DISTRIBUTED FILTERS FOR BAYESIAN NETWORK GAMES

Ceyhun Eksin, Pooya Molavi, Alejandro Ribeiro, and Ali Jadbabaie

Department of Electrical and Systems Engineering, University of Pennsylvania

ABSTRACT

We consider a repeated network game where agents' utilities are quadratic functions of the state of the world and actions of all the agents. The state of the world is represented by a vector on which agents receive private signals with Gaussian noise. We define the solution concept as Bayesian Nash equilibrium and present a recursion to compute equilibrium strategies locally if an equilibrium exists at all stages. We further provide conditions under which a unique equilibrium exists. We conclude with an example of the proposed recursion in a repeated Cournot competition game and discuss properties of convergence such as efficient learning and convergence rate.

Index Terms— repeated network games, distributed algorithms, Bayesian learning.

I. INTRODUCTION

The generic model of Bayesian games over networks includes a state of the world on which agents have partial information, and individual utility functions that depend on the state, individual action and actions of other agents. In iterative learning models, agents with asymmetric information repeatedly make observations from their neighbors, infer about the state of the world and take actions. The normative model of learning in networks assumes that agents are rational. A natural model for rationality in uncertain environments is that agents are Bayesian in the way they form their beliefs and act optimally with respect to their individual utility functions and beliefs. In this paper, we consider a Bayesian network game with quadratic individual utility functions. The optimal behavior is defined by the Bayesian Nash Equilibrium (BNE).

The main goal in iterative learning models is to characterize convergent behavior of the population and determine the transient dynamics of individual rational behavior. However, the rational behavior imposes an overwhelming computational burden on agents even for small sized networks [1]. This intractability has led to the study of simplified learning models with 'non-Bayesian' agents, payoffs that depend only on self action and the state of the world (pure *information externality*), or specific signal and network structures. In 'non-Bayesian' models, agents are assumed to make inferences on the state of the world according to some heuristic rule [2]. In models with purely *informational externalities*, the actions of other agents do not affect self payoff. This simplifies the analysis as agents' optimal actions are just a function of their belief of the world. Even in this case, without any structural assumptions on information available to agents, only asymptotic analysis of the learning dynamics with rational agents is possible [3], [4]. In learning with pure informational externalities, there exists explicit characterization of rational behavior when the signals are Gaussian [1] or when the network structure is a tree [5]. In [6], we provide an asymptotic analysis of learning dynamics when the payoffs of agents are a quadratic function of actions of other agents, i.e., when there are both *information* and *payoff* externalities. In [7], we consider the same utility function and provide a local filter that propagates beliefs and computes equilibrium actions when agents' initial estimate of the scalar state of the world follows a Gaussian distribution. In this paper, we show that the proposed local filter extends to the case when state of the world is a vector. Additionally, we provide conditions imposed on the quadratic utility function for existence and uniqueness of the BNE at each step of the game.

Specifically, we consider an iterative learning scenario where each agent makes initial private observations of the vector state of the world that is corrupted by additive Gaussian noise (Section II). We define the BNE notion to characterize myopic optimal behavior of the agents (Section II-A). We show that the posterior distribution of private signals remain Gaussian for all agents and equilibrium strategy is linear in estimates of private signals at each stage if there exists a BNE (Theorem 1). Then we show that when the Hessian of the corresponding Bayesian potential function is symmetric and positive definite, BNE is unique (Proposition 1). Furthermore, we provide a filter in which beliefs are updated in a Bayesian way and BNE actions are computed locally (Section IV). We close the paper with a numerical example on a Cournot competition game where the state of the world is 2-dimensional and discuss convergence rate of the local filter (Section V).

II. BAYESIAN NETWORK GAMES

We consider games with incomplete information in which a population $V = 1, \ldots, N$ composed of agents in a network repeatedly choose actions and receive payoffs that depend on their own actions, a real-valued parameter $\boldsymbol{\theta} \in \mathbb{R}^m$ for $m \geq 1$, and actions of everyone else. An undirected connected network with nodes V and edge set E restricts the information flow. That is, agent i can only exchange information with neighboring agents $n(i) = \{j : \{j, i\} \in E\}$ that form an edge with it. We use $d(i)$ to denote the degree of agent i, that is, the cardinality of the set $n(i)$.

Work in this paper is supported 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.

The parameter $\boldsymbol{\theta}$ is unknown to agents. At time $t = 0$, each agent receives initial private signal $\mathbf{x}_i \in \mathbb{R}^m$,

$$\mathbf{x}_i = \boldsymbol{\theta} + \boldsymbol{\epsilon}_i \tag{1}$$

where the additive noise term $\boldsymbol{\epsilon}_i \in \mathbb{R}^m$ is multivariate Gaussian with zero mean and variance-covariance matrix $C_i \in \mathbb{R}^{m\times m}$. We use $\mathbf{x}_i[n]$ to denote the nth private signal of agent i where $n \le m$. We assume that private signals are independent among agents. We define the set of all private signals as

$$\mathbf{x} := [\mathbf{x}_1^T, \dots, \mathbf{x}_N^T]^T, \tag{2}$$

where $\mathbf{x} \in \mathbb{R}^{Nm\times 1}$.

At each stage $t = 0, 1, 2, \dots$, agent i takes action $\mathbf{a}_i(t) \in \mathbb{R}^m$ and receives a payoff that is quadratic in actions of all agents and state of the world,

$$u_i(\mathbf{a}_i, \{\mathbf{a}_j\}_{j\in V\setminus i}, \boldsymbol{\theta}) = -\frac{1}{2}\mathbf{a}_i^T\mathbf{a}_i + \sum_{j\in V\setminus\{i\}}\mathbf{a}_i^T B_{ij}\mathbf{a}_j + \mathbf{a}_i^T D\boldsymbol{\theta}, \tag{3}$$

where constants B_{ij} and D belong to $\mathbb{R}^{m\times m}$.

The payoff in (3) depends on actions of other agents. Hence, agent i has to reason about not only the state of the world but also behavior of other agents based on the available information. In this paper, we require that agents can observe actions of their neighbors at each stage. We define the information available to agent i at time t as his history $h_{i,t} := \{h_{i,t-1}, \mathbf{a}_{n(i)}(t-1)\}$ where the initial history is his private signal, $h_{i,0} = \{\mathbf{x}_i\}$. The behavior of agent i is determined by his information and strategy, that is, the strategy of agent i specifies the actions agent i takes at every stage of the game as a function of the observed history. Formally, a strategy for agent i denoted by σ_i is a sequence of functions $(\sigma_{i,\tau})_{\tau=0,\dots,\infty}$ where $\sigma_{i,t} : h_{i,t} \mapsto \mathbf{a}_i(t)$. We use σ to denote the strategy profile of all agents $\{\sigma_i\}_{i\in V}$. The strategy profile played up to time t is defined as $\sigma_{0:t}$. Agents' strategies are assumed to be common knowledge which implies that agent i can determine j's behavior at time t exactly if agent i is given $h_{j,t}$. As a result, reasoning about behavior of other agents is equivalent to reasoning about information of other agents.

II-A. Equilibrium

Let $\mathbf{P}$ denote agents common prior over the state of the world and private signals, that is, $\mathbf{P} = \mu_\theta \times \mu_{x_1} \cdots \times \mu_{x_n}$, where μ_θ is the improper uniform distribution over $\mathbb{R}^m$ and μ_{x_i} is the $\boldsymbol{\theta}$ mean normal distribution with variance given by C_i. We denote the expectation operator corresponding to $\mathbf{P}$ with $\mathbf{E}[\cdot]$. Since strategies map histories to actions and actions played determine histories in turn, strategies up to time t, $\sigma_{0:t-1}$, induce a probability on the histories at time t together with the prior $\mathbf{P}$. We denote the probability induced by a strategy at time t as $\mathbf{P}_{\sigma_{0:t-1}}$ and the corresponding expectation operator as $\mathbf{E}_{\sigma_{0:t-1}}[\cdot]$. Given a strategy profile σ, the best response of agent i at time t to the strategies of other agents $\{\sigma_{j,t}\}_{j\in V\setminus i}$ is a random function $\text{BR}_{i,t} : \mathbb{R}^{m(N-1)} \to \mathbb{R}^m$ defined as

$$\begin{aligned}&\text{BR}_{i,t}(\{\sigma_{j,t}\}_{j\in V\setminus i})\\ &\quad = \operatorname*{argmax}_{\mathbf{a}_i\in\mathbb{R}^m} \mathbf{E}_{\sigma_{0:t-1}}[u_i(\mathbf{a}_i, \{\sigma_{j,t}(h_{j,t})\}_{j\in V\setminus i}, \boldsymbol{\theta})|h_{i,t}]\end{aligned} \tag{4}$$

Note that here we consider myopic best responses, that is, agents do not consider the effect of their current actions on future payoffs.

So far we have not determined the strategies that agents use. A reasonable restriction for agents' strategies is that they are best responses to the strategies of other agents. This gives rise to our notion of equilibrium. A strategy profile σ^* is a BNE if it satisfies

$$\sigma^*_{i,t}(h_{i,t}) = \text{BR}_{i,t}(\{\sigma^*_{j,t}(h_{j,t})\}_{j\in V\setminus i}) \text{ for all } \quad i \in V,\, t \in \mathbb{N}. \tag{5}$$

BNE strategy is such that there is no other strategy that agent i could unilaterally deviate to that will provide a higher payoff at any point in time. Our equilibrium notion is based on the premise that agents are choosing actions that are myopically optimal given their information.

Note that if all agents have complete knowledge, that is, if they know the private signals of all agents $\mathbf{x}$, then the set of equilibrium strategies would be fixed at each stage. On the other hand, when agents have incomplete information, they need to estimate the private signals of others at each stage based on the new observed actions. Since they are refining their estimates by accumulating information over time, the equilibrium strategies are not necessarily time invariant.

For the utility function in (3), we obtain the best response function for agent i by taking the derivative of the expected utility function with respect to $\mathbf{a}_i$, equating it to zero, and solving for $\mathbf{a}_i$. As a result, the BNE equilibrium strategy defined by equations in (5) for the quadratic utility function in (3) is the solution to the following fixed point equation

$$\sigma^*_{i,t}(h_{i,t}) = \sum_{j\in V\setminus i} B_{ij}\mathbf{E}_{i,t}[\sigma^*_{j,t}(h_{j,t})] + D\mathbf{E}_{i,t}[\boldsymbol{\theta}] \tag{6}$$

for all $i \in V$ and $t \in \mathbb{N}$ where $\mathbf{E}_{i,t}[\cdot] := \mathbf{E}_{\sigma^*_{0:t-1}}[\cdot \,|\, h_{i,t}]$. Similarly, $\mathbf{P}_{i,t}(\cdot) := \mathbf{P}_{\sigma^*_{0:t-1}}[\cdot \,|\, h_{i,t}]$. In this paper, we assume that agents play according to the BNE strategy (6). In the following sections, we develop the method to calculate the equilibrium strategies at each step locally.

III. EQUILIBRIUM COMPUTATION

According to the model in Section II, at each stage agents use the observed history to estimate the unknown parameter as well as the histories of other agents. They use the common knowledge BNE strategy and their estimates on histories of other agents to form beliefs on actions of agents $\mathbf{P}_{i,t}(\{\mathbf{a}_j(t)\}_{j\in V\setminus i})$. For an outside observer who knows the private signals $\mathbf{x}$, the network structure and BNE strategy, the trajectory of the game is deterministically known. Thus it is sufficient for agents to keep track of their estimates of $\mathbf{x}$ in order to form beliefs on the histories and actions of other agents. In this section, we provide a method for an outside observer to track the trajectory of the game given $\mathbf{x}$.

Given agent i's estimate of all the private signals at time t, $\mathbf{E}_{i,t}[\mathbf{x}]$, we define the corresponding error covariance matrix for agent i, $M^i_{\mathbf{xx}}(t) \in \mathbb{R}^{Nm\times Nm}$, as $M^i_{\mathbf{xx}}(t) := \mathbf{E}_{i,t}[(\mathbf{x} - \mathbf{E}_{i,t}[\mathbf{x}])(\mathbf{x} - \mathbf{E}_{i,t}[\mathbf{x}])^T]$. Further, we denote agent i's estimate of $\boldsymbol{\theta}$ at time t as $\mathbf{E}_{i,t}[\boldsymbol{\theta}]$, and the corresponding error variance value, $M^i_{\boldsymbol{\theta\theta}}(t) \in \mathbb{R}^{m\times m}$, as $M^i_{\boldsymbol{\theta\theta}}(t) := \mathbf{E}_{i,t}[(\boldsymbol{\theta} - \mathbf{E}_{i,t}[\boldsymbol{\theta}])(\boldsymbol{\theta} - \mathbf{E}_{i,t}[\boldsymbol{\theta}])^T]$. Agent i's estimate of $\boldsymbol{\theta}$ is based on her estimate of the private signals; therefore, agent i also needs to keep track of the error covariance between her estimate of $\boldsymbol{\theta}$ and her estimate of $\mathbf{x}$, $M^i_{\boldsymbol{\theta}\mathbf{x}}(t) \in \mathbb{R}^{m\times Nm}$, defined as $M^i_{\boldsymbol{\theta}\mathbf{x}}(t) := \mathbf{E}_{i,t}[(\boldsymbol{\theta} - \mathbf{E}_{i,t}[\boldsymbol{\theta}])(\mathbf{x} - \mathbf{E}_{i,t}[\mathbf{x}])^T]$.

At node i, based on signal model (1), the initial estimates of $\boldsymbol{\theta}$ and $\mathbf{x}$ are Gaussian with means that are expressed as linear combinations of $\mathbf{x}_i$. Hence, the initial expected mean at i can be expressed as a linear combination of all the private signals by sparse estimation matrices

$$\mathbf{E}_{i,0}[\boldsymbol{\theta}] = Q_{i,0}\mathbf{x}, \text{ and } \mathbf{E}_{i,0}[\mathbf{x}] = L_{i,0}\mathbf{x}. \tag{7}$$

where $Q_{i,0} \in \mathbb{R}^{m\times Nm}$ and $L_{i,0} \in \mathbb{R}^{Nm\times Nm}$. Both of these estimation matrices have zeros except in the ith m column block. Initial error variance-covariance matrices can also be defined accordingly.

Our goal is to characterize Bayesian update of estimates and calculation of equilibrium actions. We start by assuming that at time t agents' estimates of $\boldsymbol{\theta}$ and $\mathbf{x}$ are normally distributed with means

$$\mathbf{E}_{i,t}[\boldsymbol{\theta}] = Q_{i,t}\mathbf{x}, \text{ and } \mathbf{E}_{i,t}[\mathbf{x}] = L_{i,t}\mathbf{x}, \tag{8}$$

where $Q_{i,t} \in \mathbb{R}^{m\times Nm}$ and $L_{i,t} \in \mathbb{R}^{Nm\times Nm}$ are known estimation weights.

We further assume that there exists an equilibrium strategy that is linear in expectations of private signals,

$$\sigma^*_{i,t}(h_{i,t}) = U_{i,t}\mathbf{E}_{i,t}[\mathbf{x}] \quad \text{for all } i \in V, \tag{9}$$

for action coefficients $U_{i,t} \in \mathbb{R}^{m\times Nm}$. By substituting the candidate strategies in (9) to the BNE condition in (6) for all $i \in V$, we obtain the following equations

$$U_{i,t}\mathbf{E}_{i,t}[\mathbf{x}] = \sum_{j\in V\setminus i} B_{ij}\mathbf{E}_{i,t}[U_{j,t}\mathbf{E}_{j,t}[\mathbf{x}]] + D\mathbf{E}_{i,t}[\boldsymbol{\theta}]. \tag{10}$$

for all $i \in V$. After using the fact that $\mathbf{E}_{i,t}[\mathbf{E}_{j,t}[\mathbf{x}]] = L_{j,t}\mathbf{E}_{i,t}[\mathbf{x}]$ with mean estimate assumptions in (8) for the corresponding terms in (10) and ensuring that the strategies in (9) satisfy the equilibrium equations for any realization of history by equating coefficients that multiply each component of $\mathbf{x}$, we obtain the set of equations given by

$$L^T_{i,t}U^T_{i,t} = \sum_{j\in V\setminus i} L^T_{i,t}L^T_{j,t}U^T_{j,t}B^T_{ij} + Q^T_{i,t}D^T \text{ for all } i \in V \tag{11}$$

which we can solve to get the action coefficients $\{U_{i,t}\}_{i\in V}$. The existence of a linear equilibrium strategy means that the set of linear equations in (11) has at least one solution. In Section III-A, we provide conditions for existence and uniqueness of a solution.

For a linear equilibrium strategy, the actions can be written as a linear combination of the private signals using (8), that is, the action of agent i at time t is given by

$$\mathbf{a}_i(t) = U_{i,t}L_{i,t}\mathbf{x} \qquad \text{for all } i \in V. \tag{12}$$

Being able to express actions as in (12) permits writing observations of agents in linear form. From the perspective of an observer, the action $\mathbf{a}_j(t)$ is equivalent to observing a linear combination of private signals. As a result, we can represent observation vector of agent i, $\mathbf{a}_{n(i)}(t) := \left[\mathbf{a}_{j_1}(t), \ldots, \mathbf{a}_{j_{d(i)}}(t)\right]^T \in \mathbb{R}^{md(i)}$ in linear form as

$$\mathbf{a}_{n(i)}(t) = H^T_{i,t}\mathbf{x} = [U_{j_1,t}L_{j_1,t}; \ldots; U_{j_{d(i)},t}L_{j_{d(i)},t}]\mathbf{x} \tag{13}$$

where $H^T_{i,t} = [U_{j_1,t}L_{j_1,t}; \ldots; U_{j_{d(i)},t}L_{j_{d(i)},t}] \in \mathbb{R}^{md(i)\times Nm}$ is the observation matrix of agent i. Agent i's belief of $\mathbf{x}$ at time t is Gaussian by assumption, and at time $t+1$ agent i observes a linear combination of $\mathbf{x}$. Hence, agent i's belief at time $t+1$ can be obtained by a sequential LMMSE update. As a result, mean estimates remain weighted sums of private signals as in (8). In the following lemma, we explicitly present the way we compute the estimation weights, $L_{i,t+1}$ and $Q_{i,t+1}$, at time $t+1$.

Lemma 1 *Consider a Bayesian game with quadratic function as in* (3) *with the belief and BNE strategy assumptions as above. Further define the gain matrices as*

$$K^i_{\mathbf{x}}(t) := M^i_{\mathbf{xx}}(t)H_{i,t}\left(H^T_{i,t}M^i_{\mathbf{xx}}(t)H_{i,t}\right)^{-1}, \tag{14}$$
$$K^i_{\boldsymbol{\theta}}(t) := M^i_{\boldsymbol{\theta}\mathbf{x}}(t)H_{i,t}\left(H^T_{i,t}M^i_{\mathbf{xx}}(t)H_{i,t}\right)^{-1}. \tag{15}$$

If agents play according to a linear equilibrium strategy then agent i's posterior $\mathbf{P}_{i,t+1}([\boldsymbol{\theta}^T, \mathbf{x}^T])$ is Gaussian with means that are linear combination of private signals,

$$\mathbf{E}_{i,t+1}[\boldsymbol{\theta}] = Q_{i,t+1}\mathbf{x}, \text{ and } \mathbf{E}_{i,t+1}[\mathbf{x}] = L_{i,t+1}\mathbf{x}, \tag{16}$$

where the estimation matrices are given by

$$L_{i,t+1} = L_{i,t} + K^i_{\mathbf{x}}(t)\left(H^T_{i,t} - H^T_{i,t}L_{i,t}\right), \tag{17}$$
$$Q_{i,t+1} = Q_{i,t} + K^i_{\boldsymbol{\theta}}(t)\left(H^T_{i,t} - H^T_{i,t}, L_{i,t}\right), \tag{18}$$

and the covariance matrices are further given by

$$M^i_{\mathbf{xx}}(t+1) = M^i_{\mathbf{xx}}(t) - K^i_{\mathbf{x}}(t)H^T_{i,t}M^i_{\mathbf{xx}}(t), \tag{19}$$
$$M^i_{\boldsymbol{\theta\theta}}(t+1) = M^i_{\boldsymbol{\theta\theta}}(t) - \left[K^i_{\boldsymbol{\theta}}(t)^T H^T_{i,t}M^i_{\mathbf{x}\boldsymbol{\theta}}(t)\right]^T, \tag{20}$$
$$M^i_{\boldsymbol{\theta}\mathbf{x}}(t+1) = M^i_{\theta\mathbf{x}}(t) - K^i_{\boldsymbol{\theta}}(t)H^T_{i,t}M^i_{\mathbf{xx}}(t). \tag{21}$$

Proof sketch[1]*:* Since observations of i, $\mathbf{a}_{n(i)}(t)$, are linear combinations of private signals $\mathbf{x}$ which are Gaussian, observations of i are also normally distributed from the perspective of i. Furthermore, by assumption in (8), the prior distribution $\mathbf{P}_{i,t}(\mathbf{x})$ is Gaussian. Hence, the posterior distribution is also Gaussian. Specifically, the mean of the posterior distribution corresponds to the LMMSE estimator with gain matrix in (14); that is,

$$\mathbf{E}_{i,t+1}[\mathbf{x}] = \mathbf{E}_{i,t}[\mathbf{x}] + K^i_{\mathbf{x}}(t)\left(\mathbf{a}_{n(i)}(t) - \mathbf{E}_{i,t}[\mathbf{a}_{n(i)}(t)]\right). \tag{22}$$

Because $\boldsymbol{\theta}$ and $\mathbf{x}$ are jointly Gaussian at time t, $\boldsymbol{\theta}$ and $\mathbf{a}_{n(i)}(t)$ are also jointly Gaussian. Consequently, the estimate of $\boldsymbol{\theta}$ is given by a sequential LMMSE estimator with gain matrix in (15),

$$\mathbf{E}_{i,t+1}[\boldsymbol{\theta}] = \mathbf{E}_{i,t}[\boldsymbol{\theta}] + K^i_{\boldsymbol{\theta}}(t)\left(\mathbf{a}_{n(i)}(t) - \mathbf{E}_{i,t}\left[\mathbf{a}_{n(i)}(t)\right]\right). \tag{23}$$

Substituting mean estimate assumptions at time t (8) and observation matrix (13) inside (22) and (23) and by grouping terms that multiply $\mathbf{x}$, we obtain the estimation weights recursion in (17) and (18). Similarly, the updates for error covariance matrices

[1]Proofs of results in this paper are available in [8].

are as given in (19)–(21) following standard LMMSE updates. ∎

Lemma 1 shows that when mean estimates are linear combinations of private signals at time t, they remain that way at time $t+1$. In the next theorem, we show that assumption in (8) is indeed true for all time by realizing that the estimates at time $t=0$ are linear combinations of private signals.

Theorem 1 *Given the quadratic utility function in* (3)*, if there exists a linear equilibrium strategy σ_t^* as in* (9) *for $t \in \mathbb{N}$, then the action coefficients $U_{i,t}$ can be computed by solving the system of linear equations in* (11)*, and further, agents' estimates of $\mathbf{x}$ and $\boldsymbol{\theta}$ are linear combinations of private signals as in* (8) *with estimation matrices computed recursively using* (14)-(15) *and* (17)-(21) *with initial mean values* (7)*.*

Theorem 1 shows that an external observer can compute the equilibrium actions explicitly when the strategy profile and network structure is known, and there exists a linear equilibrium strategy at each time. Next we provide conditions for the uniqueness and existence of linear equilibrium strategies at all times. Then in the following section we present a local algorithm for individual agents to propagate their beliefs and compute equilibrium strategies.

III-A. Existence and uniqueness of linear equilibrium strategy

The existence and uniqueness of linear equilibrium strategy for Bayesian quadratic games have been studied in economics literature [9], [10] for single stage games. By using these existing results with Theorem 1, we show that under certain conditions on utility function (3), there exists a unique linear equilibrium strategy at all times.

Proposition 1 *Given the utility function in* (3)*, define the matrix $B \in \mathbb{R}^{Nm \times Nm}$ with $m \times m$ diagonal blocks where i,jth $m \times m$ block $B[i,j] := [-B_{ij}]$ for $i \in V$, $j \in V \setminus i$ and $B[i,i] = I$. If B is a symmetric positive definite matrix then there exists a unique equilibrium action linear in expectations of private signals at each stage $t \in \mathbb{N}$.*

Proof Sketch: First, we define the following value function,

$$v(\mathbf{a},\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i \in V}\mathbf{a}_i^T\mathbf{a}_i + \frac{1}{2}\sum_{i\in V}\sum_{j \in V\setminus i}\mathbf{a}_i^T B_{ij}\mathbf{a}_j + \sum_{i \in V}\mathbf{a}_i^T D\boldsymbol{\theta}$$
$$= -\frac{1}{2}\mathbf{a}^T B\mathbf{a} + \mathbf{a}^T(\mathbf{1}\otimes D)\boldsymbol{\theta} \quad (24)$$

where $\mathbf{1} \in \mathbb{R}^{N\times 1}$ is a vector of ones and $\mathbf{a} = [\mathbf{a}_1^T, \ldots, \mathbf{a}_N^T]^T$. Note that $\partial \mathbf{E}_{i,t}[u_i(\mathbf{a},\boldsymbol{\theta})]/\partial \mathbf{a}_i = \partial \mathbf{E}_{i,t}[v(\mathbf{a},\boldsymbol{\theta})]/\partial \mathbf{a}_i$ for all $i \in V$ and $h_{i,t}$ where $u_i(\cdot)$ is as in (3). In other words, $v(\cdot)$ is a Bayesian potential function for the game $\{u_i(\cdot)\}_{i\in V}$. Furthermore, the BNE of the Bayesian potential function in (24) correspond to the BNE of the game with utility functions given by $\{u_i(\cdot)\}_{i\in V}$, see Lemma 5 in [10].

Theorem 5 in [9] states that if individual signals are jointly Gaussian as in (1) and B is positive definite, for the 'team' payoff function in (24) the unique equilibrium is linear in self private signals. Hence, the game in (3) has a unique equilibrium strategy linear in self private signal at time $t=0$. As a result, agents' observations at time $t=1$, that is, $\mathbf{a}_{n(i)}(0)$, are also Gaussian random variables by Theorem 1. Consequently, the signals observed by agents $\{h_{j,1}\}_{j\in V}$ are jointly Gaussian. Again by Theorem 5 in [9], the equilibrium strategy at time $t=1$ is a linear combination of observed signals. Furthermore, there exists some weighting matrix $\Phi_{i,1} \in \mathbb{R}^{(d(i)+1)m\times Nm}$ such that $h_{i,1} = \Phi_{i,t}\mathbf{E}_{i,t}[\mathbf{x}]$ for $i \in V$. Henceforth the equilibrium action can be written as a linear combination of expectations of private signals as in (9). The induction argument can be completed by assuming the actions up to time t, $\{\mathbf{a}_i^*(s) : i \in V, s < t\}$ are of the form in (9) and following the same reasoning. ∎

Algorithm 1: QNG filter for $\boldsymbol{\theta} \in \mathbb{R}^m$

1 *Initialization:* Set posterior distribution on $\boldsymbol{\theta}$ and $\mathbf{x}$

$$\begin{bmatrix}\boldsymbol{\theta}\\ \mathbf{x}\end{bmatrix} \mid h_{i,0} \sim \mathcal{N}\left(\begin{bmatrix}Q_{i,0}\mathbf{x}\\ L_{i,0}\mathbf{x}\end{bmatrix}, \begin{pmatrix}M^i_{\boldsymbol{\theta\theta}}(0), M^i_{\boldsymbol{\theta}\mathbf{x}}(0)\\ M^i_{\mathbf{x}\boldsymbol{\theta}}(0), M^i_{\mathbf{xx}}(0)\end{pmatrix}\right)$$

and $\{L_{j,0}, Q_{j,0}\}_{j\in V}$ according to signal model (1).

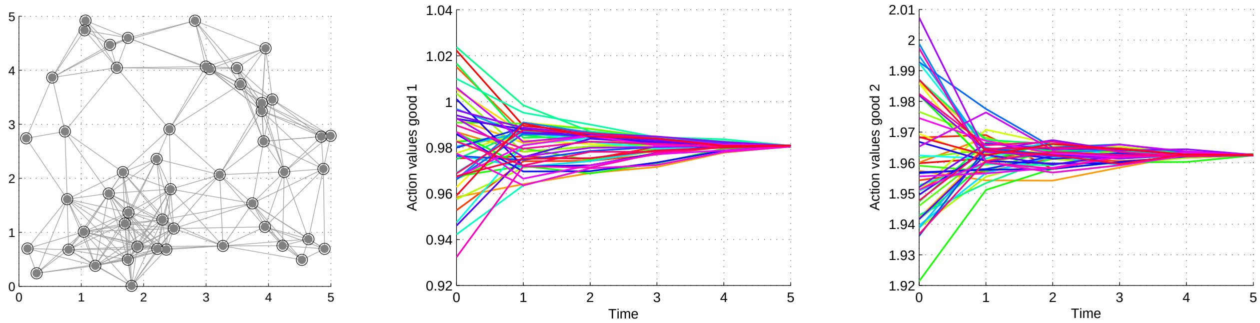
2 **For** $t = 0, 1, 2, \ldots$

1) *Equilibrium strategy:* Solve for $\{U_{j,t}\}_{j\in V}$ using the set of equations in (11).
2) *Play and observe:* Take action $\mathbf{a}_i(t) = U_{i,t}\mathbf{E}_{i,t}[\mathbf{x}]$ and observe $\mathbf{a}_{n(i)}(t)$.
3) *Observation matrix:* Construct $H_{i,t}$ using (13).
4) *Bayesian estimates:* Calculate $\mathbf{E}_{i,t+1}[\mathbf{x}]$ and $\mathbf{E}_{i,t+1}[\boldsymbol{\theta}]$ using (22) and (23), respectively. Update error covariance matrices using (19)–(21).
5) *Estimation weights:* Construct $\{H_{j,t}\}_{j\in V}$ using (13) and update $\{L_{j,t+1}, Q_{j,t+1}\}_{j\in V}$ using (17)–(18).

IV. QUADRATIC NETWORK GAME FILTER

In the QNG filter summarized in Algorithm 1, we provide a sequential local algorithm for agent i to calculate updates for $\boldsymbol{\theta}$ and $\mathbf{x}$ and to act according to equilibrium strategy. In the calculation of action and estimation coefficients in QNG filter, we make use of the assumptions that signal and network structure, and the strategy profile are common knowledge.

The QNG filter entails a full network simulation in which agent i maintains individual beliefs while keeping track of computations of all the agents in the network. Initially, agent i knows estimation weights of all agents which are $\{L_{j,0}, Q_{j,0}\}_{j\in V}$ given common knowledge of the signal structure. Note that this does not imply that agent i knows private signals of other agents. Using the estimation weights at time $t = 0, 1, 2, \ldots$ agent i constructs the system of equations in (11) and solves for individual action coefficients $\{U_{j,t}\}_{j\in V}$ – see Step 1 in QNG filter. Note that the solution for the action coefficients in (11) does not depend on the realization of private signals. In step 2, agent i multiples her private signal estimate $\mathbf{E}_{i,t}[\mathbf{x}]$ by the vector $U_{i,t}$ to determine self equilibrium play. Unlike the action coefficients, the actions realized depend on the observed history, and hence on the realization of the private signals. Next agent i observes



(a) (b) (c)

Fig. 1. Geometric network with $N = 50$ agents (a): Agents are randomly placed on a 5×5 square. There exists an edge between any pair of agents with distance less than 1.5 meter apart in the geometric network. Number of good 1 and 2 produced by each agent are depicted in (b) and (c), respectively.

actions of neighboring agents and uses $\{U_{j,t}, L_{j,t}\}_{j\in n(i)}$ to form the observation matrix $H_{i,t}$. She uses the observation matrix to refine her estimates on $\mathbf{x}$ and $\boldsymbol{\theta}$ according to a sequential LMMSE update using (21) and (22) – see step 4. Note that the estimation weights $L_{i,t}$ and $Q_{i,t}$ cannot be used to calculate the mean estimates provided by Theorem 1, unless the private signals $\mathbf{x}$ are exactly known. Finally in step 5, agent i constructs observation matrices of all the agents via (13) and updates estimation coefficients of all the agents $\{L_{j,t}, Q_{j,t}\}_{j\in V}$ using (17)–(18) which are necessary to compute equilibrium action coefficients of the next step.

V. COURNOT COMPETITION GAME

Consider a competition model in which N firms compete on the amount of goods they produce. There are m goods and each firm's decision of how much to produce of a certain good affects the price of that good. Specifically, the selling unit price for good n decreases linearly by the total amount produced $\mathbf{p}[n] - \sum_{j\in V} \mathbf{a}_j[n]$ where $\mathbf{p}[n]$ is the constant market price when zero good n is produced. Further, each unit of good n produced has a fixed unit cost $\mathbf{c}[n]$ that is identical for all firms. The profit of firm i for production levels $\mathbf{a} \in \mathbb{R}^m$ is given by the utility

$$u_i(\mathbf{a}_i, \{\mathbf{a}_j\}_{j\in V\setminus i}, \boldsymbol{\theta}) = \Big(\boldsymbol{\theta} - \mathbf{a}_i - \sum_{j\in V\setminus i} \mathbf{a}_j\Big)^T \mathbf{a}_i \qquad (25)$$

where we define $\boldsymbol{\theta}[n] := \mathbf{p}[n] - \mathbf{c}[n]$ as the effective unit profit of good n. By rearranging terms one can see that this utility function is of the same form in (3).

In this example, we consider $m = 2$ goods with effective unit profits $\boldsymbol{\theta} = [50, 100]^T$. Agent i makes private observations $\mathbf{x}_i$ on $\boldsymbol{\theta}$ based on (1) with $\boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, I)$. We evaluate convergence behavior in geometric network with $N = 50$ agents; see Fig. 1 (a). The network has a diameter of $d = 5$. The action values of each agent for good 1 and 2 are depicted in Fig. 1 (b)-(c), respectively. The results show that agents' actions converge to the Nash equilibrium with complete information at time $t = 5$, that is, nth action of i at time $t = 5$ is

$$\mathbf{a}_i(5)[n] = \frac{\mathbf{E}\left[\boldsymbol{\theta}[n] \,\middle|\, \mathbf{x}\right]}{N+1} \quad i \in V,\ n = 1, 2. \qquad (26)$$

This implies that agents learn the sufficient statistic to calculate the best estimates of effective unit profit for each good in the amount of time it takes for an information to propagate through the entire network.

VI. REFERENCES

[1] E. Mossel and O. Tamuz, "Efficient Bayesian learning in social networks with Gaussian estimators," *ArXiv e-prints*, April 2010.

[2] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-Bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 4, pp. 210–225, 2012.

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

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

[5] Y. Kanoria and O. Tamuz, "Tractable bayesian social learning on trees," in *Proc. of the IEEE Intl. Symp. on Information Theory (ISIT)*, 2012, pp. 2721–2725.

[6] 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.

[7] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie, "Bayesian quadratic network game filters," in *Proc. Int. Conf. Acoustics Speech Signal Process. (to appear)*, Vancouver, Canada, 2013.

[8] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie, "Bayesian quadratic network game filters," *ArXiv e-prints*, February 2013.

[9] R. Radner, "Team decision problems," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 857–881, 1962.

[10] T. Ui, "Bayesian potentials and information structures: Team decision problems revisited," *International Journal of Economic Theory*, vol. 5, no. 3, pp. 271–291, 2009.