A Model of Belief Influence in Large Social Networks*

Antonio Jiménez-Martínez†

January 2015

Abstract

This paper develops a model of belief influence through communication in an exogenous social network. The network is weighted and directed, and it enables individuals to listen to others' opinions about some exogenous parameter of interest. Agents use Bayesian updating rules. The weight of each link is exogenously given and it specifies the quality of the corresponding information flow. We explore the effects of the network on the agents' first-order beliefs about the parameter and investigate the aggregation of private information in large societies. We begin by characterizing an agent's limiting beliefs in terms of some entropy-based measures of the conditional distributions available to him from the network. Our results on consensus and correctness of limiting beliefs are in consonance with some of the literature on opinion influence under non-Bayesian updating rules. First, we show that the achievement of a consensus in the society is closely related to the presence of prominent agents who are able to crucially change the evolution of other agents' opinions over time. Secondly, we show that the correct aggregation of private information is facilitated when the influence of the prominent agents is not very high.

Keywords: Communication networks, opinion influence, Bayesian updating rules, private signals, private messages, consensus, correct limiting beliefs.

JEL Classification: D82, D83, D85.

---

*A companion paper appears in *Emile Borel and the Notion of Strategy: Ninetieth Anniversary* under the title "Evolution of Beliefs in Networks." This project has benefited greatly from very useful conversations with Rabah Amir, Luciana Moscoso-Boedo, Alejandro Manelli, Larry Samuelson, and Myrna Wooders. For helpful comments and suggestions, I am also grateful to two anonymous referees, an editor, and seminar audiences at CIDE and ITAM. Any remaining errors are my own.

†División de Economía, Centro de Investigación y Docencia Económicas (*CIDE*).

1 Introduction

In social environments where individual decisions depend on uncertain parameters, coordination is often facilitated when agents reach similar beliefs about the actual values of such parameters.[1] Examples of these decisions include consumption, occupational, investment, and voting choices. In practice, the evolution of our beliefs about uncertain variables usually depend on (a) our own private learning about the variables (e.g., based on private observations or research), and (b) how we gather information from neighbors, friends, co-workers, local leaders, and political actors. The links of a social network are channels that transmit opinions about products, job vacancies, investment opportunities, and political programs.[2]

In recent years, the size of most social networks has become increasingly large. Through the new phone technology, email, social media, and the explosive expansion of the internet, the world has experienced profound improvements in our abilities to access the opinions of greater numbers of people, even when they are located at great distances. Nowadays, we are able to cheaply keep in touch with others and find a broad variety of opinions and advise about almost any type of matter from any place in the globe.

By focusing on relatively large networked societies, the aim of this paper is to explore the relation between the network that connects a group of agents and the evolution of their first-order beliefs about some common parameter of interest. We develop a stylized model of network-based dynamic belief formation, with Bayesian updating rules, where there are two types of information transmission: (a) each agent receives private information about the parameter from an external (idiosyncratic) source and (b) there is communication between connected agents about the information they are obtaining from their sources.

More specifically, consider a group of agents who care about some uncertain exogenous parameter.[3] Each of them begins with some initial priors and observes over time a sequence of private signals about the parameter. The informativeness of this stream of signals describes the quality of the agent's private learning about the parameter, which intuitively could be associated with, e.g., the technology of his source or the attention level that he is able to put on the source. In addition, the agents are connected through an exogenous (directed and weighted) social network that specifies a pattern of relations where each agent can listen to the opinions of others. Each directed link is characterized by an exogenously given weight that describes

[1] To fix ideas, consider, e.g., the sort of environments which are commonly modeled as beauty-contest games.

[2] For instance, some recent papers highlight the importance of social networks in the generation of social capital (see, e.g., Jackson, Rodríguez-Barraquer, and Tan, 2012; Campbell, 2014).

[3] The parameter could correspond to some economic, social, or political variable. Examples include the profitability of an investment, the effects of some public policy, the ideology of a certain politician, or whether a certain social movement will spread out.

the quality of the information transmission from the speaker to the listener. Intuitively, this weight could be determined, e.g., by the speaker's communication skills, the listener's level of attention or understanding capabilities, or the information transmission technology of the link. In particular, at each date, each agent receives a *non-strategic* message from each agent to whom he has a directed link. Such a message is correlated with the sender's signal so that it conveys some information about the signal that the speaker is privately observing.[4] Also, through a link, each agent can receive messages from other indirectly connected agents. We assume that the signals and messages that each agent receives are, conditionally on the parameter realization, independent both across time and across senders. Using this framework, we explore the conditions on the network structure and, in particular, on its weights, under which the agents reach a consensus in their first-order beliefs about the parameter value. We also investigate the conditions under which the agents aggregate correctly the decentralized information that they obtain from their sources.[5] This model's assumptions about what is privately and commonly known between the agents, how they process the information they receive, and the type of inferences that they make are motivated and justified by our earlier mentioned emphasis on large networked societies.

The amount of information that is transmitted from the speaker to the listener is given by the exogenous weight of their directed link and cannot be manipulated neither by the sender nor by the receiver.[6] We assume that the weight of each link is constant over time, which leads to stationary updating processes.

To complete the groundwork for our analysis, we need to address two key modeling assumptions. First, following recent developments on the ranking of information, we choose an entropy-based measure, namely, the *power measure* to measure of the informativeness of sources and links. The power measure is the average of the relative entropy of the posterior in the first period with respect to the prior so that it captures, from an ex-ante viewpoint, the gain of information in moving from the prior to the posterior.[7]

Secondly, we need to adopt a notion of correct beliefs for our framework. In our model, the

[4] At a more intuitive level, the network describes exogenously given conduits through which the agents listen to others speak about their private learning. As a motivating example, consider a group of investors deciding their investment in a collective fund. Each investor begins with some priors about the potential profitability of the fund and collects over time some further information by studying privately a number of characteristics of the fund. In addition, through communication, each investor has, in each period, some (noisy) access to the private analyses of the fund features made by other investors.

[5] Since Condorcet (1785)'s seminal essay, the problem of whether a group of agents who have dispersed information will be able to aggregate their pieces of information and reach a correct consensus has been the focus of a large body of mathematical and philosophical work.

[6] Thus, in this paper we are not interested in the rich strategic interactions present in a sender-receiver game.

[7] In contrast with other measures extensively used in decision theory, such as the Blackwell (1953)'s ordering, the power measure induces *complete orders* over sets of information structures.

beliefs of an outside Bayesian observer who begins with some priors and can use over time the sources available to *all* agents converge almost surely to some limiting beliefs. This observer's beliefs ignore the flows of information through the network. On the other hand, each agent using only his own source and the information he obtains from others also converges almost surely to some limiting beliefs. Suppose that all the agents' beliefs converge to some consensus limiting beliefs. Then, we ask which network features facilitate that the agents' limiting beliefs coincide with the observer's. A central observation that justifies our approach is that, by a law of large numbers argument, the aggregation of the decentralized information sources provides us with an estimate (from an ex-ante viewpoint) of the true parameter value which becomes arbitrarily accurate as the number of agents in the society grows (in the limit, tending to infinity).

Our results begin by identifying the presence of decay in the flow of information along the links of the network, in Lemma 1.[8] We then turn to provide a simple but complete characterization of an agent's limiting beliefs, in Theorem 1. Armed with this result, we proceed by identifying necessary and sufficient conditions, in Theorem 2, under which an agent j is able to influence another agent i in a way such that both of them end up with the same limiting beliefs that agent j would reach without communication (i.e., using only his source).[9] Our characterization result is provided in terms of the power measure of the connections between the agents in the society and of other entropy-related measures. The advantages of using entropy-related measures are that they summarize very precisely a number of features of priors, sources, and links (which, in fact, constitute the primitives of the model), and that they allow for insights for a fairly general class of distributions.

The main message that arises from Theorem 2 is that the weight of the connections (either direct or indirect) from i to j must be sufficiently high for agent i to be influenced by agent j, and that the weight of the connections from agent j to any agent k in the society must be sufficiently low. The latter condition prevents j from being influenced by any other agent.

Our characterizations of limiting beliefs and of opinion influence (Theorems 1 and 2), combined with our results on the decay of the flow of information (Lemma 1), allow one to identify those prominent agents in the society who might influence crucially others, and to assess whether a consensus could finally be achieved.[10]

[8]Including some exogenous decay in the flow of information across connections has been common in the economic literature on networks since the seminal papers by Jackson and Wolinsky (1996), and by Bala and Goyal (2000). An interesting feature of our model is that the presence of decay can be described very precisely in terms of the power measures of sources and links.

[9]This can be naturally interpreted as agent j being able, as time evolves, to convince agent i to share his views about the uncertain parameter.

[10]For applications, this approach seems very useful in those cases where observables could be used to estimate distributions over signals and messages. In these cases, the power measures proposed in this paper can be used as a proxy to describe the strength of connexions in networks. Some recent empirical papers (Banerjee, Chandrasekhar,

Proposition 1 provides a sufficient condition on the weights of the connections under which, provided that there is a consensus, the agents' limiting beliefs aggregate correctly the information available to all of them through their sources. We show that a society with consensus attains correct limiting beliefs if the influence of the prominent agents is not so large so as to distort the evolution of beliefs that results by aggregating the sources. The intuition is that some limits on the influence of the prominent agents helps preventing cases of "large-scale manipulation" in which such agents could use their "favorable" positions in the network to bias everyone towards opinions that do not necessarily correspond to the aggregation of all the agents' sources. Our last result, Proposition 2, considers a particular network architecture whereby one agent enjoys the most favorable position to influence everyone else and, at the same time, not being influenced by the others, namely, the *center-directed star network*. We show that, provided that the central agent is able to influence the others, correct limiting beliefs are precluded for societies large enough.

Our model is closely related to, and motivated by, the opinion influence literature that builds on the DeGroot (1974)'s benchmark. This literature assumes that agents are non-Bayesian and use some "rule of thumb" to incorporate others' opinions into their belief updating.[11] Within this literature, perhaps the paper closest to ours in terms of the questions asked is Golub and Jackson (2010). They show that limiting beliefs arbitrarily close to the truth are obtained when the influence of the most influential agent vanishes as the size of the society tends to infinity. Our results bear a clear resemblance with theirs. Although agents use Bayesian rules in our model, we also obtain their message that, to attain consensus and correct limiting beliefs, a certain level of popularity is quite convenient whereas a disproportionate popularity could be harmful. Given the similarity in the main insights, our model can be viewed as a methodological contribution, with a Bayesian foundation, that reinforces the general picture that emerges from these rather tractable non-Bayesian models of opinion influence.

The rest of the paper is organized as follows. Section 2 comments further on the related literature, Section 3 presents the model, Section 4 analyses the attainment of consensus and of correct limiting beliefs, and Section 5 concludes with a discussion of the results and of possible extensions. The proofs of all the results are grouped together in the Appendix.

Duflo, and Jackson, 2013 and 2014) have obtained estimations of the strength of connexions in particular social communication networks.

[11] In the DeGroot's model, agents update their beliefs by averaging their neighbors' beliefs according to some exogenous weights that describe the intensity of the links between the agents. While a major advantage of these models lies in their tractability, common features with the present paper are that the weights of the links are exogenous and constant over time, and that the induced belief-revision processes are stationary. A classical contribution in this literature is DeMarzo, Vayanos, and Zwiebel (2003) who propose a network-based explanation for the emergence of "unidimensional" opinions.

2 Related Literature

This paper is related to the theoretical literature that, considering higher-order beliefs, asks whether a group of agents commonly learn the true parameter.[12] Our approach is different in that, motivated by our interest on large networked societies, we focus only on the agents' first-order posteriors about the parameter when they start with possibly different priors.[13] Another important difference between the current paper and the models within the learning literature with higher-order beliefs (e.g., Parikh and Krasucki, 1990; Heifetz, 1996; Koessler, 2001; Steiner and Stewart, 2011; Cripps, Mailath, Ely, and Samuelson, 2008 and 2013) is in the fact that this literature evaluates the correctness of beliefs by conditioning the posteriors on a given value of the parameter, which is taken as the actual value.[14]

The present paper relates also to several branches of the literature on influence in networks with non-Bayesian rules other than the one that stems from the DeGroot's benchmark. Acemoglu, Ozdaglar, and ParandehGheibi (2010) consider that the agents meet pairwise and adopt the average of their pre-meeting beliefs. They study how the presence of agents who influence others, but do not change their own beliefs, interferes with the spread of information along the network. Although they do not focus on consensus in particular, our model allows for insights with a similar flavor since some spread of beliefs among agents with different opinions is required for consensus in our paper. Also, the question of whether consensus is attained under non-Bayesian updating rules is analyzed by Acemoglu, Como, Fagnani, and Ozdaglar (2013). They distinguish between "regular" agents, who update their beliefs according to the information they receive from their neighbors, and "stubborn" agents, who never update their beliefs. They show that consensus is never obtained when the society contains stubborn agents with different opinions. Again, this insight bears some resemblance with ours when the connections

[12] In a setting without communication among the agents, Cripps, Ely, Mailath, and Samuelson (2008) show that (approximate) common learning of the parameter is attained when signals are sufficiently informative and the sets of signals are finite. They assume that the agents start with common priors and ask whether each agent not only assigns sufficiently high probability to some given parameter value but also to the event that each other agent assigns high probability to such a value, and so on, ad infinitum.

[13] Thus, we do not consider ex-ante probabilistic assessments that the agents could make over the histories underlying their beliefs as we do not explore their higher-order beliefs. Importantly, the result of common learning attainment by Cripps, Mailath, Ely, and Samuelson (2008) mentioned in footnote 12 requires that the sets of signals and messages be finite. This is not surprising since they assume that each agent is able to keep track of the higher-order beliefs of *all* agents about the signals each of them is receiving at each period. Clearly, this approach is less appealing when one considers a society where the number of its members is large. In fact, the argument given by Rubinstein (1989) in his celebrated email game suggests that common learning of the true parameter is precluded with arbitrarily large societies.

[14] In other words, this strand of the literature uses an ex-post perspective to regard parameter values as being correct while we use an ex-ante viewpoint. Our model can be regarded as an attempt to introduce Bayesian updating rules into the DeGroot's framework of opinion influence and evolution of first-order beliefs. Accordingly, as in the approach pursued by DeMarzo, Vayanos, and Zwiebel (2003), and by Golub and Jackson (2010), our notion of correctness asks whether the network structure allows for the aggregation of the decentralized sources of private information.

of some agent do not allow him to be sufficiently influenced by others.

Using a Gaussian information structure, belief dynamics and the attainment of correct consensus beliefs in social networks are also recently explored, in way complementary to the one carried out in this paper, by Azomahou and Opolot (2014). Their model considers both Bayesian and non-Bayesian updating rules. For the Bayesian case, they obtain the interesting result that, under certain conditions, if the agents start with common priors, then consensus and correct aggregation of the decentralized information always follow, regardless of the network structure. The present paper is different from theirs mainly in our assumptions on the the agents' information processing capabilities and on the way in which they transmit information to others.

At a more instrumental level, the current paper is related to the literature on *strategic communication* initiated by Crawford and Sobel (1982) since the transmission of information in our model through signals and messages is modeled, for each period, exact the same way in which a sender transmits information to a receiver in a sender-receiver or *cheap talk* game.[15]

Finally, this paper is also related to the statistical branch of information theory and to a growing diverse economic literature that uses entropy-based measures. The concept of power of a signal that we use was originally proposed by Shannon (1948) in his celebrated paper. Entropy-based measures have been subsequently used by applied mathematicians to model a number of aspects of communication, ranging from data compression and coding to channel capacity or distortion theory. Nevertheless, such measures have remained seldom used by economists for decades. Recently, a few papers have incorporated entropy-based measures to account for informativeness levels in several economic phenomena. For example, Sciubba (2005) uses the power of a signal to rank information in her work on survival of traders in financial markets. Cabrales, Gossner, and Serrano (2013) propose, for a class of no-arbitrage investment problems, an entropy-based measure, which formally coincides with the power measure, and that they term as *entropy informativeness*.

3 The Model

We use $\Delta(X)$ throughout the paper to denote the set of all Borel probability distributions on a given set X. Also, for a probability distribution P, we use $E_P[\,\cdot\,]$ to denote the expectation operator with respect to P.

There is a finite set of agents $N = \{1, 2, \ldots, n\}$ who care about an exogenous parameter $\theta \in \Theta = \{\theta_1, \theta_2, \ldots, \theta_L\}$.[16] Time is discrete and indexed by $t \in \{0, 1, 2, \ldots\}$. The true value of

[15] As mentioned earlier, the crucial difference is that the amount of information transmitted in our model is not endogenously chosen but it is exogenously given by the description of sources and links.

[16] Although the parameter space is assumed to be finite, the extension of our main results to a compact, but

θ is selected by nature in period $t = 0$. Each agent i begins with a prior distribution $p_i \in \Delta(\Theta)$ that describes his (subjective) beliefs about the parameter in period $t = 0$.

Following a classical argument, forcefully constructed by Savage (1954, p. 48), most Bayesian learning models assume that all agents share the same priors. Some recent theoretical research, however, concludes that Savage (1954)'s Bayesian foundations are not necessarily very strong even in individual learning problems.[17] Our view in this paper is that assuming that the agents already start with common priors is quite restrictive to explore complex belief dynamics and to understand where the limiting beliefs of a set of individuals come from. As required in any standard model of belief dynamics, the agents first have to start with some priors, but we do not restrict them to be necessarily common.[18]

3.1 External Sources

The realized parameter value θ is not observed directly by any agent. Instead, each agent obtains some private (noisy) information about the parameter through an external (idiosyncratic) source of information. The interpretation of an agent's source is that of a mean, technology or institution through which he carries out his private research over time about the parameter. To fix ideas, we can think of a source as a newspaper to which the agent is subscribed. Say, for example, that one agent i obtains information about some financial variable by reading the Wall Street Journal (WSJ) every week while other agent j is subscribed to the Financial Times (FT).[19]

Agent i's source generates in each period $t \geq 1$ a *signal realization* $s_{it} \in S = \{s_1, s_2, \ldots, s_L\}$ which is privately observed by i. When the true parameter value is θ, agent i's information source delivers signal $s \in S$ with probability $\phi_i^\theta(s)$. We use $\phi_i(s)$ to denote the corresponding unconditional distribution. A *signal profile* in period t is denoted by $\boldsymbol{s}_t = (s_{it})_{i \in N} \in S^n$. An *external source* for agent i is a set of conditional distributions over signals

$$\Phi_i := \left\{ \phi_i^\theta \in \Delta(S) \,:\, \theta \in \Theta \right\}.$$

Throughout the paper, we impose the following technical assumption on the distributions that specify the external sources.

not necessarily finite, parameter space would only change sums to integrals in the appropriate formulae.

[17]For example, Acemoglu, Chernozhukov, and Yildiz (2009) show that, under mild assumptions, Bayesian updating from signals does not necessarily lead to agreement about the parameter true value. This result challenges the classical justification for the common priors premise.

[18]Nevertheless, to ease the technical details and notational burden, the result in Lemma 1 and all our examples are presented for the common priors case.

[19]By construction, the information that the agent receives through his source does not include any information that he can receive from other agents in the society.

Assumption 1. For each agent i, there exists at least a parameter value θ such that the conditional distribution $\phi_i^\theta \in \Delta(S)$ has full support.

Assumption 1 above guarantees that the agents' limiting beliefs are well defined. In addition, some results of the paper will require that we strengthen further our assumptions on these distributions by imposing the following requirement instead.

Assumption 2. For each agent i and each parameter value θ, the conditional distribution $\phi_i^\theta \in \Delta(S)$ has full support.

Using Bayes' rule, a source Φ_i enables us to update, in each period $t \geq 1$, any belief about the parameter θ. Because distributions over signals are constant over time, this updating process is *stationary.* Let $s_i^t = (s_{i1}, \ldots, s_{it})$ be a *sequence of signals observed by agent i up to period t.* Then, when agent i uses the source Φ_i, we denote by $q_i^{s_i^t} \in \Delta(\Theta)$ his posteriors about the parameter upon observing the sequence of signals s_i^t. By considering that agents learn in an imprecise way using their sources, we attempt to capture practical situations in which either the quality of the reports provided by the source, the available information transmission technology, the agent's degree of attention or his understanding/cognitive capabilities do not allow him to fully learn the parameter value.[20]

3.2 The Social Network and its Directed Links

We consider that the agents receive information not only from their sources but they can also listen to the opinions of others. More precisely, the agents are connected through an exogenous social network which, in each period t, allows them to pay attention to the opinions about θ that others are forming themselves using their own sources. We focus on directed networks where links are one-sided.

A directed link from agent i to agent j is denoted by Ψ_{ij} and it allows i to receive messages from j over time. Specifically, we assume that, in each period $t \geq 1$, each agent i receives a (private) *message realization* $m_{ijt} \in M = \{m_1, m_2, \ldots, m_L\}$ from each agent j to whom he has a directed link.[21] The message m_{ijt} is correlated with the signal s_{jt} so that it conveys some (noisy) information about the signal that the sender j is observing at that t. Then, agent i can use such information about s_{jt} to update his beliefs about θ. Intuitively, through this type of communication, agent i has some (noisy) access to the private research about θ that agent j

[20] Nevertheless, our model also includes the possibility that the agents do obtain full information about θ using their sources. Following the terminology of sender-receiver games, this is the extreme case described by a *completely separating* information source.

[21] We assume that $|S| = |M| = |\Theta|$ in order to allow both an information source and a directed link for full information disclosure.

conducts using his source Φ_j. In the next subsection, we will formally present the definition of a directed link Ψ_{ij}, the particular assumptions imposed on the agents' informational capabilities, and the way in which they use a directed link to update their beliefs.

Besides direct (perhaps noisy) attention to the sources of others, we consider that the network also allows for indirect attention. More precisely, we assume that messages can be transmitted indirectly through directed links. Hence, given two directed links Ψ_{ik} and Ψ_{kj}, agent k can receive a message from agent j and then pass it through to agent i. Through the links Ψ_{ik} and Ψ_{kj}, agent i receives, in each period, two different messages, one *direct* message from agent k (m_{ik}) and one *indirect* message from agent j (m_{ij}, which was previously received by agent k from agent j). The message m_{ik} conveys information about the signal s_k observed by agent k while the message m_{ij} conveys information about the signal s_j observed by agent j. Nonetheless, for the clarity of exposition, it is convenient to focus first on the description of the transmission only of direct messages through links.

3.3 Information Transmission (only) with Direct Messages

Suppose for the moment that the agents receive only direct messages so that an agent cannot pass to another agent a message that he has received from a third agent. A *message vector*[22] received by agent i in period t is denoted by $\boldsymbol{m}_{it} = (m_{ijt})_{j \in N \setminus \{i\}}$ and a *message profile* in period t is denoted by $\boldsymbol{m}_t = (\boldsymbol{m}_{it})_{i \in N}$. For each period $t \geq 1$, the distribution over messages observed by agent i, conditional on agent j's signal realization $s_{jt} = s$, is denoted by σ_{ij}^s and the corresponding unconditional distribution is denoted by σ_{ij}. We then define a *message protocol* from agent j to agent i as a set of conditional distributions over messages $\Sigma_{ij} := \left\{ \sigma_{ij}^s \in \Delta(M) \,:\, s \in S \right\}$.

Bayes' rule allows us to use the message protocol Σ_{ij} to obtain, in each period t, some posteriors about j's private signal s_{jt} by observing the message m_{ijt}. More importantly, by combining the source Φ_j with the message protocol Σ_{ij} according to the total probability rule, we can obtain Bayesian posteriors about θ. As in the case of external sources, this updating process is stationary since distributions over messages are constant over time.

The *directed link* from agent i to agent j associated with the signal Φ_j and the message protocol Σ_{ij} is the set of conditional distributions over messages

$$\Psi_{ij} := \left\{ \psi_{ij}^{\theta} \in \Delta(M) \,:\, \theta \in \Theta, \;\; \text{such that} \;\; \psi_{ij}^{\theta}(m) = \sum_{S} \sigma_{ij}^{s}(m) \phi_j^{\theta}(s) \right\}. \tag{1}$$

[22] In principle, our description of message vectors captures a situation where each agent receives messages from each other agent in the society. Nevertheless, the specification of a link Ψ_{ij} will determine the degree of informativeness of the messages m_{ijt} that flow through it. In some cases, the corresponding degree of informativeness may be null, which is interpreted as if there is actually no directed link from agent i to agent j and, therefore, as if i receives no message whatsoever from j.

Intuitively, the role of a directed link Ψ_{ij} is that of allowing agent i to have some access to agent j's private signal and, by doing so, to update his beliefs about the parameter θ.

Let $m_{ij}^t = (m_{ij1}, \ldots, m_{ijt})$ be a *sequence of messages received by agent i from agent j up to period t*. We use $q_i^{m_{ij}^t} \in \Delta(\Theta)$ to denote agent i's posteriors about the parameter upon receiving the sequence of messages m_{ij}^t from agent j.

3.3.1 The Informational Capabilities of the Agents

At this point, we describe the method of information processing followed by the agents. Specifically, we describe what knows privately each individual and what is commonly known between some agents. Also, we describe how each agent combines his priors and the information he receives, both from his source and from his communication with others, to obtain his posteriors at each date. Unlike most Bayesian learning models, this paper does not assume that all agents have the same beliefs about the informational primitives of the model (i.e., priors p_i, external sources Φ_i, and message protocols Σ_{ij}, for each $i, j \in N$). Neither it does assume that such beliefs coincide with the true data generating processes nor that there is common knowledge that all agents share such beliefs. Our view is that while these informational requirements are stringent even in individual learning problems, they become exponentially complex and demanding in social contexts.[23] The assumption that each agent is only uncertain about the value of θ and that there is no doubt about the underlying "model of the world" is perhaps not a good approximation to reality in social environments. Furthermore, given the emphasis of this paper on real-world large networked societies, the degree of complexity necessary for each agent be able to form higher-order accurate conjectures about the private signals that each other agent is receiving in each period seems too high. Given these challenges of the Bayesian method in social situations, we take a pragmatic approach to consider the agents' informational capabilities when it comes to complex issues.

First, we assume that priors are private information and that agents do not incorporate others' priors into their own information updating processes. By construction, agents are not interested in making any inferences about others' priors. In short, we assume that each agent begins with his priors, updates them over time using his own source, and, under the restrictions imposed by the network structure, makes some inferences over time about the information that other agents are receiving from their sources.

Secondly, we assume that for each directed link Ψ_{ij}, the conditional distributions associated

[23] For a discussion both of (a) the challenges implied by this type of informational requirements in social situations and of (b) the of restrictiveness of the assumption of common priors and common knowledge of the true generating data processes see, e.g., Acemoglu and Ozdaglar (2011)'s excellent survey of Bayesian and non-Bayesian models of opinion influence in social networks.

to the source Φ_j and the message protocol Σ_{ij} are commonly known *only* between agents i and j.[24] Therefore, using the total probability rule, as expressed in (1) above, agents i and j have common knowledge about Ψ_{ij}. Agent i is then able to use privately each conditional probability law ψ_{ij}^{θ} ($\theta \in \Theta$) to update his priors about θ in each $t \geq 1$. The interpretation is that agent i has an understanding of what type of messages to expect from j when the true parameter value is θ. However, in general, agent i cannot use source Φ_j with the same precision as j does because he is making noisy inferences about the signals s_{jt} that j is receiving.[25] Then, agent i places himself in j's position and, under the restrictions imposed by the inferences that he makes about the signals s_{jt}, uses Φ_j to update his own priors p_i, without making any use of j's priors. Going back to our newspaper subscription example, the intuition is that the WSJ's subscriber cannot observe directly what the FT's subscriber is actually learning by reading each weekly issue of the FT. Instead, the WSJ's subscriber receives (noisy) messages from the FT's subscriber that gives him some information about what the latter is learning by reading his weekly issues. In short, the WSJ's subscriber cannot read directly the FT but can hear the FT's subscriber speak about what he is reading.

Given the message protocol Σ_{ij}, the information transmission process from j to i about the signals s_{jt} through the messages m_{ijt} is modeled as in the canonical *cheap talk* framework.[26] Therefore, this model does include the possibility that agent j transmits to agent i, with full precision, the signal s_{jt} that he is observing.[27] However, the possibility that message protocols do not transmit signals with full precision is also allowed. The idea here is to capture practical situations where, because of our understanding capabilities or the available information transmission technology, we are not able, by listening to someone, to learn as much as he actually does using his own information source. In many real-world situations, some information is often lost in the inter-person communication process. This decay result is formally obtained in Lemma 1. The amount of information that agent i receives from j using the directed link Ψ_{ij} is described by its weight and, in the next subsection, we introduce formally our approach to measure such weight.

Because of the above assumptions on the agents' informational capabilities, we consider that each agent forms only first-order beliefs about θ using the Bayesian updating rules attached to his source and links. However, he is not able to form any accurate higher-order conjectures about

[24]Thus, agents i and j commonly know the main features of the link that connects them but such information remains unknown to any other agent.

[25]Recall that, although the conditional distributions associated with Φ_j are known by an agent i who has a link Ψ_{ij}, the particular signal realizations s_{jt} remain j's private information.

[26]See, e.g., the classical sender-receiver framework introduced by Crawford and Sobel (1982).

[27]Using the terminology of sender-receiver games, this is the extreme case described by a *completely separating* message protocol Σ_{ij}.

the histories underlying the signals that all agents are privately observing. Thus, given these constraints imposed their information processing, agents are not fully rational in this model. Yet, with regards to the information transmission through each link, they update their beliefs in a Bayesian manner (in the same fashion as in a sender-receiver game). This gives us a tractable reference framework to which the non-Bayesian models related to the DeGroot's benchmark of opinion influence can be compared.

Finally, we do not allow for the possibility that the agents manipulate strategically the messages they send, neither that they lie or withhold any information they posses about their signals. The information they send to others is noisy but exogenously determined. A plausible interpretation is that the agents make some investments in their links (e.g., investments in technological connexions, friendship or club relations, social networks through mobile devices or the internet) in a way such that only *hard information* can subsequently flow through them. Then, once the links are formed, their weights remain fixed and cannot be altered neither by senders nor by receivers.

3.4 Measuring the Informativeness of Sources and Links

To measure the degree of informativeness about θ attached to the agents' sources and to the communication between them, we use some entropy-based concepts.

Definition 1. Let X be a finite set. The *entropy* (or *Shannon entropy*) of a probability distribution $P \in \Delta(X)$ is[28]

$$H(P) := -\sum_{X} P(x) \log P(x).$$

The entropy of a distribution is always nonnegative and measures the average information content one is missing from the fact that the true realization of the associated random variable is unknown. In other words, it measures the ex-ante uncertainty of the corresponding random variable. To measure the information content of sources and links, we rely on the concept of relative entropy between distributions.

Definition 2. Let X be a finite set and let $P, Q \in \Delta(X)$. The *relative entropy* (or *Kullback-Leiber distance*) of P with respect to Q is[29]

$$D(P \,||\, Q) := \sum_{X} P(x) \log \frac{P(x)}{Q(x)}.$$

[28]In Definition 1, it follows the convention $0 \log 0 = 0$, which is justified by continuity.

[29]The following conventions are used: $0 \log(0/0) = 0$ and, based on continuity arguments, $0 \log(0/a) = 0$ and $a \log(a/0) = \infty$.

The relative entropy is not a metric[30] but it constitutes a formal measure of the gain of information in moving from distribution Q to P. The relative entropy is always nonnegative and equals zero if and only if $P = Q$ almost everywhere.

We apply the relative entropy to the agents' posteriors in period $t = 1$ with respect to their priors and, specifically, define the *power of the external source* Φ_i as the expectation (over the possible signals $s_i^1 \in S$ that i can observe) of the relative entropy of $q_i^{s_i^1}$ with respect to p_i.

Definition 3 (**Power of the external source).**

$$\mathbb{P}(\Phi_i) := \sum_S \phi_i(s_i^1) D\big(q_i^{s_i^1} \,||\, p_i\big). \tag{2}$$

The power measure allows us to rank completely any set of sources according to their degree of informativeness. Then, we say that Φ_i *is at least as informative as* Φ_i' if $\mathbb{P}(\Phi_i) \geq \mathbb{P}(\Phi_i')$. Notice that the power of a signal is a relevant measure to study how an agent's beliefs evolve using only his private learning, when he does not listen to anyone else, i.e., to study the evolution of the posteriors $q_i^{s_i^t}$. In our social context, however, agents do hear others' opinions and, therefore, the relevant measure to study the beliefs dynamics of an agent when communication is allowed for must take into account the messages that he receives. To this end, we define, in a way analogous to the power of a source, the *power of the directed link* Ψ_{ij} as the expectation (over the possibles messages $m_{ij}^1 \in M$ that i can receive from j) of the relative entropy of the posterior $q_i^{m_{ij}^1}$ with respect to the prior p_i.

Definition 4 (**Power of the directed link).**

$$\mathbb{P}(\Psi_{ij}) := \sum_M \psi_{ij}(m_{ij}^1) D\big(q_i^{m_{ij}^1} \,||\, p_i\big). \tag{3}$$

Using the power of a directed link, we say that Ψ_{ij} *is at least as informative as* Ψ_{ij}' if $\mathbb{P}(\Psi_{ij}) \geq \mathbb{P}(\Psi_{ij}')$ (i.e., the weight of Ψ_{ij} is at least as larger as the weight of Ψ_{ij}'). If $\mathbb{P}(\Psi_{ij}') = 0$, then we interpret this as if there is actually no directed link from agent i to agent j. To define $\mathbb{P}(\Phi_i)$ and $\mathbb{P}(\Psi_{ij})$, we focus on the discrepancy only between posteriors in $t = 1$ and priors since the associated Bayesian updating processes are stationary.

The role of the link Ψ_{ij} is to enable i to update his beliefs about θ by listening to j. Recall that the link does so by allowing i to obtain some (perhaps noisy) information about the signals that j is observing. Therefore, the power (or weight) $\mathbb{P}(\Psi_{ij})$ depends positively on both the quality of message protocol Σ_{ij} and the quality of the source Φ_j (which is, in turn, measured by $\mathbb{P}(\Phi_j)$). This can be noted from the expression in (1), which implies that the

[30] In particular, the relative entropy is not symmetric and it does not satisfy the triangle inequality either.

distributions $\psi_{ij}^{\theta}, \psi_{ij} \in \Delta(M)$, as well as the induced Bayesian posteriors $q_i^{m_{ij}^t} \in \Delta(\Theta)$, depend on the (exogenous) specifications of Σ_{ij} and Φ_j. To see formally that the power of a link Ψ_{ij} is positively related to the power of the source Φ_j, let us consider for technical simplicity the case with common priors. Thus, if all agents begin with some priors $p \in \Delta(\Theta)$, then it can be verified that $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j) + R(\Sigma_{ij})$, where the term $R(\Sigma_{ij})$ is specified as

$$R(\Sigma_{ij}) := \sum_{\Theta} \sum_{M} \sum_{S} p(\theta)\phi_j^{\theta}(s)\sigma_{ij}^{s}(m) \log \frac{\phi_j(s) \sum_S \sigma_{ij}^{s'}(m)\phi_j^{\theta}(s')}{\phi_j^{\theta}(s) \sum_S \sigma_{ij}^{s'}(m)\phi_j(s')}.$$

The term $R(\Sigma_{ij})$ is always nonpositive and it equals zero if and only if Σ_{ij} is a *completely separating* message protocol (i.e., when Σ_{ij} fully reveals to agent i the signal that agent j observes).[31] Intuitively, a high power of a link directed to a FT's subscriber is obtained when either the particular subscription provides him with good information (because of the available technology or the attention that the subscriber puts on his reading), or we obtain good information from the FT's subscriber about what he is reading (because of our attention to what he says or his communication skills, or the technology through which we communicate), or both.

We are ready now to define a directed network. A *directed network* Ψ is a set of directed weighted links which connects the agents in the society:

$$\Psi := \{\Psi_{ij} \,:\, i, j \in N, i \neq j, \ \text{ such that } \ \mathbb{P}(\Psi_{ij}) > 0\}.$$

We turn now describe how information is transmitted through indirect messages.

3.5 Information Transmitted (both) with Direct and Indirect Messages

To describe the transmission of information through indirect connexions, we now extend some of the concepts introduced in the previous subsection. A *directed path* from agent i to agent j is a sequence $\gamma_{ij} = (\Psi_{ii_1}, \Psi_{i_1 i_2}, \dots, \Psi_{i_K j})$ of directed links such that $\mathbb{P}(\Psi_{ii_1}) > 0$, $\mathbb{P}(\Psi_{i_K j}) > 0$, and $\mathbb{P}(\Psi_{i_k i_{k+1}}) > 0$ for each $k \in \{1, \dots, K-1\}$. We use $\Gamma_{ij}[\Psi]$ to denote the set of all directed paths from agent i to agent j under network Ψ and $N_i := \{j \in N \,:\, \text{there is some } \gamma_{ij} \in \Gamma_{ij}[\Psi]\}$ to denote set of agents to whom agent i has a directed path. A directed network Ψ is *connected* if, for each agent $i \in N$, there is at least one directed path $\gamma_{ij} \in \Gamma_{ij}[\Psi]$ to each other agent $j \in N \setminus \{i\}$. Intuitively, a network is connected if it allows each agent to hear (either directly or indirectly) the opinions of each other agent in the society. Network connectedness can be regarded as a basic prerequisite to study the achievement of a consensus.

For a network Ψ, an agent i may receive messages from other agent j through (possibly) multiple paths $\gamma_{ij} \in \Gamma_{ij}[\Psi]$. To avoid informational redundancies, we will restrict attention to

[31] The technical arguments under these claims appear in the proof of Lemma 1.

those paths which convey the highest amount of information.[32]

For the transmission of indirect messages, we make the natural assumption that an agent k uses the same conditional distribution $\sigma_{ik}^{s_k} = \sigma_{ik}^{m_{kj}}$ to transmit information to agent i both about the signal s_k that he observes and about the message m_{kj} that he receives from another agent j. The interpretation is that we consider the existence of a common technology for information transmission, which is equally used for both signals and messages that pass from one agent to another.

Let $\psi_{ij}^{\theta}[\gamma_{ij}] \in \Delta(M)$ denote the distribution over messages received by i from j through the directed path γ_{ij}, conditional on the parameter value being θ. Let the corresponding unconditional distribution be denoted by $\psi_{ij}[\gamma_{ij}]$. For a directed path γ_{ij}, let $\pi_{\gamma_{ij}}^{s} \in \Delta(M)$ denote the conditional distribution over messages received by agent i from agent j, conditional on agent j observing signal $s_j = s$. Then, for a directed path $\gamma_{ij} = (\Psi_{ii_1}, \Psi_{i_1 i_2}, \ldots, \Psi_{i_K j})$, using the total probability rule, we obtain, for the message $m_{ij} = m$:

$$\pi_{\gamma_{ij}}^{s}(m) = \sum_{M} \cdots \sum_{M} \sigma_{ii_1}^{m_{i_1 i_2}}(m) \prod_{k=1}^{K-2} \sigma_{i_k i_{k+1}}^{m_{i_{k+1} i_{k+2}}}(m_{i_k i_{k+1}}) \sigma_{i_{K-1} i_K}^{m_{i_K j}}(m_{i_{K-1} i_K}) \sigma_{i_K j}^{s}(m_{i_K j}). \tag{4}$$

Using the distribution $\pi_{\gamma_{ij}}^{s}$ specified above, we then obtain the conditional probability that agent i receives indirectly a message m through the path $\gamma_{ij} = (\Psi_{ii_1}, \Psi_{i_1 i_2}, \ldots, \Psi_{i_K j})$ when the true parameter value is θ, i.e., $\psi_{ij}^{\theta}[\gamma_{ij}](m) = \sum_{S} \pi_{\gamma_{ij}}^{s}(m) \phi_j^{\theta}(s)$. We use $q_i^{m_{ij}^t}[\gamma_{ij}] \in \Delta(\Theta)$ to denote agent i's posteriors about the parameter, conditional on receiving the sequence of messages m_{ij}^t from agent j through the path γ_{ij}. We can extend straightforwardly the concept of power of a link to a path. The *power of the path* γ_{ij} is defined as

$$\mathbb{P}(\gamma_{ij}) := \sum_{M} \psi_{ij}[\gamma_{ij}](m_{ij}^1) D\big(q_i^{m_{ij}^1}[\gamma_{ij}] \,||\, p_i\big). \tag{5}$$

Note that large values $\mathbb{P}(\gamma_{ij})$ of the power of a path are associated with a highly informative source Φ_j and/or highly informative message protocols $\Sigma_{ii_1}, \Sigma_{i_1 i_2}, \ldots, \Sigma_{i_K j}$. Formally, from the expression in (4), the distributions $\psi_{ij}^{\theta}[\gamma_{ij}], \psi_{ij}[\gamma_{ij}] \in \Delta(M)$, and the induced Bayesian posteriors $q_i^{m_{ij}^t}[\gamma_{ij}] \in \Delta(\Theta)$, depend on the description of j's source and of the message protocols throughout the path γ_{ij}.

[32]Note that the information that an agent i receives from another agent j through two different paths $\gamma_{ij}, \gamma'_{ij} \in \Gamma_{ij}[\Psi]$ refers only to the information provided by the same source Φ_j to agent j. Thus, the information that flows through any of these paths does not include any information attached to the sources of the agents located along any of the two paths. Suppose, e.g., that $\mathbb{P}(\gamma_{ij}) > \mathbb{P}(\gamma'_{ij})$. This corresponds intuitively to a situation where the information about j's private learning that flows through γ'_{ij} is relatively more affected by decay (as formally described in Lemma 1) than the information that flows through γ_{ij}. Thus, γ_{ij} and γ'_{ij} would provide i with two different Bayesian updating processes about j's private learning from Φ_j. Yet, since these two processes cannot be combined to obtain a more informative updating process, we choose to focus only on the path γ_{ij}.

As mentioned earlier, agent i can receive indirect messages from another agent j through several different paths in the network. We then restrict attention to those paths in the set

$$\left\{\widehat{\gamma}_{ij} \in \Gamma_{ij}[\Psi] \,:\, \mathbb{P}(\widehat{\gamma}_{ij}) \geq \mathbb{P}(\gamma_{ij}) \quad \forall \gamma_{ij} \in \Gamma_{ij}[\Psi]\right\}.$$

If the set above is not singleton, then we randomly pick one of its elements as our path of interest and denote it by $\widehat{\gamma}_{ij}$. For future reference, we will denote $\psi^{\theta}_{ij}[\widehat{\gamma}_{ij}] =: \widehat{\psi}^{\theta}_{ij}$ and $\psi_{ij}[\widehat{\gamma}_{ij}] =: \widehat{\psi}_{ij}$.

An implication of our characterization of limiting beliefs, Theorem 1, is that the agents' priors do not influence their limiting beliefs. In this model, an agent i's limiting beliefs are entirely determined by the influences or biases described by both his information source Φ_i and by all his directed paths $\widehat{\gamma}_{ij}$. Nevertheless, as we will discuss in Subsection 4.2, agents' priors do affect the speed of convergence of their first-order posteriors to their limiting beliefs.

3.6 Evolution of Beliefs, Consensus, and Correct Beliefs

Let us introduce a few additional concepts to analyze the evolution of the agents' first-order posteriors. A *period-t history for agent i* is a vector of sequences $h^t_i := \left(s^t_i, (m^t_{ij})_{j\neq i}\right)$ of signals and messages vectors received by agent i up to period t. Let H^t_i be all histories of length t for player i and let $H_i = \cup_{t\geq 1} H^t_i$ be all histories for player i and let $h_i \in H_i$ denote a generic history for agent i. We impose the following assumption on the (conditional) independence of the signals and messages that any agent receives.

Assumption 3. For each agent $i \in N$, each history $h_i \in H_i$ is independent, conditional on the parameter realization θ, both across periods $t \geq 1$ and across senders $j \neq i$.

Agent i's posteriors about θ are then given by the random variable $q^{h^t_i}_i(\theta) : \Theta \to [0, 1]$. For each agent i and each value of the parameter θ, the sequence of random variables $\left\{q^{h^t_i}_i(\theta)\right\}^{\infty}_{t=1}$ is a bounded martingale,[33] which ensures that the agents' posterior beliefs converge almost surely (see, e.g., Billingsley, 1995, Theorem 35.5).

Under our assumptions on the agents' informational capabilities and on the conditional independence of signals both across periods and agents, Assumption 3, Bayes' rule gives us:

$$q^{h^1_i}_i(\theta) = \frac{\phi^{\theta}_i(s_{i1}) \prod_{j\in N_i} \widehat{\psi}^{\theta}_{ij}(m_{ij1}) p_i(\theta)}{\sum_{\theta'\in\Theta} \phi^{\theta'}_i(s_{i1}) \prod_{j\in N_i} \widehat{\psi}^{\theta'}_{ij}(m_{ij1}) p_i(\theta')}$$

and

$$q^{h^t_i}_i(\theta) = \frac{\phi^{\theta}_i(s_{it}) \prod_{j\in N_i} \widehat{\psi}^{\theta}_{ij}(m_{ijt}) q^{h^{t-1}_i}_i(\theta)}{\sum_{\theta'\in\Theta} \phi^{\theta'}_i(s_{it}) \prod_{j\in N_i} \widehat{\psi}^{\theta'}_{ij}(m_{ijt}) q^{h^{t-1}_i}_i(\theta')} \qquad \forall t > 1.$$

[33] More formally, $\left\{q^{h^t_i}_i(\theta)\right\}^{\infty}_{t=1}$ is a bounded martingale with respect to the (conditional) measure on Θ which is induced by the priors p_i, $i \in N$, and the conditional distributions ϕ^{θ}_i, $\widehat{\psi}^{\theta}_{ij}$, for $i, j \in N$.

Then, by iterating recursively the expressions above, we have:

$$q_i^{h_i^t}(\theta) = \frac{\prod_{\tau=1}^{t} \phi_i^{\theta}(s_{i\tau}) \prod_{j\in N_i} \widehat{\psi}_{ij}^{\theta}(m_{ij\tau}) p_i(\theta)}{\sum_{\theta'\in\Theta} \prod_{\tau=1}^{t} \phi_i^{\theta'}(s_{i\tau}) \prod_{j\in N_i} \widehat{\psi}_{ij}^{\theta'}(m_{ij\tau}) p_i(\theta')}.$$

Furthermore, the expression above can be conveniently rewritten as

$$q_i^{h_i^t}(\theta) = \left[1 + \sum_{\theta'\neq\theta} \prod_{\tau=1}^{t} \frac{\phi_i^{\theta'}(s_{i\tau}) \prod_{j\in N_i} \widehat{\psi}_{ij}^{\theta'}(m_{ij\tau}) p_i(\theta')}{\phi_i^{\theta}(s_{i\tau}) \prod_{j\in N_i} \widehat{\psi}_{ij}^{\theta}(m_{ij\tau}) p_i(\theta)} \right]^{-1}, \tag{6}$$

which will be our key equation to study the agents' limiting beliefs.

Definition 5. A *consensus is (asymptotically) achieved in the society* if the posterior beliefs of all agents converge to the same value regardless of their priors, that is, if for each $i \in N$, each $p_i \in \Delta(\Theta)$, and for some (common) probability distribution $p \in \Delta(\Theta)$, we have $\lim_{t\to\infty} q_i^{h_i^t} = p$.

Our notion of correct beliefs requires that the network permits the aggregation of the pieces of information transmitted by the agents' sources. Let us consider an external observer who has access to the sources available to all agents in the society but cannot use any connexion in the network. The observer's priors are given by a distribution $p_{\text{ob}} \in \Delta(\Theta)$. A *period-t history for the external observer* is a sequence $h^t := (\boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_t)$ of signal profiles. Let H^t be the set of all histories of length t for the external observer and $H = \cup_{t\geq 1} H^t$ be the set of all histories for the external observer. The posteriors of the external observer about θ are given by the random variable $q_{\text{ob}}^{h^t}(\theta) : \Theta \to [0,1]$.[34] With these preliminaries in hand, correct limiting beliefs require that the communication processes allowed by the network structure aggregate the diverse information obtained by the agents (from their sources), exactly such as the external observer does.

A key observation to justify our approach to correct limiting beliefs is that, for large enough societies, the observer's limiting beliefs are arbitrarily accurate estimates of the true parameter value. It follows from a standard law of large numbers argument[35] that, according to the information obtained by aggregating all distributions ϕ_i^{θ} ($i \in N$), the resulting posteriors converge to put probability one on the true parameter value. Of course, this approach is intuitively more compelling for large enough societies.

Definition 6. The directed network Ψ attains *correct limiting beliefs* if a consensus is achieved in the society and, in addition, for each $i \in N$, we have $\lim_{t\to\infty} q_i^{h_i^t} = \lim_{t\to\infty} q_{\text{ob}}^{h^t}$.

[34] Again, for each value of the parameter θ, the sequence of random variables $\{q_{\text{ob}}^{h^t}(\theta)\}_{t=1}^{\infty}$ is a bounded martingale so that the external observer's posteriors converge almost surely.

[35] For instance, Doob's (1994) consistency Theorem.

The external observer's posteriors are obtained analogously to those of any agent (as derived in expression (6)) and, therefore, we omit the details.

Finally, note that this paper makes use of the notion of surely convergence (or pointwise convergence) of random variables in order to study whether posteriors converge to some limiting beliefs. Since surely convergence implies almost surely convergence, if a consensus occurs and the number of agents is sufficiently large, then our notion of correct limiting beliefs agrees with the notion of *asymptotic learning*[36] used by some of the related literature (see, e.g., Acemoglu, Dahleh, Lobel, and Ozdaglar, 2011).

4 Results

4.1 Decay in the Flow of Information

Informativeness levels decrease as information flows from one link to another. By construction, this model captures the presence of some (exogenous) decay in the transmission of information since it allows for sources and message protocols that do not fully disclose information. Only when the message protocol Σ_{ij} is *completely separating*, there is no loss of information and thus agent i obtains, through the link Ψ_{ij}, exactly the same amount of information about θ as agent j does using directly his source Φ_j. With the same logic, decay is also present when information is conveyed by indirect messages. For the common priors case, the following lemma provides these intuitive results in terms of the power measure.

Lemma 1 (**Decay in the Transmission of Information**). *(a) Consider a directed link in a social network $\Psi_{ij} \in \Psi$, and suppose that agents i and j have the same priors p. Then, $\mathbb{P}(\Psi_{ij}) \leq \mathbb{P}(\Phi_j)$. Moreover, $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j)$ if and only if the message protocol Σ_{ij} associated with the directed link Ψ_{ij} is such that agent i fully learns the signals observed by agent j from the external source associated with Φ_j. (b) Consider a directed path in a social network $\gamma_{ij} = (\Psi_{ik}, \Psi_{kj}) \in \Gamma_{ij}(\Psi)$. Suppose that agents i and j have the same priors p. Then, $\mathbb{P}(\gamma_{ij}) \leq \mathbb{P}(\Psi_{kj})$. Moreover, $\mathbb{P}(\gamma_{ij}) = \mathbb{P}(\Psi_{kj})$ if and only if the message protocol Σ_{ik} associated with the directed link Ψ_{ik} is such that agent i fully learns the messages received by agent k from agent j.*

For the heterogeneous priors case, the insights provided by Lemma 1 (a) continue to hold if the result is rephrased as follows. Suppose that, as an alternative to his private signal Φ_j, agent

[36] There are different formulations of asymptotic learning in social contexts. For the benchmark proposed in this paper, asymptotic learning would require that, as the number of agents tends to infinity, the average of the posteriors converge almost surely to some beliefs that place probability one on the true parameter value. If a consensus occurs, then the average of the limiting beliefs trivially coincides with the consensus beliefs. As a consequence, the surely convergence criterion used in our notion of correct limiting beliefs would imply the almost surely convergence required for asymptotic learning.

j places himself in agent i's position and uses the directed link Ψ_{ij} to update his beliefs about the parameter. Then, the information about θ that agent j receives through this directed link Ψ_{ij} is less precise than the information that he would obtain using directly his signal Φ_j. The insights provided by Lemma 1 (b) also continue to hold under the analogous restatement of the result. Nevertheless, if two agents i and j begin with different priors and we simply ask about the relation between $\mathbb{P}(\Psi_{ij})$ and $\mathbb{P}(\Phi_j)$, then it could well be the case that $\mathbb{P}(\Psi_{ij}) > \mathbb{P}(\Phi_j)$. This is due to the role that the agents' priors have on the power measures of sources and links.

What can we say about the power of a source Φ_j and/or of a directed link Ψ_{ij} in the particular cases where one, or both of them, allow for full information disclosure?[37] To avoid the formal difficulties implied in the heterogenous priors case, suppose that the agents begin with some common priors p. Then, for any agent $j \in N$, it can be verified that $\mathbb{P}(\Phi_j) = H(p) - \sum_{s_j^1 \in S} \phi_j(s_j^1) H(q_j^{s_j^1})$ so that $\mathbb{P}(\Phi_j) \leq H(p)$. Thus, $\mathbb{P}(\Phi_j) = H(p)$ if and only if the average entropy of agent j's posteriors (obtained only from his source) vanishes. In other words, $\mathbb{P}(\Phi_j) = H(p)$ in the particular case in which j obtains full information about the parameter from his source. In addition, for agent i to obtain full information about the parameter from a directed link to agent j, it must be the case that (a) agent i obtains full information about the learning process carried out by j using his source (i.e., $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j)$) and that (b) agent j obtains full information about the parameter from his source (i.e., $\mathbb{P}(\Phi_j) = H(p)$). Therefore, from the result in Lemma 1, we observe that $\mathbb{P}(\Psi_{ij}) \leq H(p)$ for each $i, j \in N$, i.e., for the common priors case, the entropy $H(p)$ constitutes an upper bound on the degree of informativeness about θ that any agent in the society can obtain, regardless of the network structure.[38]

4.2 Characterizing Limiting Beliefs and Consensus

For an agent i, we define the function $G_i : \Theta \to \mathbb{R}$ as

$$G_i(\theta) := \sum_S \phi_i(s) \log \phi_i^\theta(s). \tag{7}$$

The value $G_i(\theta)$ is always negative and describes the (average) likelihood that the source Φ_i assigns to θ being the true parameter value. Let $\Theta_i \subseteq \Theta$ be the set specified as $\Theta_i := \arg\max_{\theta \in \Theta} G_i(\theta)$. To account for the information that an agent i receives from another agent

[37] This corresponds formally to *completely separating* Φ_j and/or Σ_{ij}. In intuitive terms, it describes situations where agent j learns from his source without any noise and/or he directly transmits to agent i directly the signals that he observes, instead of the (noisy) messages.

[38] For the heterogenous priors case, an analogous upper bound on $\mathbb{P}(\Psi_{ij})$ that depends on both entropies $H(p_i)$ and $H(p_j)$ can be derived. We do not provide the details since it only implies a more sophisticated mathematical expression which, however, conveys no further intuitions.

j, we define the function $F_{ij} : \Theta \to \mathbb{R}$ as

$$F_{ij}(\theta) := \sum_{M} \widehat{\psi}_{ij}(m_{ij}) \log \widehat{\psi}^{\theta}_{ij}(m_{ij}). \tag{8}$$

The value $F_{ij}(\theta)$ is always negative and describes the (average) likelihood that the most informative directed path from agent i to agent j, $\widehat{\gamma}_{ij}$, assigns to θ being the true parameter value. For an agent i, we then specify the set $\Theta^*_i \subseteq \Theta$ as $\Theta^*_i := \arg\max_{\theta \in \Theta} \left\{ G_i(\theta) + \sum_{j \in N_i} F_{ij}(\theta) \right\}$.

We obtain the interesting feature that the various influences/biases imposed on an agent's beliefs by his source and by the opinions that he hears from others are additively aggregated to determine his limiting beliefs.

Theorem 1. *Consider a social network Ψ and suppose that Assumptions 1 and 3 hold. Then, for any history $h^t_i \in H_i$, agent i's limiting beliefs satisfy:*
(i) $\lim_{t \to \infty} q_i^{h^t_i}(\theta) = 0$ *for each* $\theta \notin \Theta^*_i$;
(ii) if Θ^*_i is singleton so that $\Theta^*_i = \{\theta^*\}$ for some $\theta^* \in \Theta$, then $\lim_{t \to \infty} q_i^{h^t_i}(\theta) = 1$;
(iii) if Θ^*_i is not singleton, then $\lim_{t \to \infty} q_i^{h^t_i}(\theta) = p_i(\theta) / \sum_{\Theta^*_i} p_i(\theta')$ *for each* $\theta \in \Theta^*_i$.

From the expressions of the likelihood functions G_i and F_{ij} given in (7) and (8), we observe that priors play no role in the agents' limiting beliefs. To fix ideas, consider, e.g., that two agents i and j start with very different priors p_i and p_j about θ but use a common source Φ and receive information from others through a common directed path $\widehat{\gamma}$. Then, the results of Theorem 1 imply that both agents achieve some common limiting beliefs. In this case, their priors will determine the speed of convergence to such limiting beliefs. On the other hand, two agents starting with common priors may end up either with different or common limiting beliefs. In short, Theorem 1 implies that the slope of the trajectory of an agent i's posteriors to his limiting beliefs depends on his priors p_i, on his information source Φ_i, and on all the paths $\widehat{\gamma}_{ij} \in \Gamma_{ij}[\Psi]$, for each $j \in N_i$. Nevertheless, his limiting beliefs do not depend on his priors. Even if agent i receives no information from others and his priors put a large weight on some value θ, he may end up with some limiting beliefs that put probability one to some other value $\theta' \neq \theta$. This will happen if his source Φ_i biases him sufficiently towards θ' in his private learning.

Consider now the case where no agent receives any information whatsoever from any other agent. In particular, this situation is formally obtained if each message protocol Σ_{ij} is such that messages do not depend on observed signals, i.e., $\sigma^s_{ij}(m) = \sigma_{ij}(m)$ for each $s \in S$.[39] If this is the case, then it follows from the expression in (4) that $\pi^s_{\gamma_{ij}}(m) = \pi_{\gamma_{ij}}(m)$ for any directed path

[39] Following the terminology of sender-receiver games, this case corresponds to a *pooling* message protocol Σ_{ij}. Notice that this extreme case can be alternatively obtained if we simply exclude the possibility of network connections.

γ_{ij} so that $\psi_{ij}[\gamma_{ij}](m) = \pi_{\gamma_{ij}}(m)$. As a consequence, $F_{ij}(\theta) = -H(\widehat{\psi}_{ij})$ for each pair of different agents in the society. Since each F_{ij} does not depend on θ in this extreme case, we have $\Theta_i = \Theta_i^*$ for each agent $i \in N$. In this case, using Theorem 1, we obtain the intuitive insight that an agent i's limiting beliefs are governed only by his source Φ_i. We thus naturally interpret each Θ_i as the set of parameter values that agent i favors due only to his private learning, without communication, and each Θ_i^* as the set of values that he favors using both his source and the information that he receives from others. Then, to explore the achievement of a consensus, we would like to study the conditions on the network structure under which some agents are able to influence others' opinions in a way such that all of them end up with the same limiting beliefs. Specifically, we wish to identify the features of the network and its weights which, starting from a reference situation $\Theta_i \neq \Theta_j$, induce $\Theta_i^* = \Theta_j^*$, with the additional requirement that $\Theta_j^* = \Theta_j$.[40]

Definition 7. Given a connected social network Ψ, and two distinct agents $i, j \in N$ such that $\Theta_i \neq \Theta_j$, we say that agent j influences agent i if $\Theta_i^* = \Theta_j^* = \Theta_j$.

In the definition above, for an agent to influence another, we require that he must not be in turn influenced by other agents. Also, from the results of Theorem 1, we note that if the sets Θ_i^* and Θ_j^* satisfying $\Theta_i^* = \Theta_j^*$ are not singleton and agents i and j begin with different priors, then their limiting beliefs are indeed different. In this case, agents i and j do agree on the set of parameter values that have positive probability of occurrence. However, each Θ_i^* is generically singleton because non-singleton sets Θ_i^* are not robust to small perturbations of the network.[41]

Consider now the extreme case in which all message protocols are completely separating so that all agents reveal with full precision their signals to their neighbors. Suppose, without loss of generality, that each message protocol Σ_{ij} is specified as $\sigma_{ij}^{s_l}(m_l) = 1$ for each $l \in \{1, \dots, L\}$. Then, by using the expression in (4), we obtain that, for any directed path γ_{ij} in the network, $\pi_{\gamma_{ij}}^{s_l}(m_l) = 1$ for each $l \in \{1, \dots, L\}$. As a consequence, $F_{ij} = G_j$ for each pair of agents $i, j \in N$ such that i has a directed path to j. Then, using the result in Theorem 1, we see that an agent i's limiting beliefs favor the set of parameters $\arg\max_{\theta \in \Theta} \left\{ G_i(\theta) + \sum_{j \in N_i} G_j(\theta) \right\}$. In short, in this case where information flows through links without any decay, how agent j influences agent i through any directed path γ_{ij} in the network depends only on the functions G_i and G_j or, in other words, on the qualities of the sources Φ_i and Φ_j. Notice, though, that the network architecture continues to be crucial to determine how a particular agent might be influenced by

[40] The interpretation is that the agents j are able to influence agents i's opinions so that all them put positive probability in the long-run to the same parameter values that agents j considered with positive probability based solely on their sources.

[41] The set of networks for which some Θ_i^* is not singleton has Lebesgue measure zero in the set of all possible networks.

others.

The following example illustrates (a) how limiting beliefs are obtained for the case without communication and (b) how an agent can influence the evolution of others' beliefs in a way such that a consensus is finally achieved in the society.

Example 1. Consider a set of $n = 4$ agents who care about two possible parameter values, i.e., $\Theta = \{\theta_1, \theta_2\}$. The agents are connected through a social network $\Psi = \{\Psi_{13}, \Psi_{21}, \Psi_{24}, \Psi_{32}, \Psi_{43}\}$, which is depicted in Figure 1. The agents begin with the (common) priors $p(\theta_1) = p(\theta_2) = 1/2$, so that $H(p) = -\log(1/2)$. The agents' private signals are specified as: (agent 1) $\phi_1^{\theta_1}(s_1) = 1/6$ and $\phi_1^{\theta_2}(s_1) = 1/2$; (agent 2) $\phi_2^{\theta_1}(s_1) = 1/3$ and $\phi_2^{\theta_2}(s_1) = 2/3$; (agent 3) $\phi_3^{\theta_1}(s_1) = 2/5$ and $\phi_3^{\theta_2}(s_1) = 9/10$; (agent 4) $\phi_4^{\theta_1}(s_1) = 2/3$ and $\phi_4^{\theta_2}(s_1) = 1/3$. With this information, we can compute: (agent 1) $G_1(\theta_1) = -0.7188$ and $G_1(\theta_2) = -0.6931$ so that $\Theta_1 = \{\theta_2\}$; (agent 2) $G_2(\theta_1) = -0.752$ and $G_2(\theta_2) = -0.752$ so that $\Theta_2 = \{\theta_1, \theta_2\}$; (agent 3) $G_3(\theta_1) = -0.7744$ and $G_3(\theta_2) = -0.8744$ so that $\Theta_3 = \{\theta_1\}$; (agent 4) $G_4(\theta_1) = -0.752$ and $G_4(\theta_2) = -0.752$ so that $\Theta_4 = \{\theta_1, \theta_2\}$. Therefore, when the agents use only their sources, there is some discrepancy in their limiting beliefs. In particular, agents 2 and 4 end up with their initial priors, agent 1 favors the parameter value θ_2, and agent 3 favors the parameter value θ_1.

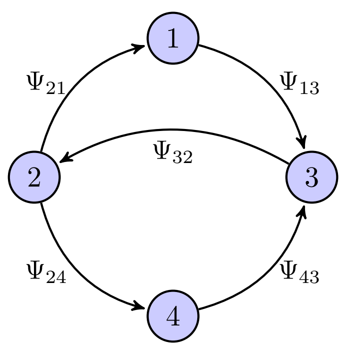

FIGURE 1

To describe the links of the directed network, we specify the corresponding message protocols as: (link 13) $\sigma_{13}^{s_1}(m_1) = 1$ and $\sigma_{13}^{s_2}(m_1) = 0$; (link 21) $\sigma_{21}^{s_1}(m_1) = 4/5$ and $\sigma_{21}^{s_2}(m_1) = 1/5$; (link 24) $\sigma_{24}^{s_1}(m_1) = 1/4$ and $\sigma_{24}^{s_2}(m_1) = 0$; (link 32) $\sigma_{32}^{s_1}(m_1) = 1/3$ and $\sigma_{32}^{s_2}(m_1) = 2/3$; (link 43) $\sigma_{43}^{s_1}(m_1) = 9/10$ and $\sigma_{43}^{s_2}(m_1) = 1/10$. With this information, we can obtain the associated distributions $\widehat{\psi}_{ij}^{\theta} \in \Delta(\{m_1, m_2\})$, for $\theta \in \{\theta_1, \theta_2\}$. Observe that the network in Figure 1 is connected so that each agent can listen to the opinions of each other agent through some directed path. Also, some agents are connected through several paths. In particular, agent 2 can listen to agent 3 through the paths $\gamma_{23} = (\Psi_{21}, \Psi_{13})$ and $\gamma'_{23} = (\Psi_{24}, \Psi_{43})$. By computing the power of each path, we pick the paths which transmit the highest amount of information.

For the directed links, we obtain: (link 13) $F_{13}(\theta_1) = -0.7744$ and $F_{13}(\theta_2) = -0.8744$; (link 21) $F_{21}(\theta_1) = -0.6956$ and $F_{21}(\theta_2) = -0.6931$; (link 24) $F_{24}(\theta_1) = -0.3835$ and $F_{24}(\theta_2) = -0.3867$; (link 32) $F_{32}(\theta_1) = -0.6993$ and $F_{32}(\theta_2) = -0.6993$; (link 43) $F_{43}(\theta_1) = -0.7448$ and $F_{43}(\theta_2) = -0.7747$.

For the directed paths which transmit the highest amount of information, we obtain: (path 12) $F_{12}(\theta_1) = -0.6993$ and $F_{12}(\theta_2) = -0.6993$; (path 14) $F_{14}(\theta_1) = -0.662$ and $F_{14}(\theta_2) = -0.662$; (path 23) $F_{23}(\theta_1) = -0.7221$ and $F_{23}(\theta_2) = -0.7299$; (path 31) $F_{31}(\theta_1) = -0.6932$ and $F_{31}(\theta_2) = -0.6931$; (path 34) $F_{34}(\theta_1) = -0.662$ and $F_{34}(\theta_2) = -0.662$; (path 41) $F_{41}(\theta_1) = -0.6931$ and $F_{41}(\theta_2) = -0.6931$; (path 42) $F_{42}(\theta_1) = -0.6971$ and $F_{42}(\theta_2) = -0.6971$.

We observe that, for each agent $i \neq 3$, the value of $F_{i3}(\theta_1)$ is higher than the value of $F_{i3}(\theta_2)$. This indicates that, through communication, agents place a relatively high intensity on the parameter value that agent 3 favors in the case without communication. Then, we analyze whether agent 3 can be influential in this society. This turns out to be the case: by computing the corresponding values of $G_i(\theta) + \sum_{j \neq i} F_{ij}(\theta)$, for each $i = 1, \ldots, 4$ and each $\theta \in \{\theta_1, \theta_2\}$, using the values above, we obtain $\Theta_i^* = \{\theta_1\}$ for each $i = 1, \ldots, 4$. Thus, the society achieves a consensus in which each agent believes in the long-run with probability one that θ_1 is the true parameter value.

A complementary way to study how influential are some agents and the achievement of a consensus would naturally involve to use the power of the paths. We next provide two necessary and sufficient conditions, in terms of the power of the paths of the network, under which an agent is able to influence another. For an agent j to influence another agent i, it would be natural, on the one hand, to require that the informativeness of the (most informative) path from i to j be sufficiently high. On the other hand, it would be also natural to require that the informativeness of the (most informative) path from agent j to any other agent in the society be sufficiently low in order to prevent j from being influenced. This turns out to be the case, and such conditions are stated formally in Theorem 2 below. Other intuitive message of Theorem 2 is that some j is more likely to influence another agent i when agent j's private learning places high intensity on some parameter values. This can be interpreted as agent j being very convinced of his opinion about the true parameter values from his private learning. Such an agent j can be viewed as a "self-confident" agent.

Before stating Theorem 2, we need to introduce an additional entropy-related measure. For two distinct agents $i, j \in N$, let

$$Q(p_i, p_j) := -\sum_{\Theta} p_i(\theta) \log p_i(\theta) \sum_{M} \widehat{\psi}_{ij}^{\theta}(m) \left[\frac{\sum_{\theta'} \widehat{\psi}_{ij}^{\theta'}(m) p_j(\theta')}{\sum_{\theta'} \widehat{\psi}_{ij}^{\theta'}(m) p_i(\theta')} \right].$$

The measure $Q(p_i, p_j)$ describes how far are agent i's priors with respect to the uniform case, compared to the discrepancy of j's priors relative to uniformity. For the particular case where agents i and j share common priors, $p_i = p_j = p$, it can be easily verified that $Q(p,p) = H(p)$. Also, from the expression above, it can be checked that $Q(p_i, p_j)$ increases with $H(p_i)$, for any given priors p_j.

Theorem 2. *Consider a connected social network Ψ and two different agents $i, j \in N$ such that $\Theta_i \neq \Theta_j$. Suppose that Assumptions 2 and 3 hold, then agent j influences agent i if and only if Ψ satisfies the following conditions:*
(i) for agents i and j:

$$\mathbb{P}(\widehat{\gamma}_{ij}) > [G_i(\theta_i) - G_i(\theta_j)] + \max_{\theta \notin \Theta_j} \sum_{h \in N_i} [F_{ih}(\theta) - F_{ih}(\theta_j)] + Q(p_i, p_j) - E_{\widehat{\psi}_{ij}} \left[H(q_i^{m_{ij}^1}[\widehat{\gamma}_{ij}]) \right], \quad (7a)$$

for any $\theta_i \in \Theta_i$ and any $\theta_j \in \Theta_j$.
(ii) for agent j (with respect to the rest of the society):

$$\begin{aligned} &\max_{k \in N_j} \left\{ \mathbb{P}(\widehat{\gamma}_{jk}) + E_{\widehat{\psi}_{jk}} \left[H(q_j^{m_{jk}^1}[\widehat{\gamma}_{jk}]) \right] \right\} \\ &\qquad < G_j(\theta_j) + \sum_{h \in N_j} F_{jh}(\theta_j) - \max_{\theta \notin \Theta_j} \left[G_j(\theta) + \sum_{h \in N_j} F_{jh}(\theta) \right] + Q(p_j, p_k), \end{aligned} \quad (7b)$$

for any $\theta_j \in \Theta_j$.

Condition (i) of Theorem 2 identifies a lower bound on the informativeness levels of the (most informative) directed path from agent i to agent j under which j is able to affect i's beliefs in a way such that i ends up favoring the same parameter values that j favors in the absence of communication. On the other hand, condition (ii) identifies an upper bound on the level of informativeness of the (most informative) directed path from agent j to any other agent in the society, which characterizes the situation where j continues to believe in the long-run that his most-favored parameter values in the absence of communication, Θ_j, are also the most likely ones after listening to others' opinions.

From inequality (7a) above, we observe that the required lower bound on $\mathbb{P}(\widehat{\gamma}_{ij})$ increases with the difference $G_i(\theta_i) - G_i(\theta_j)$. Intuitively, it is easer for us to influence some agent when the intensity that he puts on the parameter values that he considers the most likely ones does not differ much from the intensity that he puts on the values that we consider to be the most likely ones. Inequality (7a) also states that the required lower bound on $\mathbb{P}(\widehat{\gamma}_{ij})$ decreases with each $F_{ih}(\theta_j)$, $h \in N_i$, that is, it is easer for us to influence some agent when the information that he receives through communication place a high intensity on the parameters value that we consider to be the most likely ones. Furthermore, such a lower bound increases with $\max_{\theta \notin \Theta_j} \sum_{h \in N_i} F_{ih}(\theta)$, which can be interpreted as a measure of the highest intensity that the information received by agent i from the rest of the society places on a parameter value other than the ones favored by agent j. Then, we obtain that it is easer for agent j to influence agent i when i's message protocols do not place a large intensity on parameter values different from those in Θ_j. Finally, inequality (7a) also states that the required lower bound on $\mathbb{P}(\widehat{\gamma}_{ij})$ increases with $Q(p_i, p_j)$ and decreases with the average entropy $E_{\widehat{\psi}_{ij}}[H(q_i^{m_{ij}^1}[\widehat{\gamma}_{ij}])]$. Therefore, for agent j to influence agent i, we need lower values of $\mathbb{P}(\widehat{\gamma}_{ij})$ when agent i's priors display little uncertainty ex-ante[42] and/or when agent i's posteriors, based solely on the information that he receives from agent j, have in average high uncertainty.[43]

On the other hand, from inequality (7b), we observe that high values of the informativeness of agent j's path to another agent k in the society are compatible with j not being influenced by k when: (a) j's source and/or the opinions that he receives from k put a high intensity on the parameter values that he considers the most likely ones without communication (i.e., high values of $G_j(\theta_j)$ and/or of $F_{jh}(\theta_j)$), (b) j's source and/or the opinions that he receives from other agents $h \in N_i$ do not place a high intensity on parameter values different from the ones that he favors without communication (i.e., low values of $\max_{\theta \notin \Theta_j}\{G_j(\theta) + \sum_{h \in N_j} F_{jk}(\theta)\}$), (c) j's priors are very uncertain ex-ante (i.e., high values of $Q(p_j, p_k)$),[44] and (d) the ex-ante uncertainty in average of j's posteriors, conditioned on the messages that he receives from k, is low (i.e., low values of $E_{\widehat{\psi}_{ij}}[H(q_i^{m_{ij}^1}[\widehat{\gamma}_{ij}])]$).

Example 2. Consider again the social network analyzed in Example 1. Figure 2 depicts the power of each directed link in the network. The power measures $\mathbb{P}(\Psi_{ij})$ below are computed using the set of primitives described in Example 1. Recall that, in the case without communication, agents 2, 3, and 4 favor parameter value θ_1 but, unlike agent 3, agents 2 and 4 consider that parameter value θ_2 has also some positive probability of occurrence. Based only on their sources, the discrepancy of opinions is relatively higher between agents 1 and 3. Agent 1 favors θ_2 with probability one while agent 3 favors θ_1 with probability one. The result obtained in Example 1 that agent 3 influences the rest of the society so as to achieve a consensus is not surprising now

[42] It can be easily verified that, for any given priors p_j, lower values of $Q(p_i, p_j)$ are associated with agent i's priors which are close to the uniform case $\overline{p}_i(\theta) = 1/L$ for each $\theta \in \Theta$.

[43] Higher values of $E_{\widehat{\psi}_{ij}}[H(q_i^{m_{ij}^1}[\widehat{\gamma}_{ij}])]$ are associated with posteriors which put large probabilities on a few parameter values.

[44] Since $Q(p_j, p_k)$ increases with $H(p_j)$, the message here is that it is easer for an agent to influence others when he begins with priors that put relatively large probability weights on a small number of parameter values. A natural interpretation, compared to priors more close to the uniform case, is that the agent begins with "strong opinions" about which parameter values are most likely. If, in addition, the network does not allow this agent to hear other agents who have strong beliefs in favor of different parameter values, then we would obtain the interpretation that this is a "stubborn agent," hardly influenced by others in the social group.

if we note the weights described in Figure 2.

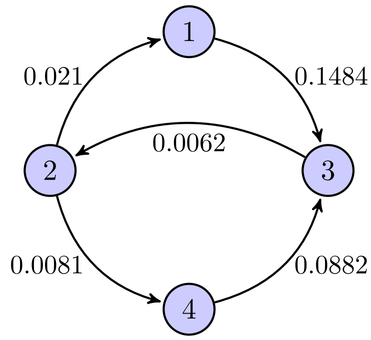

FIGURE 2

We observe that the intensity with which agent 1 listens to agent 3's opinions is the highest in the society (0.1484). Also, agent 3 listens to the others' opinions exclusively through his link with agent 2, and the power of this link is the lowest in the society (0.0062). In addition, we obtain $\mathbb{P}(\widehat{\gamma}_{23}) = 0.0474$ while $\mathbb{P}(\Psi_{21}) = 0.021$. In other words, through agent 1, agent 2 pays more attention to the opinions of agent 3 than to the opinions of agent 1 himself. On the other hand, using the result in Lemma 1, we know that $\mathbb{P}(\widehat{\gamma}_{41}) < 0.021$ so that, given that $\mathbb{P}(\Psi_{43}) = 0.0882$, we observe that agent 4 also pays more attention to the opinions of agent 3 than to the opinions of agent 1. In short, using the power measure, we observe that agent 3 is a good candidate to influence the opinions of the rest of the society. Clearly, he is the agent in the best position, according to the architecture of the directed links and to their weights, to do so.

Then, we examine whether condition (7a) of Theorem 2 holds for for the directed link Ψ_{13}. Recall from Example 1 that $H(p) = -\log(1/2) = 0.6931$ Also, from the computations of Example 1, we observe that

$$G_1(\theta_2) - G_1(\theta_1) = 0.0257$$

and

$$F_{12}(\theta_2) - F_{12}(\theta_1) = 0, \quad F_{13}(\theta_2) - F_{13}(\theta_1) = -0.1, \quad \text{and} \quad F_{14}(\theta_2) - F_{14}(\theta_1) = 0.$$

Thus, to verify whether whether condition (7a) holds for the directed link Ψ_{13}, we only need to compute the expected entropy $E_{\psi_{13}}[H(q_{13}^{m})]$. Using the message protocols specified in Example 1, we compute: $q_{13}^{m_1}(\theta_1) = 4/13$ and $q_{13}^{m_2}(\theta_1) = 6/7$. With these posteriors, we easily obtain $E_{\psi_{13}}[H(q_{13}^{m})] = 0.5447$. Then, according to condition (7a), for agent 3 to influence agent 1, we need that the intensity of the directed link Ψ_{13} be above the bound

$$0.0257 - 0.1 + 0.6931 - 0.5447 = 0.0471.$$

This intensity is clearly exceeded in our example since we have $\mathbb{P}(\Psi_{13}) = 0.1484$.

Now, we turn to examine whether condition (7b) holds for agent 3 so that he is not influenced by any of the three other agents. First, from the computations of Example 1, we observe that

$$G_3(\theta_1) + \sum_{k\neq 3} F_{3k}(\theta_1) - \Big[G_3(\theta_2) + \sum_{k\neq 3} F_{3k}(\theta_2)\Big] + H(p) = 0.7932. \tag{9}$$

Second, we can easily compute

$$E_{\widehat{\psi}_{31}}[H(q^m_{31}[\widehat{\gamma}_{31}])] = 0.6909, \quad E_{\psi_{32}}[H(q^m_{32})] = 0.6869, \quad \text{and} \quad E_{\widehat{\psi}_{34}}[H(q^m_{34}[\widehat{\gamma}_{34}])] = 0.6002.$$

We, therefore, obtain

$$\mathbb{P}(\Psi_{32}) + E_{\psi_{32}}[H(q^m_{32})] = 0.0062 + 0.6869 = 0.6932.$$

Furthermore, by using the result in Lemma 1, we know that

$$\mathbb{P}(\widehat{\gamma}_{31}) + E_{\widehat{\psi}_{31}}[H(q^m_{31}[\widehat{\gamma}_{31}])] < 0.021 + 0.6909 = 0.7119, \quad \text{and}$$
$$\mathbb{P}(\widehat{\gamma}_{34}) + E_{\widehat{\psi}_{34}}[H(q^m_{34}[\widehat{\gamma}_{34}])] < 0.0081 + 0.6002 = 0.6083.$$

Since $\max_{k\neq 3}\Big\{\mathbb{P}(\widehat{\gamma}_{3k}) + E_{\widehat{\psi}_{3k}}[H(q^m_{3k}[\widehat{\gamma}_{3k}])]\Big\}$ is less than 0.7119, which exceeds not the required value 0.7932, identified in (9) above, we obtain that condition (7b) holds for agent 3.

In this example, one can analogously analyze the conditions in Theorem 2 for the most informative paths $\widehat{\gamma}_{23}$ and $\widehat{\gamma}_{43}$ to conclude that these conditions are satisfied in a way such that agent 3 influences agents 2 and 4 as well, and a consensus is achieved.

An obvious implication of Theorem 2 is the following sufficient condition that guarantees the achievement of a consensus in the society.

Corollary 1. *Consider a connected social network Ψ. A consensus is attained in the society if there exists a set of agents $\overline{N} \subset N$ such that (i) for each $j \in \overline{N}$, we have $\Theta_j = \{\theta^*\}$ for some $\theta^* \in \Theta$, and (ii) for each $i \in N \setminus \overline{N}$ there is some agent $j \in \overline{N}$ such that agent j influences agent i.*

Following the related literature, if a set of agents $\overline{N}$ as the one described in Corollary 1 exists, then we refer to it as a set of *prominent agents*.

Some insights about the role of network centrality on consensus emerge straightforwardly by combining our results on the weights of paths required for an agent to be able to influence others, Theorem 2, with our previous results on decay provided by Lemma 1. The message that more central agents can be especially influential is well-established in the literature on influence in networks. Our model delivers this idea too as it provides precise conditions, in terms of our entropy-based measures, for an agent to critically influence others. Since the informativeness of messages decreases as they pass from one agent to another, these conditions depend crucially on how central such an agent is.

More specifically, for an agent to be a prominent agent in our model, it helps that he enjoys high centrality following the sort of criteria considered by the measures of *closeness centrality* or *information centrality*.[45] Under a number of different specifications, this class of centrality measures tracks how close a given agent is to any other agent. Within this class, a particular subset of centrality measures that appears naturally well suited to our model is that of *decay centrality* measures. The general idea here is to measure the proximity between a given agent and any other agent weighted by the decay of the path that connects them. The particular notion of decay would depend on the context analyzed. In fact, our model provides some interesting entropy-based tools that allow us to readily formalize this idea in our benchmark. In addition, since networks are directed in this benchmark, we wish that our decay centrality measure also captures the idea that an agent is central when he pays little attention to the other agents. Specifically, let

$$C_j(\Psi) = \sum_{\{i \in N \,:\, \widehat{\gamma}_{ij} \in \Gamma_{ij}[\Psi]\}} \mathbb{P}(\widehat{\gamma}_{ij}) \quad - \sum_{\left\{h \in N \,:\, \widehat{\gamma}_{jh} \in \Gamma_{jh}[\Psi]\right\}} \mathbb{P}(\widehat{\gamma}_{jh}) \tag{10}$$

be the *decay centrality* measure of agent j in the social network Ψ. The measure $C_j(\Psi)$ computes, in terms of the quality of the information that flows through the (most informative) directed paths that connects them, how close is any agent i to agent j. Note that this centrality measure depends crucially on the decay along each directed path to agent j since each $\mathbb{P}(\widehat{\gamma}_{ij})$ is negatively related to the decay accumulated along $\widehat{\gamma}_{ij}$. In addition, $C_j(\Psi)$ substracts the distance, in terms of the power of the corresponding link, of agent j to any other agent h.[46]

Intuitively, using the decay centrality measure $C_j(\Psi)$ specified in (10) above, an agent j enjoys high centrality if he is accessed by many agents through relatively short paths with very informative message protocols and, at the same time, he listens to others only through long paths with low informative message protocols. It can be easily verified that, for any given configuration of sources and message protocols, $C_j(\Psi)$ is maximized when agent j is the central agent in a *center-directed star network*. Without loss of generality, a *center-directed star network* has the form $\Psi^s = \{\Psi_{j1} : j \in N \setminus \{1\}\}$, where agent 1 has been chosen to have the central position. In Proposition 2, we consider a center-directed star network and study the conditions under which the central agent can become prominent and, at the same time, preclude the attainment

[45] The information centrality measure was introduced by Stephenson and Zelen (1989) with the motivation that information flows through a social network. This measure is specified as the harmonic average of the distance between a given agent and any other agent.

[46] A typical formulation of a decay centrality measure for non-directed social networks would only consider, in our benchmark, the positive component $\sum_{\left\{i \in N \,:\, \widehat{\gamma}_{ij} \in \Gamma_{ij}[\Psi]\right\}} \mathbb{P}(\widehat{\gamma}_{ij})$.

of correct limiting beliefs.

In organizations or social groups, the natural interpretation of a prominent central agent would then be that of some agent j with good communication skills for general audiences (so that $\mathbb{P}(\widehat{\gamma}_{ij})$ be relatively high for as many agents i as possible) and with a position that allows him to speak directly to as many other agents as possible (so that decay does not affect crucially the informativeness level of the information that he is transmitting). In other words, other agents in the organization should be able to hear this agent's opinions through "good quality channels," and as directly as possible, without the need of many intermediaries to receive his messages. Also, it helps for an agent to be prominent that he has access to others' opinions mainly through "bad quality channels" and/or through long paths of directed links (so that $\mathbb{P}(\widehat{\gamma}_{jh})$ be relatively low for as many agents h as possible).

4.3 Correct Limiting Beliefs and Influence of Prominent Agents

Suppose that a consensus is achieved because a group of prominent agents is able, through the paths of the network, to influence any other agent in the society. However, it could well be the case that the limiting beliefs of these prominent agents differ from the limiting beliefs associated with the aggregation of all the private sources of information in the society. In this case, the intuition is that the prominent agents are able to perform a "large-scale manipulation" by using the biases attached to their sources, their positions in the network, and/or the weights of the links that connects them to other agents. The next proposition provides a sufficient condition on the informativeness levels of the links of the network under which correct limiting beliefs are attained in the society.

Proposition 1. *Consider a connected social network Ψ and suppose that a consensus is achieved in the society in a way such that, for some $\theta^* \in \Theta$, we have $\Theta_i^* = \{\theta^*\}$ for each $i \in N$. If the following condition*

$$\sum_{i \in N} \sum_{j \in N_i} \left[F_{ij}(\theta^*) - F_{ij}(\theta) \right] < 0,$$

is satisfied for each $\theta \in \Theta \setminus \{\theta^*\}$, then the social network Ψ attains correct limiting beliefs.

The sufficient condition identified in Proposition 1 is intuitive. Suppose that the aggregation of the pieces of information obtained from the sources of all the agents leads one to believe in the long-run that a given parameter value θ^* is the true one. Then, the condition above imposes some restrictions on the influence of prominent agents. It requires that there is no agent whose influence on others be such that some agents' limiting posteriors favor alternative parameter values $\theta \in \Theta \setminus \{\theta^*\}$. Thus, we obtain that the attainment of correct beliefs is facilitated if the influence of the prominent agents is not too high.

The message conveyed by the sufficient condition of Proposition 1 above is reminiscent of the main results obtained by Golub and Jackson (2010) in their work without Bayesian updating (Propositions 2 and 3). Although they use a notion of correct beliefs that differs slightly from ours,[47] correctness of beliefs requires in their model that the influence of prominent agents vanish as the size of the society grows. In our setting, as well as in theirs, a disproportionate popularity by some agent(s) could turn into an obstacle to achieve correct limiting beliefs.

However, the fact that the condition stated in Proposition 1 is only sufficient is illustrated in the following example. In Example 3 below, the influence of a prominent agent is relatively high, which leads to a consensus, and yet correct limiting beliefs are obtained.

Example 3. Consider again the social network described in Example 1. Recall that this society achieves a consensus in which all the agents' beliefs converge to a distribution that places probability one on the parameter value θ_1. This consensus was propitiated by the fact that agent 3 was able to influence the rest of agents in the society. Using the computations of the functions F_{ij} provided in Example 1, it is easy to verify that

$$\begin{aligned}\sum_{i=1}^{4}\sum_{j\neq i}\left[F_{ij}(\theta_1)-F_{ij}(\theta_2)\right] &= [-0.7744+0.8744]+[-0.6956+0.6931] \\ &\quad +[-0.7221+0.7299]+[-0.3835+0.3867]+[-0.6932+0.6931] \\ &= 0.1305>0,\end{aligned}$$

so that the sufficient condition of Proposition 1 is not satisfied. Nevertheless, using Theorem 1, we can still check directly whether the consensus beliefs coincide with the limiting beliefs of the external observer. Note that, from the result in Theorem 1, the parameter values that are favored by the external observer are those in the set $\arg\max_{\Theta}\sum_{i\in N}G_i(\theta)$. Then, using the computations of the functions G_i provided in Example 1, we obtain $\sum_{i=1}^{4}G_i(\theta_1)=-2.9972$ and $\sum_{i=1}^{4}G_i(\theta_1)=-3.0719$ so that, for our social network, we have $\lim_{t\to\infty}q_{\text{ob}}^{h^t}(\theta_1)=1$. Thus, although the sufficient condition in Proposition 1 is not satisfied, the influence of agent 3 does not interfere with the limiting beliefs that are obtained by aggregating the external sources, and correct limiting beliefs are attained in this social network.

Under which conditions will then prominent agents be able to manipulate beliefs and propagate misinformation in the society? To provide some answers to this question, we next analyze a particular network structure in which, compared to any other possible network architecture, a

[47]Their definition of belief correctness also requires that some external observer aggregates the pieces of information initially held by the agents.

given agent is in the best position to influence anyone else. By combining the results of Lemma 1 and of Theorem 2, it follows that this reference network is the *center-directed star network* $\Psi^s = \{\Psi_{j1} : j \in N \setminus \{1\}\}$, where (without loss of generality) agent 1 has the central position. In this network the decay centrality measure $C_1(\Psi)$ proposed in (10) is maximized for any given configuration of sources and message protocols. Notice that, while the central agent is listened by each other agent through a single directed link, he does not pay attention to anyone. In addition, any peripheral agent only receives messages from the central agent. Provided that a consensus is attained, the following proposition shows that, if the size of the society is sufficiently large, then the central agent manipulates anyone else in a way that precludes the correct aggregation of disperse information.

Proposition 2. *Consider the center-directed star network $\Psi^s = \{\Psi_{j1} : j \in N \setminus \{1\}\}$. Suppose that Assumptions 1 and 3 hold and that a consensus is achieved in a way such that agent* 1 *influences any other agent's limiting beliefs to put probability one on some parameter value $\theta^* \in \Theta$. Then, there is some finite size of the society $n^* \in \{1, 2, \dots\}$ such that correct limiting beliefs are precluded for each $n \geq n^*$.*

The following example illustrates the result of Proposition 2 above by using some of the entropy-based measures proposed in this paper.

Example 4. Suppose that $\Theta = \{\theta_1, \theta_2\}$ and consider a set of $n = 5$ agents who are connected through the center-directed star network $\Psi^s = \{\Psi_{21}, \Psi_{31}, \Psi_{42}, \Psi_{51}\}$.

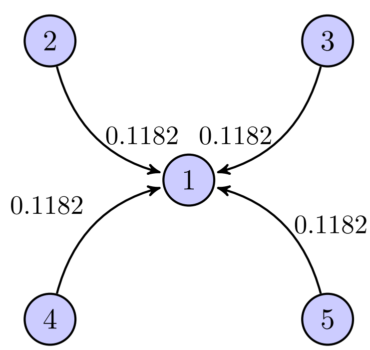

FIGURE 3

The agents begin with the (common) priors $p(\theta_1) = p(\theta_2) = 1/2$ and their private signals are specified as: (agent 1) $\phi_1^{\theta_1}(s_1) = 1/2$ and $\phi_1^{\theta_2}(s_1) = 1/12$; (each agent $j \neq 1$) $\phi^{\theta_1}(s_1) = 1/6$ and $\phi^{\theta_2}(s_1) = 1/2$. Then, it can be verified that $G_1(\theta_1) - G_1(\theta_2) = 0.0932 > 0$ and $G(\theta_2) - G_1(\theta_1) = 0.0256 > 0$ so that, by using the results of Theorem 1, we know that, without communication among the agents, the central agent would favor parameter value θ_1 while any of the peripheral agents would favor parameter value θ_2. Now, suppose that the communication from the central to each peripheral agent is described by a common message protocol σ, specified as $\sigma^{s_l}(m_l) = 1$ for each $l \in \{1, 2\}$. Note that this message protocol is completely separating so that there is no decay in the transmission of information from the center of the star to any agent in the periphery. The power of the directed links are depicted in Figure 3.

In this case, we obtain that $F_{j1}(\theta) = G_1(\theta)$ for each $j \neq 1$ and each $\theta \in \Theta$. Therefore, we have

$$F_{j1}(\theta_1) + G(\theta_1) - [F_{j1}(\theta_2) + G(\theta_2)] = 0.0932 - 0.0256 > 0,$$

so that each peripheral agent is influenced by agent 1, and their limiting beliefs put probability one on θ_1 being the true parameter value. However, by aggregating the information sources available to all agents, we obtain

$$G_1(\theta_2) + 4 \cdot G(\theta_2) - [G_1(\theta_1) + 4 \cdot G(\theta_1)] = -0.0932 + 4 \cdot 0.0256 = 0.0092 > 0,$$

so that correct limiting beliefs put probability one on θ_2 being the truth. In this example, the power of the source of each peripheral agent $j \neq 1$ is $\mathbb{P}(\Phi) = 0.0647$. Not surprisingly, since $\mathbb{P}(\Psi_{j1}) = \mathbb{P}(\Phi_1) = 0.1182$ for each $j \neq 1$, we observe that the quality of the communication from the central agent to any peripheral agent is higher than the quality of the information that the peripheral agents receive from their sources.

Since limiting beliefs are additively determined using the influences that the agents receive from their sources and from communication with others, this model easily accommodates phenomena of diffusion of "extreme opinions."[48] Given this feature, an interesting insight that arises from Proposition 2 is that a central and prominent agent is more likely to induce incorrect consensus beliefs as the size of the society grows. In other words, if the available communication technology is sufficiently good so that the quality of the information from central-prominent agents to other peripheral agents does not diminish as the society grows, then manipulation follows more easily for large networked societies.[49]

5 Concluding Comments

Using Bayesian updating rules, this paper has developed a mathematically tractable model of beliefs evolution under opinion influence to which non-Bayesian models can be compared. The

[48] This feature contrast sharply with the insights of the DeGroot's model, wherein extreme opinions are smoothed down as time evolves.

[49] Although Proposition 2 provides formally this insight only for the specific case of a star network, given the logic behind the result, the message that it conveys is robust under more general network structures in which prominent agents enjoy positions with high centrality.

assumptions on the agents' informational capabilities, the focus on first-order beliefs, and the notion of belief correctness are more appealing when one considers societies large enough. For small societies, the use of a belief correctness notion based on conditioning posteriors on a given parameter value, together with allowing agents to compute higher-order beliefs, would deliver the message that they always learn the truth because both signals and received messages are independent over time in our model. This implication would follow rather directly from the main result of Cripps, Ely, Mailath, and Samuelson (2008). Nevertheless, for the approach usually pursued in the learning literature with higher-order beliefs, recent research (e.g., Parikh and Krasucki, 1990; Heifetz, 1996; Koessler, 2001; Steiner and Stewart, 2011) shows that the presence of communication among the agents may in some cases preclude common learning of the parameter. In particular, Cripps, Ely, Mailath, and Samuelson (2013) show that common learning is precluded when the messages that the agents receive are correlated across time. Analyzing consensus and the evolution of correct higher-order beliefs for small societies when messages follow time dependence patterns remains an interesting open question.

Finally, an obvious interesting extension of the model would be that of endogenizing the listening behavior. To follow this approach, more structure should be added to the model so as to consider that the agents pursue the maximization of a payoff that depends on the unknown parameter. Then, by characterizing listening structures that are "stable," one could obtain some insights into the formation of communication networks in a dynamic framework of belief evolution.

References

[1] Acemoglu, D., Chernozhukov, V., and M. Yildiz (2009): "Fragility of Asymptotic Agreement under Bayesian Learning," mimeo.

[2] Acemoglu, D., Como, D., Fagnani, D., and A. Ozdaglar (2013): "Opinion Fluctuations and Disagreement in Social Networks," *Mathematics of Operations Research*, 38(1), 1-27.

[3] Acemoglu, D., Dahleh, M. A., Lobel, I., and A. Ozdaglar (2011): "Bayesian Learning in Social Networks," *Review of Economic Studies*, 78(4), 1201-1236.

[4] Acemoglu, D. and A. Ozdaglar (2011): "Opinion Dynamics and Learning in Social Networks," *Dynamic Games and Applications*, 1(1), 3-49.

[5] Acemoglu, D., Ozdaglar, A., and A. ParandehGheibi (2010): "Spread of (Mis)information in Social Networks," *Games and Economic Behavior*, 70, 194-227.

[6] Azomahou, T. T., and D. C. Opolot (2014): "Beliefs Dynamics in Communication Networks," UNU-MERIT WP.

[7] Bala, V. and S. Goyal (2000): "A Non-Cooperative Model of Network Formation," *Econometrica*, 68(5), 1181-1229.

[8] Banerjee, A., Chandrasekhar, A. G., Duflo, E., and M. O. Jackson (2013): "The Diffusion of Microfinance," *Science*, July, 341.

[9] Banerjee, A., Chandrasekhar, A. G., Duflo, E., and M. O. Jackson (2014): "Gossip: Identifying Central Individuals in a Social Network," mimeo.

[10] Billingsley, P. (1995): *Probability and Measure* (Third Ed.), Wiley & Sons.

[11] Blackwell, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265-272.

[12] Cabrales, A., Gossner, O., and R. Serrano (2013): "Entropy and the Value of Information for Investors," *American Economic Review*, 103, 360-377.

[13] Campbell, A. (2014): "Signaling in Social Network and Social Capital Formation," *Economic Theory*, 57, 303-337.

[14] Condorcet, N. C. de (1785): "Essai sur l'Application de l'Analyse a la Probabilite des Decisions Rendues a la Pluralite des Voix," *Imprimerie Royale*, Paris, France.

[15] Crawford, V. P. and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6) 1431-1451.

[16] Cripps, M. W., Ely, J. C., Mailath, G. J., and L. Samuelson (2008): "Common Learning," *Econometrica*, 76(4) 909-933.

[17] Cripps, M. W., Ely, J. C., Mailath, G. J., and L. Samuelson (2013): "Common Learning with Intertemporal Dependence," *International Journal of Game Theory*, 42(1), 55-98.

[18] DeGroot, M. H. (1974): "Reaching a Consensus," *Journal of the American Statistical Association*, 69, 345, 118-121.

[19] DeMarzo, P., Vayanos, D., and J. Zwiebel (2003): "Persuasion Bias, Social Influence, and Unidimensional Opinions," *Quarterly Journal of Economics*, 118(3), 909-968.

[20] Doob, J. L., (1949): "Application of the Theory of Martingales," in *Le Calcul des Probabilitiés et ses Applications. Colloques Internationaux du Centre National de le Recherche Scientifique,* 13: 23-27. Centre National de la Recherche Scientifique, Paris.

[21] Golub, B. and M. O. Jackson (2010): "Naïve Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, 2(1), 112-149.

[22] Heifetz A. (1996): "Comment on Consensus without Common Knowledge," *Journal of Economic Theory*, 70, 273-277.

[23] Jackson, M. O., Rogríguez-Barraquer, T., and X. Tan (2012): "Social Capital and Social Quilts: Network Pattern of Favor Exchange," *American Economic Review*, 102(5), 1857-1897.

[24] Jackson, M. O. and A. Wolinsky (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44-74.

[25] Koessler, F. (2001): "Common Knowledge and Consensus with Noisy Communication," *Mathematical Social Sciences*, 42, 139-159.

[26] Parikh, R., and P. Krasucki (1990): "Communication, Consensus, and Knowledge," *Journal of Economic Theory*, 52(1), 178-189.

[27] Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge'," *American Economic Review*, 79(3), 385-391.

[28] Savage, L. J. (1954): *The Foundations of Statistics*, Dover reprint, New York, 1972.

[29] Sciubba, E. (2005): "Asymmetric Information and Survival in Financial Markets," *Economic Theory*, 25, 353-379.

[30] Shannon, C. E. (1948): "A Mathematical Theory of Communication," *The Bell System Technical Journal*, 27, 379-423.

[31] Steiner, J. and C. Stewart (2011): "Communication, Timing, and Common Learning," *Journal of Economic Theory*, 146(1), 230-247.

[32] Stephenson, K. and M. Zelen (1989): "Rethinking Centrality: Methods and Examples," *Social Networks*, 11, 1-37.

Appendix

Proof of Lemma 1. (a) Consider a social network Ψ. Take two different agents $i, j \in N$ and a directed link $\Psi_{ij} \in \Psi$ from agent i to agent j. Suppose that the agents i and j to begin with some common priors p. Using the definition of power of a directed link in (3), we have

$$\begin{aligned}
\mathbb{P}(\Psi_{ij}) &= \sum_M \psi_{ij}(m) D\big(q_i^m \,||\, p\big) \\
&= \sum_M \psi_{ij}(m) \sum_\Theta q_i^m(\theta) \log \frac{q_i^m(\theta)}{p(\theta)} \\
&= \sum_M \psi_{ij}(m) \sum_\Theta \frac{\psi_{ij}^\theta(m) p(\theta)}{\psi_{ij}(m)} \log \frac{\psi_{ij}^\theta(m)}{\psi_{ij}(m)} \\
&= \sum_\Theta \sum_M p(\theta) \sum_S \sigma_{ij}^s(m) \phi_j^\theta(s) \log \frac{\sum_S \sigma_{ij}^{s'}(m) \phi_j^\theta(s')}{\sum_S \sigma_{ij}^{s'}(m) \phi_j(s')}.
\end{aligned} \tag{11}$$

Now, by applying, for each given $\theta \in \Theta$ and each given $m \in M$, the log-sum inequality to the expression in (11) above, we obtain

$$\mathbb{P}(\Psi_{ij}) \leq \sum_\Theta \sum_M p(\theta) \sum_S \sigma_{ij}^s(m) \phi_j^\theta(s) \log \frac{\phi_j^\theta(s)}{\phi_j(s)}. \tag{12}$$

On the other hand, using the definition of power of a source in (2), we have

$$\begin{aligned}
\mathbb{P}(\Phi_j) &= \sum_S \phi_j(s) D\big(q_j^s \,||\, p\big) \\
&= \sum_S \phi_j(s) \sum_\Theta q_j^s(\theta) \log \frac{q_j^s(\theta)}{p(\theta)} \\
&= \sum_S \phi_j(s) \sum_\Theta \frac{\phi_j^\theta(s) p(\theta)}{\phi_j(s)} \log \frac{\phi_j^\theta(s)}{\phi_j(s)} \\
&= \sum_\Theta \sum_S p(\theta) \phi_j^\theta(s) \log \frac{\phi_j^\theta(s)}{\phi_j(s)}.
\end{aligned} \tag{13}$$

By combining the inequality in (12) with the expression in (13) above, we obtain

$$\begin{aligned}
\mathbb{P}(\Psi_{ij}) &\leq \sum_\Theta \sum_M p(\theta) \sum_S \sigma_{ij}^s(m) \phi_j^\theta(s) \log \frac{\phi_j^\theta(s)}{\phi_j(s)} \\
&= \sum_\Theta \sum_S p(\theta) \phi_j^\theta(s) \log \frac{\phi_j^\theta(s)}{\phi_j(s)} \Big[\sum_M \sigma_{ij}^s(m) \Big] \\
&= \sum_\Theta \sum_S p(\theta) \phi_j^\theta(s) \log \frac{\phi_j^\theta(s)}{\phi_j(s)} = \mathbb{P}(\Phi_j),
\end{aligned}$$

as stated.

Moreover, by combining the expressions in equations (11) and (13), we obtain

$$\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j) + R(\Sigma_{ij}), \tag{14}$$

where

$$R(\Sigma_{ij}) := \sum_{\Theta} p(\theta) \sum_{M} \sum_{S} \sigma_{ij}^{s}(m) \phi_j^{\theta}(s) \log \frac{\phi_j(s) \sum_S \sigma_{ij}^{s'}(m) \phi_j^{\theta}(s')}{\phi_j^{\theta}(s) \sum_S \sigma_{ij}^{s'}(m) \phi_j(s')}.$$

Since $\sum_S \sigma_{ij}^{s'}(m)\phi_j^{\theta}(s')$ gives us the probability of agent i receiving message m from agent j conditional on the parameter value being θ while $\sum_S \sigma_{ij}^{s'}(m)\phi_j(s')$ gives the corresponding unconditional probability, it follows that $R(\Sigma_{ij}) \leq 0$ for any message protocol $R(\Sigma_{ij})$. Now, note that the message protocol Σ_{ij}, associated with the directed link Ψ_{ij}, allows agent i to learn fully the signal that agent j observes if and only if Σ_{ij} *completely separates* all the signal realizations $s \in S$ available to agent j. Without loss of generality, Σ_{ij} completely separates all the signal realizations in S if and only if $\sigma_{ij}^{s_l}(m_l) = 1$ for each $l \in \{1, \dots, L\}$. In this case, for each $\theta \in \Theta$, we obtain

$$\sum_{M} \sum_{S} \sigma_{ij}^{s}(m) \phi_j^{\theta}(s) \log \frac{\phi_j(s) \sum_S \sigma_{ij}^{s'}(m) \phi_j^{\theta}(s')}{\phi_j^{\theta}(s) \sum_S \sigma_{ij}^{s'}(m) \phi_j(s')} = \sum_{l=1}^{L} \phi_j^{\theta}(s_l) \log \frac{\phi_j(s_l) \phi_j^{\theta}(s_l)}{\phi_j^{\theta}(s_l) \phi_j(s_l)} = 0$$
$$\Leftrightarrow R(\Sigma_{ij}) = 0.$$

Therefore, from the expression in (14), we obtain that the message protocol Σ_{ij} allows agent i to fully learn about the signal observed by agent j if and only if $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j)$.

(b) The proof of part (b) uses exactly the same arguments given above for part (a). The only difference is that the role played in (a) by the source Φ_j is now played by the directed link Ψ_{kj}. All the formal expressions required would replicate the previous ones used in (a) upon adaptation to the appropriate formulae. Therefore, we forego a formal statement. ∎

Proof of Theorem 1. Consider a given social network Ψ and take an agent $i \in N$. For a history $h_i^t \in H_i$, let $\alpha(s; h_i^t)$ be the number of periods in which agent i has observed signal s up to period t and let $\beta_j(m; h_i^t)$ be the number of periods in which agent i has received message m from agent j (through the directed path which transmits the highest amount of information from j to i) up to period t. Consider a history $h_i^t \in H_i$ and a given $\theta \in \Theta$. From the expression derived in (6), we obtain

$$q_i^{h_i^t}(\theta) = \left[1 + \sum_{\theta' \neq \theta} \frac{p_i(\theta')}{p_i(\theta)} \left(\prod_{s \in S} \left(\frac{\phi_i^{\theta'}(s)}{\phi_i^{\theta}(s)} \right)^{\alpha(s;h_i^t)/t} \prod_{j \in N_i} \prod_{m \in M} \left(\frac{\widehat{\psi}_{ij}^{\theta'}(m)}{\widehat{\psi}_{ij}^{\theta}(m)} \right)^{\beta_j(m;h_i^t)/t} \right)^{t} \right]^{-1}.$$

Since observed frequencies approximate distributions, i.e., $\lim_{t\to\infty} \alpha(s; h_i^t)/t = \phi_i(s)$ and $\lim_{t\to\infty} \beta_j(m; h_i^t)/t = \widehat{\psi}_{ij}(m)$, we have

$$\lim_{t\to\infty} q_i^{h_i^t}(\theta) = \left[1 + \sum_{\theta'\neq\theta} \frac{p_i(\theta')}{p_i(\theta)} \lim_{t\to\infty} \left(\prod_{s\in S} \left(\frac{\phi_i^{\theta'}(s)}{\phi_i^{\theta}(s)}\right)^{\phi_i(s)} \prod_{j\in N_i} \prod_{m\in M} \left(\frac{\widehat{\psi}_{ij}^{\theta'}(m)}{\widehat{\psi}_{ij}^{\theta}(m)}\right)^{\widehat{\psi}_{ij}(m)} \right)^t \right]^{-1}.$$

Therefore, studying the converge of $q_i^{h_i^t}(\theta)$ reduces to studying whether each term, for $\theta' \neq \theta$,

$$\prod_{s\in S} \left(\frac{\phi_i^{\theta'}(s)}{\phi_i^{\theta}(s)}\right)^{\phi_i(s)} \prod_{j\in N_i} \prod_{m\in M} \left(\frac{\widehat{\psi}_{ij}^{\theta'}(m)}{\widehat{\psi}_{ij}^{\theta}(m)}\right)^{\widehat{\psi}_{ij}(m)}$$

exceeds or not one. By taking logs, this is equivalent to studying whether, for each $\theta' \neq \theta$, the expression

$$\sum_{s\in S} \phi_i(s) \log \frac{\phi_i^{\theta'}(s)}{\phi_i^{\theta}(s)} + \sum_{j\in N_i} \sum_{m\in M} \widehat{\psi}_{ij}(m) \log \frac{\widehat{\psi}_{ij}^{\theta'}(m)}{\widehat{\psi}_{ij}^{\theta}(m)}$$

exceeds or not zero. Then, using the definitions of G_i and of F_{ij} in (7) and in (8), respectively, we obtain that:

(i) $\lim_{t\to\infty} q_i^{h_i^t}(\theta) = 0$ if

$$G_i(\theta) + \sum_{j\neq i} F_{ij}(\theta) < G_i(\theta') + \sum_{j\neq i} F_{ij}(\theta') \;\text{ for some } \theta' \neq \theta \quad \Leftrightarrow \quad \theta \notin \Theta_i^*;$$

(ii) $\lim_{t\to\infty} q_i^{h_i^t}(\theta) = 1$ if

$$G_i(\theta) + \sum_{j\neq i} F_{ij}(\theta) > G_i(\theta') + \sum_{j\neq i} F_{ij}(\theta') \;\text{ for each } \theta' \in \Theta\setminus\{\theta\} \quad \Leftrightarrow \quad \Theta_i^* = \{\theta\};$$

(iii)

$$\lim_{t\to\infty} q_i^{h_i^t}(\theta) = \left[1 + \sum_{\theta'\in\Theta_i^*\setminus\{\theta\}} \frac{p_i(\theta')}{p_i(\theta)}\right]^{-1} = \frac{p_i(\theta)}{\sum_{\theta'\in\Theta_i^*} p_i(\theta')}$$

if Θ_i^* is not singleton and $\theta \in \Theta_i^*$. ∎

Proof of Theorem 2. Consider a given social network Ψ, and take two different agents $i, j \in N$ and the directed path $\widehat{\gamma}_{ij} \in \Gamma_{ij}[\Psi]$ which conveys the highest amount of information from agent j to agent i. Using the definition of power of a directed path in (5), we have:

$$\begin{aligned}
\mathbb{P}(\widehat{\gamma}_{ij}) &= \sum_M \widehat{\psi}_{ij}(m) D\big(q_i^m[\widehat{\gamma}_{ij}] \,||\, p_i\big) = \sum_M \widehat{\psi}_{ij}(m) \sum_\Theta q_i^m[\widehat{\gamma}_{ij}](\theta) \log \frac{q_i^m[\widehat{\gamma}_{ij}](\theta)}{p_i(\theta)} \\
&= \sum_M \widehat{\psi}_{ij}(m) \sum_\Theta q_i^m[\widehat{\gamma}_{ij}](\theta) \log q_i^m[\widehat{\gamma}_{ij}](\theta) \\
&\quad - \sum_\Theta p_i(\theta) \log p_i(\theta) \sum_M \widehat{\psi}_{ij}^{\theta}(m) \left[\frac{\sum_{\theta'} \widehat{\psi}_{ij}^{\theta'}(m) p_j(\theta')}{\sum_{\theta'} \widehat{\psi}_{ij}^{\theta'}(m) p_i(\theta')}\right] \\
&= Q(p_i, p_j) - E_{\widehat{\psi}_{ij}}\big[H(q_i^m[\widehat{\gamma}_{ij}])\big].
\end{aligned} \tag{15}$$

Using Definition 7, it follows that agent j influences agent i if and only if the two following conditions are satisfied:

(i) $\Theta_i^* = \Theta_j^*$. This condition is satisfied if and only if for any $\theta \in \Theta_j$,

$$G_i(\theta) + \sum_{h \in N_i} F_{ih}(\theta) \geq G_i(\theta') + \sum_{h \in N_i} F_{ih}(\theta') \quad \forall \theta' \in \Theta.$$

Since we know that, for each $\theta \in \Theta_i$, $G_i(\theta) \geq G_i(\theta')$ for each $\theta' \in \Theta$, the above condition is equivalent to require that for any $\theta_j \in \Theta_j$ and any $\theta_i \in \Theta_i$

$$G_i(\theta_j) + \sum_{h \in N_i} F_{ih}(\theta_j) \geq G_i(\theta_i) + \sum_{h \in N_i} F_{ih}(\theta) \quad \forall \theta \in \Theta.$$

$$\Leftrightarrow G_i(\theta_j) + \sum_{h \in N_i} F_{ih}(\theta_j) > G_i(\theta_i) + \max_{\theta \notin \Theta_j} \sum_{h \in N_i} F_{ih}(\theta).$$

By adding the identity obtained in (15) to both sides of the inequality above, we obtain the following necessary and sufficient condition for $\Theta_i^* = \Theta_j^*$:

$$\mathbb{P}(\widehat{\gamma}_{ij}) > G_i(\theta_i) - G_i(\theta_j) + \max_{\theta \notin \Theta_j} \sum_{h \in N_i} F_{ih}(\theta) - \sum_{h \in N_i} F_{ih}(\theta_j) + Q(p_i, p_j) - E_{\widehat{\psi}_{ij}}\big[H(q_{ij}^m[\widehat{\gamma}_{ij}])\big],$$

which coincides with the condition stated in (7a).

(ii) $\Theta_j^* = \Theta_j$. This condition is satisfied if and only if, for any $\theta_j \in \Theta_j$,

$$G_j(\theta_j) + \sum_{h \in N_j} F_{jh}(\theta_j) \geq G_j(\theta) + \sum_{h \in N_j} F_{jh}(\theta) \quad \forall \theta \in \Theta.$$

$$\Leftrightarrow G_j(\theta_j) + \sum_{h \in N_j} F_{jh}(\theta_j) > \max_{\theta \notin \Theta_j} \big[G_j(\theta) + \sum_{h \in N_j} F_{jh}(\theta)\big].$$

By adding the identity obtained in (15) (upon changing the agents' subscripts to consider $\mathbb{P}(\widehat{\gamma}_{jk})$), where $k \in N_j$, to both sides of the inequality above, we obtain the condition

$$\mathbb{P}(\widehat{\gamma}_{jk}) < G_j(\theta_j) + \sum_{h \in N_j} F_{jh}(\theta_j) - \max_{\theta \notin \Theta_j} \big[G_j(\theta) + \sum_{h \in N_j} F_{jh}(\theta)\big] + Q(p_j, p_k) - E_{\widehat{\psi}_{jk}}\big[H(q_{jk}^m[\widehat{\gamma}_{ik}])\big],$$

for each $k \in N_j$, which, by rearranging terms, coincides with the condition stated. ∎

Proof of Proposition 1. First, note that application of the result in Theorem 1 to the external observer leads directly to the result that, for each history h_t, $\lim_{t \to \infty} q_{\mathrm{ob}}^{h^t}(\theta^*) = 1$ if and only if $\arg\max_{\theta \in \Theta} \sum_{i \in N} G_i(\theta)$ is a singleton with $\arg\max_{\theta \in \Theta} \sum_{i \in N} G_i(\theta) = \{\theta^*\}$.

Second, suppose that a consensus is achieved in the society in a way such that, for some $\theta^* \in \Theta$, we have $\lim_{t \to \infty} q_i^{h_i^t}(\theta^*) = 1$ for each history h_i^t, for each agent $i \in N$. Then, by using the result in Theorem 1, it follows that, for each agent $i \in N$,

$$G_i(\theta^*) + \sum_{j \in N_i} F_{ij}(\theta^*) \geq G_i(\theta) + \sum_{j \in N_i} F_{ij}(\theta) \quad \forall \theta \in \Theta,$$

which, by summing over all agents, implies

$$\sum_{i \in N} G_i(\theta^*) - \sum_{i \in N} G_i(\theta) \geq - \sum_{i \in N} \sum_{j \in N_i} [F_{ij}(\theta^*) - F_{ij}(\theta)]. \tag{16}$$

Therefore, provided that the consensus described above is achieved in the society, if

$$\sum_{i \in N} \sum_{j \in N_i} [F_{ij}(\theta^*) - F_{ij}(\theta)] < 0 \quad \forall \theta \in \Theta \setminus \{\theta^*\}$$

holds, then the condition in (16) above implies that $\sum_{i \in N} G_i(\theta^*) \geq \sum_{i \in N} G_i(\theta)$ for each $\theta \in \Theta$, with strict inequality if $\theta \neq \theta^*$. As a consequence, correct limiting beliefs are attained in the society. ∎

Proof of Proposition 2. Consider the center-directed star network $\Psi^s = \{\Psi_{j1} : j \in N \setminus \{1\}\}$ and suppose that a consensus is achieved in a way such that, from the results of Theorem 1, for each agent $j \in N \setminus \{1\}$, we have $\Theta^*_j = \Theta_1 = \{\theta^*\}$ for some given parameter value $\theta^* \in \Theta$. Let us define, for $\theta \in \Theta \setminus \{\theta^*\}$, $\overline{\eta}(\theta) := \max_{j \in N \setminus \{1\}} [G_j(\theta) - G_j(\theta^*)]$ and $\underline{\eta}(\theta) := \min_{j \in N \setminus \{1\}} [G_j(\theta) - G_j(\theta^*)]$.

First, note that by applying the log-sum inequality for each given $m \in M$, we know that

$$\begin{aligned} F_{j1}(\theta^*) - F_{j1}(\theta) &= \sum_M \sum_S \sigma^s_{j1}(m) \phi^\theta_1(s) \log \frac{\sum_S \sigma^{s'}_{j1}(m) \phi^{\theta^*}_1(s')}{\sum_S \sigma^{s'}_{j1}(m) \phi^\theta_1(s')} \\ &\leq \sum_S \phi^\theta_1(s) \log \frac{\phi^{\theta^*}_1(s)}{\phi^\theta_1(s)} \left[\sum_M \sigma^s_{j1}(m) \right] \\ &= G_1(\theta^*) - G_1(\theta) \qquad \forall \theta \in \Theta \setminus \{\theta^*\}, \quad \forall j \in N \setminus \{1\}. \end{aligned} \tag{17}$$

Since we are supposing that the central agent is able to influence each other agent j so that all agents' limiting beliefs put probability one on θ^* being the true parameter value (i.e., $\Theta^*_j = \theta^*$ for each $j \in N \setminus \{1\}$), then it must be the case that $F_{j1}(\theta^*) - F_{j1}(\theta) > G_j(\theta) - G_j(\theta^*)$ for each parameter value $\theta \in \Theta \setminus \{\theta^*\}$ and each agent $j \in N \setminus \{1\}$. It then follows from the inequality in (17) above that $G_1(\theta^*) - G_1(\theta) > G_j(\theta) - G_j(\theta^*)$ for each $\theta \in \Theta \setminus \{\theta^*\}$ and each $j \in N \setminus \{1\}$. This condition is equivalent to require $G_1(\theta^*) - G_1(\theta) > \overline{\eta}(\theta)$ for each $\theta \in \Theta \setminus \{\theta^*\}$.

Secondly, correct limiting beliefs put probability one on some parameter value $\widehat{\theta} \neq \theta^*$ being the true one if and only if

$$\sum_{j \in N \setminus \{1\}} \left[G_j(\widehat{\theta}) - G_j(\theta^*) \right] > G_1(\theta^*) - G_1(\widehat{\theta}).$$

A straightforward sufficient condition for the requirement above to be satisfied is

$$\underline{\eta}(\widehat{\theta}) > (n-1)^{-1} \left[G_1(\theta^*) - G_1(\widehat{\theta}) \right].$$

Therefore, since considering that the influence of the central agent leads to a consensus in which all agents put probability one to θ^* being the truth necessarily requires that $G_1(\theta^*) - G_1(\theta) >$

$\overline{\eta}(\theta)$ for any $\theta \neq \theta^*$, it follows that correct limiting beliefs are not attained if, for some $\widehat{\theta} \neq \theta^*$, the following condition is satisfied.

$$G_1(\theta^*) - G_1(\widehat{\theta}) > \overline{\eta}(\widehat{\theta}) > \underline{\eta}(\widehat{\theta}) > (n-1)^{-1} \left[G_1(\theta^*) - G_1(\widehat{\theta}) \right].$$

The proof is completed by noting that $\overline{\eta}(\widehat{\theta}) > \underline{\eta}(\widehat{\theta})$ is satisfied by construction and that, in addition, we can always find some large enough finite $n^* \geq 1$ such that, for each $n \geq n^*$, the sufficient condition above holds. ∎