Learning in Network Games with Incomplete Information

Ceyhun Eksin, Pooya Molavi, Alejandro Ribeiro, and Ali Jadbabaie
Department of Electrical and Systems Engineering, University of Pennsylvania,
200 South 33rd Street, Philadelphia, PA 19104.
Email:{ceksin, pooya, aribeiro, jadbabai}@seas.upenn.edu.

The role of social networks in learning and opinion formation has been demonstrated in a variety of scenarios such as the dynamics of technology adoption [1], consumption behavior [2], organizational behavior [3] and financial markets [4]. Emergence of network-wide social phenomena from local interactions between connected agents have been studied using field data [5]–[7] as well as lab experiments [8], [9]. Interest in opinion dynamics over networks is further amplified by the continuous growth in the amount of time that individuals spend on social media websites and the consequent increase in the importance of networked phenomena in social and economic outcomes. As quantitative data become more readily available, a research problem is to identify metrics that could characterize emergent phenomena such as conformism or diversity in individuals' preferences for consumer products or political ideologies [10]. With these metrics available, a natural follow up research goal is the study of mechanisms that lead to diversity or conformism and the role of network properties like neighborhood structures on these outcomes. All of these questions motivate the development of theoretical models of opinion formation through local interactions in different scenarios.

The canonical model of learning in networks considers a set of connected agents each endowed with private information regarding a *common* underlying random state. Each agent uses his private information to form a probability distribution on the state of the world and selects an action from an allowable set that is optimal with respect to this belief. The definition of optimality with respect to the belief varies but a general model is to postulate the existence of a utility function that depends on the selected action, the state of the world, and possibly on the actions selected by other members of the network. If the state of the world were known, we would say that we have complete information and select the action that maximizes the utility. However, information is incomplete because only a belief on the state is available. Therefore, agents proceed to select actions that maximize the expectation of their utilities with respect to their beliefs. In a networked setting agents further observe actions taken by agents in their connectivity neighborhoods. These observations contain information on the state that a rational agent would feel compelled to incorporate into his belief leading to the selection of a different optimal action. This phenomenon of observations of neighboring actions affecting decision-making of agents is called *information externality*. In

Work in this paper is supported by ARO P-57920-NS, NSF CAREER CCF-0952867, NSF CCF-1017454, and AFOSR MURI FA9550-10-1-0567.

general, at a given point in time, any given agent has seen a history of neighboring actions that he combines with his private information to update the probability distribution on the state of the world. This belief determines an action that is optimal with respect to the expected utility. As time progresses, agents learn the state of the world ω in the sense that they refine their knowledge – i.e., the mass of the belief becomes more concentrated – through the observation of neighboring actions. The focus is often on the characterization of asymptotic properties of the belief and the actions of agents as well as algorithmic considerations.

When utilities of agents depend not only on the unknown state of the world but also on the unknown actions of other agents, we say that there are *payoff externalities*. In most social learning scenarios payoff externalities and information externalities coexist in that the action chosen by agent i is determined by both an informational component pertaining to agents' beliefs about the underlying state of the world and a payoff externality corresponding to their beliefs about the actions taken by all the other agents in the network. In stock markets, for instance, the actions of each individual affect the utility of all the other agents and agents respond strategically to actions based on their beliefs. At the same time these actions contain information about stocks' intrinsic valuations that market participants are intent on learning. By selecting certain actions agents are revealing, perhaps unwillingly, pieces of private information about the true value of the stock [11], [12]. The focus is the study of asymptotic behavior of agents' actions and their beliefs given a fixed network that determines the flow of information; see Section II.

There exists an extensive literature on Bayesian learning over networks for scenarios without payoff externalities [13]–[16]. One may think of this problem as a variant of distributed estimation since agents intend to compute an estimate based on global information by aggregating local information and successively refining their estimates using those of their neighbors. Linear and nonlinear estimation problems are well-studied in the signal processing literature; see e.g., [17]. The main difference between distributed estimation problems and the ones considered here is that in the former network nodes may exchange observations, estimates, and even some auxiliary variables [18]–[26]. In the problems considered here, on the other hand, agents try to infer the state of the world by observing actions of neighboring nodes. The former is a suitable model for algorithm and protocol design, but the latter is a more appropriate model of social and economic interactions. Besides signal processing, models with purely informational externalities have been studied in economics [14], [15], [27], [28], computer science [29], statistics [30], and control theory [31]–[34].

Even though Bayesian learning stands as the normative behaviorial model for agents in social networks, 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 his neighbors but also that of the neighbors of his neighbors and so on. Because of such computational intractability little is known about Bayesian learning besides the asymptotic behavior. However, under some structural assumptions on distribution of information [30] or the network structure [29], Bayesian learning is shown to be tractable in absence of payoff externalities. In this paper we present a tractable algorithm for the case when agents also face payoff externalities, assuming that agents' initial private signals are normally distributed; see Section III. We use the algorithm to numerically study the effect of the network structure on convergence time; see Section IV.

I. Bayesian learning in networks

The network learning models considered in this paper comprise of an unknown state of the world $\omega \in \Omega$ and a group of agents $\mathcal{N} = \{1, \ldots, N\}$ whose interactions are characterized by a network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. At subsequent points in time $t = 0, 1, 2, \ldots$, agents in the network observe private signals $s_{i,t}$ that carry information about the state of the world ω and decide on an action $a_{i,t}$ belonging to some common compact metric action space A that they deem optimal with respect to a utility function of the form

$$u_i\big(\omega, a_{i,t}, \{a_{j,t}\}_{j \in \mathcal{N} \setminus i}\big). \tag{1}$$

Besides his action $a_{i,t}$, the utility of agent i depends on the state of the world ω and the actions $\{a_{j,t}\}_{j \in \mathcal{N} \setminus i}$ of *all* other agents in the network. This dependence tries to capture tradeoffs that arise in social and economic networks. For example, the state of the world ω may represent the inherent value of a service, the private signals $s_{i,t}$ quality perceptions after use, and $a_{i,t}$ decisions on how much to use the service. The utility of a person derives from the use of the service depending not only on the inherent quality ω but also on how much other people use the service.

Deciding optimal actions $a_{i,t}$ would be easy if all agents were able to coordinate their actions. All private signals $s_{i,t}$ could be combined to form a single probability distribution on the state of the world ω and that common belief used to select $a_{i,t}$. Agents could act together and combine their utilities into a social objective or they could exhibit strategic behavior and select game equilibrium actions. Whether there is payoff externality or not, global coordination is an implausible model of social and economic behavior. We therefore consider agents that act independently of each other and couple their behavior through observation of the action history of agents in their network neighborhood $\mathcal{N}_i$.

To be more precise say that at time $t = 0$, there is a common initial belief among agents about the unknown parameter ω. This common belief is represented by a probability distribution P. At time $t = 0$, each agent observes his own private signal $s_{i,0}$ which he uses in conjunction with the prior belief P to choose and execute action $a_{i,0}$. Upon execution of $a_{i,0}$ actions $\{a_{j,0}\}_{j \in \mathcal{N}_i}$ of neighboring agents become known to node i. Knowing the actions of its neighbors provides agent i with information about the neighboring private signals $\{s_{j,0}\}_{j \in \mathcal{N}_i}$, which in turn refines his belief about the state of the world ω. This new knowledge prompts a re-evaluation of the optimal action $a_{i,1}$ in the subsequent time slot. In general, at stage t, agent i has acquired knowledge in the form of the history

$h_{i,t}$ of past and present private signals $s_{i,\tau}$ for $\tau = 0, \ldots, t$ and past actions of neighboring agents $\{a_{j,\tau}\}_{j\in\mathcal{N}_i}$ for times $\tau = 1, \ldots, t-1$. This history is used to determine the action $a_{i,t}$ for the current slot. In going from stage t to stage $t+1$, neighboring actions $\{a_{j,t}\}_{j\in\mathcal{N}_i}$ become known and incorporated into the history of past observations. We can thus formally define the history $h_{i,t}$ by the recursion

$$h_{i,t+1} = \big(h_{i,t}, \{a_{j,t}\}_{j\in\mathcal{N}_i}, s_{i,t+1}\big). \tag{2}$$

The component of the game that determines action of agent i from observed history $h_{i,t}$ is his strategy $\sigma_{i,t}$. A pure strategy is a function that maps any possible history to an action, $\sigma_{i,t} : h_{i,t} \mapsto a_{i,t}$. The value of a strategy function $\sigma_{i,t}$ associated with the given observed history $h_{i,t}$ is the action of agent i, $a_{i,t}$. Given his strategy $\sigma_i := \{\sigma_{i,u}\}_{u=0,\ldots,\infty}$, agent i knows exactly what action to take at any stage upon observing the history at that stage. Hence, the (pure) strategies of all the agents across time $\sigma := \{\sigma_{j,u}\}_{j\in\mathcal{N},u=0,\ldots,t}$, namely, the strategy profile determines the path of play, that is, the sequence of histories each agent will observe. As a result, if agent i at time t knows the information set at time t, i.e., $h_t = \{h_{1,t}, \ldots, h_{N,t}\}$, then he knows the continuation of the game from time t onwards given knowledge of the strategy profile σ.

When agents have (common) prior P on the state of the world at time $t = 0$, the strategy profile σ induces a belief $P_\sigma(\cdot)$ on the path of play. That is, $P_\sigma(h)$ is the probability associated with reaching an information set h when agents follow the actions prescribed by σ. Therefore, at time t, the strategy profile determines the prior belief $q_{i,t}$ of agent i given $h_{i,t}$, that is,

$$q_{i,t}(\cdot) = P_\sigma(\cdot|h_{i,t}). \tag{3}$$

The prior belief $q_{i,t}$ puts a distribution on the set of possible information sets h_t at time t given that agents played according to $\sigma_{0,\ldots,t-1}$ and i observed $h_{i,t}$. Furthermore, the strategies from time t onwards $\sigma_{t,\ldots,\infty}$ permit the transformation of beliefs on the information set into a distribution over respective upcoming actions $\{a_{j,u}\}_{j\in\mathcal{N},u=t,\ldots,\infty}$. As a result, upon observing $\{a_{j,t}\}_{j\in\mathcal{N}_i}$ and $s_{i,t}$, agent i updates his belief using Bayes' rule,

$$q_{i,t+1}(\cdot) \;=\; P_\sigma\big(\cdot \,\big|\, h_{i,t+1}\big) \;=\; P_\sigma\big(\cdot \,\big|\, h_{i,t}, s_{i,t+1}, \{a_{j,t}\}_{j\in\mathcal{N}_i}\big) \;=\; q_{i,t}\big(\cdot \,\big|\, s_{i,t+1}, \{a_{j,t}\}_{j\in\mathcal{N}_i}\big). \tag{4}$$

Since the belief is a probability distribution over the set of possible actions in the future, agent i can calculate expected payoffs from choosing an action. A rational behavior for agent i is to select the action $a_{i,t}$ that maximizes the expected utility given his belief $q_{i,t}$,

$$a_{i,t} \in \operatorname*{argmax}_{\alpha_i\in A} \mathbb{E}_\sigma\left[u_i\big(\omega, \alpha_i, \{\sigma_{j,t}(h_{j,t})\}_{j\in\mathcal{N}\setminus i}\big) \,\big|\, h_{i,t}\right] := \operatorname*{argmax}_{\alpha_i\in A} \int_h u_i\big(\omega, \alpha_i, \{\sigma_{j,t}(h_{j,t})\}_{j\in\mathcal{N}\setminus i}\big) q_{i,t}(h) \tag{5}$$

where we have defined conditional expectation operator $\mathbb{E}_\sigma[\cdot \,|\, h_{i,t}]$ with respect to the conditional distribution

$P_\sigma(\cdot \,|\, h_{i,t})$. The rational action $a_{i,t}$ in (5) is optimal given strategy profile σ; as a result, $a_{i,t}$ is a function of the strategy profile σ.

So far we have not imposed any constraints on the strategy profile σ. According to the definition of rational behavior in (5), all agents should maximize the expected value of self utility function. With this in mind we define the Bayesian Nash equilibrium (henceforth, BNE) to be the strategy profile of a rational agent. A BNE strategy σ^* is a best response strategy such that no agent can expect to increase his utility by unilaterally deviating from his strategy $\sigma^*_{i,t}$ given that the rest of the agents play equilibrium strategies $\{\sigma^*_{j,t}\}_{j\in\mathcal{N}\setminus i}$; that is, σ^* is BNE if for each $i \in \mathcal{N}$ and $t = 0, 1, 2, \ldots$, the strategy $\sigma^*_{i,t}$ maximizes the expected payoff:

$$\sigma^*_{i,t}(h_{i,t}) \in \operatorname*{argmax}_{\alpha_i \in A} \mathbb{E}_{\sigma^*}\left[u_i(\omega, \alpha_i, \{\sigma^*_{j,t}(h_{j,t})\}_{j\in\mathcal{N}\setminus i}) \,\middle|\, h_{i,t}\right]. \tag{6}$$

We emphasize that (6) needs to be satisfied for all possible histories $h_{i,t}$ and not just for the history realized in a particular game realization. This is necessary because agent i does not know the history observed by agent j but rather has a probability distribution on histories. Thus, to evaluate the expectation in (5) agent i needs a representation of the equilibrium strategy for all possible histories $h_{j,t}$.

In this paper we restrict our attention to the equilibrium notion where agents choose myopically optimal actions as in (5). It is also possible to define BNE for non-myopic agents that discount future payoffs. Agents exhibiting non-myopic behavior might *experiment* to obtain valuable information to be used in the future. In rest of the paper we consider myopic agents playing with respect to BNE strategy σ^*_i. To simplify future notation, we define the expectation operator

$$\mathbb{E}_{i,t}\big[\ \cdot\ \big] := \ \mathbb{E}_{\sigma^*}\big[\ \cdot\ \mid h_{i,t}\big], \tag{7}$$

to represent expectations with respect to the local history $h_{i,t}$ when agents play according to the equilibrium strategy σ^*.

BNE is an extension of Nash equilibrium to games with incomplete information. In this solution concept we assume that agents interpret actions of their neighbors knowing that they play according to the BNE strategy, i.e. BNE is common knowledge. Note that while defining rational behavior in (5), we have not specified how agent i models actions of other agents. In order to calculate his expected utility in (5), agent i needs to have a model of strategies of other agents. Common knowledge of BNE strategies and rationality is a particular model of agents' behavior in which agent i believes, correctly so, that agent j is rational. In other words, agent i's model of behavior of agent $j \in \mathcal{N} \setminus \{i\}$ is that j also maximizes expected payoff as in (6) and further that agent i can correctly guess j's actions if he had access to j's history $h_{j,t}$. In a networked setting, agents also require knowledge of the network in order to infer about information of other agents. Hence, we also assume network structure is common knowledge. Notice that this equilibrium notion couples beliefs and strategies in a consistent way in the sense that strategies induce beliefs and the beliefs determine optimal strategy. This rational model provides a benchmark for comparison with other behavioral models.

A. Notions of Consensus

An important research question when studying social learning models is whether agents reach consensus, and if they do, whether the outcome is efficient according to some criterion. Several different notions of consensus have been studied in the literature. Agents are said to reach consensus in their actions if they all eventually take the same action, or more formally, if the distance between $a_{i,t}$ and $a_{j,t}$ goes to zero as time goes to infinity where the distance is defined using the metric on the action space A. If agents' utility functions are the same, then reaching consensus in the actions implies that agents obtain the same utility; however, the converse is not necessarily true. A different characterization of consensus is in terms of agents' beliefs. Agents i and j are said to reach consensus in their beliefs, if the distance (in total variation) between probability measures $q_{i,t}(\cdot)$ and $q_{j,t}(\cdot)$ goes to zero as t goes to infinity. If this is true for any two pair of agents, we say that all agents reach consensus. The consensus belief however might be inaccurate in the sense of not corresponding correctly to the agents' observations. Another notion of convergence considers expected payoffs. We say that agents are expected to perform equally well asymptotically if

$$\lim_{t\to\infty} \mathbb{E}\left[u_i(\omega, a_{i,t}, \{a_{k,t}\}_{k\in\mathcal{N}\setminus i})\right] = \lim_{t\to\infty} \mathbb{E}\left[u_j(\omega, a_{j,t}, \{a_{k,t}\}_{k\in\mathcal{N}\setminus j})\right], \tag{8}$$

where the expectation is over all possible realizations of the state of the world ω. The result in (8) establishes a form of consensus that is attained in the limit. It is possible that for individual realizations of the parameter ω, the expected payoffs of different agents are different; however, if we consider an average across realizations of ω, the payoffs asymptotically coincide. We can interpret this result as stating that *ex ante* all agents are expected to obtain the same payoff.

Each of the notions of agreement discussed above might be relevant in certain applications. Moreover, they do not necessarily coincide. Agents might reach consensus in their actions without having the same beliefs if agents' actions do not completely reflect the beliefs held by them. On the other hand, agents might reach consensus in their beliefs (and even learn all the information) and yet take disparate actions.

In the following sections we consider learning in a class of games in which actions are on the real line and the utility function is quadratic in agents' actions and the state of the world. In Section II we survey a recent result that shows convergence in expected payoffs. This result is essential in proving that agents' reach consensus in their actions in the limit. In Section III we derive tractable recursions for rational learning given that agents' private signals are normally distributed.

II. Learning with payoff externalities

Presence of payoff externalities adds another layer of complexity to the learning process compared to models with purely informational externalities, since it prohibits agents from interpreting the actions of their neighbors as solely revealing information about the true state of the world. Instead, they have to keep track of motives of other agents and at the same time incorporate the new information effectively. The interested readers can refer to the callout for an illustration of the rational learning process with both payoff and informational externalities. An example of learning with only informational externalities is given in [16]. In this section, we introduce games with utility functions that are quadratic both in the state of the world and agents' actions. We then exemplify this quadratic form in the context of financial markets. Finally, we provide asymptotic convergence results for learning in quadratic games over networks.

A. Quadratic Games

At any time t, selection of actions $\{a_i := a_{i,t} \in \mathbb{R}\}_{i\in\mathcal{N}}$ when the state of the world is $\omega \in \mathbb{R}$ results in agent i receiving a payoff,

$$u_i(\omega, a_i, \{a_j\}_{j\in\mathcal{N}\setminus i}) = -\frac{1}{2}\sum_{j\in\mathcal{N}} a_j^2 + \sum_{j\in\mathcal{N}\setminus\{i\}} \beta_{ij} a_i a_j + \ \delta a_i \omega + c\omega^2, \tag{9}$$

where β_{ij}, δ and c are real valued constants. The constant β_{ij} measures the effect of j's action on i's utility. For notational convenience we let $\beta_{ii} = 0$ for all $i \in \mathcal{N}$.

Since u_i is a strictly concave function of a_i (i.e., $\partial^2 u_i/\partial a_i^2 < 0$), the myopically optimal action can be computed explicitly by taking the derivative with respect to a_i, equating the result to zero, and solving for a_i. As a result, the rational action, defined in (5), for agent i in response to any strategy $\{\sigma_{j,t}\}_{j\in\mathcal{N}\setminus i}$ is a linear function of the strategies of other agents and the underlying parameter:

$$a_{i,t} = \sum_{j\in\mathcal{N}\setminus\{i\}} \beta_{ij} \mathbb{E}_{i,t}[\sigma_{j,t}(h_{j,t})] + \delta\, \mathbb{E}_{i,t}[\omega]. \tag{10}$$

According to the equilibrium definition (6) in Section I, at each state agents play a myopic best response given the observed history against other agents' actions, which in turn are myopic best responses. Consequently, for the quadratic utility function an equilibrium strategy profile $\{\sigma^*_{i,t}\}_{i\in\mathcal{N}}$ solves the following set of equations

$$\sigma^*_{i,t}(h_{i,t}) = \sum_{j\in\mathcal{N}\setminus\{i\}} \beta_{ij} \mathbb{E}_{i,t}[\sigma^*_{j,t}(h_{j,t})] + \delta\mathbb{E}_{i,t}[\omega], \tag{11}$$

for all $i \in \mathcal{N}$.

B. Coordination Games

Keynes in his General Theory of economics argues that if a person is asked to guess the beauty contest winner, he should evaluate each contestant with respect to what he thinks other people's criteria of beauty is. Similarly, an investment in the stocks of a company entails a player not only to consider his own estimate of how well the company is doing but also what everyone else thinks about the company's status [11]. Let $\omega \in \mathbb{R}$ be the true stock value of a company. In this context, 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

$$u_i(\omega, a_i, \{a_j\}_{j\in\mathcal{N}\setminus i}) = -\frac{1-\lambda}{2}(a_i-\omega)^2 - \frac{\lambda}{2(N-1)}\sum_{j\in\mathcal{N}\setminus\{i\}}(a_i-a_j)^2, \tag{12}$$

where $\lambda \in (0,1)$. This payoff function is reminiscent of that of a coordination game (or potential game) [35], [36] with the only difference being the addition of a term corresponding to an estimation problem. The first term of the payoff function measures the desire of the player to estimate the true value of the stock as the quadratic distance between i's action and ω. The second term is the coordination (or the beauty contest) term measuring the payoff associated with being close to valuations of other members of the society. It represents how the actions of others affect the payoff of agent i. The constant λ gauges the relative importance of coordination and estimation.

Using (6), the BNE strategy σ^* in this quadratic game solves the following set of equations:

$$\sigma^*_{i,t}(h_{i,t}) = (1-\lambda)\,\mathbb{E}_{i,t}[\omega] + \frac{\lambda}{n-1}\sum_{j\in\mathcal{N}\setminus\{i\}}\mathbb{E}_{i,t}\left[\sigma^*_{j,t}(h_{j,t})\right], \qquad i\in\mathcal{N} \text{ and } t\in\mathbb{N}. \tag{13}$$

Since the payoff (12) is of the form in (9), the equilibrium equations in (10) are linear in strategies of other agents as in (11).

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 [12], [37]. Each agent collects an initial noisy measurement of the true heading angle ω, 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. We assume the agents move with constant and equal velocity. The first term in (12) represents agents' goal to estimate the correct heading angle. The desire of agent i to maintain the initial formation is captured by the second term in (12).

C. Asymptotic Properties of Learning in Quadratic Games

In this section we present results from [12], [38] that focus on *symmetric*, *supermodular* and *diagonally dominant* games.

The game defined by the utility function in (9) is symmetric when the pairwise influences β_{ij} are equal for any pair; that is, $\beta_{ij} = \beta$ for constant $\beta \in \mathbb{R}$ for all $i \in \mathcal{N}$ and $j \in \mathcal{N} \setminus \{i\}$. A game is *supermodular* when agents' strategies are complementary to each other. Strategic complementarity between agents i and j means that the marginal utility of agent i's action increases with an increase in j's action. For a twice differentiable utility function, this is equivalent to requiring that $\partial^2 u_i(\cdot)/\partial a_i \partial a_j > 0$ for any two agents i and j. For our quadratic utility function in (9), the actions of i and j are strategic complements when $\beta_{ij} \geq 0$ for all $i \in \mathcal{N}$ and $j \in \mathcal{N} \setminus i$. We further restrict our attention to games in which the Hessian matrices of the utility function are strictly *diagonally dominant*. For the utility function in (9), this is equivalent to requiring that there exists $\rho < 1$ such that

$$\sum_{j \in \mathcal{N} \setminus \{i\}} \beta_{ij} \leq \rho \quad \text{for all } i \in \mathcal{N}. \tag{14}$$

The interpretation of (14) is that an agent's utility is more sensitive to changes in his own actions than to changes in the actions of other agents. Notice that the payoff function (12) of the coordination problem satisfies all of these properties.

According to the learning framework introduced in Section I, agents take actions specified by the equilibrium strategy, observe neighboring actions, update their beliefs according to the Bayes' rule, and then start the next stage as a new game with beliefs different from the previous stage. This means that the equilibrium of the new game is not necessarily the same as the equilibrium of the previous stage. However, since agents accumulate information about the unknown state over time, it is possible to show that under the equilibrium behavior in (6), agents' expected utilities converge for the utility function in (9) [12]. By the same token agents' equilibrium actions $a^*_{i,t} := \sigma^*_{i,t}(h_{i,t})$ defined in (11) converge in the limit:

$$a^*_{i,t} \to a^*_{i,\infty} \quad \text{a.s.} \quad \text{for all } \; i \in \mathcal{N}. \tag{15}$$

Existence of limit actions implies that the agents can learn their neighbors' limit actions. Since i observes actions of $j \in \mathcal{N}_i$, agent j's action at time $t-1$ is in the information set of agent i at time t; i.e., $a^*_{j,t-1} \in h_{i,t}$. This implies that $a^*_{j,t-1}$ is measurable with respect to the information of i at time t. Therefore, since $a^*_{j,t} \to a^*_{j,\infty}$ with probability one i's information is an increasing set that converges as t goes to infinity, the limit action $a^*_{j,\infty}$ is measurable with respect to i's information at infinity. In other words, agent i is able to identify the limit action of a neighboring agent j.

The fact that agents can identify the limit actions of their neighbors leads to a number of interesting conclusions by making use of the so-called *imitation principle* [14], [15]. The imitation principle states that the expected payoff of agent i with respect to his history cannot increase if he adopts an action of a neighboring agent

$$\mathbb{E}_{i,\infty}\left[u_i(\omega, a^*_{i,\infty}, \{a^*_{k,\infty}\}_{k\in\mathcal{N}\setminus i})\right] \geq \mathbb{E}_{i,\infty}\left[u(\omega, a^*_{j,\infty}, \{a^*_{k,\infty}\}_{k\in\mathcal{N}\setminus i})\right], \quad j \in \mathcal{N}_i. \tag{16}$$

The inequality in (16) is due to the definition of $a^*_{i,t}$ as the maximizing action in (6). Notice that (16) is only true for neighboring nodes because i only observes (and hence can identify) the actions of his neighbors. Actions of other agents, on the other hand, are not observed by i and may not be measurable with respect to $q_{i,\infty}$.

By applying the imitation principle and making use of the assumption on the strategic complementarity of the actions between agents i and $j \in \mathcal{N}_i$, we can show that neighboring agents are expected to receive the same payoff whether i plays his own limit action $a^*_{i,\infty}$ or any of his neighbors' limit actions $a^*_{j,\infty}$ [12]; that is,

$$\mathbb{E}[\mathbb{E}_{i,\infty}\left[u_i(\omega, a^*_{i,\infty}, \{a^*_{k,\infty}\}_{k\in\mathcal{N}\setminus i})\right]] = \mathbb{E}[\mathbb{E}_{i,\infty}\left[u_i(\omega, a^*_{j,\infty}, \{a^*_{k,\infty}\}_{k\in\mathcal{N}\setminus i})\right]], \quad \text{for all } j \in \mathcal{N}_i. \tag{17}$$

The intuitive argument behind (17) is as follows. By the symmetry property of the utility function strategic complementarity implies that unilateral deviations by i and j to each other's actions are at least as good as playing their own limit actions in expectation. However, by the imitation principle in (16), this behavior can never yield strictly higher payoffs. Hence, it must be the case that deviations to neighbors' limit actions result in the expected performance of agents. Note that this is true only for neighboring agents.

By the imitation principle in (16), the left hand side of (17) is no larger than the right hand side for all ω. Hence, it must be case that the equality holds almost surely when we remove the outer expectation in (17):

$$\mathbb{E}_{i,\infty}\left[u_i(\omega, a^*_{i,\infty}, \{a^*_{k,\infty}\}_{k\in\mathcal{N}\setminus i})\right] = \mathbb{E}_{i,\infty}\left[u_i(\omega, a^*_{j,\infty}, \{a^*_{k,\infty}\}_{k\in\mathcal{N}\setminus i})\right], \quad \text{for all } j \in \mathcal{N}_i. \tag{18}$$

According to (18), agent i expects his limit action to result in a payoff no worse than if he were to play the limit action of one of his neighbors; i.e, from the perspective of agent i agent j's limit action is just as good as self limit action.

An immediate corollary of (18) is that for a connected network agents reach consensus in their actions. The result is proved by the following argument. Given (18) the limit action of agent j is a maximizer of the expected utility of agent i; i.e., $a^*_{j,\infty} = \operatorname{argmax}_{\alpha_i \in A} \mathbb{E}_{i,\infty}\left[u(\omega, \alpha_i, \{a^*_{k,\infty}\}_{k\in\mathcal{N}\setminus i})\right]$. By strict concavity of (9) the myopically optimal action in (5) is unique. Hence, it must be the case that $a^*_{i,\infty} = a^*_{j,\infty}$ for all $i \in \mathcal{N}$ and $j \in \mathcal{N}_i$ with probability one. Given that the network is connected this implies that $a^*_{i,\infty} = a^*_{j,\infty}$ for any pair of agents. This conformity result proved in [12] extends some of the results in [15] to models with payoff externalities.

III. Gaussian Quadratic Network Games

In this section we restrict our attention to games with quadratic utility functions as in (9) and private signals that are normally distributed, and show that the equilibrium strategies can be computed explicitly. The results of this section are presented in more formally and in more detail in [37]. Assume that at time $t = 0$, agent i receives a private noisy signal $s_i \in \mathbb{R}$ about the unknown parameter:

$$s_i = \omega + \epsilon_i, \tag{19}$$

where ϵ_i is normal with mean zero and variance C_i. The grouping of all private signals is denoted by the vector of private signals $\mathbf{s} := [s_1, \ldots, s_n]^T \in \mathbb{R}^{N\times 1}$. Further, agents' common prior for ω is an (improper) uniform measure over $\mathbb{R}$. Hence, the posterior at time $t = 0$, $P(\omega, \mathbf{s}^T \mid s_i)$ is normal.

To see how equilibrium responses can be computed explicitly, assume for the sake of argument that at given time t it is possible to write the minimum mean squared error (MMSE) estimates of the state of the world θ and the private signals $\mathbf{s}$ as linear combinations of the private signals themselves; i.e., at time t there are vectors $\mathbf{k}_{i,t}$ and matrices $L_{i,t}$ for which we can write

$$\mathbb{E}_{i,t}\left[\omega\right] = \mathbf{k}_{i,t}^T\mathbf{s}, \quad \mathbb{E}_{i,t}\left[\mathbf{s}\right] = L_{i,t}\mathbf{s}, \quad \text{for all } i \in V. \tag{20}$$

Notice that (20) does not imply that MMSE estimates $\mathbb{E}_{i,t}\left[\omega\right]$ and $\mathbb{E}_{i,t}\left[\mathbf{s}\right]$ are computed as linear combinations of the private signals $\mathbf{s}$. This is not possible because agent i does not know the values of all private signals—if this were the case, there would be no game to be played and the expression $\mathbb{E}_{i,t}\left[\mathbf{s}\right] = L_{i,t}\mathbf{s}$ would be pointless. Our assumption is that whatever may be the computations that agents perform, they are equivalent to computing the linear combinations in (20).

The validity of (20)—which we assume for the moment without a proof—is instrumental in simplifying the computation of best responses and the associated fixed points that define the equilibrium actions. For that matter we solve (6) by postulating that the best response action can be written as a linear combination $a_{i,t}^* = \mathbf{v}_{i,t}^T\mathbb{E}_{i,t}[\mathbf{s}]$ of the private signals' MMSE estimates with weights given by some vector $\mathbf{v}_{i,t}$ to be determined. Given this candidate solution, we can rewrite the best response fixed point condition in (6) as

$$\mathbf{v}_{i,t}^T\mathbb{E}_{i,t}[\mathbf{s}] = \sum_{j\in V\setminus\{i\}} \beta_{ij}\mathbb{E}_{i,t}\Big[\mathbf{v}_{j,t}^T\mathbb{E}_{j,t}[\mathbf{s}]\Big] + \delta\, \mathbb{E}_{i,t}[\omega]. \tag{21}$$

Since we are assuming that (20) holds and that in particular the private signal MMSE estimate is $\mathbb{E}_{i,t}\left[\mathbf{s}\right] = L_{i,t}\mathbf{s}$, we can rewrite the double expectations inside the summation in (21) as

$$\mathbb{E}_{i,t}\Big[\mathbf{v}_{j,t}^T\mathbb{E}_{j,t}[\mathbf{s}]\Big] = \mathbb{E}_{i,t}\Big[\mathbf{v}_{j,t}^T L_{j,t}\mathbf{s}\Big]. \tag{22}$$

Substituting the expression in (22) into (21), and further noting that as per (20) the MMSE estimate of the world state is $\mathbb{E}_{i,t}[\omega] = \mathbf{k}_{i,t}^T \mathbf{s}$ and the estimate of the private signals is $\mathbb{E}_{i,t}[\mathbf{s}] = L_{i,t}^T \mathbf{s}$, we can simplify the equilibrium condition to

$$\mathbf{v}_{i,t}^T L_{i,t} \mathbf{s} = \sum_{j \in V \setminus \{i\}} \beta_{ij} \mathbf{v}_{j,t}^T L_{j,t} L_{i,t} \mathbf{s} + \delta\, \mathbf{k}_{i,t}^T \mathbf{s}, \tag{23}$$

which as (6), or (21) for that matter, we require for all agents $i \in V$. In order to solve this systems of equations we observe that it is underdetermined. Each vector $\mathbf{v}_{i,t}$ contains N elements and since there are N of this vectors there are a total of N^2 unknowns. However, there is one equation like (21) for each agent leading to a total of N equations. We can take advantage of this indeterminacy and proceed to equate the terms that multiply each individual signal s_j on each side of (21). This results in a set of N equations of the form

$$L_{i,t}^T \mathbf{v}_{i,t} = \sum_{j \in V \setminus \{i\}} \beta_{ij} L_{i,t}^T L_{j,t}^T \mathbf{v}_{j,t} + \delta\, \mathbf{k}_{i,t}, \tag{24}$$

associated with each agent i. Since we have a total of N agents, there are N^2 equations that we can use to determine the N^2 values of the vectors $\mathbf{v}_{i,t}$ for all agents i. Observe that the systems of linear equations defined by (24) does not depend on the realization of the private signals and that as a consequence neither do the coefficients $\mathbf{v}_{i,t}$. Irrespective of the realization of the private signals $\mathbf{s}$, the strategy of agent i at time t is the linear combination $a_{i,t}^* = \mathbf{v}_{i,t}^T \mathbb{E}_{i,t}[\mathbf{s}]$ with weights $\mathbf{v}_{i,t}$. An important consequence of this observation is that the coefficient $\mathbf{v}_{i,t}$ can be determined locally by each agent as long as he has access to the (known) network parameters without requiring knowledge of the (unknown) private signal values. The actions realized, on the other hand, depend on the observed history through the MMSE estimate $\mathbb{E}_{i,t}[\mathbf{s}]$ and hence on the realization of the private signals. As well they should.

For future reference stack all weighting coefficients $\mathbf{v}_{i,t}$ into the aggregate vector $\mathbf{v}_t := [\mathbf{v}_{1,t}^T, \ldots, \mathbf{v}_{N,t}^T]^T$ and all coefficients $\mathbf{k}_{i,t}$ into the aggregate $\mathbf{k}_t := [\mathbf{k}_{1,t}^T, \ldots, \mathbf{k}_{N,t}^T]^T$. Further define the matrix $L_t \in \mathbb{R}^{N^2 \times N^2}$ as the matrix with jth $N \times N$ diagonal block equal to $L_{j,t}^T$ and off diagonal blocks $-\beta_{ij} L_{i,t}^T L_{j,t}^T$

$$L_t := \begin{pmatrix} L_{1,t}^T & -\beta_{12} L_{1,t}^T L_{2,t}^T & \cdots & -\beta_{1N} L_{1,t}^T L_{N,t}^T \\ -\beta_{21} L_{2,t}^T L_{1,t}^T & L_{2,t}^T & \cdots & -\beta_{2N} L_{2,t}^T L_{N,t}^T \\ \vdots & \cdots & \ddots & \vdots \\ -\beta_{N-11} L_{N-1,t}^T L_{1,t}^T & \cdots & L_{N-1,t}^T & -\beta_{N-1N} L_{N-1,t}^T L_{N,t}^T \\ -\beta_{N1} L_{N,t}^T L_{1,t}^T & \cdots & -\beta_{NN-1} L_{N,t}^T L_{N-1,t}^T & L_{N,t}^T \end{pmatrix}. \tag{25}$$

With these definitions the system of linear equations in (24) can be written in the more compact form

$$L_t \mathbf{v}_t = \delta \mathbf{k}_t. \tag{26}$$

We have argued that if we can write MMSE estimates as linear combinations of private signals as per (20), the determination of the equilibrium strategy reduces to the solution of the system of linear equations in (26). However, is it true that we can write the estimates $\mathbb{E}_{i,t}[\omega]$ and $\mathbb{E}_{i,t}[\mathbf{s}]$ as in (20)? And if this is true, what are the values of the vectors $\mathbf{k}_{i,t}$ and the matrices $L_{i,t}$ which are needed to formulate (26)? To answer these questions we offer an inductive argument in the form of a recursive equation to update the values of the coefficients $\mathbf{k}_{i,t}$ and $L_{i,t}$.

An important consequence of the assumption in (20) which we have not emphasized is that equilibrium actions can be also written as a linear combination of the private signals. Indeed, since we can find vectors $\mathbf{v}_{i,t}^T$—as the solution of the system of equations in (26)—such that equilibrium actions are $a_{i,t}^* = \mathbf{v}_{i,t}^T \mathbb{E}_{i,t}[\mathbf{s}]$, and since we assume $\mathbb{E}_{i,t}[\mathbf{s}] = L_{i,t}\mathbf{s}$, we can write the action of agent i at time t as

$$a_{i,t}^* = \mathbf{v}_{i,t}^T L_{i,t}\mathbf{s}, \quad \text{for all } i \in \mathcal{N}. \tag{27}$$

Do note that as in the case of (20), we do not imply that agent i calculates its equilibrium action using (27). This is impossible because some private signals are unknown and the correct interpretation of (27) is that whatever computations agents perform to determine their equilibrium actions, they are equivalent to performing the linear combinations in (27).

The expression in (27) simplifies the understudying of the information revealed by the action of a user. From the perspective of agent i, observing the action $a_{j,t}$ of agent j is equivalent to observing the linear combination of private signals. Observing the composition of neighboring actions $\mathbf{a}_{\mathcal{N}_i,t}^* := \left[a_{j_1,t}, \ldots, a_{j_{d(i)},t}\right]^T$, where we use $d(i)$ to denote the cardinality of the set $\mathcal{N}_i$, is therefore equivalent to observing the vector linear combination

$$\mathbf{a}_{\mathcal{N}_i,t}^* \;=\; H_{i,t}^T\mathbf{s} \;:=\; \left[\mathbf{v}_{j_1,t}^T L_{j_1,t}; \ldots; \mathbf{v}_{j_{d(i)},t}^T L_{j_{d(i)},t}\right]\mathbf{s}, \tag{28}$$

where we have defined the observation matrix $H_{i,t}^T := \left[\mathbf{v}_{j_1,t}^T L_{j_1,t}; \ldots; \mathbf{v}_{j_{d(i)},t}^T L_{j_{d(i)},t}\right] \in \mathbb{R}^{d(i)\times N}$. If (27) is true for all times t, which implies that the same is true of (28), it follows that from the perspective of agent i estimation of the private signals $\mathbf{s}$ and of the underlying state of the world ω is a simple sequential linear (L)-MMSE estimation problem. Indeed, at time $t = 1$, the prior distribution $P(\omega, \mathbf{s}^T)$ is Gaussian, and i observes neighboring actions $\mathbf{a}_{\mathcal{N}_i,0}^*$ given by the linear combination $H_{i,0}^T\mathbf{s}$. Incorporating the information contained in this linear observation changes the posterior distribution to $P(\omega, \mathbf{s}^T \,\big|\, h_{i,1})$ but this latter distribution is also normal. At general time $t+1$, agent i has a normal prior $P(\omega, \mathbf{s}^T \,\big|\, h_{i,t})$ and observation of neighboring actions $\mathbf{a}_{\mathcal{N}_i,t}^* \;=\; H_{i,t}^T\mathbf{s}$ results in a normal posterior $P(\omega, \mathbf{s}^T \,\big|\, h_{i,t+1})$. Thus, to track the belief $P(\omega, \mathbf{s}^T \,\big|\, h_{i,t})$ it suffices to keep bearings on the corresponding means and variances which we can do using a LMMSE filter.

Specifically, consider agent i at time t and define the private signal covariance matrix $M_{\mathbf{ss}}^i(t)$ and the state-private signal cross covariance $M^i_{\omega\mathbf{s}}(t)$, respectively, defined by the expressions

$$M^i_{\mathbf{ss}}(t) := \mathbb{E}_{i,t}\Big[\Big(\mathbf{s}-\mathbb{E}_{i,t}[\mathbf{s}]\Big)\Big(\mathbf{s}-\mathbb{E}_{i,t}[\mathbf{s}]\Big)^T\Big], \tag{29}$$

$$M^i_{\omega\mathbf{s}}(t) := \mathbb{E}_{i,t}\Big[\Big(\omega-\mathbb{E}_{i,t}[\omega]\Big)\Big(\mathbf{s}-\mathbb{E}_{i,t}[\mathbf{s}]\Big)^T\Big]. \tag{30}$$

The LMMSE estimation of $\mathbf{s}$ from observation $\mathbf{a}^*_{\mathcal{N}_i,t} = H^T_{i,t}\mathbf{s}$ as given by (28) requires definition of the LMMSE gain $K^i_{\mathbf{s}}(t)$ given by the product of the cross covariance between the signal $\mathbf{s}$ and the observation $\mathbf{a}^*_{\mathcal{N}_i,t}$ times the inverse of the covariance matrix of the observation $\mathbf{a}^*_{\mathcal{N}_i,t}$. Since the covariance of $\mathbf{s}$ is $M^i_{\mathbf{ss}}(t)$ and the observation model is $\mathbf{a}^*_{\mathcal{N}_i,t} = H^T_{i,t}\mathbf{s}$, the LMMSE gain is given explicitly by

$$K^i_{\mathbf{s}}(t) = M^i_{\mathbf{ss}}(t)H_{i,t}\Big(H^T_{i,t}M^i_{\mathbf{ss}}(t)H_{i,t}\Big)^{-1}. \tag{31}$$

Using the value of the LMMSE gain $K^i_{\mathbf{s}}(t)$ in (31), the posterior mean $\mathbb{E}_{i,t+1}[\mathbf{s}]$ and posterior covariance matrix $M^i_{\mathbf{ss}}(t+1)$ after observing the neighboring actions $\mathbf{a}^*_{\mathcal{N}_i,t}$ follow from the recursive expressions

$$\mathbb{E}_{i,t+1}[\mathbf{s}] \quad = \quad \mathbb{E}_{i,t}[\mathbf{s}] + K^i_{\mathbf{s}}(t)\Big(\mathbf{a}^*_{\mathcal{N}_i,t} - \mathbb{E}_{i,t}[\mathbf{a}^*_{\mathcal{N}_i,t}]\Big), \tag{32}$$

$$M^i_{\mathbf{ss}}(t+1) \;=\; M^i_{\mathbf{ss}}(t) - K^i_{\mathbf{s}}(t)H^T_{i,t}M^i_{\mathbf{ss}}(t) \tag{33}$$

where the executed value of the observations follows from (28) as $\mathbb{E}_{i,t}[\mathbf{a}^*_{\mathcal{N}_i,t}] = H^T_{i,t}\mathbb{E}_{i,t}[\mathbf{s}]$. Likewise, for the estimation of the state ω from observations $\mathbf{a}^*_{\mathcal{N}_i,t}$ we compute the LMMSE gain

$$K^i_{\omega}(t) = M^i_{\omega\mathbf{s}}(t)H_{i,t}\Big(H^T_{i,t}M^i_{\mathbf{ss}}(t)H_{i,t}\Big)^{-1}, \tag{34}$$

given by the product of the cross covariance $M^i_{\omega\mathbf{s}}(t)H_{i,t}$ between signal ω and observation $\mathbf{a}^*_{\mathcal{N}_i,t} = H^T_{i,t}\mathbf{s}$ times the inverse of the observation's covariance $H^T_{i,t}M^i_{\mathbf{ss}}(t)H_{i,t}$. We then have that the state's posterior mean $\mathbb{E}_{i,t+1}[\omega]$ and posterior cross covariance $M^i_{\omega\mathbf{s}}(t+1)$ after observing the neighboring actions $\mathbf{a}^*_{\mathcal{N}_i,t}$ are given by the recursions

$$\mathbb{E}_{i,t+1}[\omega] \quad = \quad \mathbb{E}_{i,t}[\omega] + K^i_{\omega}(t)\Big(\mathbf{a}^*_{\mathcal{N}_i,t} - \mathbb{E}_{i,t}\left[\mathbf{a}^*_{\mathcal{N}_i,t}\right]\Big), \tag{35}$$

$$M^i_{\omega\mathbf{s}}(t+1) \;=\; M^i_{\omega\mathbf{s}}(t) - K^i_{\omega}(t)H^T_{i,t}M^i_{\mathbf{ss}}(t). \tag{36}$$

We emphasize that it is possible to write a similar variance update for the world state variance $M^i_{\omega\omega}(t) := \mathbb{E}_{i,t}\big[(\omega - \mathbb{E}_{i,t}[\omega])^2\big]$ but this is inconsequential to our argument. Further note that the somewhat unfamiliar form of the LMMSE gains $K^i_{\mathbf{s}}(t)$ in (31) and $K^i_{\omega}(t)$ in (34) are due to the fact that the observation model $\mathbf{a}^*_{\mathcal{N}_i,t} = H^T_{i,t}\mathbf{s}$ in (28) is noiseless. We have therefore concluded that if we have linear actions as per (27), which is true as long as (20) is true, the propagation of beliefs $P(\omega, \mathbf{s}^T \,\big|\, h_{i,t})$ reduces to the recursive propagation of means and covariances in (31)–(36).

From the expressions in (32) and (35) we can see that it is possible to write the state and private signal expectations at time $t+1$ as the linear combinations $\mathbb{E}_{i,t+1}[\omega] = \mathbf{k}_{i,t+1}^T \mathbf{s}$ and $\mathbb{E}_{i,t+1}[\mathbf{s}] = L_{i,t+1}\mathbf{s}$ akin to those shown in (20). For the private signal MMSE, substitute $\mathbb{E}_{i,t}[\mathbf{s}] = L_{i,t}\mathbf{s}$, $\mathbf{a}_{\mathcal{N}_i,t}^* = H_{i,t}^T \mathbf{s}$, and $\mathbb{E}_{i,t}[\mathbf{a}_{\mathcal{N}_i,t}^*] = H_{i,t}^T \mathbb{E}_{i,t}[\mathbf{s}] = H_{i,t}^T L_{i,t}\mathbf{s}$ into (32) to conclude that if it is possible to write (20) at time t, we can also write it at time $t+1$. Perhaps more importantly, these substitutions also yield a recursive formula that allows updating the matrices $L_{i,t}$ as

$$L_{i,t+1} = L_{i,t} + K_{\mathbf{s}}^i(t)\Big(H_{i,t}^T - H_{i,t}^T L_{i,t}\Big). \tag{37}$$

The same argument can be made for $\mathbb{E}_{i,t}[\omega]$ to conclude that if $\mathbb{E}_{i,t}[\omega] = \mathbf{k}_{i,t}^T \mathbf{s}$ at time t, it is also true at time $t+1$ with the linear combination coefficients adhering to the recursion

$$\mathbf{k}_{i,t+1}^T = \mathbf{k}_{i,t}^T + K_{\omega}^i(t)\Big(H_{i,t}^T - H_{i,t}^T L_{i,t}\Big). \tag{38}$$

To complete the inductive argument we need to show that (20) is true at time $t=0$, but this is obviously true because $\mathbb{E}_{i,0}[\omega] = s_i$ and $\mathbb{E}_{i,0}[s_j] = s_i$ for all agents.

The induction loop we just completed is sufficiently long so as to warrant retracing. We begin by the assumption that at time t we can write MMSE estimates of the state of the world ω and the private signals $\mathbf{s}$ as linear combinations of the private signals themselves as per (20). From this assumption it follows that equilibrium actions can be written as the linear combinations of private signals in (27). From here it follows that beliefs are propagated as per the LMMSE filter summarized in (31)–(36). A simple set of substitutions allows us to conclude that (20) is true at time $t+1$ with the vector $\mathbf{k}_{i,t+1}$ propagated as per (38) and the matrix $L_{i,t+1}$ propagated as in (38).

The expressions in (20) and (27) are neither used to propagate beliefs nor to compute equilibrium actions. The actual operations carried by each agent are summarized in Figs. 1 and 2 and described in the following section.

A. 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. 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 his equilibrium action $a_{i,t}^*$.

The QNG filter at node i is an implementation of the LMMSE filters defined by (32) and (35) followed by the play $a_{i,t}^* = \mathbf{v}_{i,t}^T \mathbf{E}_{i,t}[\mathbf{s}]$. A block diagram for this filter is shown in Fig. 1. At time t, the input to the filter is the observed actions $\mathbf{a}_{\mathcal{N}_i,t}^*$ of agent i's neighbors. The prediction $\mathbf{E}_{i,t}[\mathbf{a}_{\mathcal{N}_i,t}^*] = H_{i,t}^T \mathbf{E}_{i,t}[\mathbf{s}]$ of this vector is subtracted from the observed value and the result is fed into two parallel blocks respectively tasked with updating the belief $\mathbf{E}_{i,t}[\omega]$ on the state of the world ω and the belief $\mathbf{E}_{i,t}[\mathbf{s}]$ on the private signals $\mathbf{s}$ of other agents. To update the belief on ω we implement (35) by multiplying the innovation $\mathbf{a}_{\mathcal{N}_i,t}^* - \mathbf{E}_{i,t}[\mathbf{a}_{\mathcal{N}_i,t}^*]$ by the gain $K_{\omega}^i(t)$ and add the

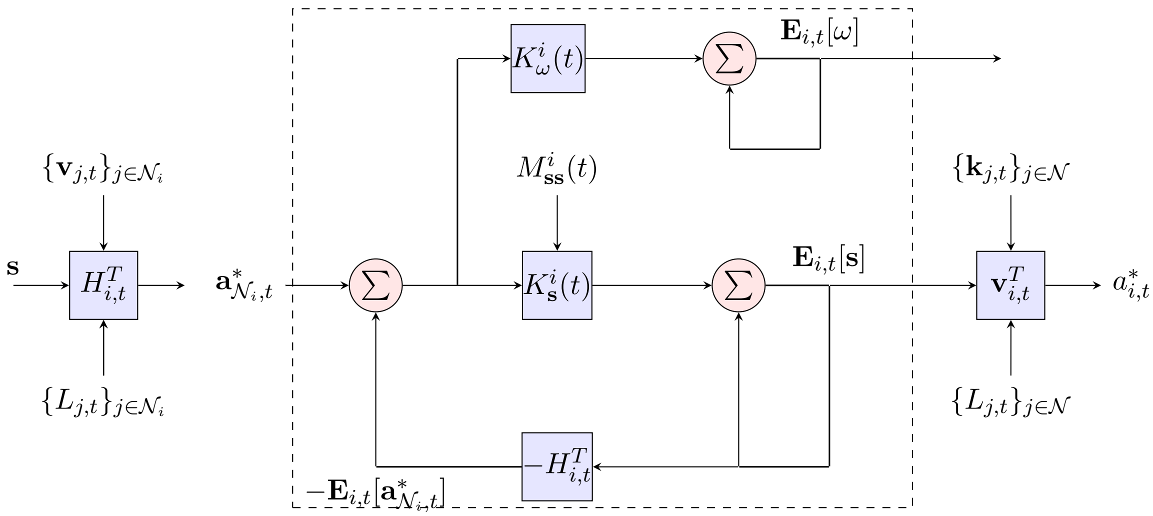


Fig. 1. Block diagram of the Quadratic Network Game (QNG) filter at agent i. The QNG filter contains a mechanism for belief propagation and a mechanism to calculate equilibrium actions. Inside the dashed box, the belief propagation feedback loops that compute the estimates of $\mathbf{s}$ and ω as linear combinations of private signals' estimates of previous time are summarized. The observation prediction is subtracted from the observation to form the prediction error. Afterwords, the belief propagation for $\mathbf{s}$ and ω follow the same steps with different gains. The prediction error is multiplied by the corresponding gain matrix, and added to the previous mean estimate to form the corrected estimate [cf. (32) and (35)]. Multiplying the corrected signal estimate with the action coefficient gives the equilibrium action. The gain coefficients are provided by the LMMSE block in Fig. 2. The observation matrix and action coefficient are fed from the game block in Fig. 2. While these coefficients can be calculated by each agent, the mean estimates $\mathbf{E}_{i,t}[\mathbf{s}]$ and equilibrium action $a^*_{i,t}$ can only be calculated by agent i.

result to the previous state estimate $\mathbf{E}_{i,t}[\omega]$. To update the belief on the private signals $\mathbf{s}$ we multiply the innovation by the LMMSE gain $K^i_{\mathbf{s}}(t)$. The result of this amplification is added to the previous private signal belief $\mathbf{E}_{i,t}[\mathbf{s}]$ as dictated by (32). In order to determine the equilibrium play we multiply the private signal estimate $\mathbf{E}_{i,t}[\mathbf{s}]$ by the vector $\mathbf{v}^T_{i,t}$ obtained by solving the system of linear equations in (26).

Observe that in the QNG filter, we do not use the fact that estimates $\mathbf{E}_{i,t}\left[\omega\right]$ and $\mathbf{E}_{i,t}[\mathbf{s}]$ as well as actions $a^*_{i,t}$ can be written as linear combinations of the private signals [cf. (20) and (27)]. While the expressions in (20) and (27) are certainly correct, they cannot be used for implementation, because $\mathbf{s}$ is partially unknown to agent i. The role of (20) and (27) is to allow derivation of recursions that we use to keep track of the gains used in the QNG filter. These recursions can be divided into a group of LMMSE updates and a group of game updates as we show in Fig. 2.

As it follows from (31), (33), (34), and (36), the update of LMMSE coefficients is identical to the gain and covariance updates of a conventional sequential LMMSE. The only peculiarity is that the observation matrix $H_{j,t}$ is fed from the game update block and is partially determined by the LMMSE gains and covariances of previous iterations. Nevertherless, this peculiarity is more associated with the game block than with the LMMSE block. The game block uses (37) and (38) to keep track of the matrices $L_{j,t}$ and the vectors $\mathbf{k}_{j,t}$. The matrices $L_{j,t}$ are used as building blocks of the matrix L_t and the vectors $\mathbf{k}_{j,t}$ are stacked in the vector $\mathbf{k}_t$ and used to formulate the systems of equations in (26). Solving this system of equations yields the coefficients $\mathbf{v}_{j,t}$ which in turn determine the observation matrix $H_{j,t}$ as per (28). As we mentioned before, the game block feeds the matrices $H_{j,t}$ to the

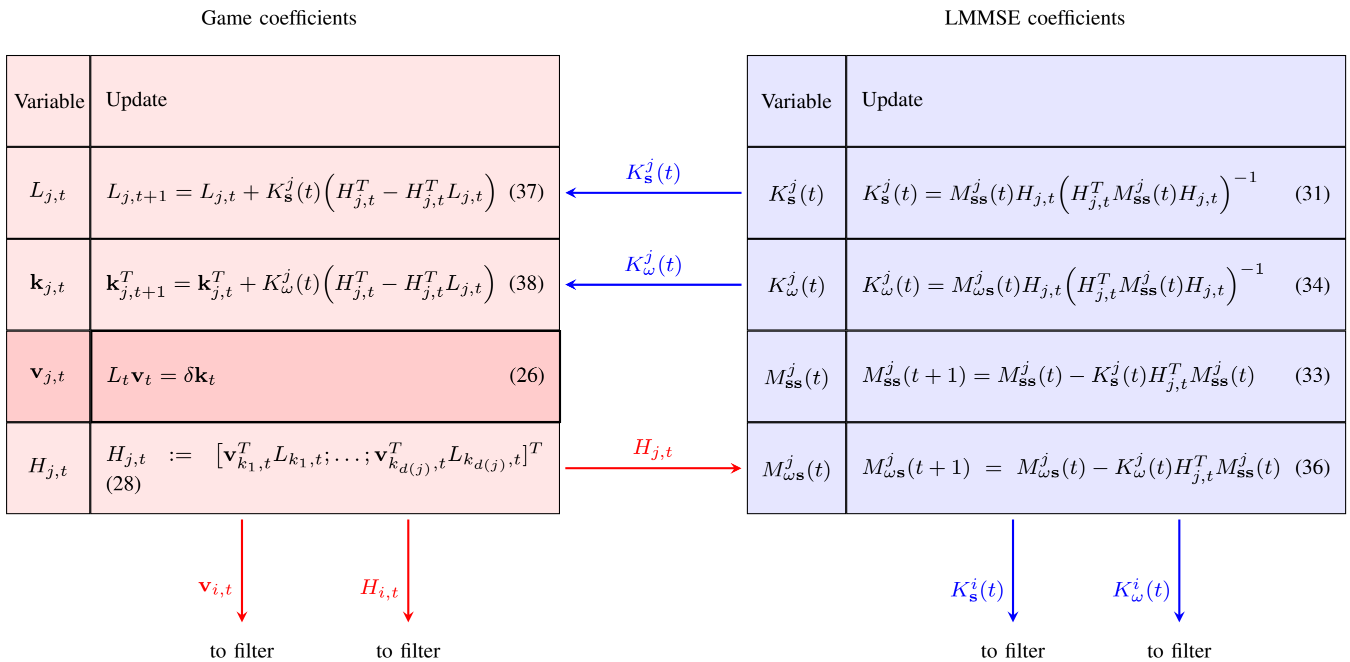


Fig. 2. Propagation of gains required to implement the QNG filter of Fig. 1. Gains are separated into interacting LMMSE and game blocks. All agents perform a full network simulation in which they compute the gains of all other agents. This is necessary because when we compute the play coefficients $\mathbf{v}_{j,t}$ in the game block, agent i builds the matrix L_t that is formed by the blocks $L_{j,t}$ of all agents [cf. (25)]. This full network simulation is possible because the network topology and private signal models are assumed to be common knowledge.

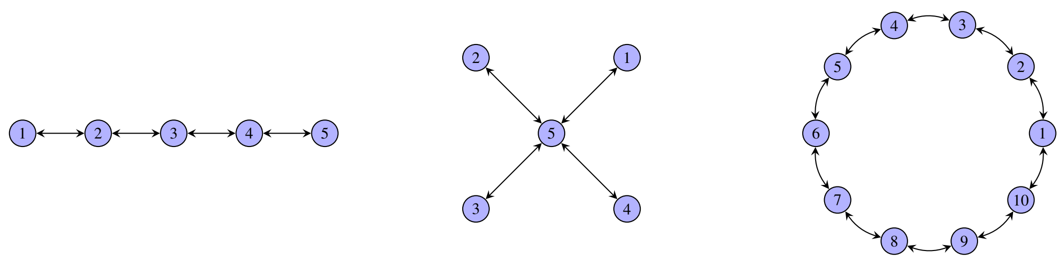


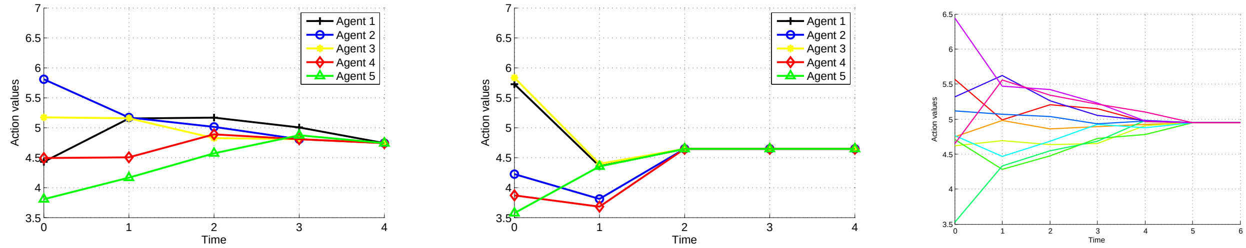
Fig. 3. Line, star and ring networks.

filter block as these are used in the LMMSE gains and covariance updates. The LMMSE block feeds the gains $K^j_{\mathbf{s}}(t)$ and $K^j_{\omega}(t)$ to the game block as these are needed to update $L_{j,t}$ and $\mathbf{k}_{j,t}$.

A fundamental observation is that agent i is keeping track of the matrices and vectors in Fig. 2 in their entirety and not only of their components corresponding to himself. The reason for this is the step in the game block in which we compute the play coefficients $\mathbf{v}_{j,t}$. To solve this system of equations we need to build the matrix L_t that is formed by the blocks $L_{j,t}$ of all agents. All of these computations for other agents are internal, however. The QNG as shown in Fig. 1 simply needs access to the LMMSE gains $K^i_{\mathbf{s}}(t)$ and $K^i_{\omega}(t)$ fed from the filter block as well as the observation matrix $H_{i,t}$ and the play coefficients $\mathbf{v}_{i,t}$ fed from the game block.

IV. Numerical Examples

We use the QNG filter derived earlier to explicitly propagate individual beliefs and compute the equilibrium actions locally for the coordination game introduced in Section II-B. Agents weight estimation and coordination



(a) (b) (c)

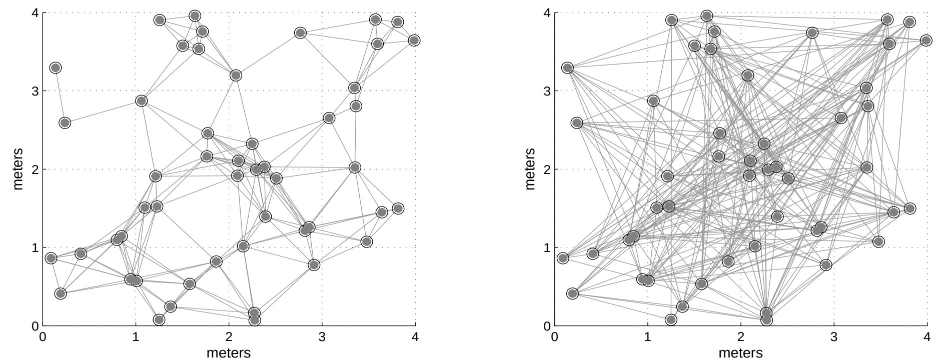
Fig. 4. Values of agents' actions over time for the coordination game and networks in Fig. 3, respectively. Each line plots agent i's action at each time. Agents reach consensus in the optimal estimate $\hat{\omega}^*$ in the number of steps equal to the diameter of the corresponding network.

components of the payoff (12) equally, that is, $\lambda = 0.5$. In all of the examples, we set the true value of the stock to be $\omega = \$5$, and signal structure is as given by (19) where ϵ_i is Gaussian with mean zero and $C_i = 1$. We test the QNG filter on various networks.

We first consider line $(N = 5)$, star $(N = 5)$ and ring $(N = 10)$ networks depicted in Fig. 3. The evolution of each agent's action over time is depicted in Figs. 4 (a)–(c) for the corresponding line, star and ring networks. The results show that agents reach consensus in their actions as indicated by the asymptotic consensus result described in Section II-C. Furthermore, the consensus action is the optimal estimate of the stock value $\hat{\omega}^* := \mathbb{E}[\omega \,|\, \mathbf{s}]$ which is also the BNE of the complete information game. Note that this does not necessarily imply that agents learn the true value of all the private signals; rather, this implies that they learn the sufficient statistic (in this case, the mean of the private signals) to calculate the optimal estimate of ω.

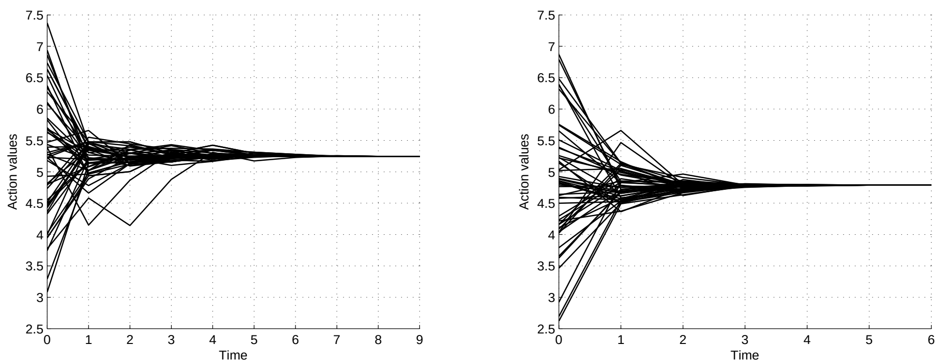
We further evaluate convergence behavior of the QNG filter in geometric and random networks shown in Figs. 5 (a)–(b), respectively. Both networks contain $N = 50$ agents. For the geometric network, agents are randomly placed on a 4m by 4m square, and then pairs that are less than 1m apart are connected. In the random network pairs are connected with probability 0.2. The evolution of each agent's action values over time is depicted in Figs. 6 (a)–(b) for the geometric and random networks, respectively. In this case, we also observe that the action consensus is achieved at $\hat{\omega}^*$ implying that consensus holds for any connected network.

Our characterization of agents' updates enables us to characterize the convergence rates based on network properties. In the three benchmark networks of Fig. 3, diameter of the network Δ is the sole determinant of the convergence rate, that is, agents' actions converge in exactly Δ steps in all these cases; see Figs. 4 (a)–(c). The diameters of the geometric and random networks are 7 and 3, respectively. Convergence to consensus action happens in $O(\Delta)$ for both networks; see Figs. 6 (a)–(b). In all of these examples, we observe that Δ is the sole determinant of convergence rate. It is shown in [30] that agents on a connected network converge to $\hat{\omega}^*$ in at most $2N\Delta$ steps for the payoff function $u(\omega, a_i) = -(\omega - a_i)^2$. In the same paper, it is also conjectured that convergence occurs in $O(N)$ steps. The model exhibiting no payoff externalities with the payoff function $u(\omega, a_i) = -(\omega - a_i)^2$ is a specific case of the general framework presented in this paper. Therefore, we expect similar results regarding



(a) (b)

Fig. 5. Geometric (a) and random (b) networks with $N = 50$ agents. Agents are randomly place on a 4meter $\times$ 4meter 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 0.2.



(a) (b)

Fig. 6. Values of agents' actions over time in the coordination game for geometric and random networks in Fig. 5 (a) and Fig. 5 (b), respectively. Each line indicates an agent's sequence of actions over the time horizon. Agents reach consensus at the optimal estimate value $\hat{\omega}^*$ in number of steps proportional to the diameter of the corresponding network.

convergence rates to hold for the QNG filter. Furthermore, our simulation results indicate that $2N\Delta$ and $O(N)$ are crude upper bounds for the convergence rate. We conjecture that convergence happens in $O(\Delta)$ steps; however, proving this remains an open problem.

V. Concluding Remarks

This article provides an overview of recent results in social learning models in presence of payoff externalities with a focus on agent behavior. We presented a framework to model repeated games of incomplete information over networks and showed that when agents' utilities are quadratic—under certain assumptions—agent over a connected network eventually reach consensus in their actions and expected payoffs.

Algorithmic aspects of rational learning received special attention. We derived the QNG filter for propagating beliefs in quadratic network games when signals are Gaussian. Numerical examples were provided for various network structures. Based on simulations, we stated and discussed results that show convergence rates based on network diameter.

VI. Callout: Example of Learning with Payoff Externality

In this section, we give an example of rational behavior in a model with both informational and payoff externalities. The example illustrates how rational agents are able to rule out possible states of the world upon observing actions of their neighbors.

There are three agents in a line network; that is, $\mathcal{N} = \{1, 2, 3\}$, $\mathcal{N}_1 = \{2\}$, $\mathcal{N}_2 = \{1, 3\}$, and $\mathcal{N}_3 = \{2\}$. The possible states of the world belong to the set, $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Agents have a common uniform prior over the possible states. At the beginning, agents receive private signals s_1, s_2, and s_3. Based on s_1, agent 1 can distinguish whether the true state is ω_3 or belongs to the set $\{\omega_1, \omega_2\}$. The private signal of s_2 does not carry any information. s_3 reveals whether the true state is ω_1 or belongs to the set $\{\omega_2, \omega_3\}$. We assume that agents know the informativeness of the private signals of all agents; i.e., the partition of the private signals is known by all agents. There are two possible actions, $A = \{l, r\}$.

Agent i's payoff depends on the actions of the other two agents $a_{\mathcal{N}\setminus\{i\},t} := \{a_{j,t}\}_{\mathcal{N}\setminus i}$ in the following way:

$$u_i(\omega, a_{i,t}, a_{\mathcal{N}\setminus\{i\},t}) = \begin{cases} 1 & \text{if} \quad \omega = \omega_1, a_{i,t} = l, a_{\mathcal{N}\setminus\{i\},t} = \{l, l\}, \\ 4 & \text{if} \quad \omega = \omega_3, a_{i,t} = r, a_{\mathcal{N}\setminus\{i\},t} = \{r, r\}, \\ 0 & \text{otherwise.} \end{cases} \tag{39}$$

According to (39), agent i earns a payoff only when all the agents choose l and the state is ω_1 or when all the agents choose r and the state is ω_3.

Initial strategies of agents consist of functions that map their observed histories at $t = 0$ (which are only their signals) to actions. Let $(\sigma^*_{1,0}, \sigma^*_{2,0}, \sigma^*_{3,0})$ be a strategy profile at $t = 0$ defined as

$$\sigma^*_{1,0}(s_1) = \begin{cases} l & \text{if} \quad s_1 = \{\omega_1, \omega_2\}, \\ r & \text{if} \quad s_1 = \{\omega_3\}, \end{cases}$$
$$\sigma^*_{2,0}(s_2) = r,$$
$$\sigma^*_{3,0}(s_3) = \begin{cases} l & \text{if} \quad s_3 = \{\omega_1\}, \\ r & \text{if} \quad s_3 = \{\omega_2, \omega_3\}. \end{cases}$$

Note that since agent 2's signal is uninformative, he needs to take the same action regardless of his signal.

Agents' strategies at a time $t \geq 1$ map their observed histories to actions. For $t \geq 1$ let the $(\sigma^*_{1,t}, \sigma^*_{2,t}, \sigma^*_{3,t})$ be a strategy profile defined as

$$\sigma^*_{1,t}(h_{1,t}) = \begin{cases} l & \text{if} \quad s_1 = \{\omega_1, \omega_2\}, \\ r & \text{if} \quad s_1 = \{\omega_3\}, \end{cases}$$
$$\sigma^*_{2,t}(h_{2,t}) = \begin{cases} r & \text{if} \quad a_{1,t-1} = a_{3,t-1} = r, \\ l & \text{otherwise}, \end{cases}$$
$$\sigma^*_{3,t}(h_{3,t}) = \begin{cases} l & \text{if} \quad s_3 = \{\omega_1\}, \\ r & \text{if} \quad s_3 = \{\omega_2, \omega_3\}. \end{cases}$$

Note that even though agents' strategies could depend on their entire histories, in the above specification agent 1 and 3's actions only depend on their private signals, whereas, agent 2's actions only depend on the last actions taken by his neighbors.

We argue that $\sigma^* = (\sigma^*_{i,t})_{i\in\mathcal{N},t=0,1,...}$ as defined above is an equilibrium strategy. We assume that the strategy profile σ^* is common knowledge and verify that agents' actions given any history maximizes their expected utilities given the beliefs induced by the Bayes' rule.

First, consider the time period $t = 0$. Suppose that agent 1 observes $s_1 = \{\omega_1, \omega_2\}$. He assigns one half probability to the event $\omega = \omega_1$ in which case—according to σ^*—agent 2 plays r and agent 3 plays l, and he assigns one half probability to state $\omega = \omega_2$ in which case agent 2 plays r and agent 3 plays r. Therefore, his expected payoff is zero regardless of the action he takes; that is, he does not have a profitable unilateral deviation from the strategy profile σ^*. Next suppose that agent 1 observes $s_1 = \{\omega_3\}$. In this case he knows for sure that $\omega = \omega_3$ and that agents 2 and 3 both play r. Therefore, the best he can do is also to play r—which is the action specified by σ^*. This argument shows that agent 1 has no profitable deviation from σ^* regardless of the realization of s_1. Next, we focus on agent 2. She has no information at $t = 0$. Therefore, he assigns one third probability to the event $\omega = \omega_1$ in which case $a_{1,0} = a_{3,0} = l$, one third probability to the event $\omega = \omega_3$ in which case $a_{1,0} = l$ and $a_{3,0} = r$, and one third probability to the event $\omega = \omega_2$ in which case $a_{1,0} = a_{3,0} = r$. Therefore, his expected payoff of taking action r is $4/3$, whereas his expected payoff of taking action l is $1/3$. Finally, considering agent 3, if he observes $s_3 = \{\omega_1\}$, he knows that agents 1 and 2 play l and r respectively, in which case he is indifferent between l and r. If he observes $s_3 = \{\omega_2, \omega_3\}$, on the other hand, he assigns one half probability to $\omega = \omega_2$ in which case $a_{1,0} = l$ and $a_{2,0} = r$, and one half probability to $\omega = \omega_3$ in which case $a_{1,0} = a_{2,0} = r$. Therefore, he strictly prefers playing r in this case. We have shown that at $t = 0$, no agent has an incentive to deviate from the actions prescribed by σ^*. We have indeed shown something stronger. Strategies $\sigma^*_{1,0}$ and $\sigma^*_{2,0}$ are *dominant strategies* for agents 1 and 3, respectively; that is, regardless of what other agents do, agents 1 and 3 have no incentive to deviate from playing these strategies.

Next, consider the time period $t = 1$. In this time period, agent 2 knowing the strategies that agents 1 and 3 used in the previous time period learns the true state; namely, if they played $\{l, l\}$ the state is ω_1, if they played $\{r, r\}$ the state is ω_3, and otherwise the state is ω_2. Also, by the above argument agents 1 and 3 will never have an incentive to change their strategies from what is prescribed by σ^*. Therefore, σ^* is consistent with equilibrium at $t = 1$ as well. The exact same argument can be repeated for $t > 1$.

Now that we have shown that σ^* is an equilibrium strategy, we can focus on the evolution of agents' expected payoffs. For the rest of the example, assume that $\omega = \omega_1$. At $t = 0$, agent 3 learns the true state. Agents 1, 2, and 3 play l, r, and l, respectively. Since agents 1 and 2 know that agent 2 will play $a_{2,0} = r$, their conditional expected payoffs at $t = 0$ are zero. Agent 2 on the other hand, assigns one third probability to the state ω_3 and action profile (r, r, r); therefore, his expected payoff is given by $4/3$. At $t = 1$, all agents play l. Agent 2 learns the true state. Since agents 2 and 3 know the true state and known that the action profile that is chosen is (l, l, l), their expected payoffs are equal to one. On the other hand, agent 1 does not know whether the state is ω_1 or ω_2 but he knows that the action profile taken is (l, l, l); therefore, his conditional expected payoff is equal to $1/2$. In later stages, agents changes neither their beliefs nor their actions.

The example illustrates an important aspect of learning in presence of payoff externalities. Agents need to infer about the actions of other agents in the next stage based on the information available to them and using the knowledge of equilibrium strategy in order to make prediction about how others would play. This inference process includes reasoning about others' reasoning about actions of self and other agents which in turn leads to the notion of equilibrium strategy that we formally defined in Section I.

REFERENCES

[1] M. O. Jackson and L. Yariv. The diffusion of behavior and equilibrium structure on social networks. *American Economic Review (papers and proceedings)*, 97:92–98, 2007.

[2] Y. Bramoullé, R. Kranton, and M. D'Amours. Strategic interaction and networks. *Working Paper*, 2009, Available at SSRN: http://ssrn.com/abstract=1612369.

[3] A. Calvó-Armengol and J.M. Beltran. Information gathering in organizations: equilibrium, welfare, and optimal network structure. *Journal of the European Economic Association*, 7(1):116–161, 2009.

[4] G.M. Angeletos and A. Pavan. Efficient use of information and social value of information. *Econometrica*, 75(4):1103–1142, 2007.

[5] O. Bandiera and I. Rasul. Social networks and technology adoption in northern mozambique. *The Economic Journal*, 116:869–902, 2006.

[6] T. Conley and C. Udry. Social learning through networks: The adoption of new agricultural technologies in ghana. *American Journal of Agricultural Economics*, 83:668–673, 2001.

[7] A. Ambrus, M. Mobius, and A. Szeidl. Consumption risk-sharing in social networks. Technical report, National Bureau of Economic Research, 2010.

[8] M. Kearns, S. Judd, J. Tan, and J. Wortman. Behavioral experiments on biased voting in networks. *Proceedings of the National Academy of Science*, January 2009.

[9] B. Latané and T. L'Herrou. Spatial clustering in the conformity game: Dynamic social impact in electronic groups. *Journal of Personality and Social Psychology*, 70:1218–1230, 1996.

[10] D. Lazer, A. Pentland, L. Adamic, S. Aral, A. Barabsi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M.V. Alstyne. Computational social science. *Science*, 323:721–723, 2009.

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

[12] 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.

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

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

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

[16] M. Mueller-Frank. A general framework for rational learning in social networks. *Working Paper*, 2011, Available at SSRN: http://ssrn.com/abstract=1690924 or http://dx.doi.org/10.2139/ssrn.1690924.

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

[18] 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.

[19] 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.

[20] I. Schizas, G. Giannakis, S. Roumeliotis, and A. Ribeiro. Consensus in ad hoc wsns with noisy links - part ii: distributed estimation and smoothing of random signals. *IEEE Trans. Signal Process.*, 56(4):1650–1666, April 2008.

[21] 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.

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

[23] J. Chen and A.H. Sayed. Diffusion adaptation strategies for distributed optimization and learning over networks. Arxiv 1111.0034v2, 2012.

[24] J. Chen, S.-Y. Tu, and A.H. Sayed. Distributed optimization via diffusion adaptation. In *Proc. IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, San Juan, Puerto Rico, December 2011.

[25] 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.

[26] V. Krishnamurthy. Quickest time herding and detection for optimal social learning. arxiv, abs/1003.4972, 2010, 2011, arXiv:1003.4972v2.

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

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

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

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

[31] V. Borkar and P. Varaiya. Asymptotic agreement in distributed estimation. *IEEE Trans. Autom. Control*, 27(3):650–655, 1982.

[32] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Autom. Control*, 48(6):988–1001, 2003.

[33] P. Molavi and A. Jadbabaie. Network topology and efficiency of observational social learning. In *Proceedings of the 51st IEEE Conference on Decision and Control*, 2012. forthcoming.

[34] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and Tsitsiklis J.N. Convergence in multiagent coordination, consensus, and flocking. In *Proc. of the 44th IEEE Conference on Decision and Control*, pages 2996–3000, Seville, Spain., 2005.

[35] D. Monderer and L.S. Shapley. Potential games. *Games and economic behavior*, 14:124–143, 1996.

[36] J.R. Marden, G. Arslan, and J.S. Shamma. Cooperative control and potential games. *IEEE Trans. Syst., Man, and Cybern. B, Cybern.*, 39(6):1393–1407, 2009.

[37] 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.

[38] 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 (invited)*, Pacific Grove, California, USA., 2012.

Ceyhun Eksin received the B.Sc. degree in Control Engineering from Istanbul Technical University, in 2005. He received his M.Sc. degree in Industrial Engineering from Bogazici University, Istanbul, in 2008. In this period, he also spent one semester at Technische Universiteit Eindhoven as Erasmus Exchange student. He joined the University of Pennsylvania in 2008 as a Ph.D. student. His research interests are in the areas of signal processing, distributed optimization and social networks. Since 2011, he has been focusing on distributed optimization and learning in social, biological and technological networks.

Pooya Molavi received his B.S. degree in Electrical Engineering from Sharif University of Technology, Tehran, Iran, in 2008. In the same year, he joined the department of Electrical and Systems Engineering and GRASP Laboratory at the University of Pennsylvania, Philadelphia, PA, as a Ph.D. student. He received the gold medal of the 16th Iranian National Physics Olympiad in 2003. His research interests are in the areas of social learning and network economics.

Alejandro Ribeiro is an assistant professor at the Department of Electrical and Systems Engineering at the University of Pennsylvania (Penn), Philadelphia, where he started in 2008. He received the B.Sc. in electrical engineering from the Universidad de la Republica Oriental del Uruguay, Montevideo, in 1998. From 2003 to 2008 he was at the Department of Electrical and Computer Engineering, the University of Minnesota, Minneapolis, where he received the M.Sc. and Ph.D. in electrical engineering. From 1998 to 2003 he was a member of the technical staff at Bellsouth Montevideo. His research interests lie in the areas of communication, signal processing, and networking. His current research focuses on network and wireless communication theory. He received the 2012 S. Reid Warren, Jr. Award presented by Penn's undergraduate student body for outstanding teaching and the NSF CAREER award in 2010. He is also a Fulbright scholar and the recipient of student paper awards at ICASSP 2005 and ICASSP 2006.

Ali Jadbabaie received his BS degree (with High honors) in Electrical Engineering from Sharif University of Technology in 1995. He received his Masters degree in Electrical and Computer Engineering from the University of New Mexico, Albuquerque in 1997 and his Ph.D. degree in Control and Dynamical Systems from California Institute of Technology (Caltech) in 2001. From July 2001-July 2002 he was a postdoctoral associate at the department of Electrical Engineering at Yale University. Since July 2002 he has been with the department of Electrical and Systems Engineering and GRASP Laboratory at the University of Pennsylvania, Philadelphia, PA, where he is currently the Founding Co-director of Singh Program in Market and Social Systems Engineering and a Professor of Electrical and Systems Engineering with Secondary appointments in the departments of Computer and Information Sciences and Operations and Information Management in the Wharton School. He is a recipient of various awards, including NSF Career, ONR Young Investigator, Best student paper award of the American Control Conference 2007 (as advisor), the O. Hugo Schuck Best Paper award of the American Automatic Control Council (joint with his student), and the George S. Axelby Outstanding Paper Award of the IEEE Control Systems Society. His research is broadly in the interface of control theory and network science, specifically, analysis, design and optimization of networked dynamical systems with applications to sensor networks, multi-robot formation control, social aggregation, network economics, distributed optimization, and other collective phenomena.