5], $f_j$ denotes the fluctuation coefficient of the jammer. The observation error coefficient from the followers to the leader can be expressed as:

$$e_r = \frac{\left|\tilde{J} - J\right|}{J}. \tag{43}$$

### A. CONVERGENCE OF THE PROPOSED GAME

The convergence to the SE can be shown in this part for both users and the jammer. As is shown in Fig. 4, there are two users in the UAV group. The iteration number is set to be 30. With the algorithm carrying out, user 1, user 2 and the jammer converge to the only equilibrium value (SE), which is consistent with the theoretical analysis that there exists only one SE in the proposed game. 3-user case and 4-user case in the UAV group have also been simulated, and the analysis can be proved to be analogous to the 2-user case.

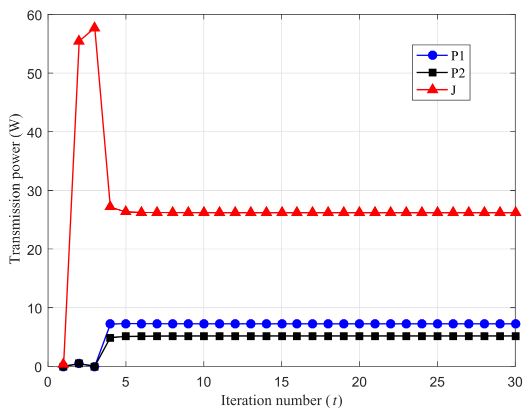



**FIGURE 4.** The convergence of the transmission power.

### B. INFLUENCE OF INCOMPLETE INFORMATION AND OBSERVATION ERROR

The influence of incomplete information and observation error on users are shown in Fig. 5 and Fig. 6. The different values of fluctuation coefficient $f_u$ and observation error coefficient $e_r$ are compared. With the increase of $f_u$, the utility of both user 1 and user 2 increased, while with the increase of $e_r$, the utility of both user 1 and user 2 will decrease instead.

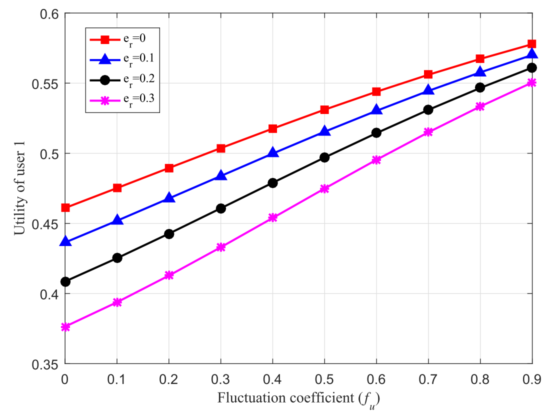



**FIGURE 5.** Influence of incomplete information and observation error on user 1.

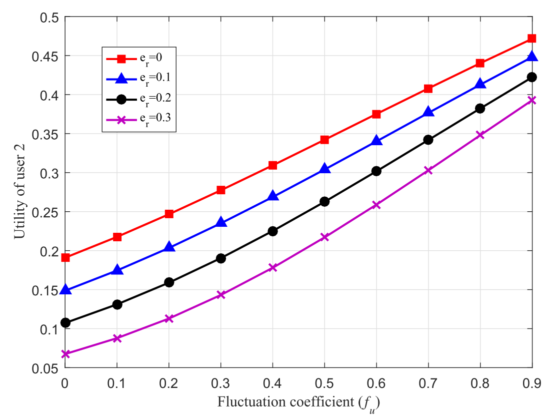



**FIGURE 6.** Influence of incomplete information and observation error on user 2.

As a result, the fluctuation coefficient $f_u$ reflects the accuracy degree of information obtained by the jammer. If $f_u$ is too large, it means the decline of the jamming ability. Observation error coefficient $e_r$ reflects the observation accuracy of user-side. The smaller the observation error is, the more accurate power control strategies of users will be. Thus, observation error on the jammer will influence the utility of users as shown in simulation results.

### C. THE INFLUENCE OF LOCATIONS

Without loss of generality, the flying path should be considered due to the fact that the distance variation between the users and jammer may influence the utility of the user.

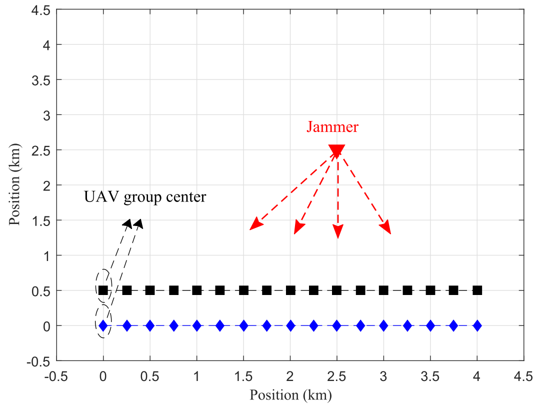



**FIGURE 7.** The path distribution of UAV group.

The mobility of users in the scenario shown in Fig. 3 is not considered, i.e., the UAV can hover motionless in the air. Thus, we evaluate the utility performance depicted in Fig. 7 where there are two paths for the UAV group. Specifically, the jammer is located in (2.5km, 2.5km). The ordinates of the UAV group center in path 1 and 2 are 0.5km and 0km respectively while the abscissa of the UAV group center is changed from 0km to 5km with interval of 0.25km.

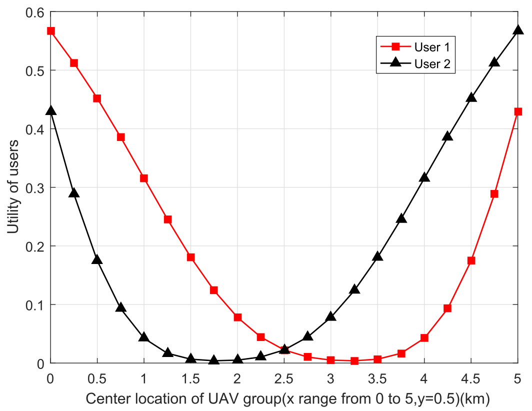



**FIGURE 8.** The utility of users in the first path.

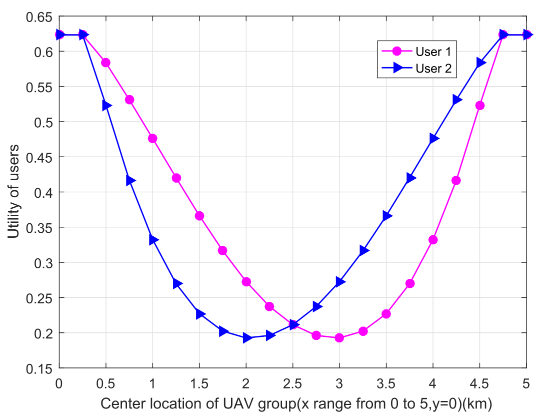



**FIGURE 9.** The utility of users in the second path.

Fig. 8 and Fig. 9 show the variation trends of users' utilities when center location changes with different flying paths. Some significant results can be observed from these two figures: i) There is a downward trend in the utilities of the users when the abscissa of the center location ranges from 0km to 2km. The reason is that the decrease of distance between users and the jammer leads to the enhancement of the jamming ability which results in the reduction of utility and vice versa. ii) When the abscissa of the UAV group center is 2.5km, the utilities of user 1 and 2 are equal since the jamming distance are totally equal. Moreover, in Fig. 9, when the jammer is too far away from UAV group, the best choice for the jammer is adjusting its transmission power equal to zero. Thus, it can be observed that the utility of user 1 and 2 are the same when the abscissa is 0km, 0.25km, 4.75km and 5km, which indicates that the smart jammer is not willing to jam anymore because of the high transmission cost.

## VI. CONCLUSION

In our paper, we mainly focused on anti-jamming transmission problem with incomplete information. A multi-user case was investigated, and co-channel mutual interference was also taken into consideration. A Bayesian-Stackelberg game was proposed in modeling the interactions between UAVs and the smart jammer. Moreover, it was proved that the formulated algorithm can converge to the Stackelberg Equilibrium. An optimal power control scheme for the UAV group was proposed and simulation results could verify the convergence. The influence of incomplete information, observation error together with the locations on user-side were also discussed in the simulation part.

## REFERENCES

[1] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123–1152, 2nd Quart., 2016.

[2] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.

[3] C. Chen, M. Song, C. Xin, and J. Backens, "A game-theoretical anti-jamming scheme for cognitive radio networks," *IEEE Netw.*, vol. 27, no. 3, pp. 22–27, May/Jun. 2013.

[4] L. Zhang, Z. Guan, and T. Melodia, "United against the enemy: Anti-jamming based on cross-layer cooperation in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5733–5747, Aug. 2016.

[5] H. Zhu, C. Fang, Y. Liu, C. Chen, M. Li, and X. S. Shen, "You can jam but you cannot hide: Defending against jamming attacks for geo-location database driven spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2723–2737, Oct. 2016.

[6] S. D'Oro, E. Ekici, and S. Palazzo, "Optimal power allocation and scheduling under jamming attacks," *IEEE/ACM Trans. Netw.*, vol. 25, no. 3, pp. 1310–1323, Jun. 2017.

[7] F. Yao, L. Jia, Y. Sun, Y. Xu, S. Feng, and Y. Zhu, "A hierarchical learning approach to anti-jamming channel selection strategies," *Wireless Netw.*, pp. 1–13, Jul. 2017, doi: 10.1007/s11276-017-1551-9.

[8] G. Ding, Q. Wu, and J. Wang, "Sensing confidence level-based joint spectrum and power allocation in cognitive radio networks," *Wireless Pers. Commun.*, vol. 72, no. 1, pp. 283–298, Jan. 2013.

[9] Y. Xu, G. Ren, J. Chen, L. Jia, and Y. Xu, "Anti-jamming transmission in UAV communication networks: A Stackelberg game approach," in *Proc. IEEE/CIC ICCC*, Oct. 2017, pp. 1–6.

[10] Z. Han *et al.*, *Game Theory Wireless Communication Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[11] Y. Xu, J. Wang, Q. Wu, Z. Du, L. Shen, and A. Anpalagan, "A game-theoretic perspective on self-organizing optimization for cognitive small cells," *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 100–108, Jul. 2015.
[12] Y. Sun, Q. Wu, J. Wang, Y. Xu, and A. Anpalagan, "VERACITY: Overlapping coalition formation-based double auction for heterogeneous demand and spectrum reusability," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2690–2705, Oct. 2016.
[13] H. Shao, Y. Sun, H. Zhao, and W. Zhong, "Locally cooperative traffic-offloading in multi-mode small cell networks: Potential games," in *Proc. WCSP*, Oct. 2015, pp. 1–5.
[14] Y. Sun, J. Wang, F. Sun, and Z. Zhang, "Local altruistic coalition formation game for spectrum sharing and interference management in hyper-dense cloud-RANs," *IET Commun.*, vol. 10, no. 15, pp. 1914–1921, Oct. 2016.
[15] Q. Wu, Y. Xu, J. Wang, L. Shen, J. Zheng, and A. Anpalagan, "Distributed channel selection in time-varying radio environment: Interference mitigation game with uncoupled stochastic learning," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4524–4538, Nov. 2013.
[16] J. Zheng, Y. Cai, Y. Xu, and A. Anpalagan, "Distributed channel selection for interference mitigation in dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4757–4762, Nov. 2014.
[17] J. Zheng, Y. Cai, W. Yang, Y. Wei, and W. Yang, "A fully distributed algorithm for dynamic channel adaptation in canonical communication networks," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 491–494, Oct. 2013.
[18] L. Xiao, *Anti-Jamming Transmissions in Cognitive Radio Networks*. Cham, Germany: Springer, 2015.
[19] L. Xiao, X. Lu, D. Xu, Y. Tang, L. Wang, and W. Zhuang, "UAV relay in VANETs against smart jamming with reinforcement learning," *IEEE Trans. Veh. Technol.*, pp. 1–12, Jan. 2018, doi: 10.1109/TVT.2018.2789466.
[20] E. Altman, K. Avrachenkov, and A. Garnaev, "A jamming game in wireless networks with transmission cost," in *Proc. NET-COOP*, 2007, pp. 1–12.
[21] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, "Anti-jamming games in multichannel cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 1, pp. 4–15, Jan. 2012.
[22] B. Wang, Y. Wu, K. J. R. Liu, and T. C. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 877–889, Apr. 2011.
[23] R. K. Sharma and D. B. Rawat, "Advances on security threats and countermeasures for cognitive radio networks: A survey," *IEEE Commun. Surv. Tuts.*, vol. 17, no. 2, pp. 1023–1043, 2nd Quart., 2015.
[24] M. G. Oskoui, P. Khorramshahi, and J. A. Salehi, "Using game theory to battle jammer in control channels of cognitive radio ad hoc networks," in *Proc. IEEE ICC*, May 2016, pp. 1–5.
[25] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Joint adaptation of frequency hopping and transmission rate for anti-jamming wireless systems," *IEEE Trans. Mobile Comput.*, vol. 15, no. 9, pp. 2247–2259, Sep. 2016.
[26] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Game theoretic anti-jamming dynamic frequency hopping and rate adaptation in wireless systems," in *Proc. WiOpt*, May 2014, pp. 247–254.
[27] M. J. Abdel-Rahman and M. Krunz, "Game-theoretic quorum-based frequency hopping for anti-jamming rendezvous in dsa networks," in *Proc. IEEE DYSPAN*, Apr. 2014, pp. 248–258.
[28] Y. Sun *et al.*, "Energy-aware joint user scheduling and power control for two-tier femtocell networks: A hierarchical game approach," *IEEE Syst. J.*, pp. 1–12, Jul. 2016, doi: 10.1109/JSYST.2016.2580560.
[29] X. Kang, R. Zhang, and M. Motani, "Price-based resource allocation for spectrum-sharing femtocell networks: A Stackelberg game approach," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 538–549, Apr. 2012.
[30] D. Yang, J. Zhang, X. Fang, A. Richa, and G. Xue, "Optimal transmission power control in the presence of a smart jammer," in *Proc. IEEE GLOBECOM*, Dec. 2012, pp. 5506–5511.
[31] D. Yang, G. Xue, J. Zhang, A. Richa, and X. Fang, "Coping with a smart jammer in wireless networks: A Stackelberg game approach," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4038–4047, Aug. 2013.
[32] Y. Li, L. Xiao, J. Liu, and Y. Tang, "Power control Stackelberg game in cooperative anti-jamming communications," in *Proc. GAMENETS*, Nov. 2014, pp. 1–6.
[33] L. Jia, F. Yao, Y. Sun, Y. Xu, S. Feng, and A. Anpalagan, "A hierarchical learning solution for anti-jamming Stackelberg game with discrete power strategies," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 818–821, Aug. 2017.
[34] Y. E. Sagduyu, R. A. Berry, and A. Ephremides, "Jamming games in wireless networks with incomplete information," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 112–118, Aug. 2011.
[35] E. Altman, K. Avrachenkov, and A. Garnaev, "Jamming in wireless networks under uncertainty," in *Proc. WiOpt*, Jun. 2009, pp. 1–7.
[36] L. Xiao, T. Chen, J. Liu, and H. Dai, "Anti-jamming transmission Stackelberg game with observation errors," *IEEE Commun. Lett.*, vol. 19, no. 6, pp. 949–952, Jun. 2015.
[37] L. Jia, F. Yao, Y. Sun, Y. Niu, and Y. Zhu, "Bayesian Stackelberg game for antijamming transmission with incomplete information," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 1991–1994, Oct. 2016.
[38] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[39] R. El-Bardan, S. Brahma, and P. K. Varshney, "Strategic power allocation with incomplete information in the presence of a jammer," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3467–3479, Aug. 2016.
[40] R. El-Bardan, S. Brahma, and P. K. Varshney, "Power control with jammer location uncertainty: A game theoretic perspective," in *Proc. CISS*, Mar. 2014, pp. 1–6.
[41] L. Yu *et al.*, "Power control games for multi-user anti-jamming communications," *Wireless Netw.*, pp. 1–10, Jan. 2018, doi: 10.1007/s11276-018-1664-9.
[42] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[43] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave N-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, Jul. 1965.
[44] Y. Xu, Q. Wu, L. Shen, J. Wang, and A. Anpalagan, "Opportunistic spectrum access with spatial reuse: Graphical game and uncoupled learning solutions," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4814–4826, Oct. 2013.

**YIFAN XU** received the B.S. degree in communication engineering from the Beijing Institute of Technology, Beijing, China, in 2016. He is currently pursuing the M.S. degree with the College of Communication Engineering, Army Engineering University of PLA. His current research interests include game theory, learning theory, and communication anti-jamming technology.

**GUOCHUN REN** received the B.S. degree in wireless communication and the M.S. degree in communications engineering and information system from the Institute of Communications Engineering, Nanjing, China, in 1989 and 1994, respectively. He is currently a Professor with the Institute of Communications Engineering, Army Engineering University of PLA, Nanjing. His current research interests include wireless communications networks and cognitive radio networks.

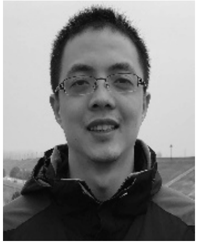
**JIN CHEN** received the B.S. degree in wireless communication, the M.S. degree, and the Ph.D. degree in communications engineering and information system from the Institute of Communications Engineering, Nanjing, China, in 1993, 1996, and 1999, respectively. She is currently a Professor with the Institute of Communications Engineering, Army Engineering University of PLA, Nanjing. Her current research interests include wireless communications networks and cognitive radio networks.

**YUNPENG LUO** received the B.S. degree in communications engineering from the Institute of Communications Engineering, Nanjing, China, in 1989. He is currently an Associate Professor with the Army Engineering University of PLA, China. His current research interests focus on learning theory and game theory for wireless communications.

**LULIANG JIA** received the B.S. degree in communications engineering from Lanzhou Jiaotong University, Lanzhou, China, in 2011, and the M.S. degree in communications and information systems from the College of Communication Engineering, PLA University of Science and Technology, Nanjing, China, in 2014. He is currently pursuing the Ph.D. degree with the College of Communication Engineering, Army Engineering University of PLA. His current research interests include game theory, learning theory, and communication anti-jamming technology.

**XIN LIU** received the B.S. degree in communications engineering, the M.S. and Ph.D. degrees in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, in 2004, 2008, and 2011, respectively. He has been with the College of Information Science and Engineering, Guilin University of Technology, since 2017, where he is currently an Assistant Professor. His research interests focus on deep learning, game theory, and soft-defined radio. He has published several papers in international conferences and reputed journals in his research area.

**YANG YANG** received the B.Eng. and Ph.D. degrees in communication and information engineering from the PLA University of Science and Technology (PLA UST), China, in 2005 and 2011, respectively. Since 2012, she has been a Teacher and a Researcher with the Institute of Communications Engineering, Army Engineering University of PLA. Her current research interests focus on UAV-assisted communication, multi-antenna transceiver optimization, dynamic spectrum sharing and power allocation in cognitive radio, and other wireless communication related topics.



**YUHUA XU** received the B.S. degree in communications engineering and the Ph.D. degree in communications and information systems from the College of Communications Engineering, PLA University of Science and Technology, in 2006 and 2014, respectively. He is currently with the College of Communications Engineering, Army Engineering University of PLA, as an Associate Professor. He has authored several papers in international conferences and reputed journals in his research area. His research interests focus on UAV communication networks, opportunistic spectrum access, learning theory, and distributed optimization techniques for wireless communications.

He received the Certificate of Appreciation as an Exemplary Reviewer for the IEEE Communications Letters in 2011 and 2012, respectively. He was selected to receive the IEEE Signal Processing Society's 2015 Young Author Best Paper Award and the Funds for Distinguished Young Scholars of Jiangsu Province in 2016. He served as an Associate Editor for *Transactions on Emerging Telecommunications Technologies* (Wiley) and *KSII Transactions on Internet and Information Systems*, and the Guest Editor for the Special Issue on The Evolution and the Revolution of 5G Wireless Communication Systems for *IET Communications*.