Dynamic games with side information in economic networks

Ceyhun Eksin, Pooya Molavi, Alejandro Ribeiro and Ali Jadbabaie

Department of Electrical and Systems Engineering, University of Pennsylvania

{ceksin, pooya, aribeiro, jadbabai}@seas.upenn.edu

***Abstract*— We consider a dynamic game with payoff externalities. Agents' utility depends on an unknown true state of the world and actions of everyone in the network. Each agent has an initial private information about the underlying state and repeatedly observes actions of its neighbors. We analyze the asymptotic behavior of agents' actions and beliefs in a connected network when it is common knowledge that the agents are myopic and rational. Given a quadratic payoff function, we provide a new proof for an existing result that claims almost sure consensus in actions asymptotically. Given consensus in actions, we show that agents have the same mean estimate of the true state of the world in the limit. We justify these results in a numerical example motivated by a socio-economic scenario.**

I. Introduction

The model discussed in this paper belongs to the class of repeated games of incomplete information. Agents repeatedly play a game where the payoffs depend on a parameter ("state of the world") as well as the actions taken by other agents. In a game of incomplete information the payoff-relevant parameter is unknown to the agents; rather, they make private noisy observations about the parameter that can be used when deciding on an action. In such a setting, on the one hand, agents' optimal actions depend on their private observations as well as how they expect others to play, and on the other hand, by selecting certain actions agents are revealing—perhaps unwillingly—pieces of private information about the unknown parameter [1]. As a result, the actions chosen by "rational" agents are influenced by both an *information externality* pertaining to the flow of information about the unknown parameter and a *payoff externality* corresponding to the dependence of the payoff on actions taken by other agents. Rational agents need to select actions that are optimal given their information while also taking into account the effect of their decisions on the future play. They also need to try the learn about the underlying parameter as much as possible by incorporating the new information revealed to them optimally, i.e., using the Bayes rule. This kind of strategic learning is relevant to the vast literature on learning in games; refer to [2] and references therein.

There exists an extensive literature on learning over networks. Bayesian learning stands as the normative behavioral model for agents in social networks; however, it is often computationally intractable even for networks with small number of agents. This is since a Bayesian update requires an agent to infer not only about the information of her neighbors but also that of the neighbors of her neighbors and so on. Because of such computational intractability, Bayesian learning models often focus on asymptotic characterizations of agents' behavior [3]–[5]. Only under some structural assumptions on the network or distribution of information, Bayesian updating can be performed tractably [6]–[8]. The mathematical intractability of Bayesian learning in the general case motivates the introduction of bounded rational approaches that resort to simplified Bayesian updates and heuristic rules to keep computations tractable. These non-Bayesian models are predicated on the claim that they replicate some observed real world behavior [9]. From an analytical perspective, the use of tractable rules permits transient convergence analysis and the determination of the effects of network structure on agents' behaviors and convergence rates [9]–[12]. The algorithms in distributed estimation problems also assume relevance to these bounded rational models, in which agents strive to recover the estimate based on global information through refining their estimates using local information [13]–[16].

In this paper, we study the asymptotic behavior of agents' actions and their beliefs given a fixed network that determines the flow of information – see Section II. We consider payoffs that are quadratic in self action, the state of the world and actions of other agents. Agents incorporate new information according to Bayes' rule and are myopically optimal at each stage of the game – see Section II-B. Given this setup, we review a recent result in [17] that shows agents' strategies converge in the limit (Lemma 1). In parallel to [17], we use this result to show that agents' limit actions are in agreement with each other (Theorem 1). There are two contributions of this paper. First, we follow a different approach in proving the consensus in actions result (see Lemma 2). While the proof in [17] is applicable to more general class of games, the proof in this paper is suitable for quadratic payoff functions. Second, we show that agents mean estimates of the state have to be equal in the limit by using the consensus in limit actions (Corollary 1). Finally, we provide an example from a socio-economic context where agents have trade-off between estimating the true value of a company's stock and estimating value estimates of other agents – see Section IV.

Research supported in parts by ARO P-57920-NS, NSF CAREER CCF-0952867, NSF CCF-1017454, ONR MURI "Next Generation Network Science", and AFOSR MURI FA9550-10-1-0567.

II. Quadratic games

Consider a repeated game with N players. Time is discrete and agents act synchronously. Hence, at each stage $t \in \mathbb{N}$, agent i takes an action $a_{i,t}$ that belongs to a compact measurable action space A_i and receives a payoff u_i based on actions of other agents $\mathbf{a}_{-i,t} = \{a_{j,t}\}_{j \neq i} \in \bigtimes_{j \neq i} A_j$ and the underlying state of the world $\theta \in \Theta \subseteq \mathbb{R}$,

$$u_i(a_{i,t}, \mathbf{a}_{-i,t}, \theta) = \alpha a_{i,t} - \frac{1}{2} \sum_{j=1,\ldots,N} a_{j,t}^2 + \beta \sum_{j \neq i} a_{i,t} a_{j,t} + \delta a_{i,t} \theta + c\theta^2. \quad (1)$$

α, β, c and δ are constants. Note that the payoff is quadratic in actions of other agents and the underlying state of the world. It is possible to include other additive terms that depend on θ in (1).

The underlying state of the world θ is unknown to agents. At time $t = 0$, agent i receives private signal s_i coming from some measurable signal set S_i. The space of private signals is $S := S_1 \times \ldots \times S_N$. Initially agents have common prior $\mathbf{P}$ over the state of the world and private signals, that is, $\mathbf{P} = \Theta \times S$.

Agents can only receive information from a certain subset of the population. We use V to denote the whole population. The subset of the population that agent i can observe is denoted by $V_i \subseteq V$. We use a graph $G = (V, E)$ with node set V and edge set E to summarize the observations within the population. Using the graph G, we formally define the observation set of i as $V_i = \{j : (j,i) \in E\}$. We assume the graph is undirected (agent i can observe j if and only if j can observe i), static (fixed over time) and connected (there is a path from $i \in V$ to $k \in V$, $p_1, \ldots, p_l$ such that $p_1 = i$, $p_n = k$ and $p_{m-1} \in V_m$ for $m = 2, \ldots, l$).

At the beginning of each stage of the game, agent i observes actions of agents in V_i from the previous stage. We use $h_{i,t}$ to denote the part of the history observed by agent i at time t. Agent i's history can be recursively written as

$$h_{i,t} = \{h_{i,t-1}, \{a_{j,t-1}\}_{j \in V_i}\} \quad (2)$$

where $h_{i,0} = \{x_i\}$. The set of all possible histories observed by agent i at time t is denoted by $H_{i,t}$; i.e., $h_{i,t} \in H_{i,t}$. The history of the game at time t is $h_t = \bigcup_{i \in V} h_{i,t}$. The game history contains underlying state of the world, private signals of agents, and the set of all actions up to time t. We denote the feasible set of plays that game history belongs to as H_t. We assume that the action space is equal for all agents, that is, $A_i = A$ for all $i \in V$. Hence, the game history at time t can be defined as $H_t := \Theta \times S \times A^{N(t-1)}$. The space of play that covers the entire horizon is denoted by $\Omega := H_\infty := \bigcup_{\tau=0}^{\infty} H_\tau$. $(\Omega, \mathcal{F})$ is the measurable space with Borel σ-algebra $\mathcal{F}$. We define the σ-field over Ω of agent i at time t as $\mathcal{H}_{i,t}$. $\mathcal{H}_{i,t}$ represents the information of agent i at time t.

So far, we have described the individual utility functions and information provided to agents. Next, we characterize agent behavior in the repeated game.

A. Strategy

Agent i's behavior is determined by a pure strategy $\sigma_i = (\sigma_{i,0}, \sigma_{i,1}, \ldots)$ where $\sigma_{i,t} : \mathcal{H}_{i,t} \mapsto A$ from $t = 0$ to $t = \infty$. Hence, the strategy of agent i determines its course of action given any possible information observed by agent i. The strategy profile $\sigma := (\sigma_1, \ldots, \sigma_N)$ generates a path of play and determines the information observed by each agent. At time $t = 0$, actions played by agents $\mathbf{a}_t \in A^N$ are

$$\mathbf{a}_0(\sigma) = (\sigma_{1,0}(x_1), \ldots, \sigma_{N,0}(x_N)). \quad (3)$$

The actions at time $t = 0$ determines the history observed by each agent as given by the recursion in (2). Hence, the actions of the subsequent step $t = 1$ are

$$\mathbf{a}_1(\sigma) = (\sigma_{1,0}(h_{1,1}), \ldots, \sigma_{N,0}(h_{N,1})) \quad (4)$$

and this process is repeated for $t > 1$ generating a realization of the game $\omega \in \Omega$. Another way to look at this process is that there is a realization of the game $\omega \in \Omega$ and the history observed by agent i is a deterministic function of ω, that is, $h_{i,t}(\omega)$.

Together with the common prior $\mathbf{P}$, the strategy profile σ induces a probability distribution over the space of plays Ω. Let P_σ be the probability distribution induced by σ. We define the corresponding expectation operator $\mathbf{E}_\sigma$ with respect to the distribution P_σ. Using P_σ measure on Ω, we define the $L^2(\Omega)$-norm for the $\mathcal{H}_{i,\infty}$ measurable functions $f_i : \Omega \mapsto A$ as

$$\|f_i\|_{\sigma,2} = \left(\int_\Omega f_i^2 d\mathbf{P}_\sigma \right)^{\frac{1}{2}}. \quad (5)$$

We define the space of functions $\mathcal{A}_i$ that consists of all f_i such that $\|f_i\|_{\sigma,2} < \infty$. Note that $\sigma_{i,t}(h_{i,t}(\omega))$ is $\mathcal{H}_{i,t}$-measurable and hence belongs to $\mathcal{A}_i$ for all $t \in \mathbb{N}$. We say that $f_i \in \mathcal{A}_i$ and $g_i \in \mathcal{A}_i$ are P_σ-a.e. equivalent when $\|f_i - g_i\|_{\sigma,2} = 0$. Similarly, we define the space of functions $\mathcal{A} := \bigtimes_{i \in V} \mathcal{A}_i$ using the following norm

$$\|f\|_{\sigma,2} = \max_{i \in V} \left(\int_\Omega f_i^2 d\mathbf{P}_\sigma \right)^{\frac{1}{2}}. \quad (6)$$

B. Myopic rationality

Given a strategy profile σ, agents form beliefs over the space of play Ω using Bayes' rule. Hence at stage $t \in \mathbb{N}$, the belief of agent is the conditional probability over Ω given information at i, that is, $\mathbf{P}_\sigma(\cdot \,|\, \mathcal{H}_{i,t})$.

Given the strategy profile σ and utility function in (1), agent i responds to the strategies of other agents at time t, that is, let $\mathbf{s}_{-i,t} \in \mathcal{A}_{-i} := \bigtimes_{j \in V \setminus i} \mathcal{A}_j$ be the strategies of other agents at time t then i's response is a function $\Phi_{i,t} : \mathcal{A}_{-i} \mapsto \mathcal{A}_i$. Furthermore, best response function for agent i is defined as

$$\Phi_{i,t}(\mathbf{s}_{-i,t}) = \operatorname*{argmax}_{a_i \in A} \mathbf{E}_\sigma \left[u_i(a_i, \mathbf{s}_{-i,t}, \theta) \,\middle|\, \mathcal{H}_{i,t} \right] \quad (7)$$

Note that this definition assumes that up until time t, agents played according to the sequence of strategies $\sigma_{0,1,\ldots,t-1}$ mapping respective histories $\{h_{j,\tau}\}_{j \in V}$ to actions $\{a_{j,\tau}\}_{j \in V}$

for $\tau < t$. Hence, agent i's best response at time t to strategies of other agents is based on the belief formed by prior strategies.

Since $u_i(\cdot)$ is continuous in its arguments and A is compact, the best response function is well defined. Furthermore, the maximizer in (7) is unique because the utility function is strictly concave—observe that $\partial^2 u_i(\cdot)/\partial a_i^2 < 0$. As a result, we obtain the best response function for the utility function in (1) by taking the derivative of $\mathbf{E}_\sigma\left[u_i(a_i, \mathbf{s}_{-i,t}, \theta) \,\middle|\, \mathcal{H}_{i,t}\right]$ with respect to a_i and solving for a_i when the resultant derivative is equated to zero,

$$\Phi_{i,t}(\mathbf{s}_{-i,t}) = \alpha + \beta \sum_{j \in V \setminus \{i\}} \mathbf{E}_\sigma[s_{j,t}|\mathcal{H}_{i,t}] + \delta\, \mathbf{E}_\sigma[\theta|\mathcal{H}_{i,t}]. \tag{8}$$

The best response function in (8) is a linear function of i's expectation of the underlying state of the world and of the strategies of other agents at time t. Note that in (8) agents are playing myopically optimal.

We define our notion of Bayesian-Nash equilibrium as the fixed point of (8) for any realization of ω and $t \in \mathbb{N}$, that is, for all $i \in V$ and $t \in \mathbb{N}$,

$$\sigma^*_{i,t}(h_{i,t}(\omega)) = \Phi_{i,t}(\sigma^*_{-i,t}) \quad \text{for all } h_{i,t} \in H_{i,t}. \tag{9}$$

A simple interpretation of (9) based on the definition of best response function in (8) is that agent i plays best response to best response strategies of all the other agents. Hence agent i not only has to form its own belief on the state of the world and beliefs of other agents but also has to consider strategies of other agents while taking actions at each step. We can equivalently write (9) in vector form using $\Phi_t : \mathcal{A} \mapsto \mathcal{A}$ which is the stacking of best responses in (8) as follows

$$\sigma^*_t = \Phi_t(\sigma^*_t) \quad \text{for all } t \in \mathbb{N}. \tag{10}$$

Myopic rationality presented in this section can be thought of as the subgame perfect Bayesian-Nash equilibrium while the agents' future discounting goes to zero.

III. CONVERGENCE RESULTS

The setup considered in Section II introduces payoff externalities to the learning process. Here, we provide an asymptotic analysis of the repeated Bayesian game. Agents' beliefs over the underlying parameter converges since it is a martingale [3]. From the learning point of view, we are interested in answering whether agents' belief over θ agree with each other in the limit or not. In this paper, we rely on existing results in [17] that shows convergence of actions in the limit. As is shown in [17], this implies that agents can identify limit strategies of agents in their observation set. We use this result along with properties of the utility function to prove that agents' reach consensus in actions (Theorem 1). While this result exists in [17], here we provide a new proof for it. The contribution of this paper is that we use these existing results to conclude that agents' mean estimates of θ are equal in the limit.

We need the following assumption on the constant β in order for our results to hold.

Assumption 1: There exists a $\kappa < 1$ such that

$$0 \leq \beta \leq \frac{\kappa}{N-1}. \tag{11}$$

Lemma 1: If σ^* is a BNE and Assumption 1 holds, then

$$\lim_{t\to\infty} \sigma^*_t = \sigma^*_\infty \quad P_\sigma \text{ a.s.} \tag{12}$$

Proof: Use triange inequality for the distance between σ^*_t and σ^*_∞,

$$\|\sigma^*_t - \sigma^*_\infty\|_{\sigma,2} \leq \|\sigma^*_t - \Phi_t(\sigma^*_\infty)\|_{\sigma,2} + \|\Phi_t(\sigma^*_\infty) - \sigma^*_\infty\|_{\sigma,2}. \tag{13}$$

Consider the first term on the right hand side of the inequality (13). Using the definition of σ^*_t in (10) and best response function in (8), the difference can be equivalently written as

$$\|\sigma^*_t - \Phi_t(\sigma^*_\infty)\|_{\sigma,2} = \max_{i\in V} \|\beta \sum_{j\in V\setminus\{i\}} \mathbf{E}_\sigma[\sigma^*_{j,t} - \sigma^*_{j,\infty} \,|\, \mathcal{H}_{i,t}]\|_{\sigma,2} \tag{14}$$

We can obtain an upper bound on (14) by moving the norm inside the sum using the triangle inequality

$$\|\sigma^*_t - \Phi_t(\sigma^*_\infty)\|_{\sigma,2} \leq \max_{i\in V} \beta \sum_{j\in V\setminus\{i\}} \|\mathbf{E}_\sigma[\sigma^*_{j,t} - \sigma^*_{j,\infty} \,|\, \mathcal{H}_{i,t}]\|_{\sigma,2} \tag{15}$$

We would only be making the right hand side of (15) larger by removing the conditional expectations inside the norms since conditional expectation is a projection in $L_2(\Omega)$. Further, we can upper bound each term in the sum by the maximum distance among all $j \in V\setminus\{i\}$ to get the following

$$\|\sigma^*_t - \Phi_t(\sigma^*_\infty)\|_{\sigma,2} \leq \max_{i\in V} \beta(N-1) \max_{j\in V\setminus\{i\}} \|\sigma^*_{j,t} - \sigma^*_{j,\infty}\|_{\sigma,2}$$

By (11), we have

$$\begin{aligned} \|\sigma^*_t - \Phi_t(\sigma^*_\infty)\|_{\sigma,2} &\leq \kappa \max_{j\in V} \|\sigma^*_{j,t} - \sigma^*_{j,\infty}\|_{\sigma,2} \\ &= \kappa \|\sigma^*_t - \sigma^*_\infty\|_{\sigma,2} \end{aligned} \tag{16}$$

where we included the ith element inside the maximum operator to obtain the first equality. Note that the second equality follows by the norm definition in (6).

Consider the second term in (13), note that by Levy's $0-1$ law $\Phi_t(\sigma^*_\infty) \to \Phi_\infty(\sigma^*_\infty)$ and by the definition in (10), $\Phi_\infty(\sigma^*_\infty) = \sigma^*_\infty$. As a result, for arbitrary $\epsilon > 0$, there exists τ such that for all $t > \tau$,

$$\|\Phi_t(\sigma^*_\infty) - \sigma^*_\infty\|_{\sigma,2} \leq \epsilon(1-\kappa). \tag{17}$$

Substituting (16) and (17) for the corresponding terms in (13), we see that $\|\sigma^*_t - \sigma^*_\infty\|_{\sigma,2}$ can be made arbitrarily small and hence (12) follows. ■

Given that there is a limit strategy that is played infinitely often by the players, agent i can identify the limit strategies of agents in V_i; i.e. $\sigma_{j,\infty} \in H_{i,\infty}$ for $j \in V_i$. To see this note by the observation model in (2), we have $\sigma^*_{j,t-1}(h_{j,t-1}(\omega)) \in h_{i,t}(\omega)$. Taking limits on both sides and using (12), we see that it is true– see [17] for a formal proof.

Next, we use this identifiability result to show consensus in limit actions. We denote the limit action of agent i as

$a_{i,\infty} := \sigma^*_{i,\infty}(h_{i,\infty}(\omega)) \in A$ for all $i \in V$ that is a resultant of its limit strategy. Before we prove this result, we present an intermediate result that will be used in proving action consensus.

Lemma 2: If σ^* is a BNE and Assumption 1 holds then for $j \in V_i$,

$$\mathbf{E}_\sigma[u_i(\mathbf{a}^*_\infty, \theta)] + \mathbf{E}_\sigma[u_j(\mathbf{a}^*_\infty, \theta)] \leq \mathbf{E}_\sigma[u_i(a^*_{j,\infty}, \mathbf{a}^*_{-i,\infty}, \theta)] + \mathbf{E}_\sigma[u_j(a^*_{i,\infty}, \mathbf{a}^*_{-j,\infty}, \theta)] \tag{18}$$

Proof: Let $\omega \in \bar{\Omega} \subseteq \Omega$ be the set in which $a^*_{i,\infty} \geq a^*_{j,\infty}$. Consider the following difference between conditional expected utilities when j deviates to i's limit strategy $\mathbf{E}_\sigma[u_i(a^*_{i,\infty}, a^*_{i,\infty}, a^*_{-i\backslash j,\infty}) \,|\, \mathcal{H}_{i,\infty}] - \mathbf{E}_\sigma[u_i(a^*_{j,\infty}, a^*_{i,\infty}, a^*_{-i\backslash j,\infty}) \,|\, \mathcal{H}_{i,\infty}]$[1]. Taking expectation of this difference, we lose the conditional expectations and obtain the following equality

$$\begin{aligned}\mathbf{E}_\sigma[u_i(a^*_{i,\infty}, a^*_{i,\infty}, a^*_{-i\backslash j,\infty})] &- \mathbf{E}_\sigma[u_i(a^*_{j,\infty}, a^*_{i,\infty}, a^*_{-i\backslash j,\infty})] \\ &= \alpha(a^*_{i,\infty} - a^*_{j,\infty}) - \frac{1}{2}(a^{*2}_{i,\infty} - a^{*2}_{j,\infty}) \\ &+ \beta(a^{*2}_{i,\infty} - a^*_{i,\infty}a^*_{j,\infty}) + \delta(a^*_{i,\infty} - a^*_{j,\infty})\mathbf{E}_\sigma[\theta]\end{aligned} \tag{19}$$

Now, consider the same difference when agent j does not deviate, that is,

$$\mathbf{E}_\sigma[u_i(a^*_\infty)] - \mathbf{E}_\sigma[u_i(a^*_{j,\infty}, a^*_{-i,\infty})]. \tag{20}$$

When we expand the terms in (20) like in (19), the only difference from right hand side of (19) is the third term which is replaced by $\beta(a^*_{i,\infty}a^*_{j,\infty} - a^{*2}_{j,\infty})$. Since $\beta \geq 0$ and $a^*_{i,\infty} \geq a^*_{j,\infty}$, $\beta(a^*_{i,\infty}a^*_{j,\infty} - a^{*2}_{j,\infty})$ is smaller than $\beta(a^{*2}_{i,\infty} - a^*_{i,\infty}a^*_{j,\infty})$. As a result, we have

$$\begin{aligned}\mathbf{E}_\sigma[u_i(a^*_\infty)] &- \mathbf{E}_\sigma[u_i(a^*_{j,\infty}, a^*_{-i,\infty})] \leq \\ \mathbf{E}_\sigma[u_i(a^*_{i,\infty}, a^*_{i,\infty}, a^*_{-i\backslash j,\infty})] &- \mathbf{E}_\sigma[u_i(a^*_{j,\infty}, a^*_{i,\infty}, a^*_{-i\backslash j,\infty})]\end{aligned} \tag{21}$$

For the utility function in (1), we have $u_i(a^*_{i,\infty}, a^*_{i,\infty}, a^*_{-i\backslash j,\infty}) = u_j(a^*_{i,\infty}, a^*_{-j,\infty})$ and $u_i(a^*_{j,\infty}, a^*_{i,\infty}, a^*_{-i\backslash j,\infty}) = u_j(a^*_\infty)$. Substituting these equalities for the corresponding terms in (21) and rearranging the terms, we get (18). Now, we obtain the same inequality when $a^*_{j,\infty} > a^*_{i,\infty}$. Hence it is true for all $\omega \in \Omega$. ■

Next, we show that there is consensus in actions.

Theorem 1: Let σ^* be a BNE strategy.If Assumption 1 holds then

$$a^*_{i,\infty} = a^*_{j,\infty} \quad \text{for all} \quad i, j \in V \qquad \mathbf{P}_\sigma\text{-a.s.} \tag{22}$$

Proof: By myopic optimality of agents ,

$$\mathbf{E}_\sigma[u_i(a^*_{j,\infty}, \mathbf{a}^*_{-i,\infty}, \theta)|\mathcal{H}_{i,\infty}] \leq \mathbf{E}_\sigma[u_i(\mathbf{a}^*_\infty, \theta)|\mathcal{H}_{i,\infty}] \tag{23}$$

for all $i \in V$. When we take expectation of both sides of (23), we are left with expectations, $\mathbf{E}_\sigma[u_i(a^*_{j,\infty}, \mathbf{a}^*_{-i,\infty}, \theta)] \leq \mathbf{E}_\sigma[u_i(\mathbf{a}^*_\infty, \theta)]$ for all $i \in V$. As a result, the first term on the left hand side of (18) is at least as large as the first term on its right hand side. The same relation holds for the second terms on both hand sides of (18). Hence, we have

$$\mathbf{E}_\sigma[u_i(a^*_{j,\infty}, \mathbf{a}^*_{-i,\infty}, \theta)] = \mathbf{E}_\sigma[u_i(\mathbf{a}^*_\infty, \theta)] \tag{24}$$

By (23), we have that $\mathbf{E}_\sigma[u_i(\sigma_\infty, \theta)|\mathcal{H}_{i,\infty}] - \mathbf{E}_\sigma[u_i(\sigma^i_\infty, \theta)|\mathcal{H}_{i,\infty}] \geq 0$ with $\mathbf{P}_\sigma$-probability one. This proves that the relation in (24) also holds for conditional expectations,

$$\mathbf{E}_\sigma[u_i(a^*_{j,\infty}, \mathbf{a}^*_{-i,\infty}, \theta) \,|\, \mathcal{H}_{i,\infty}] = \mathbf{E}_\sigma[u_i(\mathbf{a}^*_\infty, \theta) \,|\, \mathcal{H}_{i,\infty}] \tag{25}$$

By the definition of myopic rationality $a^*_{j,\infty} = \Phi_{i,\infty}(\sigma^*_{-i,\infty})$ for all $j \in V_i$. However, $\Phi_{i,\infty}(\sigma^*_{-i,\infty})$ is a singleton. Hence, it must be that $a^*_{i,\infty} = a^*_{j,\infty}$ for all $i \in V$ and $j \in V_i$. Connectivity of the network proves the result. ■

Next, we use the consensus in actions result to argue that agents mean estimate of θ is equal.

Corollary 1: If σ^* is a BNE and Assumption 1 holds then $\mathbf{E}_\sigma[\theta \,|\, \mathcal{H}_{i,\infty}] = \mathbf{E}_\sigma[\theta \,|\, \mathcal{H}_{j,\infty}]$ for all $i \in V$ and $j \in V$.

Proof: By definition of BNE, we have

$$a^*_{i,\infty} = \alpha + \beta \sum_{j \in V\backslash\{i\}} \mathbf{E}_\sigma[a^*_{j,\infty}|\mathcal{H}_{i,\infty}] + \delta\, \mathbf{E}_\sigma[\theta|\mathcal{H}_{i,\infty}] \tag{26}$$

From Theorem 1, we have consensus in actions, $a^*_{i,\infty} = a^*$ for all $i \in V$. Substituting in a^* for all the action terms in (8) for $t = \infty$ and solving for a^*, we have $a^* = f(\mathbf{E}_\sigma[\theta|\mathcal{H}_{i,\infty}])$ for some function $f(\cdot)$ for all $i \in V$ and hence the result follows. ■

Corollary 1 shows that agents have the same mean estimates of the underlying state of the world in the limit.

IV. Coordination Game

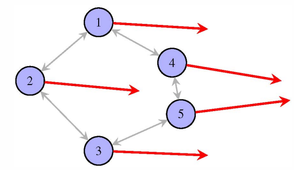

Fig. 1. Coordination game of mobile agents

An investment decision in the stocks of a company includes a player to consider both its valuation of the asset and also how everyone values the asset. In this setting, we define $\omega \in \mathbb{R}$ to be the true value of the stock and action $a_{i,t}$ represents player i's valuation of the asset; i.e., it is the price that agent i is willing to pay per stock share at time t. The payoff function for agent i is given by

$$u_i(\theta, a_i, a_{-i}) = -\frac{1-\lambda}{2}(a_i - \theta)^2 - \frac{\lambda}{2(N-1)} \sum_{j \in V\backslash\{i\}} (a_i - a_j)^2, \tag{27}$$

where $\lambda \in (0,1)$. The first term of the payoff function measures the desire of the player to estimate the true value of the stock. The desire to coordinate with other players is

[1] In this proof, we drop θ from the argument of u_i for convenience.

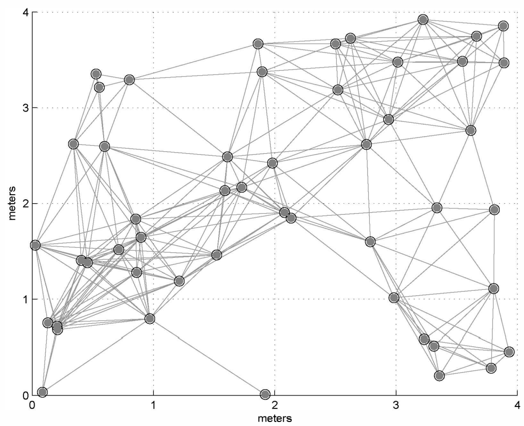

Fig. 2. Geometric network with $N = 50$ agents. Agents are randomly place on a 4 meter $\times$ 4 meter square. There exists an edge between any pair of agents with distance less than 1 meter apart.

represented by the second term. The constant λ gauges the relative importance of coordination and estimation. Note that this utility function satisfies Assumption 1.

The same payoff function can also be motivated by looking at coordination among a network of mobile agents starting with a certain formation trying to move toward a finish line on a straight path [17] – see Fig. 1. Each agent collects an initial noisy measurement of the true heading angle θ, that is, the angle that achieves the shortest path toward the finish line. In this example the actions of agents represent their choice of heading direction or movement angle. The first and second terms in (27) again correspond to the estimation and formation coordination payoffs, respectively.

A. Numerical example

We exemplify the game where the true value of the stock is $\theta = 10\$$. We let $\lambda = 0.5$. Agents initially receive private signals that are corrupted with zero mean Gaussian noise,

$$x_i = \theta + \epsilon_i, \tag{28}$$

where $\epsilon_i \sim \mathcal{N}(0, 1)$. ϵ_i is independent across agents. There exists a local recursion for individual belief updates and myopic rational behavior when the initial private signals are jointly Gaussian [8].

Based on this result, we evaluate convergence behavior in a geometric network with $N = 50$ agents (Fig. 2). The geometric network has a diameter of $d_g = 5$. The action values of each agent is depicted in Fig. 3.

The results show that agents' actions $\mathbf{a}_i$ converge to the best estimates of true state of the world, that is, $\mathbf{E}_\sigma[\theta \mid \mathcal{H}_\infty]$. This means that $\mathbf{E}_\sigma[\theta \mid \mathcal{H}_\infty] = \mathbf{E}_\sigma[\theta \mid \mathcal{H}_{i\infty}]$ for all $i \in V$. Furthermore, convergence rate is fast in the order of the diameter of the network.

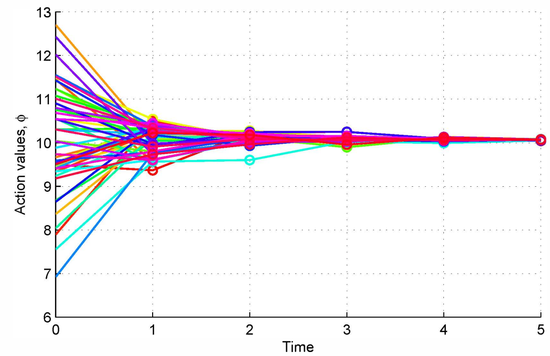

Fig. 3. Values of agents' actions over time for the geometric network.

REFERENCES

[1] S. Morris and H.S. Shin. The social value of public information. *American Economic Review*, 92:1521–1534, 2002.

[2] J. Sobel. Economists' models of learning. *Journal of Economic Theory*, 94(2):241–261, 2000.

[3] D. Gale and S. Kariv. Bayesian learning in social networks. *Games Econ. Behav.*, 45:329–346, 2003.

[4] D. Rosenberg, E. Solan, and N. Vieille. Informational externalities and emergence of consensus. *Games Econ. Behav.*, 66:979–994, 2009.

[5] P.M. Djuric and Y. Wang. Distributed bayesian learning in multiagent systems. *IEEE Signal Process. Mag.*, 29:65–76, March, 2012.

[6] Y. Kanoria and O. Tamuz. Tractable bayesian social learning. *Preprint*, 2011.

[7] E. Mossel and O. Tamuz. Efficient bayesian learning in social networks with gaussian estimators, 2011, arXiv: 1002.0747v2.

[8] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie. Bayesian learning in gaussian quadratic network games. *Working Paper*, 2012, Available at: http://www.seas.upenn.edu/∼ceksin/preprints/GQNG-IEEE.pdf.

[9] P. M. DeMarzo, D. Vayanos, and J. Zwiebel. Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, 118:909–968, 2003.

[10] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi. Non-Bayesian social learning. *Games and Economic Behavior*, 76(4):210–225, 2012.

[11] V. Bala and S. Goyal. Learning from neighbours. *Review of Economic Studies*, 65(3), 1998.

[12] B. Golub and M.O. Jackson. Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2:112–149, 2010.

[13] M. Rabbat, R. Nowak, and J. Bucklew. Generalized consensus computation in networked systems with erasure links. In *Proc. of IEEE 6th Workshop on the Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1088–1092, New York, NY, USA., June 2005.

[14] I. Schizas, A. Ribeiro, and G. Giannakis. Consensus in ad hoc wsns with noisy links - part i: distributed estimation of deterministic signals. *IEEE Trans. Signal Process.*, 56(1):1650–1666, January 2008.

[15] S. Kar and J. M. Moura. Convergence rate analysis of distributed gossip (linear parameter) estimation: Fundamental limits and tradeoffs. *IEEE J. Sel. Topics Signal Process.*, 5(4):674–690, 2011.

[16] F.S. Cattivelli and A.H. Sayed. Modeling bird flight formations using diffusion adaptation. *IEEE Trans. Signal Process.*, 59(5):2038–2051, 2011.

[17] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie. Learning in linear games over networks. In *Proceedings of the 50th Annual Allerton Conference on Communications, Control, and Computing (to appear)*, Allerton, Illinois, USA., 2012.