Removing Phase Transition in Distributed Spectrum Sharing Games Via Congestion Advertisement

Mahdi Azarafrooz and R. Chandramouli
Department of Electrical and Computer Engineering
Stevens Institute of Technology

Abstract—Distributed spectrum sharing has been studied via game theoretical tools. We seek answer to the question that how long does it take to reach good equilibrium in a spectrum sharing game $G$. Mainly we are motivated to answer the question that given that $n$ secondary users are located on the vertex of a general graph $G$, whether there exists any scheme to reach the good equilibrium in polynomial (in $n$) time. Answer to the problem has been provided by establishing the connections with the concepts of maximum entropy correlated equilibria (MECE) and phase transition.

A MECE logit-response learning mechanism has been introduced that maps the dynamic of the game to the glauber dynamic of anti-ferromagnetic Ising models.

Using this mapping it has been shown that there exists graphs that it is impossible to achieve a good equilibrium within polynomial steps for the game $G$. It also has been proved that there is a trade off between the capacity of graph $G$ in providing good social welfare and fast convergence properties.

It has been shown how base stations can improve the convergence rate by modifying the dynamics of the system through a spectrum congestion advertisement so that independent of the graph $G$ it is possible to obtain the good equilibrium solution in polynomial time.

The effect of congestion advertisement on dynamics of network has been verified via simulations. Appropriate advertisement leads to improved price of anarchy and faster convergence.

Index Terms—Spectrum Sharing, Game Theory, Graph Theory, Spin Systems, Logit-Response Mechanism, Phase Transition.

I. INTRODUCTION

Cognitive Radios (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 appropriately [1-2], etc. Therefore, every action of a specific 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 the spatial re-use. This was modeled 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 mechanism design and an iterative technique to compute an efficient Nash equilibrium solution with fast convergence properties. 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’ payoff 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 by defining a secondary user specific utility as a function of 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 modeled as a minority game where the networks try to minimize their cost in finding a clear band. A graphical game model for competitive spectrum access is discussed in [9].

This paper is directly related to [10-12]. In [10] the convergence rate of congestion games for reaching a good equilibrium is studied. Convergence of coordination games in a social networking context is presented in [11]. In [12] a statistical physic framework for graphical games with global interactions is considered.
II. SPECTRUM SHARING AS A GRAPHICAL ANTI-COORDINATION GAME

Consider a CR network scenario where $n$ secondary users are placed on an undirected graph $G = (V, E)$, where $|V| = n$ and $E$ is the set of edges. Let $N_i$ 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 $A_{1 \times n}$ be the action vector where the $i$th element $a_i \in \{1, ..., B\}$ 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., depending 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., $\theta_l$, $l \in \{1, 2, ..., B\}$ denotes the quality of the $l$th spectral band. For example, this could be a function of primary user activity, required data rate, etc. The higher the value of $\theta_l$ the more desirable that band is. let $I_{a_i}$ 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 $U^i(I_{a_i}, \theta_{a_i})$. That is, the utility function depends on the interference level as well as the quality of the operating band. Then the social welfare of the network is:

$$U(A, \Theta) = \sum_i U^i(I_{a_i}, \theta_{a_i})$$

(1)

The optimal solution $A^*$ to the spectrum sharing problem is then given by:

$$A^* = \arg\max_A U(A, \Theta)$$

(2)

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

Consider a simple scenario as seen in Fig. 1 which 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 have a cost of -1 otherwise they obtain 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 strategy vector of $A = (a_1 = W, a_2 = B, a_3 = W)$ has been selected by the users. Then user 3 has a cost -1 from being in the same color of user 1. Also it obtains 1 from playing the game with user 2 since they have chosen different colors. Therefore user 3 obtains utility $U^3 = 0$ from playing the composite game with it’s neighbors. A potential function for user 3 can be provided so that it could be possible to correspond the improvement of it’s utility from one strategy configuration to another. For example given $a_1 = B, a_2 = B user 3 can improve it utility from -2 to 2 by changing it’s strategy from B to W which corresponds to the the same change of it’s potential function value.

Therefore, to generalize this example, the elements of a spectrum sharing graphical game $G$ are:

1. Players are the secondary users $i \in V$.
2. Set of pure strategies for user (vertex) $i$ is the set $a_i = \{1, ..., B\}$. Then the joint action strategy space for the entire network is $A = \{1, ..., B\}^n$. Let us denote the joint-action by $A \in A$ and let $A(i : a_i') \equiv (a_1, ..., a_{i-1}, a_i', ..., a_n)$. The mixed strategy set for user $i$ is then the probability mass function $p(a_i) \equiv p(ia_i), \sum_{a_i=1} p(ia_i) = 1$.

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

$$U^i(A(i : a_i)) = \theta_{a_i} + \sum_{j \in N_i} M(a_i, a_j)$$

(3)

where $j \in N_i$ if there is an edge $e(i, j) \in E$ and $M$ is the following anti-coordination payoff matrix:

$$M = \begin{pmatrix} -1, -1 & 0, 0 & ... & 0, 0 \\ 0, 0 & -1, -1 & ... & 0, 0 \\ ... & ... & ... & ... \\ 0, 0 & 0, 0 & ... & -1, -1 \end{pmatrix}$$

(4)

From $M$ we see that if two users choose the same band then their respective payoff is -1 otherwise it is 0 for each.

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

In this paper we need to more equilibrium definitions.

**Definition 2.2** [4] An correlated equilibrium (CE) for game $G$ is a joint-probability distribution $Q$ over the joint action space $A$ such that for every user $i,$ and every action pair
\((j,k) \in A^2, j \neq k,\)
\[
\sum_{A \in A_1} Q(A)U^i(A) \geq \sum_{A \in A_1} Q(A)U^i(A(i:j)) \tag{5}
\]

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 \(H(Q) \equiv -Q(A) \ln(Q(A))\) denote the Shannon entropy, then a maximum entropy correlated equilibrium (MECE) is the joint mixed strategy
\[
Q^* = \arg\max_{Q \in CE} H(Q) \tag{6}
\]

**Theorem 2.1.** The spectrum sharing game \(G\) has a potential function \(\Phi : A \rightarrow \mathbb{R}\) given by:
\[
\Phi(A) = \sum_i^a \theta_{a_i} + \sum_{i,j \in N_i} H(a_i, a_j) \tag{7}
\]

where \(H\) is:
\[
H = \begin{pmatrix}
-1 & 0 & \ldots & 0 \\
0 & -1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & -1
\end{pmatrix} \tag{8}
\]

**Proof:** If a matrix game \(M\) has potential function \(H\), then so does the associated graphical game with the following potential function \(\Phi'\):
\[
\Phi'(A) = \sum_i \sum_{j \in N_i} H(a_i, a_j) \tag{9}
\]

To see this suppose that user \(i\) deviates, say by choosing strategy \(a_i'\). Then,
\[
U^i(A(i:a_i)) - U^i(A(i:a_i')) = 
\sum_{j \in N_i} [M(a_i,a_j) - M(a_i',a_j)] = 
\sum_{j \in N_i} [H(a_i, a_j) - H(a_i', a_j)] \tag{10}
\]

From this it is now easy to see that the potential function \(H\) is
\[
H = \begin{pmatrix}
-1 & 0 & \ldots & 0 \\
0 & -1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & -1
\end{pmatrix} \tag{11}
\]

Therefore it follows that:
\[
\Phi(A) = \sum_i^a \theta_{a_i} + \sum_{i,j \in N_i} H(a_i, a_j) \tag{12}
\]

Theorem 2.1 proves the existing of pure Nash strategies for \(G\). Let \(E(G) \subseteq A\) denote the set of pure NE equilibria.

**Definition 2.4** [13]. The price of anarchy \(PoA(G)\) is:
\[
PoA(G) := \max_{A \in E} \frac{U(A)}{\min_{A \in E} U(A)} \geq 1 \tag{13}
\]

Assume a special case of the spectrum sharing game \(G\) when there are \(B = 2\) channels available and \(\Theta_{1 \times 2} = 0\) [14]. It can be shown that \(PoA(G)\) can be \(\Omega(n^2)\) worse than \(PoS(G)\) [15]. For example, consider \(G\) to be the complete bipartite graph \(G = K_{\frac{n}{2}, \frac{n}{2}}\). Since \(G\) is bipartite graph \(PoS(G) = 1\). To show that \(PoA(G)\) can be \(\Omega(n^2)\) worse it is enough to notice that one Nash equilibrium can be realized when half of the users on left occupy a same channel and the other half occupy the other channel. This implies there are both high and low efficient Nash equilibria in spectrum sharing games. Let us call an equilibrium \(A\) a good if \(PoA(G)\) is small and bad otherwise. In this situation a central authority is needed to move the system behavior from a bad to a good equilibrium. In [15], a central authority advertises the approximate optimal equilibrium. 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 suggest using a distributed learning method such as Log-linear mechanism [17-18] modified by a congestion advertisement for spectrum sharing. 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 that of [15] implausible. Primary users evacuate and occupy their band continually. Autonomous secondary users join or leave the network too. Moreover the network structure can also be unknown. Therefore in the next section we first describe the logit-response learning.

A. Logit-Response Dynamic

We proposed a synchronous logit-response in [3] for spectrum sharing games. In this work an asynchronous logit-
response has been considered to reflect reality more closely. In asynchronous logit-response [17], it is assumed that players are equipped with i.i.d rate 1 “Poisson alarm clocks” and when their alarm goes off they revise their strategy according to a noisy best response. Poisson distribution assumption implies that exactly one player at a time is allowed to update its strategy (asynchronous). In another way the times between consecutive revision opportunities are independent and distributed with the exponential distribution of mean 1. When the user i alarm goes off he selects the strategy j with probability \( p_{ij} \) according to a noisy best response mechanism of the following:

\[
p_{ij} = \frac{\exp(\beta U^j(A(i:j)))}{\sum_{k=1}^{B} \exp(\beta U^k(A(i:k)))}
\]

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

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

\[
\pi(A) = \frac{\exp(\beta \Phi(A))}{Z}
\]

where \( Z = \sum_{A \in A} \exp(\beta \Phi(A)) \) and as \( \beta \to \infty \), \( \pi(A) \) is concentrated on a Nash equilibrium.

Moreover it turns out that the achieved equilibrium \( A^* \) for \( \beta \to \infty \) using the logit-response is a good equilibrium. That is, the price of anarchy is small for the achieved equilibrium \( A^* \). Recall the high price of anarchy of \( \Omega(n^2) \) in bipartite graph \( G = K_{n,n} \). It has been observed in simulations that the worst case equilibrium realizes when \( \beta \) is not large enough. This is due to the fact that the logit-learnings employ the so called “simulated annealing” technique of [28] that enables the system to achieve the global maximum of the potential function \( \Phi(A) \) and therefore avoiding local maxima that lead to bad equilibria.

However the main problem with this mechanism is the slow convergence rate. Therefore, in the next chapter we first introduce the fastest logit-response mechanism. It turns out that under the fast logit-response dynamic, \( G \) dynamics can be studied using the existing literatures on statistical physic which enables us to answer our main question that whether it is possible to achieve a good equilibrium in polynomial steps.

### III. MECE Logit-Response Mechanism

In the standard logit-response mechanism users find the chance to update their strategies with a fixed rate independent of the their positions in 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. User 1 as the first player selects strategy B since there is no other neighbor has selected any strategy yet. Then user 2 selects W since user 3 has not selected yet and there is no other neighbor for him. Then there is no payoff dominant strategy for user 3 at this stage and he randomly selects a strategy for example W. The same process repeats another round so that users reach the NE. However if user 3 selects his strategy before others as in Fig.2b the game ends up in NE within the first round. Assume that the graph to be more complex than this. It can be expected that game might converge to NE with exponential number of iterations. This implies that the standard logit-response should be modified with respect to some parameters such as the position of the users in the graph and dynamic of the system.

Consider the MECE logit learning as is shown in Algorithm 1. In MECE learning mechanism each user has equipped with Poisson alarm clocks with time/position varying rates of \( nq_i \) where \( \sum_i q_i = 1 \). The \( \{q_i, \forall i \in n\} \) are set according to \( Q^* \) as described in (6). When the users finds the chance they updates according to (15).

**Theorem 3.1** MECE logit-response is the fastest logit-response.

*Proof:* As we showed in the previous example the order of learning induces bad effects in the convergence speed. In fact the effect of these orderings falls in the term \( Z \) that appears in the stationary distribution \( \pi(A) \) of (16). It has been shown that removal of the term \( Z \) from stationary distribution can make the dynamics exponentially faster (refer to example 2 in [10] for more discussion). In the next lemma we show how the MECE logit satisfies such a situation.

**Lemma 3.1** The stationary distribution for \( G \) with the potential function \( \Phi \) under the modified logit-response is \( \pi(A) \propto \exp(\beta \Phi(A)) \).

*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. In fact the i.i.d assumption on the alarm distribution in the standard logit-response implies that it implements the NE with \( Q = \prod_{i=1}^{n} q_i \) with \( q_i = \frac{1}{n} \). Using this intuition and corollary 4.1 in [20] we can observe that there exits a joint probability distribution \( Q^* \) as that of (6) so that it removes the term \( Z \) from stationary distribution \( \pi(A) \).

Therefore the dynamics of the MECE Logit learning is the fastest any logit-learning can achieve.

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.

So far it can be concluded that the MECE logit-response mechanism has the fastest dynamic among other logit-response mechanisms and can be exponentially faster than the standard logit-response. However the main importance of the modified mechanism for this paper is that it maps the dynamic of the game \( G \) to that of Ising models as is described in the next theorem. By this mapping we then answer the main question of this work that how hard is to achieve a good equilibrium for \( G \) in polynomial time.
Algorithm 1 MECE Logit Response Learning

Initialization: $t = 0.$

1. A central authority solves the alarm clock probability distribution $Q^* = \arg\max_{Q \in C} \sum_{A} U(Q(A))$.
   
   a) Each user $i$ is equipped with Poisson alarm clocks of rate $nq_i$.
   
2. When the alarm goes off for a player $i$, $t = t + 1$
   
   Update strategy $j$ with probability $p_{ij} = \frac{\exp(\beta U_j(A(i:j)))}{\sum_{k=1}^{N_n} \exp(\beta U_k(A(i:k)))}$.

3. If $\forall_i \left| U_{t+1}^i - U_1^i \right| \leq \epsilon_{\text{stop}}$, Stop.

   Otherwise go back to step 1.

From now on up to the simulation section we consider a simplified version of game $G$ for ease of analysis. We consider $B = 2$ and $\Theta = 0$. Moreover assume if user $i$ is transmitting in channel 1, $a_i = -1$ and if it is transmitting in channel 2, $a_i = 1$.

**Theorem 3.2** The dynamic of the game with the modified learning mechanism coincides with the Glauber dynamics for the anti-ferromagnetic Ising models.

**Proof:** Proof of this theorem is straightforward by defining the Glauber dynamic of Ising models. Consider the strategy $\Phi(\mathbf{a})$ as the following procedure: starting from any initial condition, repeatedly choose a site $i \in V$ uniformly at random, replacing the spin of the site $i$ with one sampled from $\pi(A)$ conditioned on the spins of currently assigned to $N_i$. The Ising model is called a anti-ferromagnetic for $\pi(A) \propto -\sum_{i \in V} \sum_{j \in N_i} a_i a_j$, $\forall i \in V$. Ising is called ferromagnetic for $\pi(A) \propto \sum_{i \in V} \sum_{j \in N_i} a_i a_j$, $\forall i \in V$ [21].

Because of the memoryless properties of Poisson distribution and $\sum_i q_i = 1$, we can assume alarm clock distribution of MECE logit-learning is equivalent to i.i.d. rate 1 Poisson clocks. Using the simplified assumptions we can rewrite $\Phi(\mathbf{a}) = -\sum_{i \in V} \sum_{j \in N_i} a_i a_j$, $\forall i \in V$. Then the proof is complete by using Lemma 3.1.

So far we mapped the $G$ to an anti-ferromagnetic Ising model. In the next section we then establish the connection between the anti-ferromagnetic Ising model and the well discussed model of (ferromagnetic) Ising to clarify the effect of graph structure $G$ on dynamic of the system.

**IV. 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 $\beta < \beta_c$ where $\beta_T$ 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. This yields the next theorem.

**Theorem 4.1** For spectrum sharing game $G$ there is a graph $G$ that it is impossible to achieve a good equilibrium in polynomial steps.

**Proof:** We explained in the section II-A that to achieved a good equilibrium under logit-learning it is required that $\beta \to \infty$. On the other hand the Lemma 3.2 shows the game borrows the dynamic characteristic of Ising model and therefore a phase transition exists, i.e, for $\beta > \beta_c$ exponential steps in $n$ is needed for convergence. This implies a form of uncertainty principle in achieving a good equilibrium in polynomial steps in $n$ for spectrum sharing game $G$.

The previous theorem then arises the question that what kind of graphs show better behavior in terms of convergence speed. The answer is hidden in the Isoperimetric function of graphs. Considers all the possible subdivisions of the graph in two disjoint subsets of vertices: $S$ and its corresponding complement $V \setminus S$. The Isoperimetric function of graph $C(G)$ 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, $C(G) = \min_{S \subseteq V: |S|=\frac{n}{2}} \frac{\text{cut}(S,V \setminus S)}{|S|}$.

**Theorem 4.2** The smaller the value of Isoperimetric function of a graph $C_G$ is the faster anti-ferromagnetic Ising model dynamics converge.

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

$$\Phi(\mathbf{a}) = -\sum_{i \in V} \sum_{j \in N_i} W_{ij} a_i a_j, \forall i \in V$$

(17)

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

$$\Phi(\mathbf{a}) = \sum_{i \in V} \sum_{j \in N_i} W_{ij}^c a_i a_j - \sum_{i \in V} \sum_{j \in N_i} I_{ij} a_i a_j, \forall i \in V$$

(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 model.
over the complete graph. Now first let explain that the second term is independent of the phase transition analysis since we can look at it as a symmetric congestion game which has been proved to show no phase transition behavior [10]. Therefore the phase transition behavior of the anti-ferromagnetic Ising can be concluded using the phase transition behavior of Ising model on complementary graph \( G^c \). The first term phase transition behavior has been studied via Isoperimetric function of graphs in [11-27]. Using the result of [11-27] over complementary graph \( G^c \) that results in the absence of exponential mixing time and therefore the absence of phase transition.

Intuitively the previous theorem states that the more connected is the graph of interaction \( G \), the faster the game reaches an equilibrium.

A. Normalized Social welfare Optimization

The previous discussions we were addressing the question that what kinds of graphs show faster dynamic in achieving their good equilibrium. This should not be confused with the optimization problem of finding \( G \) that can achieve the maximal social welfare \( \sum_i U^i \). In fact we show there is a trade of between achieving the normalized maximal social welfare and the rate of convergence. Consider the optimization problem of (2). Our objective here is to find the graphs \( G \) that provide high social welfare \( U(\mathbf{A}) = \sum_{i \in V} U^i(\mathbf{A}(i, a_i)) \) therefore the optimization problem should be modified with a new notion of normalized social welfare defined as:

\[
\max_{\mathbf{A} \in \{-1, 1\}^n} \frac{U(\mathbf{A})}{2|E|} \tag{19}
\]

This is because we are looking for the graphs that have high capacity for achieving the optimal social welfare and therefore it should be normalized with respect to the number of edges \(|E|\).

**Theorem 4.3** For the spectrum sharing game \( \mathcal{G} \) there is a trade off between achieving the high rate of convergence and obtaining the optimal normalized social welfare for any graph \( G \).

**Proof:** Instead of solving the linear optimization problem of (19) let consider solving the quadratic problem by rewriting 
\( U(\mathbf{A}) = \sum_i U^i \) as 
\( U(\mathbf{A}) = -\sum_i \sum_{j \in N_i} (a_i - a_j)^2 \). We can rewrite in the graphical format 
\( U(\mathbf{A}) = -\sum_i \sum_{j \in N_i} (a_i - a_j)^2 = -\mathbf{A}^T \mathcal{L}_G \mathbf{A} \)

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

\[
\min_{\mathbf{A} \in \{-1, 1\}^n} \frac{\mathbf{A}^T \mathcal{L}_G \mathbf{A}}{\mathbf{A}^T \mathbf{A}} \tag{21}
\]

We assume \( G \) to be any arbitrary connected graph.

**Lemma 4.1** Let \( \lambda \) be the smallest nonzero eigenvalue of \( \mathcal{L}_G \), then

\[
\lambda = \min_{\mathbf{A} \in \mathbb{R}^n} \frac{\mathbf{A}^T \mathcal{L}_G \mathbf{A}}{\mathbf{A}^T \mathbf{A}} \leq \min_{\mathbf{A} \in \{-1, 1\}^n} \frac{\mathbf{A}^T \mathcal{L}_G \mathbf{A}}{\mathbf{A}^T \mathbf{A}} \tag{22}
\]

That is the optimal normalized social welfare is bounded below by \( \lambda \).

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

**Lemma 4.2** (Cheegers inequality [29]) Let \( \lambda \) be the smallest nonzero Laplacian eigenvalue of graph \( \varsigma \)-expander graph \( G \) then

\[
\varsigma^2 \leq \lambda \leq 2\varsigma \tag{23}
\]

That is \( \mathcal{C}(G) = \varsigma \) is upper bounded with \( \lambda \).

(22) states the graphs with lower value of \( \lambda \) provide larger normalized social welfare but (23) shows to have faster convergence larger \( \lambda \) are desired. **This shows a trade off between the normalized social welfare and the convergence speed.**

V. Spectrum Sharing Via Congestion

We saw in previous section that there exists graph \( G \) such as those who do not belong to expander graph family so that it is impossible to achieve a good equilibrium with polynomial mixing time. In this section we are investigating a method to reach a good equilibrium in any general graph. That is the mixing time for logit response is polynomial for \( \beta \to \infty \) independent of graph \( G \). The next theorem addresses the ultimate objective of this paper which is proposing a method to achieve the good equilibrium with polynomial mixing time for any graph \( G \).

**Theorem 5.1** Assume that each user \( i \) evaluates it’s utility as:

\[
U^i_h = -\sum_{j \in N_i} a_i a_j + a_i \epsilon \tag{24}
\]

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

**Proof:** To show the proof we follow these steps of reasoning:
_First_, Theorem 3.2 showed that dynamic of \( G \) can be analyzed by studying the dynamics of Glauber algorithms for anti-ferromagnetic Ising model. Moreover Theorem 4.2 maps the anti-ferromagnetic dynamics with graph \( G \) to that of ferromagnetic Ising models over complementary graph \( G^c \).

_Second_, with utility function of (23) then the best response strategy of user \( i \) can be written as \( \text{sign}(\sum_{j \in N_i} a_j + \epsilon) \). In this
way the term $\epsilon$ makes one specific strategy for user $i$ “risk dominant”. The risk dominant strategy for user $i$ is the one yielding the highest payoff when user $i$ have no information about its neighbors $N_i$. That is if half of the $N_i$ are active in a channel strategy and the other half are active in the other, the user $i$ will select the risk dominant one. [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 strategy. Let 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.

**Third**, We can write the social welfare as $U^i(A) = -(n_{11} + n_{22})$ 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 do not experience interference add zero value to the social welfare. We noticed that the logit-response leads to good equilibrium for $\beta \rightarrow \infty$ that is in the stationary distribution $n_{11} + n_{22}$ will be be minimized. However making one strategy risk-dominant affect the quality of achieved equilibrium solution. In fact the risk dominant strategy should be selected in the favor of the strategy with less energy defined as $E_1 = \sum_i \sum_{j \in N_i} \delta_{a_i, -1}\delta_{a_j, -1}$.

$E_2 = \sum_i \sum_{j \in N_i} \delta_{a_i, 1}\delta_{a_j, 1}$. By energy we mean the value of concentration of user on a specific strategy. For example in a complete graph due to symmetry of strategy configuration, the channel with more number of active users has higher energy. Energy also depends on the graphical characteristics of the users that have occupied a strategy. Then it is enough to flip the sign of $\epsilon$ in the favor of the strategy with less energy. For example 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 $(n_{11} + n_{22})$ since it prevents from existing any permanent risk dominant strategy in the network. Notice that if one strategy becomes risk dominant permanently without considering the energy difference then in the stationary state every user, with probability converging to 1 choses the same strategy and therefore increasing the existing interfering links.

The previous theorem states that to have an exponentially faster spectrum sharing users need to arrive an arbitrary small risk dominant term $\epsilon$ in their decision as far as the sign of $\epsilon$ favors the appropriate channel strategy (in the sense of energy). For example this sign can be adjusted through an advertisement entity who has access to some general knowledge of the $G$. The parameter $h > 0$ represents the trade off between the quality of the achieved equilibrium solution and convergence rate (by changing the scale of $\epsilon$) pretty similar to the temperature $\frac{1}{\beta}$. More interestingly we also observe in the simulation section (Fig. 6) that similar to the parameter $\beta$ a phase transition behavior occurs for $h > h_T$.

Producing these suitable risk-dominant terms can be based upon an advertisement entity such as a Base station. The exact solution for the right risk dominant terms for an anti-ferromagnetic Ising model is a hard problem but it can be approximated by a simple 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 decision time slot. In this way when users experience the same signal interference for both available channels, they select the one with less congestion.

Let’s $\rho_{a_i} = \frac{1}{|N_i|} \sum_{j \in V} \delta_{a_i, a_j}$ be associated with action profile $A$ and advertisement control parameter $h > 0$. Then $i$’s user utility is a measurable mapping of the following form:

$$U^i_h(A(i, a_i)) = U^i(A(i, a_i)) - h\rho_{a_i}$$

(25)

In fact in this way we generalize the (24) by presenting the energy difference as the congestion term $\rho$. We have applied this congestion advertisement method to the spectrum sharing game $G$ with utility format of (25) to validate the theoretical part. The results are explained in the next section.

**VI. Simulation Results**

Simulations have been done for more generalized versions than the theoretical part. It shows so many of our theoretical results maintain their validity even with change of some assumptions. There generalized assumptions are:

- Learning method is a Responsive Learning Automata (RLA) [24] with learning parameter $\alpha 
\in (0, 1)$. 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 $\theta$.
- Utility format is of the form $U^i(A(i, a_i)) = \frac{\theta_{a_i}}{\rho_{a_i}(1+|I_i|)}$.
- $G$ is a geometric graph.

A. Price of Anarchy

Fig. 3 shows the improvement of spectrum sharing with congestion advertisement mechanism in terms of network interference avoidance. Also it can be seen that by injecting congestion incentive into users utility herding to a spectrum with the higher quality happens with less probability which in turn reduces the interference.

B. Convergence

To show the convergence improvement the volatility concept has been plotted. 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 theorem 4.2. Fig. 5 also shows congestion advertisement method enhances the dynamics of $G$. 


Fig. 3. Price of Anarchy Improvement, $B = 4$ and the simulation has averaged over $100$ realizations, best $\alpha$ for RLA learning for each case has selected to make the comparison independent of learning process, the network interference avoidance is:

$$\frac{1}{|V|} \sum_{i \in V} \frac{1}{1 + |I_i|}.$$

C. Phase transition

We have run several simulations for over different values of learning parameter $\alpha$ and we have selected the best $\alpha$ that bring the highest social welfare in terms of signal interference $\frac{1}{1 + |I_i|}$ for different values of $h$. This resulted behavior shows a transition point Fig. 6. As can be seen at the beginning with increase in $h$ the exploration rate required to find the good equilibrium reduces. This improve the convergence rate. However keep increasing $h$ the system arrives at a transition point $h_T$ where users herd on the channel with less congestion and therefore starts to increase the signal interference which needs higher level of irrationality $\alpha$. This behavior is similar to the behavior of the Glauber dynamics for $\beta > \beta_T$.

REFERENCES


The total variation distance between two probability distributions $\mu$ and $\nu$ on $\mathcal{A}$ is defined by

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{i \in \mathcal{A}} |\mu_i - \nu_i|.$$  \hfill (26)

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

$$\tau(\epsilon) = \min_t \{\|\mu(t) - \nu\|_{TV} \leq \epsilon\}.$$  \hfill (27)

**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:

$$p^i_{t+1} = p^i_t + \alpha r^i_t \sum_{j \neq i} s^j_j p^j_t$$

$$\forall j \neq i, p^i_{t+1} = p^i_t - \alpha r^i_t s^j_j p^j_t$$

$$s^j_j = \min\{1, \frac{p^j_t - \alpha/2}{\alpha p^j_j r^i_t}\}$$

### VII. APPENDIX

**A. Laplacian Graph**

**Definition A.1** The Laplacian of the graph $G$ is defined as the $n \times n$ matrix $L_G = (L_{ij})$ in which $L_{ij} = \begin{cases} d_{ij} & i = j, \\ -w_{ij} & i \neq j. \end{cases}$

**B. Markov**

**Definition A.2** The total variation distance between two probability distributions $\mu$ and $\nu$ on $\mathcal{A}$ is defined by

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{i \in \mathcal{A}} |\mu_i - \nu_i|.$$  \hfill (26)