Learning Networks*

Promit K. Chaudhuri† Jonghyun Park† Sudipta Sarangi†
Hector Tzavellas†

September 2026

Abstract

We study repeated strategic interaction when agents observe only their local neighborhoods and are uncertain about the network in which they are embedded. The interaction structure is both the object of learning and the structure that governs strategic interaction, local observation, and the flow of information. Agents infer the network from locally observed equilibrium actions and update beliefs using Bayes' rule. We show that learning terminates in finite time and that equilibrium behavior converges to the complete-information Nash equilibrium of the realized network. A central implication is that correct behavior and correct knowledge are distinct outcomes: agents may behave exactly as if the network were known while still failing to identify it. We characterize when identification occurs and show that, in the linear-quadratic network game, networks with distinct walk profiles are generically identified. Finally, we show that learning speed is jointly determined by network topology and the ability of equilibrium behavior to transmit information.

JEL Classifications: C73, D83, D85

Keywords: Bayesian learning, network games, local observation, network uncertainty

*We would like to thank Patrick Allmis, Christophe Bravard, Niloy Bose, Anindya Chakrabarti, Pritha Dev, Bhaskar Dutta, Srijita Ghosh, Ben Golub, Ahmed S. Mahmud, Matt Jackson, Luca Merlino, Debraj Ray, Ravi Ashok Satpute as well as seminar and conference participants of the University of Virginia Biocomplexity Institute, CRETE 2025, India Management Research Conference IIMA, 19th Annual Conference on Economic Growth and Development ISI New Delhi, Fifth International Conference on Issues in Economic Theory & Policy Presidency University Kolkata, Tenth Annual Conference on Network Science and Economics, the Midwest International Trade & Theory conference 2025 Spring, Southern Economic Association 95th meeting for their insightful comments and suggestions. Sudipta Sarangi would like to acknowledge the support of *Collegium de Lyon* and the *French Institutes of Advanced Study* for this research, specifically for funding received from the Horizon Europe Research and Innovation Programme (Marie Skłodowska-Curie Actions (MSCA COFUND GA-945408)).

†Department of Economics, Virginia Tech. Email: (1) promit@vt.edu, (2) jongh2419@vt.edu, (3) ssarangi@vt.edu, (4) hectortz@vt.edu

1 Introduction

Many economic interactions take place in networked settings. Often members of the network may not know the entire network: they only observe what happens locally and may know little about who their friends are connected to and how information reaches others. In social and economic networks for example, individuals see the behavior of their neighbors but not who influences whom beyond their immediate peers. In R&D collaboration firms may know their own collaborators, but not all the other partners of their collaborators. In organizations, workers observe close collaborators or team members without knowing the broader structure through which incentives and information propagate. Agents in these settings must form expectations about others from what they observe together with an incomplete understanding of what others observe. Consistent with this premise, Breza et al. (2018) document that individuals' knowledge of network links is highly localized and declines sharply with network distance. This raises a natural question: how do agents learn and behave when the structure of interaction itself is uncertain and must be inferred from local observations?

Existing work on Bayesian learning in repeated games typically sets this question aside by maintaining common knowledge of the interaction structure. Under perfect observability, all players observe every action, and learning concerns only others' behavior within a known environment (Jordan, 1991, 1995; Kalai and Lehrer, 1993a,b). Similarly, under imperfect monitoring, players observe signals of actions, but the signal structure, whether a general distribution or a fixed monitoring graph, is commonly known (Renault and Tomala, 1998, 2004). More generally, in this literature the interaction structure whether global or local is a fixed primitive and common knowledge. This leaves open a more basic question. When the interaction structure itself is unknown, what must agents learn in order to behave optimally? We show that long-run behavior and long-run knowledge are fundamentally different objects. Players may eventually behave exactly as if the network were known while still failing to identify the network itself.

We model this type of uncertainty directly. A network is drawn from a commonly known prior and remains fixed throughout the interaction, but its realization is unknown to the agents. Each agent observes only her local neighborhood and the actions of those with whom she interacts directly. Starting from these local observations, agents form beliefs about the set of feasible networks and update those beliefs through repeated play using Bayes' rule. The object of learning is therefore the interaction structure itself, rather than an exogenous state or the strategies of other players. This allows us to study how decentralized observations shape inference about the broader environment in which agents are embedded.

Two features distinguish this setting. First, the network that agents seek to learn is the same object that determines strategic interactions, local observations, and the paths through which information can spread. Thus the unknown state not only affects behavior, but also constrains the process by which it can be learned. Second, the information used to learn the network is generated endogenously through equilibrium behavior. Agents infer the structure from observed actions, but those actions are chosen by players who are simultaneously making inferences about the same unknown network from their own local observations. As a result, learning depends not only on what agents observe, but also on how equilibrium behavior maps network structures into observable actions. The framework therefore links decentralized observation, strategic interaction, and network learning in a common Bayesian environment.

The environment generates a decentralized learning problem. Agents begin with only local information about the realized network and use observed actions to refine their beliefs about the broader interaction structure. Because observations are local, different agents may learn different things from play. What can be inferred about the network therefore depends on the information conveyed by the actions that agents observe.

Our first result characterizes the convergence of equilibrium behavior. We consider myopic agents who update beliefs about the interaction structure based on locally observed actions. As observations accumulate, the set of networks consistent with an agent's history shrinks monotonically and eventually stabilizes. We show that once learning stabilizes, equilibrium play coincides with a complete-information Nash equilibrium of the realized network. A key feature of this result is that convergence does not require identification of the network. Agents may continue to assign positive probability to multiple interaction structures, yet all remaining structures induce the same payoff-relevant environment and therefore the same best-response problem. Consequently, uncertainty about the interaction structure becomes behaviorally irrelevant even though the structure itself may not be learned. This contrasts with standard Bayesian learning environments, where players form beliefs over others' strategies under global observability and convergence is tied to the accuracy of those beliefs. Here, inference is directed at the interaction structure itself. Because a common prior over networks aligns beliefs about behavior conditional on a candidate structure, local observations are sufficient for equilibrium behavior to converge even when the underlying structure remains uncertain.

While behavioral convergence does not require identification of the interaction structure, full learning requires more. Our second result characterizes when the realized network is identified. The key requirement is that equilibrium behavior be sufficiently informative to distinguish among observationally equivalent structures. If the complete-information equilib-

rium correspondence assigns different actions to a player or to one of her observed neighbors across candidate networks, then these actions separate the competing structures. As observations accumulate, all inconsistent networks are eliminated and the realized network is identified once learning stabilizes.

The result highlights that network identification is ultimately governed not by observation alone, but by the equilibrium mapping from interaction structures into observable behavior. Behavioral convergence requires only that the remaining candidate structures induce the same best-response problem. Identification is stronger: it requires that they induce different observable actions. Thus local observations may be sufficient for correct long-run behavior even when the network remains unknown, but full learning occurs whenever equilibrium behavior separates all structures that are observationally equivalent from the player's perspective.

We illustrate the identification result in the linear-quadratic network game of Ballester et al. (2006), where complete-information equilibrium actions admit a closed-form expression in terms of Katz-Bonacich centrality. These actions aggregate walks of all lengths originating from each player, making the mapping from networks to actions explicit. We show that if two networks differ in the number of walks of any length from a given player, then their equilibrium actions differ for almost all parameter values. Thus, whenever a player or one of her observed neighbors faces observationally indistinguishable networks with distinct walk profiles, equilibrium actions separate them once learning stabilizes. This allows agents to identify the realized network.

Lastly, we study the speed of learning. Since agents refine beliefs by eliminating structures inconsistent with observed behavior, learning concludes in finite time. We first establish a coarse bound that depends only on the size of the candidate class, and then derive a sharper bound that exploits a distinctive feature of the environment. The network is both the object being learned and the structure through which information about it must travel. Because agents observe only locally, information about distant parts of the network can reach them only through a sequence of intermediate observations and actions. As a result, the speed of learning is governed by the distance over which information must propagate. More generally, learning depends not only on the accumulation of observations, but also on how effectively equilibrium behavior transmits information through the interaction structure.

Related literature

A natural benchmark for our paper is the Bayesian learning literature, which primarily focuses on the issue of convergence to Nash equilibrium under repeated play and perfect observability. Jordan (1991) shows that myopic Bayesian play in finite normal-form games converges almost surely to the Nash set, while Jordan (1995) extends this to forward-looking players, showing that play asymptotically converges to the Nash equilibrium of the repeated game. Kalai and Lehrer (1993a) establish the same conclusion when each player's prior assigns positive probability to strategies other players employ during the game, while Kalai and Lehrer (1993b) relax this informational restriction by showing that beliefs need only correctly predict other players' actions on the equilibrium path. We depart from this literature in two ways. First, the object being learned in our game is the interaction network rather than players' strategies. Second, in our model, convergence occurs in finite time.

A related strand examines repeated games with imperfect monitoring, where players observe signals rather than others' actions, and characterizes the set of sustainable equilibrium payoffs under public monitoring (Fudenberg et al., 1994), private monitoring with communication (Compte, 1998; Kandori and Matsushima, 1998), and without communication (Hörner and Olszewski, 2006; Sugaya, 2022). In these models, the monitoring technology, understood as the rule mapping actions into observed signals, is typically taken as a commonly known primitive. Fudenberg and Yamamoto (2010) instead allow players to be uncertain about the monitoring structure itself. Our setting differs along another dimension: the unknown object is the interaction network, which determines both payoffs and local observation, and the signals from which it is learned are equilibrium actions generated endogenously through play.

When the observation structure itself is uncertain, the literature has also considered weaker equilibrium notions. Self-confirming equilibrium requires beliefs about others' strategies to be consistent only with observations along the equilibrium path (Fudenberg and Levine, 1993), a notion Battigalli et al. (2023) extend to network games in which players may not know the interaction network and observe only their realized payoffs. Our agents instead maintain a common prior over candidate networks and update it by Bayes' rule using locally observed equilibrium actions. In the peer-confirming equilibrium of Lipnowski and Sadler (2019), a network encodes which players hold correct conjectures about each other's strategies. We share the view that networks restrict what each player can know about others, but here the network is itself the latent state being learned, while equilibrium behavior provides the endogenous signals through which that learning occurs.

A separate literature studies repeated games with local (rather than global) monitoring, re-

stricting observation or communication links, but assumes this structure is common knowledge (Renault and Tomala, 1998; Ben-Porath and Kahneman, 1996; Laclau, 2014). Closer to our setting, Nava and Piccione (2014) analyze fixed networks where each player observes only her neighborhood, while Wolitzky (2013) considers networks drawn each period from a known distribution; both study which beliefs and structural features sustain cooperation. We likewise examine how beliefs over the observation structure shape long-run play, but focus on whether the structure itself can be learned. We analyze how repeated local observations refine beliefs about the structure over time and affect equilibrium behavior.

Our payoff environment belongs to the class of network games with local spillovers, in which a player's payoff depends only on her own action and the actions of their neighbors (Goyal and Joshi, 2006, 2003; Ballester et al., 2006). A separate body of work introduces incomplete information into such games in static settings. Galeotti et al. (2010) study network games in which agents only know their degree, while de Martí and Zenou (2015) analyze incomplete information about payoff-relevant parameters and characterize Bayesian-Nash equilibria. In a formation setting, Foerster et al. (2021) let some links go unobserved and characterize stable networks when players hold conjectures about the hidden structure. Beliefs there are not refined through play, whereas here repeated observation of actions resolves the uncertainty. Closest to our informational environment is Chaudhuri et al. (2026), who study static network games under network uncertainty when agents know only their local connections. We move beyond this static benchmark by making equilibrium actions observable through repeated play. The network then becomes something agents learn from the signals those actions provide, rather than a fixed object they form beliefs about in a one-shot interaction.

Our paper is also related to the literature on Bayesian learning in networks. Renault and Tomala (2004) and Li and Tan (2020) study how agents learn an unknown state under graph-based monitoring. Macault et al. (2022) study a known routing network with an unknown persistent cost state, distinguishing between learning the state itself and convergence to the complete-information flow. Gale and Kariv (2003), Mueller-Frank (2013), and Huang et al. (2024) analyze how agents infer information from neighbors' actions, while Dasaratha and He (2026) quantify how network structure shapes the efficiency of signal aggregation. We differ on the object of inference: the latent state in our setting is the network itself, and informative signals take the form of endogenous equilibrium actions that reflect agents' positions in interaction structure. Convergence therefore need not take the form of consensus, as in the imitation principle of Mueller-Frank (2013); instead, the limiting outcome is contained in the complete-information Nash equilibrium set of the realized network game.

Finally, our paper is also complementary to the large literature on boundedly rational or non-

Bayesian learning in networks (Bala and Goyal, 1998; Golub and Jackson, 2010; Acemoglu et al., 2010; Dasaratha and He, 2020). That literature typically asks whether decentralized updating aggregates dispersed information. Instead we maintain full Bayesian rationality and study a different friction: players observe only a local subset of equilibrium actions and use those observations to learn an initially unknown interaction structure. Our contribution is to characterize when this local inference process stabilizes, when it identifies the true network, and when the resulting behavior coincides with the complete-information Nash equilibrium of the realized network game.

The rest of the paper is structured as follows. Section 2 sets up the game. Section 3 presents the convergence and network identification results. Section 4 illustrates the results in a linear quadratic game. Section 5 characterizes the speed of learning. Section 6 concludes.

2 A Repeated Game under Network Uncertainty

We study a repeated game of strategic interaction in a finite population. In each period, a player's stage payoff depends on the actions of a subset of other players with whom she interacts, and at the end of the period she observes only the actions of those same players. The interaction structure (or network) is drawn once at the beginning of the game and remains fixed thereafter. It is not common knowledge: each player observes only her own local neighborhood and holds beliefs about the rest of the structure. Over time, players update these beliefs using signals generated by repeated play.

The distinctive feature of the model is that the unknown state is the interaction structure itself. Players learn from signals that take the form of locally observed actions. These signals are equilibrium choices, so both their realization and their informational content are determined endogenously with behavior. A candidate structure is eliminated only when it predicts a different action by an observed player and thus the informativeness of observation depends on equilibrium behavior. Moreover, a neighbor's action conveys information about the network through the inferences she has drawn from her own local observations. Information about distant parts of the structure consequently reaches a player through the optimizing behavior of those along the path. This combination of strategic interaction and limited observation produces decentralized learning. Public-observability environments give all players access to a common signal or the full action profile after each period. In the present model, each player's information depends on her position in the interaction structure and on the actions she observes over the course of play.

Primitives

There is a finite set of players $N = \{1, \ldots, n\}$. For each $i \in N$, let $A_i \subseteq \mathbb{R}_+$ be a nonempty action set. For any subset $S \subseteq N$, let $A_S := \prod_{i \in S} A_i$, and write $A := A_N$.

Time is discrete and indexed by $\tau = 0, 1, 2, \ldots$. In each period $\tau \geq 1$, players choose actions simultaneously. Let $a_i^\tau \in A_i$ denote player i's action in period τ, and let $a^\tau = (a_i^\tau)_{i \in N} \in A$ denote the resulting action profile.

Interaction Structure

An *interaction structure* is a profile $\mathbf{g} = \big(N_i(\mathbf{g})\big)_{i \in N}$, where for each $i \in N$, $N_i(\mathbf{g}) \subseteq N \setminus \{i\}$. The set $N_i(\mathbf{g})$ represents the players with whom i interacts directly. The interaction structure determines both whose actions affect player i's payoff and whose actions player i observes at the end of each period.[1] This formulation captures environments in which the same network of local relationships determines both strategic spillovers and the information available through repeated observation.

Let $\mathcal{G}_n$ be a finite set of feasible interaction structures. At $\tau = 0$, nature draws $\mathbf{g} \in \mathcal{G}_n$ according to a commonly known prior $p \in \Delta(\mathcal{G}_n)$ with full support, and the realized structure remains fixed thereafter.

After $\mathbf{g}$ is realized, player i observes only her local interaction set $N_i(\mathbf{g})$. Her initial belief at $\tau = 1$ is therefore the prior p conditioned on the set of structures consistent with $N_i(\mathbf{g})$, that is, $p(\cdot \mid N_i(\mathbf{g}))$.

This informational environment is closely related to Chaudhuri et al. (2026), who study static network games under network uncertainty in which agents observe their local connections and update their beliefs about the rest of the network using Bayes' rule. We adopt the same informational premise but ask a different question. There, beliefs rest on local information alone; here, repeated play allows equilibrium actions to be locally observable, and these actions become the signals through which the network is progressively learned. Our focus is therefore not on equilibrium under fixed network uncertainty, but on whether and how repeated local observation resolves it over time.

Stage Payoffs

For each player i and network structure $\mathbf{g} \in \mathcal{G}_n$, stage payoffs are described by a *local spillover function*

$$v_i^{\mathbf{g}} : A_i \times A_{N_i(\mathbf{g})} \to \mathbb{R},$$

[1] More generally, one could distinguish between an influence set $N_i^u(\mathbf{g})$, which determines payoff relevance, and an observation set $N_i^z(\mathbf{g})$, which determines observed signals. We focus on the benchmark case in which these sets coincide.

with $v_i^{\mathbf{g}} = v_i^{\mathbf{g}'}$ whenever $N_i(\mathbf{g}) = N_i(\mathbf{g}')$. Given an action profile $a \in A$, player i's payoff is

$$u_i(a, \mathbf{g}) = v_i^{\mathbf{g}}\big(a_i, a_{N_i(\mathbf{g})}\big).$$

Thus, player i's payoff depends only on her own action and the actions of her neighbors. In particular,

$$u_i(a, \mathbf{g}) = u_i(a, \mathbf{g}') \quad \text{whenever } N_i(\mathbf{g}) = N_i(\mathbf{g}').$$

The formulation nests standard linear-quadratic network games(Ballester et al., 2006), and more generally captures local-spillover environments in which payoff-relevant interactions are mediated through direct links (Bramoullé et al., 2014). In R&D collaboration networks, for example, a firm's research payoff may depend on the R&D effort, cost reductions, or innovation outcomes of its collaborators, as in models of collaboration networks in oligopoly and R&D network formation (Goyal and Joshi, 2003, 2006; Billand et al., 2019, 2026). Similar local payoff externalities arise in network games with peer effects, workplace interactions, and local public-good environments, where an agent's payoff-relevant environment is determined by the behavior of those with whom she directly interacts (Jackson and Zenou, 2015; Calvó-Armengol et al., 2009; Lindquist et al., 2024).

We impose the following standard regularity conditions on stage payoffs.

Assumption 1. *For every player $i \in N$ and every network structure $\mathbf{g} \in \mathcal{G}_n$: (i) A_i is compact and convex, (ii) $v_i^{\mathbf{g}}$ is continuous, and (iii) for every $a_{N_i(\mathbf{g})} \in A_{N_i(\mathbf{g})}$, the map $a_i \mapsto v_i^{\mathbf{g}}(a_i, a_{N_i(\mathbf{g})})$ is concave on A_i.*

Local Observation and Private Histories

At the end of each period, player i does not observe the full action profile. Instead, she observes only her own action and the actions of the players in her neighborhood $N_i(\mathbf{g})$.[2]

Formally, player i's observation is given by a signal function

$$z_i : A \times \mathcal{G}_n \to Z_i, \quad z_i(a, \mathbf{g}) = \big(a_i, (a_j)_{j \in N_i(\mathbf{g})}\big),$$

where $Z_i := \bigcup_{S \subseteq N \setminus \{i\}} (A_i \times A_S)$. If a^τ is the action profile in period τ, the realized signal is $z_i^\tau = z_i(a^\tau, \mathbf{g})$.

[2]This relates to Bayesian learning in social networks, where agents observe their neighbors' actions to learn an underlying state (Gale and Kariv, 2003; Mueller-Frank, 2013; Huang et al., 2024). There the network channels information about an exogenous state. Here it is the unknown state itself: observed actions are informative because the structure that generates them is the one that determines payoffs.

Player i's private history at the beginning of period $\tau \geq 1$ is

$$h_i^\tau = \begin{cases} N_i(\mathbf{g}) & \text{if } \tau = 1, \\ \big(N_i(\mathbf{g}), z_i^1, \ldots, z_i^{\tau-1}\big) & \text{if } \tau \geq 2. \end{cases}$$

Thus, a private history consists of the observed neighborhood together with the sequence of locally observed signals. In particular, at $\tau = 1$, the history reduces to the observed neighborhood.

The set of all possible private histories for player i at period τ is

$$\mathcal{H}_i^\tau = \bigcup_{S \subseteq N \setminus \{i\}} \{S\} \times (A_i \times A_S)^{\tau-1}.$$

Under a realized network structure $\mathbf{g}$ and a profile of past play, player i's realized history satisfies $h_i^\tau \in \mathcal{H}_i^\tau$.

Strategies and Induced Play

A pure strategy for player i is a sequence $\sigma_i = (\sigma_i^\tau)_{\tau \geq 1}$, where

$$\sigma_i^\tau : \mathcal{H}_i^\tau \to A_i$$

assigns an action to each period-τ private history. A pure-strategy profile is denoted by $\sigma = (\sigma_i)_{i \in N}$.

Given a network structure $\mathbf{g} \in \mathcal{G}_n$ and a strategy profile σ, play is defined recursively. For each player i, let

$$h_i^1(\mathbf{g}, \sigma) := N_i(\mathbf{g}).$$

For each period $\tau \geq 1$, define

$$a_i^\tau(\mathbf{g}, \sigma) = \sigma_i^\tau\big(h_i^\tau(\mathbf{g}, \sigma)\big), \qquad a^\tau(\mathbf{g}, \sigma) = \big(a_k^\tau(\mathbf{g}, \sigma)\big)_{k \in N} \qquad ,$$

and

$$z_i^\tau(\mathbf{g}, \sigma) = z_i\big(a^\tau(\mathbf{g}, \sigma), \mathbf{g}\big).$$

The private history then evolves according to

$$h_i^{\tau+1}(\mathbf{g}, \sigma) = \big(N_i(\mathbf{g}), z_i^1(\mathbf{g}, \sigma), \ldots, z_i^\tau(\mathbf{g}, \sigma)\big).$$

Thus, for each $\mathbf{g}$ and σ, the model induces a unique sequence

$$\left\{(h^\tau(\mathbf{g},\sigma), a^\tau(\mathbf{g},\sigma), z^\tau(\mathbf{g},\sigma))\right\}_{\tau\geq 1}.$$

Distinct interaction structures $\mathbf{g}, \mathbf{g}' \in \mathcal{G}_n$ may generate identical private histories for player i under σ, and are therefore indistinguishable from her perspective. This observation motivates the following definition.

Definition 1. *Fix a strategy profile σ, a player i, and a period $\tau \geq 1$. Two interaction structures $\mathbf{g}, \mathbf{g}' \in \mathcal{G}_n$ are observationally equivalent for player i at period τ under σ, written $\mathbf{g} \sim_{i,\tau} \mathbf{g}'$, iff $h_i^\tau(\mathbf{g},\sigma) = h_i^\tau(\mathbf{g}',\sigma)$. We denote by $[\mathbf{g}]_{i,\tau}$ the corresponding equivalence class.*[3]

Beliefs

Fix a player i and a period $\tau \geq 1$. Since player i's private history at the beginning of period τ is determined by the realized interaction structure together with play in periods $1, \ldots, \tau-1$, the prior p and the strategy profile σ induce a probability distribution over $\mathcal{H}_i^\tau$. For any history $h_i^\tau \in \mathcal{H}_i^\tau$, define

$$P_{i,\sigma}(h_i^\tau) := \sum_{\mathbf{g}\in\mathcal{G}_n} p(\mathbf{g})\, \mathbb{I}\{h_i^\tau(\mathbf{g},\sigma) = h_i^\tau\}.$$

Thus, $P_{i,\sigma}(h_i^\tau)$ is the ex ante probability that player i's realized private history at the beginning of period τ is h_i^τ.

The strategy profile σ therefore induces, for each player i and each period τ, a posterior belief over interaction structures at every history in $\mathcal{H}_i^\tau$. Let the induced belief system be denoted by

$$\mu_\sigma = \left(\mu_{i,\sigma}^\tau\right)_{i\in N,\ \tau\geq 1},$$

where

$$\mu_{i,\sigma}^\tau(\,\cdot \mid h_i^\tau) \in \Delta(\mathcal{G}_n)$$

denotes player i's posterior at history $h_i^\tau \in \mathcal{H}_i^\tau$. At every history such that $P_{i,\sigma}(h_i^\tau) > 0$, these beliefs are obtained by Bayes' rule:

$$\mu_{i,\sigma}^\tau(\mathbf{g} \mid h_i^\tau) = \frac{p(\mathbf{g})\, \mathbb{I}\{h_i^\tau(\mathbf{g},\sigma) = h_i^\tau\}}{P_{i,\sigma}(h_i^\tau)}, \qquad \text{for all } \mathbf{g} \in \mathcal{G}_n.$$

At histories $h_i^\tau \in \mathcal{H}_i^\tau$ such that $P_{i,\sigma}(h_i^\tau) = 0$, beliefs may be specified arbitrarily.

[3] The dependence of the relation $\sim_{i,\tau}$ and $[\mathbf{g}]_{i,\tau}$ on σ is suppressed whenever the strategy profile is fixed.

Define the path of play of player i under σ by

$$\mathcal{P}_i(\sigma) := \bigcup_{\tau \geq 1} \{h_i^\tau \in \mathcal{H}_i^\tau : P_{i,\sigma}(h_i^\tau) > 0\},$$

and define

$$\mathcal{P}(\sigma) := \bigcup_{i \in N} \mathcal{P}_i(\sigma)$$

as the path of play induced by σ.

Because the prior p has full support, for every $\mathbf{g} \in \mathcal{G}_n$, every player i, and every period τ, the induced history $h_i^\tau(\mathbf{g}, \sigma)$ satisfies

$$P_{i,\sigma}\big(h_i^\tau(\mathbf{g}, \sigma)\big) \geq p(\mathbf{g}) > 0.$$

Hence, every history generated by some interaction structure under σ lies on the path of play.

Equilibrium

Players are myopic in the sense that, in each period, they choose actions to maximize current expected payoff given their posterior beliefs. This corresponds to the dynamic counterpart of Jordan (1991), where the unknown state is the network structure $\mathbf{g}$ and beliefs are updated through locally observed signals z_i^τ. In contrast to models with forward-looking behavior (e.g., Kalai and Lehrer (1993a); Jordan (1995)), repeated interaction here serves to generate information rather than intertemporal incentives. We focus on myopic behavior in order to isolate learning generated through observed actions. Under myopia, actions reflect agents' current beliefs about the interaction structure rather than attempts to influence future beliefs or continuation play.

Definition 2. *A strategy profile σ^* is a myopic Bayesian equilibrium if, for every player i, every period $\tau \geq 1$, and every history $h_i^\tau \in \mathcal{P}_i(\sigma^*)$,*

$$\sigma_i^{*\tau}(h_i^\tau) \in \arg\max_{a_i \in A_i} \sum_{\mathbf{g} \in \mathcal{G}_n} \mu_{i,\sigma^*}^\tau(\mathbf{g} \mid h_i^\tau)\, u_i\left(a_i,\, \sigma_{-i}^{*\tau}\big(h_{-i}^\tau(\mathbf{g}, \sigma^*)\big);\, \mathbf{g}\right),$$

where $h_{-i}^\tau(\mathbf{g}, \sigma^*) = \big(h_j^\tau(\mathbf{g}, \sigma^*)\big)_{j \neq i}$.

The equilibrium condition constrains behavior only on the path of play $\mathcal{P}(\sigma^*)$. At histories $h_i^\tau \notin \mathcal{P}_i(\sigma^*)$, the strategy $\sigma_i^{*\tau}(h_i^\tau)$ is unconstrained, and two strategy profiles that agree on

$\mathcal{P}(\sigma^*)$ are equivalent. Under Assumption 1, a myopic Bayesian equilibrium exists[4].

The myopic equilibrium abstracts from intertemporal incentives and, in particular, from the use of actions to influence future beliefs or continuation play. When players maximize current expected payoffs, they form beliefs given current action. Given these beliefs, their actions depend only on what they have inferred about the interaction structure, and so convey that information to those who observe them. This is essential in our setting, where players learn the structure precisely by observing their neighbors' actions. Under forward-looking behavior, actions could instead be chosen to influence others' beliefs or subsequent play, and would no longer serve as reliable signals of what each player has learned. This feature is central to our analysis, as it isolates the informational content of play and underlies our identification and convergence results. In this sense, the equilibrium concept plays a role analogous to the best-response dynamics studied in the Bayesian learning literature (Jordan, 1991, 1995; Kalai and Lehrer, 1993b,a), adapted to an environment in which the interaction and observation structure is unknown and the signals used to infer it are generated endogenously through equilibrium behavior. The concept is also related to self-confirming equilibrium (Battigalli et al., 2023; Fudenberg and Levine, 1993), with the distinction that beliefs here are fully disciplined by a common prior and Bayes' rule at all on-path histories.[5]

3 Convergence and Identification

This section studies behavior and information once learning stabilizes. Two distinct findings arise. First, agents may eventually behave as if they know the true interaction structure, in the sense that equilibrium play coincides with the complete-information Nash equilibrium under the realized structure. This need not imply that they identify the structure itself. Second, agents may or may not be able to identify the realized interaction structure. We now show that the first property holds quite generally, whereas the second depends on whether equilibrium actions are sufficiently informative to distinguish between observationally equivalent structures.

Fix a myopic Bayesian equilibrium σ^*. Throughout this section, all histories, observational equivalence classes, and induced actions are defined relative to σ^*. For each player $i \in N$,

[4]In each period, beliefs are pinned down by Bayes' rule on the path of play, and the per-period objective is continuous in a_i and concave on the compact, convex set A_i, so a maximizer exists.

[5]With forward-looking players, actions could manipulate beliefs or sustain history-dependent play. This introduces signaling and punishment motives that confound their informational role. Section 3.1 shows our results extend under restrictions that remove such effects, so myopia serves as a tractable benchmark that isolates learning.

period $\tau \geq 1$, and network structure $\mathbf{g} \in \mathcal{G}_n$, let $[\mathbf{g}]_{i,\tau}$ denote the observational-equivalence class of $\mathbf{g}$ for player i at period τ, that is,

$$[\mathbf{g}]_{i,\tau} = \{\mathbf{g}' \in \mathcal{G}_n : h_i^\tau(\mathbf{g}', \sigma^*) = h_i^\tau(\mathbf{g}, \sigma^*)\}.$$

Because the prior p has full support, at any on-path history $h_i^\tau(\mathbf{g}, \sigma^*)$, the support of player i's posterior coincides with $[\mathbf{g}]_{i,\tau}$.

Throughout, for each $\mathbf{g} \in \mathcal{G}_n$, let $NE(\mathbf{g}) \subseteq A$ denote the set of pure strategy Nash equilibria of the complete information stage game induced by $\mathbf{g}$ and let $a^c(\mathbf{g})$ be a generic element of $NE(\mathbf{g})$ with $a_i^c(\mathbf{g})$ denoting its i^{th} component. For each player $i \in N$, write $NE_i(\mathbf{g}) := \{a_i(\mathbf{g}) : a(\mathbf{g}) \in NE(\mathbf{g})\}$ for the projection of the complete-information Nash equilibrium set under $\mathbf{g}$ onto player i's action space.

3.1 Convergence to Stage-Game Equilibrium

Observation in this environment is local: players observe only the actions of those in their interaction neighborhoods. We show that, once learning stabilizes, this local information is nonetheless sufficient to align equilibrium behavior with the complete-information benchmark.

We begin by formalizing when learning ends.

Definition 3. *Learning ends at period τ^* if, for all $\tau \geq \tau^*$,*

$$[\mathbf{g}]_{i,\tau} = [\mathbf{g}]_{i,\tau^*} \qquad \textit{for all } i \in N \textit{ and } \mathbf{g} \in \mathcal{G}_n.$$

After τ^*, no player rules out further interaction structures on the basis of observed play. This does not require that players identify the true structure, only that the equilibrium path generates no additional eliminations. In particular, if $\mathbf{g}' \in [\mathbf{g}]_{i,\tau}$ for some $\tau \geq \tau^*$, then $\mathbf{g}' \in [\mathbf{g}]_{i,\tau+1}$, so $\mathbf{g}$ and $\mathbf{g}'$ generate the same next-period signal for player i.

Lemma 1. *For every myopic Bayesian equilibrium σ^*, there exists a finite period τ^* such that learning ends, i.e. for all $i \in N$ and $\mathbf{g} \in \mathcal{G}_n$,*

$$[\mathbf{g}]_{i,\tau} = [\mathbf{g}]_{i,\tau^*} \qquad \forall \tau \geq \tau^*.$$

This proof and all other proofs are relegated to the Appendix. We now state the main result. It shows that once learning stabilizes, any residual uncertainty becomes payoff-irrelevant along the equilibrium path.

Theorem 1. *Let $\mathbf{g} \in \mathcal{G}_n$ be the realized interaction structure. If learning ends at τ^* under σ^*, then*

$$a^\tau(\mathbf{g}, \sigma^*) \in NE(\mathbf{g}) \qquad \forall \tau \geq \tau^*.$$

The logic is straightforward. Once learning ends, all interaction structures in $[\mathbf{g}]_{i,\tau^*}$ generate the same payoff-relevant local observations at the realized history. Residual uncertainty is therefore irrelevant for player i's best-response problem.

Theorem 1 reflects a distinctive feature of the environment. The interaction structure is simultaneously the object of inference, the determinant of strategic interactions, and the determinant of local observations. Players therefore learn about the very structure that governs both their incentives and the information available to them. As a result, equilibrium behavior can converge even when the network itself is not identified, since only the payoff-relevant aspects of the interaction structure need be learned. The same feature also shapes the learning process itself. As we show below, the speed of learning depends on how information propagates through the interaction structure that agents are trying to infer.

Moreover, this theorem is our core behavioral implication. Even if players may fail to identify the realized interaction structure ($|[\mathbf{g}]_{i,\tau^*}|$ may exceed one) they behave as if it were common knowledge. Uncertainty about the network structure thus becomes behaviorally irrelevant in the long run. All of this rests on myopia. Were players forward-looking, an action could be chosen for its effect on others' beliefs or on future play, and the correspondence between learning and complete-information behavior in Theorem 1 need no longer hold. We take this up at the end of the section, and find that it is restored under a mild condition.

The theorem departs from classical Bayesian learning results such as Kalai and Lehrer (1993a) and Jordan (1995), in two ways. First, we replace global observability with strictly local observability. Second, we obtain finite-time rather than asymptotic convergence.

The result also contrasts with the imitation principle in Mueller-Frank (2013). In that setting, agents observe neighbors' actions to learn about a common payoff-relevant state. Their payoffs, however, do not directly depend on those neighbors' actions. Since agents face the same decision problem, an agent who observes a neighbor can copy the neighbor's action, and this disciplines long-run disagreement. Our model differs because the realized network is the payoff-relevant state. Neighbors' actions enter payoffs through local strategic spillovers, so players in different network positions generally face different best-response problems. Therefore, learning need not produce imitation or consensus. Once learning stabilizes, play enters the set of complete-information Nash equilibria of the realized network game. Unlike the consensus outcome under the imitation principle, this profile is generally

heterogeneous across players, since agents in different network positions face different best-response problems. Moreover, note that when the stage game admits multiple equilibria, the realized profile need not be unique, but it remains a complete-information Nash equilibrium of the realized network in every period.

Consider, then, forward-looking play. Suppose σ^* is a perfect Bayesian equilibrium. Let τ^* denote the stabilization time from Lemma 1. Although $[g]_{i,\tau}$ remains constant for all $\tau \geq \tau^*$, a deviation may affect continuation values through its impact on beliefs or prescribed future play, even if it does not alter the observational-equivalence class along the equilibrium path. Residual uncertainty within $[g]_{i,\tau^*}$ may therefore be irrelevant for stage payoffs but relevant for continuation values, so equilibrium behavior need not coincide with the complete-information benchmark.

Remark. *With forward-looking players, Theorem 1 continues to hold whenever unilateral deviations do not affect posterior beliefs or continuation play.*

This is precisely the feedback that myopia targets. Suppose that for all $\tau \geq \tau^*$, continuation strategies depend only on $[g]_{i,\tau^*}$ and unilateral deviations do not affect posterior beliefs or continuation play, then continuation payoffs are invariant to deviations. Each player's problem reduces to a static best-response problem given beliefs over $[g]_{i,\tau^*}$, and

$$a_i^\tau(g, \sigma^*) \in \mathrm{BR}_i(a_{-i}^\tau(g, \sigma^*); g) \quad \text{for all } i,$$

so that $a^\tau(g, \sigma^*)$ coincides with the complete-information Nash equilibrium.

3.2 Perfect Learning

Theorem 1 shows that long-run behavior coincides with the complete-information benchmark. We now ask when a player fully identifies the realized network.

Definition 4. *Fix* $\mathbf{g} \in \mathcal{G}_n$. *Learning is* perfect *for player* i *if*

$$[\mathbf{g}]_{i,\tau^*} = \{\mathbf{g}\}.$$

Perfect learning is stronger than behavioral convergence. It requires that equilibrium behavior be sufficiently informative to eliminate all incorrect interaction structures, and not merely to pin down equilibrium actions. Identification may depend on a single observed action, but it may also arise from several actions jointly even when none identifies the network

on its own. To accommodate this possibility, for any nonempty $S \subseteq N$ and action profile $a \in A$, write $a_S := (a_j)_{j \in S} \in A_S$. Define the joint projection of the complete-information Nash equilibrium set onto the actions of players in S by

$$NE_S(\mathbf{g}) := \{a_S : a \in NE(\mathbf{g})\}.$$

For a singleton $S = \{j\}$, this reduces to the previously defined set $NE_j(\mathbf{g})$.

Theorem 1 establishes that correct behavior does not require identification of the network. We now ask when the stronger outcome obtains. Under what conditions do local observations allow agents not only to behave as if the interaction structure were known, but also to learn the realized structure itself? The answer depends on whether equilibrium behavior is sufficiently informative to distinguish among the networks that remain observationally equivalent.

Theorem 2. *Fix $\mathbf{g} \in \mathcal{G}_n$ and suppose learning ends at τ^* under σ^*. Fix $i \in N$. If there exists a nonempty set $S \subseteq \{i\} \cup N_i(\mathbf{g})$ such that*

$$NE_S(\mathbf{g}') \cap NE_S(\mathbf{g}'') = \emptyset$$

for every distinct $\mathbf{g}', \mathbf{g}'' \in \mathcal{G}_n$, then learning is perfect for player i.[6]

Theorem 2 strengthens behavioral convergence to full identification. The vector of actions observed by player i can refine her beliefs by ruling out networks inconsistent with it. Identification requires more: that this refinement eventually eliminates every incorrect structure. The theorem shows that it is enough for some collection of actions that i observes, including possibly her own action, to separate the candidate networks jointly. In particular, the previous single-action conditions are special cases: if $S = \{i\}$, player i identifies the structure from her own action, while if $S = \{j\}$ for some $j \in N_i(\mathbf{g})$, she identifies it from the action of a single neighbor. More generally, two or more observed actions may identify the network jointly even when none of them does so individually.

This extension matters because equilibrium behavior can encode network structure through combinations of actions. For example, it may be that $NE_j(\mathbf{g}') \cap NE_j(\mathbf{g}'') \neq \emptyset$ and $NE_k(\mathbf{g}') \cap$

[6] This separation condition is a sufficient identification condition rather than a restriction. Under uniqueness, it reduces to injectivity of the map $\mathbf{g}' \mapsto a_S^c(\mathbf{g}')$. With multiple equilibria, $NE_S(\mathbf{g})$ is the joint projection of the equilibrium set and need not equal the Cartesian product of the individual projections $\prod_{j \in S} NE_j(\mathbf{g})$. In environments with rich action spaces, observed actions may reveal beliefs or informational states; see Arieli and Mueller-Frank (2017) and Molavi et al. (2018). In the linear-quadratic network game (Ballester et al., 2006), equilibrium actions are generically one-to-one in the network's walk profile, as we show below.

$NE_k(\mathbf{g}'') \neq \emptyset$ for some observed neighbors j and k, while

$$NE_{\{j,k\}}(\mathbf{g}') \cap NE_{\{j,k\}}(\mathbf{g}'') = \emptyset.$$

Neither action then identifies the network by itself, but their joint realization does. Perfect learning therefore depends on the informational content of the locally observed action profile, rather than necessarily on a single player's behavior.

Together, Theorems 1 and 2 define the limits of decentralized learning here. Local observations suffice for correct long-run behavior: once learning stabilizes, agents act as if the interaction structure were known, despite incomplete network knowledge. However, identifying the structure itself requires a stronger condition, namely that equilibrium behavior be informative enough to distinguish observationally equivalent structures. Thus behavior and knowledge may diverge, since agents can act optimally relative to the true structure without ever learning it. This distinction reflects a fundamental feature of environments with endogenous and locally generated information: learning is constrained not only by what is observed, but also by how equilibrium behavior encodes underlying primitives. In this sense, information along the path of play is determined jointly by the interaction structure and the equilibrium mapping from structure to actions. The linear-quadratic application below provides an example in which players converge to the complete-information Nash equilibrium while some players nevertheless fail to identify the realized network.

4 Linear Quadratic Network Games

We illustrate the general convergence results for the workhorse linear-quadratic network game to provide a transparent interpretation of both convergence and identification. In this environment, equilibrium actions admit a closed-form representation that makes explicit how local interactions aggregate information about the global structure.

4.1 Environment

For each player $i \in N$, let $A_i = \mathbb{R}$[7] and define stage payoffs by

$$u_i(a, \mathbf{g}) = a_i - \frac{1}{2}a_i^2 + \lambda a_i \sum_{j \in N} g_{ij}(\mathbf{g}) a_j. \tag{1}$$

where $g_{ij}(\mathbf{g}) = 1$ if $j \in N_i(\mathbf{g})$, and 0 otherwise. This is the linear-quadratic network game of Ballester et al. (2006), among the most widely studied models in the networks literature. Player i chooses an action that trades off a quadratic cost against complementarities with the actions of her neighbors. The interaction structure affects payoffs only through the set of neighbors $N_i(\mathbf{g})$, so the local-spillovers property of the general model is preserved.

Define $G(\mathbf{g}) = [g_{ij}(\mathbf{g})]_{1 \le i,j \le n}$ as the adjacency matrix representation of the interaction structure. Let $\bar{\rho} := \max_{\mathbf{h} \in \mathcal{G}_n} \rho(G(\mathbf{h}))$, where $\rho(G(\mathbf{h}))$ is the spectral radius of the corresponding adjacency matrix, and suppose $0 < \lambda < 1/\bar{\rho}$. Then, for any $\mathbf{g} \in \mathcal{G}_n$, the unique pure Nash equilibrium of the complete-information stage game is given by

$$a^c(\mathbf{g}) = \big(I - \lambda G(\mathbf{g})\big)^{-1}\mathbf{1}, \tag{2}$$

which corresponds to the Katz–Bonacich centrality vector of the network $\mathbf{g}$.

We now connect this structure to the general results. Let $\mathbf{g}$ be the realized interaction structure, and suppose learning ends at period τ^* under σ^*. By Theorem 1, equilibrium play converges to the complete-information Nash equilibrium under $\mathbf{g}$. Combining this with (2) yields:

Corollary 1. *Suppose payoffs are given by* (1)*. If learning ends at period τ^* under σ^* and the realized interaction structure is $\mathbf{g}$, then the equilibrium action profile coincides with the corresponding Katz–Bonacich centrality vector of the network $\mathbf{g}$, i.e.*

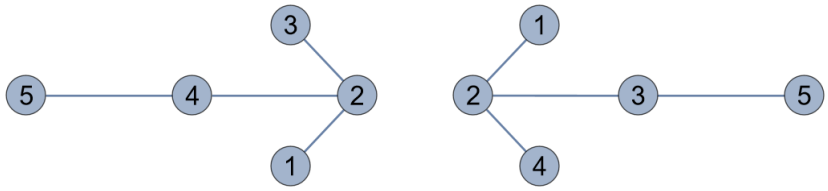
$$a_i^{\tau}(\mathbf{g}, \sigma^*) = \Big[\big(I - \lambda G(\mathbf{g})\big)^{-1}\mathbf{1}\Big]_i \qquad \forall i \in N,\ \forall \tau \ge \tau^*.$$

The corollary shows that even though players initially observe only their local neighborhoods and must infer the rest of the network indirectly, equilibrium play aggregates information across the network. Once learning stabilizes, each player behaves exactly as if the entire

[7] Although this specification does not satisfy Assumption 1(i), the results of Section 3 remain applicable conditional on a myopic Bayesian equilibrium, since compactness is not otherwise used in their proofs. Assumption 1(i) is imposed as part of sufficient conditions for equilibrium existence in the general model. In the present linear-quadratic environment, the complete-information stage game has a unique Nash equilibrium under the parameter restriction imposed below.

interaction structure were known. Any remaining uncertainty is confined to structures that are observationally and behaviorally indistinguishable.

Finally, the following example shows that information may be learned locally without being transmitted downstream through equilibrium behavior, which also show that the convergence of a Nash equilibrium may not ensure perfect learning.



(a) g_a (b) g_b

Figure 1: The two networks are isomorphic under the permutation that exchanges players 3 and 4 while fixing players 1, 2, and 5.

Consider the network class $\mathcal{G}^{iso} = \{g_a, g_b\}$ in Figure 1, and suppose that g_a is realized. The two networks are isomorphic under the permutation that exchanges players 3 and 4 while fixing players 1, 2, and 5. Player 1 has the same neighborhood in both networks, $N_1(g_a) = N_1(g_b) = \{2\}$, so initially $[g_a]_{1,1} = \{g_a, g_b\}$. Moreover, because the isomorphism fixes player 2, the complete-information equilibrium action of player 2 is the same under the two networks: $a_2^c(g_a) = a_2^c(g_b)$. Hence the only action observed by player 1 does not distinguish the two structures. Player 1's equivalence class never shrinks below $\{g_a, g_b\}$, so her posterior remains non-degenerate in every period. She plays her complete-information action throughout.

4.2 Identification via Network Walks

The centrality representation also clarifies why equilibrium actions can identify the underlying interaction structure. Expanding (2) as a Neumann series, we obtain

$$a^c(\mathbf{g}) = \sum_{k=0}^{\infty} \lambda^k G(\mathbf{g})^k \mathbf{1}. \tag{3}$$

For each player i and $k \geq 0$, define

$$W_i^k(\mathbf{g}) := \left[G(\mathbf{g})^k \mathbf{1}\right]_i.$$

Because $G(\mathbf{g})$ is the adjacency matrix, $W_i^k(\mathbf{g})$ counts the number of walks of length k that originate at player i. Thus,

$$a_i^c(\mathbf{g}) = \sum_{k=0}^{\infty} \lambda^k W_i^k(\mathbf{g}),$$

so equilibrium actions encode the entire walk structure of the network around each player.

This is where behavior meets identification. Player i's action aggregates walks of every length, so it reflects the network's higher-order structure, not just her direct links. Two networks that differ at any walk length must then differ in her action, for almost every λ[8]. Behavior separates them, and identification follows.

Proposition 1. *Fix $i \in N$ and two networks $\mathbf{g}, \mathbf{g}' \in \mathcal{G}_n$. Suppose there exists $k \geq 1$ such that $W_i^k(\mathbf{g}) \neq W_i^k(\mathbf{g}')$. Then $a_i^c(\mathbf{g}) \neq a_i^c(\mathbf{g}')$ for almost all values of λ satisfying $0 < \lambda < 1/\max\{\rho(G(\mathbf{g})), \rho(G(\mathbf{g}'))\}$.*

The intuition follows directly from (3). Equilibrium actions aggregate walk counts of all lengths. If two networks differ at any walk length, then their corresponding power series generically differ, and hence define distinct equilibrium actions for almost all admissible values of λ.

An additional implication of Proposition 1 is how sharply equilibrium actions distinguish networks in the linear-quadratic game. Although each network yields a unique Nash equilibrium, for almost all λ, networks with different walk profiles generate distinct equilibrium action profiles. Hence, equilibrium behavior is generically one-to-one in the network's walk profile. This reflects the model's core structure: equilibrium actions encode the network via walks, yielding a strong form of behavioral identifiability and the following corollary.

Corollary 2. *Fix $\mathbf{g} \in \mathcal{G}_n$ and $i \in N$. Suppose some $j \in \{i\} \cup N_i(\mathbf{g})$ can distinguish every pair of interaction structures in $\mathcal{G}_n$ through its walk profile, i.e., $\mathbf{g}' \neq \mathbf{g}'' \in \mathcal{G}_n$ implies $W_j^k(\mathbf{g}') \neq W_j^k(\mathbf{g}'')$ for some $k \geq 1$. Then, if learning terminates at period τ^* under σ^*, player i achieves perfect learning for almost all admissible values of λ.*

The linear-quadratic environment provides a concrete illustration of the general results. Information about the network propagates through walks, and equilibrium actions aggregate these effects via the Katz–Bonacich representation; once learning stabilizes, behavior reflects the entire upstream network relevant for payoffs. This is Theorem 1, local observation and

[8] In our context "almost all λ" means for all admissible λ outside a set of Lebesgue measure zero. In each of our results this exceptional set is finite and it consists of the roots of a polynomial in λ so the property holds for all but finitely many values.

equilibrium response suffice to recover the payoff-relevant structure. Proposition 1 adds that identification turns on whether these effects uniquely encode the network. When walk profiles differ, actions separate the structures for almost all λ, and learning is perfect. Convergence thus reflects how payoff-relevant influence propagates through the network; identification, whether that influence generates distinct walk profiles.

Note that the linear-quadratic utility is used only for the closed-form Katz–Bonacich representation and the walk-based identification result; convergence itself does not require interior equilibria. It applies to any network game with local payoff spillovers, including the linear-best-reply class of Bramoullé et al. (2014), which comprises quadratic games, local public-good games and R&D networks, with multiplicity in equilibria and corner solutions. In all of these, once learning terminates under a myopic Bayesian equilibrium, the realized profile is a complete-information Nash equilibrium of the realized network. Corner solutions and multiplicity therefore do not break the bridge from learning to complete-information behavior.

5 Speed of Learning

Section 3 established two conclusions. First, for any fixed myopic Bayesian equilibrium, each player's observational-equivalence class shrinks only through elimination and therefore stabilizes in finite time. Second, once learning stabilizes, equilibrium play coincides with the complete-information Nash equilibrium of the realized interaction structure. We now ask a natural follow up question: *how long can the learning process last?* Unlike most Bayesian learning environments, learning here proceeds through endogenous equilibrium actions rather than exogenously provided signals. The speed of learning therefore depends not only on the accumulation of observations, but also on how effectively information is transmitted through behavior.

Fix an arbitrary myopic Bayesian equilibrium σ^*. By Lemma 1, the earliest period at which learning ends under σ^* is well defined. Denote this time by

$$\tau^*(\sigma^*) := \min\Big\{\tau \geq 1 : [\mathbf{g}]_{i,s} = [\mathbf{g}]_{i,\tau} \text{ for all } s \geq \tau, i \in N,\ \mathbf{g} \in \mathcal{G}_n\Big\}.$$

Two forces determine $\tau^*(\sigma^*)$. The first is purely combinatorial and uses only the finiteness of $\mathcal{G}_n$. The second reflects the topology of local observation: information propagates only along directed paths in the interaction structure, so that learning speed is governed by the depth of these paths.

5.1 A Coarse Bound from Finite Refinements

For each player i and period $\tau \geq 1$, let

$$\Pi_i^\tau(\sigma^*) := \Big\{ [\mathbf{g}]_{i,\tau} : \mathbf{g} \in \mathcal{G}_n \Big\}, \qquad b_i^\tau(\sigma^*) := \big|\Pi_i^\tau(\sigma^*)\big|.$$

The collection $\Pi_i^\tau(\sigma^*)$ is the partition of $\mathcal{G}_n$ induced by player i's period-τ private histories and provides a representation of player i's information at period τ. Each element collects interaction structures that are observationally indistinguishable given i's private history. As learning proceeds, this partition becomes finer since new observations refine indistinguishable structures into smaller sets. The number of cells $b_i^\tau(\sigma^*)$ therefore proxies the amount of information player i has acquired. Bounding the speed of learning reduces to bounding how many such refinements can occur before the partition stabilizes.

From the definition of private histories, for every $i \in N$, $\tau \geq 1$, and $\mathbf{g} \in \mathcal{G}_n$, we have that

$$[\mathbf{g}]_{i,\tau+1} \subseteq [\mathbf{g}]_{i,\tau}.$$

Thus, $\Pi_i^{\tau+1}(\sigma^*)$ is a refinement of $\Pi_i^\tau(\sigma^*)$, and the sequence

$$b_i^1(\sigma^*),\, b_i^2(\sigma^*),\, b_i^3(\sigma^*), \ldots$$

is weakly increasing.

Theorem 3. *Fix a myopic Bayesian equilibrium σ^*. Suppose that σ^* satisfies the following condition: $\Pi_i^{\tau+1}(\sigma^*) = \Pi_i^\tau(\sigma^*) \Rightarrow a_i^{\tau+1}(\mathbf{g}, \sigma^*) = a_i^\tau(\mathbf{g}, \sigma^*)$, for every $i \in N$ and every $\mathbf{g} \in \mathcal{G}_n$. Then, $\tau^*(\sigma^*) \leq 1 + n(|\mathcal{G}_n| - 1)$.*

Theorem 3 follows from a simple counting argument. Each strict refinement of $\Pi_i^\tau(\sigma^*)$ increases $b_i^\tau(\sigma^*)$ by at least one, and for each player i, the number of such increases is bounded by $|\mathcal{G}_n| - 1$. Aggregating across players yields the stated bound.

Theorem 3 uses only the finiteness of the candidate network class and therefore provides a worst-case bound on learning time. It does not exploit the structure of local observation or the role of equilibrium behavior in transmitting information. In our environment, however, information reaches an agent only through the actions of players she observes, and those actions themselves reflect what those players have learned. The next subsection exploits this interaction between network topology and equilibrium behavior to obtain a sharper bound on the speed of learning.

The stationarity condition of Theorem 3 is used only to convert a bound on the number of strict refinements into a bound on calendar time. Without this condition, the finite state space still bounds the number of times information partitions can be strictly refined, but it does not bound when those refinements occur. In particular, a nonstationary equilibrium selection may keep the same information partition unchanged for several periods before revealing information that players already possess.[9]

5.2 A Distance Bound under Action Separation

A sharper bound can be found once we exploit the fact that learning in our model is local. Under local observation, player i never observes the entire interaction structure directly. Instead, she learns about distant parts of the network only indirectly through the actions of the players she observes. This induces a propagation structure: information about a remote node must first affect that node's own history, then that node's action, then a downstream player's signal, and so on. Information can thus travel at most one edge per period.

To convert this propagation structure into a sharper bound, however, it is not enough that actions respond to histories. It is important that observed actions reveal the histories that generated them. Otherwise, information may be generated upstream, but fail to be decoded by downstream players.

To formalize the resulting propagation structure further, we orient links as follows

$$j \to i \qquad \Longleftrightarrow \qquad j \in N_i(\mathbf{g}).$$

Thus, information flows toward the observer.

For a realized network structure $\mathbf{g}$, define the directed distance from j to i by

$$\mathrm{dist}_{\mathbf{g}}(j,i) := \min\Big\{\ell \in \mathbb{N}_0 : \exists \text{ a directed path } j \to \cdots \to i \text{ of length } \ell\Big\},$$

with the convention $\mathrm{dist}_{\mathbf{g}}(j,i) = \infty$ if no such path exists. For $r \geq 0$, define the upstream r-neighborhood of i by

$$N_i^r(\mathbf{g}) := \{\, j \in N : \mathrm{dist}_{\mathbf{g}}(j,i) \leq r \,\}.$$

[9]For example, suppose player j can distinguish two interaction structures $\mathbf{g}$ and $\mathbf{g}'$, while player i cannot, and i observes j. If two actions are payoff-equivalent for j, then an equilibrium selection may prescribe the same action for j under $\mathbf{g}$ and $\mathbf{g}'$ for τ periods and then prescribe different actions in period $\tau+1$. Player i's partition would not refine before period $\tau+1$, but would refine afterward.

Hence,

$$N_i^0(\mathbf{g}) = \{i\}, \qquad N_i^1(\mathbf{g}) = \{i\} \cup N_i(\mathbf{g}).$$

For $r \geq 1$, define the rooted upstream radius-r structure of player i by

$$R_i^r(\mathbf{g}) := \big(N_j(\mathbf{g})\big)_{j \in N_i^{r-1}(\mathbf{g})}.$$

This object records the incoming neighborhoods of all nodes whose directed distance to i is at most $r-1$. Equivalently, it is the subset of the interaction structure that can possibly affect player i's observations within r rounds of play. Nodes outside this rooted upstream neighborhood are too far away to have influenced what player i has observed by time r. The next lemma summarizes the informational content of a player's history.

Lemma 2. *Fix a myopic Bayesian equilibrium σ^*. For every $i \in N$, $\tau \geq 1$, and $\mathbf{g}, \mathbf{g}' \in \mathcal{G}_n$,*

$$R_i^\tau(\mathbf{g}) = R_i^\tau(\mathbf{g}') \quad \Longrightarrow \quad h_i^\tau(\mathbf{g}, \sigma^*) = h_i^\tau(\mathbf{g}', \sigma^*). \tag{4}$$

If, in addition, equilibrium actions separate informational classes, that is,

$$[\mathbf{g}]_{i,\tau} \neq [\mathbf{g}']_{i,\tau} \quad \Longrightarrow \quad a_i^\tau(\mathbf{g}, \sigma^*) \neq a_i^\tau(\mathbf{g}', \sigma^*)$$

for every $i \in N$, $\tau \geq 1$, and $\mathbf{g}, \mathbf{g}' \in \mathcal{G}_n$, then the converse of (4) also holds.

Lemma 2 has a simple economic interpretation. The separation condition requires different informational classes to induce different actions.[10] Behavior then not only responds to local information but also transmits it: observing a neighbor's action reveals what that neighbor has learned so far. By period τ, player i can therefore be affected only by portions of the network whose consequences can reach her through $\tau - 1$ rounds of local observation. When actions separate informational classes, no information is lost in this transmission, so player i's period-τ history is informationally equivalent to knowing the rooted upstream structure $R_i^\tau(\mathbf{g})$.

The condition has a particularly transparent interpretation in the linear-quadratic game of Section 4. At any on-path history $h_i^\tau(\mathbf{g}, \sigma^*)$, player i's unique myopic best response is[11]

$$a_i^\tau(\mathbf{g}, \sigma^*) = 1 + \lambda \sum_{\mathbf{g}' \in [\mathbf{g}]_{i,\tau}} \mu_{i,\sigma^*}^\tau\big(\mathbf{g}' \mid h_i^\tau(\mathbf{g}, \sigma^*)\big) \sum_{j \in N_i(\mathbf{g})} a_j^\tau(\mathbf{g}', \sigma^*).$$

[10] Equivalently, a player's equilibrium action reveals her current informational class along the equilibrium path. This is related to the idea that actions can encode beliefs or private information when the action space is sufficiently rich; see, for example, Arieli and Mueller-Frank (2017).

[11] For the details about the incomplete network games, please see Chaudhuri et al. (2026).

Since $\lambda > 0$, equilibrium actions separate informational classes precisely when distinct informational classes induce distinct posterior expectations of aggregate neighbor activity. Thus, whenever what a player has learned changes her forecast of her neighbors' aggregate action, it also changes her own action and makes that information observable to players who observe her.

For a simple example, let $\mathcal{G}_n = \{\mathbf{g}^0, \mathbf{g}^1\}$ consist of two structures on three players satisfying

$$N_1(\mathbf{g}^0) = N_1(\mathbf{g}^1) = \{2\}, \qquad N_2(\mathbf{g}^0) = \emptyset, \qquad N_2(\mathbf{g}^1) = \{3\},$$

and $N_3(\mathbf{g}^0) = N_3(\mathbf{g}^1) = \emptyset$. Player 2 knows the realized structure from her neighborhood. Since player 3 chooses the action 1 under either structure, player 2 chooses 1 under $\mathbf{g}^0$ and $1+\lambda$ under $\mathbf{g}^1$. Her action therefore reveals the realized structure to player 1. After observing player 2, player 1 chooses $1 + \lambda$ under $\mathbf{g}^0$ and $1 + \lambda + \lambda^2$ under $\mathbf{g}^1$. Thus the separation condition holds in this example, and the information initially held by player 2 reaches player 1 after one round of observation.

This equivalence yields a sharper bound on the speed of learning. For every $\mathbf{g} \in \mathcal{G}_n$, define its directed diameter by

$$\operatorname{diam}(\mathbf{g}) := \max_{i,j \in N} \left\{ \operatorname{dist}_{\mathbf{g}}(j, i) : \operatorname{dist}_{\mathbf{g}}(j, i) < \infty \right\},$$

and define the maximal directed diameter across the candidate class by

$$d(\mathcal{G}_n) := \max_{\mathbf{g} \in \mathcal{G}_n} \operatorname{diam}(\mathbf{g}).$$

Proposition 2. *Suppose equilibrium actions separate informational classes. Then*

$$\tau^*(\sigma^*) \leq d(\mathcal{G}_n) + 1.$$

The one-edge-per-period logic behind Proposition 2 is related to earlier work on how quickly information can spread through a known network, especially DeMarzo et al. (2003) and Mueller-Frank (2013).[12]

[12] DeMarzo et al. (2003) show that repeated communication among Bayesian agents can aggregate dispersed information in finitely many rounds when the communication network is known. More closely related, Mueller-Frank (2013) studies agents who repeatedly announce posterior probabilities and shows that, for almost every prior, consensus is reached within one more period than the network diameter. The key mechanism is that, for almost every prior, different information sets lead to different posterior announcements. An observed announcement therefore reveals what the sender has learned, allowing information to move one link per period. Our separation condition provides the corresponding property when agents communicate indirectly through their equilibrium actions rather than by announcing their beliefs.

Proposition 2 shows that the speed of learning is governed jointly by the topology of the interaction structure and the transmission of information through equilibrium behavior. When the candidate class has small diameter, information that can affect a player reaches her within only a few rounds. A larger diameter permits information to travel through more intermediate players before reaching an observer and therefore yields a weaker bound on learning time. Distance determines how far information must travel, while separation determines whether it survives the journey. Local observation creates the propagation friction, and equilibrium separation ensures that information is not lost as it is conveyed through players' actions. Because the bound depends on diameter rather than directly on the number of players, it can be short even in large networks with small diameter, as in small-world networks (Watts and Strogatz, 1998). The number of players nonetheless enters indirectly because it bounds the diameter itself. A direct corollary is that learning time is at most linear in the number of players.

Corollary 3. *Suppose equilibrium actions separate informational classes. Then*

$$\tau^*(\sigma^*) \leq n.$$

In summary, the two bounds are complementary. While Theorem 3 is general, resulting from the finite number of possible interaction structures, it abstracts from the structure of information flow. Proposition 2 exploits this structure to provide sharper bounds, but requires that equilibrium actions separate informational classes, ensuring observed behavior fully reveals current informational states.

6 Conclusion

This paper studies strategic interaction when the underlying interaction structure is unknown to agents. The central message is that learning how to behave and learning the interaction structure itself are distinct outcomes. Agents use local observations of equilibrium behavior to form and update beliefs about candidate networks, yet they need not identify the realized network in order to behave optimally. We show that learning terminates in finite time and that equilibrium behavior converges to the complete-information Nash equilibrium of the realized network. Identification requires more. It depends on whether equilibrium behavior is sufficiently heterogeneous to distinguish among the candidate structures that remain observationally equivalent. In the linear-quadratic environment, this heterogeneity is captured by differences in walk profiles, which generically imply differences in equilibrium actions.

The results also have implications for empirical work. First, equilibrium behavior may reveal structural information that is not directly available to any individual agent. Agents begin with only local knowledge of the network, yet observed actions aggregate information dispersed across the interaction structure. When the equilibrium mapping from networks to actions is separating, panel data on actions can therefore be used to recover the realized interaction structure. The walk-profile characterization in the linear-quadratic environment provides a concrete condition under which such identification is possible. More broadly, our results provide a behavioral foundation for the heterogeneity conditions that underlie the identification of peer effects from reduced-form outcomes.

A second implication concerns the timing of data collection in network applications. Empirical work often assumes that observed behavior corresponds to an equilibrium outcome and employs the associated best-response relationships for estimation. Our analysis highlights that this assumption requires sufficient time for learning to occur. When observations are collected too early, agents may still be acting on incomplete information about the interaction structure, and observed behavior may not yet reflect the complete-information benchmark.

The paper also has implications for policy in environments where agents do not fully observe the networks in which they operate. Local observation is sufficient for behavior to converge to outcomes consistent with the realized structure, even when the structure itself remains unknown. Consequently, if a planner cares only about behavior, there may be little value in ensuring that agents learn the entire network. If, instead, the objective is for agents to understand the structure in which they are embedded, the relevant intervention is not simply to provide more data, but to improve the extent to which equilibrium behavior separates alternative structures.

More broadly, the framework provides a tractable way to study environments in which the object being learned is also the structure through which information is observed and transmitted. While we have focused on bilateral interaction networks, the analysis relies only on a finite latent interaction structure governing both payoffs and observations. The same logic can therefore be extended to richer environments, including hypergraphs, multilayer relationships, and networks with shadow links, where learning dynamics are governed by the corresponding notions of informational distance.

In short, the paper separates what agents need to know in order to behave optimally from what they need to know in order to identify the interaction structure itself. Understanding this distinction helps clarify the relationship between behavior, learning, and information in decentralized environments with network uncertainty.

Proofs

Proof of Lemma 1.

Fix a myopic Bayesian equilibrium σ^* and a player $i \in N$. Let $\Pi_i^\tau(\sigma^*) = \Pi_i^\tau$ be the partition of $\mathcal{G}_n$ induced by the equivalence relation $\sim_{i,\tau}$ under σ^*. Formally, $\Pi_i^\tau = \left\{[\mathbf{g}]_{i,\tau} : \mathbf{g} \in \mathcal{G}_n\right\}$. We first show that $\Pi_i^{\tau+1}$ is a refinement of Π_i^τ.

For any $\tau \geq 1$ choose any equivalence class $[\mathbf{g}]_{i,\tau+1} \in \Pi_i^{\tau+1}$ arbitrarily and let $\mathbf{g}' \in [\mathbf{g}]_{i,\tau+1}$. Then, from the definition of histories we get,

$$h_i^{\tau+1}(\mathbf{g}, \sigma^*) = h_i^{\tau+1}(\mathbf{g}', \sigma^*) \Rightarrow h_i^\tau(\mathbf{g}, \sigma^*) = h_i^\tau(\mathbf{g}', \sigma^*).$$

From the last equality it follows that, $\mathbf{g} \sim_{i,\tau} \mathbf{g}' \Rightarrow \mathbf{g}' \in [\mathbf{g}]_{i,\tau}$. Hence, for any $\mathbf{g} \in \mathcal{G}_n$ we have $[\mathbf{g}]_{i,\tau+1} \subseteq [\mathbf{g}]_{i,\tau}$. Therefore, $\Pi_i^{\tau+1}$ is a refinement of Π_i^τ.

Let $b_i^\tau = |\Pi_i^\tau|$. Since $\Pi_i^{\tau+1}$ is a refinement of Π_i^τ, the sequence $(b_i^\tau)_{\tau \geq 1}$ is weakly increasing. Also, because each Π_i^τ is a partition of the finite set $\mathcal{G}_n$, we have $1 \leq b_i^\tau \leq |\mathcal{G}_n|$ for every $\tau \geq 1$. Hence $(b_i^\tau)_{\tau \geq 1}$ is an integer-valued weakly increasing sequence that is bounded above. Therefore, there exists $T_i < \infty$ such that $b_i^\tau = b_i^{T_i}$ for all $\tau \geq T_i$.

Fix any $\tau \geq T_i$. Since $\Pi_i^{\tau+1}$ refines Π_i^τ and the two partitions have the same number of elements, this refinement cannot be strict. Thus $\Pi_i^{\tau+1} = \Pi_i^\tau$. Repeating the same argument period by period, we obtain $\Pi_i^\tau = \Pi_i^{T_i}$ for every $\tau \geq T_i$. Thus the equivalence class in the partition containing $\mathbf{g}$ is the same at periods τ and T_i. Therefore, $[\mathbf{g}]_{i,\tau} = [\mathbf{g}]_{i,T_i}$ for every $\mathbf{g} \in \mathcal{G}_n$ and every $\tau \geq T_i$.

Since N is finite, let $T = \max_{i \in N} T_i$. Then $T < \infty$, and for every player i, every interaction structure $\mathbf{g} \in \mathcal{G}_n$, and every period $\tau \geq T$, we have $[\mathbf{g}]_{i,\tau} = [\mathbf{g}]_{i,T}$. Thus no player's observational-equivalence class changes after period T. Therefore, learning ends in finite time. ■

Proof of Theorem 1.

Fix the myopic Bayesian equilibrium σ^*, a realized interaction structure $\mathbf{g} \in \mathcal{G}_n$, a period $\tau \geq \tau^*$, and a player $i \in N$. Write h_i^τ for $h_i^\tau(\mathbf{g}, \sigma^*)$. Since $h_i^\tau = h_i^\tau(\mathbf{g}, \sigma^*)$ is generated by the realized interaction structure $\mathbf{g}$ under σ^*, and since the prior has full support, we have $P_{i,\sigma^*}(h_i^\tau) \geq p(\mathbf{g}) > 0$. Hence $h_i^\tau \in P_i(\sigma^*)$, so the posterior at h_i^τ is determined by Bayes' rule.

By Bayes' rule, player i assigns positive probability at h_i^τ exactly to those interaction structures that generate the same period-τ private history as $\mathbf{g}$. Thus $\mu_{i,\sigma^*}^\tau(\mathbf{g}' \mid h_i^\tau) > 0$ if and only if $\mathbf{g}' \in [\mathbf{g}]_{i,\tau}$. Moreover, because the first component of player i's private history is her observed neighborhood, every $\mathbf{g}' \in [\mathbf{g}]_{i,\tau}$ satisfies $N_i(\mathbf{g}') = N_i(\mathbf{g})$.

Now take any $\mathbf{g}' \in [\mathbf{g}]_{i,\tau}$. Since learning has ended at τ^* and $\tau \geq \tau^*$, the equivalence class of $\mathbf{g}$ for player i is the same at periods τ and $\tau + 1$. Hence $\mathbf{g}' \in [\mathbf{g}]_{i,\tau+1}$ as well. Therefore $h_i^{\tau+1}(\mathbf{g}', \sigma^*) = h_i^{\tau+1}(\mathbf{g}, \sigma^*)$. Since a period-$(\tau+1)$ history is obtained from the period-τ history by adding the period-τ signal, it follows that $z_i^\tau(\mathbf{g}', \sigma^*) = z_i^\tau(\mathbf{g}, \sigma^*)$. Together with $N_i(\mathbf{g}') = N_i(\mathbf{g})$, equality of signals implies that the actions observed by player i are the same under $\mathbf{g}'$ and under $\mathbf{g}$; that is, $a_{N_i(\mathbf{g}')}^\tau(\mathbf{g}', \sigma^*) = a_{N_i(\mathbf{g})}^\tau(\mathbf{g}, \sigma^*)$.

We can now compare player i's payoff across all interaction structures in the support of her posterior. Fix any $a_i \in A_i$ and any $\mathbf{g}'$ such that $\mu_{i,\sigma^*}^\tau(\mathbf{g}' \mid h_i^\tau) > 0$. From the previous paragraph, $\mathbf{g}'$ has the same neighborhood for player i as $\mathbf{g}$, and the corresponding local action vector is also the same. Therefore, by the local payoff-equivalence assumption,

$$u_i\Big(a_i, \sigma_{-i}^{*\tau}(h_{-i}^\tau(\mathbf{g}', \sigma^*)); \mathbf{g}'\Big) = u_i\Big(a_i, \sigma_{-i}^{*\tau}(h_{-i}^\tau(\mathbf{g}, \sigma^*)); \mathbf{g}\Big).$$

Thus every payoff term that receives positive posterior probability in player i's myopic objective is identical to the payoff she would obtain under the realized interaction structure $\mathbf{g}$.

Since the remaining interaction structures have zero posterior probability, player i's posterior expected payoff at h_i^τ is the same function of a_i as the complete-information payoff under the realized interaction structure $\mathbf{g}$, i.e.

$$\sum_{\mathbf{g}' \in \mathcal{G}_n} \mu_{i,\sigma^*}^\tau(\mathbf{g}' \mid h_i^\tau) u_i\Big(a_i, \sigma_{-i}^{*\tau}(h_{-i}^\tau(\mathbf{g}', \sigma^*)); \mathbf{g}'\Big) = u_i\Big(a_i, \sigma_{-i}^{*\tau}(h_{-i}^\tau(\mathbf{g}, \sigma^*)); \mathbf{g}\Big)$$

for every $a_i \in A_i$.

Therefore, by the myopic Bayesian equilibrium condition, $\sigma_i^{*\tau}(h_i^\tau)$ maximizes $u_i(a_i, \sigma_{-i}^{*\tau}(h_{-i}^\tau(\mathbf{g}, \sigma^*)); \mathbf{g})$ over A_i. Using the definition of induced play, this gives

$$a_i^\tau(\mathbf{g}, \sigma^*) \in \arg\max_{a_i \in A_i} u_i\big(a_i, a_{-i}^\tau(\mathbf{g}, \sigma^*); \mathbf{g}\big).$$

Hence $a_i^\tau(\mathbf{g}, \sigma^*)$ is a best response to $a_{-i}^\tau(\mathbf{g}, \sigma^*)$ in the complete-information stage game induced by $\mathbf{g}$.

Since this argument applies to every player $i \in N$, the realized action profile $a^\tau(\mathbf{g}, \sigma^*)$ is

a Nash equilibrium of the complete-information stage game under $\mathbf{g}$. Therefore, for every $\tau \geq \tau^*$,

$$a^{\tau}(\mathbf{g}, \sigma^*) \in NE(\mathbf{g}),$$

where $NE(\mathbf{g})$ denotes the set of pure Nash equilibria of the complete-information stage game induced by $\mathbf{g}$. ■

Proof of Theorem 2.

Fix the myopic Bayesian equilibrium σ^*, a realized interaction structure $\mathbf{g} \in \mathcal{G}_n$, and a player $i \in N$. Suppose learning ends at τ^* under σ^*, and let $S \subseteq \{i\} \cup N_i(\mathbf{g})$ be a nonempty set satisfying the separation condition in Theorem 2.

Suppose, toward a contradiction, that learning is not perfect for player i. Then there exists $\mathbf{g}' \neq \mathbf{g}$ such that $\mathbf{g}' \in [\mathbf{g}]_{i,\tau^*}$. Hence $h_i^{\tau^*}(\mathbf{g}', \sigma^*) = h_i^{\tau^*}(\mathbf{g}, \sigma^*)$, which in particular implies $N_i(\mathbf{g}') = N_i(\mathbf{g})$. Since σ_i^* is a pure strategy, equality of private histories also implies $a_i^{\tau^*}(\mathbf{g}', \sigma^*) = a_i^{\tau^*}(\mathbf{g}, \sigma^*)$.

Because learning ends at τ^*, we also have $\mathbf{g}' \in [\mathbf{g}]_{i,\tau^*+1}$. Therefore $h_i^{\tau^*+1}(\mathbf{g}', \sigma^*) = h_i^{\tau^*+1}(\mathbf{g}, \sigma^*)$, so player i's period-τ^* signals coincide under $\mathbf{g}'$ and $\mathbf{g}$. Together with $N_i(\mathbf{g}') = N_i(\mathbf{g})$, equality of signals implies

$$a_j^{\tau^*}(\mathbf{g}', \sigma^*) = a_j^{\tau^*}(\mathbf{g}, \sigma^*) \qquad \text{for every } j \in N_i(\mathbf{g}).$$

Since $S \subseteq \{i\} \cup N_i(\mathbf{g})$, it follows that $a_S^{\tau^*}(\mathbf{g}', \sigma^*) = a_S^{\tau^*}(\mathbf{g}, \sigma^*)$.

By Theorem 1, applied to both interaction structures,

$$a^{\tau^*}(\mathbf{g}, \sigma^*) \in NE(\mathbf{g}) \qquad \text{and} \qquad a^{\tau^*}(\mathbf{g}', \sigma^*) \in NE(\mathbf{g}').$$

Hence the common action vector on S satisfies

$$a_S^{\tau^*}(\mathbf{g}, \sigma^*) = a_S^{\tau^*}(\mathbf{g}', \sigma^*) \in NE_S(\mathbf{g}) \cap NE_S(\mathbf{g}'),$$

contradicting the assumed separation of $NE_S(\cdot)$ across distinct interaction structures. Therefore $[\mathbf{g}]_{i,\tau^*} = \{\mathbf{g}\}$, so learning is perfect for player i. ■

Proof of Corollary 1.

Fix a myopic Bayesian equilibrium σ^* and a realized interaction structure $\mathbf{g} \in \mathcal{G}_n$. Suppose learning ends at period τ^* under σ^*.

By Theorem 1, for every $\tau \geq \tau^*$, $a^\tau(\mathbf{g}, \sigma^*) \in NE(\mathbf{g})$. It remains to characterize the complete-information Nash equilibrium set for the linear-quadratic stage game.

Under (1), player i's payoff is strictly concave in a_i, and hence the first-order condition is necessary and sufficient. Thus any complete-information Nash equilibrium a under $\mathbf{g}$ satisfies $1 - a_i + \lambda \sum_{j \in N} g_{ij}(\mathbf{g}) a_j = 0$ for every $i \in N$. Stacking these equations gives

$$\big(I - \lambda G(\mathbf{g})\big) a = \mathbf{1}.$$

Since $0 < \lambda < 1/\bar{\rho}$ and $\rho(G(\mathbf{g})) \leq \bar{\rho}$, the matrix $I - \lambda G(\mathbf{g})$ is invertible. Therefore the complete-information Nash equilibrium under $\mathbf{g}$ is unique and is given by

$$a^c(\mathbf{g}) = \big(I - \lambda G(\mathbf{g})\big)^{-1} \mathbf{1}.$$

Since $a^\tau(\mathbf{g}, \sigma^*) \in NE(\mathbf{g})$ for every $\tau \geq \tau^*$, uniqueness implies

$$a_i^\tau(\mathbf{g}, \sigma^*) = \Big[\big(I - \lambda G(\mathbf{g})\big)^{-1} \mathbf{1}\Big]_i \qquad \text{for every } i \in N \text{ and every } \tau \geq \tau^*.$$

■

Proof of Proposition 1.

Fix $i \in N$ and two interaction structures $\mathbf{g}, \mathbf{g}' \in \mathcal{G}_n$. Write $G := G(\mathbf{g})$ and $G' := G(\mathbf{g}')$, and let $\rho^* := \max\{\rho(G), \rho(G')\}$. For $\lambda \in (0, 1/\rho^*)$, define

$$f_i(\lambda) := e_i^\top (I - \lambda G)^{-1} \mathbf{1} - e_i^\top (I - \lambda G')^{-1} \mathbf{1}.$$

Thus $f_i(\lambda) = a_i^c(\mathbf{g}) - a_i^c(\mathbf{g}')$ whenever the two complete-information equilibrium vectors are well defined.

Since $\lambda \rho(G) < 1$ and $\lambda \rho(G') < 1$ for every $\lambda \in (0, 1/\rho^*)$, both Neumann-series expansions are valid. Hence

$$f_i(\lambda) = \sum_{m=0}^{\infty} \lambda^m \left[(G^m \mathbf{1})_i - ((G')^m \mathbf{1})_i\right] = \sum_{m=0}^{\infty} \lambda^m \left[W_i^m(\mathbf{g}) - W_i^m(\mathbf{g}')\right].$$

By assumption, there exists $k \geq 1$ such that $W_i^k(\mathbf{g}) \neq W_i^k(\mathbf{g}')$. Therefore the above power series has a nonzero coefficient, so f_i is not identically zero on $(0, 1/\rho^*)$.

It remains to rule out infinitely many coincidences. Each coordinate of $(I - \lambda G)^{-1}\mathbf{1}$ is a rational function of λ, since

$$e_i^\top (I - \lambda G)^{-1}\mathbf{1} = \frac{P_i(\lambda)}{\det(I - \lambda G)}$$

for some polynomial P_i, and similarly for G'. Thus f_i can be written as a rational function $N_i(\lambda)/D_i(\lambda)$, where $D_i(\lambda) \neq 0$ for all $\lambda \in (0, 1/\rho^*)$. Since f_i is not identically zero, N_i is not the zero polynomial. Hence N_i has only finitely many roots.

Therefore there exists a finite and hence null set, $\Lambda_i(\mathbf{g}, \mathbf{g}') \subset (0, 1/\rho^*)$ such that $f_i(\lambda) \neq 0$ for every $\lambda \in (0, 1/\rho^*) \setminus \Lambda_i(\mathbf{g}, \mathbf{g}')$. Equivalently,

$$a_i^c(\mathbf{g}) \neq a_i^c(\mathbf{g}') \qquad \text{for all } \lambda \in (0, 1/\rho^*) \setminus \Lambda_i(\mathbf{g}, \mathbf{g}').$$

■

Proof of Corollary 2.

Fix $\mathbf{g} \in \mathcal{G}_n$ and $i \in N$. Suppose there exists $j \in \{i\} \cup N_i(\mathbf{g})$ such that the walk profile $\mathbf{h} \mapsto (W_j^k(\mathbf{h}))_{k \geq 0}$ distinguishes every pair of interaction structures in $\mathcal{G}_n$.

Take any two distinct interaction structures $\mathbf{h}, \mathbf{h}' \in \mathcal{G}_n$. By assumption, there exists $k \geq 0$ such that $W_j^k(\mathbf{h}) \neq W_j^k(\mathbf{h}')$. Since $W_j^0(\mathbf{h}) = W_j^0(\mathbf{h}') = 1$, this difference must occur at some $k \geq 1$. Proposition 1 then implies that there is a finite and hence null set, $\Lambda_j(\mathbf{h}, \mathbf{h}') \subset (0, 1/\bar{\rho})$ such that $a_j^c(\mathbf{h}) \neq a_j^c(\mathbf{h}')$ for every $\lambda \in (0, 1/\bar{\rho}) \setminus \Lambda_j(\mathbf{h}, \mathbf{h}')$.

Since $\mathcal{G}_n$ is finite, the union $\Lambda_j := \bigcup_{\mathbf{h} \neq \mathbf{h}'} \Lambda_j(\mathbf{h}, \mathbf{h}')$ is finite. Hence, for every $\lambda \in (0, 1/\bar{\rho}) \setminus \Lambda_j$, the map $\mathbf{h} \mapsto a_j^c(\mathbf{h})$ is injective on $\mathcal{G}_n$.

In the linear-quadratic stage game, the complete-information Nash equilibrium under each $\mathbf{h}$ is unique. Therefore $NE_j(\mathbf{h}) = \{a_j^c(\mathbf{h})\}$ for every $\mathbf{h} \in \mathcal{G}_n$. Thus, for every admissible $\lambda \notin \Lambda_j$, the equilibrium sets $NE_j(\mathbf{h})$ are pairwise disjoint across distinct interaction structures.

If $j = i$, condition (i) of Theorem 2 applies. If $j \in N_i(\mathbf{g})$, condition (ii) of Theorem 2 applies. Therefore, whenever learning terminates at τ^* under σ^*, player i learns the realized interaction structure perfectly: $[\mathbf{g}]_{i,\tau^*} = \{\mathbf{g}\}$. ■

Proof of Theorem 3.

Write $\Pi_i^\tau = \Pi_i^\tau(\sigma^*)$ and $b_i^\tau = b_i^\tau(\sigma^*)$. Fix a player $i \in N$. Since Π_i^τ is a partition of the finite set $\mathcal{G}_n$, we have $1 \leq b_i^\tau \leq |\mathcal{G}_n|$ for every $\tau \geq 1$. Moreover, since $\Pi_i^{\tau+1}$ refines Π_i^τ, the sequence $(b_i^\tau)_{\tau \geq 1}$ is weakly increasing. Hence, for each player i, this sequence can strictly increase at most $|\mathcal{G}_n| - 1$ times.

We now show that, before learning ends, some player's partition must strictly refine in every period. Fix any period τ. Suppose, toward a contradiction, that no player's partition strictly refines from τ to $\tau + 1$. Since each $\Pi_i^{\tau+1}$ refines Π_i^τ, this means $\Pi_i^{\tau+1} = \Pi_i^\tau$, for every $i \in N$. By the partition-stationarity condition,

$$a_i^{\tau+1}(\mathbf{g}, \sigma^*) = a_i^\tau(\mathbf{g}, \sigma^*) \qquad \text{for every } i \in N \text{ and every } \mathbf{g} \in \mathcal{G}_n.$$

Therefore, the period-$(\tau + 1)$ action profile carries exactly the same graph-dependent information as the period-τ action profile. Since the period-τ action profile did not refine any player's partition, the period-$(\tau+1)$ action profile cannot refine any player's partition either. Hence $\Pi_i^{\tau+2} = \Pi_i^{\tau+1}$, for every $i \in N$. Repeating the same argument inductively gives

$$\Pi_i^s = \Pi_i^\tau \qquad \text{for every } s \geq \tau \text{ and every } i \in N.$$

Thus learning has already stabilized at period τ.

Therefore, if $\tau < \tau^*(\sigma^*)$, it cannot be the case that all players' partitions remain unchanged from τ to $\tau + 1$. Hence there must exist some player $i \in N$ and some interaction structure $\mathbf{g} \in \mathcal{G}_n$ such that $[\mathbf{g}]_{i,\tau+1} \subset [\mathbf{g}]_{i,\tau}$. Since $\Pi_i^{\tau+1}$ is a refinement of Π_i^τ, this strict containment means that the refinement is strict, and therefore $b_i^{\tau+1} > b_i^\tau$.

Thus every period before $\tau^*(\sigma^*)$ generates at least one strict increase in one of the n sequences $(b_i^\tau)_{\tau \geq 1}$, $i \in N$. Since each such sequence can strictly increase at most $|\mathcal{G}_n| - 1$ times, the total number of strict increases across all players is at most $n(|\mathcal{G}_n| - 1)$. Hence the number of periods before $\tau^*(\sigma^*)$ is at most $n(|\mathcal{G}_n| - 1)$, and so

$$\tau^*(\sigma^*) \leq 1 + n(|\mathcal{G}_n| - 1).$$

■

Proof of Lemma 2.

We prove the two implications by induction on τ.

For $\tau = 1$, the claim follows directly from the definitions. Indeed, $h_i^1(\mathbf{g}, \sigma^*) = N_i(\mathbf{g})$, while $R_i^1(\mathbf{g})$ records exactly $N_i(\mathbf{g})$ since $N_i^0(\mathbf{g}) = \{i\}$. Hence $R_i^1(\mathbf{g}) = R_i^1(\mathbf{g}')$ if and only if $h_i^1(\mathbf{g}, \sigma^*) = h_i^1(\mathbf{g}', \sigma^*)$. Thus both implications hold at $\tau = 1$.

Fix $\tau \geq 1$ and suppose the result holds at period τ for every player. We first prove the forward implication at period $\tau + 1$. Suppose $R_i^{\tau+1}(\mathbf{g}) = R_i^{\tau+1}(\mathbf{g}')$. Then, in particular, $N_i(\mathbf{g}) = N_i(\mathbf{g}')$. For every $j \in N_i(\mathbf{g}) \cup \{i\}$ and every $\ell \in N_j^{\tau-1}(\mathbf{g})$,

$$\mathrm{dist}_{\mathbf{g}}(\ell, i) \leq \mathrm{dist}_{\mathbf{g}}(\ell, j) + \mathrm{dist}_{\mathbf{g}}(j, i) \leq \tau.$$

Hence $N_j^{\tau-1}(\mathbf{g}) \subseteq N_i^{\tau}(\mathbf{g})$. Since

$$R_i^{\tau+1}(\mathbf{g}) = \big(N_\ell(\mathbf{g})\big)_{\ell \in N_i^{\tau}(\mathbf{g})} \quad \text{and} \quad R_j^{\tau}(\mathbf{g}) = \big(N_\ell(\mathbf{g})\big)_{\ell \in N_j^{\tau-1}(\mathbf{g})},$$

the latter is determined by the former. Therefore $R_i^{\tau+1}(\mathbf{g}) = R_i^{\tau+1}(\mathbf{g}')$ implies $R_j^{\tau}(\mathbf{g}) = R_j^{\tau}(\mathbf{g}')$ for every $j \in N_i(\mathbf{g}) \cup \{i\}$. By the induction hypothesis, $h_j^{\tau}(\mathbf{g}, \sigma^*) = h_j^{\tau}(\mathbf{g}', \sigma^*)$ for every $j \in N_i(\mathbf{g}) \cup \{i\}$. Since strategies are pure, these equal histories imply equal period-τ actions for all those players. Together with $N_i(\mathbf{g}) = N_i(\mathbf{g}')$, this gives $z_i^{\tau}(\mathbf{g}, \sigma^*) = z_i^{\tau}(\mathbf{g}', \sigma^*)$. Since the period-$\tau$ histories of player i are also equal by the induction hypothesis, the recursive definition of histories gives $h_i^{\tau+1}(\mathbf{g}, \sigma^*) = h_i^{\tau+1}(\mathbf{g}', \sigma^*)$.

We now prove the converse at period $\tau + 1$ under the separation condition. Suppose $h_i^{\tau+1}(\mathbf{g}, \sigma^*) = h_i^{\tau+1}(\mathbf{g}', \sigma^*)$. Then both the period-$\tau$ histories and the period-τ signals of player i coincide under $\mathbf{g}$ and $\mathbf{g}'$. In particular, $N_i(\mathbf{g}) = N_i(\mathbf{g}')$, and equality of signals gives $a_j^{\tau}(\mathbf{g}, \sigma^*) = a_j^{\tau}(\mathbf{g}', \sigma^*)$ for every $j \in N_i(\mathbf{g}) \cup \{i\}$.

By the contrapositive of separation condition, these equal actions imply $[\mathbf{g}]_{j,\tau} = [\mathbf{g}']_{j,\tau}$ for every $j \in N_i(\mathbf{g}) \cup \{i\}$. Hence $h_j^{\tau}(\mathbf{g}, \sigma^*) = h_j^{\tau}(\mathbf{g}', \sigma^*)$ for every such j. Applying the induction hypothesis for the converse implication, we obtain $R_j^{\tau}(\mathbf{g}) = R_j^{\tau}(\mathbf{g}')$ for every $j \in N_i(\mathbf{g}) \cup \{i\}$.

Finally, $R_i^{\tau+1}(\mathbf{g})$ is determined by $N_i(\mathbf{g})$ together with the collection $\{R_j^{\tau}(\mathbf{g}) : j \in N_i(\mathbf{g}) \cup \{i\}\}$. Since each of these objects is the same under $\mathbf{g}$ and $\mathbf{g}'$, it follows that $R_i^{\tau+1}(\mathbf{g}) = R_i^{\tau+1}(\mathbf{g}')$.

By induction, the result holds for every $\tau \geq 1$. ■

Proof of Proposition 2.

Suppose the separation condition holds and let $D := d(\mathcal{G}_n)$. Fix $i \in N$, $\mathbf{g} \in \mathcal{G}_n$, and $\tau \geq D + 1$. We show that $[\mathbf{g}]_{i,\tau} = [\mathbf{g}]_{i,D+1}$.

Take any $\mathbf{g}' \in [\mathbf{g}]_{i,D+1}$. Then $h_i^{D+1}(\mathbf{g}', \sigma^*) = h_i^{D+1}(\mathbf{g}, \sigma^*)$. By Lemma 2 and the separation condition, this implies $R_i^{D+1}(\mathbf{g}') = R_i^{D+1}(\mathbf{g})$.

Now fix an arbitrary $\tilde{\mathbf{g}} \in \mathcal{G}_n$. By definition of D, every node that reaches i under $\tilde{\mathbf{g}}$ does so along a path of length at most D. Hence $N_i^r(\tilde{\mathbf{g}}) = N_i^D(\tilde{\mathbf{g}})$ for every $r \geq D$. Therefore, for every $\tau \geq D+1$,

$$R_i^\tau(\tilde{\mathbf{g}}) = \big(N_j(\tilde{\mathbf{g}})\big)_{j \in N_i^{\tau-1}(\tilde{\mathbf{g}})} = \big(N_j(\tilde{\mathbf{g}})\big)_{j \in N_i^{D}(\tilde{\mathbf{g}})} = R_i^{D+1}(\tilde{\mathbf{g}}).$$

In particular,

$$R_i^\tau(\mathbf{g}') = R_i^{D+1}(\mathbf{g}') = R_i^{D+1}(\mathbf{g}) = R_i^\tau(\mathbf{g}).$$

Lemma 2 then gives $h_i^\tau(\mathbf{g}', \sigma^*) = h_i^\tau(\mathbf{g}, \sigma^*)$, so $\mathbf{g}' \in [\mathbf{g}]_{i,\tau}$. Hence $[\mathbf{g}]_{i,D+1} \subseteq [\mathbf{g}]_{i,\tau}$.

The reverse inclusion follows from the fact that histories are nested over time: if two structures generate the same period-τ history for player i, then they also generate the same period-$(D+1)$ history since $\tau \geq D+1$. Thus $[\mathbf{g}]_{i,\tau} \subseteq [\mathbf{g}]_{i,D+1}$. Therefore $[\mathbf{g}]_{i,\tau} = [\mathbf{g}]_{i,D+1}$ for every $\tau \geq D+1$.

Since the argument is uniform over i and $\mathbf{g}$, learning stabilizes by period $D+1$. Hence

$$\tau^*(\sigma^*) \leq D+1 = d(\mathcal{G}_n) + 1.$$

■

Proof of Corollary 3.

Fix any $\mathbf{g} \in \mathcal{G}_n$ and any players i, j with

$$\mathrm{dist}_{\mathbf{g}}(j, i) < \infty.$$

By definition, $\mathrm{dist}_{\mathbf{g}}(j, i)$ is the length of a shortest directed path from j to i. Any shortest directed path is simple, meaning that it visits each player at most once., since otherwise one could delete a directed cycle and obtain a shorter directed path from j to i. A simple directed path on n players has length at most $n-1$. Thus

$$\mathrm{dist}_{\mathbf{g}}(j, i) \leq n-1.$$

Hence

$$d(\mathcal{G}_n) = \max_{\mathbf{g} \in \mathcal{G}_n} \mathrm{diam}(\mathbf{g}) \leq n-1.$$

Applying Proposition 2, we obtain

$$\tau^*(\sigma^*) \leq d(\mathcal{G}_n) + 1 \leq n.$$

■

References

Acemoglu, Daron, Asuman Ozdaglar, and Ali ParandehGheibi (2010), "Spread of (mis)information in social networks." *Games and Economic Behavior*, 70(2), 194–227.

Arieli, Itai and Manuel Mueller-Frank (2017), "Inferring beliefs from actions." *Games and Economic Behavior*, 102, 455–461.

Bala, Venkatesh and Sanjeev Goyal (1998), "Learning from neighbours." *The Review of Economic Studies*, 65(3), 595–621.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou (2006), "Who's Who in Networks. Wanted: The Key Player." *Econometrica*, 74(5), 1403–1417.

Battigalli, Pierpaolo, Fabrizio Panebianco, and Paolo Pin (2023), "Learning and selfconfirming equilibria in network games." *Journal of Economic Theory*, 212, 105700.

Ben-Porath, Elchanan and Michael Kahneman (1996), "Communication in repeated games with private monitoring." *Journal of Economic Theory*, 70(2), 281–297.

Billand, Pascal, Christophe Bravard, Jacques Durieu, and Sudipta Sarangi (2019), "Firm heterogeneity and the pattern of R&D collaborations." *Economic Inquiry*, 57(4), 1896–1914.

Billand, Pascal, Christophe Bravard, Sudipta Sarangi, and Myrna Wooders (2026), "The impact of price dispersion on R&D network formation." *Economic Theory*.

Bramoullé, Yann, Rachel Kranton, and Martin D'Amours (2014), "Strategic interaction and networks." *American Economic Review*, 104(3), 898–930.

Breza, Emily, Arun G Chandrasekhar, and Alireza Tahbaz-Salehi (2018), "Seeing the forest for the trees? an investigation of network knowledge." Working Paper 24359, National Bureau of Economic Research.

Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou (2009), "Peer effects and social networks in education." *The Review of Economic Studies*, 76(4), 1239–1267.

Chaudhuri, Promit K., Matthew O. Jackson, Sudipta Sarangi, and Hector Tzavellas (2026), "Games under network uncertainty." Working paper.

Compte, Olivier (1998), "Communication in repeated games with imperfect private monitoring." *Econometrica*, 66(3), 597–626.

Dasaratha, Krishna and Kevin He (2020), "Network structure and naive sequential learning." *Theoretical Economics*, 15(2), 415–444.

Dasaratha, Krishna and Kevin He (2026), "Aggregative efficiency of Bayesian learning in networks." ArXiv preprint arXiv:1911.10116, revised February 2026.

de Martí, Joan and Yves Zenou (2015), "Network games with incomplete information." *Journal of Mathematical Economics*, 61, 221–240.

DeMarzo, Peter M, Dimitri Vayanos, and Jeffrey Zwiebel (2003), "Persuasion bias, social influence, and unidimensional opinions." *The Quarterly journal of economics*, 118(3), 909–968.

Foerster, Manuel, Ana Mauleon, and Vincent J Vannetelbosch (2021), "Shadow links." *Journal of Economic Theory*, 197, 105325.

Fudenberg, Drew, David Levine, and Eric Maskin (1994), "The folk theorem with imperfect public information." *Econometrica*, 62(5), 997–1039.

Fudenberg, Drew and David K. Levine (1993), "Self-confirming equilibrium." *Econometrica*, 61(3), 523–545.

Fudenberg, Drew and Yuichi Yamamoto (2010), "Repeated games where the payoffs and monitoring structure are unknown." *Econometrica*, 78(5), 1673–1710.

Gale, Douglas and Shachar Kariv (2003), "Bayesian learning in social networks." *Games and Economic Behavior*, 45(2), 329–346.

Galeotti, Andrea, Sanjeev Goyal, Matthew O. Jackson, Fernando Vega-Redondo, and Leeat Yariv (2010), "Network games." *The Review of Economic Studies*, 77(1), 218–244.

Golub, Benjamin and Matthew O. Jackson (2010), "Naive learning in social networks and the wisdom of crowds." *American Economic Journal: Microeconomics*, 2(1), 112–149.

Goyal, Sanjeev and Sumit Joshi (2003), "Networks of collaboration in oligopoly." *Games and Economic Behavior*, 43(1), 57–85.

Goyal, Sanjeev and Sumit Joshi (2006), "Unequal connections." *International Journal of Game Theory*, 34(3), 319–349.

Hörner, Johannes and Wojciech Olszewski (2006), "The folk theorem for games with private almost-perfect monitoring." *Econometrica*, 74(6), 1499–1544.

Huang, Wanying, Philipp Strack, and Omer Tamuz (2024), "Learning in repeated interactions on networks." *Econometrica*, 92(1), 1–27.

Jackson, Matthew O. and Yves Zenou (2015), "Games on networks." In *Handbook of Game Theory with Economic Applications* (H. Peyton Young and Shmuel Zamir, eds.), volume 4, 95–163, Elsevier.

Jordan, James S. (1991), "Bayesian learning in normal form games." *Games and Economic Behavior*, 3(1), 60–81.

Jordan, James S. (1995), "Bayesian learning in repeated games." *Games and Economic Behavior*, 9(1), 8–20.

Kalai, Ehud and Ehud Lehrer (1993a), "Rational learning leads to Nash equilibrium." *Econometrica*, 61(5), 1019–1045.

Kalai, Ehud and Ehud Lehrer (1993b), "Subjective equilibrium in repeated games." *Econometrica*, 61(5), 1231–1240.

Kandori, Michihiro and Hitoshi Matsushima (1998), "Private observation, communication and collusion." *Econometrica*, 66(3), 627–652.

Laclau, Marie (2014), "Communication in repeated network games with imperfect monitoring." *Games and Economic Behavior*, 87, 136–160.

Li, Wei and Xu Tan (2020), "Locally Bayesian learning in networks." *Theoretical Economics*, 15(1), 239–278.

Lindquist, Matthew J., Eleonora Patacchini, Michael Vlassopoulos, and Yves Zenou (2024), "Spillovers in criminal networks: Evidence from co-offender deaths." CEPR Discussion Paper 19159, Centre for Economic Policy Research.

Lipnowski, Elliot and Evan Sadler (2019), "Peer-confirming equilibrium." *Econometrica*, 87(2), 567–591.

Macault, Emilien, Marco Scarsini, and Tristan Tomala (2022), "Social learning in nonatomic routing games." *Games and Economic Behavior*, 132, 221–233.

Molavi, Pooya, Alireza Tahbaz-Salehi, and Ali Jadbabaie (2018), "A theory of non-Bayesian social learning." *Econometrica*, 86(2), 445–490.

Mueller-Frank, Manuel (2013), "A general framework for rational learning in social networks." *Theoretical Economics*, 8(1), 1–40.

Nava, Francesco and Michele Piccione (2014), "Efficiency in repeated games with local interaction and uncertain local monitoring." *Theoretical Economics*, 9(1), 279–312.

Renault, Jérôme and Tristan Tomala (1998), "Repeated proximity games." *International Journal of Game Theory*, 27(4), 539–559.

Renault, Jérôme and Tristan Tomala (2004), "Learning the state of nature in repeated games with incomplete information and signals." *Games and Economic Behavior*, 47(1), 124–156.

Sugaya, Takuo (2022), "Folk theorem in repeated games with private monitoring." *The Review of Economic Studies*, 89(4), 2201–2256.

Watts, Duncan J. and Steven H. Strogatz (1998), "Collective dynamics of 'small-world' networks." *Nature*, 393(6684), 440–442.

Wolitzky, Alexander (2013), "Cooperation with network monitoring." *The Review of Economic Studies*, 80(1), 395–427.