Naïve Learning in Social Networks: Convergence, Influence, and the Wisdom of Crowds*

Benjamin Golub[†] and Matthew O. Jackson[‡]

January 14, 2007
Revised: May 29, 2007

Abstract

We study learning and influence in a setting where agents communicate according to an arbitrary social network and naïvely update their beliefs by repeatedly taking weighted averages of their neighbors' opinions. A focus is on conditions under which beliefs of all agents in large societies converge to the truth, despite their naïve updating. We show that this happens if and only if the influence of the most influential agent in the society is vanishing as the society grows. Using simple examples, we identify two main obstructions which can prevent this. By ruling out these obstructions, we provide general structural conditions on the social network that are sufficient for convergence to truth. In addition, we show how social influence changes when some agents redistribute their trust, and we provide a complete characterization of the social networks for which there is a convergence of beliefs. Finally, we survey some recent structural results on the speed of convergence and relate these to issues of segregation, polarization and propaganda.

JEL classification Numbers: D85, D83, A14, L14, Z13.

Keywords: social networks, learning, diffusion, bounded rationality.

---

*We thank Francis Bloch, Antoni Calvo-Armengol, Drew Fudenberg, Tim Roughgarden, and Jonathan Weinstein for helpful comments and suggestions. Financial support under NSF grant SES-0647867 is gratefully acknowledged.

[†]Division of the Humanities and Social Sciences, Caltech, Pasadena, CA, 91125.
http://www.its.caltech.edu/∼beng e-mail: beng@its.caltech.edu

[‡]Department of Economics, Stanford University, Stanford, CA, 94305.
http://www.stanford.edu/∼jacksonm e-mail: jacksonm@stanford.edu

1 Introduction

Social networks are primary conduits of information, opinions, and behaviors. They carry news about products, jobs, and various social programs; influence decisions to become educated, to smoke, and to commit crimes; and drive political opinions and attitudes toward other groups, just to mention a few of their effects. In view of this, it is important to understand how beliefs and behaviors evolve over time, how this depends on the network structure, and whether or not the resulting outcomes are fully efficient.

Given the complex forms that social networks often take, it can be difficult for the agents involved, or for a modeler, to update beliefs properly based on communication in a network. For example, Choi, Gale and Kariv (2007; 2005) find that although subjects in simple three-person networks update fairly well in some circumstances, they do not do so well in evaluating repeated observations and judging indirect information whose origin is uncertain. Given that our social networks involve repeated transfers of information among large numbers of individuals, fully rational learning becomes infeasible at best. Nonetheless, it can still be that fairly simple updating rules can lead to eventual outcomes like those achieved through fully rational learning. In this paper we examine these questions with respect to a variation of a model of network influence that has its roots in measures of centrality and prestige developed by Katz (1953) and Bonacich (1987), and which is also also related to models of social influence and persuasion by French (1956), Harary (1959), and Friedkin and Johnsen (1997). More recently, a variation on this model has been analyzed as a model of information transmission and opinion formation by DeMarzo, Vayanos and Zwiebel (2003).

In the variation of the model that we analyze, agents update their beliefs or attitudes in each period simply by taking weighted averages of their neighbors' opinions from the previous period, possibly placing some weight on their own previous beliefs. While the agents in this scenario are boundedly rational, failing to adjust correctly for repetitions and dependencies in information that they hear multiple times,[1] we show that this process can still lead them to fully accurate beliefs in the limit as society grows large. Moreover, the limiting properties of this process are not only useful as a model of belief evolution, but also as a basis for analyzing the influence or power of the different individuals comprising a network.[2]

Our contributions are outlined as follows, in the order in which they appear in the paper.

[1]See Friedkin and Johnsen (1997) and DeMarzo, Vayanos and Zwiebel (2003) for more background discussion on this type of bounded rationality.

[2]For instance, it can also be viewed as a myopic best-response dynamic for a game where agents care about matching the behavior of those in their social network (possibly placing some weight on themselves).

First, much of the previous literature has worked under an assumption that at least some agents always place some weight on their own opinions when updating. This is often for convenience, as that assumption guarantees convergence of beliefs by appealing to some standard results on Markov chains. While we might expect this to be true for some agents in some situations, there are clearly many applications where agents start without information or believe that others may be better informed and thus defer to the opinions of others. This turns out to be important in determining whether or not beliefs converge. Our first contribution is to provide a complete characterization of convergence. That is, we identify the conditions that are necessary and sufficient for convergence of an iteration of an arbitrary stochastic matrix. This characterization applies not only to the belief-updating application of this paper, but also to the stability of other measures of centrality and other Markov chain applications. While the basic idea of the condition is known in the Markov chain literature, we have not seen it stated in the form we give, which is particularly relevant in the social network context. When beliefs do converge, they converge to a certain weighted average of agents' initial beliefs, and the weights correspond naturally to a measure of social influence or importance.

In the second section, which contains the main novel theoretical results of the paper, we study when a large society of naïve updaters will actually converge so that all individuals learn the true state of nature, assuming they all start with different noisy signals about this state. There is a difference between having all agents converge to the same belief and having all agents converge to the correct belief. If the limiting belief tends to the correct belief as a society grows large, we call the society wise. The section contains a complete characterization of wisdom in terms of influence weights: a society is wise if and only if the influence of the most influential agent is vanishing as the society grows. Building on this characterization, we then focus on the relationship between social structure and wisdom. Using simple examples, we identify two main obstructions that can prevent a society from being wise. One of them is the existence of extreme imbalances in trust, with some groups getting a very disproportionate share of attention. The other main obstruction is a lack of dispersion: when small groups do not pay sufficient attention to the rest of the world. Examples illustrate that either problem, even in the absence of the other, can prevent wisdom. With this in mind, we can formulate general structural conditions which rule out these obstructions. The first type of condition requires a minimal amount of balance, and the second type of condition requires a minimal degree of dispersion. Assuming that precise versions of these conditions hold, we prove that as a society grows, its limiting beliefs become arbitrarily accurate. That is, they correctly aggregate the information that might initially be dispersed throughout the network. These results are in contrast with Theorem 2 of DeMarzo, Vayanos and Zwiebel (2003), which says that consensus beliefs (for a fixed population of n agents) are correct only if a knife-edge restriction on the weights holds. More generally, our conclusions differ from a long line of previous work which suggests that the sufficient conditions for naïve learning to happen are very strong.[3] We show that beliefs can be correct in the large-society limit for a fairly broad collection of networks.

Third, we apply the model to study the effects of changes in the weights agents give to their neighbors' beliefs. We give a simple matrix calculation which can always be used to determine how social influence changes when some agents redistribute their trust. We obtain two interpretable corollaries by considering specific perturbations, showing that an agent's social influence weakly increases when some agents listen to him more at the expense of others. Moreover, we show quite generally that the impact of a redistribution of trust on others' social influences is directly proportional to the influence of the agent doing the redistributing. In the process, we derive some simple and implementable summation formulas for computing perturbations of the limiting distribution of strongly connected Markov chains, and find the signs of these perturbations. To our knowledge, these mathematical results have not appeared elsewhere and extend the theory of Markov processes.

Finally, we survey several recent and useful results on the dynamics of the updating process studied here. In particular, for situations where there is convergence, we give explicit upper and lower bounds on the rate of convergence using standard results related to the second largest eigenvalue of the matrix of network interactions. We then describe how a theorem on Markov chains by Hartfiel and Meyer (1998) can be used to relate second eigenvalues to the structure of society. Building on this, we can deduce that convergence is "slow" if society is divided into several mutually distrustful factions. We also study some results on fast convergence, and finally discuss how these conclusions provide a quantitative explanation for some forms of polarization and propaganda. Most of the material in this section is drawn from recent mathematical results throughout the Markov chain and computer science literatures. As these have not previously been collected and discussed in relation to models of information transmission or social centrality, our contribution in this part is primarily expository.

Our work relates to several lines of research other than the ones already discussed. There is a large theoretical literature on social learning, both fully and boundedly rational, and a number of studies investigating learning in the context of social networks. Similarly to Ellison and Fudenberg (1993; 1995), we examine updating that is somewhat myopic. This

[3]See Sobel (2000) for a comprehensive survey.

is also in the spirit of Bala and Goyal (1998; 2001) in allowing arbitrarily complex network structures.

In many of these and other social learning models (e.g., Banerjee (1992), Gale and Kariv (2003), Celen and Kariv (2004), and Banerjee and Fudenberg (2004)), agents converge to holding the same belief or at least the same judgment as to an optimal action. These conclusions often rely on observational learning, so that agents are observing choices or payoffs over time and updating accordingly.[4] Our results are quite different from these. In contrast to the observational learning models, convergence and the efficiency of learning in our model depend critically on the details of the network architecture and on the influence of various agents.

In addition to the study by DeMarzo, Vayanos and Zwiebel (2003) that we have already discussed, there is work by Collignon and Al-Sadoon (2006) the examines a similar model. They concentrate on when it is that each individual exerts some influence in a society, so their focus differs from ours. The closest point of overlap is that they present some simulations related to rates of convergence, an issue which we discuss from a theoretical perspective in Section 6.

In sociology, since the work of Katz (1953), French (1956), and Bonacich (1987), eigenvector-like notions of centrality and prestige have been analyzed.[5] As some such models are based on convergence of iterated influence relationships, our results provide foundations for understanding when convergence is obtained.

Finally, there is an enormous empirical literature about the influence of social acquaintances on behavior and outcomes that we will not attempt to survey here, but simply point out that our model provides testable predictions about the relationships between social structure and social learning.

2 The Model

2.1 Agents and Interaction

A finite set $A = \{1, 2, \ldots, n\}$ of agents interact according to a social network. The agents are the nodes of a directed graph. The interaction patterns are captured through an $n \times n$ nonnegative matrix $\mathbf{T}$. The matrix $\mathbf{T}$ may be directed so that $T_{ij} > 0$ while $T_{ji} = 0$. We refer to $\mathbf{T}$ as the *interaction* matrix.

[4]For a general version of the observational learning approach, see Rosenberg, Solan and Vieille (2006).

[5]See also Wasserman and Faust (1994), Bonacich and Lloyd (2001) and Jackson (2007) for more recent elaborations.

In what follows, we examine the case where $\mathbf{T}$ is a stochastic matrix, so that it its entries across each row sum to one – so the assumption amounts to a normalization. As discussed below, this type of matrix is particularly relevant in a situation where agents are updating beliefs by taking weighted averages; the entry T_{ij} is the weight or trust that agent i places on the current belief of agent j in forming his or her belief for the next period.

2.2 Updating Processes

Let us discuss the application of this framework to belief updating. DeMarzo, Vayanos and Zwiebel (2003) examine a model where the agents in the network are trying to estimate some unknown parameter μ. The belief held by agent i at time t is $p_i(t)$, and the vector of beliefs at time t is written $\mathbf{p}(t)$. The updating rule is:

$$\mathbf{p}(t) = \left[(1-\lambda_t)\mathbf{I} + \lambda_t \widehat{\mathbf{T}}\right] \mathbf{p}(t-1)$$

where $\lambda_t \in (0,1]$ and $\widehat{\mathbf{T}}$ is a stochastic matrix. In the case where λ_t is constant over time, this can be written as

$$\mathbf{p}(t) = \mathbf{T}\mathbf{p}(t-1) = \mathbf{T}^t \mathbf{p}(0), \tag{1}$$

where $\mathbf{T}$ has strictly positive diagonal entries.[6] This will be the updating rule studied in this paper, though in the ensuing analysis, we will drop the assumption on the diagonal entries of $\mathbf{T}$.

The evolution of beliefs can be motivated by the following Bayesian setup discussed by DeMarzo, Vayanos and Zwiebel (2003). At the beginning of the evolution, $t = 0$, each agent receives a noisy signal $p_i(0) = \mu + e_i$ where $e_i \in \mathbb{R}$ is a noise term with expectation zero. Agent i hears the opinions of the agents with whom he interacts, and assigns precision π_{ij} to each one of them. These subjective estimates may, but need not, coincide with the true precisions of their signals. If agent i does not listen to agent j, he gives him precision $\pi_{ij} = 0$. In the case where the signals are normal, Bayesian updating from independent signals at $t = 1$ entails the rule (1) with $T_{ij} = \pi_{ij} / \sum_{k=1}^{n} \pi_{ik}$.

The key behavioral assumption is that the agents continue using this rule throughout the evolution. That is, they do not adjust the precision estimates or account for the possible

[6] DeMarzo, Vayanos and Zwiebel (2003) examine the case in which $p_i(t)$ is a vector instead of a scalar, to permit the analysis of multidimensional opinions. Then each agent has a vector of beliefs, not just a single estimate. The main conclusion of the study is that all the components of such a belief vector, in the long run, behave in essentially the same way. As the focus of our analysis is different, it is sufficient to consider a single dimension and the extension to many dimensions is direct.

repetition of information and for the "cross-contamination" of their neighbors' signals. This bounded rationality arising from persuasion bias is discussed at length by DeMarzo, Vayanos and Zwiebel (2003), and so we do not reiterate that here.

Friedkin and Johnsen (1997) examine a related model where social attitudes depend on the attitudes of neighbors and evolve over time. Let $\mathbf{p}(0) \in \mathbb{R}^n_+$ be a vector indicating agents' beliefs or attitudes. Let $\mathbf{D}$ be an $n \times n$ matrix where entries are only positive along the diagonal, and $D_{ii} \in (0,1)$ indicates the extent to which agent i pays attention to others' attitudes. The evolution is described by

$$\mathbf{p}(t) = \mathbf{DTp}(t-1) + (\mathbf{I} - \mathbf{D})\mathbf{p}(0)$$

So, agents start with attitudes $\mathbf{p}(0)$ and then mix in some of their neighbors' recent attitudes with their starting attitudes. We rewrite this as

$$\mathbf{p}(t) = [\mathbf{DT}]^t \, \mathbf{p}(0) + \sum_{i=0}^{t-1} [\mathbf{DT}]^i \, (\mathbf{I} - \mathbf{D}) \, \mathbf{p}(0),$$

and so the behavior of $\mathbf{p}(t)$ depends on the powers of $\mathbf{DT}$. Thus, to understand the behavior of $\mathbf{p}(t)$ over time it is important to understand the properties of $(\mathbf{DT})^t$, which are the object of study in this paper.

It is important to note that other applications also have the same form as that here. For instance, Google's "PageRank" system is analagous to the influence vectors from Theorem 3 below, where the $\mathbf{T}$ matrix is the normalized link matrix.[7] This corresponds to defining the influence of node i as the limit of $\mathbf{T}^t$ times a unit vector $\mathbf{p}(0)$ where the 1 is placed in the i-th entry and other entries are set to 0. Other citation and influence measures also have similar such eigenvector bases (e.g., see Palacios-Huerta and Volij (2004)).[8]

2.3 Walks, Paths and Cycles

The following are standard graph-theoretic definitions applied to the directed graph of connections induced by the interaction matrix $\mathbf{T}$.

A *walk* in $\mathbf{T}$ is a sequence of nodes $B = i_1, i_2, \ldots, i_K$, not necessarily distinct, such that

[7] So $T_{ij} = 1/\ell_i$ if page i has a link to page j, where ℓ_i is the number of links that page i has to other pages.

[8] We also see iterative interaction matrices in of recursive utility (e.g., Rogers (2006)) and in strategic games played by agents on networks where influence measures turn out to be important (e.g., Ballester, Calvò-Armengol and Zenou (2006)). In such applications understanding the properties of $\mathbf{T}^t$ and related matrices is critical.

$T_{i_k i_{k+1}} > 0$ for each $k \in \{1, \ldots, K-1\}$. We write $i \in B$ for a node i and walk B if i is a node in the sequence B, and we say that a walk $B = i_1, i_2, \ldots, i_K$ goes *from* i_1 *to* i_K. The length of the walk is defined to be $K-1$.

A *path* in $\mathbf{T}$ is a walk consisting of distinct nodes.

A *cycle* is a walk $i_1, i_2, \ldots, i_K$ such that $i_1 = i_K$. The *length* of the cycle is defined to be $K-1$. A cycle is *simple* if the only node appearing twice in the sequence is the starting (ending) node.

The matrix $\mathbf{T}$ is *strongly connected* if there is path relative to $\mathbf{T}$ from any node to any other node. Similarly, we say that $A' \subset A$ is strongly connected if $\mathbf{T}$ restricted to A' is. This is true if and only if the nodes in A' all lie on a cycle involves only them.

A group of nodes $A' \subset A$ is *closed* relative to $\mathbf{T}$ if $i \in A'$ and $T_{ij} > 0$ implies that $j \in A'$.

A closed group of nodes $A' \subset A$ is *minimal* relative to $\mathbf{T}$ if no nonempty strict subset is closed.

Observe that any minimal closed group is strongly connected.

3 Convergence of Beliefs Under Naïve Updating

3.1 Definitions and Examples

Consider an arbitrary matrix iteration process characterized by an updating rule of the form

$$\mathbf{p}(t) = \mathbf{T}\mathbf{p}(t-1) = \mathbf{T}^t \mathbf{p}(0),$$

where $\mathbf{T}$ is a row-stochastic matrix. We now provide a full characterization of the interaction matrices for which there is convergence in the sense that $\lim_{t\to\infty} \mathbf{T}^t \mathbf{p}$ exists for all vectors $\mathbf{p}$. We will specialize to the language of the belief-updating model discussed in Section 2.2, but the mathematical result applies equally well in other settings.

DEFINITION. A matrix $\mathbf{T}$ is *convergent* if $\lim_{t\to\infty} \mathbf{T}^t \mathbf{p}$ exists for all vectors $\mathbf{p}$.

A condition ensuring convergence in strongly connected matrices is aperiodicity.

DEFINITION. The matrix $\mathbf{T}$ is *aperiodic* if the greatest common divisor of the lengths of its simple cycles is 1.

It is well-known that if $\mathbf{T}$ is strongly connected (also referred to as being irreducible) and aperiodic, then it is convergent (e.g., see Meyer (2000)). In fact, studies of social networks involving convergence of a matrix generally assume that $\mathbf{T}$ is strongly connected and that $T_{ii} > 0$ for some or all i, which implies that the matrix is aperiodic.

To see what can go wrong when aperiodicity fails, let us examine a simple example.

Example **1.**

$$\mathbf{T} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Clearly,

$$\mathbf{T}^t = \begin{cases} \mathbf{T} & \text{if } t \text{ is odd} \\ \mathbf{I} & \text{if } t \text{ is even.} \end{cases}$$

In particular, if $p_1(0) \neq p_2(0)$, then the belief vector never reaches a steady state; the two agents keep switching beliefs.

However, it is not necessary to have $T_{ii} > 0$ for even a single i in order to ensure convergence.

Example **2.** Consider

$$\mathbf{T} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Here,

$$\mathbf{T}^t \to \begin{pmatrix} 2/5 & 2/5 & 1/5 \\ 2/5 & 2/5 & 1/5 \\ 2/5 & 2/5 & 1/5 \end{pmatrix}.$$

Even though $\mathbf{T}$ has cycles and has 0's in its diagonals, it is aperiodic and converges. If we change to

$$\mathbf{T} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

then $\mathbf{T}$ is periodic as all of its cycles are of even lengths and $\mathbf{T}$ is no longer convergent.

3.2 A Characterization of Convergence

As mentioned before, it is well-known that aperiodicity is sufficient for convergence in the case where where $\mathbf{T}$ is strongly connected. The following theorem shows that, in this case, aperiodicity is also a necessary condition for convergence. We give a simple constructive proof that for a given strongly connected matrix which is not aperiodic, there is an unstable initial vector $\mathbf{p}$. The idea of this result is standard in the Markov chain literature,[9] but since the graph induced by $\mathbf{T}$ is a less important object in most Markov chain models than it is here, we have not seen the fact stated in terms of simple cycles, as it is here. This result is then quite useful in developing our full characterization of convergence when we also consider $\mathbf{T}$'s that are not strongly connected.

THEOREM **1.** *If a stochastic matrix* $\mathbf{T}$ *is strongly connected, then it is convergent if and only if it is aperiodic.*

While strongly connected interaction matrices occur in some settings, most social interactions will not involve strong connection. Thus, it is important to have a more general characterization. We now use the above fact to give a complete characterization of convergence.

DEFINITION**.** The matrix $\mathbf{T}$ is *strongly aperiodic* if it is aperiodic when restricted to any closed group of nodes.

THEOREM **2.** *A stochastic matrix* $\mathbf{T}$ *is convergent if and only if it is strongly aperiodic.*

Theorem 2 is not a simple extension of Theorem 1. The main issue is the following. We can think of decomposing a society into minimal closed groups (i.e. maximal groups that are strongly connected) and then the set of remaining agents. We know from Theorem 1 that convergence holds when we restrict attention to the strongly connected agents under aperiodicity. The challenge of the proof is to show that this is in fact all that is needed to imply that the beliefs of all the other agents must also converge. Since all agents are path connected to some agent in a minimal closed group, the proof boils down to showing that an agent with some weight on a path that goes to an agent with convergent beliefs must also have beliefs that converge. This relies on limiting properties of the interaction matrix, which can be seen by rearranging it into a suitable block form.

We emphasize that Theorem 2 is useful beyond the application to beliefs, but also in understanding centrality measures and a variety of other Markov chain applications.

3.3 Influence and the Limiting Beliefs

Beyond knowing whether or not beliefs converge, we are also interested in characterizing what beliefs converge to when they do converge. The following is an easy extension of Theorem 10

[9]For example, see Kemeny and Snell (1960, p. 6–7 and p. 35–37).

in DeMarzo, Vayanos and Zwiebel (2003). They consider a case where $\mathbf{T}$ has positive entries on the diagonal, but their proof is easily extended to the case with 0 entries on the diagonal. We let $\mathcal{M}$ be the collection of minimal closed groups, and define $M = \bigcup_{B \in \mathcal{M}} B$. A subscript B indicates projection (resp. restriction) of vectors (resp. operators) to the subspace of $\mathbb{R}^n$ corresponding to the set of agents in B.

THEOREM **3.** *A stochastic matrix $\mathbf{T}$ is convergent if and only if there is a nonnegative row vector $\mathbf{s} \in \mathbb{R}^n$, and for each $j \notin M$ a vector $\mathbf{w}^j \geq 0$ with $|\mathcal{M}|$ entries that sum to 1 such that*

1. $\sum_{i \in B} s_i = 1$ *for any minimal closed group B,*
2. $s_i = 0$ *if i is not in a minimal closed group,*
3. $\mathbf{s}_B$ *is the left eigenvector of $\mathbf{T}_B$ corresponding to the eigenvalue 1,*
4. *for any vector $\mathbf{p}$ and any minimal closed group B, $(\lim_{t\to\infty} \mathbf{T}_B^t \mathbf{p})_B = \mathbf{s}_B \mathbf{p}_B$,*
5. *for any $j \notin M$, $(\lim_{t\to\infty} \mathbf{T}^t \mathbf{p})_j = \sum_{B \in \mathcal{M}} w_B^j \mathbf{s}_B \mathbf{p}_B$.*

This result says that, when beliefs converge, all agents in any closed group will eventually come arbitrarily close to holding the same belief. This belief will be a weighted average of the initial beliefs of the agents in that group. The weights are the entries of the vector $\mathbf{s}$, and the weight of any agent in a closed group is positive. We refer to s_i the *influence weight* or simply the *influence* of agent i. Agents who are not in M have no influence and their beliefs converge to weighted sums of the beliefs of the agents whom they observe.

The most important thing to note about the vector $\mathbf{s}$ is that $s_j = \sum_{i \in A} T_{ij} s_i$ for all j, so that the influence of an agent is the sum of the influences of those who trust him, weighted by their trust for him. This is a very natural property for a measure of influence to have: influential people are those who are trusted by other influential people.

4 The Wisdom of Crowds: Convergence to Accurate Beliefs in Large Societies

In this section, we examine sequences of convergent matrices $(\mathbf{T}_n)_{n=1}^{\infty}$ indexed by n, the number of agents in each. This may be viewed as a sequence of successive snapshots of a growing network, as we add one agent at a time; or simply as a standard tool for understanding what happens in "large" societies. We are interested in the conditions under which agents communicating through the network converge to hold the "correct" belief.

Throughout this section, we maintain the assumption that $\mathbf{T}_n$ is convergent for each n without repeating it in every result.

4.1 Information and Examples of Unwise Societies

Suppose that the true state of nature is μ, and let each agent i in network n see a signal ϕ_i^n that is distributed with mean μ, a finite variance of at least $\sigma^2 > 0$, and support that is a subset of a compact set $[-M, M]$. Suppose that signals are independently (but not necessarily identically) distributed.

Let $\mathbf{s}^n$ be the influence vector corresponding to $\mathbf{T}_n$, as defined in Theorem 3. We write the (i, j) entry of $\mathbf{T}_n$ as T_{ij}^n, and the belief of agent i at time t as $p_i^n(t)$. In each network, order the agents so that $s_i^n \geq s_{i+1}^n \geq 0$ for each i and n, where $\sum_i s_i^n = 1$.

In each network of the sequence, the limiting belief of each agent i in network n approaches some limit $p_i^n(\infty)$. We say the sequence of networks is *wise* when, for each i, this belief converges in probability to the true state μ as $n \to \infty$.

Definition. The sequence $(\mathbf{T}_n)_{n=1}^{\infty}$ is *wise* if

$$\operatorname*{plim}_{n\to\infty} p_i^n(\infty) = 0$$

for each i.

To get some feeling for which societies are wise, start by supposing, for a moment, that the communication structure is a strongly complete graph – i.e., every possible link is present and that all links have equal weight: i.e., $T_{ij}^n = 1/n$ for all $i, j \in A$. In this situation, it is easy to see that $s_i^n = 1/n$ for each $i \in A$. Indeed, after the first period every agent holds the average belief of the society. By a law of large numbers, as n grows the beliefs become arbitrarily accurate.

Obviously, this question becomes substantially more complicated when the communication structure has less symmetry. Different agents can have different influence weights, and so certain signals will affect the final state more than others. Can the fundamental idea of the above example can be carried through in more general networks? The answer is sometimes, but not always, yes.

To get some idea of the challenge faced in discerning when a wise crowds result holds, let us examine an example.

EXAMPLE **3.** Consider the following network, defined for arbitrary n. Fix $\delta, \varepsilon \in (0,1)$ and define, for each $n \geq 1$, an n-by-n interaction matrix

$$\mathbf{T}_n := \begin{bmatrix} 1-\delta & \frac{\delta}{n-1} & \frac{\delta}{n-1} & \cdots & \frac{\delta}{n-1} \\ 1-\varepsilon & \varepsilon & 0 & \cdots & 0 \\ 1-\varepsilon & 0 & \varepsilon & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1-\varepsilon & 0 & 0 & \cdots & \varepsilon \end{bmatrix}.$$

The network is shown in Figure 1 for $n = 6$ agents.

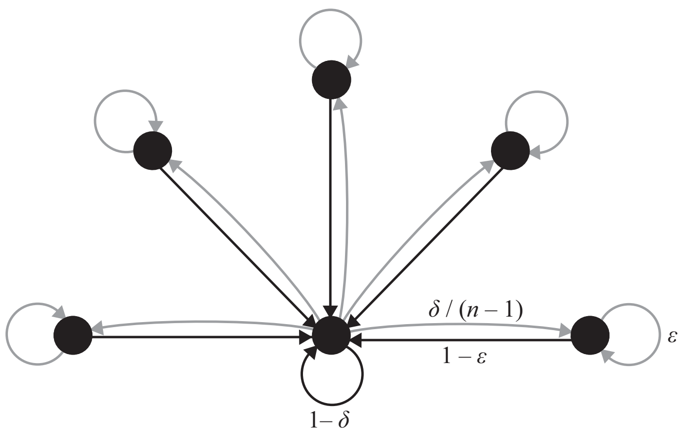

Figure 1: The unbalanced star network (shown here for $n = 6$ agents), which is one example demonstrating that the limiting belief is not always accurate.

We find that

$$s_i^n = \begin{cases} \frac{1-\varepsilon}{1-\varepsilon+\delta} & \text{if } i = 1 \\ \frac{\delta}{(n-1)(1-\varepsilon+\delta)} & \text{if } i > 1. \end{cases}$$

This network will not always converge to the truth. Observe that at stage n, the limiting belief is $s_1^n \phi_1^n$ plus some other independent random variables that have mean μ. As s_1^n is bounded away from 0, the variance of of the limiting belief remains bounded away from 0 for all n. So beliefs will not generally converge to truth. The intuition is simply that the leader's information – even when it is far from the mean – is weighted heavily enough that it biases the final belief, and the followers' signals cannot do much to correct it. Indeed, Proposition 1 below shows that as long as some agent's influence is bounded away from 0 for all n, convergence to true beliefs will not generally occur.

Note that even if we let $1-\varepsilon$ approach 0 at any rate we like, so that people are not weighting the center very much, the center has nonvanishing influence as long as δ is of the same order as $1-\varepsilon$. Thus, it is not simply the total weight on a given indivdiual that matters, but the relative weights that matter.

One thing that goes wrong in this example is that the central agent receives an high amount of trust relative to the amount given back to others, making him unduly influential. However, this is not the only obstruction to convergence to true beliefs. There are examples in which the trust coming into any node is bounded relative to the trust going out, and there is still an extremely influential agent who can keep society's beliefs away from the true state. This shows how indirect weight can matter.

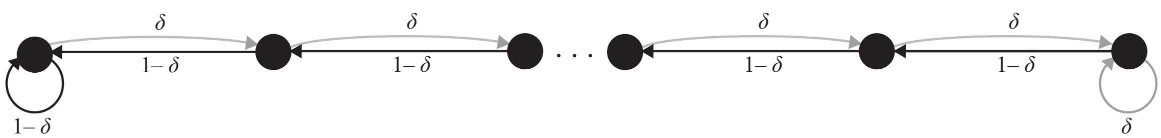

Figure 2: The unbalanced line, which demonstrates that a network may not converge to truth even if every agent's incoming trust is bounded. Agents are numbered from left to right.

Example 4. Fix $\delta \in (0, 1/2)$ and define, for each $n \geq 1$, an n-by-n interaction matrix by

$$\begin{cases} T^n_{11} = 1-\delta & \\ T^n_{i,i-1} = 1-\delta & \text{if } i \in \{2, \ldots, n\} \\ T^n_{i,i+1} = \delta & \text{if } i \in \{1, \ldots, n-1\} \\ T^n_{nn} = \delta & \\ T^n_{ij} = 0 & \text{otherwise.} \end{cases}$$

The network is shown in Figure 2. It is simple to verify that

$$s^n_i = \left(\frac{\delta}{1-\delta}\right)^{i-1} \cdot \frac{1-\left(\frac{\delta}{1-\delta}\right)}{1-\left(\frac{\delta}{1-\delta}\right)^{n+1}}.$$

In particular, $\lim_{n\to\infty} s^n_1$ can be made as close to 1 as desired by choosing a small δ. As in the previous example, it can then be shown that the system will not generally converge to true beliefs. The reason for the leader's undue influence here is somewhat more subtle than in Example 3: it is not the trust he directly receives, but indirect trust accruing to him due to his privileged position in the network. Thus, while he only receives twice as much direct trust as the typical agent, his influence can exceed the sum of all other influences by a huge factor for small δ. This shows that it can be extremely misleading to measure agents' influence based on direct incoming trust; instead, the entire structure of the network is relevant.

4.2 Wisdom in Terms of Influence: A Law of Large Numbers

We now seek to investigate the question outlined above more generally. We first develop a variation on a standard law of large numbers that is helpful in our setting, as we are working with weighted averages and potentially non-identically distributed random variables. The following result will be used to completely characterize wisdom in terms of influence weights.

Lemma **1.** *If* $(\mathbf{T}_n)_{n=1}^{\infty}$ *is a sequence of strongly connected convergent matrices, then*

$$\operatorname*{plim}_{n\to\infty} \sum_{i\in A_n} s_i^n \phi_i^n = \mu$$

if and only if $\max_i s_i^n \to 0$.[10]

This says that, in a strongly connected society, the limiting belief of all agents, $p(\infty) = \sum_{i\in A_n} s_i^n \phi_i^n$ will converge to truth if and only if the most important agent's influence tends to 0. With slightly more careful analysis, this lemma implies an important result.

Proposition **1.** *If* $(\mathbf{T}_n)_{n=1}^{\infty}$ *is an arbitrary sequence of convergent stochastic matrices, then*

$$\operatorname*{plim}_{n\to\infty} p_i^n(\infty) = \mu$$

for all i *if and only if* $\max_i s_i^n \to 0$.

Thus, Proposition 1 implies that $(\mathbf{T}_n)_{n=1}^{\infty}$ is wise if and only if the influence of the most important agent in the whole society tends to 0. This result is natural in view of the examples in Section 4.1, where saw that a society can be led astray if the leader has too much influence.

[10]We remark that since $\sum_{i\in A_n} s_i^n \phi_i^n$ is bounded, this is equivalent to having $\operatorname{plim}_{n\to\infty} \left(|\sum_{i\in A_n} s_i^n \phi_i^n - \mu|^r\right) = 0$ for all $r > 0$.

4.3 Wisdom in Terms of Social Structure: Sufficient Conditions

The characterization found above is useful, but still quite abstract. It is interesting to ask what is required, in more concrete terms, for wisdom. We now provide structural sufficient conditions for a society to be wise. First, we note that when studying wisdom, we can choose the most convenient power of the interaction matrices to work with – that is, we can study direct influences or indirect influences at any level.

PROPOSITION **2.** *If, for all n there exists k_n such that $\mathbf{R}_n = \mathbf{T}_n^{k_n}$, then $(\mathbf{T}_n)_{n=1}^{\infty}$ is wise if and only if $(\mathbf{R}_n)_{n=1}^{\infty}$ is wise.*

To show this, note that $\lim_{t\to\infty} \mathbf{T}_n^t = \lim_{t\to\infty} \mathbf{R}_n^t$, so that for every n, the influence vectors will be the same for both matrices by an easy application of Theorem 3.

Our first sufficient condition for wisdom is straightforward:

PROPOSITION **3.** *If $(\mathbf{T}_n)_{n=1}^{\infty}$ is a sequence of strongly connected,*[11] *aperiodic, and row-stochastic matrices such that each column sums to one, then it is wise.*[12]

Proposition 3 follows directly from fact that if $\mathbf{T}$ is both row and column stochastic, and strongly connected, then it has (left and right) unit eigenvectors of $\mathbf{s} = (\frac{1}{n}, \ldots, \frac{1}{n})$, and so then the influence of each agent in the society is equal. This makes clear how strong it is to have each agent receiving the same total weight in such a social network. An obvious sufficient condition for this is to have the matrix be symmetric, so that pairs of agents have the same trust for each other.

We now consider other, less restrictive assumptions that generate the same conclusion. Since wisdom is a notion defined in the large society limit, we are led to consider asymptotic properties of social groups. In what follows, $(B_n)_{n=1}^{\infty}$ denotes a arbitrary sequence of sets such that $B_n \subseteq A_n$ for each $n \in \mathbb{N}$. This type of sequence should be viewed as a subset of society, possibly growing and changing as society grows.

DEFINITION**.** The sequence $(B_n)_{n=1}^{\infty}$ of sets of nodes is *small* if $\lim_{n\to\infty} \frac{|B_n|}{n} = 0$.

DEFINITION**.** The sequence $(B_n)_{n=1}^{\infty}$ of sets of nodes is *finite* if there is a k such that $\sup_n |B_n| \leq k$.

[11] Instead of strong connectedness, it suffices for it to be possible to partition the agents into strongly connected subsets that are growing in size plus some subsets that receive no trust from any of the strongly connected ones.

[12] We thank Jonathan Weinstein for suggesting this proposition.

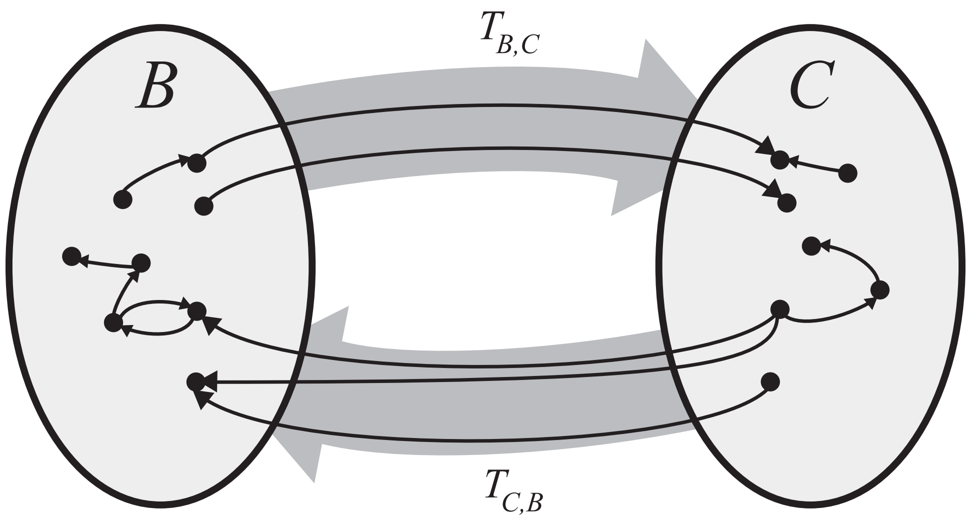

Figure 3: The large arrows illustrate the concept of the weight of one group on another.

The conditions will be stated in terms of weight that certain groups have for other groups, so we make the following definition. For any two sets $B, C \subseteq A$, let

$$T_{B,C} = \sum_{\substack{i \in B \\ j \in C}} T_{ij}.$$

This is the weight of B on C, also called the trust of B for C. The concept is illustrated in Figure 3.

The first family of sufficient conditions for wise conditions is as follows.

Property **1 (Balance).** If $|B_n| \leq n/2$

$$\sup_n \frac{T^n_{A_n - B_n, B_n}}{T^n_{B_n, A_n - B_n}} < \infty.$$

The balance condition says that any group of agents who involves less than half of the society cannot be getting infinitely more trust from the remaining agents than they give to the remaining agents. This rules out situations like Example 3 above. It also excludes situations in which some group receives a bounded amount of trust but has vanishing trust for the rest of the world.

Property **2 (Minimal Out-Dispersion).** There is a $q \in \mathbb{N}$ and $r > 0$ such that if B_n is finite, $|B_n| \geq q$, and $|C_n| \geq n/2$, then $T^n_{B_n, C_n} > r$ for all large enough n.

The minimal out-dispersion condition requires that any large enough group must give at least some minimal trust to groups that contain more than half of the agents. This rules out situations like Example 4 above, in which there are agents that ignore the vast majority of society.

Having stated these two conditions, we can give the first main result on wise crowds, which states that the conditions are sufficient for wisdom.

THEOREM **4.** *If $(\mathbf{T}_n)_{n=1}^{\infty}$ is a sequence of convergent stochastic matrices satisfying balance and minimal out-dispersion, then it is wise.*

Note, however, that neither condition is sufficient on its own. Example 4 satisfies the first property but not the second. The square of the matrix in Example 3 satisfies the second but not the first. In both examples the society fails to be wise. (This relies on an appeal to Proposition 2 in the latter case.)

Theorem 4 suggests that there are two important ingredients in wisdom: a lack of extreme imbalances in the trust structure and also a lack of local self-centered groups which pay very little attention to the outside world. To explore this idea further, we formulate different conditions in the same spirit which also generate wisdom. The essential difference is that the notion of dispersion now focuses on links coming into a certain type of group as opposed to ones going out.

PROPERTY **3 (Balance for Small Groups).** If $(B_n)_{n=1}^{\infty}$ is small, then

$$\sup_n \frac{T^n_{A_n - B_n, B_n}}{T^n_{B_n, A_n - B_n}} < \infty.$$

This property weakens the balance condition to only hold for small groups.

PROPERTY **4 (Minimal In-Dispersion).** There is a $q \in \mathbb{N}$ and an $r < 1$ such that if $|B_n| = q$ and $C_n \subseteq A_n - B_n$ is finite then $T^n_{C_n, B_n} \leq r T^n_{B_n, A_n - B_n}$ for all large enough n.

This condition requires that the trust coming into a finite group not be too concentrated. The finite group B_n cannot have a finite neighborhood which gives B_n as much trust, asymptotically, as B_n gives out. This essentially requires influential groups to have a broad base of support, and rules out situations like Example 4 above. Indeed, along with Property 3, it is enough to generate wisdom.

THEOREM **5.** *If $(\mathbf{T}_n)_{n=1}^{\infty}$ is a sequence of convergent stochastic matrices satisfying balance for small groups and minimal in-dispersion, then it is wise.*

The proofs of Theorem 4 and 5 are technical, but the intuition behind them is not difficult. Suppose, by way of contradiction, that the wisdom conclusion does not hold. Then there must be a group of agents that have positive influence as $n \to \infty$, and a remaining uninfluential group. Since the sum of influences must add up to 1, having some very influential agents requires having a great number of uninfluential agents. In particular, the influential group must be fairly small. As a result, it can only give out a limited amount of trust, and thus can only have a similarly limited amount of trust coming in, using one of the balance conditions. Recall that the influence of an agent is a trust-weighted sum of the influences of those who trust him. The contradiction comes from the fact that the uninfluential group does not have enough influence to support the high influence of the influential group, since it can give this group only a limited amount of trust. But neither can the influential group get all its support from inside itself, because the minimal out- and in-dispersion conditions require it to send some of its trust outside, or to get a nontrivial fraction of its support from outside, respectively.

It turns out that this informal argument is challenging to convert to a formal one, because the array of influence weights s_i^n as n and i range over all possible values has some surprising and difficult properties. Nevertheless, the basic ideas outlined above can be carried through successfully.

5 Comparative Statics: Changes in Trust

As we have seen, for a convergent system, the distribution of influence weights is an eigenvector of the interaction matrix. While this characterization is very handy mathematically, it is still somewhat abstract and so we now provide comparative statics which illustrate the relationship between local trust and global influence more concretely.

There are some very easy conclusions that we can reach based on the fact that $s_j = \sum_i T_{ij} s_i$. For instance, if agent j gets at least as much trust from each other agent as agent k does (so that $T_{ij} \geq T_{ik}$ for all i), then j has at least as much influence as k (so that $s_j \geq s_k$). Similarly, holding else is equal, it is better to obtain trust from an agent who has more influence. That is, If $T_{ij} = T_{ik}$ for all $i \neq \ell, m$ and $T_{\ell j} = T_{mk} > T_{\ell k} = T_{mj}$, then $s_\ell > s_m$ implies $s_j > s_k$.

To move beyond these observations, we need to derive how $\mathbf{s}$ changes as $\mathbf{T}$ changes. For instance, suppose that a particular agent redistributes his or her trust. That is, he or she trusts one acquaintance more and another acquaintance less – the latter being necessary because the weights any given agent assigns to his or her contacts must sum to 1. Intuitively, one might guess that the influence of the agent to whom more trust has just been allocated would increase, but this is not entirely obvious, especially as there are many indirect effects. Nevertheless, a corollary of the main result of this section is that, the intuitive prediction is correct.

First, we build on a result of Schweitzer (1968) to give two characterizations of how changes in trust affect influence weights, and then we deduce the proposition claimed above. We consider general perturbations of the interaction matrix – i.e., changing $\mathbf{T}$ to $\mathbf{T} + \delta\mathbf{C}$, where $\mathbf{C}$ is arbitrary subject to the condition that the resulting matrix still be stochastic. This requires that each row of $\mathbf{C}$ sum to 0.

THEOREM **6.** *For any strongly connected, convergent* $\mathbf{T}$*, suppose that* $\mathbf{C}$ *is a matrix whose rows each sum to* 0*. Let*

$$\tilde{\mathbf{T}}(\delta) = \mathbf{T} + \delta\mathbf{C},$$

and let $\tilde{\mathbf{s}}(\delta)$ *be the vector of influence weights corresponding to* $\tilde{\mathbf{T}}(\delta)$*, supposing that* $\tilde{\mathbf{T}}(\delta)$ *is nonnegative for small enough* δ*. Then*

$$\tilde{\mathbf{s}}'(0) = \mathbf{s}\mathbf{C}\left(\mathbf{I} - \mathbf{T} + \mathbf{e}\mathbf{s}\right)^{-1},$$

and the inverse on the right hand side exists. The derivative can also be written as

$$\tilde{\mathbf{s}}'(0) = \sum_{k=0}^{\infty} \mathbf{s}\mathbf{C}\mathbf{T}^k, \tag{2}$$

and the series converges.

This comparative static is still somewhat abstract, but easy to compute once the perturbation of the interaction matrix is known, especially using the infinite series above, which we have not seen elsewhere[13]. We obtain two interpretable corollaries by considering specific perturbations.

Now we can characterize changes in influence upon perturbation of $\mathbf{T}$ using indirect influences, and explicitly give the signs of certain influence changes. Define the *t-step weight of i on j* as the (i, j) entry of $\mathbf{T}^t$, which we write as $T_{ij}^{(t)}$.

COROLLARY **1.** *Consider any strongly connected, convergent* $\mathbf{T}$*. If* $C_{ij} = 1$*;* $C_{ik} = -1$*; and*

[13]The series expression is particularly attractive because for networks with second eigenvalues that are not too large, it converges very quickly, so it is only necessary to compute a few terms.

C *has* 0 *entries elsewhere, then* $\tilde{s}'_j(0) \geq 0$ *and* $\tilde{s}'_k(0) \leq 0$. *Moreover,*

$$\tilde{s}'_j(0) = s_i \left[1 + \sum_{t=1}^{\infty} \left(T^{(t)}_{jj} - T^{(t)}_{kj} \right) \right]$$

and

$$\tilde{s}'_k(0) = -s_i \left[1 + \sum_{t=1}^{\infty} \left(T^{(t)}_{kk} - T^{(t)}_{jk} \right) \right].$$

Corollary **2.** *Retain the assumptions of the previous corollary. If* $C_{ij} > 0$; $C_{ik} \leq 0$ *for all all* $k \neq j$; *and* 0 *elsewhere, then* $\tilde{s}'_j(0) \geq 0$.

The first corollary says that if agent i redistributes his or her trust, placing more weight on the opinion of agent j and correspondingly less on that of agent k, then agent j becomes weakly more influential and agent k becomes weakly less influential in the network. The second corollary says that agent j becomes more influential even if the new weight he or she receives is taken from several agents, not just one.

Moreover, Corollary 1 shows that, quite generally, it is better to be trusted more by a more influential agent. That is, the improvement in one's influence arising from a favorable redistribution of trust by some agent is directly proportional to the influence of the agent redistributing it. This holds regardless of the structure of **T**, and generalizes our observation at the start of this section.

6 Dynamics: How Long Does Disagreement Last?

While Theorem 3 pins down the limiting behavior of beliefs, it does not illuminate the question of how long convergence takes or how this depends on the specifics of the belief matrix. Since disagreement is often observed in practice, there may be networks in which convergence takes a a very long time.

As in DeMarzo, Vayanos and Zwiebel (2003), in this section, we develop variations on results from spectral theory that relate rates of convergence to the size of eigenvalues. Beyond that, we then relate the bounds on convergence rates to the structure of the interaction. Applying a theorem on second eigenvalues, we can conclude that slow convergence corresponds to the case where society is factious – divided into several mutually distrustful components. There is also a useful sufficient condition for agreement to happen quickly.

The contribution in this section is not to present new mathematical results, as the results here easily follow from existing results from various literatures; but rather to collect and adapt the results to the current setting.

The first proposition, which is a standard Markov chain convergence result, gives an upper bound on the difference between the current belief vector and the limiting one, thus describing a condition under which convergence is fast. A proof can be found in Seneta (1973, p. 7).

Proposition 4. *Fix any norm $\|\cdot\|$ on $\mathbb{R}^n$. Let $\mathbf{T}$ be strongly connected and aperiodic, and let $\lambda(\mathbf{T})$ be the second-largest eigenvalue, in magnitude, of $\mathbf{T}$. Then $|\lambda_2(\mathbf{T})| < 1$ and there exist positive real constants C and K (which depend only on the matrix $\mathbf{T}$) such that for each $i \in A$:*

$$|p_i(\infty) - p_i(t)| \leq Ct^K \left|\lambda_2(\mathbf{T})\right|^t \cdot \|\mathbf{p}(0)\| . \tag{3}$$

Note that the exponential decay of $|\lambda_2(\mathbf{T})|^t$ overpowers the polynomial growth of t^K, so the system converges, as we already know. This proposition says that when $|\lambda_2(\mathbf{T})|$ is small, the system is guaranteed to converge quickly to its steady state.

Reversing the inequality is not possible in general. However, a companion proposition, whose proof can be found in Karlin and Taylor (1975, p. 542–551), gives a partial converse.

Proposition 5. *Retain the assumptions of Proposition 4. There exists an initial vector $\mathbf{p}(0) \in \mathbb{R}^n$, a positive real constant C, and an agent $i \in A$, such that*

$$|p_i(\infty) - p_i(t)| \geq C \left|\lambda_2(\mathbf{T})\right|^t . \tag{4}$$

This proposition says that if $|\lambda_2(\mathbf{T})|$ is large – i.e., close to 1 – then for some initial belief vectors, the system will converge to its steady state quite slowly. It is clear that this statement can only hold for some initial belief vectors, as one can always start agents out with identical beliefs in which case convergence is instantaneous.

The two propositions taken together allow us to use $|\lambda_2(\mathbf{T})|$ as a proxy for the system's tendency to equilibrate. If $|\lambda_2(\mathbf{T})|$ is small, then all the agents quickly reach agreement. If $|\lambda_2(\mathbf{T})|$ is large, then convergence can take a long time. Note that if some agents are to one side of the limiting belief, then some others must clearly be on the other side – otherwise, a process of averaging could never arrive at the limiting belief. Thus, while the eventual beliefs will always coincide, this might take happen slowly enough that diversity of opinion is observed for a long time, even in strongly connected networks.

While useful, these results leave something to be desired. In particular, the second largest eigenvalue of the interaction matrix is a rather abstract invariant of the system. What does it mean, in more concrete terms, for the interaction matrix to have a small or large second eigenvalue?

In fact, we can give an explicit necessary and sufficient condition for convergence to take a long time. To this end, we provide a few definitions.

Define the *cohesion*[14] of $\mathbf{T}$, following Hartfiel and Meyer (1998), as

$$\sigma(\mathbf{T}) = \min_{\substack{\varnothing \neq B, C \subsetneq A \\ B \cap C = \varnothing}} \left(T_{B,A-B} + T_{C,A-C}\right). \tag{5}$$

The minimum is taken over all pairs of disjoint subsets of agents such that neither subset is empty. Let $\mathbf{T}$ be rearranged (by permuting the labeling of the agents) so that the above minimum is achieved at $B' = \{1, \ldots, k_1\}$ and $C' = \{k_1 + 1, \ldots, k_2\}$. Then we can partition $\mathbf{T}$ as

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \mathbf{T}_{13} \\ \mathbf{T}_{21} & \mathbf{T}_{22} & \mathbf{T}_{23} \\ \mathbf{T}_{31} & \mathbf{T}_{32} & \mathbf{T}_{33} \end{bmatrix}, \tag{6}$$

where $\mathbf{T}_{11}$ is a k_1-by-k_1 matrix and $\mathbf{T}_{22}$ is has dimensions $(k_2 - k_1)$-by-$(k_2 - k_1)$. We can think of $\sigma(\mathbf{T})$ as the sum of the entries in $\mathbf{T}_{12}, \mathbf{T}_{13}, \mathbf{T}_{21}$, and $\mathbf{T}_{23}$. The cohesion $\sigma(\mathbf{T})$ being small corresponds to each agent in B' and C' having very little weight on those outside his or her group.

The cohesion measure of a stochastic matrix is closely related to the second largest eigenvalue, which is a result due to Hartfiel and Meyer (1998). A slight variation of their result is the following theorem which leads to implications for speed of convergence from our perspective.

THEOREM **7.** *[Hartfiel and Meyer (1998)] Consider a stochastic, strongly connected matrix* $\mathbf{T}$*. Having a low cohesion implies having a large second eigenvalue in the sense that for any* $\varepsilon > 0$*, there exists a* $\delta > 0$ *so that* $\sigma(\mathbf{T}) < \delta$ *implies* $|\lambda_2(\mathbf{T}) - 1| < \varepsilon$*. Conversely, having a large second eigenvalue implies a low cohesion in the following sense: for any* $\varepsilon > 0$*, there exists a* $\delta > 0$ *such that* $|\lambda_2(\mathbf{T}) - 1| < \delta$ *implies* $\sigma(\mathbf{T}) < \varepsilon$*.*

As claimed earlier, Theorem 7, together with Propositions 4 and 5, imply that convergence of beliefs is slow if there are at least two distrustful factions – groups who have little trust for those outside them, and in particular for each other. Conversely, if convergence is sufficiently slow, then it must be possible to partition society in this way.

[14] Hartfiel and Meyer (1998) call this quantity the *uncoupling measure*. In this terminology, systems with low uncoupling measures are very uncoupled. We have chosen the alternative term *cohesion* as it seems more descriptive.

Corollary **3.** *For any $\varepsilon > 0$, there exists a $\delta > 0$ such that such that if $\mathbf{T}$ is strongly connected and aperiodic with cohesion less than δ then we can find a an initial vector $\mathbf{p}(0) \in \mathbb{R}^n$, an agent $i \in A$, and a $C \in \mathbb{R}$ such that*

$$|p_i(\infty) - p_i(t)| \geq C(1-\varepsilon)^t. \tag{7}$$

While Corollary 3 provides conditions for convergence to be slow, we might also be interested in sufficient conditions for fast convergence. Here, general conditions seem to be harder to find, especially since fast convergence is required for all initial beliefs. Nevertheless, there are some situations where we can deduce conditions that ensure fast convergence.

The following proposition builds on results about *expander graphs* (e.g., see Hoory, Linial, and Wigderson (2006)). Consider a $\mathbf{T}$ which is d-regular and symmetric: that is, such that there exists $d \geq 1$ such that each i has $T_{ij} = \frac{1}{d}$ for d agents $j \neq i$, and where $T_{ij} = T_{ji}$. The expansion ratio of a symmeric and K-regular $\mathbf{T}$ is defined by

$$h(\mathbf{T}) = \min_{B:|B| \leq n/2} \frac{T_{B,B^C}}{|B|} \tag{8}$$

This keeps track of how many agents outside of B are trusted by the agents inside of B relative to B's size, and then finds the smallest ratio of this rate of "expansion". It is clear that h lies between 0 and 1. Theorems on expander graphs relate this ratio to second eigenvalues. Building from that theory, we can prove the following:

Proposition **6.** *If $\mathbf{T}$ is strongly connected, d-regular, and symmetric, then there exist positive real constants C and K (which depend only on the matrix $\mathbf{T}$) such that for each $i \in A$:*

$$|p_i(\infty) - p_i(t)| \leq Ct^K \left(1 - \frac{(h(\mathbf{T}))^2}{2}\right)^t \cdot \|\mathbf{p}(0)\|. \tag{9}$$

This is derived directly from Proposition 4 and the fact that $\lambda_2(\mathbf{T}) \leq 1 - \frac{(h(\mathbf{T}))^2}{2}$ (see Theorem 2.4 in Hoory, Linial, and Wigderson (2006)).[15] The proposition implies that if groups of agents are looking sufficiently outwards to sets of other agents, then convergence of beliefs is fast. For instance, if each set of agents B placed high enough weight on B^C so that $h(\mathbf{T})$ is at least $1-\varepsilon$, then $|p_i(\infty) - p_i(t)| \leq Ct^K \left(\frac{1+\varepsilon}{2}\right)^t \cdot \|\mathbf{p}(0)\|$.

We also mention a convenient sufficient condition for convergence to be fast, which is due

[15]The definition of h is adjusted here for the stochastic nature of the matrix, and the bound on the eigenvalue is adjusted accordingly.

to Haveliwala and Kavmar (2006).

Consider a case where $\mathbf{T}$ can be decomposed as follows:

$$\mathbf{T} = u\mathbf{U} + (1-u)\mathbf{H}, \tag{10}$$

where $u \in [0,1]$; $\mathbf{U}$ is a stochastic matrix such that all the entries in any given column are equal; and $\mathbf{H}$ is any stochastic matrix. The matrix $\mathbf{U}$ is the *uniform component* of $\mathbf{T}$ and the matrix $\mathbf{H}$ the *heterogeneous component* of $\mathbf{T}$. Define the *uniformity* of $\mathbf{T}$ as the largest u such that the decomposition in (10) is possible.[16] We say that $\mathbf{U}$ is uniform because, for each j, all entries in column j of $\mathbf{U}$ are equal, say to U_{*j}. This corresponds to all agents trusting agent j at least uU_{*j}. Since $\mathbf{U}$ is stochastic, it follows that everyone shares a baseline distribution of trust across some agents; the uniform component of each agent's trust distribution adds up to u.

Haveliwala and Kavmar (2006) show that if $\mathbf{T}$ is a stochastic matrix that is decomposed as in (10), Then $|\lambda_2(\mathbf{T})| \leq 1-u$. Thus, if the uniformity of $\mathbf{T}$ is high enough, then convergence is guaranteed to be fast. Intuitively, uniformity is high when a significant portion of everyone's information comes from a common set of agents. For example, it is high if everyone has a significant degree of trust for some set of media organizations. In terms of cohesion, we can see that having uniformity above some level then leads to a cohesion above some level. So intuitively, one expects faster convergence. However, the results on cohesion only hold as cohesion approaches 0, and so the uniformity results of Haveliwala and Kavmar (2006) show that, at least in some special cases, the results extend beyond the limiting extremes.

7 Conclusion

There are several testable empirical implications to be drawn from the results presented here.

First, the results on necessary and sufficient conditions for convergence to common beliefs suggest that the topological details of network structure can have a large qualitative impact: they determine whether the agents ever come to agree. In particular, if the network is regular in the sense of all cycle lengths having a common factor, then beliefs may cycle indefinitely. This is in contrast with previous results on learning in networks, in which the precise small-scale topological structure of the network does not typically play such a key role. On the other hand, the result is generally an optimistic one for long-term convergence: networks for which convergence fails are quite special, and many networks arising from stochastic

[16]The fact that the set of possible values of u is compact proves that this maximum exists.

processes would satisfy the sufficient conditions for convergence.

The main topic of this paper, explored in Section 4, concerns whether large societies whose agents get noisy estimates of the truth converge to true beliefs. We show that they do under certain assumptions about social structure. The flavor of the main condition is that no group of agents (unless it is large) should get very much more trust than it gives back. As long as this holds, and one of several additional conditions regarding dispersion is satisfied, it follows that sufficiently large societies will come arbitrarily close to the truth. These results suggest two insights. First, excessive attention to small groups of pundits or opinion-makers who are not reciprocally attentive to group opinion is bad for convergence to truth. On the other hand, social cohesion – in the sense of not having segments of society that essentially ignore each other's views – is good.

In our context, these conclusions provide an answer to a broad question asked by Joel Sobel (2000): can large societies whose agents are fairly naïve individually be smart in the aggregate? In this model, they can, if there is enough dispersion in who they listen to, and if they avoid concentrating too much on any small group of agents.[17] This conclusion contrasts with the very special conditions required for naïve learning presented by DeMarzo, Vayanos and Zwiebel (2003). In this sense, there seems to be more hope for boundedly rational social learning than has previously been believed. On the other hand, our sufficient conditions are fairly strong in the sense that they can fail if there is just one group which receives too much trust or is too insular. This raises a natural question: which processes of network formation satisfy the sufficient conditions we have set forth? How must agents dynamically allocate trust to ensure that no group obtains an excessive share of influence in the long run? These are potential directions for future work.

Our results on comparative statics show that when agents redistribute their trust, perturbations in the global social influences can readily be computed if we understand agents' indirect weights on each other's opinions. Importantly, sometimes only a few levels of indirect weights are required to get a very good approximation to the true perturbation. Thus, these results provide a means of testing the theory with only local information about details of the network structure.

The results that we surveyed regarding convergence rates provide some insight into polarization and propaganda. We should expect long-term polarization on an issue when the social structure describing how people discuss that issue splits into several mutually distrustful groups. This would mean that an agent's discussion partners are mostly restricted

[17]This is similar to the discussion in Bala and Goyal (1998) of what can go wrong when there is a commonly observed "royal family" under a different model of observational learning.

to the group in which he or she is located, and that there is little trust across party lines. Such properties have been studied in the political science literature: see, e.g., Huckfeldt and Sprague (1987). In contrast to the slow convergence of an incohesive society, elimination by authoritarian regimes of all but a few official media outlets leads to greater cohesion. Consider *Pravda* in the former Soviet Union, or the blocking of many foreign news sources in China, etc. One obvious reason for this behavior is to control access to information. This also increases the influence of those news sources, as well as helping in terms of a rate of convergence.

To finish, we mention some obvious extensions of the project. First, the theory can be applied to a variety of strategic situations in which social networks play a role. For instance, consider an election in which two political candidates are trying to convince voters. While the voters remain nonstrategic about their communications, the politicians (who may be viewed as being outside the network) can be quite strategic about how they attempt to shape beliefs. A salient question is whom the candidates would choose to target. The social network would clearly be an important ingredient. A related application would consider firms competitively selling similar products (such as Coke and Pepsi).[18] Here, there would be some benefits to one firm of the other firms' advertising. These complementarities, along with the complexity added by the social network, would make for an interesting study of marketing. Second, it would be interesting to involve heterogeneous agents in the network. In this paper, we have focused on nonstrategic agents who are all boundedly rational in essentially the same way. We might consider how the theory changes if there are some fully rational agents in the network. Can a small mixture of different agents significantly change the group's behavior? Such extensions would be a step toward connecting fully rational and boundedly rational models, and would open the door to a more robust understanding of social learning.

Appendix: Proofs

Proof of Theorem 1:

LEMMA **2.** *If* $\mathbf{T}$ *is strongly connected and aperiodic, then it is primitive (i.e.,* $\mathbf{T}^k > 0$ *for some finite* k*).*

Lemma 2 is a standard corollary to the Perron-Frobenius Theorem (see, e.g., Horn and Johnson (1985, Theorem 8.5.3)).

[18]See Galeotti and Goyal (2007) for a one-firm model of optimal advertising on a network.

LEMMA **3.** *[The Stochastic Matrix Theorem] If* $\mathbf{T}$ *is stochastic and primitive, then it converges.*

A proof can be found in Meyer (2000, Section 8.3).

LEMMA **4.** *If* $\mathbf{T}$ *is strongly connected, stochastic and not aperiodic, then it does not converge.*

Proof of Lemma 4: First we introduce some notation. Denote the length of a walk or cycle B by $|B|$. Suppose $B = i_1, i_2, \dots, i_b$ and $C = j_1, j_2, \dots, j_c$ are walks such that $i_b = j_1$. Then define $B + C$, called the *concatenation* of B and C, as $i_1, i_2 \dots, i_b, j_2, \dots, j_c$. It is easy to check that when two walks are concatenated, the length of the resulting walk is the sum of the lengths of its constituents. That is, $|B + C| = |B| + |C|$, and the same is also true for cycles.

Choose any node i. As $\mathbf{T}$ is strongly connected, i is on at least one cycle. Since $\mathbf{T}$ is not aperiodic, there is an integer $d > 1$ such that any cycle D containing i has length divisible by d.

Let Y be the set of all nodes j such that some path from i to j has length divisible by d.
Claim: If $j \in Y$ then all walks from i to j have length divisible by d.
Proof of Claim: Suppose $j \in Y$. Let B be a walk from i to j whose length is divisible by d. Let B' be another walk from i to j. We will show that the length of B' is divisible by d. By strong connectedness, there is a walk E from j to i. Since $B + E$ is a cycle through i, it follows that d divides $|B + E| = |B| + |E|$. The fact that d divides $|B|$ implies that d divides $|E|$. But $B' + E$ is another cycle through i, so d divides $|B' + E| = |B'| + |E|$. Since we saw d divides the second summand and the left hand side, it must divide $|B'|$, as desired. This shows the claim.

To prove the lemma, define $\mathbf{p}(0)$ by

$$p_j(0) = \begin{cases} 0 & \text{if } j \in Y \\ 1 & \text{if } j \notin Y \end{cases}.$$

Write

$$\mathbf{p}(t) = \mathbf{T}^t \mathbf{p}(0).$$

We claim $p_i(t) = 0$ whenever $d \mid t$ and $p_i(t) = 1$ whenever $d \nmid t$.

For the first part, suppose $d \mid t$ and let $\mathbf{R} = \mathbf{T}^t$. Then

$$\mathbf{p}(t) = \mathbf{R}\mathbf{p}(0)$$

and so

$$p_i(t) = \sum_{j=1}^{n} R_{ij} p_j(0).$$

Using a fact from Meyer (2000, p. 672) we can rewrite this as

$$p_i(t) = \sum_{j=1}^{N} p_j(0) \sum_{i_2,\ldots,i_t} T_{ii_2} T_{i_2 i_3} \cdots T_{i_t j}.$$

The coefficient of $p_j(0)$ is nonzero if and only if there is a walk from i to j of length t. This happens if and only if $j \in Y$, by definition of Y. But $p_j(0) = 0$ whenever $j \in Y$. This shows $p_i(t) = 0$.

To show $p_i(t) = 1$ whenever $d \nmid t$, let $\mathbf{R} = \mathbf{T}^t$ again and, as before write

$$p_i(t) = \sum_{j=1}^{n} p_j(0) \sum_{i_2,\ldots,i_t} T_{ii_2} T_{i_2 i_3} \cdots T_{i_t j}.$$

The coefficient of $p_j(0)$ is nonzero if and only if there is a walk from i to j of length t. By Claim 1, this happens only if $j \notin Y$. For all such j, we have $p_j(0) = 1$. Therefore

$$p_i(t) = 1 \sum_{j=1}^{n} \sum_{i_2,\ldots,i_t} T_{ii_2} T_{i_2 i_3} \cdots T_{i_t j}.$$

This is the sum of the ith row of $\mathbf{R} = \mathbf{T}^t$, and this matrix is still stochastic, so $p_i(t) = 1$. ∎

The theorem follows directly from these lemmas. ∎

Proof of Theorem 2 and 3:

The backward implication of Theorem 2 follows from Lemma 4 after observing that if $\mathbf{T}$ is not strongly aperiodic, then some minimal closed group of $\mathbf{T}$ must fail to be aperiodic. The backward implication of Theorem 3 is immediate.

We will prove the forward implication of Theorem 2 and also Theorem 3 at the same time, via the following standard lemma.

Lemma 5. *If* $\mathbf{T}$ *is strongly connected and aperiodic, then there is a row vector* $\mathbf{s} > 0$ *such that for any* $\mathbf{p}$,

$$\lim_{t\to\infty} \mathbf{T}^t \mathbf{p} = \mathbf{s}\mathbf{p}.$$

This vector is the left eigenvector of $\mathbf{T}$ *corresponding to the eigenvalue* 1. *In particular, all entries of the limit are the same.*

Proof of Lemma 5: Standard facts about Markov matrices, which can be found in Meyer (2000, Section 8.3) and Berman and Plemmons (1979, Chapter 2) imply the following facts, which we collect here for later reference. The spectral radius of the matrix $\mathbf{T}$, denoted $\rho(\mathbf{T})$, is 1 as $\mathbf{T}$ is stochastic. In fact, 1 is the unique eigenvalue of $\mathbf{T}$ with magnitude 1, and all other eigenvalues are strictly smaller in magnitude. By Meyer (2000, Section 8.3), the following limit expression holds:

$$\lim_{t\to\infty} \frac{\mathbf{T}^t}{\rho(\mathbf{T})} = \frac{\mathbf{es}}{\mathbf{se}}, \tag{11}$$

where $\mathbf{e}$ is the column vector of ones and $\mathbf{s}$ is the unique, positive left-hand Perron (row) eigenvector of $\mathbf{T}$ corresponding to eigenvalue 1. We may scale $\mathbf{s}$ so that its entries sum to 1.

Since the right side of the above limit equation is $\mathbf{e}$ times a 1-by-1 matrix, it follows that all the entries in the limiting vector are the same, namely $\mathbf{sp}$. This proves all the claims in the lemma. ∎

This proves the theorem for a strongly connected interaction matrix. Now suppose that the matrix is not strongly connected, so that some proper subset of agents is not closed. Then by relabeling agents, it can be transformed into

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{0} & \mathbf{T}_{22} \end{bmatrix}, \tag{12}$$

where the bottom right block corresponds to all agents in M, i.e. all agents in any minimal closed group, and the rows above it correspond to agents who are in no minimal closed group. We may further decompose

$$\mathbf{T}_{22} = \begin{bmatrix} \mathbf{T}_{B_1} & & \\ & \ddots & \\ & & \mathbf{T}_{B_m} \end{bmatrix},$$

where each B_k is minimally closed. Each will also be aperiodic, because $\mathbf{T}$ is strongly aperiodic. Lemma 5 shows that for each k,

$$\lim_{t\to\infty} \mathbf{T}^t_{B_k}\mathbf{p}_{B_k} = \mathbf{s}_{B_k}\mathbf{p}_{B_k}, \tag{13}$$

where $\mathbf{s}_{B_k}$ is the unique left eigenvector of $\mathbf{T}_{B_k}$, scaled so that its entries sum to 1. Define $\mathbf{s} = \mathbf{0} \oplus \mathbf{s}_{B_1} \oplus \cdots \oplus \mathbf{s}_{B_m}$, where $\mathbf{0}$ is a zero row vector such that $\mathbf{s} \in \mathbb{R}^n$. This shows (1–3) of Theorem 3.

For the remaining parts of Theorem 3, we note by Meyer (2000, Section 8.4) that the decomposition in (12) entails

$$\lim_{t\to\infty} \mathbf{T}^t \mathbf{p} = \begin{bmatrix} \mathbf{0} & \mathbf{Z} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \mathbf{p}, \tag{14}$$

where

$$\mathbf{E} = \begin{bmatrix} \mathbf{es}_{B_1} & & \\ & \ddots & \\ & & \mathbf{es}_{B_m} \end{bmatrix}.$$

The block-diagonal form of $\mathbf{E}$, along with (13), immediately implies (4) of Theorem 3.

Since powers of stochastic matrices are stochastic, $\mathbf{Z}$ has rows summing to 1. For each $j \notin M$, define $\mathbf{w}^j \in \mathbb{R}^{|\mathcal{M}|}$ by $w_k^j = \sum_{i\in B_k} Z_{ji}$. Then $\sum_{k=1}^m w_k^j = 1$. Note that

$$\lim_{t\to\infty} \mathbf{T}^t \mathbf{p} = \lim_{r\to\infty} \mathbf{T}^r \left(\lim_{t\to\infty} \mathbf{T}^t \mathbf{p} \right),$$

and so the matrix on the right hand side of (14) is idempotent. Then (14) can be written as

$$\lim_{t\to\infty} \mathbf{T}^t \mathbf{p} = \begin{bmatrix} \mathbf{0} & \mathbf{Z} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \mathbf{q}, \tag{15}$$

where

$$\mathbf{q} = \begin{bmatrix} \mathbf{0} & \mathbf{Z} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \mathbf{p}.$$

Since $\mathbf{E}_{B_k}\mathbf{p}_{B_k} = \mathbf{s}_{B_k}\mathbf{p}_{B_k}$, it follows that $q_i = \mathbf{s}_{B_k}\mathbf{p}_{B_k}$ if $i \in B_k$. Now, for each $j \notin M$, we have

$$\left(\lim_{t\to\infty} \mathbf{T}^t \mathbf{p} \right)_j = \sum_{i\in M} Z_{ji} q_i = \sum_{k=1}^m w_k^j \mathbf{s}_{B_k} \mathbf{p}_{B_k}$$

by definition of $\mathbf{q}$ and $\mathbf{w}^j$. This completes the proof of (5) in Theorem 3. The forward direction of Theorem 2 follows immediately. ∎

Proof of Lemma 1:

We know that the variance of each ϕ_i^n lies between σ^2 and $\overline{\sigma}^2$ for some $\overline{\sigma}^2$.

Let $X^n = \sum_i s_i^n \phi_i^n$. Then $\mathrm{Var}(X^n) \le \overline{\sigma}^2 \sum_i s_i^2$.

First, suppose $s_1^n \to 0$. Since $s_n^i \ge s_{i+1}^n \ge 0$ for all i and n, it follows that

$$\mathrm{Var}(X^n) \le \overline{\sigma}^2 \sum_i s_i^2 \le \overline{\sigma}^2 s_1^n \sum_i s_i = \overline{\sigma}^2 s_1^n \to 0.$$

By Chebychev's inequality, fixing any $\varepsilon > 0$

$$\mathbb{P}\left[\left|\sum_i s_i^n \phi_i^n - \mu\right| > \varepsilon\right] \leq \frac{\mathrm{Var}(X^n)}{\varepsilon^2} \to 0.$$

For the converse, suppose (taking a subsequence if neccesary) $s_1^n \to s > 0$. Since each X^n has support in $[-M, M]$, a variance bounded below, and mean μ, it then follows that there exists $\delta > 0$ such that $\mathrm{Var}(X^n) > \delta$ for all n. But this implies directly that there exists ε and $x > 0$ such that, for each n,

$$\mathbb{P}\left[\left|\sum_i s_i^n \phi_i^n - \mu\right| > \varepsilon\right] > x.$$

This completes the proof. ∎

Proof of Proposition 1: First we prove that if the condition $s_1^n \to 0$ holds, then convergence to truth occurs. By Theorem 3, agents with no influence converge to weighted averages of beliefs of agents with influence, so it suffices to show that if $i_n \in A_n$ is any sequence of agents in minimal closed groups, then $\mathrm{plim}_{n\to\infty} p_{i_n}^n(\infty) = \mu$. Let B_n be the minimal closed group of i_n. Without loss of generality, we may replace $\mathbf{T}_n$ with induced interaction matrix on the agents in B_n. Now, by the lemma, all that is required for every agent in B_n to converge to true beliefs is that $|B_n| \to \infty$ and the most influential agent in B_n have influence converging to 0. The second fact follows because the most influential agent in A_n has influence converging to 0, and *a fortiori* the same must hold for the leader in B_n. The first fact follows directly from this, for the influences of agents in B_n now converge to 0 but sum to 1, which is impossible if the number of agents is bounded by a finite number.

Conversely, if the influence of some agent remains bounded above 0, then we may restrict attention to his closed group and conclude from the argument of the lemma that convergence to truth is not generally guaranteed. ∎

Proof of Theorem 4:

Recall that we have ordered the agents so that $s_i^n \geq s_{i+1}^n$ for all i. In the proof of this theorem and the next, all unadorned limits are taken as $n \to \infty$. Suppose to the contrary that there is a subsequence where $s_1^n \to s > 0$.

Take the subsequence to be the sequence.

Let k_n be a sequence such that $\lim k_n s_{k_n}^n \to 0$ and $k_n \leq n/2$. To see that such a sequence exists, consider a countable sequence of $x \to 0$. Let us first argue that for each x there is at most a finite set of set of n, such that $is_i^n \geq x$ for all $i \leq n/2$. Suppose to the contrary that there exists $x > 0$ such that for an infinite set of n, $is_i^n \geq x$ for all $i \leq n/2$. Thus, for these n, $\sum_{i \leq n/2} is_i^n \geq \sum_{i \leq n/2} \frac{x}{i} \to \infty$, which is a contradiction. Thus, for each x there is n_x such that for every $n > n_x$, the set $C_{x,n} = \{i : is_i^n < x\}$ is nonempty. The n_x form a nondecreasing sequence as $x \to 0$. Select the sequence k_n by choosing from $C_{x',n}$ where x' is the largest x such that $n_x \leq n$.

For each n, let $H_n = \{1, \ldots, k_n\}$ and $L_n = A_n - H_n$. Observe that since $\mathbf{s}^n$ is a left hand eigenvector of $\mathbf{T}_n$, we have

$$\sum_{j \in H_n} s_j^n = \sum_{i \in H_n} \sum_{j \in H_n} T_{ij}^n s_i^n + \sum_{i \in L_n} \sum_{j \in H_n} T_{ij}^n s_i^n$$

Rewrite this as

$$\sum_{i \in H_n} s_j^n \left(1 - \sum_{i \in H_n} T_{ji}^n\right) = \sum_{i \in L_n} \sum_{j \in H_n} T_{ij}^n s_i^n$$

or

$$\sum_{j \in H_n} \left(\sum_{i \in L_n} T_{ji}^n s_j^n\right) = \sum_{i \in L_n} \left(\sum_{j \in H_n} T_{ij}^n s_i^n\right). \tag{16}$$

Let

$$s_H^n = \sum_{j \in H_n} s_j^n \cdot \frac{\sum_{i \in L_n} T_{ji}^n}{T_{H_n, L_n}^n}$$

and

$$s_L^n = \sum_{i \in L_n} s_i^n \cdot \frac{\sum_{j \in H_n} T_{ij}^n}{T_{L_n, H_n}^n}.$$

We rewrite (16) as

$$s_H^n T_{H_n, L_n}^n = s_L^n T_{L_n, H_n}^n. \tag{17}$$

Taking $B_n = \{1, \ldots, q\}$ and $C_n = A_n - B_n$ in the statement of the minimal out-dispersion property, we have for large enough n,

$$s_H^n \geq \frac{s_q^n r}{q k_n}$$

for a natural number q and a positive real r.

Also, $s_L^n \leq s_{k_n}^n$. Thus, (17) implies that

$$r s_q^n T_{H_n, L_n}^n \leq q s_{k_n}^n k_n T_{L_n, H_n}^n. \tag{18}$$

Since $T^n_{L_n,H_n}/T^n_{H_n,L_n}$ is bounded (by balance), this implies that $\lim_n s^n_q = 0$.

So, consider a case where $\lim s^n_q = 0$. Let k be the largest i such that $\lim_n s^n_i = 0$. Let $H_n = \{1, 2, \ldots, k\}$. Then, as above, we have the following facts:

$$\begin{aligned}
\sum_{i\in H_n} s^n_i \sum_{j\in L_n} T^n_{ij} &= \sum_{i\in L_n}\sum_{j\in H_n} T^n_{ij}s^n_i \\
s^n_k \sum_{i\in H_n}\sum_{j\in L_n} T^n_{ij} &\le s^n_{k+1}\sum_{i\in L_n}\sum_{j\in H_n} T^n_{ij} \qquad \text{by the ordering of the } s^n_i \\
\frac{s^n_k}{s^n_{k+1}} &\le \frac{T^n_{A_n-H_n,H_n}}{T^n_{H_n,A_n-H_n}}.
\end{aligned}$$

The left side will have supremum ∞ over all n because $s^n_{k+1} \to 0$ while s^n_k has positive lim sup. The right side, however, is bounded using the balance property. This is a contradiction, and therefore this case is complete. ∎

Proof of Theorem 5: By Proposition 1 and the ordering we have chosen for $\mathbf{s}^n$, it suffices to show that

$$\lim_{n\to\infty} s^n_1 = 0 \tag{19}$$

Suppose otherwise.

We proceed by cases. First, assume that there are only finitely many i such that $\lim_{n\to\infty} s^n_i > 0$. Then we can proceed as at the end of the proof of Theorem 4 to reach a contradiction. Note that only balance for finite groups is needed, which is implied by balance for small groups.

From now on, we may assume that there are infinitely many i such that $\lim_{n\to\infty} s^n_i > 0$. In particular, if we take the q guaranteed by Property 4 and set $B_n = \{1, 2, \ldots, q\}$, then we know that $\lim_{n\to\infty} s^n_i > 0$ for each $i \in B_n$. Now, fix a function $g : \mathbb{N} \to \mathbb{N}$ and define $C_n = \{q + 1, \ldots, q + g(n)\}$. Finally, put $D_n = \{q + g(n) + 1, q + g(n) + 2, \ldots, n\}$.

We claim g can be chosen such that

$$\limsup \frac{T_{C_n,B_n}}{T_{B_n,A_n-B_n}} \le r$$

and $\lim g(n) = \infty$, i.e. g is a divergent function. Let $C^k_n = \{q + 1, q + 2, \ldots, q + k - 1\}$. By Property 4, there exists an n_1 such that for all $n \ge n_1$, we have

$$\frac{T_{C^1_n,B_n}}{T_{B_n,A_n-B_n}} \le r.$$

Having chosen $n_1, \ldots, n_{k-1}$, there exists an $n_k > n_{k-1}$ such that for all $n \geq n_k$ we have

$$\frac{T_{C_n^k, B_n}}{T_{B_n, A_n - B_n}} \leq r.$$

Define

$$g(n) = \max\{k : n_k \leq n\}.$$

Since $n_1, n_2, \ldots$ is an increasing sequence of integers, the set whose maximum is being taken is finite. It is also nonempty for $n \geq n_1$, so g is well defined there. For $n < n_1$, let $g(n) = 1$. Next, observe $\lim g(n) = \infty$. For if not, there is some k' so that $n_k \leq n_{k'}$ for all k, which is false since $n_1, n_2, \ldots$ is an increasing sequence of integers. Finally, since C_n defined above is equal to $C_n^{g(n)}$ and

$$\frac{T_{C_n^{g(n)}, B_n}}{T_{B_n, A_n - B_n}} \leq r$$

for all $n \geq n_1$ by construction, it follows that

$$\limsup \frac{T_{C_n, B_n}}{T_{B_n, A_n - B_n}} \leq r. \tag{20}$$

This shows our claim about the choice of g.

Now we have the following string of implications:

$$\begin{aligned}
\sum_{i \in B_n} s_i^n &= \sum_{i \in B_n} \sum_{j \in A_n} T_{ji}^n s_j^n \\
\sum_{i \in B_n} s_i^n &= \sum_{i \in B_n} \sum_{j \in B_n} T_{ji}^n s_j^n + \sum_{i \in B_n} \sum_{j \in C_n} T_{ji}^n s_j^n + \sum_{i \in B_n} \sum_{j \in D_n} T_{ji}^n s_j^n \\
\sum_{i \in B_n} s_i^n &= \sum_{i \in B_n} \sum_{j \in B_n} T_{ij}^n s_i^n + \sum_{i \in C_n} \sum_{j \in B_n} T_{ij}^n s_i^n + \sum_{i \in D_n} \sum_{j \in B_n} T_{ij}^n s_i^n \\
\sum_{i \in B_n} s_i^n \sum_{j \notin B_n} T_{ij}^n &= \sum_{i \in C_n} s_i^n \sum_{j \in B_n} T_{ij}^n + \sum_{i \in D_n} s_i^n \sum_{j \in B_n} T_{ij}^n.
\end{aligned}$$

Rearranging,

$$\sum_{i \in B_n} s_i^n \sum_{j \notin B_n} T_{ij}^n - \sum_{i \in C_n} s_i^n \sum_{j \in B_n} T_{ij}^n = \sum_{i \in D_n} s_i^n \sum_{j \in B_n} T_{ij}^n. \tag{21}$$

Using the ordering of the s_i^n, the first double summation on the left side satisfies

$$\sum_{i \in B_n} s_i^n \sum_{j \notin B_n} T_{ij}^n \geq s_q^n \sum_{i \in B_n} \sum_{j \notin B_n} T_{ij}^n = s_q^n T_{B_n, A_n - B_n}^n.$$

Similarly, the second summation on the left side of (21) satisfies

$$\sum_{i\in C_n} s_i^n \sum_{j\in B_n} T_{ij}^n \leq s_{q+1}^n \sum_{i\in C_n} \sum_{j\in B_n} T_{ij}^n = s_{q+1}^n T_{C_n,B_n}^n.$$

Finally, the summation on the right side of (21) satisfies

$$\sum_{i\in D_n} s_i^n \sum_{j\in B_n} T_{ij}^n \leq s_{q+g(n)+1}^n \sum_{i\in D_n} \sum_{j\in B_n} T_{ij}^n = s_{q+g(n)+1}^n T_{A_n-B_n\cup C_n,B_n}^n.$$

For notational ease, put $f(n) = q + g(n) + 1$. Combining the above facts with (21), we find

$$s_q^n T_{B_n,A_n-B_n}^n - s_{q+1}^n T_{C_n,B_n}^n \leq s_{f(n)}^n T_{A_n-B_n\cup C_n,B_n}^n.$$

By the ordering of the s_i^n, it follows that

$$s_{q+1}^n T_{B_n,A_n-B_n}^n - s_{q+1}^n T_{C_n,B_n}^n \leq s_{f(n)}^n T_{A_n-B_n\cup C_n,B_n}^n. \tag{22}$$

By Property 4, there is a $r < 1$ so that for all large enough n, we have

$$T_{C_n,B_n}^n < rT_{B_n,A_n-B_n}^n.$$

Using this and a trivial bound on the right hand side of (22), we may rewrite (22) as

$$s_{q+1}^n (1-r) T_{B_n,A_n-B_n}^n \leq s_{f(n)}^n T_{A_n-B_n,B_n}^n. \tag{23}$$

To finish the proof, we need two observations. The first is that $s_{f(n)}^n \to 0$. Suppose not, so that it exceeds some $a > 0$ for infinitely many n. Then for all such n, we use the ordering of the s_i^n to find

$$\sum_{i=1}^{f(n)} s_i^n \geq af(n),$$

and this quantity diverges, contradicting the fact that

$$\sum_{i=1}^{n} s_i^n = 1.$$

The second observation is that we may, without loss of generality, assume g is a divergent function satisfying $\lim_{n\to\infty} \frac{g(n)}{n} = 0$, so that $(C_n)_{n=1}^\infty$ is small. For if we have a g so that this condition does not hold, it is easy to verify that reducing g to some smaller divergent function for which the condition does hold cannot destroy the property in (20).

Now we rewrite (23) as

$$(1-r)\frac{s^n_{q+1}}{s^n_{f(n)}} \leq \frac{T^n_{A_n-B_n,B_n}}{T^n_{B_n,A_n-B_n}}.$$

By an argument very similar to the previous case, the observations we have just derived along with Property 3 generate the needed contradiction. ∎

Proof of Theorem 6 The equation

$$\tilde{\mathbf{s}}'(0) = \mathbf{sC}\left(\mathbf{I} - \mathbf{T} + \mathbf{es}\right)^{-1}$$

is a well-known result of Schweitzer (1968, equation 15). Multiplying both sides of it on the right by $(\mathbf{I} - \mathbf{T} + \mathbf{es})$, one obtains

$$\tilde{\mathbf{s}}'(0)\left(\mathbf{I} - \mathbf{T} + \mathbf{es}\right) = \mathbf{sC}.$$

However, the entries of $\tilde{\mathbf{s}}'(0)$ sum to zero, while all the rows of $\mathbf{es}$ are the same, so the corresponding product vanishes and

$$\tilde{\mathbf{s}}'(0) = \tilde{\mathbf{s}}'(0)\mathbf{T} + \mathbf{sC}.$$

Replacing $\tilde{\mathbf{s}}'(0)$ on the right hand side with the entire right hand side repeatedly, we find, for all $r \geq 1$

$$\tilde{\mathbf{s}}'(0) = \tilde{\mathbf{s}}'(0)\mathbf{T}^{r+1} + \sum_{t=0}^{r} \mathbf{sCT}^t.$$

Taking the limit as $r \to \infty$, the first term on the right vanishes by the argument we just gave, because $\mathbf{T}^{r+1} \to \mathbf{es}$. The summation on the right converges because the left hand side is a well-defined vector by the first part of the proof. This establishes (2). ∎

Proof of Corollary 1 The summation formulas follow immediately from (2). To establish that the signs are as claimed, define, for every $\delta \in (\frac{1}{2}, 1)$,

$$\mathbf{v}^{(\delta)} = \sum_{t=0}^{\infty} \mathbf{sC}(\delta\mathbf{T})^t. \tag{24}$$

We will show $v_j^{(\delta)} > 0$ and $v_k^{(\delta)} < 0$ for each such δ. Additionally, we will check that we can interchange the limit as $\delta \to 1$ from below with the above summation. Using the summation expression for the derivative from (2), this will prove the corollary.

The first claim is that

$$\sum_{t=0}^{\infty} \lim_{\delta \to 1^-} \mathbf{sC}(\delta \mathbf{T})^t = \lim_{\delta \to 1^-} \sum_{t=0}^{\infty} \mathbf{sC}(\delta \mathbf{T})^t. \tag{25}$$

To see (25), we will use the Lebesgue dominated convergence theorem. Observe that

$$\left\| \mathbf{sC}(\delta \mathbf{T})^t \right\| \leq \left\| \mathbf{sCT}^t \right\|.$$

The terms on the right hand side are those of an absolutely convergent series, as follows. By Seneta (1973, Theorem 1.2) and the triangle inequality, we have

$$\sum_{t=0}^{\infty} \|\mathbf{sCT}^t\| \leq \sum_{t=0}^{\infty} \|\mathbf{sCes}\| + \sum_{t=0}^{\infty} \alpha q^t t^K \|\mathbf{sC}\|, \tag{26}$$

where $q = |\lambda_2(\mathbf{T})| < 1$ and α, K are constants. Now $\mathbf{sCes} = \mathbf{0}$ since all rows of $\mathbf{es}$ are equal, while each row of $\mathbf{sC}$ sums to 0. Choose $p \in (q, 1)$ and observe

$$\sum_{t=0}^{\infty} \alpha q^t t^K \|\mathbf{sC}\| \leq \alpha' + \sum_{t=0}^{\infty} \alpha p^t \|\mathbf{sC}\|,$$

where α' is another constant, since eventually $p^t > t^K q^t$. But the series on the right hand side is geometric with ratio less than 1, so it converges. Combining these facts, (26) converges as claimed and (25) is established.

Now, we will show that $\tilde{s}_j'(0) \geq 0$ and $\tilde{s}_k'(0) \leq 0$. By the above claim about interchanging sums and limits, it suffices to see $v_j^{(\delta)} > 0$ and $v_k^{(\delta)} < 0$ for each $\delta \in (\frac{1}{2}, 1)$. To this end, observe that the series defining $\mathbf{v}^{(\delta)}$ is absolutely convergent, so we may rearrange the order of summation and write

$$\mathbf{v}^{(\delta)} = \mathbf{sC} \sum_{t=0}^{\infty} (\delta \mathbf{T})^t.$$

Now $\delta \mathbf{T}$ has spectral radius $\delta < 1$, so the Neumann series guarantees

$$\mathbf{v}^{(\delta)} = \mathbf{sC} \left(\mathbf{I} - \delta \mathbf{T} \right)^{-1}. \tag{27}$$

Write $\mathbf{X} = \mathbf{I} - \delta \mathbf{T}$ and $\mathbf{Y} = \mathbf{X}^{-1}$. It is easy to see that for each i, we have $X_{ii} > \sum_{j \neq i} |X_{ij}|$. Then Fiedler and Pták (1967, Theorem 3.5) guarantees that for each j and for each $i \neq j$, $|Y_{jj}| > |Y_{ij}|$. Moreover, by Berman and Plemmons (1979, Lemma 2.1), $\mathbf{Y}$ has only positive entries, so in fact $Y_{jj} > Y_{ij} > 0$. In (27), note that $\mathbf{sC}$ has s_i in column j, has $-s_i$ in column k, and has 0 elsewhere. This combined with the facts about $\mathbf{Y}$ now establishes the claim about the signs of $v_j^{(\delta)}$ and $v_k^{(\delta)}$, by inspection of (27). ∎

Proof of Corollary 2 The proof proceeds exactly as in the previous corollary, except in this case, $\mathbf{sC}$ has $C_{ij}s_i$ in column j while the rest of its columns are nonpositive numbers summing to $-C_{ij}s_i$. Then once again inspection of (27) along with the facts derived above about $\mathbf{Y}$ complete the proof. ∎

References

V. Bala and S. Goyal. Learning from neighbours. *The Review of Economic Studies*, 65(3): 595–621, 1998.

V. Bala and S. Goyal. Conformism and diversity under social learning. *Economic Theory*, 17:101–120, 2001.

C. Ballester, A. Calvó-Armengol, and Y. Zenou. Who's who in networks. Wanted: the key player. *Econometrica*, 74:1403–1417, 2006.

A. V. Banerjee. A simple model of herd behavior. *Quarterly Journal of Economics*, 107(3): 797–817, 1992.

A. V. Banerjee and D. Fudenberg. Word-of-mouth learning. *Games and Economic Behavior*, 46:1–22, 2004.

A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, 1979.

P. B. Bonachich. Power and centrality: a family of measures. *American Journal of Sociology*, 92:1170–1182, 1987.

P. Bonacich and P. Lloyd. Eigenvector-like measures of centrality for asymmetric relations. *Social Networks*, 23(3):191–201, 2001.

B. Çelen and S. Kariv. Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review*, 94(3):484–497, 2004.

S. Choi, D. Gale, , and S. Kariv. Behavioral aspects of learning in social networks: An experimental study. In J. Morgan, editor, *Advancies in Applied Microeconomics*. BE Press, Berkeley, 2005.

S. Choi, D. Gale, , and S. Kariv. Social learning in networks: A quantal response equilibrium analysis of experimental data. 2007.

S. Collignon and M. Al-Sadoon. A theory of stochastic consensus. Preprint, July 2006.

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

G. Ellison and D. Fudenberg. Rules of thumb for social learning. *Journal of Political Economy*, 101(4):612–643, 1993.

G. Ellison and D. Fudenberg. Word-of-mouth communication and social learning. *Journal of Political Economy*, 111(1):93–125, 1995.

M. Fiedler and V. Pták. Diagonally dominant matrices. *Czechoslovak Mathematical Journal*, 17(92):420–433, 1967.

J. French. A formal theory of social power. *Psychological Review*, 63(3):181–194, 1956.

N. Friedkin and E. Johnsen. Social positions in influence networks. *Social Networks*, 19: 209–222, 1997.

D. Gale and S. Kariv. Bayesian learning in social networks. *Games and Economic Behavior*, 45(2):329–346, 2003.

A. Galeotti and S. Goyal. A theory of strategic diffusion. Preprint, available at `http://privatewww.essex.ac.uk/~agaleo/`, February 2007.

F. Harary. Status and contrastatus. *Sociometry*, 22:23–43, 1959.

D. J. Hartfiel and C. D. Meyer. On the structure of stochastic matrices with a subdominant eigenvalue near 1. *Linear Algebra and Its Applications*, 272(1):193–203, 1998.

T. H. Haveliwala and S. D. Kamvar. The second eigenvalue of the Google matrix. Preprint, available at `http://nlp.stanford.edu/pubs/secondeigenvalue.pdf`, 2006.

S. Hoory, N. Linial, and A. Widgerson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

R. Huckfeldt and J. Sprague. Networks in context: The social flow of political information. *The American Political Science Review*, 81(4):1197–1216, 1987.

M. O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, 2007. Forthcoming.

S. Karlin and H. Taylor. *A First Course in Stochastic Processes*. Academic Press, San Diego, 1975.

L. Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18:39–43, 1953.

J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. van Nostrand, Princeton, 1960.

C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, 2000.

I. Palacios-Heurta and O. Volij. The measurement of intellectual influence. *Econometrica*, 72(3):963–977, 2004.

B. Rogers. A strategic theory of network status. Preprint, under revision, 2006.

D. Rosenberg, E. Solan, and N. Vieille. Informational externalities and convergence of behavior. Preprint, October 2006.

P. J. Schweitzer. Perturbation theory and finite Markov chains. *Jounral of Applied Probability*, 5(2):401–413, 1968.

E. Seneta. *Non-Negative Matrices*. John Wiley & Sons, New York, 1973.

J. Sobel. Economists' models of learning. *Journal of Economic Theory*, 94(1):241–261, 2000.

S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, 1994.