# Repeated Open Spectrum Sharing Game with Cheat-Proof Strategies

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

***Abstract*—Dynamic spectrum access has become a promising approach to improve spectrum efficiency by adaptively coordinating different users' access according to spectrum dynamics. However, users who are competing with each other for spectrum may have no incentive to cooperate, and they may even exchange false private information about their channel conditions in order to get more access to the spectrum. In this paper, we propose a repeated spectrum sharing game with cheat-proof strategies. By using the punishment-based repeated game, users get the incentive to share the spectrum in a cooperative way; and through mechanism-design-based and statistics-based approaches, user honesty is further enforced. Specific cooperation rules have been proposed based on the maximum total throughput and proportional fairness criteria. Simulation results show that the proposed scheme can greatly improve the spectrum efficiency by alleviating mutual interference.**

***Index Terms*—Cognitive radio, open spectrum sharing, repeated game, cheat-proof strategy.**

## 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 the government agencies such as Federal Communications Commission (FCC) to review their policies [1][2]. The traditional rigid allocation policies by FCC have severely hindered the efficient utilization of scarce spectrum. Hence, dynamic spectrum access, with the aid of cognitive radio technology [3], has become a promising approach by breaking the paradigm and enabling wireless devices to utilize the spectrum adaptively and efficiently. There are two major classes of dynamic spectrum access: flexible access in the licensed band, and open sharing in the unlicensed band. The industrial, scientific, and medical (ISM) radio band, in which WLAN networks, bluetooth systems, cordless phones, and other novel wireless devices coexist, demonstrates success and importance of open sharing. Nevertheless, unlicensed sharing without regulation usually leads to the overuse of the time/frequency/power units, or the so-called "tragedy of the commons" [4]. To avoid such inefficient usage of the spectrum resources, as suggested in [5], basic open access protocols/etiquettes have to be set by either government or industry standardization, and the right to use the spectrum should be only shared among those wireless users who conform with the unlicensed protocols/etiquettes.

In order to fully utilize the limited spectrum resources, efficiently and fairly sharing the spectrum among multiple users becomes an important issue. There are some previous efforts addressing this issue. The work in [6] examined the secondary user access patterns to propose a feasible spectrum sharing scheme. The authors in [7] proposed a primary prioritized Markov dynamic spectrum access scheme to optimally coordinate secondary users' spectrum access and achieve a good statistical tradeoff between efficiency and fairness. In [8], open access of a group of links was scheduled in a centralized way by a spectrum server. However, since multiple users compete for the spectrum resources, they may have no incentive to cooperate with each other, and can behave selfishly. Therefore, game theory [9] is a proper and flexible tool to analyze the interactions among selfish users [10]. In [11], game theory was employed to resolve channel conflicts distributively by associating the Nash equilibrium with a maximal coloring problem for spectrum sharing. A local bargaining mechanism was proposed in [12] to distributively optimize the efficiency of spectrum allocation and maintain bargaining fairness among secondary users. Auctions were proposed for sharing spectrum among multiple users such that the interference was below a certain level [13]. A real-time spectrum auction framework was proposed in [14] to achieve a conflict-free allocation which maximizes auction revenue and spectrum utilization. Based on double-auction rules, belief-based dynamic pricing approaches were developed in [15] and [16] to optimize the overall spectrum efficiency while keeping the participating incentives of the selfish users. In [17], the authors proposed a repeated game approach to increase efficiency when multiple primary users sell their bands. The repeated game was also employed to model the open spectrum sharing problem in [18] with the assumption that the channels are time-invariant. In [19], iterative waterfilling power allocation was proposed for Gaussian interference channels with frequency-selective fading, and convergence was discussed in [20] under a more general assumption that users updated the powers in an asynchronous way. Some practical difficulties of the iterative waterfilling method were circumvented in [21] by exchanging "interference price" which took mutual interference into

Manuscript received February 8, 2008; revised June 23, 2008 and July 23, 2008; accepted July 29, 2008. The associate editor coordinating the review of this paper and approving it for publication was Prof. E. Hossain.

Y. Wu, B. Wang, and K. J. R. Liu are with the Department of Electrical and Computer Engineering, and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA (e-mail: {wuyl, bebewang, kjrliu}@umd.edu).

T. C. Clancy is with Laboratory for Telecommunications Sciences, US Department of Defense, College Park, MD 20740, USA (e-mail: clancy@LTSnet.net).

This paper was presented in part at the IEEE International Conference on Communications (ICC), Beijing, May 2008.

Digital Object Identifier 10.1109/TWC.2008.080182.

consideration.

Although existing dynamic spectrum access schemes based on game theory have successfully enhanced spectrum efficiency, in order to achieve more flexible spectrum access in long-run scenarios, some basic questions still remain unanswered. First, the spectrum environment is constantly changing and there is no central authority to coordinate the spectrum access of different users. Thus, the spectrum access scheme should be able to distributively adapt to the spectrum dynamics, e.g., channel variations, with only local observations. Moreover, users competing for the open spectrum may have no incentive to cooperate with each other, and they may even exchange false private information about their channel conditions in order to get more access to the spectrum. Therefore, cheat-proof spectrum sharing schemes should be developed to maintain the efficiency of the spectrum usage.

Motivated by the preceding, in this paper we propose a cheat-proof etiquette for unlicensed spectrum sharing by modeling the distributed spectrum access as a repeated game. In the proposed game, punishment will be triggered if any user deviates from cooperation, and hence users are enforced to access the spectrum cooperatively. We propose two sharing rules based on the maximum total throughput and proportional fairness criteria, respectively; accordingly, two cheat-proof strategies are developed: one provides players with the incentive to be honest based on mechanism design theory [22], and the other makes cheating nearly unprofitable by statistical approaches. Therefore, the competing users are enforced to cooperate with each other honestly. The simulation results show that the proposed scheme can greatly improve the spectrum efficiency under mutual interference.

The remainder of this paper is organized as follows. In Section II, the system model for open spectrum sharing is described. In Section III, we develop a punishment-based repeated game for open spectrum sharing. The specific design of cooperation rules and misbehavior detection are discussed in Section IV. In Section V, we develop two cheat-proof strategies for the proposed spectrum sharing rules. Simulation results are shown in Section VI, and Section VII concludes this paper.

## II. System Model

We consider a situation where $K$ groups of unlicensed users coexist in the same area and compete for the same unlicensed spectrum band, as shown in Fig. 1. The users within the same group attempt to communicate with each other, whose usage of the spectrum will introduce interference to other groups. For simplicity, we assume that each group consists of a single transmitter-receiver pair, and that all the pairs are fully loaded, i.e., they always have data to transmit. At time slot $n$, all pairs are trying to occupy the spectrum, and the received signal at the $i$-th receiver $y_i[n]$ can be expressed as

$$y_i[n] = \sum_{j=1}^{K} h_{ji}[n]x_j[n] + w_i[n],\ i = 1, 2, \dots, K, \quad (1)$$

where $x_j[n]$ is the transmitted information of the $j$-th pair, $h_{ji}[n] (j = 1, 2, \dots, K; i = 1, 2, \dots, K)$ represents the channel gain from the $j$-th transmitter to the $i$-th receiver, and

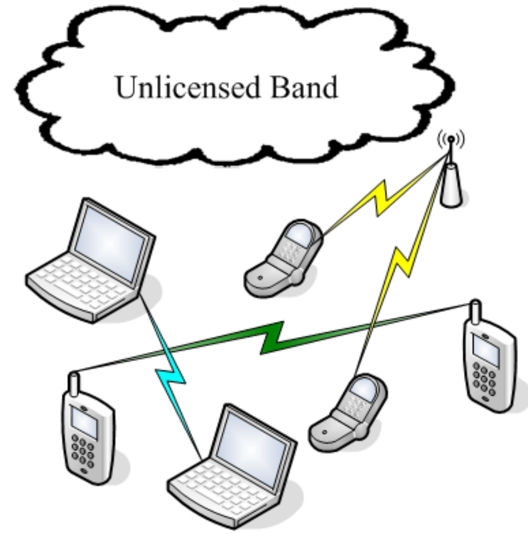


Fig. 1. Illustration of open spectrum sharing.

$w_i[n]$ is the white noise at the $i$-th receiver. In the rest of the paper, the time index $n$ will be omitted wherever no ambiguity is caused. We assume the channels are Rayleigh fading, i.e., $h_{ji} \sim \mathcal{CN}(0, \sigma_{ji}^2)$, and distinct $h_{ji}$'s are statistically independent. The channels are assumed to remain constant during one time slot, and change independently from slot to slot. The noise is independently identically distributed (i.i.d.) with $w_i \sim \mathcal{CN}(0, N_0)$, where $N_0$ is the noise power. Limited by the instrumental capability, the transmission power of the $i$-th user cannot exceed his/her own peak power constraint $P_i^M$, i.e., $|x_i[n]|^2 \le P_i^M$ at any time slot $n$.

Usually, there is no powerful central unit to coordinate the spectrum access in the unlicensed band, and different coexisting systems do not share a common goal to help each other voluntarily. It is reasonable to assume that each transmitter-receiver pair is selfish: pursuing higher self-interest is the only goal for the wireless users. Such selfish behaviors can be well analyzed by game theory. Therefore, we propose to model the spectrum sharing game as follows:

- Players: the $K$ transmitter-receiver pairs,
- Actions: each player can choose the transmission power level $p_i$ in $[0, P_i^M]$,
- Payoffs: $R_i(p_1, p_2, \dots, p_K)$, the gain of transmission achieved by the $i$-th player after power levels $p_1, p_2, \dots, p_K$ have been chosen by individual players.

In general, the gain of transmission is a non-negative increasing function of data throughput. For simplicity, we assume that all the players share the same valuation model that the gain of transmission equals data throughput. The results can be easily extended to cases with different valuation models. The averaged payoff of the $i$-th player can be approximated by

$$R_i(p_1, p_2, \dots, p_K) = \log_2\left(1 + \frac{p_i|h_{ii}|^2}{N_0 + \sum_{j\ne i} p_j|h_{ji}|^2}\right), \quad (2)$$

when mutual interference is treated as Gaussian noise, e.g., when the code division multiple access (CDMA) technique is employed.

## III. Repeated Spectrum Sharing Game

In this section, we find the equilibria of the proposed spectrum sharing game. We assume that all the players are selfish and none is malicious. In other words, players aim to maximize their own interest, but will not jeopardize others or even the entire system at their own cost. Because all the selfish players try to access the unlicensed spectrum as much as possible, severe competition often leads to strong mutual interference and low spectrum efficiency. However, since wireless systems coexist over a long period of time, the spectrum sharing game will be played for multiple times, in which the undue competition could be resolved by mutual trust and cooperation. We propose a punishment-based repeated game to boost cooperation among competing players.

### A. One-shot game

First, we look into the one-shot game where players are myopic and only care for the current payoff. The vector of power levels $(p_1^*, p_2^*, \ldots, p_K^*)$ is called a Nash equilibrium [9] if and only if for all $i = 1, 2, \ldots, K$ and all possible power level choices $p_i' \in [0, P_i^M]$,

$$R_i(p_1^*, p_2^*, \ldots, p_i^*, \ldots, p_K^*) \geq R_i(p_1^*, p_2^*, \ldots, p_i', \ldots, p_K^*) \quad (3)$$

always holds. The Nash equilibrium, from which no individual would have the incentive to deviate, provides a stable point in which the system resides. For this one-shot spectrum sharing game, the equilibrium occurs when every pair transmits at the highest power level, as shown in the following proposition.

*Proposition 1:* The only Nash equilibrium for this one-shot game is $\left(P_1^M, P_2^M, \ldots, P_K^M\right)$.

*Proof:* First, we show that $\left(P_1^M, P_2^M, \ldots, P_K^M\right)$ is a Nash equilibrium. According to the definition of the payoff (2), when $p_1, p_2, \ldots, p_{i-1}, p_{i+1}, \ldots, p_K$ are fixed, and hence the interference power is fixed, the $i$-th player's payoff $R_i(p_1, p_2, \ldots, p_K)$ grows as the power level $p_i$ increases. Therefore, for any player $i$, deviating from $P_i^M$ to any lower value will decrease the payoff, which makes $\left(P_1^M, P_2^M, \ldots, P_K^M\right)$ a Nash equilibrium.

Then, we show by contradiction that no other equilibria exist. Assume that $(p_1^*, p_2^*, \ldots, p_K^*)$ is any equilibrium other than $\left(P_1^M, P_2^M, \ldots, P_K^M\right)$, which means at least one entry is different, say $p_i^* \neq P_i^M$. However, this player can always get better off by deviating from $p_i^*$ to $P_i^M$, which violates the definition of a Nash equilibrium. ■

When channel states are fixed, substituting the equilibrium strategy $p_i = P_i^M$ for all $i$ into (2) yields

$$R_i^S(h_{1i}, h_{2i}, \ldots, h_{Ki}) = \log_2\left(1 + \frac{P_i^M |h_{ii}|^2}{N_0 + \sum_{j \neq i} P_j^M |h_{ji}|^2}\right), \quad (4)$$

where the superscript '$S$' stands for "selfish". This is indeed the only possible outcome of the one-shot game with selfish players. Furthermore, when channel fading is taken into account, the expected payoff can be calculated by averaging over all channel realizations,

$$r_i^S = E_{\{h_{ji}, j=1,\ldots,K\}}\left[R_i^S(h_{1i}, h_{2i}, \ldots, h_{Ki})\right]. \quad (5)$$

In this paper, the payoff represented by the upper-case letter is the utility under a specific channel realization, whereas the payoff in the lower-case letter is the utility averaged over all channel realizations.

Proposition 1 implies that the common open spectrum is excessively exploited owing to lack of cooperation among the selfish players. In order to maximize their own profit, all the players always occupy the spectrum with maximum transmission power, which, in turn, makes everyone suffer from strong mutual interference. If the players can somehow share the spectrum in a more cooperative and regulated fashion, everyone will get better off because interference has been greatly reduced. Since spectrum sharing lasts over quite a long period of time, it can be seen as a game played for numerous rounds, in which cooperation is made possible by established individual reputation and mutual trust.

### B. Repeated game

In open spectrum sharing, players cannot be "forced" to cooperate with each other; instead, they must be self-enforced to participate in cooperation. We propose a punishment-based repeated game to provide players with the incentive for cooperation.

First of all, we have to define the payoff for the repeated game. Since players view the multiple rounds in the game as a whole, the payoff of a repeated game is defined as the sum of payoffs discounted over time [9],

$$U_i = (1 - \delta) \sum_{n=0}^{+\infty} \delta^n R_i[n], \quad (6)$$

where $R_i[n]$ is player $i$'s payoff at the $n$-th time slot, and $\delta\,(0 < \delta < 1)$ is the discount factor. When $\delta$ is closer to 1, the player is more patient. Because players value not only the current payoff but also the rewards in the future, they have to constrain behavior in the present to keep a good credit history; otherwise, a bad reputation may cost even more in the future.

In general, if players do not cooperate with each other, the only reasonable choice is the one-shot game Nash equilibrium with the expected payoff $r_i^S$ given in (5). However, if all the players follow some predetermined rules to share the spectrum, higher expected one-slot payoffs $r_i^C$ ('$C$' stands for "cooperation") may be achieved, i.e., $r_i^C > r_i^S$ for $i = 1, 2, \ldots, K$. For example, the cooperation rule may require only several players access the spectrum simultaneously, and hence mutual interference is greatly reduced. Nevertheless, without any commitment, selfish players always want to deviate from cooperation. One player can take advantage of others by transmitting in the time slots which he/she is not supposed to, and the instantaneous payoff at one specific slot is a random variable denoted by $R_i^D$ ('$D$' stands for "deviation").

Although it is not a stable equilibrium in the one-shot game, cooperation is an equilibrium in the repeated game enforced by the threat of punishment. Specifically, every player states the threat to others: if anyone deviates from cooperation, there will be no more cooperation forever. Such a threat, also known as the "trigger" punishment [9], deters deviation and helps maintain cooperation. For example, assume that player $i$ hesitates whether to deviate or not. Denote the discounted payoff with deviation as $U_i^D$, and that without deviation as $U_i^C$. As shown by the following proposition, the payoffs strongly converge to constants regardless of the channel realizations. Then, for the sake of the player's own benefit, it is better not to deviate as long as $r_i^C > r_i^S$.

*Proposition 2:* As $\delta \to 1$, $U_i^D$ converges to $r_i^S$ almost surely, and $U_i^C$ converges to $r_i^C$ almost surely.

*Proof:* First, we show that as $\delta \to 1$, the discounted payoff defined in (6) is asymptotically equivalent to the average of the one-time payoffs. By switching the order of operations, we have

$$\begin{aligned}\lim_{\delta\to 1} U_i &= \lim_{\delta\to 1}\lim_{N\to+\infty}\frac{1-\delta}{1-\delta^{N+1}}\sum_{n=0}^{N}\delta^n R_i[n]\\ &= \lim_{N\to+\infty}\sum_{n=0}^{N}\left(\lim_{\delta\to 1}\frac{\delta^n-\delta^{n+1}}{1-\delta^{N+1}}\right)R_i[n] \qquad (7)\\ &= \lim_{N\to+\infty}\frac{1}{N+1}\sum_{n=0}^{N}R_i[n],\end{aligned}$$

where the last equality holds according to L'hospital's rule.

Assume that player $i$ deviates at time slot $T_0$. Then, the payoffs $\{R_i[n],\, n = 0,1,\ldots,T_0-1\}$ are i.i.d. random variables with mean $r_i^C$, whose randomness comes from the i.i.d. channel variations. Similarly, the payoffs $\{R_i[n],\, n = T_0+1, T_0+2,\ldots\}$ are i.i.d. random variables with mean $r_i^S$. Deviating only benefits at time slot $T_0$. According to the strong law of large numbers [23], the payoff $U_i^D$ converges to its mean $r_i^S$ almost surely.

On the other hand, if no deviation ever happens, the repeated game always stay in the cooperative stage. By using the same argument, $U_i^C$ converges to $r_i^C$ almost surely. ■

Because the selfish players always choose the strategy that maximizes their own payoffs, they will keep cooperation if $U_i^C(=r_i^C) > U_i^D(=r_i^S)$, that is, all players are self-enforced to cooperate in the repeated spectrum sharing game because of punishment after any deviation.

Nevertheless, such a harsh threat is neither efficient nor necessary. Note that not only the deviating player gets punished, but the other "good" players also suffer from the punishment. For example, if one player deviates by mistake or punishment is triggered by mistake, there will be no cooperation due to punishment, which results in lower efficiency for all players. We have to review the purpose of the punishment. The aim of punishment is more like "preventing" the deviating behaviors from happening rather than punishing for revenge after deviation. As long as the punishment is long enough to negate the reward from a one-time deviation, no player has an incentive to deviate. The new strategy, called "punish-and-forgive", is stated as follows: the game starts from the cooperative stage, and will stay in the cooperative stage until some deviation happens. Then, the game jumps into the punishment stage for the next $T-1$ time slots before the misbehavior is forgiven and cooperation resumes from the $T$-th time slot. $T$ is called the duration of punishment. In the cooperative stage, every player shares the spectrum in a cooperative way according to their agreement; while in the punishment stage, players occupy the spectrum non-cooperatively as they would do in the one-shot game. The following proposition shows that cooperation is a subgame perfect equilibrium, which ensures the Nash optimality for subgames starting from any round of the whole game.

*Proposition 3:* Provided $r_i^C > r_i^S$ for all $i = 1,2,\ldots,K$, there is $\bar{\delta} < 1$, such that for a sufficiently large discount factor $\delta > \bar{\delta}$, the game has a subgame perfect equilibrium with discounted utility $r_i^C$, if all players adopt the "punish-and-forgive" strategy.

*Proof:* Because $\left(P_1^M, P_2^M, \ldots, P_K^M\right)$ is a Nash equilibrium for the one-shot game, transmitting with the power vector $\left(P_1^M, P_2^M, \ldots, P_K^M\right)$ in every time slot is one of the Nash equilibria for the repeated game. Then, the proof of this proposition follows the Folk Theorem with Nash threats [9]. The theorem states that the "punish-and-forgive" strategy yields a subgame perfect equilibrium for sufficiently patient players (i.e., $\delta$ close to 1), whenever the game has a pure one-shot Nash equilibrium. Using the one-shot Nash equilibrium strategy as punishment and cooperating otherwise will enforce all the players to cooperate. ■

The parameter $T$ can be determined by analyzing the incentive of the players. For example, we investigate under what condition player $i$ will lose the motivation to deviate at time slot $T_0$. Although cooperation guarantees an average payoff $r_i^C$ at each time slot, the worst-case instantaneous payoff could be 0. On the contrary, deviation will prompt the instantaneous payoff at that slot. Assume the maximal profit obtained from deviation is $\overline{R_i^D}$. If player $i$ chooses to deviate, punishment stage will last for the next $T-1$ slots; otherwise, cooperation will always be maintained. Thus, the expected payoffs with and without deviation are bounded by

$$\begin{aligned}u_i^D \triangleq E[U_i^D] \le (1-\delta)\cdot \\ \left(\sum_{n=0}^{T_0-1}\delta^n r_i^C + \delta^{T_0}\overline{R_i^D} + \sum_{n=T_0+1}^{T_0+T-1}\delta^n r_i^S + \sum_{n=T_0+T}^{+\infty}\delta^n r_i^C\right),\end{aligned} \qquad (8)$$

and

$$u_i^C \triangleq E[U_i^C] \ge (1-\delta)\cdot\left(\sum_{n=0}^{T_0-1}\delta^n r_i^C + 0 + \sum_{n=T_0+1}^{+\infty}\delta^n r_i^C\right), \qquad (9)$$

respectively. From selfish players' point of view, the one with the higher payoff is clearly the better choice. $T$ should be large enough to deter players from deviating such that $u_i^C > u_i^D$ for all $i = 1,2,\ldots,K$. Then, the necessary condition for $T$ can be solved as

$$T > \max_i \frac{\log\left(\delta - \frac{(1-\delta)\overline{R_i^D}}{r_i^C - r_i^S}\right)}{\log\delta}, \qquad (10)$$

which can be further approximated by

$$T > \max_i \frac{\overline{R_i^D}}{r_i^C - r_i^S} + 1, \qquad (11)$$

by L'Hopital's rule when $\delta$ is close to 1. If the tendency to deviate is stronger (i.e., $\overline{R_i^D}/(r_i^C - r_i^S)$ is larger), the punishment should be harsher (longer duration of punishment) to prevent the deviating behavior.

## IV. Cooperation with Optimal Detection

In this section, we will discuss the specific design of the cooperation rules for spectrum sharing, as well as the method to detect deviation. When designing the rules, we assume that players can exchange information over a common control channel. Based on the information, each individual can independently determine who is eligible to transmit in the current time slot according to the cooperation rule, and thus the proposed scheme does not require a central management unit.

Cooperative spectrum sharing can be designed in the following way: in one time slot, only a few players with small mutual interference can access the spectrum simultaneously. In the extreme case, only one player is allowed to occupy the spectrum at one time slot, and the mutual interference can be completely prevented. In this paper, we will limit our attention to such orthogonal channel allocation for the following reasons, and more general cooperation rules will be studied in the future.

- It is quite simple, and the performance is good in an environment where the interference level is medium to high, as illustrated by the simulation results. This is the case where wireless users concentrate in a small area, e.g., a cluster of users inside an office building or within a coffee house.
- If several players are allowed to access the spectrum simultaneously, they will have to negotiate how much power each one can use. However, for the orthogonal assignment, if one player gets the exclusive right to occupy the channel, the maximum power will be used. Therefore, the action space boils down from a continuum of power levels to a binary choice (either $0$ or $P_i^M$), which simplifies the problem.
- In order to decide who can access the spectrum, information like the channel gains is needed. If multiple players are allowed to transmit in one time slot, the whole channel state matrix $\{h_{ji}, j = 1, 2, \ldots, K, i = 1, 2, \ldots, K\}$ is necessary to decide which players can be grouped together. The total amount of exchanged information is $O(K^2)$. On the contrary, the orthogonal assignment is interference-free, and only the diagonal entries $\{h_{ii}, i = 1, 2, \ldots, K\}$ have to be exchanged, which means the overhead reduces to $O(K)$.
- The orthogonal allocation also facilitates the detection of deviating behavior. In general, the detector is required to catch the deviation by distinguishing the ineligible players from the players allowed to access the spectrum. The detection becomes much easier in the orthogonal assignment case. The only eligible player in the current time slot will declare an event of deviation and trigger the punishment once he/she finds that someone else is also active in the unlicensed band.

The slot structure for the spectrum sharing is shown in Fig. 2. Every slot is divided into three phases: in the first phase, each player broadcasts information to others, such as channel gains; in the second phase, each player collects all the necessary information and decides whether to access the spectrum or not, according to the cooperation rule; then the eligible player will occupy the spectrum in the third phase of the slot. If the channel does not change too rapidly, the length of a slot can be designed long enough to make the overhead (the first and second phases) negligible. Since it is necessary to detect the potential deviating behavior and punish correspondingly, the eligible player cannot transmit all the time during the third phase. Instead, the player has to suspend his/her own transmission sometimes and listens to the channel to catch the deviators. The eligible player transmits and detects during the third phase: a portion of time is reserved for detection, while the rest can be used for transmission. When to perform detection during the slot is kept secret by individuals; otherwise, the other players may take advantage by deviating when the detector is not operating. Finally, if detection shows someone is deviating, an alert message will be delivered in the first phase of the next time slot.

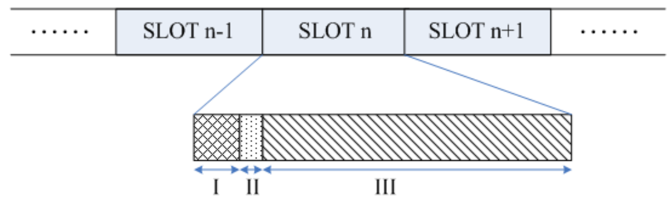


Fig. 2. Proposed slot structure for spectrum sharing. Phase I: exchange information; phase II: make decision; phase III: transmit and detect.

### A. Cooperation Criteria

There are numerous cooperation rules to decide which players can have exclusive priority to access the channel, such as the time division multiple access (TDMA). Out of many possible choices, the cooperation rules must be reasonable and optimal under some criteria, such as the maximum total throughput criterion [24] and the proportional fairness criterion [25].

Given a cooperation rule $d$, player $i$ would have an expected discounted payoff $r_i^{Cd}$. Denote $\mathcal{D}$ as the set of all possible cooperation rules. The maximum total throughput criterion aims to improve the overall system performance by maximizing the sum of individual payoffs,

$$d_{Max} = \underset{d \in \mathcal{D}}{\operatorname{argmax}} \sum_{i=1}^{K} r_i^{Cd}, \tag{12}$$

whereas the proportional fairness criterion is known to maximize their product,

$$d_{PF} = \underset{d \in \mathcal{D}}{\operatorname{argmax}} \prod_{i=1}^{K} r_i^{Cd}. \tag{13}$$

The rule based on the maximum total throughput criterion (MTT) is quite straightforward. In order to maximize the total throughput, each time slot should be assigned to the player that makes best use of it. Denote $g_i[n] = P_i^M |h_{ii}[n]|^2$ as the instantaneous received signal power of the $i$-th player at time slot $n$, and $\{g_i[n]\}$ are i.i.d. exponentially-distributed random variables with mean $P_i^M \sigma_{ii}^2$ according to the assumption about $\{h_{ii}[n]\}$. The allocation rule is to assign the channel to the player with the highest instantaneous received signal power, i.e.,

$$d_1(g_1, g_2, \ldots, g_K) = \operatorname*{argmax}_i g_i. \tag{14}$$

Since only the information of the current time slot is necessary and the same rule applies to every time slot, the time index $n$ has been omitted. The expected payoff is

$$r_i^{C1} = \int_0^{+\infty} \log_2\left(1 + \frac{g_i}{N_0}\right) \Pr(g_i > max_{j\neq i} g_j) f(g_i) \mathrm{d}g_i, \tag{15}$$

where $f(g_i) = \frac{1}{P_i^M \sigma_{ii}^2} \exp(-\frac{g_i}{P_i^M \sigma_{ii}^2})$ is the probability density function of the random variable $g_i$, and $\Pr(\cdot)$ denotes the probability that the statement within the parenthesis holds true.

The maximum total throughput criterion is optimal from the system designer's perspective; however, in a heterogeneous situation where some players always have better channels than others, the players under poor channel conditions may have little chance to access the spectrum. To address the fairness problem, another rule is proposed which allocates the spectrum according to the normalized channel gain $\bar{g}_i = g_i/E[g_i]$ instead of the absolute values,

$$d_2(\bar{g}_1, \bar{g}_2, \ldots, \bar{g}_K) = \operatorname*{argmax}_i \bar{g}_i. \tag{16}$$

Note that all $\{\bar{g}_i, i = 1, 2, \ldots, K\}$ are exponentially-distributed random variables with mean 1, the symmetry of which implies that every player will have an equal chance $(1/K)$ to access the spectrum.

*Proposition 4:* The closed-form payoff with the proposed rule (16) used can be shown as follows

$$r_i^{C2} = \int_0^{+\infty} \log_2\left(1 + \frac{P_i^M \sigma_{ii}^2 \bar{g}}{N_0}\right) \exp(-\bar{g}) \left(1 - \exp(-\bar{g})\right)^{K-1} \mathrm{d}\bar{g}. \tag{17}$$

*Proof:* The probability distribution function of each $\bar{g}_i$ is $F(\bar{g}_i) = 1 - \exp(-\bar{g}_i)$. By order statistics [26], the maximum among the $K$ i.i.d. random variables $\{\bar{g}_i, i = 1, 2, \ldots, K\}$ has the distribution function $F_M(\bar{g}) = (1 - \exp(-\bar{g}))^K$. Since each player can be the one with the largest $\bar{g}_i$ with probability $1/K$ due to symmetry, the expected payoff is

$$r_i^{C2} = \int_0^{+\infty} \log_2\left(1 + \frac{P_i^M \sigma_{ii}^2 \bar{g}}{N_0}\right) \frac{1}{K} \mathrm{d}F_M(\bar{g}). \tag{18}$$

Substituting $F_M(\bar{g})$ yields the form of the payoff in (17). ■

*Remarks:*

- The proposed rule (16) is an approximation to the proportional fairness criterion (13). $g_i$ can be decomposed into a fixed component $E[g_i]$ and a fading component $\bar{g}_i$. When the channel is constant without fading, i.e., $g_i = E[g_i]$, the proportional fairness problem becomes

$$\max_{\{\omega_i\}} \prod_{i=1}^{K} \omega_i \log_2\left(1 + \frac{E[g_i]}{N_0}\right) \quad \text{s.t.} \sum_{i=1}^{K} \omega_i \leq 1, \tag{19}$$

  where $\omega_i$ is the probability that the $i$-th player should occupy the channel. The optimal solution is $\omega_i = 1/K$ for any $i$, which means an equal share is proportionally fair. On the other hand, when only the fading part is considered, since $\bar{g}_i$ is completely symmetric for all players, assigning resources to the player with the largest $\bar{g}_i$ will maximize the product of payoffs. The two aspects suggest that rule (16) is a good approximation which requires only the information of the current time slot, and we will refer to it as the APF (approximated proportional fairness) rule in the rest of the paper.
- The rule (16) can be extended to a more general case which allocates the band according to weighted normalized channel gain $\pi_i \bar{g}_i$, where $\pi_i$ is a weight factor reflecting a player's priority for heterogeneous applications.

### B. Optimal detection

The punishment-based spectrum sharing game can provide all players with the incentive to obey the rules, since deviation is deterred by the threat of punishment. Detection of the deviating behavior is necessary to ensure the threat to be credible; otherwise, selfish players will tend to deviate knowing their misbehavior will not be caught. Because only one player can occupy the spectrum at one time slot according to the proposed cooperation rules, if that player finds that any other player is deviating, the system will be alerted into the punishment phase. There are several ways to detect whether the spectrum resources are occupied by others; in this paper we assume the player can listen to the channel from time to time using an energy detector [27].

The detectors are generally imperfect, and some detection errors are inevitable. There is the possibility that the detector believes someone else is using the channel although in fact nobody is. Triggering the game into punishment phase by mistake, this false alarm event reduces the system efficiency, and hence the probability of false alarm should be kept as low as possible. Generally speaking, the performance of the detector can be improved by increasing the detection time. Nevertheless, the player cannot transmit and detect at the same time because one cannot easily distinguish one's own signal from other players' signal in the same spectrum. Therefore, there is a tradeoff between transmission and detection: the more time one spends on the detection, the less time one reserves for data transmission.

Assume all the other parameters have been fixed, such as the length of one time slot. Then, the question is how much time in a slot should be used for detection. Let $\alpha$ denote the ratio of time for detection, $T_s$ the length of one slot, $W_s$ the bandwidth, and assume an energy detector with a threshold $\lambda$ is used, then the false alarm probability is [27]

$$\xi(\alpha) = \frac{\Gamma(\alpha T_s W_s, \alpha\lambda/2)}{\Gamma(\alpha T_s W_s)}, \tag{20}$$

where $\Gamma(\cdot)$ and $\Gamma(\cdot, \cdot)$ are the gamma function and incomplete gamma function, respectively.

We have shown that the expected discounted payoff $u_i$ equals $r_i^C$ without considering the detection error. When the imperfect detector is taken into account, the modified discounted payoff, denoted by $\check{u}_i(\alpha)$, will depend on $\alpha$. The expected throughput from the current time slot is $(1-\alpha) r_i^C$, since only the remaining $(1-\alpha)$ part of the duration can be employed for transmission. The system will jump into the punishment stage with probability $\xi(\alpha)$ due to the false alarm event, and stay in the cooperative stage with probability $1-\xi(\alpha)$. If the system stays in the cooperative stage, the expected payoff in the future is $\breve{u}_i(\alpha)$ discounted by one time unit; otherwise, the expected throughput in each time slot is $r_i^S$ until cooperation resumes from the $T$-th slot, which yields the payoff $\breve{u}_i(\alpha)$ discounted by $T$ time units. Overall, the modified discount utility should satisfy the following equation,

$$\begin{aligned}\breve{u}_i(\alpha) =&(1-\delta)(1-\alpha)r_i^C + (1-\xi(\alpha))\delta\breve{u}_i(\alpha)\\ &+\xi(\alpha)\left((1-\delta)\sum_{n=1}^{T-1}\delta^n r_i^S + \delta^T\breve{u}_i(\alpha)\right),\end{aligned} \tag{21}$$

from which $\breve{u}_i(\alpha)$ can be solved as

$$\breve{u}_i(\alpha) = \frac{(1-\delta)(1-\alpha)r_i^C + (\delta-\delta^T)\xi(\alpha)r_i^S}{1-\delta+(\delta-\delta^T)\xi(\alpha)}. \tag{22}$$

Note that the discounted payoff $\breve{u}_i(\alpha)$ is a convex combination of $(1-\alpha)r_i^C$ and $r_i^S$, and thus $r_i^S < \breve{u}_i(\alpha) < r_i^C$ for all $0 < \alpha < 1 - r_i^S/r_i^C$. Therefore, the imperfect detector will reduce the utility from $r_i^C$ to a smaller value $\breve{u}_i(\alpha)$. However, $\breve{u}_i(\alpha)$ is always larger than $r_i^S$, which means that the players still have the incentive to join in this repeated game and cooperate.

The optimal $\alpha^*$ that maximizes the modified discounted payoff (22) can be found by the first order condition

$$\frac{\partial\breve{u}_i(\alpha)}{\partial\alpha} = 0. \tag{23}$$

Or equivalently, $\alpha^*$ is the solution to the following equation

$$(1-\delta+(\delta-\delta^T))r_i^C + ((1-\alpha)r_i^C - r_i^S)(\delta-\delta^T)\frac{\xi'(\alpha)}{\xi(\alpha)} = 0, \tag{24}$$

where $\xi'(\alpha)$ is the derivative of $\xi(\alpha)$ with respect to $\alpha$. Note that by replacing $r_i^C$ with $\breve{u}_i(\alpha^*)$, the impact of imperfect detection is incorporated into the game, and requires no further considerations.

## V. Cheat-Proof Strategies

The repeated game discussed so far is based on the assumption of complete and perfect information. Nevertheless, information, such as the power constraints and channel gains, is actually private information of each individual player, and thus there is no guarantee that players will reveal their private information honestly to others. If cheating is profitable, selfish players will cheat in order to get a higher payoff. As the proposed cooperation rules always favor the player with good channel conditions, selfish players will tend to exaggerate their situations in order to acquire more opportunities to occupy the spectrum. Therefore, enforcing truth-telling is a crucial problem, since distorted information would undermine the repeated game.

In [18], a delicate scheme is designed to testify whether the information provided by an individual player has been revealed honestly. However, the method is complex and difficult to implement, especially under time-varying channels. In our proposed allocation rules, much easier strategies can be employed to induce truth-telling. When the MTT rule is used for spectrum sharing, we design a mechanism to make players self-enforced to reveal their true private information, and when the APF rule is adopted, a scheme based on statistical properties is proposed to discourage players from cheating.

### A. Mechanism-design-based strategy

Since the MTT sharing rule assigns the spectrum resources to the player who claims the highest instantaneous received signal power, players tend to exaggerate their claimed values. To circumvent the difficulty to tell whether the exchanged information has been distorted or not, a better way is to make players self-enforced to tell the truth.

Mechanism design is employed to provide players with incentives to be honest. To be specific, the players claiming high values are asked to pay a tax, and the amount of the tax will increase as the claimed value increases, whereas the players reporting low values will get some monetary compensation. This is called "transfer" in Bayesian mechanism design theory [22]. When the transfer of a player is negative, he/she has to pay others; otherwise, he/she gets compensation from others. Because players care for not only the gain of data transmission but also their monetary balance, the overall payoff is gain of transmission plus the transfer. In other words, after introducing transfer functions, the spectrum sharing game actually becomes a new game with original payoffs replaced by the overall payoffs. By appropriately designing the transfer function, the players can get the highest payoff only when they claim their true private values.

According to the cooperative allocation rule, the private information $\{g_1, g_2, \ldots, g_K\}$ has to be exchanged among players. Assume at one time slot, $\{\tilde{g}_1, \tilde{g}_2, \ldots, \tilde{g}_K\}$ is a realization of the random variables $\{g_1, g_2, \ldots, g_K\}$. Observing his/her own private information, the $i$-th player will claim $\hat{g}_i$ to others, which may not be necessarily the same as the true value $\tilde{g}_i$. All the players claim the information simultaneously. Since $\{\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K\}$ is common knowledge but $\{\tilde{g}_1, \tilde{g}_2, \ldots, \tilde{g}_K\}$ is not, the allocation decision and transfer calculation have to be based on the claimed rather than the true values. In the MTT spectrum sharing game, the player with index $d_1(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K)$ defined in (14) can access the channel, and thus data throughput at the current time slot can be written in a compact form

$$\begin{aligned}&R_i(\tilde{g}_i, d_1(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K))\\ &\quad= \begin{cases} \log_2(1+\frac{\tilde{g}_i}{N_0}) & \text{if } d_1(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K) = i \\ 0 & \text{otherwise} \end{cases}.\end{aligned} \tag{25}$$

The transfer of the $i$-th player in the proposed cheat-proof strategy is defined as

$$t_i(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K) \triangleq \Phi_i(\hat{g}_i) - \frac{1}{K-1}\sum_{j=1, j\neq i}^{K}\Phi_j(\hat{g}_j), \tag{26}$$

where

$$\Phi_i(\hat{g}_i) \triangleq E\left[\sum_{j=1,j\neq i}^{K} R_j(g_j, d_1(g_1, g_2, \ldots, g_K))\middle| g_i = \hat{g}_i\right]. \tag{27}$$

Note the expectation is taken over all realizations of $\{g_1, g_2, \ldots, g_K\}$ except $g_i$, since the player has no knowledge about others of the current time slot when deciding what to claim. $\Phi_i(\hat{g}_i)$ is the sum of all other players' expected data throughput given that player $i$ claims a value $\hat{g}_i$. Intuitively, if user $i$ claims a higher $\hat{g}_i$, he/she will gain a greater chance to access the spectrum, and all the other players will have a smaller spectrum share. However, higher payment may negate the additional gain from more spectrum access through bragging the channel gain. On the contrary, if claiming a smaller $\hat{g}_i$, user $i$ will receive some compensation at the cost of less chance to occupy the spectrum. Therefore, it is an equilibrium that each user reports his/her true private information. A rigorous proof is provided in the following proposition.

*Proposition 5:* In the proposed mechanism, it is an equilibrium that each player reports his/her true private information, i.e., $\hat{g}_i = \tilde{g}_i$, $i = 1, 2, \ldots, K$.

*Proof:* To prove the equilibrium, it suffices to show that for any $i \in \{1, 2, \ldots, K\}$, if all players except player $i$ reveal their private information without distortion, the best response of player $i$ is also to report the true private information. Without loss of generality, we assume player 2 through player $K$ report true values $\hat{g}_i = \tilde{g}_i$, $i = 2, 3, \ldots, K$.

Then, the expected overall payoff of player 1 is the expected data throughput plus the transfer. The expectation is taken over all realizations of $\{g_2, g_3, \ldots, g_K\}$ throughout the proof. When claiming $\hat{g}_1$, player 1 gets the expected overall payoff

$$r_1^t(\hat{g}_1) \triangleq E\left[R_1(\tilde{g}_1, d_1(\hat{g}_1, g_2, \ldots, g_K)\right] + t_1(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K) = E\left[R_1(\tilde{g}_1, d_1(\hat{g}_1, g_2, \ldots, g_K)) + \sum_{j=2}^{K} R_j(g_j, d_1(\hat{g}_1, g_2, \ldots, g_K))\right] - \frac{1}{K-1}\sum_{j=2}^{K}\Phi_j(\hat{g}_j). \quad (28)$$

From analysis of incentive compatibility, player 1 will claim a distorted value $\hat{g}_1$ instead of $\tilde{g}_1$ if and only if reporting $\hat{g}_1$ results in a higher payoff, i.e., $r_1^t(\tilde{g}_1) < r_1^t(\hat{g}_1)$, or equivalently,

$$E\left[R_1(\tilde{g}_1, d_1(\tilde{g}_1, g_2, \ldots, g_K)) + \sum_{j=2}^{K} R_j(g_j, d_1(\tilde{g}_1, g_2, \ldots, g_K))\right] < E\left[R_1(\tilde{g}_1, d_1(\hat{g}_1, g_2, \ldots, g_K)) + \sum_{j=2}^{K} R_j(g_j, d_1(\hat{g}_1, g_2, \ldots, g_K))\right]. \quad (29)$$

Note that the MTT rule maximizes the total throughput, that is, for any realization of $\{g_2, g_3, \ldots, g_K\}$, $\sum_{i=1}^{K} R_i(\tilde{g}_i, d_1(\tilde{g}_1, \tilde{g}_2, \ldots, \tilde{g}_K)) > \sum_{i=1}^{K} R_i(\tilde{g}_i, d^o)$ for any other possible allocation strategy $d^o$. After taking the expectation, we have

$$E\left[R_1(\tilde{g}_1, d(\tilde{g}_1, g_2, \ldots, g_K)) + \sum_{j=2}^{K} R_j(g_j, d(\tilde{g}_1, g_2, \ldots, g_K))\right] > E\left[R_1(\tilde{g}_1, d^o) + \sum_{j=2}^{K} R_j(g_j, d^o)\right] \text{ for any } d^o, \quad (30)$$

which contradicts (29). Therefore, player 1 is self-enforced to report the true value, i.e., $\hat{g}_1 = \tilde{g}_1$. Hence, in the equilibrium, all players will reveal their true private information. ■

The proposition proves that by adopting the proposed mechanism-based strategy with transfer function defined in (26), every player gets the incentive to reveal true private information to others. For the homogenous case where $P_i^M = P$, $h_{ii} \sim \mathcal{CN}(0, 1)$ for all $i$, the transfer function can be further simplified into the following form by order statistics

$$t_i(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K) = \sum_{j=1}^{K} \int_{\hat{g}_i/P}^{\hat{g}_j/P} \log_2\left(1 + \frac{Pg}{N_0}\right) \exp(-g)\left(1 - \exp(-g)\right)^{K-2} \mathrm{d}g. \quad (31)$$

Moreover, with the proposed transfer functions, all players' payment/income adds up to 0 at any time slot:

$$\sum_{i=1}^{K} t_i(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K) = \sum_{i=1}^{K}\left(\Phi_i(\hat{g}_i) - \frac{1}{K-1}\sum_{j=1, j\neq i}^{K}\Phi_j(\hat{g}_j)\right) = \sum_{i=1}^{K}\Phi_i(\hat{g}_i) - \sum_{j=1}^{K}\Phi_j(\hat{g}_j) = 0. \quad (32)$$

It means that the monetary transfer is exchanged only within the community of cooperative players without either surplus or deficit at any time. This property is very suitable for the open spectrum sharing scenario. Vickrey-Clark-Groves (VCG) mechanism [28], another well-known mechanism, can also enforce truth-telling, but it cannot keep the budget balanced. If the VCG mechanism is used, at each slot some players will have to pay a third party outside the community (e.g., a spectrum manager), which goes against the intention of the unlicensed band. Furthermore, paying for the band may make players less willing to access the spectrum. Despite that the VCG mechanism is a good choice for auctions in licensed spectrum, for the unlicensed band, as our goal is increasing spectrum efficiency and enforcing truth-telling rather than making money out of the spectrum resources, the proposed mechanism is more appropriate.

### *B. Statistics-based strategy*

For the APF rule, every player reports the normalized channel gain, and the player with the highest reported value will get access to the spectrum. Since the normalized gains are all exponentially distributed with mean 1, if the true values are reported, the symmetry will result in an equal share of the time slots in the long run, i.e., each player will have $1/K$ fractional access to the spectrum.

If player $i$ occupies the spectrum more than $(1/K + \varepsilon)$ of the total time, where $\varepsilon$ is a pre-determined threshold, it is highly possible that he/she may have cheated. Consequently, the selfish players, in order not to be caught as a cheater, can only access up to $(1/K + \varepsilon)$ of all the time slots even if they distort their private information. Thus, the statistics-based cheat-proof strategy for the APF spectrum sharing rule can be designed as follows. Everyone keeps a record of the spectral usage in the past. If any player is found to overuse the spectrum, i.e., transmitting for more than $(1/K + \varepsilon)$ of the entire time, that player will be marked as a cheater and get punished. In this way, the profit of cheating, defined as the ratio of the extra usage over the normal usage, is greatly limited.

*Proposition 6:* The profit of cheating is bounded when the statistics-based strategy is employed; furthermore, the profit approaches 0 as $n \to \infty$.

*Proof:* The worst case is that the cheater gets exactly $(1/K + \varepsilon)$ portion of resources without being caught. The profit of cheating is at most $\frac{\varepsilon}{1/K} = K\varepsilon$, which is bounded.

Moreover, the threshold $\varepsilon$ can shrink with time; to make it explicit, we use $\varepsilon[n]$ to denote the threshold at slot $n$. At one time slot, the event that a particular player accesses the spectrum is a Bernoulli distributed random variable with mean $1/K$. Then, the $n$-slot averaged access rate of a player is the average of $n$ i.i.d. Bernoulli random variables, since the channel fading is independent from slot to slot. According to the central limit theorem [23], the average access rate converges in distribution to a Gaussian random variable with mean $1/K$, whose variance decays with rate $\frac{1}{n}$. To keep the same false alarm rate, $\varepsilon[n]$ can be chosen to decrease with rate $\frac{1}{\sqrt{n}}$. Then, the upper bound of the cheating profit $K\varepsilon[n]$ will decay to 0 as $n \to \infty$. ■

Therefore, we can conclude from the proposition that the benefit to the cheater, or equivalently speaking, the harm to the others, is quite limited. As a result, this statistics-based strategy is cheat-proof.

## VI. Simulation Results

In this section, we conducted numerical simulations to evaluate the proposed spectrum sharing game with cheat-proof strategies.

We first look into the simplest case with two players ($K = 2$) to get some insight. We assume the two players are homogeneous with $P_1^M = P_2^M = P$, $\{h_{11}, h_{22}\} \sim \mathcal{CN}(0, 1)$, and $\{h_{12}, h_{21}\} \sim \mathcal{CN}(0, \gamma)$, where $\gamma = E[|h_{12}|^2]/E[|h_{11}|^2]$ reflects the relative strength of interference over the desired signal powers, and we call it the interference level. The prerequisite for the players to join the game is that each individual can obtain more profit by cooperation ($r_i^C > r_i^S$); however, cooperation is unnecessary in the extreme case when there is no mutual interference ($\gamma = 0$). Therefore, we want to know under what interference level $\gamma$ the proposed cooperation is profitable. Fig. 3 shows the cooperation payoff $r_i^C$ and non-cooperation payoff $r_i^S$ versus $\gamma$ when the averaged SNR $= P/N_0 = 15dB$. Since the two rules are equivalent in the homogeneous case, only the MTT rule is demonstrated. Under cooperative spectrum sharing, since only one player gets the transmission opportunity in each time slot, the expected payoff is independent of the strength of interference, and thus is a horizontal line in the figure. The non-cooperation payoff drops significantly as interference strength increases. From the figure, we can see that the payoff of cooperation is larger than that of non-cooperative for a wide range of the interference level ($\gamma > 0.15$). Therefore, the proposed cooperation is profitable for medium to high interference environment, which is typical for an urban area with high user density.

In Fig. 4, we illustrate the idea of the punishment-based repeated game. Assume player 1 deviates from cooperation at slot 150, and the duration of the punishment stage is $T = 150$.

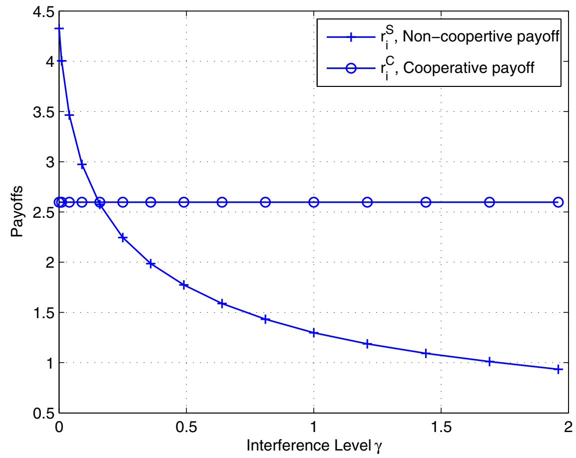


Fig. 3. Comparison of payoffs when the players share the spectrum either cooperatively or non-cooperatively.

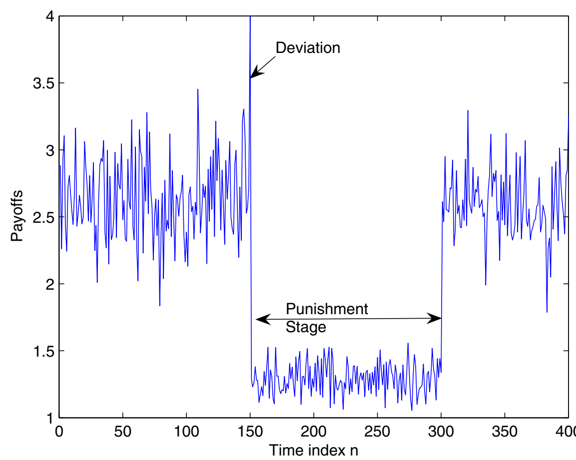


Fig. 4. Illustration of the punishment-based repeated game.

According to the "punish-and-forgive" strategy, the game will stay in the punishment stage from time slot 151 to 300. The figure shows an averaged result over 100 independent runs. We can see that although the player gets a high payoff at time slot 150 by deviation, the temporary profit will be negated in the punishment stage. Hence, considering the consequence of deviation, players have no incentive to deviate.

The effect of $\delta$ is demonstrated in Fig. 5. We assume all players have the same tendency of deviation, i.e., $\tau_i = \tau$, $\forall i$ where $\tau_i \triangleq \overline{R_i^D}/(r_i^C - r_i^S)$. We evaluate the effect of $\delta$ when $\tau$ equals 5, 10, and 20, respectively. Given a fixed $\delta$, any punishment duration $T$ above the curve can deter the deviation, and the duration should be longer for larger $\tau$ as players have greater incentive to deviate. In addition, when $\delta$ is close to 1, the minimal duration goes to $\tau + 1$ as in (11), and $\delta$ has to be larger than $\bar{\delta} = \tau/(1 + \tau)$ to guarantee that punishment can prevent players from deviating ($\delta > \bar{\delta}$ is necessary to make (10) valid). In other words, in situations where players are impatient ($\delta$ is far away from 1) and tendency to deviate is strong, it is impossible to maintain cooperation using the "punish-and-forgive" strategy with

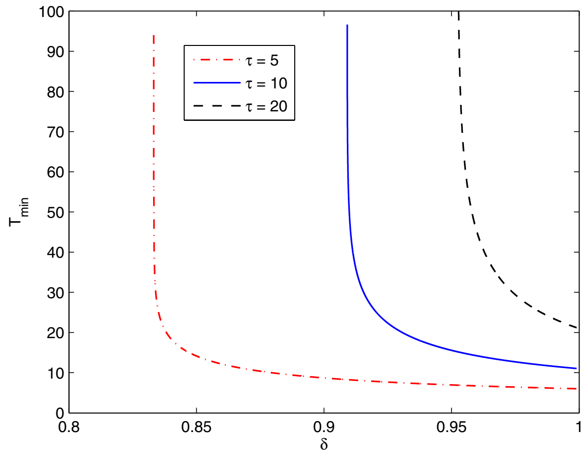


Fig. 5. Effect of the discount factor $\delta$ on the punishment duration $T$.

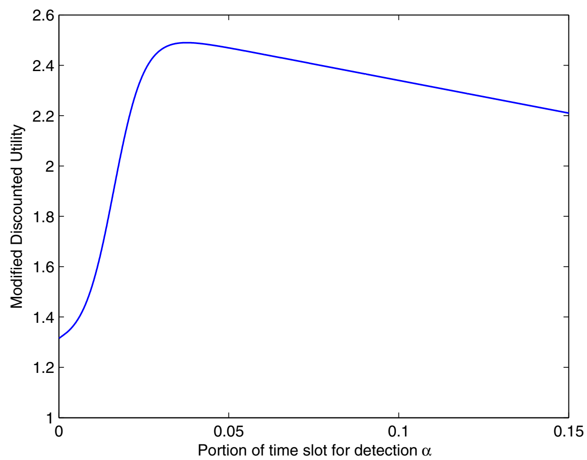


Fig. 6. Effect of the length of detection on the discounted utility.

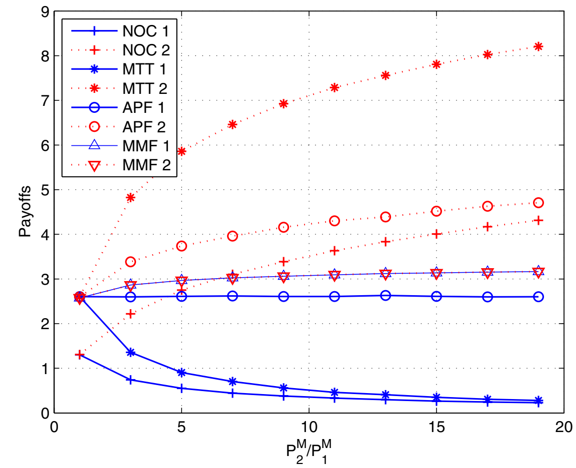


Fig. 7. The payoffs under a heterogeneous setting with different cooperation rules.

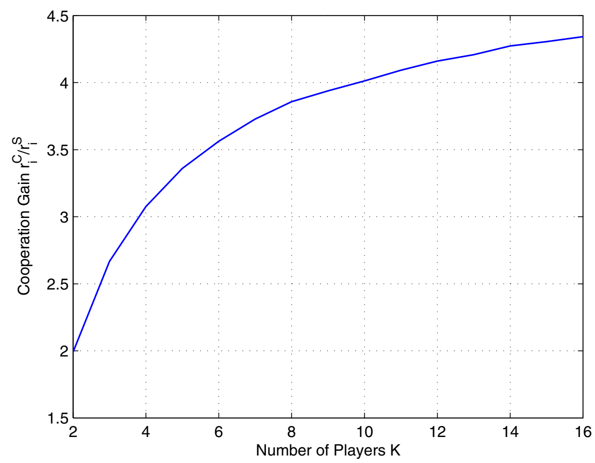


Fig. 8. The cooperation gain in a $K$-player spectrum sharing game.

repeated-game modeling.

Now, we take imperfect detection into consideration. Fig. 6 shows how a player's discount utility $\breve{u}(\alpha)$ is affected by $\alpha$ when an energy detector with a fixed detection threshold is employed. We can see that when the detection time is short, the utility is quite low due to the high false alarm rate. On the other hand, when the detection time is too long, a significant portion of the transmission opportunity is wasted. Therefore, $\alpha$ should be carefully designed to achieve the optimal tradeoff that maximizes the utility.

Next, we show the payoffs of proposed cooperation rules in a heterogeneous environment. By heterogeneity, we mean that different players may differ in power constraints, averaged direct-channel gains $\{h_{ii}, i = 1, 2, \ldots, K\}$, averaged cross-channel gains $\{h_{ij}, i \neq j\}$, or combination of them. Here we only illustrate the results when the power constraints are different, since other types of asymmetry have similar results. In the simulation, we fix the power constraint of player 1, and increase $P_2^M$, the power constraint of player 2. The payoffs with the MTT and APF rules are demonstrated in Fig. 7, where '1' and '2' refer to the payoffs of player 1 and player 2, respectively. As benchmarks, the payoffs without cooperation and payoffs using the max-min fairness criterion (another resource allocation criterion sacrificing efficiency for absolute fairness, see [29]) are provided, denoted by "NOC" and "MMF", respectively. Since player 2 has more power to transmit data, he/she can be seen as a strong player in this heterogeneous environment. As seen from the figure, both the MTT and APF rules outperform the non-cooperation case, which means players have the incentive to cooperate no matter which rule is used. Furthermore, the MTT rule favors the strong player in order to maximize the system efficiency, and the APF rule achieves a tradeoff between efficiency and fairness. The MMF curves show that the strong user is inhibited in order to reach the absolute fairness, which might conflict with selfish users' interest.

We also conduct simulations for spectrum sharing with more than two users. In Fig. 8, the cooperation gain, characterized by the ratio of $r_i^C/r_i^S$, is plotted versus the number of the players $K$. We assume a homogeneous environment with a fixed interference level $\gamma = 1$. Since the allocation rules can reap multiuser diversity gains, the cooperation gain increases as more players are involved in the sharing game.

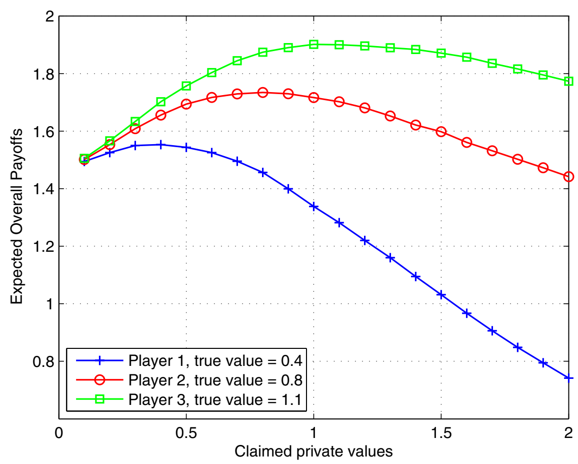


Fig. 9. The expected overall payoffs versus different claimed values.

Finally, we examine the proposed mechanism-design-based cheat-proof strategy. We assume a 3-user spectrum sharing game with the MTT rule. At one specific time slot, for example, the true private values are $\tilde{g}_1 = 0.4$, $\tilde{g}_2 = 0.8$, and $\tilde{g}_3 = 1.1$. In Fig. 9, the expected overall payoffs (throughput plus transfer) versus the claimed values are shown for each player, given the other two are honest. From the figure, we see that the payoff is maximized only if the player honestly claims his/her true information. Therefore, players are self-enforced to tell the truth with the proposed mechanism.

## VII. Conclusions

We have proposed a novel spectrum sharing scheme with cheat-proof strategies to improve the efficiency of open spectrum sharing. The spectrum sharing problem is modeled as a repeated game where any deviation from cooperation will trigger the punishment. We propose two cooperation rules with efficiency and fairness considered, and optimize the detection time to alleviate the impact due to imperfect detection of the selfish behavior. Moreover, two cheat-proof strategies based on mechanism design and properties of channel statistics are proposed to enforce that selfish users report their true channel information. Simulation results show that the proposed scheme has efficiently improved the spectrum usage by alleviating the mutual interference.

## References

[1] Federal Communications Commission, "Spectrum policy task force report," *FCC Document ET Docket No. 02-135*, Nov. 2002.
[2] Federal Communications Commission, "Facilitating opportunities for flexible, efficient and reliable spectrum use employing cognitive radio technologies: notice of proposed rule making and order," *FCC Document ET Docket No. 03-108*, Dec. 2003.
[3] J. Mitola III, "Cognitive radio: an integrated agent architecture for software defined radio," Ph.D. Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2000.
[4] G. Hardin, *The tragedy of the commons*, 1968.
[5] W. Lehr, and J. Crowcroft, "Managing shared access to a spectrum commons," in *IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN'05)*, Baltimore, Nov. 2005.
[6] S. Keshavamurthy, and K. Chandra, "Multiplexing analysis for spectrum sharing," in *Proc. IEEE MILCOMM'06*, Washington D.C., Oct. 2006.
[7] B. Wang, Z. Ji, and K. J. R. Liu, "Primary-prioritized Markov approach for dynamic spectrum access," in *IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN'07)*, Dublin, Apr. 2007.
[8] C. Raman, R. Yates, and N. Mandayam, "Scheduling variable rate links via a spectrum server," in *IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN'05)*, pp. 110-118, Baltimore, Nov. 2005.
[9] M. J. Osborne, *An introduction to game theory*, Oxford University Press, 2004.
[10] Z. Ji, and K. J. R. Liu, "Dynamic spectrum sharing: a game theoretical overview," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 88-94, May 2007.
[11] M. M. Halldorson, J. Y. Halpern, L. Li, and V. S Mirrokni,"On spectrum sharing games," in *Proc. ACM on Principles of distributed computing*, pp. 107-114, 2004.
[12] L. Cao, and H. Zheng, "Distributed spectrum allocation via local bargaining," in *Proc. IEEE SECON 2005*, pp. 475-486, Santa Clara, Sept. 2005.
[13] J. Huang, R. Berry, and M. L. Honig, "Auction-based spectrum sharing," *ACM/Springer Mobile Networks and Apps.*, vol. 11, no. 3, pp. 405-418, June 2006.
[14] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, and S. Suri, "A general framework for wireless spectrum auctions," in *IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN'07)*, Dublin, Apr. 2007.
[15] Z. Ji, and K. J. R. Liu, "Belief-assisted pricing for dynamic spectrum allocation in wireless networks with selfish users," in *Proc. IEEE Int'l Conference on Sensor, Mesh, and Ad Hoc Commun. and Networks (SECON)*, pp. 119-127, Reston, Sept. 2006.
[16] Z. Ji, and K. J. R. Liu, "Multi-stage pricing game for collusion-resistant dynamic spectrum allocation," *IEEE J. on Select. Areas in Commun.*, vol. 26, no. 1, pp. 182-191, Jan. 2008.
[17] D. Niyato, and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: dynamic game, inefficiency of Nash equilibrium, and collusion," *IEEE J. on Select. Areas in Commun.*, vol. 6, no. 1, pp. 192-202, Jan. 2008.
[18] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE J. on Select. Areas in Commun.*, vol. 25, no. 3, pp. 517-528, Apr. 2007.
[19] S. T. Chung, S. J. Kim, J. Lee, and J. M. Cioffi, "A game-theoretic approach to power allocation in frequency-selective Gaussian interference channels," in *Proc. IEEE International Symposium on Inform. Theory*, Pacifico Yokohama, June 2003.
[20] K. W. Shum, K. K. Leung, and C. W. Sung, "Convergence of iterative waterfilling algorithm for Gaussian interference channels," *IEEE J. on Select. Areas in Commun.*, vol. 25, no. 6, pp. 1091-1100, Aug. 2007.
[21] J. Huang, R. Berry and M. Honig, "Distributed interference compensation for wireless networks," *IEEE J. on Select. Areas in Commun.*, vol. 24, no. 5, pp. 1074-1084, May 2006.
[22] D. Fudenburg, and J. Tirole, *Game theory*, MIT Press, 1993.
[23] G. Grimmett, and D. Stirzaker, *Probability and random processes*, Oxford University Press, 2001.
[24] Z. Han, and K. J. R. Liu, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*, Cambridge University Press, 2008.
[25] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *J. Operational Research Society*, vol. 49, pp. 237–252, 1998.
[26] H. A. David, *Order statistics*, New York: Wiley, 1970.
[27] F. F. Digham, M.-S. Alouini and M. K. Simon, "On the energy detection of unknown signals over fading channels," in *Proc. IEEE International Conference on Commun. (ICC'03)*, Anchorage, May 2003.
[28] W. Vickrey, "Counterspeculation, auctions, and competitive sealed tenders," *J. of Finance*, vol. 16, no. 1, pp. 8-37, Mar. 1961.
[29] F. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33-37, Jan. 1997.

**Yongle Wu** (S'08) received the B.S. (with highest honor) and M.S. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2003 and 2006, respectively. He is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Maryland, College Park.

His current research interests are in the areas of wireless communications and networks, including cognitive radio techniques, dynamic spectrum access, and network security. Mr. Wu received the Graduate School Fellowship from the University of Maryland in 2006.

**Beibei Wang** (S'07) received her B.S. degree (with highest honor) from the University of Science and Technology of China, Hefei, China, in July 2004. She is currently a PhD student in the Department of Electrical and Computer Engineering, University of Maryland, College Park.

Her research interests include resource allocation and management in cognitive radio systems, wireless communications and networking, game theory, wireless multimedia, and security. Ms. Wang received the Graduate School Fellowship from the University of Maryland in 2004–2006.

**K. J. Ray Liu** (F'03) is Distinguished Scholar-Teacher of University of Maryland, College Park. He leads the Maryland Signals and Information Group conducting research encompassing broad aspects of information technology including communications and networking, information forensics and security, multimedia signal processing, biomedical imaging, and bioinformatics.

Dr. Liu is the recipient of numerous honors and awards including best paper awards from IEEE Signal Processing Society (twice), IEEE Vehicular Technology Society, and EURASIP; IEEE Signal Processing Society Distinguished Lecturer, EURASIP Meritorious Service Award, and National Science Foundation Young Investigator Award. He also received various teaching and research recognitions from University of Maryland including university-level Invention of the Year Award, and both the Outstanding Research Award and Poole and Kent Senior Faculty Teaching Award from A. James Clark School of Engineering Faculty.

Dr. Liu is Vice President – Publications and on the Board of Governor of IEEE Signal Processing Society. He was the Editor-in-Chief of IEEE Signal Processing Magazine and the founding Editor-in-Chief of EURASIP Journal on Applied Signal Processing.

His recent books include *Cooperative Communications and Networking*, Cambridge University Press, 2008; *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*, Cambridge University Press, 2008; *Ultra-Wideband Communication Systems: The Multiband OFDM Approach*, IEEE-Wiley, 2007; *Network-Aware Security for Group Communications*, Springer, 2007; *Multimedia Fingerprinting Forensics for Traitor Tracing*, Hindawi, 2005.

**T. Charles Clancy** is a senior researcher with the Laboratory for Telecommunications Sciences, and an adjunct professor of Electrical Engineering at the University of Maryland. He received his M.S. in Electrical Engineering from the University of Illinois and Ph.D. in Computer Science from the University of Maryland. He chairs the HOKEY working group and is the security advisor to the CAPWAP and EAP working groups of the IETF. His research interests include next-generation wireless networks and security.