Evolutionary Cooperative Spectrum Sensing Game: How to Collaborate?

Beibei Wang, *Student Member, IEEE,* K. J. Ray Liu, *Fellow, IEEE,* and T. Charles Clancy, *Member, IEEE*

***Abstract*—Cooperative spectrum sensing has been shown to be able to greatly improve the sensing performance in cognitive radio networks. However, if cognitive users belong to different service providers, they tend to contribute less in sensing in order to increase their own throughput. In this paper, we propose an evolutionary game framework to answer the question of "how to collaborate" in multiuser de-centralized cooperative spectrum sensing, because evolutionary game theory provides an excellent means to address the strategic uncertainty that a user/player may face by exploring different actions, adaptively learning during the strategic interactions, and approaching the best response strategy under changing conditions and environments using replicator dynamics. We derive the behavior dynamics and the evolutionarily stable strategy (ESS) of the secondary users. We then prove that the dynamics converge to the ESS, which renders the possibility of a de-centralized implementation of the proposed sensing game. According to the dynamics, we further develop a distributed learning algorithm so that the secondary users approach the ESS solely based on their own payoff observations. Simulation results show that the average throughput achieved in the proposed cooperative sensing game is higher than the case where secondary users sense the primary user individually without cooperation. The proposed game is demonstrated to converge to the ESS, and achieve a higher system throughput than the fully cooperative scenario, where all users contribute to sensing in every time slot.**

***Index Terms*—Spectrum sensing, cognitive radio networks, game theory, behavior dynamics.**

I. Introduction

With the emergence of new wireless applications and devices, the last decade has witnessed a dramatic increase in the demand for radio spectrum, which has forced government regulatory bodies, such as the Federal Communications Commission (FCC), to review their policies. Since the frequency bands allocated to some licensed spectrum holders experience very low utilization [1], the FCC has been considering opening the under-utilized licensed bands to secondary users on an opportunistic basis with the aid of cognitive radio technology [2]. When the licensed spectrum holders (primary users) are sensed as inactive, the secondary users can operate in the licensed spectrum, if they do not interfere with the primary user.

Since primary users should be carefully protected from interference due to secondary users' operation, spectrum sensing has become an essential function of cognitive radio devices [3]. Recently, cooperative spectrum sensing with relay nodes' help and multi-user collaborative sensing has been shown to greatly improve the sensing performance [4]-[10]. In [4], the authors proposed collaborative spectrum sensing to combat shadowing/fading effects. The work in [5] proposed light-weight cooperation in sensing based on hard decisions to reduce the sensitivity requirements. The authors of [6] showed that cooperative sensing can reduce the detection time of the primary user and increase the overall agility. How to choose proper secondary users for cooperation was investigated in [7]. The authors of [8] studied the design of sensing slot duration to maximize secondary users' throughput under certain constraints. Two energy-based cooperative detection methods using weighted combining were analyzed in [9]. Spatial diversity in multiuser networks to improve spectrum sensing capabilities of centralized cognitive radio networks were exploited in [10].

In most of the existing cooperative spectrum sensing schemes [4]-[10], a fully cooperative scenario is assumed: all secondary users voluntarily contribute to sensing and fuse their detection results in every time slot to a central controller (e.g., secondary base station), which makes a final decision. However, sensing the primary band consumes a certain amount of energy and time which may alternatively be diverted to data transmissions, and it may not be optimal to have all users participate in sensing in every time slot, in order to guarantee a certain system performance. Moreover, with the emerging applications of mobile ad hoc networks envisioned in civilian usage, the secondary users may be selfish and not serve a common goal. If multiple secondary users occupy different sub-bands of one primary user and can overhear the others' sensing outcomes, they tend to take advantage of the others and wait for the others to sense so as to reserve more time for their own data transmission. Therefore, it is of great importance to study the dynamic cooperative behaviors of selfish users in a competing environment while boosting the system performance simultaneously.

In this paper, we model cooperative spectrum sensing as an evolutionary game, where the payoff is defined as the throughput of a secondary user. Evolutionary games have been previously applied to modeling networking problems, such as resource sharing mechanism in P2P networks [11] and congestion control [12] using behavioral experiments. In this paper, we incorporate practical multiuser effect and constraints into the spectrum sensing game. The secondary users want to fulfill a common task, i.e., given a required detection probability to protect the primary user from interference, sense the primary band collaboratively for the sake of getting a high throughput by sharing the sensing cost. The users who do not take part in cooperative sensing can overhear the sensing results and have more time for their own data transmission. However, if no user spends time in sensing the primary user, all of them may get a very low throughput. Therefore, secondary

Part of this work was presented in *IEEE Global Communications Conference (Globecom'08)* [23].

users need to try different strategies at each time slot and learn the best strategy from their strategic interactions using the methodology of understanding-by-building.

In order to study the evolution of secondary users' strategies and answer the question that how to cooperate in the evolutionary spectrum sensing game, we propose to analyze the process of secondary users updating their strategy profile with replicator dynamics equations [16], since a rational player should choose a strategy more often if that strategy brings a relatively higher payoff. We derive the evolutionarily stable strategy (ESS) of the game, and prove the convergence to the ESS through analyzing the users' behavior dynamics. Then we extend our observation to a more general game with heterogeneous users, analyze the properties of the ESSs, and develop a distributed learning algorithm so that the secondary users approach the ESS only with their own payoff history. Simulation results show that as the number of secondary users and the cost of sensing increases, the users tend to have less incentive to contribute to cooperative sensing. However, in general they can still achieve a higher average throughput in the spectrum sensing game than that of the single-user sensing. Furthermore, using the proposed game can achieve a higher total throughput than that of asking all users to contribute to sensing at every time slot.

The rest of this paper is organized as follows. In Section II, we present the system model and formulate the multiuser cooperative spectrum sensing as a game. In Section III, we introduce the background on evolutionary game theory, analyze the behavior dynamics and the ESS of the proposed game, and develop a distributed learning algorithm for ESS. Simulation results are shown in Section IV. Finally, Section V concludes the paper.

II. System Model and Spectrum Sensing Game

A. Hypothesis of Channel Sensing

When a secondary user is sensing the licensed spectrum channel in a cognitive radio network, the received signal $r(t)$ from the detection has two hypotheses when the primary user is present or absent, denoted by H_1 and H_0, respectively. Then, $r(t)$ can be written as

$$r(t) = \begin{cases} hs(t) + w(t), & \text{if } H_1; \\ w(t), & \text{if } H_0. \end{cases} \tag{1}$$

In (1), h is the gain of the channel from the primary user's transmitter to the secondary user's receiver, which is assumed to be slow flat fading; $s(t)$ is the signal of the primary user, which is assumed to be an i.i.d. random process with mean zero and variance σ_s^2; and $w(t)$ is an additive white Gaussian noise (AWGN) with mean zero and variance σ_w^2. Here $s(t)$ and $w(t)$ are assumed to be mutually independent.

Assume we use an energy detector to sense the licensed spectrum, then the test statistics $T(r)$ is defined as

$$T(r) = \frac{1}{N}\sum_{t=1}^{N} |r(t)|^2, \tag{2}$$

where N is the number of collected samples.

The performance of licensed spectrum sensing is characterized by two probabilities. The probability of detection, P_D, represents the probability of detecting the presence of primary user under hypothesis H_1. The probability of false alarm, P_F, represents the probability of detecting the primary user's presence under hypothesis H_0. The higher the P_D, the better protection the primary user will receive; the lower the P_F, the more spectrum access the secondary user will obtain.

If the noise term $w(t)$ is assumed to be circularly symmetric complex Gaussian (CSCG), using central limit theorem the probability density function (PDF) of the test statistics $T(r)$ under H_0 can be approximated by a Gaussian distribution $\mathcal{N}(\sigma_w^2, \frac{1}{N}\sigma_w^4)$ [8]. Then, the probability of false alarm P_F is given by [8][14]

$$P_F(\lambda) = \frac{1}{2}\text{erfc}\left(\left(\frac{\lambda}{\sigma_w^2} - 1\right)\sqrt{\frac{N}{2}}\right), \tag{3}$$

where λ is the threshold of the energy detector, and $\text{erfc}(\cdot)$ denotes the complementary error function, i.e.,

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-t^2} dt.$$

Similarly, if we assume the primary signal is a complex PSK signal, then under hypothesis H_1, the PDF of $T(r)$ can be approximated by a Gaussian distribution $\mathcal{N}((\gamma+1)\sigma_w^2, \frac{1}{N}(2\gamma+1)\sigma_w^4)$ [8], where $\gamma = \frac{|h|^2\sigma_s^2}{\sigma_w^2}$ denotes the received signal-to-noise ratio (SNR) of the primary user under H_1. Then, the probability of detection P_D can be approximated by [8][14]

$$P_D(\lambda) = \frac{1}{2}\text{erfc}\left(\left(\frac{\lambda}{\sigma_w^2} - \gamma - 1\right)\sqrt{\frac{N}{2(2\gamma+1)}}\right). \tag{4}$$

Given a target detection probability $\bar{P}_D$, the threshold λ can be derived, and the probability of false alarm P_F can be further rewritten as

$$P_F(\bar{P}_D, N, \gamma) \triangleq \frac{1}{2}\text{erfc}\left(\sqrt{2\gamma+1}\,\text{erf}^{-1}(1-2\bar{P}_D) + \sqrt{\frac{N}{2}}\gamma\right), \tag{5}$$

where $\text{erf}^{-1}(\cdot)$ denotes the inverse function of the error function $\text{erf}(\cdot)$.

B. Throughput of a Secondary User

When sensing the primary user's activity, a secondary user cannot simultaneously perform data transmission. If we denote the sampling frequency by f_s and the frame duration by T, then the time duration for data transmission is given by $T - \delta(N)$, where $\delta(N) = \frac{N}{f_s}$ represents the time spent in sensing.

When the primary user is absent, in those time slots where no false alarm is generated, the average throughput of a secondary user is

$$R_{H_0}(N) = \frac{T - \delta(N)}{T}(1 - P_F)C_{H_0}, \tag{6}$$

where C_{H_0} represents the data rate of the secondary user under H_0. When the primary user is present, and not detected by the secondary user, the average throughput of a secondary user is

$$R_{H_1}(N) = \frac{T - \delta(N)}{T}(1 - P_D)C_{H_1}, \tag{7}$$

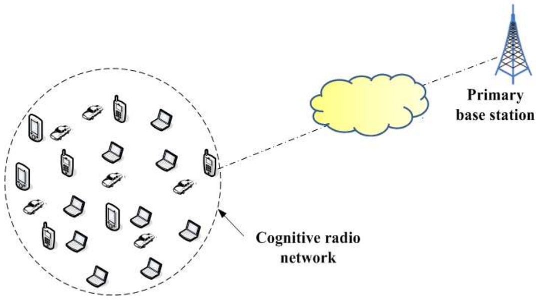


Fig. 1: System model

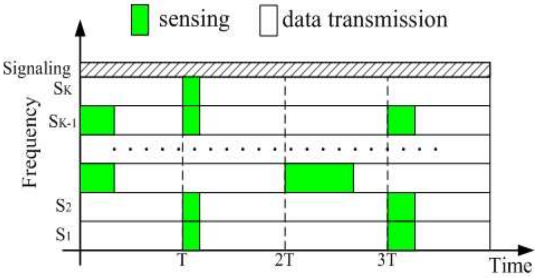


Fig. 2: Cooperative spectrum sensing

where C_{H_1} represents the data rate of the secondary user under H_1.

If we denote P_{H_0} as the probability that the primary user is absent, then the total throughput of a secondary user is

$$R(N) = P_{H_0} R_{H_0}(N) + (1 - P_{H_0}) R_{H_1}(N). \quad (8)$$

In dynamic spectrum access, it is required that the secondary users' operation should not conflict or interfere with the primary users, and P_D should be one in the ideal case. According to (5), however, P_F is then also equal to one, and the total throughput of a secondary user (8) is zero, which is impractical. Hence, a primary user who allows secondary spectrum access usually predetermines a target detection probability $\bar{P}_D$ very close to one [8], under which we assume the secondary spectrum access will be prohibited as a punishment. Then, from the secondary user's perspective, he/she wants to maximize his/her total throughput (8), given that $P_D \geq \bar{P}_D$. Since the target detection probability $\bar{P}_D$ is required by the primary user to be very close to 1, and we usually have $C_{H_1} < C_{H_0}$ due to the interference from the primary user to the secondary user, the second term in (8) is much smaller than the first term and can be omitted. Therefore, (8) can be approximated by

$$\tilde{R}(N) \approx P_{H_0} R_{H_0}(N) = P_{H_0} \frac{T - \delta(N)}{T} (1 - P_F) C_{H_0}. \quad (9)$$

We know from (5) that given a target detection probability $\bar{P}_D$, P_F is a decreasing function of N. As a secondary user reduces N (or $\delta(N)$) in the hope of having more time for data transmission, P_F will increase. This indicates a tradeoff for the secondary user to choose an optimal N that maximizes the throughput $\tilde{R}(N)$. In order to reduce both P_F and N, i.e., keep low false alarm P_F with a smaller N, a good choice for a secondary user is to cooperatively sense the spectrum with the other secondary users in the same licensed band.

C. Spectrum Sensing Game

A diagram of a cognitive radio network where multiple secondary users are allowed to access one licensed spectrum band is shown in Fig. 1, where we assume that the secondary users within each others' transmission range can exchange their sensory data about primary user detection. The cooperative spectrum sensing scheme is illustrated in Fig. 2. We assume that the entire licensed band is divided into K sub-bands, and each secondary user operates exclusively in one of the K sub-bands when the primary user is absent. Transmission time is slotted into intervals of length T. Before each data transmission, the secondary users need to sense the primary user's activity. Since the primary user will operate in all the sub-bands once becoming active, the secondary users within each other's transmission range can jointly sense the primary user's presence, and exchange their sensing results via a narrow-band signalling channel, as shown in Fig 2. In this way, each of them can spend less time detecting while enjoying a low false alarm probability P_F via some decision fusion rule [13], and the spectrum sensing cost ($\delta(N)$) can be shared by whoever is willing to contribute (C).

However, according to their locations and quality of the received primary signal, it may not be optimal to have all secondary users participate in spectrum sensing at every time slot, in order to guarantee certain system performance. Moreover, all secondary users cooperating in sensing may be difficult, if the users do not serve a common authority, and instead act selfishly to maximize their own throughput. In this case, once a secondary user is able to overhear the detection results from the other users, he/she can take advantage of that by refusing to take part in spectrum sensing, called denying (D). Although each secondary user in the cognitive radio network still achieves the same false alarm probability P_F, the users who refuse to join in cooperative sensing have more time for their own data transmission. The secondary users get a very low throughput if no one senses the spectrum, in the hope that someone else does the job.

Therefore, we can model the spectrum sensing as a non-cooperative game. The players of the game are the secondary users, denoted by $\mathcal{S} = \{s_1, \cdots, s_K\}$. Each player s_k has the same action/strategy space, denoted by $\mathcal{A} = \{$C,D$\}$, where "C" represents pure strategy *contribute* and "D" represents pure strategy *refuse to contribute (denying)*. The payoff function is defined as the throughput of the secondary user. Assume that secondary users contributing to cooperative sensing forms a set, denoted by $\mathcal{S}_c = \{s_1, \cdots, s_J\}$. Denote the false alarm probability of the cooperative sensing among set $\mathcal{S}_c$ with fusion rule "RULE" and a target detection probability $\bar{P}_D$ by $P_F^{\mathcal{S}_c} \triangleq P_F(\bar{P}_D, N, \{\gamma_i\}_{i \in \mathcal{S}_c}, \text{RULE})$. Then the payoff for a contributor $s_j \in \mathcal{S}_c$, can be defined as

$$\tilde{U}_{C,s_j} = P_{H_0}\left(1 - \frac{\delta(N)}{|\mathcal{S}_c|T}\right)(1 - P_F^{\mathcal{S}_c})C_{s_j}, \quad \text{if } |\mathcal{S}_c| \in [1, K], \tag{10}$$

where $|\mathcal{S}_c|$, i.e., the cardinality of set $\mathcal{S}_c$, represents the number of contributors, and C_{s_j} is the data rate for user s_j under hypothesis H_0. Therefore, if user s_j chooses to cooperate, then he/she will share the sensing time with the other cooperative users, and the cost is divided equally by all cooperative users. In (10), we assume that the spectrum sensing cost is equally divided among all the contributors; otherwise, there may be fairness issue. The payoff for a user $s_i \notin \mathcal{S}_c$, who selects strategy D, is then given by

$$\tilde{U}_{D,s_i} = P_{H_0}(1 - P_F^{\mathcal{S}_c})C_{s_i}, \quad \text{if } \ |\mathcal{S}_c| \in [1, K-1], \tag{11}$$

since s_i will not spend time sensing. Therefore, if user s_j chooses not to contribute to sensing, he/she will rely on the contributors' decision, have more time for data transmission and can expect a higher throughput. If no secondary user contributes to sensing and waits for the others to sense, i.e., $|\mathcal{S}_c| = 0$, from (5), we know that $\lim_{N\to 0} P_F = 1$, especially for the low received SNR regime and high $\bar{P}_D$ requirement. In this case, the payoff for a denier becomes

$$\tilde{U}_{D,s_i} = 0, \quad \text{if } \ |\mathcal{S}_c| = 0. \tag{12}$$

The decision fusion rule can be selected as the logical-OR rule, logical-AND rule, or majority rule [8]. In this paper, we use the majority rule to derive the $P_F^{\mathcal{S}_c}$, though the other fusion rules could be similarly analyzed. Denote the detection and false alarm probability for a contributor $s_j \in \mathcal{S}_c$ by P_{D,s_j} and P_{F,s_j}, respectively. Then, under the majority rule we have the following

$$P_D = \Pr[\text{at least half users in } \mathcal{S}_c \text{ report } H_1|H_1], \tag{13}$$

and

$$P_F = \Pr[\text{at least half users in } \mathcal{S}_c \text{ report } H_1|H_0], \tag{14}$$

Hence, given a $\bar{P}_D$ for set $\mathcal{S}_c$, each individual user's target detection probability $\bar{P}_{D,s_j}$ can be obtained by solving the following equation

$$\bar{P}_D = \sum_{k=\lceil \frac{1+|\mathcal{S}_c|}{2} \rceil}^{|\mathcal{S}_c|} \binom{|\mathcal{S}_c|}{k} \bar{P}_{D,s_j}^k (1 - \bar{P}_{D,s_j})^{|\mathcal{S}_c|-k}, \tag{15}$$

where we assume each contributor $s_j \in \mathcal{S}_c$ takes equal responsibility in making the final decision for fairness concern and therefore $\bar{P}_{D,s_j}$ is identical for all s_j's. Then, from (5) we can write P_{F,s_j} as

$$P_{F,s_j} = \frac{1}{2}\text{erfc}\left(\sqrt{2\gamma_{s_j} + 1}\,\text{erf}^{-1}(1 - 2\bar{P}_{D,s_j}) + \sqrt{\frac{N}{2|\mathcal{S}_c|}}\gamma_{s_j}\right), \tag{16}$$

and can further obtain $P_F^{\mathcal{S}_c}$ by substituting (16) in (14).

Since secondary users try to maximize their own payoff values, i.e., the average throughput, given the three possible outcomes in (10)-(12), the selfish users' behaviors are highly unpredictable. Contributing to cooperative sensing can provide a stable throughput, however, the stable throughput is achieved at the cost of less time for data transmission; being a free-rider may save more time for useful data transmission, but the secondary users also face the risk of having no one sense the spectrum and get zero throughput. Therefore, *how should a selfish but rational secondary user collaborate with other selfish users in cooperative spectrum sensing?* Always contribute to sensing, or always free ride, or neither? In the next, we will answer this question by analyzing the rational secondary users' behavior dynamics and derive the equilibrium strategy, with the aid of evolutionary game theory.

III. Evolutionary Sensing Game and Strategy Analysis

In this section, we first introduce the concept of evolutionarily stable strategy (ESS), and then use *replicator dynamics* to model and analyze the behavior dynamics of the secondary users in the sensing game.

A. Evolutionarily Stable Strategy

Evolutionary game theory provides a good means to address the strategic uncertainty that a player faces in a game by taking out-of-equilibrium behavior, learning during the strategic interactions, and approaching a robust equilibrium strategy. Such an equilibrium strategy concept widely adopted in evolutionary game theory is the *Evolutionarily Stable Strategy (ESS)* [22], which is "a strategy such that, if all members of the population adopt it, then no mutant strategy could invade the population under the influence of natural selection". Let us define the expected payoff as the individual fitness, and use $\pi(p, \hat{p})$ to denote the payoff of an individual using strategy p against another individual using strategy $\hat{p}$. Then, we have the following formal definition of an ESS [22].

Definition 1 *A strategy p^* is an ESS if and only if, for all $p \neq p^*$,*

1) $\pi(p, p^*) \le \pi(p^*, p^*)$, *(equilibrium condition)*
2) *if $\pi(p, p^*) = \pi(p^*, p^*)$, $\pi(p, p) < \pi(p^*, p)$ (stability condition).*

Condition 1) states that p^* is the best response strategy to itself, and is hence a Nash equilibrium (NE). Condition 2) is interpreted as a stability condition. Suppose that the incumbents play p^*, and a collection of mutants play p. Then conditions 1) and 2) ensure that as long as the fraction of mutants playing p is not too large, the average payoff to p will fall short of that to p^*. Since strategies with a higher fitness value are expected to propagate faster in a population through strategic interactions, evolution will cause the population using mutation strategy p to decrease until the entire population uses strategy p^*.

Since data transmission for each secondary user is continuous, the spectrum sensing game is played repeatedly and evolves over time. Moreover, new secondary users may join in the spectrum sensing game from time to time, and the existing secondary users may even be unaware of their appearance and strategies. Hence, a stable strategy which is robust to mutants using different strategies is especially preferred. Therefore, we propose to use evolutionary game theory [17] to analyze the behavior dynamics of the players and further derive the ESS as the secondary users' optimal collaboration strategy in cooperative spectrum sensing.

B. Evolution Dynamics of the Sensing Game

When a set of rational players are uncertain of each other's actions and utilities, they will try different strategies in every play and learn from the strategic interactions using the methodology of understanding-by-building. During the process, the percentage (or population share) of players using a certain pure strategy may change. Such a population evolution is characterized by *replicator dynamics* in evolutionary game theory. Specifically, consider a population of homogeneous individuals with identical data rate C_{s_i} and received primary SNR γ_i. The players adopt the same set of pure strategies $\mathcal{A}$. Since all players have the same C_{s_i} and γ_i, payoffs for playing a particular strategy depend only on the other strategies employed, not on who is playing them. Therefore, all players have the same payoff function U. At time t, let $p_{a_i}(t) \geq 0$ be the *number* of individuals who are currently using pure strategy $a_i \in \mathcal{A}$, and let $p(t) = \sum_{a_i \in \mathcal{A}} p_{a_i}(t) > 0$ be the total population. Then the associated *population state* is defined as the vector $x(t) = \{x_{a_1}(t), \cdots, x_{|\mathcal{A}|}(t)\}$, where $x_{a_i}(t)$ is defined as the population share $x_{a_i}(t) = p_{a_i}(t)/p(t)$. By replicator dynamics, at time *t* the evolution dynamics of $x_{a_i}(t)$ is given by the following differential equation [17]

$$\dot{x}_{a_i} = \epsilon[\bar{U}(a_i, x_{-a_i}) - \bar{U}(x)]x_{a_i}, \tag{17}$$

where $\bar{U}(a_i, x_{-a_i})$ is the average payoff of the individuals using pure strategy a_i, x_{-a_i} is the set of population shares who use pure strategies other than a_i, $\bar{U}(x)$ is the average payoff of the whole population, and ϵ is some positive number representing the time scale. The intuition behind (17) is as follows: if strategy a_i results in a higher payoff than the average level, the population share using a_i will grow, and the *growth rate* $\dot{x}_{a_i}/x_{a_i}$ is proportional to the difference between strategy a_i's current payoff and the current average payoff in the entire population. By analogy, we can view $x_{a_i}(t)$ as the probability that one player adopts pure strategy a_i, and $x(t)$ can be equivalently viewed as a mixed strategy for that player. If a pure strategy a_i brings a higher payoff than the mixed strategy, strategy a_i will be adopted more frequently, and thus $x_{a_i}(t)$, the probability of taking a_i, will increase. The rate of the probability increase $\dot{x}_{a_i}$ is proportional to the difference between pure strategy a_i's payoff and the payoff achieved by the mixed strategy.

For the spectrum sensing game with heterogeneous players, whose γ_i and/or C_{s_i} are different from each other, denote the probability that user s_j adopts strategy $h \in \mathcal{A}$ at time t by $x_{h,s_j}(t)$, then the time evolution of $x_{h,s_j}(t)$ is governed by the following dynamics equation [17]

$$\dot{x}_{h,s_j} = \left[\bar{U}_{s_j}(h, x_{-s_j}) - \bar{U}_{s_j}(x)\right] x_{h,s_j}, \tag{18}$$

where $\bar{U}_{s_j}(h, x_{-s_j})$ is the average payoff for player s_j using pure strategy h, x_{-s_j} is the set of strategies adopted by players other than s_j, and $\bar{U}_{s_j}(x)$ is s_j's average payoff using mixed strategy x_{s_j}. Eq. (18) indicates that if player s_j achieves a higher payoff by using pure strategy h than using his/her mixed strategy x_{s_j}, strategy h will be adopted more frequently, the probability of using h will increase, and the growth rate of x_{h,s_j} is proportional to the excess of strategy h's payoff and the payoff of the mixed strategy $\bar{U}_{s_j}(x)$.

C. Analysis of Sensing Game with Homogeneous Players

A strategy is ESS if and only if it is asymptotically stable to the replicator dynamics [17][21]. Therefore, we can derive the ESS of the proposed spectrum sensing game by proving its asymptotical stability. In this subsection, we study the ESS of games with homogeneous players, and will discuss the heterogeneous case in the next.

As shown in Fig. 1, players of the sensing game are secondary users within each other's transmission range. If the transmission range is small, we can approximately view that all the received γ_{s_j}'s are very similar to each other. As the γ_{s_j}'s are usually very low, in order to guarantee a low P_F given a target $\bar{P}_D$, the number of sampled signals N should be large. Under these assumptions, we can approximately view $P_F^{\mathcal{S}_c}$ as the same for different $\mathcal{S}_c$'s, denoted by $\hat{P}_F$. Further assume that all users have the same data rate, i.e. $C_{s_i} = C$, for all $s_i \in \mathcal{S}$. Then, the payoff functions defined in (10)-(12) become

$$U_C(J) = U_0\left(1 - \frac{\tau}{J}\right), \quad \text{if } J \in [1, K], \tag{19}$$

and

$$U_D(J) = \begin{cases} U_0, & \text{if } J \in [1, K-1]; \\ 0, & \text{if } J = 0, \end{cases} \tag{20}$$

where $U_0 = P_{H_0}(1 - \hat{P}_F)C$ denotes the throughput achieved by a free rider who relies on the contributors' sensing outcomes, $J = |\mathcal{S}_c|$ denotes the number of contributors, and $\tau = \frac{\delta(N)}{T}$ denotes the fraction of the entire sensing time shared by all contributors over the duration of a time slot. It can be seen from (19) and (20) that, when there is more than one contributor, if a player chooses to contribute to sensing, the payoff $U_C(J)$ is in general smaller than a free-rider's payoff $U_D(J)$, due to the sensing cost $\frac{\tau}{J}$. However, in the worst case, when no one contributes to sensing ($J = 0$), the payoff $U_D(J)$ is the smallest.

As the secondary users are homogeneous players, (17) can be applied to the special case as all players have the same evolution dynamics and equilibrium strategy. Denote x as the probability that one secondary user contributes to spectrum sensing, then the average payoff for pure strategy C can be obtained as

$$\bar{U}_C(x) = \sum_{j=0}^{K-1} \binom{K-1}{j} x^j (1-x)^{K-1-j} U_C(j+1)^1. \tag{21}$$

Similarly, the average payoff for pure strategy D is given by

$$\bar{U}_D(x) = \sum_{j=0}^{K-1} \binom{K-1}{j} x^j (1-x)^{K-1-j} U_D(j). \tag{22}$$

Since the average payoff $\bar{U}(x) = x\bar{U}_C + (1-x)\bar{U}_D$, then (17) becomes

$$\dot{x} = \epsilon x(1-x)\big[\bar{U}_C(x) - \bar{U}_D(x)\big]. \tag{23}$$

In equilibrium x^*, no player will deviate from the optimal strategy, indicating $\dot{x}^* = 0$, and we obtain $x^* = 0$, or 1, or the solution of $\bar{U}^*_C(x) = \bar{U}^*_D(x)$. Then, by equating (21) and (22), we can have the following K^{th}-order equation (see Appendix VI-A)

$$\tau(1-x^*)^K + Kx^*(1-x^*)^{K-1} - \tau = 0, \tag{24}$$

and further solve the remaining equilibrium.

Next we show that the dynamics defined in (17) converge to the above-mentioned equilibriums, which are asymptotically stable and hence the ESS. Note that the variable in (17) is the probability that a user chooses strategy $a_i \in \{C, D\}$, so we need to guarantee that $x_C(t) + x_D(t) = 1$ in the dynamic process. We show this in the following proposition.

Proposition 1 *The sum of the probability that a secondary user chooses strategy "C" and "D" is equal to one in the replicator dynamics of a symmetric sensing game.*

Proof: See Appendix VI-B. ■

In order to prove that the replicator dynamics converge to the equilibrium, we first show that all non-equilibria strategies of the sensing game will be eliminated during the dynamic process. It suffices to prove that (17) is a *myopic adjustment dynamic* [16].

Definition 2 *A system is a* myopic adjustment dynamic *if*

$$\sum_{h\in\mathcal{A}} \bar{U}_{s_j}(h, x_{-s_j})\dot{x}_{h,s_j} \geq 0, \ \forall s_j \in \mathcal{S}. \tag{25}$$

Inequality (25) indicates that the average utility of a player will not decrease in a myopic adjustment dynamic system. We then prove that the dynamics (17) satisfy Definition 2.

Proposition 2 *The replicator dynamics (17) are* myopic adjustment dynamics.

Proof: See Appendix VI-C. ■

In the following theorem, we show that the replicator dynamics in (17) converge to the ESS.

Theorem 1 *Starting from any interior point $x \in (0,1)$, the replicator dynamics defined in (17) converge to the ESS x^*. In specific, when $\tau = 1$, the replicator dynamics converge to $x^* = 0$; when $\tau = 0$, the replicator dynamics converge to $x^* = 1$; when $0 < \tau < 1$, the replicator dynamics converge to the solution of (24).*

Proof: See Appendix VI-D. ■

In practice, the time spent in sensing should be a positive value which is smaller than the duration of a time slot, i.e., we have $0 < \delta(N) < T$ and $0 < \tau = \frac{\delta(N)}{T} < 1$. Therefore, the optimal strategy for the secondary users is to *contribute to sensing with probability x^*, where x^* is the solution of (24).*

[1]Since the average payoff for pure strategy C is the payoff of a player choosing C against another $K-1$ players, who contribute to sensing with probability x, $\bar{U}_C(x)$ can be expressed as $\bar{U}_C(x) = \sum_{j=0}^{K-1} U_C(j+1)\Pr(j)$, where $\Pr(j)$ denotes the probability that there are in total j contributors among $K-1$ other players. Because $\Pr(j) = \binom{K-1}{j}x^j(1-x)^{K-1-j}$, we can obtain $\bar{U}_C(x)$ as shown in (21).

D. Analysis of Sensing Game with Heterogeneous Players

For games with heterogeneous players, it is generally very difficult to represent $\bar{U}_{s_j}(h, x_{-s_j})$ in a compact form, and directly obtain the ESS in closed-form by solving (18). Therefore, we first analyze a two-user game to gain some insight, then generalize the observation to a multi-user game.

*1) **Two-Player Game**:* When there are two secondary users in the cognitive radio network, i.e., $\mathcal{S} = \{s_1, s_2\}$, according to equations (10)-(12) we can write the payoff matrix as in Table I, where for simplicity we define $A \triangleq 1 - P_F^{\mathcal{S}_c}$, with $\mathcal{S}_c = \{s_1, s_2\}$, $B_i \triangleq 1 - P_{F,s_i}$, $D_i \triangleq P_{H_0}C_i$, and $\tau = \frac{\delta(N)}{T}$.

TABLE I: Payoff table of a two-user sensing game

	C	D
C	$D_1A(1-\frac{\tau}{2})$,$D_2A(1-\frac{\tau}{2})$	$D_1B_1(1-\tau)$,D_2B_1
D	D_1B_2,$D_2B_2(1-\tau)$	0,0

Let us denote x_1 and x_2 as the probability that user 1 and user 2 take action "C", respectively, then we have the expected payoff $\bar{U}_{s_1}(C, x_2)$ when user 1 chooses to contribute to sensing as

$$\bar{U}_{s_1}(C, x_2) = D_1A(1-\frac{\tau}{2})x_2 + D_1B_1(1-\tau)(1-x_2), \tag{26}$$

and the expected payoff $\bar{U}_{s_1}(x)$ as

$$\begin{aligned}\bar{U}_{s_1}(x) =& D_1A(1-\frac{\tau}{2})x_1x_2 + D_1B_1(1-\tau)x_1(1-x_2)\\ &+ D_1B_2(1-x_1)x_2.\end{aligned} \tag{27}$$

Thus we get the replicator dynamics equation of user 1 according to (18) as

$$\dot{x}_1 = x_1(1-x_1)D_1\Big[B_1(1-\tau) - E_1x_2\Big], \tag{28}$$

where $E_1 = B_2 + B_1(1-\tau) - A(1-\frac{\tau}{2})$. Similarly the replicator dynamics equation of user 2 is written as

$$\dot{x}_2 = x_2(1-x_2)D_2\Big[B_2(1-\tau) - E_2x_1\Big], \tag{29}$$

where $E_2 = B_1 + B_2(1-\tau) - A(1-\frac{\tau}{2})$.

At equilibrium we know $\dot{x}_1 = 0$ and $\dot{x}_2 = 0$, then from (28) and (29) we get five equilibria: $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$, and the mixed strategy equilibrium $\left(\frac{B_2(1-\tau)}{E_2}, \frac{B_1(1-\tau)}{E_1}\right)$.

According to [20], if an equilibrium of the replicator dynamics equations is a locally asymptotically stable point in a dynamic system, it is an ESS. So we can view (28) and (29) as a nonlinear dynamic system and judge whether the five equilibria are ESSs by analyzing the Jacobian matrix. By taking partial derivatives of (28) and (29), we obtain the Jacobian matrix as

$$J_m = \begin{bmatrix} D_1(1-2x_1)E_{11} & -x_1(1-x_1)D_1E_1 \\ -x_2(1-x_2)D_2E_2 & (1-2x_2)D_2E_{22} \end{bmatrix}, \tag{30}$$

where $E_{11} = B_1(1-\tau) - E_1x_2$, and $E_{22} = B_2(1-\tau) - E_2x_1$.

The asymptotical stability requires that $\det(J_m) > 0$ and $\text{tr}(J_m) < 0$. Substituting the five equilibria to (30), we can obtain the ESS according to different values of A, B_1, and

B_2 and conclude the following optimal collaboration strategy for cooperative sensing game with two heterogeneous players:

1) *When* $A(1-\frac{\tau}{2}) < B_1$*, there is one ESS* $(1,0)$*, and the strategy profile user 1 and user 2 adopt converges to (C,D);*
2) *When* $A(1-\frac{\tau}{2}) < B_2$*, there is one ESS* $(0,1)$*, and the strategy profile converges to (D,C);*
3) *When* $A(1-\frac{\tau}{2}) > B_2$ *and* $A(1-\frac{\tau}{2}) > B_1$*, there is one ESS* $(1,1)$*, and the strategy profiles converges to (C,C);*
4) *When* $A(1-\frac{\tau}{2}) < B_1$ *and* $A(1-\frac{\tau}{2}) < B_2$*, there are two ESSs* $(1,0)$ *and* $(0,1)$*, and the strategy profile converges to (C,D) or (D,C) depending on different initial strategy profiles.*

In order to explain the above-mentioned conclusions and generalize them to a multi-player game, we next analyze the properties of the mixed strategy equilibrium, although it is not an ESS. Let us take the derivative of $x_1^* = \frac{B_2(1-\tau)}{E_2}$ with respect to the performance of a detector (A, B_2) and the sensing cost τ, then we get[2]

$$\frac{\partial x_1^*}{\partial A} = \frac{B_2(1-\tau/2)(1-\tau)}{E_2^2} > 0, \tag{31}$$

$$\frac{\partial x_1^*}{\partial B_2} = \frac{[A(1-\tau/2)-B_1](1-\tau)}{E_2^2} < 0, \tag{32}$$

and

$$\frac{\partial x_1^*}{\partial \tau} = \frac{(A/2-B_1)B_2}{E_2^2} < 0. \tag{33}$$

From (31) we know that when cooperative sensing brings a greater gain, i.e., as A increases, x_1^* (and x_2^*) increases. This is why when $A(1-\frac{\tau}{2}) > B_i$, $i=1,2$, the strategy profile converges to (C,C). From (32) we find that the incentive of a secondary user s_i contributing to cooperative sensing decreases as the other user s_j's detection performance increases. This is because when user s_i learns through repeated interactions that s_j has a better B_j, s_i tends not to sense the spectrum and enjoys a free ride. Then s_j has to sense the spectrum; otherwise, he is at the risk of having no one sense and receiving a very low expected payoff. That is why when $A(1-\frac{\tau}{2}) < B_1$ (or $A(1-\frac{\tau}{2}) < B_2$), the strategy profile converges to (C,D) (or (D,C)). When the sensing cost (τ) becomes higher, the secondary users will be more reluctant to contribute to cooperative sensing and x_1^* decreases, as shown in (33).

*2) **Multi-Player Game**:* From the above-mentioned observation, we can infer that if some user s_i has a better detection performance B_i, the other users tend to take advantage of s_i. If there are more than two users in the sensing game, the strategy of the users with worse B_i's (and γ_i's) will converge to "D". Using replicator dynamics, *users with better detection performance tend to contribute to spectrum sensing (i.e., choose C)*, because they are aware of the low throughput if no one senses the spectrum. Similarly, if the secondary users have different data rates, *the user with a lower rate* C_{s_j} *tends to take advantage of those with higher rates (i.e., they choose D)*, since the latter suffer relatively heavier losses if no one contributes to sensing and they have to become more active in sensing.

The work in [7] discussed how to select a proper subset of secondary users in cooperative sensing so as to optimize detection performance. However, their approach assumes that the information about the received SNR's (γ_i's) is available at the secondary base station. In our evolutionary game framework, the secondary users can learn the ESS by using replicator dynamics only with their own payoff history. Therefore, it is suitable for distributed implementation when there exists no secondary base station and the secondary users behave selfishly. In the next section we propose a distributed learning algorithm and further justify the convergence with computer simulations.

E. Learning Algorithm for ESS

In the above cooperative sensing games with multiple players, we have shown that the ESS is solvable. However, solving the equilibrium requires the knowledge of utility function as well as exchange of private information (e.g., γ_{s_j} and C_{s_j}) and strategies adopted by the other users. This results in a lot of communication overhead. Therefore, a distributed learning algorithm that gradually converges to the ESS without too much information exchange is preferred.

From (18), we can derive the strategy adjustment for the secondary user as follows. Denote the pure strategy taken by user s_j at time t by $A_{s_j}(t)$. Define an indicator function $\mathbf{1}_{s_j}^h(t)$ as

$$\mathbf{1}_{s_j}^h(t) = \begin{cases} 1, & \text{if } A_{s_j}(t) = h; \\ 0, & \text{if } A_{s_j}(t) \neq h. \end{cases} \tag{34}$$

At some interval mT, we can approximate $\bar{U}_{s_j}(h, x_{-s_j})$ by

$$\bar{U}_{s_j}(h, x_{-s_j}) \doteq \frac{\sum_{0\le t\le mT} \tilde{U}_{s_j}(A_{s_j}(t), A_{-s_j}(t))\mathbf{1}_{s_j}^h(t)}{\sum_{0\le t\le mT}\mathbf{1}_{s_j}^h(t)}, \tag{35}$$

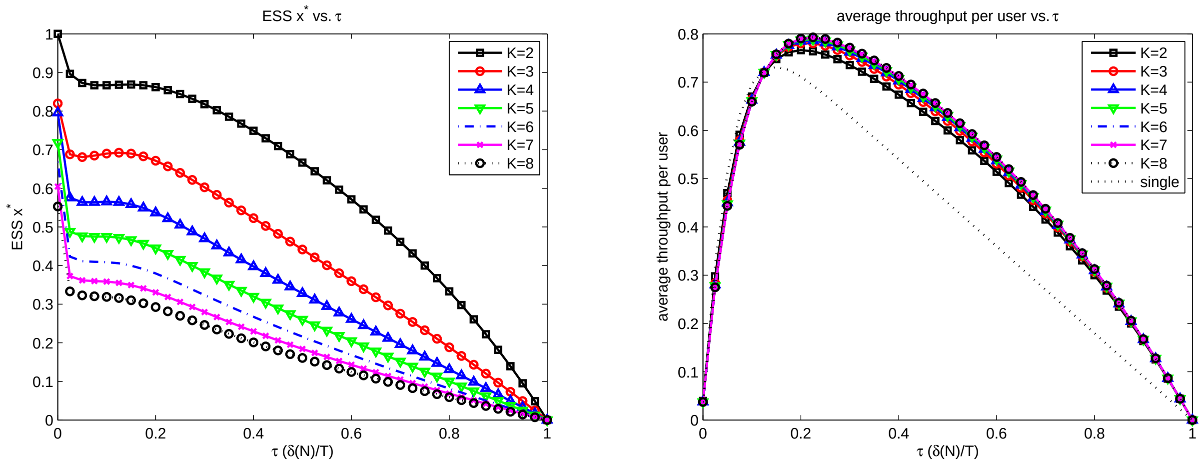
where $\tilde{U}_{s_j}(A_{s_j}(t), A_{-s_j}(t))$ is the payoff value for s_j as determined by (10)-(12). The numerator on the right hand side of (35) denotes the cumulative payoff of user s_j when s_j chooses pure strategy h from time 0 to mT, while the denominator denotes the cumulative total of the number of times when strategy h has been adopted by user s_j during this time period. Hence, (35) can be used to approximate $\bar{U}_{s_j}(h, x_{-s_j})$, and the approximation is more precise as $m \to \infty$. Similarly, $\bar{U}_{s_j}(x)$ can be approximated by the average payoff of user s_j from time 0 to mT

$$\bar{U}_{s_j}(x) \doteq \frac{1}{m}\sum_{0\le t\le mT} \tilde{U}_{s_j}(A_{s_j}(t), A_{-s_j}(t)). \tag{36}$$

Then, the derivative $\dot{x}_{h,s_j}(mT)$ can be approximated by substituting the estimations (35) and (36) into (18). Therefore, the probability of user s_j taking action h can be adjusted by

$$x_{h,s_j}((m+1)T) = x_{h,s_j}(mT) + \eta_{s_j}\left[\bar{U}_{s_j}(h, x_{-s_j}) - \bar{U}_{s_j}(x)\right] x_{h,s_j}(mT), \tag{37}$$

[2] Inequality (32) holds because $A(1-\tau/2) - B_1 < 0$; otherwise $x_1^* = \frac{B_2(1-\tau)}{E_2} > 1$, which is impractical. Inequality (33) holds because in practical applications, we have $P_{F,s_i} < 0.5$, $B_i = 1 - P_{F,s_i} > 0.5$, and $A < 1$; therefore, $\frac{A}{2} < B_i$, and $\frac{\partial x_1^*}{\partial \tau} < 0$.



(a) Probability of being a contributor

(b) Average throughput per user

Fig. 3: ESS and average throughput vs. τ.

TABLE II: Learning algorithm for ESS

1. Initialization: ◇ for $\forall s_j$, choose a proper stepsize η_{s_j}; ◇ for $\forall s_j$, $h \in \mathcal{A}$, let $x(h, s_j) \leftarrow 1/\lvert\mathcal{A}\rvert$.
2. During a period of m slots, in each slot, each user s_j: ◇ chooses an action h with probability $x(h, s_j)$; ◇ receives a payoff determined by (10)-(12); ◇ records the indicator function value by (34). 3. Each user s_j approximates $\bar{U}_{s_j}(h, x_{-s_j})$ and $\bar{U}_{s_j}(x)$ by (35) and (36), respectively. 4. Each user s_j updates the probability of each action by (37).
5. Go to Step 2 until converging to a stable equilibrium.

with η_{s_j} being the step size of adjustment chosen by s_j.

Eq. (37) can be viewed as a discrete-time replicator dynamic system. It has been shown in [18] that if a steady state is hyperbolic and asymptotically stable under the continuous-time dynamics, then it is asymptotically stable for sufficiently small time periods in corresponding discrete-time dynamics. Since the ESS is the asymptotically stable point in the continuous-time replicator dynamics and also hyperbolic [16], if a player knows precise information about $\dot{x}_{h,s_j}$, adapting strategies according to (37) can converge to an ESS. With the learning algorithm, users will try different strategies in every time slot, accumulate information about the average payoff values based on (35) and (36), calculate the probability change of some strategy using (18), and adapt their actions to an equilibrium. The procedures of the proposed learning algorithm are summarized in Table II.

By summarizing the above learning algorithm and analysis in this section, we can arrive at the following cooperation strategy in the de-centralized cooperative spectrum sensing:

Cooperation Strategy in Cooperative Spectrum Sensing: Denote the probability of contributing to sensing for user $s_i \in \mathcal{S}$ by x_{c,s_i}, then the following strategy will be used by s_i:

- if starting with a high x_{c,s_i}, s_i will rely more on the others and reduce x_{c,s_i} until further reduction of x_{c,s_i} decreases his throughput or x_{c,s_i} approaches 0.
- if starting with a low x_{c,s_i}, s_i will gradually increase x_{c,s_i} until further increase of x_{c,s_i} decreases his throughput or x_{c,s_i} approaches 1.
- s_i shall reduce x_{c,s_i} by taking advantage of those users with better detection performance or higher data rates.
- s_i shall increase x_{c,s_i} if cooperation with more users can bring a better detection performance than the case of single-user sensing without cooperation.

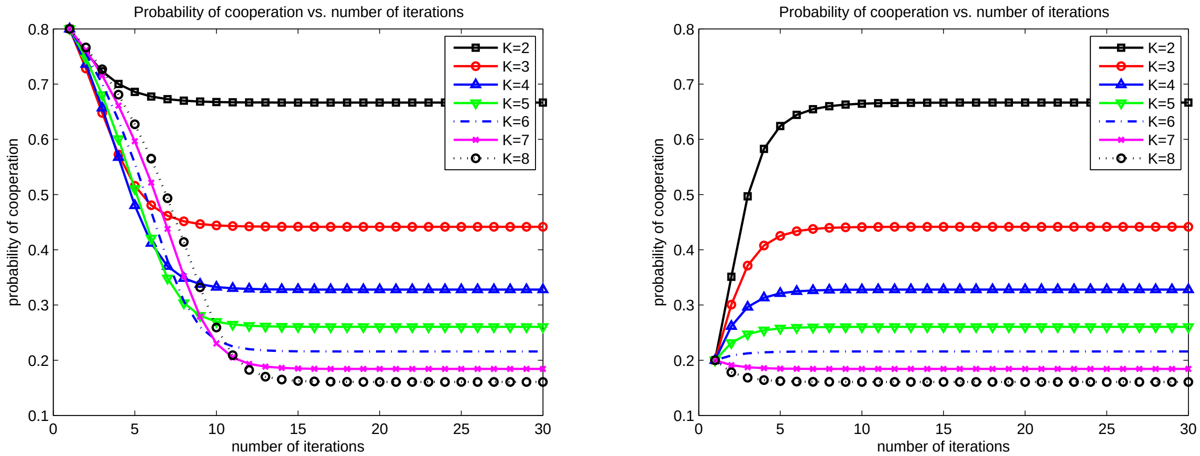
In the next section, we will demonstrate the convergence to ESS of the proposed distributed learning algorithm through simulations.

IV. Simulation Results and Analysis

The parameters used in the simulation are as follows. We assume that the primary signal is a baseband QPSK modulated signal, the sampling frequency is $f_s = 1$MHz, and the frame duration is $T = 20$ ms. The probability that the primary user is inactive is set as $P_{H_0} = 0.9$, and the required target detection probability $\bar{P}_D$ is 0.95. The noise is assumed to be a zero-mean CSCG process. The distance between the cognitive radio network and the primary base station is very large, so the received γ_{s_j}'s are in the low SNR regime, with an average value of -12 dB.

A. Sensing Game with Homogeneous Players

We first illustrate the ESS of the secondary users in a homogeneous K-user sensing game as in Section III-C, where the data rate is $C = 1$ Mbps. In Fig. 3(a), we show the equilibrium probability of being a contributor x^*. The x-axis represents $\tau = \frac{\delta(N)}{T}$, the ratio of sensing time duration over the frame duration. From Fig. 3(a), we can see that x^* decreases as τ increases. For the same τ, x^* decreases as the number of secondary users increases. This indicates that the incentive of contributing to cooperative sensing drops as the cost of sensing increases and more users exist in the network. This is because the players tend to wait for someone else to sense the spectrum and can then enjoy a free ride, when they are faced with a high sensing cost and more counterpart players. In Fig. 3(b), we show the average throughput per user when all users adopt the equilibrium strategy. We see that there is a tradeoff between the cost of sensing and the throughput for an arbitrary



(a) initial $x = 0.8$

(b) initial $x = 0.2$

Fig. 4: Behavior dynamics of a homogeneous K-user sensing game

number of users, and the optimal value of τ is around 0.25. For comparison, we also plot the throughput for a single-user sensing (dotted line "single"), where the optimal value of τ is around 0.15. Although the cost of sensing increases, we see that as more users share the sensing cost, the average throughput per user still increases, and the average throughput values for the cooperative sensing game are higher than that of the single-user sensing case.

B. Convergence of the Dynamics

In Fig. 4, we show the replicator dynamics of the game with homogeneous users, where $\tau = 0.5$. We observe in Fig. 4(a) that starting from a high initial probability of cooperation, all users gradually reduce their degree of cooperation, because being a free-rider more often saves more time for one's own data transmission and brings a higher throughput. However, too low a degree of cooperation greatly increases the chance of having no user contribute to sensing, so the users become more cooperative starting from a low initial probability of cooperation as shown in Fig. 4(b). It takes less than 20 iterations to attain the equilibrium by choosing a proper step size $\eta_{s_i} = 3$.

In Fig. 5, we show the replicator dynamics for the game with three heterogeneous players, using the learning algorithm discussed in Section III-E. We choose $\tau = 0.5$, $\gamma_1 = -14$ dB, $\gamma_2 = -10$ dB, and $\gamma_3 = -10$ dB. As expected, starting from a low initial probability of cooperation, the users tend to increase the degree of cooperation. During the iterations, the users with a worse γ_i (user 1) learn that listening to the detection results from the users with a better γ_i can bring a higher throughput. Hence, user 1's strategy converges to "D" in the long run, while the users with better detection performance (user 2 and user 3) have to sense the spectrum to guarantee their own throughput.

C. Comparison of ESS and Full Cooperation

In Fig. 6, we compare the total throughput of a 3-user sensing game using their ESS and the total throughput when the users always participate in cooperative sensing and share the sensing cost, i.e., $x_{s_i} = 1$. In the first four groups of comparison we assume a homogeneous setting, where γ_i of each user takes value from $\{-13, -14, -15, -16\}$ dB, respectively. In the last four groups, a heterogeneous setting is assumed, where γ_1 equals to $\{-12, -13, -14, -15\}$ dB, respectively, and γ_2 and γ_3 are kept the same as in the homogeneous setting. We find in the figure that using ESS has better performance than all secondary users cooperating in sensing at every time slot. This is because under ESS, the users can take turns to jointly complete the common task, and on average contribute less time to sensing and enjoy a higher throughput. This indicates that in order to guarantee a certain detection performance, it is not necessary to force all users to contribute in every time slot, and ESS can achieve a satisfying system performance even when there exist selfish users.

V. Conclusion

Cooperative spectrum sensing with multiple secondary users has been shown to achieve a better detection performance than single-user sensing without cooperation. However, how to collaborate in cooperative spectrum sensing over de-centralized cognitive radio networks is still an open problem, as selfish users are not willing to contribute their energy/time to sensing. In this paper, we propose an evolutionary game-theoretic framework to develop the best cooperation strategy for cooperative sensing with selfish users. Using replicator dynamics, users can try different strategies and learn a better strategy through strategic interactions. We study the behavior dynamics of secondary users, derive and analyze the property of the ESSs, and further propose a distributed learning algorithm that aids the secondary users approach the ESSs only with their own payoff history. From simulation results we find that the proposed game has a better performance than having all secondary users sense at every time slot, in terms of total throughput. Moreover, the average throughput per user in the sensing game is higher than in the single-user sensing case without user cooperation.

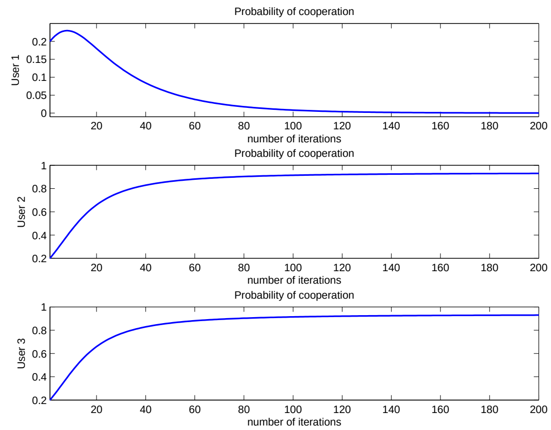


Fig. 5: Behavior dynamics of a heterogeneous 3-user sensing game

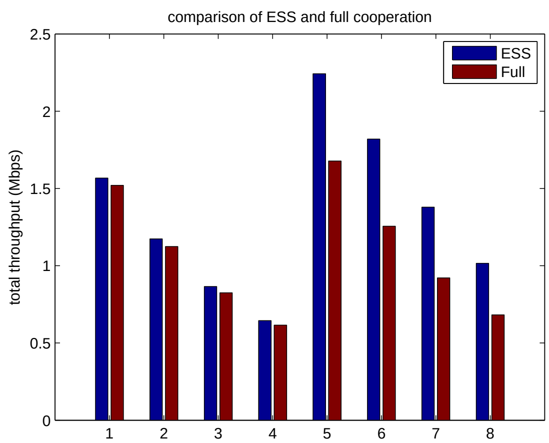


Fig. 6: Comparison of ESS and full cooperation

VI. Appendix

A. Proof of Equation (24)

Subtracting $\bar{U}_D(x)$ from $\bar{U}_C(x)$ we get

$$
\begin{aligned}
&\bar{U}_C(x) - \bar{U}_D(x) \\
=&\sum_{j=0}^{K-1} \binom{K-1}{j} x^j (1-x)^{K-1-j} [U_C(j+1) - U_D(j)] \\
=&\sum_{j=0}^{K-1} \binom{K-1}{j} x^j (1-x)^{K-1-j} [U_0(1-\frac{\tau}{j+1}) - U_0] + M_t \\
=&-U_0\tau \sum_{j=1}^{K-1} \binom{K-1}{j} x^j (1-x)^{K-1-j} \frac{1}{j+1} + M_t \\
=&-\frac{\tau U_0}{xK} \sum_{j=1}^{K-1} \frac{K!}{(j+1)!(K-j-1)!} x^{j+1} (1-x)^{K-j-1} + M_t \\
=&-\frac{\tau U_0}{xK} \sum_{j=2}^{K} \binom{K}{j} x^j (1-x)^{K-j} + M_t \\
=&\frac{\tau U_0}{xK} [(1-x)^K + Kx(1-x)^{K-1} - 1] + M_t \\
=&\frac{U_0}{K} \left[\frac{\tau(1-x)^K + Kx(1-x)^{K-1} - \tau}{x} \right],
\end{aligned}
\tag{38}
$$

with $M_t = (1-x)^{K-1} U_0 (1-\tau)$. By using L'Hôpital's rule, we know that $\lim_{x\to 0} \bar{U}_C(x) - \bar{U}_D(x) = \lim_{x\to 0} \frac{U_0}{K} [-K\tau(1-x)^{K-1} + K(1-x)^{K-1} - Kx(K-1)(1-x)^{K-2}] = U_0(1-\tau) > 0$. Thus, $x = 0$ is not a solution to equation $\bar{U}_C(x) - \bar{U}_D(x) = 0$, and the solution must satisfy $\tau(1-x)^K + Kx(1-x)^{K-1} - \tau = 0$.

B. Proof of Proposition 1

Summing x_{a_i} in (17) over a_i yields

$$
\dot{x}_C + \dot{x}_D = \epsilon[x_C \bar{U}(C, x_D) + x_D \bar{U}(D, x_C) - (x_C + x_D)\bar{U}(x)].
\tag{39}
$$

Since $\bar{U}(x) = x_C \bar{U}(C, x_D) + x_D \bar{U}(D, x_C)$, and initially a user chooses $x_C + x_D = 1$, (39) is reduced to $\dot{x}_C + \dot{x}_D = 0$. Therefore, $x_C(t) + x_D(t) = 1$ holds at any t during the dynamic process. A similar conclusion also holds in an asymmetric game.

C. Proof of Proposition 2

Substituting (17) into (25), we get

$$
\begin{aligned}
&\sum_{a_i \in \mathcal{A}} \dot{x}_{a_i} \bar{U}(a_i, x_{-a_i}) \\
=&\sum_{a_i \in \mathcal{A}} \epsilon \bar{U}(a_i, x_{-a_i}) [\bar{U}(a_i, x_{-a_i}) - \bar{U}(x)] x_{a_i} \\
=&\,\epsilon \sum_{a_i \in \mathcal{A}} x_{a_i} \bar{U}^2(a_i, x_{-a_i}) - \epsilon \Big[\sum_{a_i \in \mathcal{A}} x_{a_i} \bar{U}(a_i, x_{-a_i}) \Big]^2.
\end{aligned}
\tag{40}
$$

According to Jensen's inequality, we know (40) is non-negative, which completes the proof. In addition, we can show (25) also holds for a game with heterogeneous players in a similar way.

D. Proof of Theorem 1

From the simplified dynamics (23), we know that the sign of $\dot{x}_C(t)$ is determined by the sign of $\bar{U}_C(x) - \bar{U}_D(x)$, given $x \in (0,1)$ and $\epsilon > 0$. $\bar{U}_C(x)$ and $\bar{U}_D(x)$ are simplified as the following

$$
\begin{aligned}
\bar{U}_C(x) =& U_0 - U_0(1-x)^{K-1}\tau \\
&- U_0 \sum_{j=1}^{K-1} \binom{K-1}{j} x^j (1-x)^{K-j-1} \frac{\tau}{j+1}, \\
\bar{U}_D(x) =& U_0 - U_0(1-x)^{K-1}.
\end{aligned}
\tag{41}
$$

Furthermore, the difference $\bar{U}_C(x) - \bar{U}_D(x)$ is calculated in Appendix VI-A as

$$
\bar{U}_C(x) - \bar{U}_D(x) = \frac{U_0}{K} \left[\frac{\tau(1-x)^K + Kx(1-x)^{K-1} - \tau}{x} \right].
\tag{42}
$$

According to different values of parameter τ, we prove the theorem in three different cases.

Case I ($\tau = 1$): from (41) we know $\bar{U}_C(x) < \bar{U}_D(x)$, $\frac{dx}{dt} < 0$, and the replicator dynamics converge to $x^* = 0$.

Case II ($\tau = 0$): from (41) we have $\bar{U}_C(x) > \bar{U}_D(x)$, $\frac{dx}{dt} > 0$, and the replicator dynamics converge to $x^* = 1$.

Case III ($0 < \tau < 1$): Define $\Phi(x) = \bar{U}_C(x) - \bar{U}_D(x) = \frac{U_0}{Kx} f(x)$, with $f(x) = \tau(1-x)^K + Kx(1-x)^{K-1} - \tau$.

When $x \rightarrow 0$, using L'Hôpital's rule, we know from (42) that $\lim_{x\rightarrow 0}\Phi(x) = (1-\tau)U_0 > 0$. When $x \rightarrow 1$, $\lim_{x\rightarrow 1}\Phi(x) = -\frac{\tau}{K} < 0$. Since $\Phi(0) > 0$, $\Phi(1) < 0$, and $\Phi(x)$ is a continuous function of x in $(0,1)$, then $\Phi(x)$ must have at least one intersection with the x-axis, i.e., $\exists\tilde{x}$, such that $\Phi(\tilde{x}) = 0$. If there is only one such $\tilde{x}$, then we can infer that $\Phi(x) > 0$ when $x < \tilde{x}$, and $\Phi(x) < 0$ when $x > \tilde{x}$. Since $\Phi(x)$ has the same sign as $f(x)$ when $0 < x < 1$, it suffices to prove that there exists only one solution in $(0,1)$ to equation $f(x) = 0$. Taking derivative of $f(x)$ with respect to x, we get

$$\frac{df(x)}{dx} = (1-x)^{K-2}\big[-(K-\tau)x + (1-\tau)\big]. \quad (43)$$

When $x = \frac{1-\tau}{K-\tau}$, $\frac{df(x)}{dx} = 0$. Observing (43) we find that $f(x)$ is increasing when $0 < x < \frac{1-\tau}{K-\tau}$ with $f(0) = 0$, while decreasing when $\frac{1-\tau}{K-\tau} < x < 1$ with $f(1) = -\tau < 0$. This means equation $f(x) = 0$ has only one root x^* in $(0,1)$, which is the equilibrium solved in (24). When $0 < x < x^*$, $f(x) > 0$; and when $x^* < x < 1$, $f(x) < 0$. Since $\Phi(x)$ has the same sign as $f(x)$, we can conclude that for $0 < x < x^*$, $\Phi(x) > 0$, i.e., $\frac{dx}{dt} > 0$; for $x^* < x < 1$, $\Phi(x) < 0$, i.e., $\frac{dx}{dt} < 0$. Thus, the replicator dynamics converge to the equilibrium x^*.

Therefore, we have proved the convergence of replicator dynamics to the ESS x^*.

REFERENCES

[1] Federal Communications Commission, "Spectrum policy task force report," *FCC Document ET Docket No. 02-155*, Nov. 2002.

[2] J. Mitola III, "Cognitive radio: an integrated agent architecture for software defined radio," Ph.D. Thesis, KTH, Stockholm, Sweden, 2000.

[3] A. Sahai and D. Cabric, "A tutorial on spectrum sensing: fundamental limits and practical challenges," *IEEE DySPAN 2005*, Baltimore, MD, Nov. 2005.

[4] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing in cognitive radio networks," *IEEE DySPAN 2005*, pp. 131-136, Baltimore, MD, Nov. 2005.

[5] S. M. Mishra, A. Sahai, and R. W. Brodensen, "Cooperative sensing among cognitive radios," *IEEE ICC 2006*, pp. 1658-1663, Istanbul, Turkey, June 2006.

[6] G. Ganesan and Y. Li, "Cooperative spectrum sensing in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2204-2222, June 2007.

[7] E. Peh and Y.-C. Liang, "Optimization for cooperative sensing in cognitive radio networks," *IEEE WCNC 2007*, pp.27-32, Hongkong, Mar. 2007.

[8] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE ICC 2007*, pp. 5330-5335, Glasgow, Scotland, June 2007.

[9] F. E. Visser, G. J. Janssen, and P. Pawelczak, "Multinode spectrum sensing based on energy detection for dynamic spectrum access," *IEEE VTC 2008-Spring*, Singapore, May 2008.

[10] G. Ganesan, Y. Li, B. Bing, and S. Li, "Spatiotemporal sensing in cognitive radio networks," *IEEE J. Selected Areas in Communications*, vol. 26, no. 1, pp. 5-12, Jan. 2008.

[11] Q. Zhang, H. Xue, and X. Kou, "An evolutionary game model of resources-sharing mechanism in P2P networks," *Workshop on Intelligent Information Technology Application (IITA) 2007*, pp. 282-285, Dec. 2007.

[12] E. H. Watanabe, D. S. Menasché, E. de Souza e Silva, and R. M. Leão, "Modelling resource sharing dynamics of VoIP users over a WLAN using a game-theoretic approach," *IEEE INFOCOM*, pp. 915-923, Phoenix, AZ, Apr. 2008.

[13] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. on Aerospace and Elect. Syst.*, vol. 22, pp. 98-101, Jan. 1986.

[14] H. V. Poor, *An introduction to signal detection and estimation*, 2nd ed. Springer-Verlag, New York, 1994.

[15] C.-H. Chan, et al., "Evolution of cooperation in well-mixed N-person snowdrift games," *Physica A (2008)*, doi:10.1016/j.physa.2008.01.035.

[16] D. Fudenberg and D. K. Levine, The theory of learning in games, MIT Press, 1998.

[17] J. W. Weibull, Evolutionary game theory, MIT Press, 1995.

[18] M. Hirsch and S. Smale, Differential equations, dynamical systems, and linear algebra, New York: Academic Press, 1974.

[19] D. Fudenberg and D. K. Levine, Game theory, MIT Press, 1991.

[20] R. Cressman, Evolutionary dynamics and extensive form games, MIT Press, 2003.

[21] L. Samuelson, Evolutionary games and equilibrium selection, MIT press, 1998.

[22] J. Maynard Smith, Evolution and the theory of games, Cambridege University Press, 1982.

[23] B. Wang, K. J. R. Liu, and T. C. Clancy, "Evolutionary game framework for behavior dynamics in cooperative spectrum sensing," *IEEE Globecom*, pp. 1-5, New Orleans, LA, Nov. 2008.