Distributed learning-based spectrum allocation with noisy observations in cognitive radio networks

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This article was published in the journal IEEE Transactions on Vehicular Technology [© IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works]. The definitive version is available at: http://dx.doi.org/10.1109/TVT.2014.2309120

Metadata Record: https://dspace.lboro.ac.uk/2134/23112

Version: Accepted for publication

Publisher: © IEEE

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Distributed Learning-Based Spectrum Allocation with Noisy Observations in Cognitive Radio Networks

Mahsa Derakhshani, Member, IEEE, and Tho Le-Ngoc, Fellow, IEEE

Abstract—This paper studies the medium access design for secondary users (SUs) from a game-theoretic learning perspective. In consideration of the random return of primary users, a distributed SU access approach is presented based on an adaptive CSMA scheme, in which each SU accesses multiple idle frequency slots of a licensed frequency band with adaptive activity factors. The problem of finding optimal activity factors of SUs is formulated as a potential game, and the existence, feasibility and optimality of Nash Equilibrium (NE) are analyzed. Furthermore, to achieve NEs of the formulated game, learning-based algorithms are developed in which each SU independently adjusts its activity factors. Convergence properties of best-response dynamics and log-linear dynamics are studied. Subsequently, by learning other SUs’ behavior from locally available information, the convergence with probability 1 to an arbitrarily small neighborhood of the globally optimal solution is investigated by both analysis and simulation.

I. INTRODUCTION

In hierarchical cognitive radio networks, the design of an efficient distributed medium access scheme for secondary users (SUs) is challenging due to the conflicts with licensed primary users (PUs), and hence, has received considerable attention [1]–[3]. When studying medium access design in a cognitive radio network, a key concern is how to manage spectrum sharing among SUs in a distributed manner. The simplicity, success and coexistence capability of the listen-before-talk random access scheme has the potential to be applied for the early cognitive systems [4]. For instance, Carrier Sense Multiple Access (CSMA) has been used in the literature to randomly allocate channel times among competing SUs (e.g., [5]–[7]). Furthermore, the new wireless standard IEEE 802.11af, which has been recently introduced to implement an enhanced type of Wi-Fi using TV white spaces, will also use 802.11af, which has been recently introduced to implement an opportunistic CSMA-based protocol for spectrum sharing [8]–[10].

However, a random access scheme (e.g., CSMA) suffers from low channel utilization because of inevitable contention among users. To improve the efficiency, efforts have been exerted on optimizing random access performance (e.g., [11], [12]). Regarding IEEE 802.11 MAC operation, in [13], an opportunistic CSMA (O-CSMA) scheme is proposed for a WLAN to improve throughput by exploiting multi-user diversity gain. The proposed O-CSMA adopts a channel-adaptive back-off which prioritizes the users having high-SNR, and hence, significantly improves throughput [13]. Then, [14] proposes a multi-channel O-CSMA scheme in which high-SNR channels are selected with higher probabilities for OFDMA systems. By arranging transmission on favorable channels for multiple users, the proposed O-CSMA reaches the full multi-user diversity gain despite the distributed operation.

To offer an efficient CSMA scheme which could be compatible with requirements in a cognitive radio network, in addition to channel conditions, PU activity characteristics should be taken into consideration. Furthermore, it is important to perform fair resource allocation among SUs while avoiding collision with PUs. Thus, a more cognitive random access scheme is required to consider conflicts with PUs and fairness among SUs. Accordingly, in [15], we present an adaptive multi-channel CSMA scheme in which adaptive access probabilities (called activity factor) are adopted for SUs, taking into account PU return probabilities and spectrum sharing incentives in addition to the channel quality functions. Subsequently, the SU activity factor optimization problem for maximizing the overall SU throughput is formulated and medium access algorithms are developed for SUs.

In this paper, the adaptive CSMA-based access design for SUs is studied in a game-theoretic framework. There are three key reasons for using a game-theoretic approach. First, since game theory explicitly recognizes interactions among autonomous SUs, it enables providing distributed algorithms. Second, game-theoretic algorithms empower us to accelerate the convergence in comparison with the algorithm proposed in [15]. Third, game theory offers a useful tool to predict, analyze and characterize the long-run behavior of the system, specifically in comparison with the globally optimal solution.

In the proposed game-theoretic algorithm, activity factor selections are made independently and dynamically by each SU, which aims to satisfy its own demand despite imposed sharing incentives. More specifically, the activity factor optimization
problem is formulated as an exact potential game. Via potential game framework, it is proved that the formulated game admits at least one pure-strategy NE, i.e., the stable operating point of the system. Subsequently, the feasibility and efficiency of NE are investigated. To address incomplete information about the game structure, learning approaches—which can respond optimally to the history information and achieve NE points—are studied in terms of information requirements and convergence properties. We establish the convergence of the best-response iterations to a pure-strategy NE which is not essentially the globally optimal solution.

Aiming to enable equilibrium selection, we introduce the log-linear learning process which is basically a perturbed best-response process. The log-linear learning is one of the few learning approaches which assures convergence to the most efficient NE which is globally optimal. By introducing noise into decision making process, the log-linear iterations diverge from the suboptimal NE while moving in the direction toward the globally optimal NE which is robust to noisy perturbations [17], [18].

Inspired by how the log-linear learning works, we propose a fully-distributed algorithm following best-response dynamics in which each SU adjusts its activity factors independently by learning from locally available information. Taking advantage of the noisy observations, it is shown that the best-response iterations will finally stay in a neighborhood of the globally optimal NE with probability of one. This can be explained by the fact that the globally optimal NE of the formulated game is the single stochastically stable NE. In comparison with the learning-based algorithm in [15], the proposed game-theoretic algorithm in this paper appears to have much faster convergence.

The rest of this paper is organized as follows. In Section II, we review the related work. Section III presents an overview of the system model and introduces the activity factor optimization problem. In Section IV, the medium access design is formulated as an exact potential game. Then, the existence, feasibility and efficiency of NE of the formulated game are analyzed. Section V investigates the convergence properties of learning approaches including the best-response dynamics and the log-linear dynamics. Finally, Section VI presents the concluding remarks.

II. RELATED WORK

There are several works in literature addressing distributed medium access design in hierarchical cognitive radio networks from non-game-theoretic and game-theoretic perspectives. In [5], [6], [19]–[21], the problem of channel selection is studied, assuming that both PUs and SUs have the same transmission time-slot structure. However, such an assumption of synchronous slotted transmission structure between PU and SU networks is not sensible since it needs well coordination in time between PU and SU networks. In this paper, under assumption of no time coordination between SUs and PUs, we investigate medium access design for SUs in consideration of possible collision between PU and SU due to PU return.

When studying access design under the effect of PU return, most existing works considered a single SU or a simple mechanism for random sharing among SUs without taking into account the competition or coordination among SUs [7], [22]. However, a key challenge is how to share opportunities among SUs to achieve a network-level objective. In this paper, some sharing incentives among SUs are applied to avoid channel degradation due to crowding effects. This coordination can lead to higher network performance compared to when each SU picks actions independently and competes randomly to capture the channel.

To satisfy the imposed sharing constraints in a distributed manner, a game-theoretic approach is studied and learning algorithms, including best-response dynamics and log-linear dynamics, are investigated to reach to the globally optimal solution. In the spectrum sharing literature, there are a few works studying log-linear dynamics to learn equilibrium points. For instance, [23] studies log-linear equilibriums (or logit equilibriums) for a stochastic game and [24] presents a log-linear learning algorithm for a strategic game with a unique pure-strategy NE. However, in this paper, log-linear learning is introduced and investigated for a strategic potential game with multiple pure-strategy NEs aiming to enable equilibrium selection.

Subsequently, with the aid of an adaptive CSMA scheme, the fully-distributed learning-based access algorithm is proposed in which each SU can locally and independently update its own strategy. The proposed access algorithm distinguishes this paper in coordinating spectrum access among SUs from the previously proposed channel assignment schemes for SUs with support of a central controller (e.g., [25]), exploiting a common control channel (e.g., [26]) or a simple random access (e.g., [5]).

III. SYSTEM MODEL

We study a cognitive radio network with $N_s$ SUs looking for temporal spectrum availabilities in $N_c$ frequency slots (or channels), licensed to PUs. It should be noted that these channels can be considered as narrowband subcarriers in a multicarrier modulation scheme (e.g., OFDM). Each SU is assumed to independently follow its own transmission time frame that is simply divided into equal time intervals; each interval of duration $T$ has 2 phases: spectrum-sensing phase of $\tau$ followed by spectrum-access phase of $(T - \tau)$. Although each SU has the same values of $T$ and $\tau$, time synchronization is not necessarily required among different SUs since spectrum sharing in each channel is based on a random-access scheme, i.e., the adaptive CSMA scheme. In this paper, we assume sufficiently accurate sensing with negligible PU miss-detection, focusing on the effects of random PU returns. Perfect sensing could be a sensible assumption in certain scenarios, e.g., applications in which SUs are located inside the service area of the PU transmitter (e.g., [27]).

Let $\mathcal{N}_a := \{1, \ldots, N_a\}$ be the set of $N_a$ channels that are detected idle in the spectrum-sensing, and hence, can be utilized by $N_s$ SUs and $C_{k}^{i}$ be the channel quality function for SU $k$ in channel $i$. A simple example of the channel quality is the transmission capacity represented by $C_{k}^{i} = B^i \log (1 + P_{k} g_{k}^{i}/n_{i}^{k})$ where $B^i$ denotes the bandwidth, $g_{k}^{i}$
represents the channel power gain, $P_k$ denotes the signal power and $n_k$ represents the noise power for the SU $k$ in the channel $i$. Note that we assume block flat fading in which $g_{k,k}$ remains unchanged during a given time-interval but independently varies from one time-interval to another.

From the SU viewpoint, the PU activity in a given channel can be modeled as a two-state continuous-time random process with the OFF (0) and ON (1) states representing the idle and busy periods of the PU. Hence, there will be a non-zero probability of PU return (i.e., $\alpha_i$) during a spectrum-access phase to the channel $i$ while it was detected idle by SUs in the spectrum-sensing phase.

Regarding spectrum sharing, it is assumed that SUs share idle channels using an adaptive CSMA approach. In the proposed CSMA, each SU $k$ enters a competition to access an idle channel $i \in N_a$ during a spectrum-access phase with a certain probability (called intended activity factor), $\beta_k (0 \leq \beta_k \leq 1)$, to be determined based on channel qualities, PU return probabilities and sharing incentives. Adaptive activity factors enable prioritizing SUs who gain most from using a channel, and hence, improving channel utilization in comparison with a simple random access scheme. Note that $\beta_k$ is restricted to a binary value in previous studies [5]–[7], [19]–[21], such that $\beta_k = 0$ expresses that the SU $k$ does not transmit in the channel $i$ and $\beta_k = 1$ represents that the SU $k$ transmits in the channel $i$ for the entire spectrum-access phase. Therefore, the proposed transmission scheme with $0 \leq \beta_k \leq 1$ generalizes the existing schemes.

More specifically, in the proposed CSMA scheme, each spectrum-access phase is divided into $S$ equal sub-slots with length $\frac{T_{\text{OFF}}}{S}$, labeled $t_1, \ldots, t_S$. For every sub-slot $t_j$ in channel $i$, the SU $k$ performs the following steps:

1- Generate a Bernoulli random variable $x_k^i(t_j)$ with success probability $\beta_k^i$ to be determined. If $x_k^i(t_j) = 0$, the SU $k$ will not transmit in subslot $t_j$. If $x_k^i(t_j) = 1$, the SU $k$ will proceed to the next step.

2- Generate a back-off time $\delta_k^i(t_j)$ according to a uniform distribution in the interval $(0, \delta_{\text{max}})$.

3- After expiry of back-off time, sense the channel $i$, if it is idle, transmit.

In the proposed CSMA scheme, one SU with the smallest back-off time among the SUs who compete for the same sub-slot (i.e., $x_k^i(t_j) = 1$) will succeed and transmit in this sub-slot. Let $y_k^i(t_j)$ be a binary random variable representing the capturing status, i.e., $y_k^i(t_j) = 1$ if the SU $k$ captures channel $i$ in sub-slot $t_j$; otherwise, $y_k^i(t_j) = 0$.

Assuming that each SU $k$ keeps track of its capturing status feedbacks, $y_k^i(t_j)$, the achieved activity factor of SU $k$ in the channel $i$ (i.e., the average time proportion of a spectrum-access phase that a SU successfully occupies the channel $i$, given the competition among SUs) is obtained as [15],

$$\bar{\beta}_k^i = \frac{E \left[ \sum_{j=1}^{S} x_k^i(t_j) y_k^i(t_j) \right]}{S}$$  \hspace{1cm} (1)

Then,

$$\bar{\beta}_k^i = P[y_k^i(t_j) = 1, x_k^i(t_j) = 1]$$

Fig. 1: Probability of getting the smallest back-off time vs. number of SUs $N_a$

\begin{equation}
P[y_k^i(t_j) = 1|x_k^i(t_j) = 1] = \frac{1}{1 + \sum_{k=1, k \neq k}^{N_a} x_k^i(t_j)} \hspace{1cm} (2)
\end{equation}

Applying the Jensen’s inequality, $P[y_k^i(t_j) = 1|x_k^i(t_j) = 1] \geq \frac{1}{1 + \sum_{k=1, k \neq k}^{N_a} \beta_k^i}$.

The above expression indicates that the achieved activity factor of each SU using the proposed CSMA scheme depends on the intended activity factors of other SUs.

Since the SU $k$ in channel $i$ transmits partially with achieved activity factor $\bar{\beta}_k^i$, its normalized transmission rate is $\bar{\beta}_k^i C_k^i$. Subsequently, the throughput of SU $k$ in the idle channel $i$, is defined as its successful transmission rate $f_k^i = \bar{\beta}_k^i C_k^i (1 - \bar{\beta}_k^i \alpha_i)$ where $\bar{\beta}_k^i \alpha_i$ expresses the probability that the SU $k$ experiences transmission loss due to PU return in idle channel.
Throughput (bits/s/Hz) which reflects channel conditions as well. Instead, in this paper, we consider the channel quality function channel in one time-slot is proportional to its bandwidth. Taking the derivative of $f_k$ with respect to $s_k$, we have

$$\frac{\partial f_k}{\partial s_k} = \frac{\beta_k C_k (2 \beta_k \alpha_i - 1 - s_k)}{(1 + s_k)^3}$$

Based on (5), it is clear that $f_k$ reaches its maximum value at $s_k^* = \max(0, 2 \beta_k \alpha_i - 1)$. Since $-1 \leq 2 \beta_k \alpha_i - 1 \leq 1$, it can be concluded that optimal value of $s_k^*$ happens between 0 and 1, i.e., $0 < s_k^* < 1$. In other words, $f_k$ is a decreasing function of $s_k$ for $s_k > 1$. This means that the achievable throughput of each individual SU decreases, while the contention among SUs is increasing (for $s_k > 1$).

Thus, to control contention among SUs and achieve higher throughput for each individual SU, we limit the sum of all activity factors of different SUs in each idle channel $i$ to one, i.e., $\sum_{k=1}^{N_s} \beta_k^i \leq 1, \quad i = 1, \ldots, N_a$. By keeping $\sum_{k=1}^{N_s} \beta_k^i \leq 1, \quad i = 1, \ldots, N_a$, it can be guaranteed that $s_k^* \leq 1, \quad k = 1, \ldots, N_a, \quad i = 1, \ldots, N_a$. Furthermore, the sum of all activity factors of each SU $k$ over all idle channels $i = 1, \ldots, N_a$ is set equal to one, i.e., $\sum_{i=1}^{N_a} \beta_k^i = 1$. Generally, these constraints enable us to preserve fairness among different SUs.

Aiming to determine the optimal activity factors, the optimization problem is formulated to maximize the overall throughput of all SUs under constraints of $\sum_{k=1}^{N_s} \beta_k^i = 1, \quad i = 1, \ldots, N_a$. However, to reduce computational complexity, we set the utility function as $U_1 = \sum_{i=1}^{N_s} \sum_{k=1}^{N_a} \beta_k^i C_k (1 - \beta_k^i \alpha_i)$ while keeping $\sum_{k=1}^{N_s} \beta_k^i \leq 1$, instead of $U_2 = \sum_{k=1}^{N_s} \sum_{i=1}^{N_a} \beta_k^i C_k (1 - \beta_k^i \alpha_i)$. More specifically,

$$\max_{\beta} \quad \sum_{k=1}^{N_s} \sum_{i=1}^{N_a} \beta_k^i C_k (1 - \beta_k^i \alpha_i)$$

subject to

$$\sum_{k=1}^{N_s} \beta_k^i \leq 1, \quad i = 1, \ldots, N_a$$

$$\sum_{i=1}^{N_a} \beta_k^i = 1, \quad k = 1, \ldots, N_s$$

$$0 \leq \beta_k^i \leq 1, \quad i = 1, \ldots, N_a, \quad k = 1, \ldots, N_s$$

The formulated activity factor optimization problem is feasible if $N_s \leq N_a$. Note that this paper introduces the successful transmission rate to establish the utility function of each SU in each channel. However, most existing works, e.g., [5], [6], [19]–[21], ignore the effects of random PU returns on the SU transmission and the utility function of each SU in each channel is defined as its transmission rate. Furthermore, in previous studies, [5], [6], [19]–[21], the transmission rate of each SU in each channel is simply defined as the channel bandwidth assuming that number of bits delivered over a channel in one time-slot is proportional to its bandwidth. Instead, in this paper, we consider the channel quality function which reflects channel conditions as well.

In order to investigate the effect of choosing $U_1$ along with $\sum_{k=1}^{N_s} \beta_k^i \leq 1$ instead of $U_2$ to find optimal activity factors, we plot the overall throughput of all SUs (i.e., $\sum_{k=1}^{N_s} \sum_{i=1}^{N_a} \beta_k^i C_k (1 - \beta_k^i \alpha_i)$) versus $\alpha_i$ and $N_s$ for both scenarios (i.e., using $U_1$ and $U_2$) in Figures 2 and 3. In this example, we assume independent channels with the same bandwidth $B^i = 1$ and same $\alpha_i$. Please note that we used exhaustive search to find the optimal activity factors while optimizing $U_2$. It is shown that the throughput improvement by using $U_2$ instead of $U_1$ (along with $\sum_{k=1}^{N_s} \beta_k^i \leq 1$) to find optimal activity factors is small. This fact can be explained by the effective role of keeping $\sum_{k=1}^{N_s} \beta_k \leq 1$ to reduce competition among SUs. On the other hand, using $U_2$ makes the optimization problem non-convex and computationally complex. It is worth mentioning that we have studied the activity factor allocation problem based on $U_2$ in [28]. Since the optimization problem is generally non-convex, the medium access design is cast in a game theoretic framework. It is shown that the NE of the formulated game can be highly inefficient in terms of total profit for all SUs. However, by setting the utility function as $\sum_{k=1}^{N_s} \sum_{i=1}^{N_a} \beta_k^i C_k (1 - \beta_k^i \alpha_i)$
while keeping $\sum_{k=1}^{N_s} \beta_k^i \leq 1$, in this paper, it will be shown that the globally optimal solution is achievable through a fully-distributed game-theoretic algorithm.

IV. GAME-THEORETIC DESIGN

In a highly dynamic environment such as cognitive radio networks, it is practically essential to find a reasonably good solution which can be obtained in a sufficiently fast manner. To this end, in this section, the CSMA-based design is studied in a game-theoretic framework which enables distributed implementation and fast convergence to a reasonably good solution. More specifically, we consider a strategic non-cooperative game in which the players are SUs.

According to the optimization problem in (6), each SU could simply maximize its transmission rate (i.e., $\sum_{i=1}^{N_s} \beta_k^i C_k^i (1 - \beta_k^i \alpha_i)$). However, SUs cannot select activity factors which violate the coupling constraints in (6b). Since it is difficult for SUs to identify feasible activity factors in advance, we consider a penalty mechanism embedded in the payoff function of each SU as a cost function. The penalty mechanism is aimed to satisfy the coupling constraints among different SUs in (6b) by severely punishing the SU who violates each of them. Specifically, we focus on a linear penalty mechanism. Thus, we construct an alternative payoff function for SU $k$ as

$$u_k = \sum_{i=1}^{N_s} \beta_k^i C_k^i (1 - \beta_k^i \alpha_i) - \sum_{i=1}^{N_s} \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_j^i - 1 \right) \tag{7}$$

where $\Theta(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases}$ and $\mu_i$ is the penalty factor in channel $i$ which is a positive scalar. The second term of (7) represents the penalty mechanism in the payoff function of each SU. More specifically, we define the penalty in channel $i$ as

$$\mu_i(\sum_{k=1}^{N_s} \beta_k^i - 1), \begin{cases} \sum_{k=1}^{N_s} \beta_k^i > 1 \\ \sum_{k=1}^{N_s} \beta_k^i \leq 1 \end{cases} \tag{10}$$

In other words, on one hand, SU $k$ will not be penalized in channel $i$ if $\beta_k^i$ is chosen such that $\sum_{k=1}^{N_s} \beta_k^i \leq 1$. On the other hand, SU $k$ has to pay a penalty equal to $\mu_i(\sum_{k=1}^{N_s} \beta_k^i - 1)$ if $\beta_k^i$ is chosen such that $\sum_{k=1}^{N_s} \beta_k^i > 1$. Note that SUs need to use the same values of penalty factors, which can be easily set up during initialization.

Let $B_k = \{\beta_k^1, \ldots, \beta_k^{N_s}\}$ be the pure strategy at SU $k$ and $B_{-k}$ be the pure strategy of all SUs excluding the SU $k$. A pure strategy defines a specific action that a player can pick from its own strategy set. Such actions are not drawn from a probability distribution, as in the case of mixed strategies. Subsequently, the admissible pure strategies of SU $k$ is defined as

$$B_k = \left\{ \beta_k : \beta_k^i \in \{0, \frac{1}{S}, \frac{2}{S}, \ldots, 1\}, \forall i \in \mathcal{N}_s, \sum_{i=1}^{N_s} \beta_k^i = 1 \right\} \tag{8}$$

Note that $\beta_k^i$ takes discrete values in the proposed CSMA-based algorithm. Then, we can define the non-cooperative game for the medium access of SUs as

$$g = [\mathcal{N}_s, \{B_k\}_{k \in \mathcal{N}_s}, \{u_k\}_{k \in \mathcal{N}_s}] \tag{9}$$

where $\mathcal{N}_s = \{1, \ldots, N_s\}$ is the set of players of the game (i.e. SUs), $B_k$ is the activity factor strategy set of the SU $k$ and $u_k$ is the corresponding payoff function of the SU $k$ defined on the set of pure-strategy profiles $B = B_1 \times \ldots \times B_{N_s}$.

An exact potential game is a strategic game in which the incentive of all players to change their strategies can be expressed in a global potential function. The potential games are easy to analyze since improving each player’s utility also increases the value of a potential function [29]. In the following theorem, we demonstrate that the game $G$ falls into the framework of exact potential games.

**Theorem 1:** The formulated game $G$ is an exact potential game with the potential function,

$$\Phi = \sum_{j=1}^{N_s} \sum_{i=1}^{N_s} \beta_j^i C_j^i (1 - \beta_j^i \alpha_i) - \sum_{i=1}^{N_s} \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_j^i - 1 \right) \tag{10}$$

**Proof:** It is clear that the game $G$ satisfies the exact potential game definition [29],

$$u_k(\beta_k, \beta_{-k}) - u_k(\beta_k^*, \beta_{-k}) = \Phi(\beta_k, \beta_{-k}) - \Phi(\beta_k^*, \beta_{-k}), \forall \beta_k, \beta_k^* \in B_k, \forall k \in \mathcal{N}_s \tag{11}$$

Thus, $G$ is an exact potential game and $\Phi$ is the potential function of $G$. ■

Conceptually, a strategic game can reach a steady-state Nash Equilibrium (NE) point, if it exists, from which no player can improve its utility by changing its own strategy unilaterally [30]. In other words, a pure-strategy profile $\beta^* = \{\beta_k^*\}_{k=1}^{N_s} \in B$ is a pure-strategy NE if and only if

$$u_k(\beta_k^*, \beta_{-k}^*) \geq u_k(\beta_k^*, \beta_{-k}'), \forall \beta_k^* \in B_k, \forall k \in \mathcal{N}_s \tag{12}$$

We are interested in investigating the existence and characteristics including the feasibility and efficiency of NE of the game $G$.

A. Existence

First of all, the existence of NE of the game $G$ is studied in the following theorem based on the properties of the potential games.

**Theorem 2:** The game $G$ admits at least one pure-strategy NE.

**Proof:** This theorem comes directly from Corollary 4 in [31], which states every finite potential game $G$ has at least one pure-strategy NE. ■

It should be noted that the payoff function of each SU (i.e., $u_k$) in (7) is not a continuous function over the strategy set $B$. Thus, payoff functions are not strictly concave. As a result, it cannot be concluded that the pure-strategy NE of the formulated game $G$ is always unique.

**Remark 1:** In general, the pure-strategy NE of game $G$ may not be unique.
B. Feasibility

Since the optimization problem in (6) has the coupling constraints in (6b) which are merged in the payoff functions in the formulated game $G$, it is required to verify if an arbitrary pure-strategy NE is feasible, i.e., satisfying the constraints $\sum_{k=1}^{N_s} \beta_k \leq 1, i = 1, \ldots, N_s$. Thus, the following theorem presents conditions that assure the feasibility of pure-strategy NEs.

**Theorem 3:** All pure-strategy NEs of the game $G$ must be feasible if

$$\mu_i > \mu_{th}, \forall i \in N_a$$  \hspace{1cm} (13)

where

$$\mu_{th} = \max_{k \in N_s, \beta_k \in B_k} \left( S \sum_{i=1}^{N_s} \beta_k^i C_k^i \left( 1 - \beta_k^i \alpha_i \right) \right)$$  \hspace{1cm} (14)

**Proof:** See Appendix A for the proof.

Theorem 3 assures that, by properly designing $\mu_i$’s, the payoff functions in (7) can guarantee the feasibility of the steady states of the system.

C. Efficiency

The other aspect that we are interested to study is how efficient the NE of game $G$ is in comparison with the globally optimal solution of (6). In other words, NE efficiency is measured comparing to the globally optimal solution of (6). The following theorem specifies the relationship between the globally optimal solution and the NE of the game $G$.

**Theorem 4:** The globally optimal solution of (6) is the Pareto-optimal pure-strategy NE of $G$ if $\mu_i > \mu_{th}, \forall i \in N_a$.

**Proof:** See Appendix B for the proof.

Theorem 4 shows that the globally optimal solution constitutes a pure-strategy NE. From Theorem 4, it follows that the NE must be the globally optimal solution when the NE is unique. However, when multiple NEs exist in the formulated game $G$, a randomly chosen NE may not be the globally optimal solution. That is why we need to study different learning algorithms to find an iterative algorithm which can guarantee convergence to the most efficient NE (i.e., the globally optimal solution).

**Remark 2:** In the next section, learning-based iterative algorithms are proposed which enable the convergence to the globally optimal pure-strategy NE.

V. LEARNING EQUILIBRIA

Assuming that the rationality of players and the structure of the game are common knowledge, equilibrium can be observed as a result of analysis and introspection of the players. Otherwise, under assumption of bounded rationality or partial information, equilibrium may arise as a consequence of a long-run learning process. In this section, aiming to achieve an equilibrium of the game $G$, learning approaches are discussed in terms of information requirements and convergence properties.

A. Best-Response Dynamics

To reach an equilibrium of the game $G$, first, a simple learning algorithm for activity factor selection is presented, based on the asynchronous best-response dynamics. In particular, at each time $t \in \{0, 1, 2, \ldots\}$, exactly one SU $k \in N_s$, is randomly selected to revise its activity factors based on the best-response dynamics defined as

$$\beta_k[t] = \arg \max_{\beta_k \in B_k} u_k(\beta_k, \beta_{-k}[t-1])$$  \hspace{1cm} (15)

Under perfect knowledge of current strategies of other SUs (i.e., $\beta_{-k}[t-1]$), the convergence of the proposed game-theoretic algorithm is established in the following theorem.

**Theorem 5:** The learning algorithm under asynchronous best-response dynamics converges with probability of one to a pure-strategy NE of the game $G$ from any initial strategy point.

**Proof:** Based on Theorem 19 in [30], in a finite exact potential game, best-response dynamics will converge with probability of one to a pure-strategy NE in finite steps. Accordingly, in game $G$, the best-response iterations will converge to a pure-strategy NE.

To verify convergence of the best-response dynamics, a numerical result is also provided. In this example, we assume independent channels with the same bandwidth $B^1 = 1$ and the same $\alpha_i$. We set the similar SNR = $\frac{P_i}{n_i} = 10$dB for individual SUs. Furthermore, we assume $N_s = 3$, $N_a = 3$ and $\alpha_i = 0.1$. Note that we use the same example for all simulation results.

Figure 4 demonstrates the convergence process of the activity factors of three SUs in the first idle licensed channel for a certain channel realization. All SUs start by setting their initial values of their activity factors to zero. In each iteration, they sequentially play to optimize their own payoff functions, and update their activity factors based on the best-response dynamics in (15) (e.g., in the following order: SU3, SU2, SU1). Note that each SU updates its activity factors in three different channels simultaneously. However, due to the limited space, the simulation results for only channel 1 are provided. In iteration 1, SU3 then SU2 then SU1 play and obtain $\beta^3[1] = 0.5$, $\beta^3[1] = 0.33$, and $\beta^3[1] = 0.53$, which result in $\sum_{k=1}^{3} \beta_k^3 \geq 1$. In iteration 2, SU3 is penalized for this excessive amount, and, for its specific channel realization, Figure 4 indicates that the SU 3 is forced to reduce its activity factor to $\beta^3[2] = 0.14$ in order to keep $\sum_{k=1}^{3} \beta_k^3 \leq 1$. Subsequently, SU2 and SU1 do not need to change their strategies, and maintain $\beta^2[2] = 0.33$, and $\beta^1[2] = 0.53$. From iteration 3, since $\sum_{k=1}^{3} \beta_k^3 \leq 1, i = 1, \ldots, 3$, SUs actually optimize their own throughput, i.e., their own payoff functions become $u_k = \sum_{i=1}^{3} \beta_k^3 C_k(1 - \beta_k^3 \alpha_i)$. Thus, their strategies do not change any more. Note that, for different channel realizations, it is possible that all SUs change their activity factors to satisfy $\sum_{k=1}^{N_s} \beta_k^3 \leq 1$ in iteration 2. Furthermore, the process of reducing $\sum_{k=1}^{N_s} \beta_k^3$ may take more than two iterations. Additionally, Figure 5 shows the convergence of potential function. As evident from Figures 4 and 5, in less than 10 iterations, activity factor selection converges to a NE.
Potential function
Activity factor ($k$)

B. Log-Linear Dynamics

According to Theorem 5, the best-response iterations will converge to a pure-strategy NE which is not necessarily the globally optimal solution of (6). Since NE could be highly inefficient with regards to the network-level objective, it is thus crucial to find a learning process to reach the globally optimal solution.

B. Log-Linear Dynamics

Since the basic myopic best-response dynamics suffer from multiple rest points (e.g., any pure-strategy NE), it is essential to introduce a learning process that can select an appropriate equilibrium. Aiming to enable equilibrium selection, the log-linear learning has been proposed as a perturbed best-response process which guarantees convergence to the most efficient NE for potential games [17], [18]. The basic idea behind the equilibrium selection in log-linear learning is to introduce noise into decision making process which enables categorizing equilibria based on their stability characteristics. This noise allows players to select suboptimal actions (i.e., non-best responses) with a certain probability which is attached with the magnitude of the payoff difference of the best response and the suboptimal action.

In the log-linear learning, SUs are assumed to be myopic and boundedly rational. At each time $t > 0$, exactly one SU $k \in N_s$ is randomly selected to update its action, using a probability distribution over its strategy set in response to the current strategy profile. Let $k$ be the player chosen at time $t$ to revise its action. Then, SU $k$ will choose action $\beta_k$ given the current strategy profile $\beta_{-k}[t-1]$ with a probability based on the log-linear choice rule

$$
\pi_{\beta_k}[t] = \frac{e^{\frac{1}{\epsilon} u_k(\beta_k, \beta_{-k}[t-1])}}{\sum_{\beta'_k \in B_k} e^{\frac{1}{\epsilon} u_k(\beta'_k, \beta_{-k}[t-1])}}
$$

where $0 < \frac{1}{\epsilon} < \infty$. The scalar $\frac{1}{\epsilon}$ can be interpreted as the level of rationality of the SUs. In other words, $\epsilon$ shows the level of noise in the SUs’ decisions and determines how often SUs choose their best responses. The described rule is called log-linear since the log-likelihood ratio of selecting between two actions is linearly proportional to the difference of corresponding payoffs, given other SUs’ actions [32]. As $\epsilon \to 0$, the log-linear rule approaches to the best-response rule. However, the SU $k$ will choose any action $\beta_k \in B_k$ with equal probability as $\epsilon \to \infty$. Thus, for any $0 < \epsilon < \infty$, SUs explore non-best responses with non-zero probabilities which are exponentially smaller for actions yielding smaller payoffs [17], [18].

Our goal is to characterize the long-run behavior of asynchronous log-linear learning process for game $G$. To this end, the log-linear dynamic adjustment process is represented as an irreducible and aperiodic Markov chain $\{X_t\}_{t \in \mathbb{N}}$ on the set of strategy profiles of the game [17], [18]. Subsequently, based on the study in [17], the stationary distribution, i.e., limiting distribution in a Markov chain, is studied to explain the long-run behavior of this update rule. Accordingly, the stationary distribution of asynchronous log-linear learning process in game $G$ is presented as

$$
\mu^*(\beta) = \frac{e^{\frac{1}{\epsilon} \Phi(\beta)}}{\sum_{\beta' \in B} e^{\frac{1}{\epsilon} \Phi(\beta')}}
$$

According to (17), $\mu^*(\beta)$ (i.e., the probability that $X_t = \beta$ for sufficiently large times $t > 0$) can be expressed as an explicit function of the potential function. A strategy profile $\beta = \{\beta_k\}_{k=1}^{N_s}$ is said to be stochastically stable if $\lim_{t \to 0} \mu(\beta) > 0$. Consequently, the stochastically stable strategy can be computed based on (17).

Corollary 1: In game $G$, the only stochastically stable strategy profile of asynchronous log-linear learning process is the maximizer of potential function $\Phi$ which is the NE of the potential game $G$ based on Theorem 2 in [29]. Assuming $\mu_i \geq \mu_{ih}, \forall i \in N_c$, the globally optimal solution of (6) is equal to the maximizer of the potential function. As a result, the globally optimal solution of (6) is the only stochastically stable NE.

The importance of this Corollary—which comes directly from Corollary 1 in [17] for exact potential games—is two-fold, i.e., convergence and equilibrium selection. In particular, this
Potential function $\beta_i$ Activity factor $(\beta i)_k$ in [33], assuming $\mu_i > \mu_{th}, \forall i \in N_a$ with sufficiently large rationality level. In other words, by properly designing the rationality level, log-linear dynamics will ultimately converge to the globally optimal NE in game $G$.

To verify the convergence properties of log-linear learning, Figures 6 and 7 show the convergence process of the activity factors of three SUs in the first idle licensed channel for a certain channel realization and the convergence of the potential function, respectively. They confirm that the log-linear iterations lead to the globally optimal solution despite the best-response dynamics. As can be observed, it takes merely around 20 iterations to quickly converge to the globally optimal solution. The other aspect which affects the practicality of a learning process is the convergence speed. From Proposition 152 in [33], assuming $\mu_i > \mu_{th}$, the convergence time of the log-linear learning to be $\eta$-close to the globally optimal solution is in the order of

$$N_s \log \log(N_s) + \log \left( \frac{1}{\eta} \right)$$  \hspace{1cm} (18)

for any initial condition if the rationality level (i.e., $\frac{1}{\eta}$) is sufficiently large. According to (18), the convergence time is linearly proportional to the number of SUs using the log-linear learning in game $G$.

Note that, in the proposed adaptive CSMA scheme, a player (SU) updates its strategy for its iterative learning algorithm (using best-response or log-linear dynamics) after a successful transmission and does not need any coordination with other SUs.

Considering the proposed best-response dynamics and log-linear dynamics, each SU (e.g., SU $k$) needs the sum of current activity factors of all other SUs in each channel (i.e., $s_k^l[t] = \sum_{k=1, k \neq k}^{N_s} \beta_k^l[t-1], i = 1, \ldots, N_a$) to update its own strategy. Up to this point, we assume that SU $k$ can obtain $s_k^l[t]$ accurately, with the aim of a central coordinator or an information exchange between SUs, i.e., SU $k$ has perfect knowledge of $s_k^l[t], i = 1, \ldots, N_a$.

In the next section, to relieve the need for frequent information exchange in the proposed learning algorithms, an alternative algorithm has been proposed based on best-response dynamics using locally available information. In the proposed algorithm, SU $k$ estimates $s_k^l[t]$ based on the CSMA feedbacks. Since this estimation with limited samples suffers from random errors, each SU would have only noisy estimation of $s_k^l[t]$ in the proposed algorithm. Then, considering such noisy information, convergence of the proposed iterative algorithm following best-response dynamics is studied. Since each SU can locally and independently update its own strategy without any information exchange among SUs, this proposed iterative algorithm can be implemented in a distributed manner.

VI. BEST-RESPONSE DYNAMICS WITH NOISY OBSERVATIONS

In [15], we study how to use the capturing status feedbacks of CSMA scheme, $y^c_k(t_j)$, to estimate the sum of activity factors of all SUs in each channel. It is shown that $\beta^t = \sum_{k=1}^{N_a} \beta_k^t$ can be updated after each window of $S'$ sub-slots as

$$\hat{\beta}^t \simeq \beta_k^t + \left( \frac{s_k^t \cdot \beta_k^t}{\sum_{i=1}^{S'} y_i^c(t_i)} \right) - 1.$$  \hspace{1cm} (15)

Since estimation with a limited number of samples suffers from random errors, it is shown that $\hat{\beta}^t = \beta^t + w$ where $E[w]$ and $\text{var}[w]$ are of $O \left( (S')^{-1} \right)$ [15].

From (7), it is clear that the estimation noise of $\beta^t$ will cause a bias (i.e., $b = \max(0, w)$) in $u_k$, and hence, best-response iterations in (15) will also involve random errors. Since the first derivative of $u_k$ is finite, the bias and variance of the random noise in best-response iterations should be also of $O \left( (S')^{-1} \right)$.

As shown in log-linear learning, adding noise to the decision making process enables equilibrium selection taking advantage of the fact that the globally optimal NE is the only stochastically stable NE in potential games. Accordingly, in Theorem 3 of [34], it is shown that a bounded noise will asymptotically ensure the convergence of the best-response iterations to a neighborhood of the globally optimal solution in potential games. That is because suboptimal NE points are less stable than the globally optimal NE (i.e., global optimum) in a
sense that a small noise can cause the best-response iterations diverge from the suboptimal NE while moving in the direction toward the globally optimal NE.

Similarly, in the proposed algorithm, best-response iterations involve errors although they are random with bounded bias and variance. With a sufficiently large or increasing estimation window (i.e., \( S' \)), the random noise can be approximated as a bounded noise. Therefore, it is expected that the best-response iterations converge to the global optimum. Mathematically, this can be presented as the following claim.

**Claim 1:** \( \forall \varepsilon > 0 \), an estimation window size can be selected (i.e., \( \exists S' > 0 \)) such that

\[
\lim_{t \to \infty} \inf \Phi(\beta[t]) \geq \Phi_{\max} - \varepsilon \quad \text{with probability of 1.}
\]

This claim declares that, by properly designing an estimation window size (i.e., \( S' \)), best-response iterations can get arbitrarily close to the globally optimal solution of (6) which is also the maximizer of the potential function (i.e., \( \Phi \)), assuming \( \mu_i > \mu_{th}, \forall i \in N_a \).

To confirm convergence of noisy best-response iterations to the globally optimal solution, Figure 8 demonstrates the convergence process of the activity factors of three SUs in the first idle licensed channel. In addition, the convergence process based on the proposed learning-based non-game-theoretic algorithm in [15] is illustrated in Figure 8. It is clear that the game-theoretic algorithm accelerate the convergence in comparison with the algorithm proposed in [15]. Furthermore, Figure 9 shows the convergence of the potential function. They confirm that the best-response iterations will stay in a neighborhood of the global optimum.

**VII. CONCLUSIONS**

In this paper, using a potential game model, we developed a distributed adaptive CSMA-based access design in which activity factors of SUs over multiple idle channels are adaptively adjusted. Via potential game framework, it is established that the formulated game admits at least one pure-strategy NE. In consideration of coupled constraints among SUs, sufficient conditions are presented to assure the feasibility of pure-strategy NEs. In addition, it is proved that the globally optimal solution is the Pareto-optimal NE.

Furthermore, the convergence properties of the best-response dynamics and log-linear dynamics of the formulated game are investigated. Assuming perfect knowledge of moves previously made by all SUs, it is proved that best-response iterations converge to a pure-strategy NE which is not essentially the globally optimal solution. However, the log-linear process enables equilibrium refinement and convergence to the most socially desirable solution. Subsequently, in a game with noisy observations, it is shown that best-response iterations also converge with probability of one to a neighborhood of the global optimum.

**APPENDIX**

A. Proof of Theorem 3

Suppose that \( \beta = \{ \beta_k \}_{k=1}^{N_s} \), where \( \beta_k \in B_k \) is a pure-strategy NE of game \( G \), but it is not feasible, i.e., it violates at least one of the constraints in (6b). Further, suppose that \( m \) is the index of channel which has the most severe violation, i.e., \( m = \arg \max_{i \in N_a} \{ \sum_{j=1}^{N_s} \beta_{ij} - 1 \} \). Thus, \( \sum_{j=1}^{N_s} \beta_{jm} > 1 \). Considering \( k \in N_s \) such that \( \beta_k^m > 0 \), from (7), we have

\[
u_k(\beta_k, \beta_{-k}) = \sum_{i=1}^{N_s} \beta_{ik} C_k^i (1 - \beta_k^m \alpha_i) - \mu_m \left( \sum_{j=1}^{N_s} \beta_{jm} - 1 \right)
- \sum_{i=1,i \neq m}^{N_s} \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_{ij} - 1 \right)
\]

Assuming that \( \sum_{j=1}^{N_s} \beta_{jm}^m = 1 + \epsilon \) where \( \epsilon \neq 0 \), there must exist a channel \( n \) such that \( \sum_{j=1}^{N_s} \beta_{jn}^m \leq \max(0, 1 - \epsilon) \) since \( N_a \geq N_s \). Then, an alternative admissible strategy \( \tilde{\beta}_k = \left[ \tilde{\beta}_k^i \right]_{i=1}^{N_s} \) can be constructed such that

\[
\tilde{\beta}_k^i = \begin{cases} 
\beta_k^i, & i \neq m, n \\
\beta_k^i - \delta, & i = m \\
\beta_k^i + \delta, & i = n
\end{cases}
\]

where \( \delta \) is a non-zero discrete
value (i.e., $\delta \in \{\frac{1}{2}, \frac{2}{3}, \ldots, 1\}$) satisfying $\delta \leq \min(1, \epsilon)$, $0 \leq \beta_k^m - \delta \leq 1$ and $0 \leq \beta_k^m + \delta \leq 1$. Note that $\sum_{j=1}^{N_s} \beta_j^m - 1 \leq \delta \leq \min(1, \epsilon)$, and hence, $\Theta \left( \sum_{j=1}^{N_s} \beta_j^m - 1 \right) = \Theta \left( \sum_{j=1}^{N_s} \beta_j^m - 1 \right)$. Then, $u_k(\tilde{\beta}_k, \beta_{-k})$ can be obtained as

$$ u_k(\tilde{\beta}_k, \beta_{-k}) = \sum_{i=1}^{N_s} \tilde{\beta}_i^m C_i^k (1 - \tilde{\beta}_i^m) - \sum_{i=1}^{N_s} \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_j^m - 1 - \delta \right) $$

$$ - \sum_{i=1, i \neq m}^{N_s} \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_j^m - 1 \right) \tag{20} $$

Thus, from (19) and (20),

$$ u_k(\tilde{\beta}_k, \beta_{-k}) - u_k(\tilde{\beta}_k, \beta_{-k}) = \sum_{i=1}^{N_s} \tilde{\beta}_i^m C_i^k (1 - \tilde{\beta}_i^m) - \sum_{i=1}^{N_s} \beta_i^m C_i^k (1 - \beta_i^m) $$

$$ + \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_j^m - 1 - \delta \right) \tag{21} $$

Since $\delta \leq \epsilon$, we have

$$ \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_j^m - 1 - \delta \right) - \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_j^m - 1 \right) = \mu_i \Theta (\epsilon - \delta) - \mu_i \epsilon \leq -\mu_i \delta \tag{22} $$

Subsequently, from (21) and (22),

$$ u_k(\tilde{\beta}_k, \beta_{-k}) - u_k(\tilde{\beta}_k, \beta_{-k}) \leq \sum_{i=1}^{N_s} \tilde{\beta}_i^m C_i^k (1 - \tilde{\beta}_i^m) - \sum_{i=1}^{N_s} \beta_i^m C_i^k (1 - \beta_i^m) $$

$$ \leq \sum_{i=1}^{N_s} \tilde{\beta}_i^m C_i^k (1 - \tilde{\beta}_i^m) - \sum_{i=1}^{N_s} \beta_i^m C_i^k (1 - \beta_i^m) \tag{23} $$

Under assumption of $\mu_i > \mu_{th}, \forall i \in N_o$, from (14), we have

$$ \sum_{i=1}^{N_s} \tilde{\beta}_i^m C_i^k (1 - \tilde{\beta}_i^m) > \sum_{i=1}^{N_s} \beta_i^m C_i^k (1 - \beta_i^m) \tag{24} $$

Since $\delta \geq \frac{1}{2}$, from (24), it can be concluded that

$$ \sum_{i=1}^{N_s} \beta_i^m C_i^k (1 - \beta_i^m) - \mu_i \delta < 0 \tag{25} $$

Thus, based on (23) and (25),

$$ u_k(\beta_k, \beta_{-k}) - u_k(\tilde{\beta}_k, \beta_{-k}) \leq 0 \tag{26} $$

This contradicts the assumption that $\beta = \{\beta_k\}_{k=1}^{N_s}$ is a pure-strategy NE of game $G$ according to the definition of NE in (12). Thus, $\beta$ is not a pure-strategy NE of $G$. Hence, it can be concluded that all pure-strategy NE must be feasible if $\mu_i > \mu_{th}, \forall i \in N_o$. ■

B. Proof of Theorem 4

Let assume that $\beta = \{\beta_k\}_{k=1}^{N_s}$ is the optimal solution of (6). Assuming $\mu_i > \mu_{th}, \forall i \in N_o$, $\beta$ is the maximizer of the potential function $\Phi$. Based on Theorem 2 in [29], the maximizer of the potential function is the NE of the potential game. Hence, $\beta$ is the NE of the game $G$.

Subsequently, we need to establish that $\beta$ is the Pareto-optimal NE. Let assume that $\beta$ is not Pareto optimal, and then, there exists an arbitrary strategy profile $\beta' = \{\beta_k\}_{k=1}^{N_s}$ such that

$$ u_k(\beta_k', \beta_{-k}) \geq u_k(\beta_k, \beta_{-k}), \forall k \in N_s, k \neq j \tag{27} $$

and, for some $j$,

$$ u_j(\beta_j', \beta_{-j}) > u_j(\beta_j, \beta_{-j}) \tag{28} $$

As a result,

$$ \sum_{k=1}^{N_s} u_k(\beta) > \sum_{k=1}^{N_s} u_k(\beta) \tag{29} $$

Since $\beta$ is a NE of the game $G$, it is feasible based on Theorem 3 (i.e., $\sum_{i=1}^{N_s} \mu_i \Theta \left( \sum_{j=1}^{N_s} \beta_j^m - 1 \right) = 0$). Then,

$$ \Phi(\beta) = \sum_{k=1}^{N_s} u_k(\beta) \tag{30} $$

Furthermore, considering that $\mu_i$’s are positive scalars, from (7) and (10), for an arbitrary $\beta'$, we have

$$ \Phi(\beta') > \sum_{k=1}^{N_s} u_k(\beta') \tag{31} $$

Consequently, based on (29), (30) and (31),

$$ \Phi(\beta') > \Phi(\beta) \tag{32} $$

This contradicts the fact that $\beta$ is the maximizer of the potential function $\Phi$. Thus, the optimal solution of (6) is the Pareto-optimal pure-strategy NE of the game $G$. ■

REFERENCES


