BAYESIAN QUADRATIC NETWORK GAME FILTERS

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

Department of Electrical and Systems Engineering, University of Pennsylvania

ABSTRACT

A repeated network game where agents' utilities depend on information and payoff externalities is considered. Agents play Bayesian Nash Equilibrium strategies with respect to their beliefs on the state of the world and the actions of all other nodes in the network. These beliefs are refined over subsequent stages based on the observed actions of neighboring peers. This paper introduces the Quadratic Network Game (QNG) filter that agents can run locally to update their beliefs, select corresponding optimal actions, and eventually learn a sufficient statistic of the network's state. The QNG filter is demonstrated on a coordination game.

Index Terms— Network Games, Bayesian Nash Equilibrium, Linear Filtering, Distributed Signal Processing.

1. INTRODUCTION

Consider a network of autonomous agents intent on selecting actions that maximize local quadratic utilities that depend on an unknown state of the world – information externalities – and also the unknown actions of all other agents – payoff externalities. In a Bayesian setting agents form a belief on the actions of their peers and select an action that maximizes the expected payoff with respect to those beliefs. In turn, forming these beliefs requires that each network element make a model of how other members will respond to their local beliefs. The natural assumption is that they exhibit the same behavior, namely that they are also making a model of other nodes' responses. But that means we need a model of their model which shall include their model of our model of their model and so on. The fixed point of this iterative modeling effort is a Bayesian Nash Equilibrium (BNE). Here, we consider repeated versions of this game in which agents observe the actions taken by neighboring agents at a given time. Observation of neighboring actions alters agents' beliefs leading to the selection of new actions which become known at the next play prompting further reevaluation of beliefs and corresponding actions. In this context we talk of Bayesian learning as the agents' goal which can be interpreted as the eventual learning of peers' actions. This paper introduces the Quadratic Network Game (QNG) filter that agents can run locally to update their beliefs, select corresponding optimal actions, and eventually learn a sufficient statistic of the network's state.

The burden of computing a BNE in repeated games is, in general, overwhelming even for small sized networks [1]. This intractability has led to the study of simplified models in which agents are non-Bayesian but update their beliefs according to some heuristic rule [2–6]. A different simplification is obtained in models with pure information externalities where payoffs depend on the local action and an underlying state but not on the actions of others. This is reminiscent of distributed estimation [7–12] since agents figure out the state of the world from observed neighboring actions without strategic considerations on the actions of peers. Computations are still intractable in the case of pure information externalities and for the most part only asymptotic analysis of the learning dynamics with rational agents is possible [13–15]. However, there exist explicit characterizations when signals are Gaussian [1] or when the network structure is a tree [16]. For the network games considered here in which there are information as well as payoff externalities not much is known besides an asymptotic analysis of learning dynamics [17–19].

Work in this paper is supported by ARO W911NF-10-1-0388, NSF CAREER CCF-0952867, NSF CCF-1017454, and AFOSR MURI FA9550-10-1-0567.

The QNG filter provides a mechanism to compute BNE actions for quadratic games in which the initial condition of each agent is a single private observation of the state of the world corrupted by additive Gaussian noise (Section 2). We begin by showing that the propagation of local posterior probability distributions on the system's state as new neighboring actions are observed follows a Gaussian distribution (Section 3). We then employ a complete induction argument to derive an explicit recursion for the tracking of means and covariances of these distributions (Theorem 1) which we leverage to develop the QNG filter (Section 4). We apply the QNG filter on a vector coordination game (Section 5).

2. GAUSSIAN QUADRATIC GAMES

We consider games with incomplete information in which N identical agents in a network repeatedly choose actions and receive payoffs that depend on their own actions, an unknown scalar parameter $\theta \in \mathbb{R}$, and actions of all other agents. The network is represented by an undirected connected graph $G = (V, E)$ with node set $V = 1, \ldots, N$ and edge set E. The network structure restricts the information available to agent i who is assumed to observe actions of agents j in its neighborhood $n(i) := \{j : \{j, i\} \in E\}$ composed of agents that share an edge with him. The degree of node i is given by the cardinality of the set $n(i)$ and denoted as $d(i) := \#n(i)$. The neighbors of i are denoted $j_{i,1} <, \ldots, < j_{i,d(i)}$. We assume the network graph G is known to all agents.

At time $t = 0$ agent i observes a private signal $x_i \in \mathbb{R}$ which we model as being given by the unknown parameter θ contaminated with zero mean additive Gaussian noise ϵ_i,

$$x_i = \theta + \epsilon_i, \tag{1}$$

The noise variances are denoted as $C_i := \mathbf{E}\left[\epsilon_1^2\right]$ and grouped in the vector $\mathbf{C} := [C_1, \ldots, C_N]^T$ which is assumed known to all agents. The noise terms ϵ_i are further assumed independent across agents. For future reference define the vector of private signals $\mathbf{x} := [x_1, \ldots, x_N]^T \in \mathbb{R}^{N\times 1}$ grouping all local observations.

Consider a discrete time variable $t = 0, 1, 2, \ldots$ to index subsequent stages of the game. At each stage t agent i takes scalar action $a_i(t) \in \mathbb{R}$. The selection of agent i, along with the concurrent selections $a_j(t)$ of all other agents $j \in V \setminus i$ results in a payoff $u_i(a_i(t), \{a_j(t)\}_{j\in V\setminus i}, \theta)$ that agent i wants to make as large as possible. In this paper we restrict attention to quadratic payoffs. Specifically, selection of actions $\{a_i = a_i(t)\}_{i\in V}$ when the state of the world is θ results in agent i receiving

$$u_i(a_i, \{a_j\}_{j\in V\setminus i}, \theta) = -\frac{1}{2}a_i^2 + \sum_{j\in V\setminus\{i\}} \beta_{ij} a_i a_j + \delta a_i \theta + f(\{a_j\}_{j\in V\setminus i}, \theta), \tag{2}$$

where $\beta_{ij} \in \mathbb{R}$ for all $i \in V$, $j \in V \setminus i$ and $\delta \in \mathbb{R}$ are constants and $f(\cdot)$ is some function with arguments $\{a_j\}_{j\in V\setminus i}$ and θ. Notice that since $\partial^2 u_i/\partial a_i^2 = -1 < 0$ the payoff function in (2) is strictly concave with respect to the self action a_i of agent i.

Agent i cannot select the action $a_i(t)$ that maximizes the payoff in (2) because neither θ nor the actions $\{a_j(t)\}_{j\in V\setminus i}$ are available to him. Hence, agent i needs to reason about his beliefs on state θ and actions $\{a_j(t)\}_{j\in V\setminus i}$ based on its available information. At the playing of stage $t-1$, agent i observes the actions $\mathbf{a}_{n(i)}(t-1) := [a_{j_{i,1}}(t -$

$1), \ldots, a_{j_{i,d(i)}}(t-1)]^T \in \mathbb{R}^{d(i)\times 1}$ of all agents in his neighborhood. In general, at any point in time t the history of observations $h_{i,t}$ is augmented to incorporate the actions of neighbors in the previous stage,

$$h_{i,t} := \big\{h_{i,t-1}, \mathbf{a}_{n(i)}(t-1)\big\} = \big\{x_i, \mathbf{a}_{n(i)}(u), u < t\big\}. \quad (3)$$

Observed action history $h_{i,t}$ is then used to update estimates of the world state θ and the upcoming actions $\{a_j(t)\}_{j\in V\setminus i}$ of all other agents leading to the selection of the action $a_i(t)$ in the current stage of the game.

Next, we introduce the strategy $\sigma_{i,t}$ of agent i that is used to map histories to actions. The strategy $\sigma_{i,t}$ of agent i at time t is a σ-algebra measurable with respect to the history $h_{i,t}$. In this paper we focus on pure strategies that can be written as functions that map $h_{i,t}$ to $a_i(t)$, that is, $\sigma_{i,t} : h_{i,t} \mapsto a_i(t)$. We emphasize the difference between strategy and action. An action $a_i(t)$ is the play of agent i at time t, whereas strategies $\sigma_{i,t}$ refer to the map of histories to actions. We can think of the action $a_i(t) = \sigma_{i,t}(h_{i,t})$ as the value of the strategy function $\sigma_{i,t}$ associated with the given observed history $h_{i,t}$. Further define the strategy of agent i as the concatenation $\sigma_i := \{\sigma_{i,u}\}_{u=0,\ldots,\infty}$ of strategies that agent i plays at all times. Use $\sigma_t := \{\sigma_{i,t}\}_{i\in V}$ to refer to the strategies of all players at time t, $\sigma_{0:t} := \{\sigma_u\}_{u=0,\ldots,t}$ to represent the strategies played by all players between times 0 and t, and $\sigma := \{\sigma_u\}_{u=0,\ldots,\infty} = \{\sigma_i\}_{i\in V}$ to denote the strategy profile for all agents $i \in V$ and $t \in \mathbb{N}$. As in the case of the network topology, the strategy σ is also assumed to be known to all agents. We study mechanisms for the construction of strategies in the following section.

2.1. Bayesian Nash equilibria

Given that agent i wants to maximize the utility in (2) but has access to the partial information available in the observed history $h_{i,t}$ in (3) a reasonable strategy $\sigma_{i,t}$ is to select the action $a_i(t)$ that maximizes the expected utility with respect to $h_{i,t}$. This expected utility depends on strategies $\sigma_{0:t-1}$ played in the past by all agents and on strategies $\{\sigma_{j,t}\}_{j\in V\setminus i}$ that all other agents are to play at time t. Fix then the past strategies $\sigma_{0:t-1}$ and the upcoming strategies $\{\sigma_{j,t}\}_{j\in V\setminus i}$, and define the best response of player i at time t as

$$\mathrm{BR}_{i,t}\Big(\sigma_{0:t-1}, \{\sigma_{j,t}\}_{j\in V\setminus i}\Big) \quad (4)$$
$$:= \operatorname*{argmax}_{a_i\in\mathbb{R}} \mathbf{E}_{\sigma_{0:t-1}}\Big[u_i(a_i, \{\sigma_{j,t}\}_{j\in V\setminus i}, \theta) \,\Big|\, h_{i,t}\Big].$$

The strategies $\sigma_{0:t-1}$ played at previous times mapped respective histories $\{h_{j,u}\}_{j\in V}$ to actions $\{a_j(u)\}_{j\in V}$ for $u < t$. Therefore, the past strategies $\sigma_{0:t-1}$ determine the manner in which agent i updates his beliefs on the state of the world θ and on the histories $\{h_{j,t}\}_{j\in V\setminus i}$ observed by other agents. The strategy profiles $\{\sigma_j(t)\}_{j\in V\setminus i}$ of other players in the current stage permit transformation of history beliefs $\{h_{j,t}\}_{j\in V\setminus i}$ into a probability distribution over respective upcoming actions $\{a_j(t)\}_{j\in V\setminus i}$. The resulting joint distribution on $\{a_j(t)\}_{j\in V\setminus i}$ and θ permits evaluation and maximization of the expectation in (4).

One can think of the profiles $\{\sigma_j(t)\}_{j\in V\setminus i}$ played by other agents in the upcoming stage as the model agent i makes of the behavior of other agents. In that sense the sensible assumption is that other agents are also playing best response to a best response model of other agents. This modeling assumption leads to the definition of Bayesian Nash equilibria (BNE) as the solution to the fixed point equation

$$\sigma^*_{i,t}(h_{i,t}) = \mathrm{BR}_{i,t}\Big(\sigma^*_{0:t-1},\ \{\sigma^*_{j,t}\}_{j\in V\setminus i}\Big), \quad \text{for all } h_{i,t}, \quad (5)$$

where we have also added the restriction that an equilibrium strategy $\sigma^*_{i,u}$ has been played for all times $u < t$. We emphasize that (5) needs to be satisfied for all possible histories $h_{i,t}$ and not just for the history realized in a particular game realization.

If all agents play their BNE strategies as defined in (5), there is no strategy that agent i could unilaterally deviate to that provides a higher expected payoff than $\sigma^*_{i,t}$ [cf. (4)]. In this paper we restrict attention to games in which all agents play the BNE strategy $\sigma^*_{i,t}$ at all times. To simplify future notation define the expectation operator $\mathbf{E}_{i,t}[\cdot] := \mathbf{E}_{\sigma^*_{0:t-1}}[\cdot\,|h_{i,t}]$, that represents expectation with respect to the local history $h_{i,t}$ when agents played the equilibrium strategy $\sigma^*_{0:t-1}$ in earlier stages of the game. Similarly we define the conditional probability distribution of agent i at time t given past strategies $\sigma^*_{0:t-1}$ and his information $h_{i,t}$ by $\mathbf{P}_{i,t}(\cdot) := \mathbf{P}_{\sigma^*_{0:t-1}}\big(\cdot \,\big|\, h_{i,t}\big)$.

Since utility in (2) is a strictly concave quadratic function of a_i, the same is true of the expected utility that is maximized in (4) to obtain the best response. We can then rewrite (4) by nulling the derivative of the expected utility with respect to a_i. Performing this operation for σ^*_t in (5), the set of equations in (5) can be rewritten as

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

that need to be satisfied for all possible histories $h_{i,t}$ and agents $i \in V$. We pursue our goal to develop a filter that agents can use to compute their equilibrium actions $a^*_i(t) := \sigma^*_{i,t}(h_{i,t})$ given their observed history $h_{i,t}$ in the following section.

3. PROPAGATION OF PROBABILITY DISTRIBUTIONS

According to the model in (6), at each stage of the game agents use the observed history $h_{i,t}$ to estimate the unknown parameter θ as well as the histories $\{h_{j,t}\}_{j\in V\setminus i}$ observed by other agents. They use the latter and the known BNE strategy $\{\sigma^*_{j,t}(h_{j,t})\}_{j\in V\setminus i}$ to form a belief on the actions $\mathbf{P}_{i,t}(\{a^*_j(t)\}_{j\in V\setminus i})$ which they use to compute their equilibrium action $a^*_j(t)$ at time t. Observe that if the vector of private signals $\mathbf{x}$ is given – not to the agents but to an outside observer – the trajectory of the game is completely determined as there are no random decisions. Thus, agent i can form beliefs on the histories $\{h_{j,t}\}_{j\in V\setminus i}$ and actions $\{a^*_j(t)\}_{j\in V\setminus i}$ of other agents if it keeps a local belief on the vector of private signals $\mathbf{x}$; i.e., $\mathbf{P}_{i,t}(\mathbf{x})$. A method to track this probability distribution is derived in this section using a complete induction argument.

Start by making the assumption that at time t, the posterior distribution $\mathbf{P}_{i,t}(\mathbf{x})$ is normal with mean equal to $\mathbf{E}_{i,t}[\mathbf{x}]$. Define the corresponding error covariance matrix $M^i_{\mathbf{xx}}(t) \in \mathbb{R}^{N\times N}$ as $M^i_{\mathbf{xx}}(t) := \mathbf{E}_{i,t}[(\mathbf{x}-\mathbf{E}_{i,t}[\mathbf{x}])(\mathbf{x}-\mathbf{E}_{i,t}[\mathbf{x}])^T]$. Although agent i's probability distribution for $\mathbf{x}$ is sufficient to describe its belief on the state of the system, subsequent derivations are simpler if we keep an explicit belief on the state of the world θ. Therefore, we also assume that agent i's beliefs on θ and $\mathbf{x}$ are jointly Gaussian given history $h_{i,t}$. The mean of θ is $\mathbf{E}_{i,t}[\theta]$ and the corresponding variance is $M^i_{\theta\theta}(t) := \mathbf{E}_{i,t}[(\theta - \mathbf{E}_{i,t}[\theta])(\theta - \mathbf{E}_{i,t}[\theta])^T]$. The cross covariance $M^i_{\theta\mathbf{x}}(t) \in \mathbb{R}^{1\times N}$ between the world state θ and the private signals $\mathbf{x}$ is $M^i_{\theta\mathbf{x}}(t) := \mathbf{E}_{i,t}[(\theta - \mathbf{E}_{i,t}[\theta])(\mathbf{x} - \mathbf{E}_{i,t}[\mathbf{x}])^T]$. We further make the stronger assumption that the means of this joint Gaussian distribution can be written as linear combinations of the private signals. In particular, we assume that for some known matrix $L_{i,t} \in \mathbb{R}^{N\times N}$ and vector $\mathbf{k}_{i,t} \in \mathbb{R}^{N\times 1}$ we can write

$$\mathbf{E}_{i,t}[\mathbf{x}] = L_{i,t}\mathbf{x}, \qquad \mathbf{E}_{i,t}[\theta] = \mathbf{k}^T_{i,t}\mathbf{x}. \quad (7)$$

Observe that the assumption in (7) is not that the estimates $\mathbf{E}_{i,t}[\mathbf{x}]$ and $\mathbf{E}_{i,t}[\theta]$ are computed as linear combinations of the private signals $\mathbf{x}$ – indeed, $\mathbf{x}$ is not known by agent i in general. The assumption is that from the perspective of an external observer the actual computations that agents do are equivalent to the linear transformations in (7).

Under the complete induction hypothesis of Gaussian posterior beliefs at time t with expectations as in (7) we show that linear equilibrium strategies of the form

$$\sigma^*_{i,t}(h_{i,t}) = \mathbf{v}^T_{i,t}\mathbf{E}_{i,t}[\mathbf{x}], \quad (8)$$

for some action coefficients $\mathbf{v}_{i,t} \in \mathbb{R}^{N\times 1}$ that vary across agents but are independent of the observed history $h_{i,t}$ can be found by solving a system of linear equations. We do this in the following lemma [1].

Lemma 1 *Consider a Bayesian game with quadratic utility as in* (2)*. Suppose that for all agents i, the joint posterior beliefs on the state of the world θ and the private signals $\mathbf{x}$ given the local history $h_{i,t}$ at time t, $\mathbf{P}_{i,t}([\theta, \mathbf{x}^T])$, are Gaussian with means expressed as the linear combinations of private signals in* (7) *for some known vectors $\mathbf{k}_{i,t}$ and matrices $L_{i,t}$. Define the aggregate vector $\mathbf{k}_t := [\mathbf{k}_{1,t}^T, \ldots, \mathbf{k}_{N,t}^T]^T \in \mathbb{R}^{N^2\times 1}$ stacking the state estimation weights of all agents and the block matrix $L_t \in \mathbb{R}^{N^2\times N^2}$ with $N\times N$ diagonal blocks $((L_t))_{jj} = L_{j,t}^T$ and off diagonal blocks $((L_t))_{ij} = -\beta_{ij} L_{i,t}^T L_{j,t}^T$. If there exists a linear equilibrium strategy as in* (8) *the action coefficients $\mathbf{v}_t := [\mathbf{v}_{1,t}^T, \ldots, \mathbf{v}_{N,t}^T]^T \in \mathbb{R}^{N^2}$ can be obtained by solving the system of linear equations*

$$L_t \mathbf{v}_t = \delta \mathbf{k}_t. \tag{9}$$

Lemma 1 provides a mechanism to determine the strategy profiles $\sigma_{i,t}^*(h_{i,t})$ of all agents through the computation of the action vectors $\mathbf{v}_{i,t}$ as a block of the vector $\mathbf{v}_t$ that solves (9). We emphasize that the value of the weight vector $\mathbf{v}_t$ in (9) does not depend on the realization of private signals $\mathbf{x}$. This is in accord with the postulated equilibrium strategy in (8) which assumes the action weights $\mathbf{v}_{i,t}$ are independent of the observed history. A consequence of this fact is that the action coefficients $\{\mathbf{v}_{i,t}\}_{i\in V}$ of all agents can be determined locally by all peers as long as the matrices $L_{i,t}$ and vectors $\mathbf{v}_{i,t}$ are common knowledge. The equilibrium actions $a_i^*(t)$, however, do depend on the observed history because the equilibrium action $a_i^*(t) = \sigma_{i,t}^*(h_{i,t})$ – see Section 4.

At time t agent i computes its action vector $\mathbf{v}_{i,t}$ which it uses to select the equilibrium action $a_i^*(t) = \mathbf{v}_{i,t}^T \mathbf{E}_{i,t}[\mathbf{x}]$ as per (8). Since we have also hypothesized that $\mathbf{E}_{i,t}[\mathbf{x}] = L_{i,t}\mathbf{x}$ as per (7) the action of agent i at time t is given by

$$a_i(t) = \mathbf{v}_{i,t}^T L_{i,t} \mathbf{x}. \tag{10}$$

We emphasize that as in (7) the expression in (10) is not the computation made by agent i but an equivalent computation from the perspective of an external omniscient observer.

The actions $\mathbf{a}_{n(i)}(t) := [a_{j_{i,1}}(t), \ldots, a_{j_{i,d(i)}}(t)]^T \in \mathbb{R}^{d(i)\times 1}$ of neighboring agents $j \in n(i)$ become part of the observed history $h_{i,t+1}$ of agent i at time $t+1$ [cf. (3)]. The important consequence of (10) is that these observations are a linear combination of private signals $\mathbf{x}$. In particular, by defining the matrix $H_{i,t}^T := [\mathbf{v}_{j_{i,1},t}^T L_{j_{i,1},t}; \ldots; \mathbf{v}_{j_{i,d(i)},t}^T L_{j_{i,d(i)},t}] \in \mathbb{R}^{d(i)\times N}$ we can write

$$\mathbf{a}_{n(i)}(t) = H_{i,t}^T \mathbf{x} \tag{11}$$

Agent i's belief of $\mathbf{x}$ at time t is normally distributed and when we go from time t to time $t+1$ agent i observes a linear combination, $\mathbf{a}_{n(i)}(t) = H_{i,t}^T \mathbf{x}$, of private signals. Thus, the propagation of the probability distribution when the history $h_{i,t+1}$ incorporates the actions $\mathbf{a}_{n(i)}(t)$ is a simple sequential LMMSE estimation problem [21, Ch. 12]. In particular, the joint posterior distribution of $\mathbf{x}$ and θ given $h_{i,t+1}$ remains Gaussian and the expectations $\mathbf{E}_{i,t+1}[\mathbf{x}]$ and $\mathbf{E}_{i,t+1}[\theta]$ remain linear combinations of private signals $\mathbf{x}$ as in (7) for some matrix $L_{i,t+1}$ and vector $\mathbf{k}_{i,t+1}$ which we compute explicitly in the following lemma.

Lemma 2 *Consider a Bayesian game with quadratic utility as in* (2) *and the same assumptions and definitions of Lemma 1. Further define the observation matrix $H_{i,t}^T \in \mathbb{R}^{d(i)\times N}$ as in* (11) *and the LMMSE gains*

$$K_{\mathbf{x}}^i(t) := M_{\mathbf{xx}}^i(t) H_{i,t} \big(H_{i,t}^T M_{\mathbf{xx}}^i(t) H_{i,t}\big)^{-1}, \tag{12}$$

$$K_{\theta}^i(t) := M_{\theta\mathbf{x}}^i(t) H_{i,t} \big(H_{i,t}^T M_{\mathbf{xx}}^i(t) H_{i,t}\big)^{-1}, \tag{13}$$

[1] Proofs of results in this paper are available in [20]

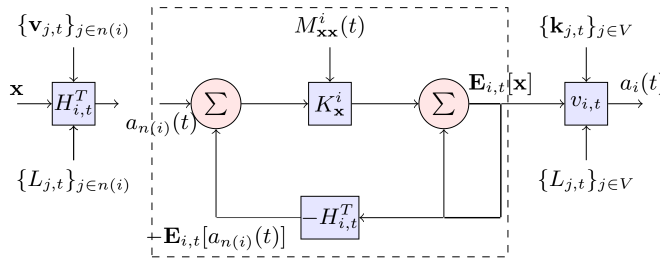

Fig. 1. Linear network game filter at agent i.

and assume that agents play the linear equilibrium strategy in (8)*. Then, $\mathbf{P}_{i,t+1}([\theta, \mathbf{x}^T])$ is Gaussian with means that can be expressed as the linear combination of private signals $\mathbf{E}_{i,t+1}[\mathbf{x}] = L_{i,t+1}\mathbf{x}$, and $\mathbf{E}_{i,t+1}[\theta] = \mathbf{k}_{i,t+1}^T \mathbf{x}$, where the matrix $L_{i,t+1}$ and vector $\mathbf{k}_{i,t+1}$ are given by*

$$L_{i,t+1} = L_{i,t} + K_{\mathbf{x}}^i(t)\Big(H_{i,t}^T - H_{i,t}^T L_{i,t}\Big), \tag{14}$$

$$\mathbf{k}_{i,t+1}^T = \mathbf{k}_{i,t}^T + K_{\theta}^i(t)\Big(H_{i,t}^T - H_{i,t}^T L_{i,t}\Big). \tag{15}$$

The posterior covariance matrix $M_{\mathbf{xx}}^i(t+1)$ for the private signals $\mathbf{x}$ the variance $M_{\theta\theta}^i(t+1)$ of the state θ and the cross covariance $M_{\theta\mathbf{x}}^i(t+1)$ are further given by

$$M_{\mathbf{xx}}^i(t+1) = M_{\mathbf{xx}}^i(t) - K_{\mathbf{x}}^i(t) H_{i,t}^T M_{\mathbf{xx}}^i(t), \tag{16}$$

$$M_{\theta\theta}^i(t+1) = M_{\theta\theta}^i(t) - K_{\theta}^i(t)^T H_{i,t}^T M_{\mathbf{x}\theta}^i(t), \tag{17}$$

$$M_{\theta\mathbf{x}}^i(t+1) = M_{\theta\mathbf{x}}^i(t) - K_{\theta}^i(t) H_{i,t}^T M_{\mathbf{xx}}^i(t). \tag{18}$$

Under the inductive hypotheses of Gaussian beliefs and linear estimates as per (7), lemmas 1 and 2 show how agents determine optimal actions given available information and propagate posterior mean and variance of the beliefs. This permits closing the inductive loop to establish the following theorem for recursive computation of BNE of repeated games with quadratic objectives (we use $\mathbf{e}_i$ to denote the ith element of the standard orthonormal basis of $\mathbb{R}^N$ and $\bar{\mathbf{e}}_i := \mathbf{1} - \mathbf{e}_i$ to write an all-one vector with the ith component nulled).

Theorem 1 *Consider a repeated Bayesian game with the quadratic utility function in* (2) *and assume that linear strategies $\sigma_{i,t}^*(h_{i,t}) = \mathbf{v}_{i,t}^T \mathbf{E}_{i,t}[\mathbf{x}]$ as in* (8) *exist for all times t. Then, the action coefficients $\mathbf{v}_{i,t}$ can be computed by solving the system of linear equations in* (9) *with $\mathbf{v}_t$, $\mathbf{k}_t$ and L_t as defined in Lemma 1. The matrices $L_{i,t}$ and the vectors $\mathbf{k}_{i,t}$ are computed by recursive application of* (12)-(13) *and* (14)-(18) *with initial values $L_{i,0} = \mathbf{1}\mathbf{e}_i^T$ and $\mathbf{k}_{i,0} = \mathbf{e}_i$, and initial covariance matrix $M_{\mathbf{xx}}^i(0) = \text{diag}(\bar{\mathbf{e}}_i)\text{diag}(\mathbf{C}) + \bar{\mathbf{e}}_i\bar{\mathbf{e}}_i^T C_i$, initial variance $M_{\theta\theta}^i(0) = C_i$, and initial cross covariance $M_{\theta\mathbf{x}}^i(0) = C_i\bar{\mathbf{e}}_i^T$.*

Theorem 1 shows that the estimates of θ and $\mathbf{x}$ remain Gaussian for all agents when agents play according to a linear equilibrium strategy as in (8) at each stage. This assumption is true when there exists a solution to the set of linear equations in (9) at all stages. Since the equations in (9) are altered at each stage, the assumption that there exists a linear equilibrium strategy at all stages seems stringent. However, notice that the stage game at time $t+1$ is exactly the same game as the stage game at time t with identical payoff functions and information structure given by the belief updates in Lemma 2. Hence, if the stage game at time $t=0$ has a linear equilibrium strategy, then it must be true that there exists a linear equilibrium in the following stages – see [20] for a formal discussion on existence and uniqueness of linear equilibrium strategies.

Note that the estimation weights $L_{i,t}$ and $\mathbf{k}_{i,t}$ cannot be used to calculate the mean estimates provided by Theorem 1, unless the private

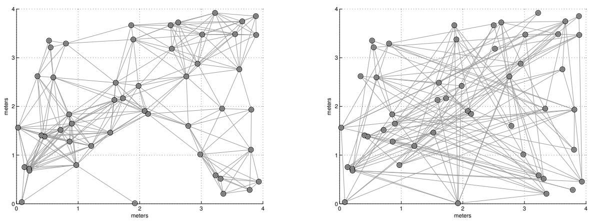

Fig. 2. Geometric (left) and random (right) networks with $N = 50$ agents. Agents are randomly place on a 4 meter $\times$ 4 meter square. There exists an edge between any pair of agents with distance less than 1 meter apart in the geometric network. In the random network, the connection probability between any pair of agents is independent and equal to 0.1.

signals $\mathbf{x}$ are exactly known which will absolve agent i from responsibility of the estimation process entirely. We cannot use (10) to compute the equilibrium actions for similar reasons. We summarize the actual operations carried by each agent in fig. 1 and in the following section.

4. QUADRATIC NETWORK GAME FILTER

In order to compute and play BNE strategies each node runs a quadratic network game (QNG) filter. This filter entails a full network simulation in which agent i maintains beliefs on the state of the world and the private signals of all other agents, $\mathbf{P}_{i,t}([\theta, \mathbf{x}^T])$. These joint beliefs allow agent i to form an implicit belief on all other actions $a_{j,t}$ for all $j \in V$ which he uses to find its equilibrium action $a_i(t)$.

The QNG filter at node i implements the sequential LMMSE estimator updates to keep track of local mean estimates of $\mathbf{x}$ and θ, respectively. These updates are illustrated in Fig. 1 inside the dashed box. At time t, the input to the filter is the observed actions $\mathbf{a}_{n(i)}(t)$ of agent i's neighbors. The prediction $\mathbf{E}_{i,t}[\mathbf{a}_{n(i)}(t)] = H_{i,t}\mathbf{E}_{i,t}[\mathbf{x}]$ of this vector is subtracted from the observed value and the resultant error is fed into the block tasked with updating the belief on the private signals $\mathbf{x}$. This error is multiplied by the gain $K^i_{\mathbf{x}}(t)$ and the resultant innovation is added to the previous mean estimate to correct the estimate of $\mathbf{x}$,

$$\mathbf{E}_{i,t+1}[\mathbf{x}] = \mathbf{E}_{i,t}[\mathbf{x}] + K^i_{\mathbf{x}}(t)\big(\mathbf{a}_{n(i)}(t) - H_{i,t}\mathbf{E}_{i,t}[\mathbf{x}]\big). \tag{19}$$

A similar correction is done on the estimate of θ by using the innovation obtained by multiplying the error by the gain $K^i_{\theta}(t)$,

$$\mathbf{E}_{i,t+1}[\theta] = \mathbf{E}_{i,t}[\theta] + K^i_{\theta}(t)\big(\mathbf{a}_{n(i)}(t) - H_{i,t}\mathbf{E}_{i,t}[\mathbf{x}]\big). \tag{20}$$

The mean updates in (19)-(20) and covariance updates (16)-(18) are admissible since they depend on previous mean estimates, observed actions and LMMSE gains.

In order to determine the equilibrium play, agent i multiples her private signal estimate $\mathbf{E}_{i,t}[\mathbf{x}]$ by the vector $\mathbf{v}_i(t)$ obtained by solving the system of linear equations in (9). In order to form the matrix L_t in (9), agent i needs to compute estimation weights $\{L_{j,t}, \mathbf{k}_{j,t}\}_{j\in V}$. Note that agent i does not need to access observations of other agents in order to calculate the updates for estimation weights using (14)-(15). It only needs to know the previous weight vector and the observation matrix. Consequently, the recursions for estimation weights (14)-(15) are useful for agent i to keep track of how other agents are calculating $\{L_{j,t}, \mathbf{k}_{j,t}\}_{j\in V}$ without making the observations other agents are making. This information is used in solving the set of equations given by (9). Using this solution, agent i can compute the action weights in (8) which she can use to compute her action, the observation matrix (11) and the LMMSE gains.

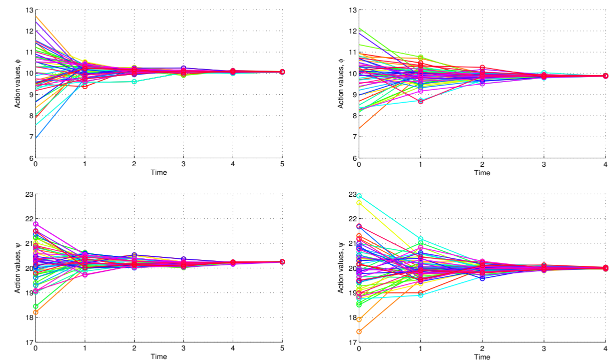

Fig. 3. Agents' actions over time for heading angle ϕ_i (top) and take-off angle ψ_i in geometric (left) and random (right) networks respectively.

5. COORDINATION GAME

Consider a network of mobile agents that want to align themselves so that they move towards a goal on three dimensional space following a straight path. While global coordination is required, communication is possible with neighboring agents only. The direction of movement is characterized by the heading angle on the $x-y$ plane $\phi \in [0°, 180°]$ and the take-off angle on the $x-z$ plane $\psi \in [0°, 180°]$. We denote the correct movement direction by $\boldsymbol{\theta} = [\phi, \psi]^T$. Agents also have the goal of maintaining the starting formation while moving at equal speed by coordinating their angle of movement with other agents. In this context, agent i's decision $\mathbf{a}_i \in [0°, 180°] \times [0°, 180°]$ represents the heading and take-off angles in the direction of movement. It is possible to formulate the objective of agent i as maximization of the payoff

$$u_i(\mathbf{a}_i, \mathbf{a}_{-i}, \boldsymbol{\theta}) = -\frac{1-\lambda}{2}(\mathbf{a}_i - \boldsymbol{\theta})^T(\mathbf{a}_i - \boldsymbol{\theta}) - \frac{\lambda}{2(N-1)}\sum_{j\in V\setminus\{i\}}(\mathbf{a}_i - \mathbf{a}_j)^T(\mathbf{a}_i - \mathbf{a}_j), \tag{21}$$

where $\lambda \in (0, 1)$ is a constant measuring the relative importance of estimation and coordination. The first term in (21) is the estimation error in the true heading and take-off angles. The second term is the coordination component that measures the discrepancy between the direction of movement and headings of other agents.

We set the correct movement angle to $\boldsymbol{\theta} = [10°, 20°]^T$ and let agents make private observations on ϕ and ψ, given by $\mathbf{x}_i = \boldsymbol{\theta} + \boldsymbol{\epsilon}_i$, where $\boldsymbol{\epsilon}_i$ is jointly Gaussian with mean zero and identity covariance matrix. Hence, signal for ϕ, $\mathbf{x}_i[1]$, is independent from signal for ψ, $\mathbf{x}_i[2]$, for all $i \in V$. Further, $\boldsymbol{\epsilon}_i$ is independent across agents. This simplifies the process such that agents are running two QNG filters in parallel, one for the heading angle and the other for the take-off angle. The QNG filter can be generalized to handle correlated vector states – see [20].

We let $\lambda = 0.5$ and evaluate convergence behavior in geometric and random networks with $N = 50$ agents; see Fig. 2. The geometric network has a diameter of $\Delta_g = 5$ where the random network has a diameter of $\Delta_r = 4$. The action values of each agent are depicted in Fig. 3. The top (bottom) row shows the heading angle ϕ_i (take-off angle ψ_i) in geometric and random networks on the left and right respectively. The results show that agents' actions $\mathbf{a}_i$ converge to the best estimates in heading and take-off angles denoted by $\hat{\phi}^* = \mathbf{E}[\phi \,|\, \mathbf{x}[1]]$ and $\hat{\psi}^* = \mathbf{E}[\phi \,|\, \mathbf{x}[2]]$, respectively. Agreement in actions is reasonable since agents have the incentive to agree with others in the movement direction in order to maintain the initial formation—see (21). Convergence occurs in a number of iterations in the order of the network diameter.

6. REFERENCES

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

[2] V. Bala and S. Goyal, "Learning from neighbours," *Review of Economic Studies*, vol. 65, no. 3, 1998.

[3] 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.

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

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

[6] A. Banerjee and D. Fudenberg, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[7] J.J. Xiao, A. Ribeiro, L. Zhi-Quan, and G.B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, pp. 27–41, July, 2006.

[8] J. Chen and A.H. Sayed, "Diffusion adaptation strategies for distributed optimization and learning over networks," *Signal Processing, IEEE Transactions on*, vol. 60, no. 8, pp. 4289–4305, 2012.

[9] K.R. Rad and A. Tahbaz-Salehi, "Distributed parameter estimation in networks," in *Proc. of the 49th IEEE Conference on Decision and Control*, Atlanta, GA, USA, Dec. 2010.

[10] 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.*, vol. 5, no. 4, pp. 674–690, 2011.

[11] 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.*, vol. 56, no. 1, pp. 1650–1666, January 2008.

[12] 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)*, New York, NY, USA., June 2005, pp. 1088–1092.

[13] D. Gale and S. Kariv, "Bayesian learning in social networks," *Games Econ. Behav.*, vol. 45, pp. 329–346, 2003.

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

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

[16] 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.

[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.

[18] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie, "Games with side information in economic networks," in *Proceedings of the 46th Asilomar Conference on Signals, Systems and Computers (to appear)*, Pacific Grove, California, USA., 2012.

[19] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie, "Learning in networks with incomplete information," *IEEE Signal Process. Mag.*, vol. (to appear), 2012.

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

[21] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, Englewood Cliffs, New Jersey, 1. edition, 1993.