2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Learning Behaviors for Coordination in Networked Systems and a Generalized Consensus Protocol

Pedram Hovareshti and John S. Baras

Abstract— **We consider the problem of coordination for efficient joint decision making in networks of autonomous agents. When making a decision, an agent is influenced by its knowledge about others' behaviors. Agents' understanding of others' behaviors is shaped through observing their actions over a long time. We have modeled the decision making on whether to cooperate in a group effort as a result of a series of two-person games between agents, where the payoff of each agent is computed as the sum of its payoffs from each of these games. The agents initially have different behaviors. In order to maximize their pay-off, they need to learn the others' behavior and coordinate with them. We consider a behavior learning algorithm for a class of behavior functions and study its effects on the emergence of coordination in the network. The conditions under which the learning algorithm converges are studied. We show that for a class of linear functions the learning algorithm results in an extension of non-homogeneous consensus protocol to the more general case of block-stochastic matrices.**

I. INTRODUCTION

In a network of autonomous agents the interdependence of agents can be modeled by a three level abstraction. *Connectivity graph* represents the geographical dependence of agents and is useful in determining how the agent sense each others' presence; *communication graph* determines which agents can communicate, and *collaboration graph* determines the agents that are coordinated in performing a task.

In [1] we proposed a learning algorithm based on the Cucker-Smale-Zhou paradigm of 'language learning' [2] to establish the dependence of a group's collaboration graph on its communication graph. The system consists of a group of agents. The agents can be machines or humans with different degrees of rationality. Each agent has to make a decision on whether to cooperate or not in a group effort. This is modeled by two-person coordination games between neighboring agents. The payoff of each agent is computed as the sum of the agents' payoffs from each of these games. After the games are played, a collaboration graph is formed from all the agents who have decided to collaborate. Each agent's decision on whether to collaborate is based on its personal understanding of its own behavioral tendencies as well as its neighbors'. To account for this fact we provision the agents with a behavioral variable, which indicates how risk averse an agent is. We model how the agents learn and adapt to each other's behaviors based on Cucker-Smale-Zhou model of language leaning. Agents exchange messages before playing the game and based on these exchanges try to learn their neighbors' behavior and adapt to them. The effect of the agents on each other is governed by an influence matrix which is partially derived by the communication graph's topology. If the agents are allowed to interact for a sufficiently long period of time before the game, a consensus will be reached [under certain conditions] on which equilibrium should be played. Equilibrium points corresponding to more agents collaborating are desirable.

This work focuses on two major issues: the emergence of collaboration based on the behavior adaptation and development of a 'generalized consensus protocol' when behavior functions are from a class of linear functions. The emergence of cooperation and conventions has applications in many social, economical and political studies [3], [4], [5]. It is explained in game theory literature by considering boundedly-rational agents who play a game indefinitely against fixed, or randomly matched agents, and learn from the previous outcomes to achieve a notion of optimality or equilibrium [6]. This framework is also adapted to study cooperation in networks and network formation [7], [8], [9]. It is usually assumed that conventions result as equilibria of coordination games [8], [7]. The trade-off between stability of network formations and costs of link establishment are also studied using cooperative game theory [10].

This work differs from the above-mentioned literature in at least two respects. First, we do not consider the problem of network formation, rather we focus on the collaboration networks possible given an underlying communication framework. In this respect, our objective is similar to that of [11] with the difference that [11] considers a $n-$person game with emphasis on deterministic formation of common knowledge, whereas we consider a set of $2-$person games and consider the agent's learning as a result of their observations in their neighborhoods. Second, we consider agents to learn the types and behaviors rather than the strategies as there is the case with evolutionary game theory. In other words, the adaptation provides the agents with a mechanism to understand how the other agents see the world so that they get a better chance to coordinate on a single strategy. (See [1] for more in depth discussion)

The paper is organized in the following sections. The model of the game and the learning process is provided in Section II. Section III provides the analysis of the system. Simulations of illustrating examples are included in Section IV. Section V concludes the paper.

Research supported by BAE Systems award number W911NF-08-2-0004 under the ARL MAST CTA (Collaborative Technology Alliance), the US Air Force Office of Research MURI award FA9550-09-1-0538, and by the Defence Advanced Research Projects Agency (DARPA) under award number 013641-001 for the Multi-Scale Systems Center (MuSyC), through the FRCP of SRC and DARPA.

Authors are with the Institute for Systems Research (ISR), University of Maryland College Park [hovaresp,baras]@umd.edu.

II. System Model

The agents are modeled as the nodes of a given graph. Each node has to take a decision on whether to cooperate (C) or not cooperate (NC) in a group effort. Based on its decision, a given node will acquire a payoff as in the coordination game setting of [1]. This model will be formalized in section II-A. We consider agents with a behavioral state or type, which determines the strategy they choose. We consider each agent's type to be defined as a function that maps a random input to a deterministic output. The idea is to capture the notion of how the agent interprets a random input: e.g. if the agent is told that an event occurs with a certain probability, how likely it is for them to believe that the event occurs. This will be formalized in section II-B. Using the Cucker-Smale-Zhou framework of language acquisition [2], we propose a model in which an agent's *strategy* depends on what it knows about its neighbors' *behaviors*. To this end, we consider a learning model in which by observing the previous behavior of their neighbors, each agent modifies its own behavior. If the agents are successful to reach a consensus on their types, they will get a higher expected utility and may achieve the Pareto optimal equilibrium. The learning process model is formalized in section II-C.

A. The coordination games

We consider a set of n agents and model the interconnection between them by a communication graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The nodes of the graph, $\mathcal{V} = \{1, 2, ..., n\}$ represent the agents and the undirected edges $\mathcal{E} = \{l_1, l_2, ..., l_e\} \subseteq \mathcal{V} \times \mathcal{V}$ represent the communication links. Each agent in $\mathcal{V}$ has to take a decision on whether to cooperate (C) or not cooperate (NC) in a group effort. This is modeled by considering each agent being engaged in a 2×2 coordination game with each of its neighbors. Let $a > b > c > 0$. In case that both of the agents cooperate (do not cooperation), each is rewarded by a (resp. c). If only one of the agents decides to cooperate, it will be rewarded by 0, whereas the non-cooperating agent will be rewarded by b. We denote the set of neighbors of agent i by N_i. The set of pure strategies for each agent is $S = \{C, NC\}$.

The over all pay-off of agent i is given by:

$$u_i(s_i, s_{-i}) = \begin{cases} a \sum_{j \in N_i} 1_{\{s_j = C\}}, & \text{if } s_i = C, \\ b \sum_{j \in N_i} 1_{\{s_j = C\}} + \\ c \sum_{j \in N_i} 1_{\{s_j = NC\}}, & \text{if } s_i = NC. \end{cases} \quad (1)$$

B. Behavior modeling

Consider that each agent has an evolving behavior system that decides on its level of optimism. We model the evolution of this behavior system in the Cucker-Smale framework of 'language evolution' [2].

Consider a set of uniformly distributed random variables on $X = [0, 1]$, here referred to as states. Consider $Y \subseteq R$ be the space of observations.

Definition 2.1: The behavior of an agent is a piecewise continuous function $f : X \to Y$.

Given a uniformly distributed random variable $x \in X$, $f_i(x)$ determines whether agent i expects an event that is supposed to occur with probability x, to actually happen. The idea is to consider an agent as an optimist if it assigns to small values of x large values of y.

To be able to study the convergence of behaviors, we need a notion of distance. Since x is uniformly distributed, distance between two behaviors can be defined as follows:

Definition 2.2: The distance between two behaviors f and g is determined by:

$$d(f, g) = (\int_0^1 |f(x) - g(x)|^2)^{1/2} dx$$

Remark 2.1: Here, we consider the set of linear functions $\mathcal{F}_l = \{f | f(x) = \theta x + \gamma; \theta, \gamma \in R\}$ and bounded linear functions $\mathcal{F^*}_l = \{f | f(x) = \theta x + \gamma; \theta \in [\theta_{min}, \theta_{max}]; \gamma \in [\gamma_{min}, \gamma_{max}]\}$, as sets of allowed behaviors.

C. Learning algorithm

To model the agents' partial understanding of the other agents' types, we need to model the communication that happens before the game is played. In reality, people make assumptions on peers' judgement system based on their observation of how their peers react to random events. We model this aspect by assuming that that the system evolves in a synchronized manner: at each time interval t, all the agents receive data from their neighbors in the form of $\{(x_j(t), y_j(t))\}_{j \in N_i}$. We assume that $x_j(t)$ is distributed uniformly on $X = [0, 1]$ and $y_j(t) = f_j(x_j(t))$, where $f_j(x_j(t)) \in \mathcal{F^*}_l$.

We model the relative influence and credibility of the agents as perceived by other agents by a stochastic matrix $W = [w]_{ij}$, where w_{ij} denotes the relative influence of agent j on agent i. We assume that $w_{ii} > \gamma > 0$ for a given $\gamma > 0$. Unless otherwise mentioned, in this paper we assume that all neighboring agents' influence is uniform, i.e..

$$w_{ij} = \begin{cases} 1, & \text{if } j \in N_i, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Having defined the influence matrix, the learning algorithm requires that at each time, all agents update their behavior function according to the following equation:

$$f_i(x_i(t+1)) = arg \min_{f \in \mathcal{F^*}_l} \sum_{j \in N_i} w_{ij} (f_j(x_j(t)) - y_j(t))^2,$$
$$i = 1, 2, ..., n. \quad (3)$$

The learning dynamics given by the set of equations (3) determine the evolution of the behavior of the system in the sense that it describes how each agent changes its behavior to make it more aligned with the behavior of others, upon getting more information that changes its understanding of the other agents' behaviors. Coordination on any policy happens if the agents reach a consensus on their behavior.

III. Analysis

The learning algorithm of the previous section constitutes a stochastic dynamical system in which the dynamics are determined by the samples of random inputs provided to agents. The following convergence theorem, stated without proof, follows as a corollary to Theorem 1 in [2].

Theorem 3.1: If all the agents use bounded linear behavior functions $f \in \mathcal{F}^*{}_l$, the learning algorithm defined in Section II, converges with probability 1 to a consensus on the behavior functions, provided that the matrix W is irreducible.

At each time instant t, agent i minimizes its cost function $J(i) = \sum_{j\in N_i} w_{ij}(f_j(x_j(t)) - y_j(t))^2$ over the set $\mathcal{F}^*{}_l$. The set $\mathcal{F}^*{}_l$ is bounded, closed and convex, so the minimum will be attained. Denote the minimizer parameters as $\theta^*(i)$ and $\gamma^*(i)$. The sufficient and necessary optimality conditions are:

$$\frac{\partial J(i)}{\partial \theta(i)}|_{\theta(i)=\theta^*(i)} = 0, \text{if } \theta_{min} < \theta^*(i) < \theta_{max} \quad (4)$$

$$\frac{\partial J(i)}{\partial \gamma(i)}|_{\gamma(i)=\gamma^*(i)} = 0, \text{if } \gamma_{min} < \gamma^*(i) < \gamma_{max} \quad (5)$$

$$\frac{\partial J(i)}{\partial \theta(i)}|_{\theta(i)=\theta^*(i)} \geq 0, \text{if } \theta^*(i) = \theta_{min} \quad (6)$$

$$\frac{\partial J(i)}{\partial \gamma(i)}|_{\gamma(i)=\gamma^*(i)} \geq 0, \text{if } \gamma^*(i) = \gamma_{min} \quad (7)$$

$$\frac{\partial J(i)}{\partial \theta(i)}|_{\theta(i)=\theta^*(i)} \leq 0, \text{if } \theta^*(i) = \theta_{max} \quad (8)$$

$$\frac{\partial J(i)}{\partial \gamma(i)}|_{\gamma(i)=\gamma^*(i)} \leq 0, \text{if } \gamma^*(i) = \gamma_{max} \quad (9)$$

If at time t the minimizer does not belong to the boundary of $\mathcal{F}^*{}_l$, equations 4 and 5 give linear regression formulae:

$$\theta_i(t+1) = \frac{\sum_{j\in N_i} x_j(t)y_j(t) - \frac{1}{n_i}\sum_{j\in N_i} x_j(t)\sum_{j\in N_i} y_j(t)}{\sum_{j\in N_i} x_j(t)^2 - \frac{1}{n_i}(\sum_{j\in N_i} x_j(t))^2}$$

$$\lambda_i(t+1) = \frac{1}{n_i}\sum_{j\in N_i} y_j(t) - \frac{\theta_i(t+1)}{n_i}\sum_{j\in N_i} x_j(t), \quad (10)$$

where, N_i denotes the neighborhood of agent i and n_i denotes the number of agent i's neighbors. Using the fact that $y_j(t) = \theta_j(t)x_j(t) + \gamma_j(t)$, and denoting $Q_i(t) = \sum_{j\in N_i} x_j^2(t), Q(t) = \sum_{j=1}^n x_j^2(t), S_i(t) = \sum_{j\in N_i} x_j(t)$, $S(t) = \sum_{j=1}^n x_j(t), \bar{x}_i(t) = \frac{1}{n_i}\sum_{j\in N_i} x_j(t)$, and $S_i^{(i)}(t) = \sum_{j\in N_i - \{i\}} x_j(t) = S_i(t) - x_i(t)$, we can rewrite Equation (10) as:

$$\begin{aligned}\theta_i(t+1) &= \frac{1}{Q_i(t)}\big[\sum_{j\in N_i}\theta_j(t)x_j^2(t) + \sum_{j\in N_i}\gamma_j(t)x_j(t) \\ &\quad - \gamma_i(t)S_i(t)\big] \\ \gamma_i(t+1) &= \frac{1}{n_i(t)}\big[\sum_{j\in N_i} x_j(t)\theta_j(t) + \sum_{j\in N_i}\gamma_j(t) \\ &\quad - \theta_i(t)S_i(t)\big].\end{aligned} \quad (11)$$

At each time instant all agents calculate their corresponding θ and γ using Equation (11). If the calculated θ and γ are respectively inside the intervals $[\theta_{min}, \theta_{max}]$ and $[\gamma_{min}, \gamma_{max}]$, their values are accepted, otherwise they are set to the corresponding boundary.We will now show that for the case of complete graph Theorem 3.1 still holds when we relax the constraint on the behavior functions so that $f \in \mathcal{F}_l = \{f | f(x) = \theta x + \gamma; \theta, \gamma \in R\}$.

A. The case of complete graphs

Let $\Theta = [\theta_1 \ \ \theta_2 \ \ \ldots \ \ \theta_n \ \ \gamma_1 \ \ \gamma_2 \ \ \ldots \ \ \gamma_n]^T$. The learning dynamics of the system can be written as:

$$\Theta(t+1) = F(t)\Theta(t) = \left[\begin{array}{c|c} P_1(t) & M_1(t) \\ \hline M_2(t) & P_2(t)\end{array}\right]\Theta(t), \quad (12)$$

where

$$[P_1(t)]_{ij} = \frac{x_j(t)^2}{Q_i(t)}, \quad [P_2(t)]_{ij} = \frac{1}{n},$$

$$[M_1(t)]_{ij} = \begin{cases} \frac{x_j(t)}{Q_i(t)}, & \text{if } i \neq j, \\ -\frac{S_i^{(i)}(t)}{Q_i(t)}, & \text{if i=j,}\end{cases}$$

and

$$[M_2(t)]_{ij} = \begin{cases} \frac{x_j(t)}{n(t)}, & \text{if } i \neq j, \\ -\frac{S_i^{(i)}(t)}{n(t)}, & \text{if i=j.}\end{cases}$$

The matrix $F(t)$ has some interesting properties.[1] Denote $1_n = [1 \ \ 1 \ \ \ldots \ \ 1]^T \in R^n$. It can be readily checked that:

$$P_1 1_n = 1_n, \quad M_1 1_n = 0,$$

$$M_2 1_n = 0, \quad P_2 1_n = 1_n,$$

and $F1_{2n} = 1_{2n}$. Furthermore, because for $\alpha, \beta \in R$

$$\left[\begin{array}{c|c} P_1 & M_1 \\ \hline M_2 & P_2\end{array}\right]\left[\begin{array}{c}\alpha 1_n \\ \hline \beta 1_n\end{array}\right] = \left[\begin{array}{c}\alpha 1_n \\ \hline \beta 1_n\end{array}\right], \quad (13)$$

1 is an eigenvalue of F with multiplicity of at least 2. Therefore F is a *block stochastic* matrix [12] with associated matrix $\Sigma = I_2$, i.e. F is in the form $\left[\begin{array}{c|c} F_{11} & F_{12} \\ \hline F_{21} & F_{22}\end{array}\right]$, and $F_{ij}1_n = \delta_{ij}1_n$, where δ_{ij} denotes Kronnecker delta function. The following lemma is immediate:

Lemma 3.1:
1) For all t, $F(t)$ has eigenvectors $v_1 = \left[\begin{array}{c} 1_n \\ \hline 0_n\end{array}\right]$ and $v_2 = \left[\begin{array}{c} 0_n \\ \hline 1_n\end{array}\right]$ associated with $\lambda = 1$.
2) For all n and T, $F(t)^n$ and $\prod_{i=0}^T F(t+i)$ are block stochastic matrices with associated matrix $\Sigma = I_2$.

This Lemma indicates that 1 is eigenvalue of F with multiplicity at least 2. The locus of the other $2n-2$ eigenvalues of F can be determined using the following Theorem.

Theorem 3.2: The matrix F has no eigenvalue outside the unit circle. Furthermore, with probability 1, the spectral radius of F is 1 and $\lambda = 1$ is a semisimple eigenvalue[2] with algebraic and geometric multiplicities equal to 2.

Proof: Consider the $n \times n$ matrix $K_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots & \\ 1 & 0 & 0 & \cdots & 1\end{bmatrix}$. Consider a

[1]The dependence of F and other matrices on time are suppressed but will be clear from the context.

[2]An eigenvalue is called semisimple if its algebraic and geometric multiplicities are equal.

transformation matrix $T_1 = \left[\begin{array}{c|c} K_n & 0_{n\times n} \\ \hline 0_{n\times n} & K_n \end{array}\right]$, and T_2, a column swapping transformation $T_2 = [e_1 \ \ e_{n+1} \ \ e_2 \ \ e_{n+2} \ \ e_3 \ \ e_{n+3} \ \ \ldots \ \ e_{2n}]$, where e_i refers to the i^{th} column of the identity matrix. Using the transformation $T = T_1T_2$,

$$TFT^{-1} = \left[\begin{array}{c|c} I_2 & D \\ \hline 0_{(2n-2)\times(2n-2)} & C \end{array}\right],$$

in which

$$C = \left[\begin{array}{c|c} 0_{(n-1)\times(n-1)} & -\frac{S}{Q}I_{n-1} \\ \hline -\frac{S}{N}I_{n-1} & 0_{(n-1)\times(n-1)} \end{array}\right]. \qquad (14)$$

Therefore: $\det(\lambda I_{2n} - F) = [\det(\lambda I_2 - I_2)][\det(\lambda I - C)] = [\det(\lambda I_2 - I_2)][\det(\lambda^2 I_{n-1} - \frac{S^2}{NQ}I_{n-1})]$.

The eigenvalues of matrix F are: $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = \lambda_4 = \cdots = \lambda_{n+2} = \frac{S}{\sqrt{nQ}}$, and $\lambda_{n+3} = \lambda_{n+4} = \cdots = \lambda_{2n} = -\frac{S}{\sqrt{nQ}}$. Considering two vectors $v_1 = [x_1x_2 ... x_n]^T$ and $v_2 = 1_n$ and writing Cauchy's inequality yields: $(\sum_{i=1}^{n} x_i)^2 \leq n\sum_{i=1}^{n} x_i^2$ or equivalently $S^2 \leq nQ$. Therefore for $i = 3, 4, \cdots, 2n$, $|\lambda_i| \leq 1$ and the equality holds only if $x_1 = x_2 = \cdots x_n$. Therefore the spectral radius of F is 1. Moreover, since x_is are all uniformly distributed on $[0, 1]$ with probability 1 they are not all equal and, $\{\lambda_i\}_{i=3}^{n}$ fall inside the unit circle. The algebraic multiplicity of $\lambda = 1$ is 2. Lemma 3.1 implies that the geometric multiplicity of $\lambda = 1$ is also 2. Therefore, $\lambda = 1$ is a semisimple eigenvalue with probability 1. ■

Corollary 3.1: There exist a norm $||.||_*$ such that $||C||_* < 1$ with probability 1.

Proof: Since $\rho(C) < 1$ with probability 1, there exists a matrix norm $||.||_*$ such that $\rho(C) \leq ||C||_* \leq \rho(C) + \epsilon < 1$, for a particular choice of $\epsilon > 0$. ■

We first consider the case of "one-time learning", i.e. the agents send their (x, y) data only once and then iterate to achieve consensus. Therefore, we only have a single matrix F and Equation (12) will be time invariant. The following result follows as a corollary of Theorem 3.2.

Corollary 3.2: For the "one time learning" scenario, agents will reach a consensus on θ and γ with probability 1, i.e. they will coordinate on the same behavior function $f(x) = \theta^* x + \gamma^*$.

Proof: Theorem 3.2 states that the spectral radius of matrix F is 1, and $\lambda = 1$ is the only eigenvalue on the unit circle and it is semisimple. Furthermore, Lemma 3.1 and Theorem 3.2 indicate that the nullspace of $I_{2n} - F$ is a two dimensional space spanned by $v_1 = \left[\begin{array}{c} 1_n \\ \hline 0_n \end{array}\right]$ and $v_2 = \left[\begin{array}{c} 0_n \\ \hline 1_n \end{array}\right]$. Therefore $\lim_{t\to\infty} F^t$ exists and is the projector onto nullspace of $I_{2n} - F$ along the range of $I_{2n} - F$, i.e. $\lim_{t\to\infty} F^t = \left[\begin{array}{c} 1_n \\ \hline 0_n \end{array}\right] u_1^T + \left[\begin{array}{c} 0_n \\ \hline 1_n \end{array}\right] u_2^T$, where $u_1, u_2 \in R^{2n}$ and

$$\lim_{t\to\infty} \Theta(t) = \Big(\left[\begin{array}{c} 1_n \\ \hline 0_n \end{array}\right] u_1^T + \left[\begin{array}{c} 0_n \\ \hline 1_n \end{array}\right] u_2^T \Big)\Theta(0) = \left[\begin{array}{c} \theta^* 1_n \\ \hline \gamma^* 1_n \end{array}\right].$$

■

To generalize this result for the case where samples are sent at each time, we first state a lemma.

Lemma 3.2: Any finite product of matrices of the form $H_m = F(m)F(m-1)\cdots F(1)$, with $F = \left[\begin{array}{c|c} P_1 & M_1 \\ \hline M_2 & P_2 \end{array}\right]$, where P_1, M_1, M_2, P_2 are as in Equation 12 is similar to a matrix of block upper triangular form

$$G_m = \left[\begin{array}{c|c} I_2 & D_m \\ \hline 0_{(2n-2)\times(2n-2)} & C_m \end{array}\right], \qquad (15)$$

in which C_m is a matrix of the form

$$\left[\begin{array}{c|c} 0_{(n-1)\times(n-1)} & C_{12} \\ \hline C_{21} & 0_{(n-1)\times(n-1)} \end{array}\right]$$

if the number of the terms in the product is odd and of the form

$$C_m = \left[\begin{array}{c|c} C_{11} & 0_{(n-1)\times(n-1)} \\ \hline 0_{(n-1)\times(n-1)} & C_{22} \end{array}\right]$$

if the number of the terms in the product is even.

Proof: The proof follows by induction. Consider two matrices $F(1) = \left[\begin{array}{c|c} P_1(1) & M_1(1) \\ \hline M_2(1) & P_2(1) \end{array}\right]$ and $F(2) = \left[\begin{array}{c|c} P_1(2) & M_1(2) \\ \hline M_2(2) & P_2(2) \end{array}\right]$. It is shown at the proof of Theorem 3.2 that using transformation T, we can write $F(2) = T^{-1}G(2)T$ and $F(1) = T^{-1}G(1)T$, where for $i = 1, 2$, $G(i) = \left[\begin{array}{c|c} I_2 & D(i) \\ \hline 0_{(2n-2)\times(2n-2)} & C(i) \end{array}\right]$, for some matrices $C(i)$ and $D(i)$. Direct multiplication results in $F(2)F(1) = T^{-1}G_2T$, where

$$G_2 = G(2)G(1) = \left[\begin{array}{c|c} I_2 & D(1) + D(2)C(1) \\ \hline 0_{(2n-2)\times(2n-2)} & C(2)C(1) \end{array}\right],$$

which is of the form of Equation (15). The induction step from the product of $m-1$ matrices to that of m matrices is identical to above. Therefore it follows that $F(m)F(m-1)\cdots F(1) = T^{-1}G_mT$, where

$$G_m = G(m)\cdots G(2)G(1) = \left[\begin{array}{c|c} I_2 & D \\ \hline 0_{(2n-2)\times(2n-2)} & C(m)\cdots C(2)C(1) \end{array}\right].$$

It follows from Equation (14) that for $m = 2k$,

$$C_m = C(m)\cdots C(2)C(1) = \left[\begin{array}{c|c} \frac{S(2k)\cdots S(2)S(1)}{n^k Q(2k)\cdots Q(4)Q(2)} I_{n-1} & 0_{(n-1)\times(n-1)} \\ \hline 0_{(n-1)\times(n-1)} & \frac{S(2k)\cdots S(2)S(1)}{n^k Q(2k-1)\cdots Q(3)Q(1)} I_{n-1} \end{array}\right], \qquad (16)$$

and for $m = 2k+1$,

$$C_m = C(m)\cdots C(2)C(1) = \left[\begin{array}{c|c} 0_{(n-1)\times(n-1)} & -\frac{S(2k+1)\cdots S(2)S(1)}{n^k Q(2k+1)\cdots Q(3)Q(1)} I_{n-1} \\ \hline -\frac{S(2k+1)\cdots S(2)S(1)}{n^{k+1} Q(2k)\cdots Q(4)Q(2)} I_{n-1} & 0_{(n-1)\times(n-1)} \end{array}\right] \qquad (17)$$

■

Theorem 3.3: For the general learning case, where each agent sends samples at each time instant, agents will reach a consensus on θ and γ with probability 1.

Proof: We show that for the learning dynamics iteration, $\Theta(t+1) = F(t)\Theta(t)$ the corresponding sequence $\{H_m\}_{m\geq 1} = \{F(m)F(m-1)\cdots F(1)\}_{m\geq 1}$ converges to a matrix $H = \left[\frac{1_n}{0_n}\right]u_1^T + \left[\frac{0_n}{1_n}\right]u_2^T$ with probability 1 for some vectors $u_1, u_2 \in R^{2n}$. First, it follows from Lemma 3.1 that $\forall m \geq 1$, $v_1 = \left[\frac{1_n}{0_n}\right]$ and $v_2 = \left[\frac{0_n}{1_n}\right]$ are also eigenvectors of H_m corresponding to eigenvalues $\lambda_{1,2} = 1$.

Next, we show that as $m \to \infty$, the 'third largest eigenvalue modulus' of H_m goes to 0 and the invariant space of H_m is spanned by v_1 and v_2 with probability 1. It was shown in Lemma 3.2 that for any H_m, there is a corresponding similar matrix G_m in block upper triangular form (15). We show that as $m \to \infty$ the sequence $\{C_m\}_{m\geq 1}$ converges to $\mathbf{0}_{(2n-2)\times(2n-2)}$ with probability 1 and therefore the third largest eigenvalue modulus of G_m goes to 0. The rest of proof follows immediately.

Suppose $m = 2k+1$. Equation 17 yields that the eigenvalues of C_m satisfy the following equation:

$$\det(\lambda I_{2n-2} - C_m) =$$
$$\det\big(\lambda^2 I_{n-1} - \frac{S(2k+1)^2\cdots S(3)^2 S(2)^2 S(1)^2}{n^{2k+1}Q(2k)\cdots Q(3)Q(2)Q(1)} I_{n-1}\big) = 0.$$

Therefore, C has two sets of repeated eigenvalues, each with multiplicity $n-1$, which satisfy the following equation:

$$\lambda^2 = \prod_{i=1}^{2k+1} \frac{S(i)^2}{nQ(i)}. \tag{18}$$

As was shown in the proof of Theorem 3.2, the Cauchy-Schwartz inequality implies that for all i, $S(i)^2 \leq nQ(i)$. The equality holds only if for all i and j, $x(i) = x(j)$, which is a zero probability event. Therefore for all i, $0 < \frac{S(i)^2}{nQ(i)} < 1$ with probability 1. Now consider $\epsilon > 0$ and let

$$\mathcal{S}_\epsilon(m) = \{i | 1 \leq i \leq m \text{ and } \frac{S(i)^2}{nQ(i)} < 1-\epsilon\}.$$

Then, $\forall 1 \leq i \leq 2n-2$

$$\lambda_i(m)^2 = \prod_{i=1}^{2k+1} \frac{S(i)^2}{nQ(i)} = \big(\prod_{i\in\mathcal{S}_\epsilon(m)} \frac{S(i)^2}{nQ(i)}\big)\big(\prod_{i\notin\mathcal{S}_\epsilon(m)} \frac{S(i)^2}{nQ(i)}\big). \tag{19}$$

Therefore, C_m has two repeated eigenvalues each of algebraic multiplicity $n-1$:

$$\lambda_{1,2} = \pm\big(\prod_{i\in\mathcal{S}_\epsilon(m)} \frac{S(i)^2}{nQ(i)}\big)^{1/2}\big(\prod_{i\notin\mathcal{S}_\epsilon(m)} \frac{S(i)^2}{nQ(i)}\big)^{1/2},$$

As $k \to \infty$, it can be proven by contradiction that the cardinality of the set $\mathcal{S}_\epsilon$ grows unbounded, i.e. $|\mathcal{S}_\epsilon| \to \infty$. Therefore, as $k \to \infty$, the first term in the right hand side of 19 goes to zero, while the second term remains bounded by 1, i.e. with probability 1,

$$\lim_{m\to\infty} \lambda_i(m)^2 = \lim_{m\to\infty} (1-\epsilon)^{|\mathcal{S}_\epsilon|} = 0,$$

i.e. all the eigenvalues of C_m converge to zero. Let $\alpha_m = \frac{S(2k+1)\cdots S(2)S(1)}{n^k Q(2k+1)\cdots Q(3)Q(1)}$ and $\beta_m = \frac{S(2k+1)\cdots S(2)S(1)}{n^{k+1}Q(2k)\cdots Q(4)Q(2)}$, then we can write

$$C_m = \lambda_1 G_1 + \lambda_2 G_2 =$$
$$(\sqrt{\alpha_m\beta_m}).\frac{1}{2}\left[\begin{array}{c|c} I_{n-1} & -\sqrt{\frac{\alpha_m}{\beta_m}}I_{n-1} \\ \hline -\sqrt{\frac{\beta_m}{\alpha_m}}I_{n-1} & I_{n-1}\end{array}\right] +$$
$$(-\sqrt{\alpha_m\beta_m}).\frac{1}{2}\left[\begin{array}{c|c} I_{n-1} & \sqrt{\frac{\alpha_m}{\beta_m}}I_{n-1} \\ \hline \sqrt{\frac{\beta}{\alpha}}I_{n-1} & I_{n-1}\end{array}\right].$$

It can be shown that $Prob(\frac{\alpha_m}{\beta_m} = 0 \vee \frac{\beta_m}{\alpha_m} = 0) = 0$, therefore given a norm $||.||$ on R^{2n-2},

$$\lim_{k\to\infty} ||C_{2k+1}|| = 0.$$

Also for $m = 2k$:

$$\lim_{k\to\infty} ||C_{2k}|| \leq \lim_{k\to\infty} ||C_{2k-1}||.||C|| = 0.$$

Therefore, with probability 1,

$$\lim_{m\to\infty} C_m = \mathbf{0}_{2n-2}.$$

Given $0 < \kappa < 1$, since $\lim_{m\to\infty} C_m = 0$, there exists $M < \infty$ such that

$$Prob(||C(M)C(M-1)\cdots C(0)|| \leq \kappa) = 1.$$

Now, consider

$$H_m = F(m)...F(2)F(1) =$$
$$\left[\begin{array}{c|c} I_2 & D(1)+\cdots+D(m)C(m-1)\cdots C(1) \\ \hline \mathbf{0}_{2n-2} & C(m)C(m-1)\cdots C(1)\end{array}\right].$$

We have shown that

$$\lim_{m\to\infty} C_m = \lim_{m\to\infty} C(m)C(m-1)\cdots C(1) = \mathbf{0}_{2n-2}.$$

Since $\rho(F(i)) = 1$, there is a constant $K < \infty$, such that $||D(i)|| \leq K$. Therefore with probability 1,

$$||D(M+1)C(M)\cdots C(1) + D(M+2)C(M+1)\cdots|| \leq$$
$$MK\kappa + MK\kappa^2 + \cdots = \frac{MK\kappa}{1-\kappa}.$$

Therefore with probability 1, as $m \to \infty$, $H_m = F(m)F(m-1)\cdots F(1)$ converges to a matrix H. Furthermore for any matrix F of the form (12), we have

$$FH = F\lim_{m\to\infty} H_m = lim_{m\to\infty}H_m = H,$$

i.e. the columns of F^∞ are in the nullspace of $I_{2n} - F$. But it was shown that the nullspace of F is spanned by $v_1 = \left[\frac{1_n}{0_n}\right]$ and $v_2 = \left[\frac{0_n}{1_n}\right]$ and therefore we can write $H = \left[\frac{1_n}{0_n}\right]u_1^T + \left[\frac{0_n}{1_n}\right]u_2^T$ for $u_1, u_2 \in R^{2n}$.

■

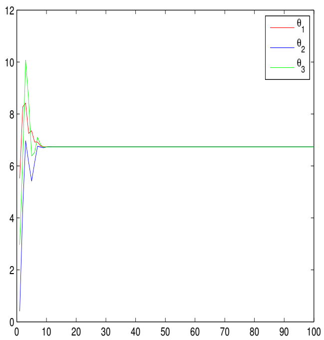

Fig. 1. Evolution of θ for a sample situation of two initially optimist and one initially pessimist agents that reach a final consensus to collaborate.

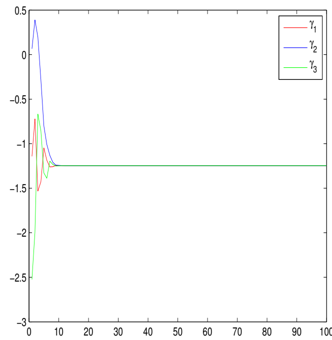

Fig. 2. Evolution of γ for a sample situation of two initially optimist and one initially pessimist agents that reach a final consensus to collaborate.

IV. Simulations

We now consider the game model of Section II and show that by learning each others' behaviors, the agents reach a higher degree of coordination resulting in higher pay-offs. The numerical values for the coordination game pay-off of Section II-A are set to $a = 5, b = 4$, and $c = 2$. We consider a complete graph of 3 agents with linear behavior functions from the set $\mathcal{F}_l$. We consider an agent to be optimist if its behavior function $f(x) = \theta x + \gamma$ intersects the $y = 1$ line for $0 \le x \le 1/2$ and pessimist if the intersection occurs at $x > 1/2$. We ran different sets of simulations with initial behaviors chosen randomly with $\theta \in [0, 8]$ and $\gamma \in [-3, 1]$ with the restriction that two pessimist and only one optimist agents initially exist. Figures 1 and 2 show an instance that starting with two pessimist agents (2 and 3) and one optimist agent (1) leads to a configuration of all optimist agents and thereby group collaboration occurs. We ran 100 sets of simulations each consisting of 1000 runs with two randomly chosen pessimist agents vs. one randomly chosen optimist agent. It was observed that in 326.09 ± 14.66 times out of the 1000 times the agents reached a consensus to cooperate with average individual pay-offs of 5956.5 ± 87.97 (out of a maximum of 10000). We then ran another set of simulations with completely random initial behaviors with initial behaviors chosen randomly with $\theta \in [0, 8]$ and $\gamma \in [-3, 1]$. This time, we ran 100 sets of simulations each consisting of 1000 runs. It was observed that in 763.50 ± 12.97 times out of the 1000 times collaboration emerged with average individual pay-off of 8581 ± 77.86 (out of a maximum of 10000).

V. Conclusions and future work

In this paper we studied the implications of a group's communication structure on the emergence of coordination among them. The problem is to make a decision on whether to cooperate or not in a group effort. This emerges as a result of a series of two-person games between agents and their neighbors. We considered agents with different behavioral variables (types) and proposed an adaptive scheme by which the agents' behaviors evolve so that they can coordinate on a common strategy. We addressed the behavior adaptation in the Cucker-Smale-Zhou framework of language evolution.

We showed that in the special case of 'linear behaviors', our scheme would result in an extension of consensus problems, in which the evolution of the system is governed by *block-stochastic matrices*. The derived linear systems may not be stable if the underlying topologies are not complete. However, this can be alleviated by scaling the non-diagonal blocks. Currently, we are working on finding extensions of the linear learning case to arbitrary topologies. Also, in the current work, there is no relationship between the pay-off values a, b, c and the initial tendency of the agents towards pessimism or optimism. We are working to 'close this loop'. Future directions include determining how the number of like-minded agents and their connectivity play a major decisive role on the attained equilibrium.

References

[1] J. S. Baras and P. Hovareshti, "Efficient Communication Infrastructures for Distributed Control and Decision Making in Networked Stochastic Systems", *Proceedings of The 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010)*, Budapest, Hungary, July 5-9, 2010.

[2] F. Cucker, S. Smale, and D. X. Zhou, "Modeling language evolution", *in Foundations of Computational Math.*, vol. 4, 2004, pp. 315–343.

[3] T. Schelling, *The Strategy of Conflict.* Harvard University Press, Cambridge, MA, 1960.

[4] D. Lewis, *Conventions: A Philosophical Study.* Harvard University Press, Cambridge, MA, 1969.

[5] R. Axelrod, *Evolution of Cooperation.* Basic Books New York, NY, 1984.

[6] D. Fudenberg and D. K. Levine, *The theory of learning in games* MIT Press, Cambridge, MA, 1998.

[7] S. Goyal and F. Vega-Redondo, "Network formation and social coordination", *Games and Economic Behavior*, vol.: 50, 2005, pp 178-207.

[8] S. K. Berninghaus and U. Schwalbe, "Conventions, local interaction, and automata networks", *Journal of Evolutionary Economics*, vol. 6, 1996, pp. 297-312.

[9] G. C. Chasparis and J. S. Shamma, "Efficient Network Formation by Distributed Reinforcement", *in Proceedings of 47th IEEE Conference on Decision and Control*, Cancun, Mexico, 2008, pp 1690-1695.

[10] M. Jackson and A. Wolinsky, "A Strategic Model of Social and Economic Networks", *Jour. of Econ. Theory*, vol. 71, 1996, pp. 44-74.

[11] M. Chwe, "Communication and Coordination in Social Networks", *Review of Economic Studies*, Vol. 67, 2000 pp. 1-16.

[12] W. Kuich and K.Walk, "Block-stochastic matrices and associated finite-state languages", *Computing* vol. 1, 1966, pp. 50–61.