Abstract—This paper is concerned with decentralized selection of multimode precoding strategy for multiple-input multiple-output (MIMO) multiple access channels. We formulate it as a discrete noncooperative game. This game is shown to possess at least one pure strategy Nash equilibrium (NE) and the optimal strategy profile which maximizes the sum rate constitutes a pure strategy NE. Then we propose a decentralized algorithm based on learning automata to achieve the NE. A repeated mechanism is introduced to improve the sum rate performance and a mechanism for adapting step size is designed to control the convergence speed. Simulation results show that the proposed algorithm, which only requires limited feedback, can achieve near optimal or optimal sum rate performance.

Index Terms—Learning automata, MIMO, multimode precoding strategy selection, multiple access channels, potential game.

I. INTRODUCTION

MULTIMODE precoding strategy selection with limited feedback for multiuser (MU) multiple-input multiple-output (MIMO) systems is of great practical importance and has received considerable attention [1]. It is known that the optimal multimode precoding strategy can be found using exhaustive search. However, the computational complexity of exhaustive search grows exponentially with the total number of the antennas and users. Low-complexity algorithms have been proposed for antenna selection [2], [3], but they can not be used for other limited feedback precoding techniques. Furthermore, most of the existing multimode precoding strategy selection algorithms are centralized and will become impractical in future MU MIMO systems where heterogeneous users dynamically share the spectrum and the base station has no prior information of the users. Therefore, an efficient decentralized multimode precoding strategy selection algorithm with low-complexity is desired.

The goal of this work is to design a low-complex and decentralized multimode precoding strategy selection scheme for the MIMO multiple access channel (MAC). The MIMO MAC arises, for example, in cognitive MIMO radio systems [4], where heterogeneous users sense an idle spectrum at the same time and attempt to access the spectrum to communicate with a same receiver simultaneously, or in uplink MIMO cellular systems.

To meet the requirement of being decentralized, we formulate the problem as a discrete noncooperative game by satisfying a competitive optimality criterion, based on the achievement of Nash equilibrium (NE). We then propose a low-complex and decentralized algorithm based on learning automata to solve the game in a self-organized manner. Moreover, different from the conventional learning automata, an adaptive step size mechanism is designed to control the convergence speed of the algorithm.

Game theory has been applied to study the radio resource allocation in MU MIMO systems. The authors in [4], [5] focused on the MIMO interference channel where the strategies of the users are continuous variables and the proposed games are not potential games. Our proposed game is, on the other hand, for MIMO MAC, and proved to be a discrete potential game. The learning automata technique has been used to solve the discrete power control potential game in [6]. However, the potential of the game is not the desired social optimum. In our work, the social optimum constitutes an NE (i.e. competitive optimum) of our proposed game while the NE may not be the social optimum.

Notations: $s^\dagger$ denotes the conjugate transpose of $s$, $|\cdot|$ denotes determination.

II. SYSTEM MODEL

We consider $n$-user MIMO MAC equipped with $M_r$ antennas at the receiver and $M_t$ transmit antennas for user $i \in \mathcal{N}$, where $\mathcal{N} = \{1, \cdots, n\}$. Each user $i$ has $M_t$ radio frequency (RF) chains, where $1 \leq M_t \leq M_r$ and $\sum_{i=1}^{n} M_t \leq M_r$. Let $P_i$ denote the signal-to-noise ratio (SNR) of user $i$ when it is active. The transmit power of each user is assumed to be uniformly allocated. Let $L_i \in \mathcal{L}_i$ denote the number of transmission data substreams (i.e., mode) of user $i$, where $\mathcal{L}_i \subseteq \{0, 1, \cdots, M_t\}$ is the set of the supported mode values of user $i$. If $L_i > 0$, an $M_t \times L_i$ identical matrix $\mathbf{W}_i$, chosen from multimode precoding strategy codebook $\mathcal{W}_i$, maps $L_i$ data substreams to $M_t$ transmit antennas of user $i$. For each user $i$, we allow $L_i = 0$ and let the multimode precoding strategy be “0” which means that user selection can be implemented. The codebook $\mathcal{W}_i = \{\mathbf{W}_{i1}, \cdots, \mathbf{W}_{ik}, \cdots, \mathbf{W}_{IL_i}\}, \forall i \in \mathcal{N}$, is a finite set of multimode precoding strategies (matrices) for all the supported modes, where $\mathbf{W}_{ik}$ is the $k$-th strategy of user $i$ and $K_i$...
is the cardinality of $\mathcal{W}_i$. Note that, using different codebooks does not affect our theoretical results in this letter.

Let $\mathbf{H}_i$ denote the $M_r \times M_r$ channel matrix of user $i$ with each entry modeled as independent identically distributed (i.i.d.) zero-mean unit-variance circularly symmetric complex Gaussian. We assume that the number of users and the channel matrices $\{\mathbf{H}_i\}_{i \in \mathcal{N}}$ are fixed during each transmission block and change independently from one block to another according to ergodic random processes. Let $\bar{\mathbf{H}}_i = \mathbf{H}_i \mathbf{W}_i$ denote the $M_r \times I_r$ matrix associated with the selected matrix $\mathbf{W}_i$ of user $i$. The MIMO MAC with multimode precoding strategy selection can be modeled as

$$
y = \sum_{i=1}^{n} \mathbf{H}_i \mathbf{W}_i x_i + z = \sum_{i=1}^{n} \bar{\mathbf{H}}_i x_i + z \tag{1}
$$

where $y(M_r \times 1)$ is the channel output with multimode precoding strategy selection, $x_i(I_r \times 1)$ is the channel input of user $i$, $z(M_r \times 1)$ is a vector of additive white Gaussian noise (AWGN) with zero mean and covariance matrix equal to $\mathbf{I}_{M_r}$.

### III. GAME-THEORETIC MULTIMODE PRECODING STRATEGY SELECTION

#### A. Problem Formulation and Game Model

Since we focus on designing decentralized algorithm, we assume that the coordination mechanism is unavailable and each user’s transmission can be looked as a source of interference for the others. Furthermore, if the users are assumed to be selfish and strictly rational agents, they will compete to maximize their individual mutual information with multiuser interference treated as noise. This motivates us to define the payoff $u_i$ for user $i$ as

$$
u_i = \frac{1}{2} \log_2 \left( \frac{p_i}{I_i \bar{\mathbf{H}}_i^H \bar{\mathbf{H}}_i + N_i} \right) - \frac{1}{2} \log_2 |N_i| \tag{2}
$$

where $N_i = \sum_{j \in \mathcal{N}, j \neq i} (p_j / I_j) \bar{\mathbf{H}}_j^H \bar{\mathbf{H}}_i + \mathbf{I}_{M_r}$. Note that this payoff function is not the actual rate achieved by the user at the receiver.

Thus, the users along with their multimode precoding strategies and payoffs constitute a discrete noncooperative game denoted as $\mathcal{G}^E = \{\mathcal{N}, \{\mathcal{W}_i\}_{i \in \mathcal{N}}, \{u_i\}_{i \in \mathcal{N}}\}$, where the rational users are regarded as the players, the codebook $\mathcal{W}_i$ is the pure strategy (i.e., action) set of user $i$, and the individual mutual information of each user is its payoff.

Formally, the game $\mathcal{G}^E$ is expressed as

$$(\mathcal{G}^E) : \max_{\mathbf{W}_i \in \mathcal{W}_i} u_i(\mathbf{W}_i, \mathbf{W}_{-i}), \text{ for all } i \in \mathcal{N} \tag{3}$$

where $\mathbf{W}_{-i}$ denotes the multimode precoding strategy selection strategies (matrices) of all the users excluding the $i$th user.

#### B. Analysis of NE

In this subsection, we analyze the NE of the proposed game. The property of the NE is studied in the following theorem.

**Theorem 1:** $\mathcal{G}^E$ is a potential game which possesses at least one pure strategy NE and the optimal solution of the sum rate maximization problem

$$(P1) \max_{\mathbf{W}_i \in \mathcal{W}_i} R_{\text{sum}} = \frac{1}{2} \log_2 \left( \sum_{i=1}^{n} \frac{p_i}{I_i \bar{\mathbf{H}}_i^H \bar{\mathbf{H}}_i + \mathbf{I}_{M_r}} \right) \tag{4}$$

constitutes a pure strategy NE of $\mathcal{G}^E$.

**Proof:** We first propose a game $\mathcal{G}^C = \{\mathcal{N}, \{\mathcal{W}_i\}_{i \in \mathcal{N}}, \{u_i^C\}_{i \in \mathcal{N}}\}$, where the payoff function of each user $i$ is defined as $u_i^C = R_{\text{sum}}$. Assume that $(\mathbf{W}_{1}^C, \ldots, \mathbf{W}_{n}^C)$ is an optimal solution of $P1$ (i.e. the maximizer of $R_{\text{sum}}$). Then, $\forall i \in \mathcal{N}$, $\mathbf{W}_i^C \in \mathcal{W}_i$; $\mathbf{W}_i^C \neq \mathbf{W}_i^*$ is an alternate strategy of user $i$, then we have

$$u_i^C (\mathbf{W}_i^C, \mathbf{W}_{-i}^C) \leq u_i^C (\mathbf{W}_i^*, \mathbf{W}_{-i}^C). \tag{5}$$

That is, if an arbitrary user changes its own strategy unilaterally, it can never get a payoff which is higher than the payoff obtained by $(\mathbf{W}_1^*, \ldots, \mathbf{W}_n^*)$. Thus, according to the definition of NE in [7], the strategy profile $(\mathbf{W}_1^*, \ldots, \mathbf{W}_n^*)$ is a pure strategy NE of $\mathcal{G}^C$.

Assume that $\mathbf{W}_i^* \in \mathcal{W}_i$ is an alternate strategy of user $i$. We define $\bar{\mathbf{H}}_i^*$ as $\bar{\mathbf{H}}_i^* = \mathbf{H}_i \mathbf{W}_i^*$ and let $L_i^*$ denote the mode corresponding to the strategy $\mathbf{W}_i^*$. Then

$$u_i(\mathbf{W}_i, \mathbf{W}_{-i}) - u_i(\mathbf{W}_i^*, \mathbf{W}_{-i})$$

$$\leq \frac{1}{2} \log_2 \left( \frac{p_i}{L_i} \bar{\mathbf{H}}_i^* \bar{\mathbf{H}}_i^* + N_i \right) - \frac{1}{2} \log_2 \left( \frac{p_i}{L_i} \bar{\mathbf{H}}_i \bar{\mathbf{H}}_i + N_i \right)$$

$$= R_{\text{sum}}(\mathbf{W}_i, \mathbf{W}_{-i}) - R_{\text{sum}}(\mathbf{W}_i^*, \mathbf{W}_{-i}). \tag{6}$$

In (6), replace the sum rate $R_{\text{sum}}$ by potential function $\Phi$, we obtain (7). Thus the game $\mathcal{G}^E$ satisfies potential game definition [8]:

$$\forall \mathbf{W}_i, \mathbf{W}_i^* \ni u_i(\mathbf{W}_i, \mathbf{W}_{-i}) - u_i(\mathbf{W}_i^*, \mathbf{W}_{-i}) = \Phi(\mathbf{W}_i, \mathbf{W}_{-i}) - \Phi(\mathbf{W}_i^*, \mathbf{W}_{-i}). \tag{7}$$

That is, $\mathcal{G}^E$ is a potential game with the sum rate $R_{\text{sum}}$ as the potential function $\Phi$. From the Lemma 2.1 in [8], the set of NEs of $\mathcal{G}^E$ coincides with the set of NEs of $\mathcal{G}^C$. Thus $(\mathbf{W}_1^*, \ldots, \mathbf{W}_n^*)$ is also the pure strategy NE of $\mathcal{G}^E$. Hence the theorem is proved.

However, in game $\mathcal{G}^E$, normally several NEs exist, and some NEs are only suboptimal solutions of $P1$. Moreover, the number of the NEs of $\mathcal{G}^E$ is difficult to be confirmed [7].

#### C. Algorithm

Assume that the knowledge of $\{\bar{\mathbf{H}}_i\}_{i \in \mathcal{N}}$ is perfectly available at the common receiver, but unknown at each user transmitter. It is known that best response dynamic (BRD) can be used to find a pure strategy NE of a potential game [8]. However, the BRD needs to know the current channel state of the active user and the interference from the other users which may cause additional overhead in the system, and therefore cannot be applied here. Then, we propose a decentralized stochastic learning algorithm which employs a team of learning automata to evolve to the NE of $\mathcal{G}^E$ with only a small amount of feedback.

To characterize the learning algorithm, we extend the game $\mathcal{G}^E$ to a mixed strategy form. Let $p_i = \{p_{i1}, \ldots, p_{iK_i}\}$ be the mixed strategy of the user $i$, where $p_{ik}$ denotes the probability with which the $k$th user chooses the $k$th pure strategy. The expected payoff $g^E$ for user $i$ is given by

$$g^E(p_1, \ldots, p_i) = E[u_i|jth \ user \ employs \ strategy \ p_j, \ 1 \leq j \leq n]$$

Authorized licensed use limited to: Shanghai Jiao Tong University. Downloaded on July 12,2010 at 04:06:48 UTC from IEEE Xplore. Restrictions apply.
denote the is a predetermined con-

If the mixed strategy multimode precoding strategy selection game is played successively, it could be modeled as a stochastic game of learning automata. Each user (i.e., player) is represented by a learning automaton and the actions of the automaton are the pure strategies of the user. The mixed strategy \( \mathbf{p}_i(t) = \{p_{i1}(t), \ldots, p_{iK_i}(t)\} \) is the action probability distribution of the \( i \)th automaton at instant \( t \). \( p_{ik}(t) \) denotes the probability with which the \( i \)th automaton chooses the \( k \)th pure strategy at instant \( t \). The normalized payoff to the \( i \)th user will be the reaction \( r_i(t) \) to the \( i \)th automaton. We let \( r_i(t) = \alpha \mathcal{C}_i(t) \) be the normalized \( t_i^c(t) \) (the payoff of the \( j \)th automaton), where \( t_i^c(t) \) is the payoff of the \( \mathcal{G}_j \) defined in the proof of Theorem 1 and \( 0 < \alpha < 1 \) is a parameter to guarantee the value of \( r_i(t) \) lies in the interval \([0,1)\). Then \( r_1(t) = \cdots = r_n(t) = r(t) = \alpha \mathcal{C}_i(t) \). When one of the automata chooses an action independently according to its current action probabilities, the game is said to be played for once. The game is played repeatedly to learn the NE. Let \( \mathbf{p}_i(0) \) denote the initial mixed strategy of user \( i \). The learning algorithm used by each of the users is given as follows.

**Decentralized Multimode Precoding Strategy Selection Algorithm (DMPSSA):**

1. Set the initial probability vector \( \mathbf{p}_i(0) \) as \( p_{ik}(0) = 1/K_i, \) \( i = 1, \ldots, n; \) \( k = 1, \ldots, K_i \).
2. At every time instant \( t \), each user chooses an action \( \mathbf{W}_i^t \) according to its action probability vector \( \mathbf{p}_i(t) \).
3. Each user obtains a reaction \( \alpha \mathcal{C}_i(t) \) based on the set of all actions from the receiver. Note that all the users have the same reaction which can be easily broadcasted from the base station to the users.
4. Each user updates its action probability through the rule

\[
\begin{align*}
p_{ik}(t+1) &= p_{ik}(t) - b_i^t(t) p_{ik}(t), \\
p_{ik}(t+1) &= p_{ik}(t) + b_i^t(t) (1 - p_{ik}(t)),
\end{align*}
\]

where \( 0 < b < 1 \) is a step size parameter.
5. If \( \mathcal{G} \in \mathcal{N} \), there exists a component of \( \mathbf{p}_i(t) \) which is larger than a value approaching one, say 0.99, stop. Otherwise, go to step 2.

The DMPSSA determines the mixed strategies for the users by the history of play. At each instant, the required information of each user includes nothing but his normalized payoff after each play and his current action. The receiver and the users need not to know any prior information of the system. The users only compute the probabilities (real numbers) of their actions and the receiver only computes a normalized common payoff. Compared with the exhaustive search, which needs to compute and compare \( \prod_{i=1}^n K_i \) combinations of rate functions, the complexity of the proposed DMPSSA is significantly reduced.

**Theorem 2:** With a sufficiently small step size \( b \), DMPSSA converges to a pure strategy NE of \( \mathcal{G}^F \). When the NE of \( \mathcal{G}^F \) is unique, DMPSSA converges to the optimal solution of \( P1 \).

**Proof:** Since \( r_i = \alpha \mathcal{C}_i = \alpha \mathcal{R}_{sum_i} \forall i \in \mathcal{N} \), then the game \( \mathcal{G}^\mathcal{V} = [\mathcal{N}, \{\mathcal{V}_i\}_{i \in \mathcal{N}}, \{r_i\}_{i \in \mathcal{N}}] \) is a game with common payoff. From theorem 4.1 in [9], the DMPSSA converges to a pure strategy NE of \( \mathcal{G}^\mathcal{V} \) with sufficiently small \( b \). Based on [7], it is easy to prove that the NEs of \( \mathcal{G}^\mathcal{V} \) coincides with the NEs of \( \mathcal{G}^\mathcal{C} \). From the Lemma 2.1 in [8], when DMPSSA finds a pure strategy NE of \( \mathcal{G}^\mathcal{C} \), it also finds a pure strategy NE of \( \mathcal{G}^F \). From Theorem 1, when the NE is unique, it must be the optimal solution of \( P1 \). Hence the theorem is proved.

From Theorem 1 and Theorem 2, when multiple pure strategy NEs exist, we can run DMPSSA several times and then choose the strategy profile with the highest payoff.

In conventional learning automata, \( \eta \) is a predetermined constant and has great impact on the convergence speed. Normally, the larger \( \eta \) is, the faster DMPSSA converges. In this paper, we design a mechanism to adjust the step size, thereby speeding the convergence while still approaching the optimal solution of \( P1 \). In specific, we define the time-varying \( b \) as:

\[
b = \begin{cases} 
  b_1, & 0 < t < T_1 \\
  b_2, & T_1 \leq t < T_2 \\
  \cdots, & \cdots \\
  b_m, & t \geq T_{m-1}
\end{cases}
\]

where \( T_1 < \cdots < T_{m-1} \) are ordered positive integers, \( T_m \) is defined as positive infinity, \( 0 < b_1 < \cdots < b_m < 1 \) denote the ordered step size values, \( m \) is a finite positive integer.

From the proof of Theorem 2, we note that this adaptive step size mechanism does not affect our theoretical results. However, the values of the parameters should be properly set to adapt the requirement in practice. The adaptive step size mechanism is very useful in practice, as it can adaptively limit the iteration time of DMPSSA under a desired level.

**IV. SIMULATION RESULTS**

In this section, the performance of the proposed algorithm is evaluated via computer simulations. The parameters of the system are set as follows: \( n = 4, M_1 = M_2 = 2, M_3 = M_4 = M_5 = 3, M_1 = M_2 = 1, M_3 = M_4 = 2, M_5 = 6, L_1 = L_2 = \{0,1\}, L_3 = L_4 = \{0,1,2\}, \alpha = 0.02, \) and \( \rho_i = \rho, \forall i,j \in \mathcal{N} \). Without loss of generality, the users are assumed to use the multimode antenna selection codebooks proposed in [10] (the simplest form of multimode precoding codebooks) which are denoted as: \( \mathcal{W}_1 = \mathcal{W}_2 = \{0,1,2\} \), \( \mathcal{W}_3 = \mathcal{W}_4 = \{0,1,0,1,0,0\} \). For a specific channel realization, Fig. 1 shows the evolution of the choice probabilities of the actions (i.e., mixed strategy) when SNR = 10 dB with \( b = 0.1 \) (upside), \( b = 0.4 \) (middle), and adaptive step size (downside). When the adaptive step size mechanism is used, we let \( m = 2, b_1 = 0.1, b_2 = 0.4, T_1 = 60 \). It is shown in Fig. 1 that DMPSSA has good convergence. For example, user 1 converges to the first strategy after about 140 iterations \( (p_{11} = 1) \) when \( b = 0.1 \), user 3 converges to the first strategy \( (p_{13} = 1) \) after about 38 iterations when \( b = 0.4 \), and the third strategy about 120 iterations \( (p_{13} = 1) \) when the adaptive step size mechanism is used. These results show that the convergence speed is fast when \( \eta \) is large and the number of the users’ strategies is small. Furthermore, DMPSSA may converge
to different NEs with different values of $b$ even for the same channel realization.

Fig. 2 plots the average sum rates versus SNR by simulating a total of $10^3$ independent trials. It shows that DMPSSA can significantly outperform the random selection scheme at all SNR region. The sum rate performance gap between the exhaustive search and DMPSSA becomes large when SNR becomes high. If we run DMPSSA two times and choose the NE with the larger payoff for each channel realization, the average sum rate performance can be improved. If DMPSSA is run six times, the sum rate performance can be improved further and even can achieve the optimal performance when $b = 0.1$.

From Fig. 2, it is also shown that the performance of $b = 0.1$ is better than the performance of $b = 0.4$. This is because multiple NEs exist in most situations and it is very likely for DMPSSA to miss the optimal (or near-optimal) NE when $b$ becomes large. This result also verifies the statement at the end of Section III-B. For DMPSSA, we can expect that the larger is the value of $b$, the worse the obtained performance becomes.

From Figs. 1, 2, we find that the proposed adaptive step size mechanism can strike a good balance between sum-rate performance and convergence speed. These results verify that the adaptive step size mechanism is efficient.

In practice, we can tune the value of the step size parameter $b$ or use adaptive step size mechanism to trade off the performance and the complexity according to the channel conditions and the requirement of the system designer. In summary, DMPSSA is flexible and efficient.

V. CONCLUSION

In this paper, we studied the decentralized multimode precoding strategy selection for MIMO MAC through a game theoretic approach. There exists pure strategy NE in our proposed game and the social optimal strategy profile constitutes an NE. A decentralized learning algorithm is proposed to obtain the NE. Repeating the proposed algorithm can improve the sum rate performance. An adaptive step size mechanism is further designed to tradeoff between performance and convergence speed. Compared to the exhaustive search, the proposed algorithm can achieve a comparable sum rate performance with low complexity.

REFERENCES