

Distributed Spectrum Sharing Games Via Congestion Advertisement

Mahdi Azarafrooz* and R. Chandramouli

Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ

Abstract

Distributed spectrum sharing via congestion advertisement is modelled and studied as a game theoretic problem. A related graphical anti-coordination game problem and a suitable logit-response learning mechanism is proposed and studied. It has been shown that introducing an arbitrary small congestion advertisement term into the users utility can improve the convergence rate of the spectrum sharing game exponentially. Finally, simulation results are presented to evaluate the price of anarchy, convergence rate and phase transition properties.

Received on 17 March 2013; accepted on 25 June 2013; published on 14 July 2014

Keywords: Spectrum Sharing, Game Theory, Learning

Copyright © 2014 M. Azarafrooz and R. Chandramouli, licensed to ICST. This is an open access article distributed under the terms of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/ws.1.2.e1

1. Introduction

Cognitive radio (CR) nodes learn to configure their transmission and reception parameters based on different cognitive processes. These cognitive processes vary from sensing an existing wireless channel, configuring a radio's parameters to accommodate the perceived wireless channel, evaluating the current situation and taking the best possible action based on this available knowledge [2], etc. Therefore, every action of a specific cognitive radio or user has an effect on the other nodes' payoffs.

In a recent work [3], we addressed the problem of distributed optimization of secondary user sharing of primary user spectrum, considering spatial re-use. This was modelled as a spatial or graphical game theoretic problem considering the radio interference induced by communication in a local neighborhood in a specific band. However as it will be shown later in this paper, this spectrum sharing game suffers from a high price of anarchy. Also, the distributed iterative algorithm to compute the equilibrium strategies for all the users is slow to converge. Therefore, we investigate the feasibility of computing a *good* (to be made precise later) equilibrium solution in polynomial steps (in the number of secondary nodes n).

Mechanism design is a tool that can be used to align incentives of the users with the system's objective. In systems where there are multiple Nash equilibria, using mechanism design, a central authority could move the system's behavior from a less efficient equilibrium to a more efficient one by promoting better user behavior. The objective of this paper is to investigate such a mechanism design and an iterative technique to compute an efficient Nash equilibrium solution with fast convergence properties. In this regard, we propose the idea of congestion advertisement by base stations as one of the mechanism design approaches.

In [5], spectrum sharing and spatial reuse in a wireless network is posed as an extended form of the congestion game where users' payoffs for using a spectrum band or channel is a function of the number of its interfering users sharing that channel. In [6], spectrum management is studied in CR networks by defining a secondary user specific utility as a function of the spectrum opportunity, congestion and bandwidth. The behavior of selfish nodes that dynamically switch their channels using broadcasted random public signal is presented in [7]. In [8], dynamic spectrum access is modelled as a minority game where the CR nodes try to minimize their cost in finding a clear band. A graphical game model for competitive spectrum access is discussed in [9].

*Corresponding author. Email: mazarafr@stevens.edu

This paper is directly related to [10-12]. In [10], the convergence rates of congestion games towards a good equilibrium is studied. Convergence of coordination games in a social networking context is presented in [11]. In [12], a framework for graphical games with global interactions is considered.

The main contributions of the paper are:

- The dynamics of the graphical spectrum sharing game is mapped to that of anti-ferromagnetic Ising model using a MECE logit-response learning mechanism.
- The effect of the graph of interaction on the convergence rate of the spectrum sharing games is studied both in terms of social welfare optimization and the convergence rate of the game.
- It has been shown that introducing an arbitrary small congestion advertisement term into the secondary users utility can improve the convergence rate exponentially.

This paper is organized as follows. In Section II a graphical anti-coordination spectrum sharing game model is proposed. A maximum entropy correlated logit response mechanism is discussed in Section III. Analysis of convergence rates of the response mechanism for specific graphs are presented in Section IV. Congestion advertisement for spectrum sharing is discussed in Section V. Simulation results are given in Section VI and conclusions in Section VII.

2. Spectrum Sharing as a Graphical Anti-Coordination Game

Consider a CR network scenario where n secondary users are placed in an undirected graph $G = (V, E)$, where $|V| = n$ and E is the set of edges. Let \mathcal{N}_i denote the neighbor set of node i . We will interchangeably use the terms node and user throughout the paper. Users are assumed to have access to B primary user bands. Let $\mathbf{A}_{1 \times n}$ be the users' action vector where the i th element $a_i \in \{1, \dots, B\}$ denotes the index of the spectrum band that user i is active in. Users can follow different approaches for evaluating the spectrum quality, e.g., based on whether data or video application needs to be supported. We assume that the evaluation approach is the same for all the users. Let $\Theta_{1 \times B}$ represent the spectrum quality vector, i.e., $\theta_l, l \in \{1, 2, \dots, B\}$ denotes the quality of the l^{th} spectral band. For example, this could be a function of the primary user activity, required data rate, etc. The higher the value of θ_l the more desirable that band is. let \mathcal{I}_{a_i} be the set of interfering transmissions with user i scheduled in band a_i .

The secondary users compete for spectrum opportunities in a decentralized non-cooperative manner. The utility obtained by secondary user i is $U^i(\mathcal{I}_{a_i}, \theta_{a_i})$. That is, the utility function depends on the interference level as well as the quality of the operating band. However from the perspective of the designer of a wireless cognitive network it is important that the system as a whole entity can achieve a good operation point. A simple metric for example is social welfare which is defined as the accumulation of all users utility in the network i.e.,

$$U(\mathbf{A}, \Theta) = \sum_i U^i(\mathcal{I}_{a_i}, \theta_{a_i}) \quad (1)$$

The optimal solution \mathbf{A}^* to the spectrum sharing problem is then given by:

$$\mathbf{A}^* = \underset{\mathbf{A}}{\operatorname{argmax}} U(\mathbf{A}, \Theta) \quad (2)$$

We first address the issue of solving this problem when users play the non-cooperative decentralized spectrum sharing game.

Consider a simple scenario seen in Fig. 1 where users 1, 2 and 3 are playing a graphical anti-coordination game as follows. Each user selects a color (spectrum band) white (W) or black (B) as their strategy based on the output of the evaluation function. Based on the color of their neighbors their utility is realized according to the payoff matrix shown in Fig.1. If two neighbors select the same color they incur a cost of -1 otherwise they get a reward 1. Moreover each user plays the game with each neighbor separately and its final decision is based on the realization of the composite game. For example, assume the users' strategy vector is $\mathbf{A} = (a_1 = W, a_2 = B, a_3 = W)$. Then user 3 obtains a cost -1 for choosing the same band as user 1. Also it obtains 1 from playing the game with user 2 since they have chosen different bands. Therefore user 3 obtains a total utility of $U^3 = 0$ from playing the composite game with its neighbors. We can also define a potential function for this game, an example is shown in Fig. 1. For example given $a_1 = B, a_2 = B$, user 3 can improve its utility from -2 to 2 by changing its strategy from B to W which corresponds to the same change in its potential function value.

We can generalize this example by defining the elements of a spectrum sharing graphical game \mathcal{G} :

1. Players are the secondary users $i \in V$.
2. Set of pure strategies for user (vertex) i is the set $a_i = \{1, \dots, B\}$. Then the joint action strategy space for the entire network is $\mathcal{A} = \{1, \dots, B\}^n$. Let us denote the joint-action by $\mathbf{A} \in \mathcal{A}$ and let $\mathbf{A}(i : a'_i) \equiv (a_1, \dots, a_{i-1}, a'_i, \dots, a_n)$ denote the vector resulting from setting the strategy of user i in vector \mathbf{A} to a'_i while keeping all other

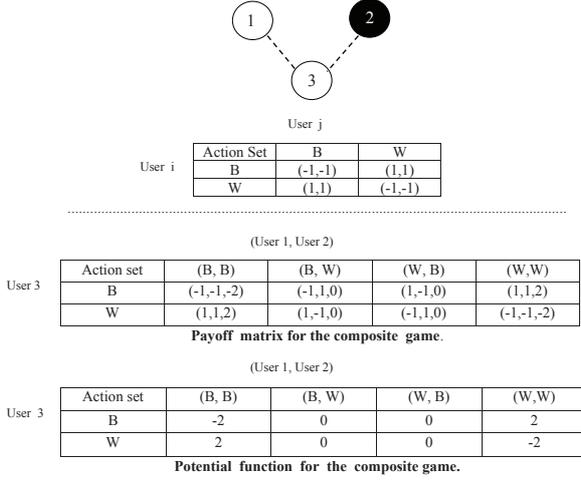


Figure 1. Simple example of a spectrum sharing graphical anti-coordination game.

strategies the same. The mixed strategy set for user i is the probability mass function $\{p(a_i) \equiv p_{ia_i}\}$, $\sum_{a_i=1}^B p_{ia_i} = 1$.

3. The utility user i receives, U^i , is given by the following linear function:

$$U^i(\mathbf{A}(i : a_i)) = \theta_{a_i} + \sum_{j \in \mathcal{N}_i} M_i(a_i, a_j) \quad (3)$$

where $j \in \mathcal{N}_i$ if there is an edge $e(i, j) \in E$ and $M_i(a_i, a_j)$ denotes the symmetric payoff user i obtains by playing the strategy a_i against the strategy a_j of user j . M_i is the following anti-coordination payoff matrix:

$$M_1 = M_2 = \dots = M = \begin{pmatrix} -1, -1 & 0, 0 & \dots & 0, 0 \\ 0, 0 & -1, -1 & \dots & 0, 0 \\ \dots & \dots & \dots & \dots \\ 0, 0 & 0, 0 & \dots & -1, -1 \end{pmatrix} \quad (4)$$

From M_i we see that if two users choose the same band then their respective payoff is -1 otherwise it is 0 for each. Also we assume the payoff Matrix to be symmetric i.e, $M_1 = M_2 = \dots = M$.

Note 1 The linear selection of users utility is for the ease of modelling of the distributed spectrum sharing games and does not reflect the actual performance of a practical communication network. However in the Simulation Results section it is shown that the result of the theoretical parts remains valid for the more generalized model where users utility follow a non-linear function of the interference.

Definition 2.1 A joint action strategy $\mathbf{A} \in \mathcal{A}$ is called a Nash equilibrium (NE) if no user $i, \forall i \in V$ has an incentive to deviate from the equilibrium strategy.

Definition 2.2 [4] A correlated equilibrium (CE) for game \mathcal{G} is a joint-probability distribution Q over the

joint action space \mathcal{A} such that for every user i , and every action pair $(j, a'_i) \in \mathbf{A}^2, j \neq a'_i$,

$$\sum_{\mathbf{A} \in \mathcal{A}} Q(\mathbf{A}) U^i(\mathbf{A}(i, a_i)) \geq \sum_{\mathbf{A} \in \mathcal{A}} Q(\mathbf{A}) U^i(\mathbf{A}(i : a'_i)) \quad (5)$$

A NE is a CE such that Q is a product distribution; that is $Q = \prod_{i=1}^n q_i$.

Definition 2.3 [20] Given a joint mixed strategy Q , let $H(Q) \equiv -Q(\mathbf{A}) \ln(Q(\mathbf{A}))$ denote the Shannon entropy. Then a maximum entropy correlated equilibrium (MECE) is the joint mixed strategy

$$Q^* = \operatorname{argmax}_{Q \in CE} H(Q) \quad (6)$$

Theorem 2.1. The spectrum sharing game \mathcal{G} has a potential function $\Phi : \mathcal{A} \rightarrow \mathbb{R}$ given by:

$$\Phi(\mathbf{A}) = \sum_i \theta_{a_i} + \sum_i \sum_{j \in \mathcal{N}_i} H(a_i, a_j) \quad (7)$$

where H is :

$$H = \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 \end{pmatrix} \quad (8)$$

Proof. We observe that if a matrix game M has a potential function H , then so does the associated graphical game with the following potential function Φ' :

$$\Phi'(\mathbf{A}) = \sum_i \sum_{j \in \mathcal{N}_i} H(a_i, a_j) \quad (9)$$

To see this suppose that user i deviates, say by choosing strategy a'_i . Then,

$$\begin{aligned} U^i(\mathbf{A}(i : a_i)) - U^i(\mathbf{A}(i : a'_i)) &= \\ \sum_{j \in \mathcal{N}_i} [M(a_i, a_j) - M(a'_i, a_j)] &= \\ \sum_{j \in \mathcal{N}_i} [H(a_i, a_j) - H(a'_i, a_j)] & \end{aligned} \quad (10)$$

From this it is now easy to see that matrix H characterizes a potential function:

$$H = \begin{pmatrix} -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 \end{pmatrix} \quad (11)$$

Therefore it follows that:

$$\Phi(\mathbf{A}) = \sum_i \theta_{a_i} + \sum_i \sum_{j \in \mathcal{N}_i} H(a_i, a_j) \quad (12)$$

□

The existence of the potential function then shows the existence of pure Nash strategies for \mathcal{G} [1]. Let $\mathcal{E}(\mathcal{G}) \subseteq \mathcal{A}$ denote the set of pure NE equilibria.

Definition 2.4 [13]. The price of anarchy $PoA(\mathcal{G})$ is:

$$PoA(\mathcal{G}) := \frac{\max_{\mathbf{A} \in \mathcal{A}} U(\mathbf{A})}{\min_{\mathbf{A} \in \mathcal{E}} U(\mathbf{A})} \geq 1 \quad (13)$$

$PoS(\mathcal{G})$ denotes the price of stability defined as:

$$PoS(\mathcal{G}) := \frac{\max_{\mathbf{A} \in \mathcal{A}} U(\mathbf{A})}{\max_{\mathbf{A} \in \mathcal{E}} U(\mathbf{A})} \geq 1 \quad (14)$$

Consider a special case of the spectrum sharing game \mathcal{G} when there are $B = 2$ available channels and $\Theta_{1 \times 2} = \mathbf{0}$ [14]. It can then be shown that $PoA(\mathcal{G})$ can be $\Omega(n^2)$ worse than $PoS(\mathcal{G})$ [15]. For example consider G to be the complete bipartite graph $G = K_{\frac{n}{2}, \frac{n}{2}}$. Since G is bipartite $PoS(\mathcal{G}) = 1$. To show that PoA can be $\Omega(n^2)$ worse it is enough to notice that one Nash equilibrium can be realized when half of the users on the left side of this bipartite graph occupy a same channel and the other half occupy the other channel. This implies that there are both good and bad Nash equilibria in spectrum sharing games. Let us call an equilibrium \mathbf{A} *good* if $PoA(\mathbf{A})$ is small and bad otherwise. In this situation a central authority can be employed to move the system behavior from a bad to a good equilibrium. For example in [15], a central authority advertises the optimal equilibria. It has been demonstrated that in a general graph G , if users employ the advertisement strategy in their best response learning mechanism, with probability more than half, the game converges to the optimal equilibrium in polynomial time. In this work we consider using a distributed learning method such as Log-linear mechanism [17-18] modified by a congestion advertisement for two reasons. First, because finding the optimal configuration (for a centralized approach) even for a simplified game model is a NP-hard problem [14]. Second, transient properties of the available spectrum opportunities in CR makes methods such as [15] not applicable for this problem. Transient properties could be of several types. Primary users evacuate and occupy their band continually. Autonomous secondary users join or leave the network. Moreover the network structure can also be unknown.

2.1. Logit-Response Dynamic

We proposed a synchronous logit-response in [3] for spectrum sharing games. In asynchronous logit-response [17], it is assumed that players are equipped with independent and identically distributed (i.i.d.) rate 1 "Poisson alarm clocks" and when their alarm goes off they revise their strategy according to a noisy

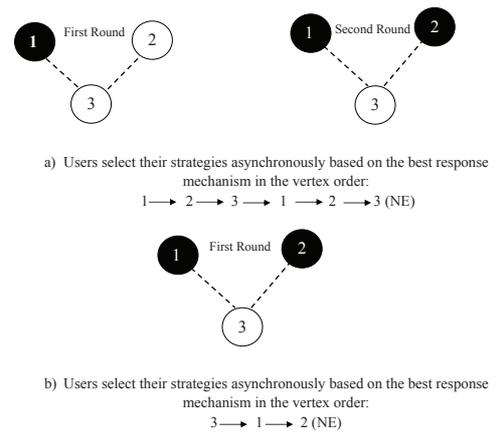


Figure 2. In graphical games how users learn is as much important as who learns first.

best response. Poisson distribution assumption implies that exactly one player at a time is allowed to update its strategy (asynchronous). Therefore the time between consecutive revision opportunities are independent and distributed with an exponential distribution of mean 1. When the user i alarm goes off it selects the strategy a_i with probability $p(a_i)$ according to a noisy best response mechanism given below::

$$p(a_i) = \frac{\exp(\beta U^i(\mathbf{A}(i : a_i)))}{\sum_{a'_i=1}^B \exp(\beta U^i(\mathbf{A}(i : a'_i)))} \quad (15)$$

where β represents the inverse temperature parameter. $\beta \rightarrow \infty$ is equivalent to the best response mechanism. For $\beta \rightarrow 0$ the dynamics are totally random.

Proposition 2.1 [17] If the game has the potential function $\Phi(\mathbf{A})$ the logit-response mechanism leads to a reversible and irreducible Markov process on the state space \mathcal{A} with the following stationary distribution:

$$\pi(\mathbf{A}) = \frac{\exp(\beta \Phi(\mathbf{A}))}{Z} \quad (16)$$

where $Z = \sum_{\mathbf{A} \in \mathcal{A}} \exp(\beta \Phi(\mathbf{A}))$ and as $\beta \rightarrow \infty$, $\pi(\mathbf{A})$ is concentrated on a Nash equilibrium. Moreover it turns out that the achieved equilibrium \mathbf{A}^* for $\beta \rightarrow \infty$ using the logit-response is a good equilibrium. That is, the price of anarchy is small for the achieved equilibrium \mathbf{A}^* .

However the main problem with this mechanism is its slow convergence rate. Therefore, in the next section we introduce the fastest logit-response mechanism.

3. MECE Logit-Response Mechanism

In the standard logit-response mechanism, users find the opportunity to update their strategies with a fixed

rate which is independent of their positions in the graph and dynamics of the system. However using a simple example we can show that the order in which users update their strategy affects the speed of convergence to a NE.

Consider Fig.2a where user 1 as the first player selects strategy B and then user 2 selects W. There is no payoff dominant strategy for user 3 at this stage. Suppose user 3 randomly selects the strategy W. The same process is then repeated for another round in order to reach the Nash equilibrium. However if user 3 selects its strategy before others as in Fig.2b the game ends up in a NE after the first round. This implies that the standard logit-response should be modified with respect to parameters such as the position of the users in the graph and the system dynamics.

Consider the maximum entropy correlated equilibrium MECE logit learning as is shown in Algorithm 1. In the MECE learning mechanism the clock alarms of the users go off according to a time varying probability distribution of Q^* as is described in (6). When a user gets the chance it updates according to (15).

As we showed in the previous example, the order of learning may cause negative effects in the convergence speed. This effect in fact can be explained via term Z in the stationary distribution $\pi(\mathbf{A})$ of (16) [10].¹

In the next lemma we show that MECE logit learning mechanism removes the term Z from the stationary distribution which can make the dynamics exponentially faster.

Lemma 3.1 The stationary distribution of \mathcal{G} , with the potential function Φ under the modified logit-response is $\pi(\mathbf{A}) \propto \exp(\beta\Phi(\mathbf{A}))$.

Proof. A correlated equilibrium can be explained conceptually by introducing a mediator who has access to a randomization device. The “alarm clocks” described in the standard logit-response mechanism is one such randomization device. The i.i.d assumption on the alarm distribution in the standard logit-response implements the NE with $Q = \prod_{i=1}^n q_i$ with $q_i = \frac{1}{n}$. The corollary 4.1 in [20] shows that there exists a joint probability distribution Q^* which removes the term Z from stationary distribution $\pi(\mathbf{A})$. \square

The modified learning approach can be thought of as a *Stackelberg* learning approach in which leaders and followers change roles along with the dynamics of the system.

The main importance of the modified mechanism is that it maps the dynamic of the game \mathcal{G} to that of

Ising models (as described in the next theorem). This demonstrates how hard it is to achieve a good equilibrium for spectrum sharing game \mathcal{G} in polynomial time.

For ease of analysis let's consider $B = 2$ and $\Theta = \mathbf{0}$. Assume if user i is transmitting in channel 1, $a_i = -1$ and if it is transmitting in channel 2, $a_i = 1$.

Theorem 3.1 The modified learning mechanism of the game coincides with the *Glauber dynamics* for the *anti-ferromagnetic Ising* models.

Proof. Consider the strategy set $a = \{-1, 1\}$ to be the set of spins. Glauber dynamics for Ising models are defined as algorithms that sample random assignments of spins to vertices V , according to a target distribution $\pi(\mathbf{A})$ using the following procedure: starting from any initial condition, repeatedly choose a site $i \in V$ uniformly at random, replace the spin of the site i with one sampled from $\pi(\mathbf{A})$ conditioned on the spins of \mathcal{N}_i . The Ising model is called an anti-ferromagnetic for $\pi(\mathbf{A}) \propto - \sum_i \sum_{j \in \mathcal{N}_i} a_i a_j, \forall i \in V$. Ising and ferromagnetic

for $\pi(\mathbf{A}) \propto \sum_i \sum_{j \in \mathcal{N}_i} a_i a_j, \forall i \in V$ [21].

Using the simplified assumptions and Theorem 2.1 we can rewrite $\Phi(\mathbf{A}) = - \sum_i \sum_{j \in \mathcal{N}_i} a_i a_j, \forall i \in V$. Then the

proof is complete by using Lemma 3.1. \square

Theorem 3.1 establishes the connection between anti-ferromagnetic Ising models and dynamics of spectrum sharing games. Propositions 4.1 and 4.2. in the next section are the direct results of this theorem.

4. Convergence Rate for Specific Graphs

It is known that the lower bound for *mixing time* (defined in appendix) of the Glauber dynamics for the Ising models is $O(n \log n)$ if the $\beta < \beta_T$ where β_T is dependent on the graph G and the model of interaction (in our case an Anti-ferromagnetic Ising model) [21]. For $\beta > \beta_T$ the mixing time is exponential.

Proposition 4.1 For spectrum sharing game \mathcal{G} there is a graph G where it is impossible to achieve a good equilibrium in polynomial steps.

Proof. As explained in section II-A in order to achieve a good equilibrium under logit-learning $\beta \rightarrow \infty$ is required. However Theorem 3.1 shows the game borrows the dynamic characteristic of the Ising model and for $\beta > \beta_T$ exponential steps in n is needed for convergence. \square

The previous theorem raises the question of what kind of graphs show better behavior in terms of the convergence speed. This can be answered using the next Proposition.

¹ We have avoided the formal discussion on the effect of term Z in the convergence rate of the logit response to keep the context as consistent as possible. In order to understand the relation between the term Z and learning dynamics please refer to example 2 in [10].

Initialization: $t = 0$.

1. A central authority solves the alarm clock probability distribution $Q^* = \underset{Q \in CE}{\operatorname{argmax}} H(Q)$.
2. When the alarm goes off for a player i , $t = t + 1$
Update strategy a_i with probability $p(a_i) = \frac{\exp(\beta U^i(\mathbf{A}(i:a_i)))}{\sum_{a'_i=1}^B \exp(\beta U^i(\mathbf{A}(i:a'_i)))}$.
3. If $\forall i, |U_{t+1}^i - U_t^i| \leq \epsilon_{\text{stop}}$, Stop .
Otherwise go back to step 1.

Algorithm 1: MECE Logit Response Learning

Consider all the possible subdivisions of the graph in two disjointed subsets of vertices: S and its corresponding complement $V \setminus S$. The Isoperimetric function of graph $\mathcal{C}(G)$ is defined as the minimum value over all possible partitions of the number of edges connecting S with $V \setminus S$ divided by the number of sites in the smallest of the two subsets. That is, $\mathcal{C}(G) = \min_{S \subset V: |S| \leq \frac{n}{2}} \frac{\text{cut}(S, V \setminus S)}{|S|}$.

Proposition 4.2 The smaller the value of Isoperimetric function of a graph \mathcal{C}_G the faster anti-ferromagnetic Ising model dynamics converges.

Proof. Let \mathbf{W} represent the adjacency matrix of G defined as the $n \times n$ matrix $\mathbf{W} = (W_{ij})$ in which $W_{ij} = \begin{cases} 1 & j \in \mathcal{N}_i, \\ 0 & \text{Otherwise.} \end{cases}$. Moreover assume \mathbf{I} to be the adjacency matrix of complete graph. Then let write $\Phi(\mathbf{A})$ as:

$$\Phi(\mathbf{A}) = - \sum_i \sum_j W_{ij} a_i a_j, \forall i \in V \quad (17)$$

Let G^c be the complementary graph of G with adjacency matrix \mathbf{W}^c , i.e $\forall i \neq j, W_{ij} + W_{ij}^c = 1$ then we can rewrite $\Phi(\mathbf{A})$ as:

$$\Phi(\mathbf{A}) = \sum_i \sum_j W_{ij}^c a_i a_j - \sum_i \sum_j I_{ij} a_i a_j, \forall i \in V \quad (18)$$

The first and second term of the right hand side of (18) can be recognized respectively as the (ferromagnetic) Ising model over the complementary graph G and anti-ferromagnetic Ising over the complete graph. The second term is independent of the phase transition analysis. This is because we can look at it as a symmetric congestion game which shows no phase transition behavior [10]. Therefore the phase transition behavior of the anti-ferromagnetic Ising can be shown by using Ising model on complementary graph G^c . The phase transition behavior of first term has been studied via Isoperimetric function of graphs in [11-27]. Using the

result of [11-27] over complementary graph G the proof is complete. \square

The previous proposition states that the more connected the graph of interaction G is, the faster the spectrum sharing game reaches an equilibrium.

4.1. Normalized Social welfare Optimization

In the previous discussions we addressed under what kinds of graphs, games converge faster to good equilibrium. This should not be confused with the problem of finding graph G whose purpose is to achieve the maximal social welfare $\sum_i U^i$.

Consider the optimization problem of (2). Our objective here is to find the graphs G which achieve high social welfare $U(\mathbf{A}) = \sum_{i \in V} U^i(\mathbf{A}(i, a_i))$. Let's modify the optimization problem with a new notion of normalized social welfare defined as:

$$\max_{\mathbf{A}: \mathbf{A} \in \{-1, 1\}^n} \frac{U(\mathbf{A})}{2|E|} \quad (19)$$

This is because we are looking for the graphs that have a high capacity for achieving the optimal social welfare. Therefore it should be normalized with respect to the number of edges $|E|$.

Instead of solving the linear optimization problem of (19) let's consider solving the quadratic problem by rewriting $U(\mathbf{A}) = \sum_i U^i$ as $U(\mathbf{A}) = - \sum_i \sum_{j \in \mathcal{N}_i} (a_i - a_j)^2$.

We can rewrite in the graphical format

$$U(\mathbf{A}) = - \sum_i \sum_{j \in \mathcal{N}_i} (a_i - a_j)^2 = -\mathbf{A}^T \mathcal{L}_G \mathbf{A} \quad (20)$$

where \mathcal{L}_G is the Laplacian graph of G (refer to Appendix for definition). Moreover notice the term $2|E| = \mathbf{A}^T \mathbf{A}$. Then the normalized social welfare maximization problem can be written as minimization problem of

$$\min_{\mathbf{A}: \mathbf{A} \in \{-1, 1\}^n} \frac{\mathbf{A}^T \mathcal{L}_G \mathbf{A}}{\mathbf{A}^T \mathbf{A}} \quad (21)$$

We assume G to be any arbitrary connected graph.

Lemma 4.1 Let λ be the smallest nonzero eigenvalue of \mathcal{L}_G , then

$$\lambda = \min_{\mathbf{A} \in \mathbb{R}^n} \frac{\mathbf{A}^T \mathcal{L}_G \mathbf{A}}{\mathbf{A}^T \mathbf{A}} \leq \min_{\mathbf{A}: \mathbf{A} \in \{-1,1\}^n} \frac{\mathbf{A}^T \mathcal{L}_G \mathbf{A}}{\mathbf{A}^T \mathbf{A}} \quad (22)$$

That is the optimal normalized social welfare is bounded below by λ .

Then to show the theorem now concentrate on the family of graphs that have high values of $\mathcal{C}(G)$ and therefore fast convergence properties. A graph G is called ζ -expander for every subset S with $|S| \leq \frac{n}{2}$, $\mathcal{C}(G) \geq \zeta$.

Lemma 4.2 (Cheeger inequality [29]) Let λ be the smallest nonzero Laplacian eigenvalue of graph ζ -expander graph G then

$$\zeta^2 \leq \lambda \leq 2\zeta \quad (23)$$

That is $\mathcal{C}(G) = \zeta$ is upper bounded with λ .

(22) states that the graphs with lower value of λ provide larger normalized social welfare. However (23) shows larger λ are needed for faster convergence. *This shows a trade off between the normalized social welfare and the convergence speed.*

5. Spectrum Sharing Via Congestion Advertisement

We saw in the previous section that it is impossible to achieve a good equilibrium with polynomial mixing time on particular graphs. Therefore we are investigating a method to reach a good equilibrium in any graph.

Theorem 5.1 Assume that each user i evaluates its utility as:

$$U_h^i = - \sum_{j \in \mathcal{N}_i} a_i a_j + a_i \epsilon \quad (24)$$

The game \mathcal{G} under the MECE-logit response converges within polynomial steps to the good equilibrium if at each stage of the game $\epsilon = \text{sign}(\mathcal{E}_2 - \mathcal{E}_1)h$ where $h > 0$ is a small value, $\mathcal{E}_1 = \sum_i \sum_{j \in \mathcal{N}_i} \delta_{a_i, -1} \delta_{a_j, -1}$, $\mathcal{E}_2 =$

$$\sum_i \sum_{j \in \mathcal{N}_i} \delta_{a_i, 1} \delta_{a_j, 1} \text{ and } \delta \text{ is the Kronecker delta .}$$

Proof. With the utility function of (24) the best response strategy of user i can be written as $\text{sign}(- \sum_{j \in \mathcal{N}_i} a_j + \epsilon)$.

By doing this the term ϵ makes one specific strategy for user i "risk dominant". The risk dominant strategy for user i is the one yielding the highest payoff and that is when user i have no information about its neighbors \mathcal{N}_i . If half of the \mathcal{N}_i are active in a channel

strategy and the other half are active in the other, user i will select the risk dominant one. Theorem 3.1 shows that dynamic of \mathcal{G} can be analysed by studying the dynamics of Glauber algorithms for anti ferromagnetic Ising model. [11] implies that polynomial mixing time for ferromagnetic Ising model over any graph G can be achieved by introducing a risk dominant term in the favor of one of the strategies. Let's show this risk dominant strategy by $a_d \in \{-1, 1\}$. Then in the stationary state with probability converging to 1 every user select strategy a_d . This is the key to polynomial mixing time.

However having a fixed risk dominant in the spectrum sharing games is disastrous as it provides wrong incentive for the users to occupy the same channel and therefore cause a large interference. Let's write the social welfare as $U^i(\mathbf{A}) = -(n_{11} + n_{22})$ where n_{11} and n_{22} represent the number of edges $e(i, j)$ in which $a_i = a_j = 1$ and $a_i = a_j = 2$ respectively. This format is clear since the other edges that don not experience interference add zero value to the social welfare. The logit-response leads to good equilibrium for $\beta \rightarrow \infty$ when the term $n_{11} + n_{22}$ will be minimized. The risk dominant strategy should be selected in favor of the appropriate strategy to minimize this term. In order to select the appropriate term, lets define the energy of a strategy in the system, by the value of concentration of users on that specific strategy, $\mathcal{E}_1 = \sum_i \sum_{j \in \mathcal{N}_i} \delta_{a_i, -1} \delta_{a_j, -1}$, $\mathcal{E}_2 = \sum_i \sum_{j \in \mathcal{N}_i} \delta_{a_i, 1} \delta_{a_j, 1}$.

For example in a complete graph due to the symmetry of strategy configuration, the channel with more active users has higher energy. Energy also depends on the graphical characteristics of the users that have occupied a strategy. Then it would be enough to flip the sign of ϵ in favour of the strategy with less energy. Assume that the energy of strategy -1 (channel 2) is more than +1 (channel 1). Then by making the $\epsilon > 0$ we actually balance the energy by making the strategy +1 risk dominant (channel 1). This reduces the term $(n_{11} + n_{22})$ as it prevents it from any existing permanent risk dominant strategy in the network. Notice that if one strategy becomes risk dominant permanently without considering the energy difference, every user with probability converging to 1 chooses the same strategy and therefore increases the existing interfering links. \square

The previous theorem states that to have an exponentially faster spectrum sharing, users need to introduce an *arbitrary small* risk dominant term ϵ in their utility. Producing these suitable risk-dominant terms can be based upon an advertisement entity such as a base station. Finding the exact risk dominant terms is a difficult problem but it can be approximated by a congestion announcement. A simple scenario of

spectrum sharing via congestion advertisement can be described as follows:

Base stations announce the number of active users in different spectrum bands during each time slot. When users experience the same signal interference for both available channels, they select the one with less congestion. Let's $\rho_{a_i} = \frac{1}{|V|} \sum_{j \in V} \delta_{a_i, a_j}$ be associated with action profile \mathbf{A} and advertisement control parameter $h > 0$. Then user's i utility can be displayed in the following form:

$$U_h^i(\mathbf{A}(i, a_i)) = U^i(\mathbf{A}(i, a_i)) - h\rho_{a_i} \quad (25)$$

This generalizes (24) by presenting the energy difference as the congestion term ρ . The congestion advertisement method has been applied to the spectrum sharing game with utility format of (25). The results are explained in the next section.

6. Simulation Results

Our simulations have been conducted for a more generalized version than the theoretical part. It shows that many of our theoretical results maintain their validity even with the change of some assumptions. These generalized assumptions are:

- Learning method is a Responsive Learning Automata (RLA) [24] with the learning parameter of $\alpha \in (0, 1)$. The description on this learning algorithm is given in Appendix.
- Users update their strategies simultaneously.
- In the simulations there are four channel strategies $B = 4$ and arbitrary Θ .
- Utility format is of the form $U^i(\mathbf{A}(i, a_i)) = \frac{\theta_{a_i}}{\rho_{a_i}(1+|\mathcal{I}_i|)}$.
- G is a geometric graph.

6.1. Price of Anarchy

In a distributed cognitive system when there are multiple Nash equilibria it is important to understand the gap between worst and best possible equilibria of the network. As it is described in Section 2 a good way for understanding this is the comparison between the price of anarchy and stability. These metrics then imply how bad or good the network output can be from its best achievable one. When the price of anarchy is much worse than the price of stability it becomes crucial for the network designer to come up with mechanisms that could move the system's behaviour from a less efficient equilibrium to a more efficient one by promoting better user behavior. This paper

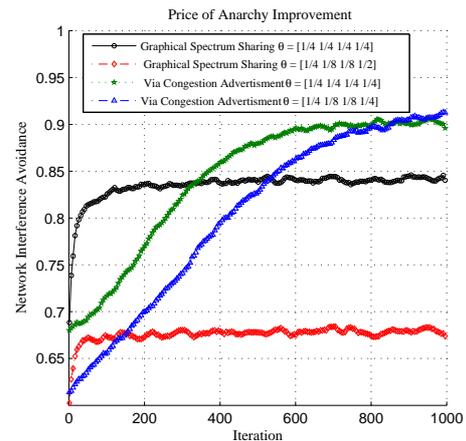


Figure 3. Price of Anarchy Improvement, $B = 4$ and the simulation has averaged over 100 realizations, best α for RLA learning for each case has selected to make the comparison independent of learning process, the network interference avoidance is: $\frac{1}{|V|} \sum_{i \in V} \frac{1}{1+|\mathcal{I}_i|}$.

introduced congestion advertisement for the spectrum sharing games as one of these mechanisms.

Fig. 3 shows the improvement of spectrum sharing with congestion advertisement mechanism in terms of network interference avoidance. As well as demonstrates that by injecting congestion incentive into users utility, there is less probability that users will herd to a spectrum with higher quality. This in turn reduces the interference.

6.2. Convergence

Network volatility has been plotted in order to show the convergence rate has improved. Volatility is defined as the variance of alternating between different strategies. Fig. 4 shows that with the increase in the communication range, the convergence rate improves since it increases the graph connectivity. This validates the result of Proposition 4.2. Fig. 5 also shows congestion advertisement method enhances the dynamics of \mathcal{G} .

6.3. Phase transition

We have run several simulations for different values of learning parameter α . We have also selected the best α which will bring the highest social welfare of $\frac{1}{1+|\mathcal{I}_i|}$ for different values of h . These results indicate a transition point in Fig 6. You can see at the beginning, with the increase in h , the exploration rate α required to find the good equilibrium reduces. This improves the convergence rate. However by continually increasing h , the system arrives at a transition point h_T . This is where users herd on the channel with less congestion which

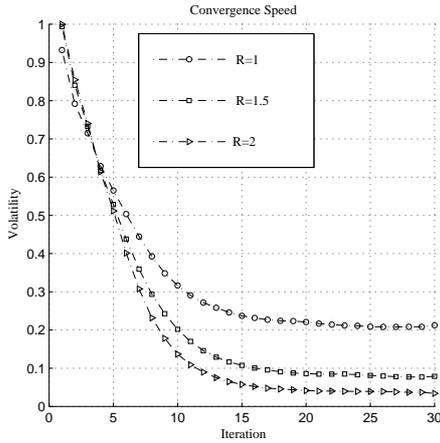


Figure 4. The simulation has run for the geometric graph. We assume a square area 100 units uniformly distributed random configuration of $n = 100$ secondary users. We consider different communication ranges R . For example $R = 1$ means that for a secondary node i all other nodes within an Euclidean distance of $R = 1$ are considered to be the neighbors \mathcal{N}_i . It shows well connected graphs have higher convergence speed.

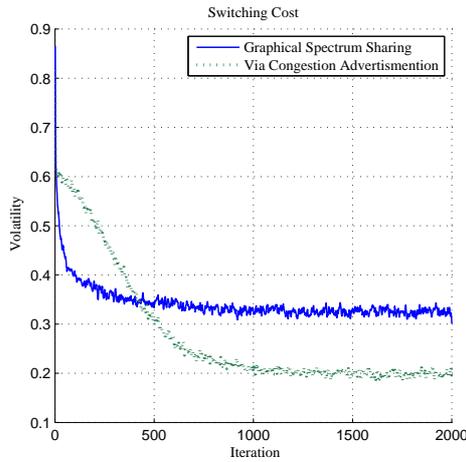


Figure 5. Network volatility improvement (as a convergence criteria) comparing with graphical spectrum sharing, $B = 4$, homogeneous spectrum quality $\theta = [1/4 \ 1/4 \ 1/4 \ 1/4]$, simulation has averaged over 15 realization, α in RLA learning for each case has selected so that it shows the best performance possible in the shown region, utility format is $U^i(\mathbf{A}(i, a_i)) = \frac{\theta_{a_i}}{\rho_{a_i}(1+|Z_i|)}$.

starts to increase the signal interference. Therefore in order to reduce the interference, a higher level of irrationality becomes necessary. This is similar to the behavior of the Glauber dynamics for $\beta > \beta_T$

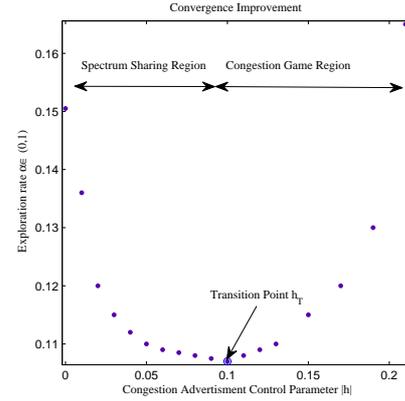


Figure 6. We have used 20 realizations for a observation window size of 500 for a network size of $n = 100$ and $B = 4$ numbers of channel. Also the simulation has run for a constant range of interaction on a random geometric graph to make the analysis independent of range of interaction.

7. Conclusion

We addressed the spectrum sharing games using graphical anti-coordination games. We showed how a modified logit learning mechanism establishes the connection between the simplified spectrum sharing games and anti-ferromagnetic Ising models. We studied the convergence rate of these spectrum sharing games and discovered the trade offs between achieving a good social welfare and convergence rate. This demonstrated spectrum sharing games under graphs which have lower isoperimetric values, tend to converge faster to equilibrium. We also showed how introducing an arbitrary small advertisement parameter into equations can enhance the convergence significantly.

Appendix A. Laplacian Graph

Definition A.1 The Laplacian of the graph G is defined as the $n \times n$ matrix $\mathcal{L}_G = (L_{ij})$ in which $L_{ij} =$

$$\begin{cases} d_{ij} & i = j, \\ -w_{ij} & i \neq j \end{cases}$$

Appendix B. Markov

Definition A.2 The total variation distance between two probability distributions μ and ν on \mathcal{A} is defined by

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{i \in \mathcal{A}} |\mu_i - \nu_i| \quad (\text{B.1})$$

Definition A.3 The time it takes for a process to reach to its stationary distribution ν is known as mixing times $\tau(\epsilon)$:

$$\tau(\epsilon) = \min_t \{\|\mu(t) - \nu\|_{TV} \leq \epsilon\} \quad (\text{B.2})$$

Appendix C. Responsive Learning Automata

Let r_i^t represents the payoff obtained at time t by playing strategy i . The update rules for responsive learning automata are:

$$p_i^{t+1} = p_i^t + \alpha r_i^t \sum_{j \neq i} s_j^t p_j^t$$

$$\forall j \neq i, p_i^{t+1} = p_i^t - \alpha r_i^t s_j^t p_j^t$$

$$s_j^t = \min\left[1, \frac{p_j^t - \alpha/2}{\alpha p_j^t r_i^t}\right]$$

where α is the learning parameter and p_i^t is the probability of playing the strategy i at time t .

References

- [1] D. Monderer and L. Shapley, "Potential Games", *Games and Economic Behavior* 14, pp.124-143, 1996.
- [2] S. Haykin, "Cognitive radio: brain-empowered wireless communications", *IEEE J. Select. Areas Commun.*, vol. 3, no. 2, pp. 201-220, 2005.
- [3] M. Azarafrooz, R. Chandramouli, "Distributed Learning in Secondary Spectrum Sharing Graphical Game". *IEEE GLOBECOM*, 2011.
- [4] R.J. Aumann, "Subjectivity and correlation in randomized strategies", *Journal of Mathematical Economics*, 1974.
- [5] M. Liu and Y. Wu, "Spectrum sharing as congestion games", *Annual Allerton Conf. Commun., Control, and Compu.*, pp. 1146 - 1153, 2008.
- [6] I. Malanchini, M. Cesana and N. Gatti, "On Spectrum Selection Games in Cognitive Radio Networks", *IEEE GLOBECOM*, pp. 1 - 7, 2009.
- [7] P. Mertikopoulos and A. Moustakas, "Correlated Anarchy in Overlapping Wireless Networks", *IEEE J. Select. Areas Commun.*, vol. 26, no. 7, pp. 1160-1169, 2008.
- [8] S. Sengupta, R. Chandramouli, S. Brahma and M. Chatterjee, "A game theoretic framework for distributed self-coexistence among IEEE 802.22 networks", *IEEE GLOBECOM*, pp. 1-6, 2008.
- [9] H. Li and Z. Han, "Competitive Spectrum Access in Cognitive Radio Networks: Graphical Game and Learning". *Wireless Communications and Networking Conference (WCNC)*, 2010.
- [10] D. Shah and J. Shin, "Dynamics in congestion games", *SIGMETRICS*, pp-107-118, 2010.
- [11] A. Montanari and A. Saberi, "The Spread of Innovations in Social Networks", *Proc. Natl. Acad. Sci.*, 2010.
- [12] A. Ramezanzpour, J. Realpe-Gomez and R. Zecchina "Statistical physics approach to graphical games: local and global interactions". *The European Physical Journal on Condensed Matter and Complex Systems*, 2011.
- [13] C. Papadimitriou, *Algorithms, games, and the internet*, In *STOC*, 2001.
- [14] G. Christodoulou, V. S. Mirrokni, A. Sidiropoulos, "Convergence and Approximation in Potential Games". In: *Durand, B., Thomas, W. (eds.) STACS 2006. LNCS, vol. 3884, pp. 3490. Springer, Heidelberg* 2006.
- [15] M. Balcan, A. Blum and Y. Mansour, "Improved equilibria via public service advertising", *SODA Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2009.
- [16] S. Kirkpatrick, Jr. Gelatt and M. Vecchi, "Optimization by Simulated Annealing" *Science*, 220, 671-680. 1983.
- [17] L. Blume, "The statistical mechanics of strategic interaction.", *Games and economic behavior*, Vol. 5, No. 3., pp. 387-424, 1993.
- [18] C. A. Ferrer and N. Netzer, "The logit-response dynamics", *TWI Research Paper Series 28, Thurgauer Wirtschafts institut, Universitat Konstanz*, 2008.
- [19] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: asynchrony, completeness, and payoff based implementation", *Games and Economic Behavior*, 2011.
- [20] L. E. Ortiz, R. E. Schapire and S. M. Kakade, "Maximum Entropy Correlated Equilibria", In *Eleventh International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2007.
- [21] T. P. Hayes and A. Sinclair, "A general lower bound for mixing of single site dynamics on graphs", In *FOCS: Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 2005.
- [22] A. Montanari and A. Saberi, "Supplementary Information for: On the Spread of Innovations in Social Networks", 2010.
- [23] K. Satyen and C. Seshadhri, "Combinatorial Approximation Algorithms for MaxCut using Random Walks", In *2nd Symposium on Innovations in Computer Science (ICS)*, 2011.
- [24] E. J. Friedman and S. Shankar "Decentralized learning and the design of Internet". *Mimeo*, 1996.
- [25] H. Young "Evolution of Conventions" *Econometrica*, Vol. 61, No. 1., pp. 57-84, 1993.
- [26] A. Galeotti. "Network Games." *The Review of Economic Studies*. 77(1), 218-244. 2010.
- [27] N. Berger, C. Kenyon, E. Mossel and Y. Peres, "Glauber dynamics on trees and hyperbolic graphs", *Probab. Theory Rel.* 131:311-340.
- [28] S. Kirkpatrick, Jr. Gelatt and M. Vecchi, "Optimization by Simulated Annealing", *Science*, 220, 671-680, 1983.
- [29] N. Alon, "Eigenvalues and expanders", *Combinatorica* 6, no. 2, 83-96, 1986.