

Local Global Games and Network Common Learning

Olga Rospuskova Omer Tamuz Jake Zhang*

July 31, 2026

Abstract

We study global games in which agents coordinate locally, with their social network neighbors, contingent on a favorable state. Before acting, agents learn the private signals of all agents within network distance r . As r grows, every agent learns the state, but efficient coordination depends on higher-order beliefs, which are shaped by the geometry of the network. We introduce network common learning, a network analogue of common learning, and show that it is attained when neighboring agents' observations differ by many signals, as on the two-dimensional grid, but fails on networks with informational bottlenecks, such as the line, where only the safe action survives in equilibrium.

1 Introduction

Global games are a successful and well-studied model of a number of important phenomena, including bank runs, protests and more (Carlsson and van Damme, 1993; Morris and Shin, 2003; Goldstein and Pauzner, 2005; Angeletos et al., 2007). The basic model is of a group of agents that need to act in unison to achieve a goal, and furthermore can only succeed when the state of nature is favorable. The important insight they provide is that even when agents have very precise information about the state, coordination can fail because of higher-order uncertainty (Rubinstein, 1989; Carlsson and van Damme, 1993; Morris and Shin, 2003). The study of global games has spurred a literature within epistemic game theory, dedicated to understanding how beliefs affect coordination. A particularly interesting concept is the idea of common learning (Cripps, Ely, Mailath, and Samuelson, 2008), which captures a convergence not just to knowing the state, but to commonly knowing it (see also Steiner and Stewart, 2011; Cripps, Ely, Mailath, and Samuelson, 2013; Frick, Iijima, and Ishii, 2023).

*Caltech. Olga Rospuskova was supported by a Citadel fellowship. Omer Tamuz was supported by a National Science Foundation CAREER award (DMS-1944153) and by a MURI award (N000142412742).

Our setting features a classic global game. There is a state that is either high or low, and each agent has to choose between a safe and a risky action. When the state is low, the safe action is optimal. When the state is high, it is better to take the risky action, but only if others take it too. We introduce a network structure to this game, so that agents do not need to coordinate with all others, but only locally, with their social network neighbors; accordingly, we call this game a *local global game*. This changes the higher-belief structure required for coordination. Instead of the common q -belief that is needed for coordination on a complete network, a general network requires a weaker condition of iterated beliefs, which we call network q -beliefs, and which we show is equivalent to the existence of an efficient equilibrium. This is similar to the conditions explored by Morris, Rob, and Shin (1995) and Kets (2011).

Our network plays an additional important role, as a conduit of information. Initially, agents are endowed with conditionally independent private signals about the state. Before choosing their actions, the agents participate in a communication stage, in which they share their signals with their neighbors, their neighbors' neighbors and so on, up to some graph distance r , which we call the observation radius. For simplicity, and since this is a coordination game, we model this communication phase non-strategically, and assume that players simply reveal their private information.

For large r , each agent learns the state with high confidence. Moreover, neighboring agents observe information that is highly overlapping, seemingly allowing their beliefs to align more easily. But does individual learning and the overlap of information allow agents to resolve strategic uncertainty and coordinate efficiently?

We show that the answer to this question depends sharply on the geometry of the network. Consider two stylized examples: the infinite line and the infinite two-dimensional grid. On both networks, increasing the observation radius allows every agent to learn the state with arbitrarily high confidence. Nevertheless, the equilibrium implications are different. On the line, when coordination requires sufficiently high confidence, taking the safe action is the only action taken in equilibrium, no matter how large the radius r is. On the two-dimensional grid, by contrast, when the radius is large there exist efficient equilibria in which the agents, with high probability, coordinate and take the risky action in the high state.

Why does the additional dimension reverse the equilibrium prediction? The key is not how much information each agent observes but how the available information changes across agents. Two neighboring agents observe largely overlapping neighborhoods, yet each also sees some signals that the other does not. We refer to signals observed by one agent but not by his neighbor as his *information boundary*. On the line, this region contains only one signal, regardless of the size of the observation radius. On the two-dimensional grid, the information boundary grows proportionally with the observation radius.

This difference determines the feasibility of coordination. When the size of the information boundary remains bounded, there is no concentration of measure, and the uncertainty about a neighbor's information never vanishes. This concern propagates

from neighbor to neighbor, eventually ruling out risky coordination. On the other hand, when the information boundary is large, the unobserved signals become predictable, because of concentration of measure. Strategic uncertainty diminishes as the radius grows, allowing confidence in risky play to sustain itself across the network.

To construct efficient equilibria, we study what we call *network q -belief* (see also Morris, Rob, and Shin, 1995; Kets, 2011). This notion is defined recursively. At the first level, an agent is required to assign probability at least q to the state being high. At each subsequent level, he needs to assign probability at least q to the joint event that the state is high and that all his neighbors assign probability q to the events defined at the preceding level. We show that the limit of this iterated belief satisfies the same fixed-point condition as the agents' best responses in equilibrium. Namely, it yields a collection of *self-sustaining events*: for each agent, whenever his corresponding event in the collection happens, he assigns probability at least q to the joint event that the state is favorable and that every neighbors' corresponding event also happens. In analogy to the idea of common learning (Cripps, Ely, Mailath, and Samuelson, 2008), we define *network common learning* to hold if for all q , network q -belief holds with high probability for all r large enough.

We use this technique to prove that coordination is possible on the two dimensional grid. The key geometric property of the grid is what we call *large information boundaries* (Definition 2), a condition that applies to many other networks. To define the large information boundaries condition, we first define the collection of *test regions*: these are regions in the social network graph that are finite intersections of the agents' observation neighborhoods. Large information boundaries are attained when every large test region that is observed by some agent i has a large information boundary: the part of it that is not seen by i 's neighbor j is large (or empty). The term "large" is used informally here, but is defined formally below.

Large information boundaries allow us to define self-sustaining events that yield efficient equilibria. The self-sustaining events require that in each test region the fraction of favorable signals is large. That is, to coordinate an agent needs to not just see many favorable signals (yielding individual learning) but see many favorable signals in each test region, which is an intersection of the regions seen by a group of agents.

Our final result provides a complementary impossibility theorem. Suppose the network contains a "bottleneck": a path along which moving from one agent to the next introduces a bounded number of new signals; this holds for the one-dimensional grid. That is, the information boundary of the observation neighborhoods along this path is of bounded size. This implies that the higher-order uncertainty generated by the boundary is significant regardless of r , and is not mitigated by concentration of measure. Consequently, when q is high enough, the collection of events that can be believed with probability q shrinks as we iterate over higher-order beliefs and move along the path, diminishing to the empty set if the path is long enough. Network q -belief is consequently empty, and the safe action is the unique equilibrium prediction.

Our paper identifies a new channel through which network geometry affects coordination. Network geometry matters not only because it determines who interacts with whom and how information is diffused, but also because it determines how much uncertainty remains when agents reason about the actions of their neighbors. When neighboring observation sets differ by many signals, the fresh information received by each agent becomes predictable in the aggregate, allowing confidence in others' actions to be sustained. When observation sets change only gradually, uncertainty about a neighbor's information cannot be ignored and may propagate over long distances, even though every individual agent is highly informed about the state.

1.1 Related literature

Our model follows the literature on global games ([Carlsson and van Damme, 1993](#); [Morris and Shin, 2003](#)). The basic lesson of this literature is that arbitrarily precise private information need not induce coordination because equilibrium behavior also hinges on higher-order beliefs ([Rubinstein, 1989](#)). We inherit this lesson but modify the object of higher-order beliefs. In the canonical model, an agent needs to coordinate with the entire population, and this imposes conditions on belief involving every agent at every epistemic order. In our model, an agent's payoff only depends on the actions of his neighbors, and so the relevant object is network q -belief, which chains beliefs along the edges of the network. Despite this difference from the classic literature, our impossibility result is in the same family as the coordinated attack problem of [Halpern and Moses \(1984\)](#): the argument in the proof of [Theorem 2](#) is a spatial version of the temporal mechanism governing the electronic mail game of [Rubinstein \(1989\)](#), with graph distance along a bottleneck path playing the role of the number of messages exchanged.

Network q -belief is a network analogue of common p -belief, as defined by [Monderer and Samet \(1989\)](#), who show that common p -belief admits a fixed-point characterization; we derive a network analogue of their fixed-point property. Common p -belief is also central to the analysis of [Kajii and Morris \(1997\)](#), who show that an equilibrium is robust to incomplete information when the event on which the game is common knowledge can be commonly p -believed, and whose contagion arguments are echoed in our impossibility result. [Kets \(2011\)](#) is closest to us in adapting the belief operator to the network structure. Beyond uncertainty about the state, [Kets \(2011\)](#) also considers uncertainty about the network structure, and applies these techniques to the study of anonymous network games.

[Morris \(2000\)](#) studies how local interactions determine whether behavior can spread contagiously through a network. In his complete information setting, agents repeatedly best respond to the actions of their neighbors. He shows that an action adopted by a finite group of agents can be spread through the entire network only when there is no cohesive group that interacts mostly within the group and resists the outside contagion. Our analysis shares the emphasized propagation and recursive structure

but the object being propagated is different. In Morris, an agent counts the proportion of his neighbors taking a certain action and the observed action diffuses through the network. In contrast, we study how uncertainty propagates through higher-order beliefs and the network structure. Network geometry therefore matters not only through who interacts with whom, as in Morris (2000), but also through how information changes across neighboring positions. Related work on local interaction dynamics and geometry includes Ellison (1993), Blume (1993) and Oyama and Takahashi (2015).

Cripps, Ely, Mailath, and Samuelson (2008) study agents who accumulate private signals over time and ask whether this leads to common p -belief. They show that when the number of signal realizations is finite, individual learning implies common learning; see also Frick, Iijima, and Ishii (2023). Our observation radius r plays the role of time in their model: as the radius r grows, the information of each agent becomes arbitrarily precise and the question of interest is whether higher-order beliefs also grow. In our model, individual learning does not always imply common learning, even with finite signals; it is the geometry of the graph which decides.

Chwe (2000) also studies coordination in which a network is used to communicate, and asks which networks are minimally sufficient for coordination. In his model communication is used to build exact common knowledge within the set of participants, so the binding constraint is whether a network has the right links.

A recent paper by Hutchcroft, Rospuskova, and Tamuz (2026) also studies coordination on social networks using local communication, but in a crucially different setting, resulting in very different conclusions. In that paper, there is no global state, and agents play a pure coordination game (e.g., choosing which side of the road to drive on). Thus, the challenge is to communicate and coordinate on how to break ties between a priori equivalent actions. In that setting, coordinating on the line graph is possible, and in fact higher efficiency can be achieved on the line than on the two dimensional grid. Amenability, the geometric property enabling coordination in that paper, is related to having small information boundaries, while in our setting large information boundaries are a sufficient condition for coordination.

2 Model

Let N be a set of agents, connected through a social network $G = (N, E)$. An edge $(i, j) \in E$ indicates that agents i and j are neighbors. We assume that the social network is undirected, so that $(i, j) \in E$ if and only if $(j, i) \in E$. For convenience, $(i, i) \notin E$. For each $i \in N$, let $N_i = \{j \in N : (i, j) \in E\}$ denote the set of i 's neighbors. We assume that every agent has finitely many neighbors, i.e. $d_i = |N_i| < \infty$ for every $i \in N$. We will be interested in the limit of very large networks, and so for simplicity we will assume that the number of agents is countably infinite.

A path from i to j is a finite sequence of agents beginning at i and ending at j , such that every two consecutive agents are neighbors. We assume that G is connected, i.e., that there is a path connecting each pair of agents. This is without loss of generality,

since otherwise one can consider each connected component in isolation. The distance between i and j is the length of the shortest path between them. Each agent observes all agents at graph distance at most r , where $r \geq 0$. We call r the observation radius. We also define $B_r(i) \subseteq N$ as the set of agents at distance at most r from i . We refer to this set as r -neighborhood of i , or as the observation neighborhood of agent i .

We consider a coordination game played on the network G . The payoff of an agent depends on the underlying state, his own action and the actions of his neighbors. The state is binary: $\theta \in \Theta$, where $\Theta = \{H, L\}$, with H and L denoting the high and low states respectively. Each agent chooses between a safe action S and a risky action R . The safe action always guarantees a payoff of zero. The risky action yields a payoff $\alpha > 0$ only if the state is H and all his neighbors coordinate on the risky action; otherwise, it incurs a loss of $\beta > 0$. Formally, the payoff of an agent i under state θ and action profile $a \in \{S, R\}^N$ is given by

$$u_i(a, \theta) = \begin{cases} 0, & \text{if } a_i = S, \\ \alpha, & \text{if } \theta = H, a_i = R \text{ and } a_j = R \forall j \in N_i, \\ -\beta, & \text{otherwise.} \end{cases}$$

The agents are uncertain about the state and share a common prior. For simplicity, we assume $\mathbb{P}(\theta = H) = \mathbb{P}(\theta = L) = 1/2$. Each agent i receives a binary signal s_i about the state. Conditional on the state, the signals are independent. We assume that the signals are symmetric with precision $p > 1/2$ and take values in $\{0, 1\}$. That is,

$$\mathbb{P}(s_i = 1 | \theta = H) = \mathbb{P}(s_i = 0 | \theta = L) = p.$$

The state and signals are random variables in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = \{H, L\} \times \{0, 1\}^N$, $\mathcal{F}$ is the product sigma-algebra, and $\mathbb{P}$ is the measure described above.

Before choosing an action, agents participate in a communication phase. In this phase, each agent i observes the private signals of all agents in $B_r(i)$, for some fixed radius of observation r . We denote the set of possible observations by $O_i^r = \{0, 1\}^{B_r(i)}$. The information available to agent i is the random variable $I_i^r = (s_j)_{j \in B_r(i)}$, which takes values in O_i^r . Thus, a pure strategy for agent i in this game is a map $t_i: O_i^r \rightarrow \{S, R\}$. This is a Bayesian game, and accordingly our solution concept is that of Bayes-Nash equilibrium.

Since the network is infinite and connected, the size of $B_r(i)$ goes to infinity as r increases, and so the number of signals observed by each agent in the communication phase increases. Since private signals are informative about the state, agents individually learn the state. In particular, as r tends to infinity, each agent's posterior belief $\mathbb{P}(\theta = H | I_i^r)$ converges to 1 if $\theta = H$, and converges to 0 if $\theta = L$, i.e. for all $\varepsilon > 0$,

$$\lim_{r \rightarrow \infty} \mathbb{P}(\mathbb{P}(\theta = H | I_i^r) > 1 - \varepsilon | \theta = H) = \lim_{r \rightarrow \infty} \mathbb{P}(\mathbb{P}(\theta = H | I_i^r) < \varepsilon | \theta = L) = 1.$$

We focus on pure Bayes-Nash equilibria, which are characterized by a collection of sets $(X_i^r)_{i \in N}$, where $X_i^r \subseteq O_i^r$. These are the sets on which i chooses the risky action, i.e., $a_i = R$ if and only if $I_i^r \in X_i^r$. The expected utility of agent i from choosing the risky action upon observing $I_i^r = x$ is

$$\alpha \mathbb{P}(\theta = H, I_j^r \in X_j^r \forall j \in N_i | I_i^r = x) - \beta(1 - \mathbb{P}(\theta = H, I_j^r \in X_j^r \forall j \in N_i | I_i^r = x)),$$

whereas choosing a safe action yields zero.

Therefore, the strategy profile induced by $(X_i^r)_{i \in N}$ is an equilibrium if and only if, for every $i \in N$,

$$\begin{aligned} \forall x \in X_i^r, \quad \mathbb{P}(\theta = H, I_j^r \in X_j^r \forall j \in N_i | I_i^r = x) &\geq \frac{\beta}{\alpha + \beta}, \\ \forall y \notin X_i^r, \quad \mathbb{P}(\theta = H, I_j^r \in X_j^r \forall j \in N_i | I_i^r = y) &\leq \frac{\beta}{\alpha + \beta}. \end{aligned}$$

Denoting $q = \frac{\beta}{\alpha + \beta}$, the risky action is optimal whenever the probability that the state is high and all neighbors choose risky action is at least q . We resolve ties in favor of R , assuming that agents take the risky action when they are indifferent.

We ask for which networks agents can coordinate efficiently in equilibrium when they observe signals within a sufficiently large radius r . That is, with high probability, they choose the risky action in the high state and the safe action in the low state.

2.1 Network q -belief and Network Common Learning

For any event $E \subseteq \Omega$, define

$$\mathcal{B}_{i,r}^q(E) = \{\omega \in \Omega : \mathbb{P}(E | I_i^r)(\omega) \geq q\}.$$

This is the event on which agent i , using signals in $B_r(i)$, assigns probability at least q to E .¹ If each neighbor j of i chooses the risky action on X_j^r , then $\mathcal{B}_{i,r}^q(H \cap (\bigcap_{j \in N_i} X_j^r))$ is the event on which agent i assigns a probability at least q to the event that the state is high and every neighbor of i chooses R .² Therefore, the equilibrium condition can equivalently be written as

$$\forall i \in N, \quad X_i^r = \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} X_j^r \right) \right). \quad (1)$$

¹Since $B_r(i)$ is finite, $\sigma(I_i^r)$ is generated by the finite partition $\{\{I_i^r = x\} : x \in O_i^r\}$, in which every atom has positive probability. Hence the conditional probability $\mathbb{P}(E | I_i^r)$ admits a canonical version, given on each atom by $\mathbb{P}(E \cap \{I_i^r = x\}) / \mathbb{P}(I_i^r = x)$, and $\mathcal{B}_{i,r}^q(E)$ is well defined as the union of the atoms on which this value is at least q .

²Here and below we use H to denote the event $\{\theta = H\}$, and use X_j^r to denote the event $\{I_j^r \in X_j^r\}$.

This is a fixed-point condition. To identify its solutions, we study iterated network beliefs. Starting from the entire sample space Ω , at each subsequent step agent i leaves only those in which he q -believes that state is H and that the previous-level condition holds for every neighbor. Formally, let $C_{i,r}^{q,0} = \Omega$. For $n \geq 1$, define

$$C_{i,r}^{q,n}(H) = \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} C_{j,r}^{q,n-1}(H) \right) \right).$$

This is the event that agent i has n^{th} order network q -belief. In particular,

$$C_{i,r}^{q,1}(H) = \mathcal{B}_{i,r}^q(H)$$

is the event in which agent i assigns probability at least q to the high state. Likewise,

$$C_{i,r}^{q,2}(H) = \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} \mathcal{B}_{j,r}^q(H) \right) \right)$$

is the event that agent i has second order network q -belief: i assigns probability at least q to the joint event that state is H and all his neighbors believe that state is H with probability at least q . At higher levels, the same condition is applied recursively to i 's neighbors, their neighbors and so on.

The event that i has network q -belief is

$$C_{i,r}^q(H) = \bigcap_{n \geq 0} C_{i,r}^{q,n}(H).$$

This is the event on which q -belief in H persists through every finite order of beliefs, starting from i . The following lemma shows that the limiting event satisfies the same fixed-point condition required for equilibrium. To simplify notation, we henceforth omit the (H) and write $C_{i,r}^{q,n}$ and $C_{i,r}^q$ to indicate $C_{i,r}^{q,n}(H)$ and $C_{i,r}^q(H)$.

Lemma 1. *For every i and r ,*

$$C_{i,r}^q = \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} C_{j,r}^q \right) \right).$$

Thus, network q -belief is *self-sustaining*: whenever it holds at i , agent i assigns probability at least q to the joint event that the state is H and network q -belief holds at each of i 's neighbors.

Since the family of sets $(C_{i,r}^q)_{i \in N}$ satisfies the equilibrium condition (1), it induces an equilibrium. The next proposition shows that this equilibrium has the largest possible risky set: whenever another equilibrium prescribes the risky action to agent i , this equilibrium does so as well.

Proposition 1. *The family $(C_{i,r}^q)_{i \in N}$ induces an equilibrium in which $X_i^r = C_{i,r}^q$. Moreover, it is the maximal risky equilibrium in the sense that in any other equilibrium, $a_i = R$ only on the event $C_{i,r}^q$.*

We now study whether network q -belief in H emerges as agents observe larger neighborhoods. Particularly, this requires that events $(C_{i,r}^q)_{i \in N}$ become very likely for every agent in state H . This motivates the following notion of network common learning.

Definition 1. *Network q -learning of H occurs if for every agent i ,*

$$\lim_{r \rightarrow \infty} \mathbb{P}(C_{i,r}^q | \theta = H) = 1.$$

Network common learning of H occurs if network q -learning of H occurs for all $q \in (0, 1)$.

So, network q -learning means that in state H network q -belief in H occurs with probability approaching 1 for every agent. It also implies asymptotic efficiency of the maximal risky equilibrium. Since the maximal equilibrium prescribes the risky action on these events, risky play becomes asymptotically likely in state H . Individual learning implies that in state L each agent chooses the safe action with high probability. For choosing the safe action, no higher-order beliefs are necessary: once the first-order belief in H is below q , agents will choose the safe action.

3 Efficient Coordination on Networks with Large Information Boundaries

In this section we introduce a geometric condition on the social network graph which we show implies that network common learning is obtained, and thus efficient coordination can be sustained in equilibrium, for all r large enough.

To define our geometric condition, recall that $B_r(i)$ is the ball of radius r around agent i . Denote by $\mathcal{Q}^r$ the family of all subsets of the agents that are (finite) intersections of balls of radius r :

$$\mathcal{Q}^r = \{\cap_{k=1}^m B_r(i_k) : m \geq 1, i_1, \dots, i_m \in N\}.$$

We call each $Q \in \mathcal{Q}^r$ a *test region*, as these will be the sets of signals that will determine if players coordinate. On the line graph, test regions are simply the finite intervals of length at most $2r + 1$. On the two dimensional grid, where balls are diamond-shaped, these are 45-degree slanted rectangles (see Figure 1). On any graph, a test region Q includes those agents whose signals are observed by some finite group of agents $i_1, \dots, i_m$, which is the reason they are useful in constructing equilibria that depend on these agents' higher-order beliefs.

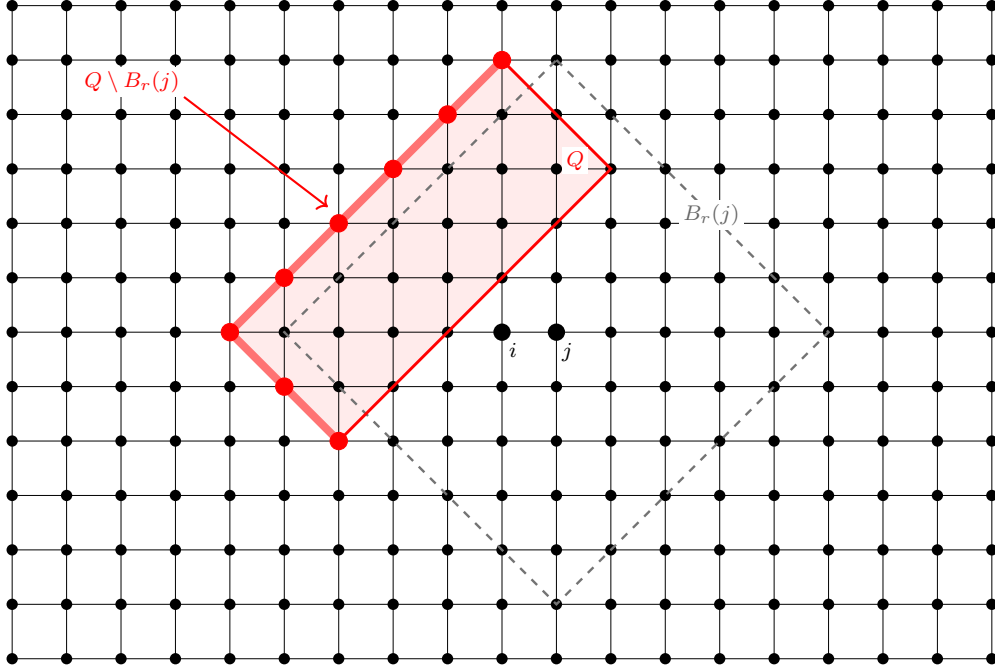


Figure 1: A test region $Q \subset B_r(i)$ is shown in red. The observation neighborhood of j is shown by a dotted line. The highlighted red agents form the nonempty (i, j) -information boundary $Q \setminus B_r(j)$.

We denote by $\mathcal{Q}_i^r \subset \mathcal{Q}^r$ the test regions that are included in $B_r(i)$:

$$\mathcal{Q}_i^r = \{Q \in \mathcal{Q}^r : Q \subseteq B_r(i)\} = \{Q \cap B_r(i) : Q \in \mathcal{Q}^r\}.$$

These are the test regions that are observed by i .

Suppose i and j are neighbors in the graph. Given a test region $Q \in \mathcal{Q}_i^r$ (observed by i), we call $Q \setminus B_r(j)$ the (i, j) -information boundary of Q (see Figure 1). This is the part of Q that is not observed by j . Our condition of *large information boundaries* requires that when Q is large, its information boundaries are large, or else the boundaries are empty.

Definition 2. *The social network graph has (ε, δ) -large information boundaries if for all $r > 0$ and neighbors i, j it holds that for every test region $Q \in \mathcal{Q}_i^r$ of size at least $\varepsilon|B_r(i)|$, the (i, j) -information boundary $Q \setminus B_r(j)$ is either empty, or else is of size at least $\delta \cdot r$.*

The social network graph has large information boundaries if for all $\varepsilon > 0$ there is a $\delta > 0$ such that the graph has (ε, δ) -large information boundaries.

In the case in which the boundary is empty, higher-order beliefs are trivial: Q is completely observed by j . When not all of Q is observed by j , this property requires that the boundary is large, which will help us control j 's higher-order beliefs.

For network common learning we will need two additional geometric conditions. The first is that the collection of test regions $\mathcal{Q}_i^r$ is not too large.

Definition 3. *The social network graph has subexponential- $\mathcal{Q}$ if the size of the collection $\mathcal{Q}_i^r$ satisfies $\lim_{r \rightarrow \infty} (\sup_{i \in N} |\mathcal{Q}_i^r|)^{1/r} = 1$.*

The second is that the size of neighborhoods does not vary too much between agents.

Definition 4. *The social network graph has similar neighborhood sizes if there is a constant $K \geq 1$ such that for all $r > 0$ and for all $i, j \in N$ it holds that $|B_r(i)|/|B_r(j)| \leq K$.*

It is straightforward to verify that all three of these conditions are satisfied by the two-dimensional grid.³ Moreover, they are satisfied by higher dimensional grids, but much more generally by the large collection of graphs that have the same global geometry as these graphs, in precisely the sense captured by these definitions.

Our main theorem is that these conditions are sufficient for network common learning.

Theorem 1. *Suppose that G satisfies the large information boundary condition, the subexponential- $\mathcal{Q}$ condition, and the similar neighborhood sizes condition. Then network common learning is obtained on G .*

The intuition behind the proof of this result is the following. Large information boundaries imply that when agent j reasons about his neighbor i 's belief, and in particular when he reasons about the signals i observed in a test region Q , a large information boundary allows for concentration of measure to reduce the uncertainty j has about the aggregate signal delivered by the unobserved signals.

In particular, to show that network q -belief is obtained, we consider the event that within the ball $B_r(i)$ observed by i , not only are there many high signals (indicating that the state is most likely high) but moreover in each test region $Q \in \mathcal{Q}_i^r$ separately there are many high signals. When the collection of test regions $\mathcal{Q}_i^r$ is small, we show that this happens with high probability. More importantly, this collection of events is self-sustaining, leading to network q -belief. We explain this construction in further detail here, and defer the full proof to the appendix.

Let $M_r = \sup_j |B_r(j)|$ be the size of the largest neighborhood of radius r . The similar neighborhood sizes property ensures that this is finite for each r . Given a finite

³In the coordinates $u = x + y$, $v = x - y$, the ball $B_r(i)$ of $\mathbb{Z}^2$ is a box $\{|u - u_i| \leq r\} \cap \{|v - v_i| \leq r\}$, so test regions are boxes $[u_1, u_2] \times [v_1, v_2]$. Similar neighborhood sizes holds with $K = 1$ since the graph looks the same everywhere. A test region contained in $B_r(i)$ is determined by four coordinates in a window of length $2r + 1$, so $|\mathcal{Q}_i^r| \leq (2r + 1)^4$, giving subexponential- $\mathcal{Q}$. For large boundaries, note that $|B_r(i)| \geq 2r^2$, so a box Q with $|Q| \geq \varepsilon |B_r(i)|$ and side lengths at most $2r + 1$ has both sides of length at least $\varepsilon r/2$. Moving to a neighbor j shifts the box $B_r(j)$ by one unit in u and in v ; hence if $Q \setminus B_r(j)$ is nonempty, Q protrudes past a face of $B_r(j)$, and the boundary contains a full side of Q , of size at least δr with $\delta = \varepsilon/4$.

subset Q of the agents, let $S(Q) = \sum_{j \in Q} s_j$ be the number of high signals observed by the agents in Q . Choosing $\varepsilon > 0$ small enough, the self-sustaining events Y_i^r are given by

$$Y_i^r = \{S(Q) \geq (p - \varepsilon)|Q| - \varepsilon M_r \quad \forall Q \in \mathcal{Q}_i^r\}. \quad (2)$$

That is, there are sufficiently many high signals in every test region observed by i .

Note that in the high state, the expectation of $S(Q)$ is $p|Q|$, which is larger than $(p - \varepsilon)|Q|$, and so for large Q this condition holds with high probability. Note also that if Q is small—in particular smaller than εM_r —then this condition holds trivially, since the right hand side of the inequality is negative. Using the fact that the size of the collection $\mathcal{Q}_i^r$ is small, we show that Y_i^r has high probability.

The crux of the proof is to show that the collection Y_i^r is self-sustaining, i.e., that

$$Y_i^r \subseteq \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} Y_j^r \right) \right).$$

For this, we use the assumption that G has large information boundaries.

4 Coordination Failure on Networks with Bottlenecks

In this section we study graphs on which the only equilibrium is the one in which agents always play the safe action, for any radius of observation. We show that a geometric condition on the graph is sufficient to guarantee this: the graph must contain a long path along which information boundaries are of uniformly bounded size. On such graphs, network common learning is not attained.

Definition 5. Fix $r, m > 0$. A path of distinct agents $i_1, i_2, \dots, i_T$ in the graph G is an (r, m) -bottleneck if

1. $|B_r(i_{t+1}) \setminus B_r(i_t)| \leq m$ for all $t < T$, i.e., agent i_{t+1} has no more than m agents in their r -neighborhood that are not in agent i_t 's r -neighborhood.
2. $B_r(i_1) \cap B_r(i_T) = \emptyset$, i.e., the first and last agent's r -neighborhoods are disjoint.

The first condition in the definition is that the information boundary $B_r(i_{t+1}) \setminus B_r(i_t)$ is of size at most m , i.e., that along the path, r -neighborhoods are highly overlapping, differing by at most m agents. The second condition ensures that the path is long enough so that all signals observed by the first agents are not observed by the last; equivalently, that i_1 and i_T are more than distance $2r$ apart. A simple example is the line graph, which contains an $(r, 1)$ -bottleneck for all r .

Our main result in this section is that on graphs with bottlenecks coordination is impossible.

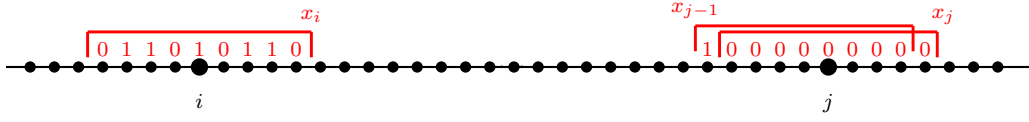


Figure 2: Proof sketch of Theorem 2. The observation x_i is arbitrary, while x_j is fixed to induce a low belief. Conditioned on $j - 1$ observing x_{j-1} , there is sufficiently high probability that j observed x_j , ruling out network q -belief by $j - 1$. By induction, this rules out network q -belief by i , for arbitrary x_i .

Theorem 2. *Suppose that G contains an (r, m) -bottleneck. If $q > 1 - (1 - p)^m$, then $\mathbb{P}(C_{i,r}^q) = 0$ for all i (i.e., network q -belief does not occur) and thus in the unique equilibrium all agents choose the safe action.*

The proof of Theorem 2 appears in the appendix. The idea behind it is the following argument, which is similar to infection arguments used in the literature (e.g., Carlsson and van Damme, 1993).

For this proof sketch, consider the one-sided line graph, i.e., $N = \mathbb{N}$ and $i, j \in N$ are neighbors if $|i - j| = 1$. Fix an $r > 0$ and an agent i , and let $x_i \in O_i^r$ be an arbitrary observation by agent i . To prove the claim, we show that after observing x_i , agent i does not have network q -belief, and therefore must take the safe action.

To this end, consider some agent j who is far enough from agent i so that their observation neighborhoods are disjoint: $B_r(i) \cap B_r(j) = \emptyset$ ($j = i + 2r + 1$ suffices). Let $x_j \in O_j^r$ be any observation by j which implies that j does not have network q -belief. For example, this holds for any x_j which induces a posterior belief that is lower than q : $\mathbb{P}(\theta = H | I_j^r = x_j) < q$. Fix any signal realizations for the agents $i + 1, i + 2, \dots, j - 1$ that are compatible with x_i and x_j , and denote the observations by these agents by $x_{i+1}, \dots, x_{j-1}$; see Figure 2.

The key observation is that because x_j does not induce network q -belief for agent j , from agent $j - 1$'s point of view, this ($I_j^r = x_j$) happens with probability at least $1 - p$: only one signal observed by j is not observed by $j - 1$, and so the probability that $I_j^r = x_j$ conditioned on $I_{j-1}^r = x_{j-1}$ is bounded away from zero, regardless of how large r is. It follows that x_{j-1} does not induce network q -belief in $j - 1$. By induction, x_i does not induce network q -belief in agent i . Since i and x_i are arbitrary, this completes the proof.

As mentioned above, the line graph has a uniform family of bottlenecks, i.e., an $(r, 1)$ -bottleneck for all r . An immediate corollary of Theorem 2 is that the existence of such a uniform family excludes network common learning.

Corollary 1. *Suppose that there exists an $m > 0$ such that for every $r > 0$ the graph contains an (r, m) -bottleneck. Then network common learning is not obtained.*

5 Conclusion

We study global games, played locally on a social network. Agents receive a large amount of information about the state by learning the private signals of agents in their r -neighborhood. As in the classic literature, we ask whether higher-order beliefs become aligned. In particular, we study network common learning, the network analogue of common learning. The answer in our case is that it depends non-trivially on the network geometry.

We do not provide a full characterization of graphs on which network common learning is obtained, instead providing separate necessary and sufficient conditions. A major point for future research is to close this gap. A natural question is whether network common learning is a *quasi-isometry* invariant. That is, is network common learning robust to local changes in the graph that do not change the global geometry?⁴

Our model focuses on a binary state and binary signals. We believe that our network common learning results extend to any finite number of states and finite number of signals. The case of infinitely many signals is more mysterious, as it is in the context of common learning.

⁴Two connected graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are isometric (or isomorphic) if there is a bijection $\varphi: V_1 \rightarrow V_2$ such that $d_1(i, j) = d_2(\varphi(i), \varphi(j))$, where d_i denote graph distances. They are quasi-isometric if there is a map $\varphi: V_1 \rightarrow V_2$ and constants $A \geq 1$, $B \geq 0$ such that for every $i, j \in V_1$

$$\frac{1}{A}d_1(i, j) - B \leq d_2(\varphi(i), \varphi(j)) \leq Ad_1(i, j) + B,$$

and every element of V_2 is of distance at most B from the image of φ .

A Proofs

A.1 Proof of Proposition 1

Proof. Firstly, note that by Lemma 1 the family $(C_{i,r}^q)_{i \in N}$ satisfies the equilibrium condition (1). Thus, it induces an equilibrium.

Next, we prove by mathematical induction that for every agent i , $X_i^r \subseteq C_{i,r}^{q,n}$ for every n and thus $X_i^r \subseteq \bigcap_{n \geq 0} C_{i,r}^{q,n} = C_{i,r}^q$.

For $n = 0$, $X_i^r \subseteq \Omega = C_{i,r}^{q,0}$. Now suppose that for every agent j , $X_j^r \subseteq C_{j,r}^k$ for $k \leq (n - 1)$. Then, $H \cap \left(\bigcap_{j \in N_i} X_j^r \right) \subseteq H \cap \left(\bigcap_{j \in N_i} C_{j,r}^{n-1} \right)$. Since the belief operator is monotone,

$$X_i^r = \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} X_j^r \right) \right) \subseteq \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} C_{j,r}^{q,n-1} \right) \right) = C_{i,r}^{q,n}.$$

Finally, consider any mixed strategy $\sigma_i : I_i^r \rightarrow [0, 1]$ with $\sigma_i(x)$ denoting the probability of choosing action R on observing x and denote $Y = \{x \in O_i^r : \sigma(x) > 0\}$ as the observation on which action R is played with positive probability, we know $Y \subseteq \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} X_j^r \right) \right)$ and the proof for the case of mixed strategy is essentially the same as above. $\square$

A.2 Proof of Theorem 1

To prove Theorem 1 we will need a lemma and a number of claims. We start with the following general lemma.

Lemma A.1. *Consider a family $(Y_i^r)_{i \in N}$ which is self-sustaining, i.e. for every i ,*

$$Y_i^r \subseteq \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} Y_j^r \right) \right).$$

Then, $Y_i^r \subseteq C_{i,r}^q$ for every i .

Proof. Similar to the proof of Proposition 1, we will prove by induction that for every agent i , $Y_i^r \subseteq C_{i,r}^{q,n}$ for every n and thus $Y_i^r \subseteq \bigcap_{n \geq 0} C_{i,r}^{q,n} = C_{i,r}^q$.

For $n = 0$, $Y_i^r \subseteq \Omega = C_{i,r}^{q,0}$. Now suppose that for every agent j , $Y_j^r \subseteq C_{j,r}^k$ for $k \leq (n - 1)$. Then, $H \cap \left(\bigcap_{j \in N_i} Y_j^r \right) \subseteq H \cap \left(\bigcap_{j \in N_i} C_{j,r}^{n-1} \right)$. Since the belief operator is monotone and $(Y_i^r)_{i \in N}$ are self-sustaining,

$$Y_i^r \subseteq \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} Y_j^r \right) \right) \subseteq \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} C_{j,r}^{q,n-1} \right) \right) = C_{i,r}^{q,n}.$$

It follows that $Y_i^r \subseteq C_{i,r}^{q,n}$ for every n and hence $Y_i^r \subseteq C_{i,r}^q$. $\square$

Suppose G has the similar neighborhood sizes property with constant K , the large information boundaries property, and subexponential- $\mathcal{Q}$. Choose $\varepsilon < (p - 1/2)/(K + 1)$ and Y_i^r as in (2), i.e.

$$Y_i^r = \left\{ S(Q) \geq (p - \varepsilon)|Q| - \varepsilon(\sup_j |B_r(j)|) \quad \forall Q \in \mathcal{Q}_i^r \right\}.$$

Since the event Y_i^r is measurable with respect to $\sigma(I_i^r)$, the information available to agent i with radius of observation r , we can identify it with a subset of $O_i^r = \{0, 1\}^{B_r(i)}$, which we also denote by Y_i^r .

Denote $M_r = \sup_j |B_r(j)|$ and $\eta = p - \varepsilon$, so that

$$Y_i^r = \{S(Q) \geq \eta|Q| - \varepsilon M_r \quad \forall Q \in \mathcal{Q}_i^r\}.$$

We first note that the set $Y_i^r \subseteq O_i^r$ is non-empty. To see this, consider the configuration $x \in O_i^r$ in which all signals are high. Then, for every $Q \in \mathcal{Q}_i^r$, we have

$$S(Q) = |Q| \geq \eta|Q| - \varepsilon M_r,$$

implying that $x \in Y_i^r$.

The next three claims establish properties of Y_i^r which are needed to show that it is self-sustaining. The first shows that on Y_i^r , agents learn the state, and furthermore learning is uniform over the agents and the elements of Y_i^r .

Claim A.1.

$$\lim_{r \rightarrow \infty} \inf_{i \in N} \inf_{x \in Y_i^r} \mathbb{P}(\theta = H \mid I_i^r = x) = 1.$$

Proof. Consider $Q = B_r(i) \in \mathcal{Q}_i^r$. Recall from the definition of similar neighborhood sizes, there exists a constant K such that for all i, j and r we have $|B_r(i)|/|B_r(j)| \leq K$. In particular, $K \geq M_r/|B_r(i)|$. Then, for every $x \in Y_i^r$, we have

$$S(B_r(i)) \geq \eta|B_r(i)| - \varepsilon M_r \geq \eta|B_r(i)| - \varepsilon K|B_r(i)| = (p - (K + 1)\varepsilon)|B_r(i)|.$$

Thus, for every $x \in Y_i^r$,

$$\frac{\mathbb{P}(\theta = H \mid I_i^r = x)}{\mathbb{P}(\theta = L \mid I_i^r = x)} = \left(\frac{p}{1 - p} \right)^{2S(B_r(i)) - |B_r(i)|} \geq \left(\frac{p}{1 - p} \right)^{(2p - 2(K + 1)\varepsilon - 1)|B_r(i)|}.$$

Note that $2p - 2(K + 1)\varepsilon - 1 > 0$ because $\varepsilon < (p - 1/2)/(K + 1)$, $|B_r(i)|$ is increasing in r (uniformly over i , by similar neighborhood sizes) and $p > 1/2$. Thus, the odds ratio goes to infinity as r increases. Hence, the posterior probability of the high state goes to 1, uniformly over i and x :

$$\lim_{r \rightarrow \infty} \inf_{i \in N} \inf_{x \in Y_i^r} \mathbb{P}(\theta = H \mid I_i^r = x) = 1.$$

□

The next claim shows that on Y_i^r , agents place high belief in Y_j^r (for a neighbor j), and that this is again uniform over the agents and over Y_i^r .

Claim A.2.

$$\lim_{r \rightarrow \infty} \inf_{i \in N} \inf_{x \in Y_i^r} \mathbb{P}(I_j^r \in Y_j^r, \forall j \in N_i \mid I_i^r = x, \theta = H) = 1.$$

Proof. Since the number of i 's neighbors is finite (and uniformly bounded over i , by similar neighborhood sizes for $r = 1$), it suffices to fix one neighbor $j \in N_i$ and show that

$$\lim_{r \rightarrow \infty} \sup_{i \in N} \sup_{x \in Y_i^r} \mathbb{P}(I_j^r \notin Y_j^r \mid I_i^r = x, \theta = H) = 0.$$

Consider $j \in N_i$. For each $Q \in \mathcal{Q}_j^r$, let

$$T = Q \cap B_r(i), \quad F = Q \setminus B_r(i).$$

Thus, T is the part of Q observed by both i and j , while F is the (j, i) -information boundary of Q , i.e. the part observed by j but not by i . Moreover, $Q = T \cup F$ and $|Q| = |T| + |F|$.

To show the desired convergence, we estimate the conditional probability that $I_j^r \notin Y_j^r$. For this to happen, some $Q \in \mathcal{Q}_j^r$ must violate the threshold property. Thus,

$$S(Q) < \eta|Q| - \varepsilon M_r.$$

We claim that this forces F to have too few high signals, namely $S(F) < \eta|F|$. Indeed, since $x \in Y_i^r$, the threshold inequality applied to T gives

$$S(T) \geq \eta|T| - \varepsilon M_r.$$

Thus, if $S(F) \geq \eta|F|$, we would have

$$S(Q) = S(T) + S(F) \geq \eta|T| - \varepsilon M_r + \eta|F| = \eta|Q| - \varepsilon M_r,$$

contradicting the fact that Q violates the threshold inequality. Hence the event $I_j^r \notin Y_j^r$ requires some $Q \in \mathcal{Q}_j^r$ with $S(F) < \eta|F|$.

Now we show that F is sufficiently large. If Q violates j 's threshold inequality, then $0 \leq S(Q) < \eta|Q| - \varepsilon M_r$ and hence

$$|Q| > \frac{\varepsilon}{\eta} M_r \geq \varepsilon M_r \geq \varepsilon |B_r(j)|.$$

Now, $F = Q \setminus B_r(i)$ is either empty or of size at least δr , where δ is given by the large information boundaries condition. Note that $F \neq \emptyset$. Otherwise, $Q = T$, and the threshold inequality for $x \in Y_i^r$ implies that Q does not violate j 's inequality, i.e. $S(T) = S(Q) \geq \eta|Q| - \varepsilon M_r$. Therefore $|F| \geq \delta r$.

Next, we show that within each set F whose size is linearly increasing in r , having too few signals is exponentially unlikely. Conditional on $\theta = H$, the high signals in F are independent and occur with probability p , so $S(F) \sim \text{Bin}(|F|, p)$. Since $\eta < p$, the Chernoff bound gives

$$\mathbb{P}(S(F) < \eta|F|) \leq \exp(-D(\eta||p)|F|) \leq \exp(-D(\eta||p)\delta r),$$

where $D(\cdot||\cdot)$ is the Kullback-Leibler divergence.

Hence, by a union bound, we get

$$\begin{aligned} \mathbb{P}(I_j^r \notin Y_j^r | I_i^r = x, \theta = H) &\leq \sum_{Q \in \mathcal{Q}_j^r} \mathbb{P}(S(Q \setminus B_r(i)) < \eta|Q \setminus B_r(i)|) \\ &\leq |\mathcal{Q}_j^r| \cdot \exp(-D(\eta||p)\delta r) \rightarrow 0 \quad \text{as } r \rightarrow \infty, \end{aligned}$$

where the convergence holds because the social network satisfies the subexponential- $\mathcal{Q}$ condition, and furthermore holds uniformly over i . □

The last of our three claims shows that the probability of Y_i^r converges to 1 in the high state, and that this convergence is uniform over i .

Claim A.3.

$$\lim_{r \rightarrow \infty} \inf_{i \in N} \mathbb{P}(I_i^r \in Y_i^r | \theta = H) = 1.$$

Proof. By a union bound and the Chernoff inequality,

$$\begin{aligned} \mathbb{P}(I_i^r \notin Y_i^r | \theta = H) &\leq \sum_{Q \in \mathcal{Q}_i^r} \mathbb{P}(S(Q) < \eta|Q| - \varepsilon M_r | \theta = H) \\ &= \sum_{\substack{Q \in \mathcal{Q}_i^r: \\ |Q| > \frac{\varepsilon}{\eta} M_r}} \mathbb{P}(S(Q) < \eta|Q| - \varepsilon M_r | \theta = H) \\ &\leq \sum_{\substack{Q \in \mathcal{Q}_i^r: \\ |Q| > \frac{\varepsilon}{\eta} M_r}} \mathbb{P}(S(Q) < \eta|Q| | \theta = H) \\ &\leq \sum_{\substack{Q \in \mathcal{Q}_i^r: \\ |Q| > \frac{\varepsilon}{\eta} M_r}} \exp(-D(\eta||p)|Q|) \\ &\leq |\mathcal{Q}_i^r| \exp\left(-D(\eta||p)\frac{\varepsilon}{\eta} M_r\right) \\ &\leq |\mathcal{Q}_i^r| \exp\left(-D(\eta||p)\frac{\varepsilon}{\eta}(r+1)\right). \end{aligned}$$

By the subexponential- $\mathcal{Q}$ condition, the last expression goes to 0 as r goes to infinity, and does so uniformly over i . □

Fix any $q \in (0, 1)$. We use the family $(Y_i^r)_{i \in N}$ constructed in (2) to establish network q -learning. First, we show that this family is self-sustaining. Using the identity

$$\mathbb{P}(H, I_j^r \in Y_j^r \forall j \in N_i | I_i^r = x) = \mathbb{P}(H | I_i^r = x) \mathbb{P}(I_j^r \in Y_j^r \forall j \in N_i | I_i^r = x, \theta = H)$$

and Claims A.1 and A.2, we obtain

$$\lim_{r \rightarrow \infty} \inf_{i \in N} \inf_{x \in Y_i^r} \mathbb{P}(H, I_j^r \in Y_j^r \forall j \in N_i | I_i^r = x) = 1.$$

Since $q < 1$, it follows that, for all sufficiently large r , $Y_i^r \subseteq \mathcal{B}_{i,r}^q \left(H \cap \left(\bigcap_{j \in N_i} Y_j^r \right) \right)$ for all i , meaning that the family $(Y_i^r)_{i \in N}$ is self-sustaining. By Lemma A.1, $Y_i^r \subseteq C_{i,r}^q$. Therefore, by Claim A.3,

$$\mathbb{P}(C_{i,r}^q | \theta = H) \geq \mathbb{P}(I_i^r \in Y_i^r | \theta = H) \rightarrow 1 \quad \text{as } r \rightarrow \infty.$$

Hence, network q -learning of H is obtained. Since $q \in (0, 1)$ was arbitrary, network common learning of H occurs. This completes the proof of Theorem 1.

A.3 Proof of Lemma 1

Proof. We first claim that the sequence of sets $(C_{j,r}^{q,n})_n$ is decreasing. This follows from the fact that the $\mathcal{B}_{i,r}^q$ operator is monotone with respect to set inclusion.

Let $E_n = H \cap \left(\bigcap_{j \in N_i} C_{j,r}^{q,n} \right)$. Since $(C_{j,r}^{q,n})_n$ is decreasing, $(E_n)_{n \geq 0}$ is also decreasing. Hence, $E_n \downarrow \cap_n E_n$. Next, by continuity from above,

$$\mathbb{P}(\cap_{n \geq 0} E_n | I_i^r) = \lim_{n \rightarrow \infty} \mathbb{P}(E_n | I_i^r) = \inf_{n \geq 0} \mathbb{P}(E_n | I_i^r).$$

Thus,

$$\begin{aligned} \omega \in \bigcap_{n \geq 0} \mathcal{B}_{i,r}^q(E_n) &\iff \forall n : \mathbb{P}(E_n | I_i^r)(\omega) \geq q \iff \inf_{n \geq 0} \mathbb{P}(E_n | I_i^r)(\omega) \geq q \iff \\ &\iff \mathbb{P}(\cap_{n \geq 0} E_n | I_i^r)(\omega) \geq q \iff \omega \in \mathcal{B}_{i,r}^q(\cap_{n \geq 0} E_n). \end{aligned}$$

So, $\cap_{n \geq 0} \mathcal{B}_{i,r}^q(E_n) = \mathcal{B}_{i,r}^q(\cap_{n \geq 0} E_n)$. Finally,

$$\begin{aligned} C_{i,r}^q &= \Omega \cap \bigcap_{n \geq 1} C_{i,r}^{q,n} = \bigcap_{n \geq 0} C_{i,r}^{q,n+1} = \bigcap_{n \geq 0} \mathcal{B}_{i,r}^q(E_n) = \mathcal{B}_{i,r}^q(\cap_{n \geq 0} E_n) = \\ &= \mathcal{B}_{i,r}^q \left(\bigcap_{n \geq 0} \left(H \cap \bigcap_{j \in N_i} C_{j,r}^n \right) \right) = \mathcal{B}_{i,r}^q \left(H \cap \bigcap_{j \in N_i} \bigcap_{n \geq 0} C_{j,r}^n \right) = \mathcal{B}_{i,r}^q \left(H \cap \bigcap_{j \in N_i} C_{j,r} \right). \end{aligned}$$

□

A.4 Proof of Theorem 2

Proof. Consider two agents i, j , and signal realizations $x_i \in O_i^r$ and $x_j \in O_j^r$ for each agent. We say that x_i and x_j are *compatible* if they agree on the overlapping region $B_r(i) \cap B_r(j)$.

By the definition of $C_{i,r}^{q,n}$, the event that i has n^{th} order network q -belief, $C_{i,r}^{q,n}$ is $\sigma(I_i^r)$ -measurable. That is, this event depends only on the information available to agent i . As such, we can identify this event with a subset of O_i^r , the possible observations of agent i . Accordingly, for $x \in O_r^i$ we write $x \in C_{i,r}^{q,n}$ whenever agent i has n^{th} order network q -belief upon observing x .

By the theorem hypothesis, there is an (r, m) -bottleneck path $i_1, i_2, \dots, i_T$. Consider two adjacent agents along the bottleneck path, i_t and i_{t+1} . Let $x_t \in O_{i_t}^r$ and $x_{t+1} \in O_{i_{t+1}}^r$ be compatible. Suppose that $x_{t+1} \notin C_{i_{t+1},r}^q$, i.e., when agent i_{t+1} observes x_{t+1} , i_{t+1} does not have network q -belief. We claim that this implies that $x_t \notin C_{i_t}^q$, i.e., agent i_t does not have network q -belief upon observing the compatible x_t . Indeed, since $|B_r(i_{t+1}) \setminus B_r(i_t)| \leq m$, and the signals observed by i_{t+1} but not by i_t are independent conditional on the state, it follows that

$$\mathbb{P}(I_{i_{t+1}}^r = x_{t+1} | I_{i_t}^r = x_t, H) = \mathbb{P}(\{s_l = x_{i_{t+1},l} \mid \forall l \in (B_r(i_{t+1}) \setminus B_r(i_t))\} | H) \geq (1-p)^m.$$

By our assumption that $x_{t+1} \notin C_{i_{t+1}}^q$,

$$\mathbb{P}(C_{i_{t+1},r}^q | I_{i_t}^r = x_t, H) \leq 1 - \mathbb{P}(I_{i_{t+1}}^r = x_{t+1} | I_{i_t}^r = x_t, H) \leq 1 - (1-p)^m,$$

and hence

$$\begin{aligned} \mathbb{P}\left(H \cap \bigcap_{j \in N_{i_t}} C_{j,r}^q \mid I_{i_t}^r = x_t\right) &\leq \mathbb{P}(C_{i_{t+1},r}^q | I_{i_t}^r = x_t, H) \mathbb{P}(H | I_{i_t}^r = x_t) \leq \\ &\leq 1 - (1-p)^m < q. \end{aligned}$$

Thus

$$x_t \notin \mathcal{B}_{i_t,r}^q \left(H \cap \left(\bigcap_{j \in N_{i_t}} C_{j,r}^q \right) \right) = C_{i_t,r}^q,$$

by the fixed-point property of $C_{i,r}^q$ (Lemma 1). We have thus shown that

$$x_{t+1} \notin C_{i_{t+1},r}^q \implies x_t \notin C_{i_t,r}^q. \quad (3)$$

Fix an arbitrary observation $x_1 \in O_{i_1}^r$ for agent i_1 . Since $B_r(i_1) \cap B_r(i_T) = \emptyset$, there exists a signal configuration on $\bigcup_t B_r(i_t)$ that agrees with x_1 on $B_r(i_1)$ and with 0 (the all zeros observation) on $B_r(i_T)$. Hence, there are $x_2, \dots, x_T$ such that $x_T = 0$ and $x_1, x_2, \dots, x_T$ are compatible. Note that $0 \notin C_{i_T,r}^q$, since when an agent observes all zero signals, their posterior belief in H is below one half, and hence below q .

Thus, using (3), it follows by induction that $x_1 \notin C_{i_1,r}^q$:

$$x_T \notin C_{i_T,r}^q \implies x_{T-1} \notin C_{i_{T-1},r}^q \implies \dots \implies x_1 \notin C_{i_1,r}^q$$

Since the choice of x_1 is arbitrary, we have shown that $C_{i_1,r}^q = \emptyset$: i_1 never has network q -belief. It then follows from the fixed point property of these sets (Lemma 1), that the same holds for i_1 's neighbors, and by induction for all the agents.

Finally, by Proposition 1, it follows that all agents play the safe action in every equilibrium. $\square$

References

- G.-M. Angeletos, C. Hellwig, and A. Pavan. Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks. *Econometrica*, 75(3):711–756, 2007. doi: 10.1111/j.1468-0262.2007.00766.x.
- L. E. Blume. The statistical mechanics of strategic interaction. *Games and economic behavior*, 5(3):387–424, 1993.
- H. Carlsson and E. van Damme. Global games and equilibrium selection. *Econometrica*, 61(5):989–1018, Sept. 1993. doi: 10.2307/2951491.
- M. S.-Y. Chwe. Communication and coordination in social networks. *The Review of Economic Studies*, 67(1):1–16, 2000.
- M. W. Cripps, J. C. Ely, G. J. Mailath, and L. Samuelson. Common learning. *Econometrica*, 76(4):909–933, 2008.
- M. W. Cripps, J. C. Ely, G. J. Mailath, and L. Samuelson. Common learning with intertemporal dependence. *International Journal of Game Theory*, 42(1):55–98, 2013.
- G. Ellison. Learning, local interaction, and coordination. *Econometrica: Journal of the Econometric Society*, pages 1047–1071, 1993.
- M. Frick, R. Iijima, and Y. Ishii. Learning efficiency of multiagent information structures. *Journal of Political Economy*, 131(12):3377–3414, 2023. doi: 10.1086/725171.
- I. Goldstein and A. Pauzner. Demand–deposit contracts and the probability of bank runs. *The Journal of Finance*, 60(3):1293–1327, 2005. doi: 10.1111/j.1540-6261.2005.00762.x.
- J. Y. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. In *Proceedings of the Third Annual ACM Symposium on Principles of Distributed Computing*, pages 50–61, 1984.

- T. Hutchcroft, O. Rospuskova, and O. Tamuz. Local coordination and the geometry of social networks. 2026.
- A. Kajii and S. Morris. The robustness of equilibria to incomplete information. *Econometrica: Journal of the Econometric Society*, pages 1283–1309, 1997.
- W. Kets. Robustness of equilibria in anonymous local games. *Journal of Economic Theory*, 146(1):300–325, 2011.
- D. Monderer and D. Samet. Approximating common knowledge with common beliefs. *Games and Economic behavior*, 1(2):170–190, 1989.
- S. Morris. Contagion. *The Review of Economic Studies*, 67(1):57–78, 2000.
- S. Morris and H. S. Shin. Global games: Theory and applications. In M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, editors, *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, volume 1, pages 56–114. Cambridge University Press, Cambridge, 2003. doi: 10.1017/CBO9780511610240.004.
- S. Morris, R. Rob, and H. S. Shin. p-dominance and belief potential. *Econometrica: Journal of the Econometric Society*, pages 145–157, 1995.
- D. Oyama and S. Takahashi. Contagion and uninvadability in local interaction games: The bilingual game and general supermodular games. *Journal of Economic Theory*, 157:100–127, 2015.
- A. Rubinstein. The electronic mail game: Strategic behavior under “almost common knowledge”. *The American Economic Review*, 79(3):385–391, 1989.
- J. Steiner and C. Stewart. Communication, timing, and common learning. *Journal of Economic Theory*, 146(1):230–247, 2011. doi: 10.1016/j.jet.2010.08.009.