Evolution of beliefs in social networks

Pushpi Paranamana
Saint Mary’s College, Notre Dame, IN, USA

Pei Wang
Rutgers University–Newark, Newark, NJ, USA

Patrick Shafto
Rutgers University–Newark, Newark, NJ, USA; Institute for Advanced Study, Princeton, NJ, USA

Abstract
Evolution of beliefs of a society are a product of interactions between people (horizontal transmission) in the society over generations (vertical transmission). Researchers have studied both horizontal and vertical transmission separately. Extending prior work, we propose a new theoretical framework which allows application of tools from Markov chain theory to the analysis of belief evolution via horizontal and vertical transmission. We analyze three cases: static network, randomly changing network, and homophily-based dynamic network. Whereas the former two assume network structure is independent of beliefs, the latter assumes that people tend to communicate with those who have similar beliefs. We prove under general conditions that both static and randomly changing networks converge to a single set of beliefs among all individuals along with the rate of convergence. We prove that homophily-based network structures do not in general converge to a single set of beliefs shared by all and prove lower bounds on the number of different limiting beliefs as a function of initial beliefs. We conclude by discussing implications for prior theories and directions for future work.

Keywords
Belief evolution, social networks, language evolution, homophily-based networks, vertical and horizontal transmission, belief control

Significance Statement
The beliefs we see in society are a product of transmission between people and over generations, thus combining both horizontal transmission and vertical transmission. Under what conditions may we expect people to converge to similar or very different beliefs? We present conditions for homogeneous belief systems and heterogeneous belief systems in the society. When social dynamics of communication are independent of beliefs, both static and randomly changing networks converge to homogeneous beliefs. In contrast, when people tend to communicate with those who have similar beliefs, long term differences in beliefs among people result. Thus, our work suggests mechanisms that one may use to balance the degree to which beliefs are shared or different in society.
Introduction

Evolution of beliefs, individual and cultural, is the result of vertical transmission between generations and horizontal transmission within a generation. Research in cognitive science has developed models of vertical transmission, through connections to probabilistic models of cognition (Chater et al., 2006) and used such models to investigate innate cognitive constraints and connections to experience (Griffiths and Kalish, 2007; Kirby et al., 2007). Separately, research in network science has developed theories that explain horizontal transmission, the social dynamics of transmission and diffusion patterns (Newman, 2003; Zhou et al., 2020). Because beliefs are shaped both by vertical and horizontal transmission, any successful theory of evolution of beliefs will need to combine aspects of both approaches. We propose a mathematical approach that enables detailed analysis of the long run consequences of vertical and horizontal transmission for individual and cultural beliefs.

Theories in cognitive science frame vertical transmission through evolution as functional adaptations of cognitive capacities, such as language, beliefs, knowledge and meta-cognition, to ancestral environment. (Griffiths and Kalish, 2007; Kirby et al., 2007; Suchow et al., 2017; Whalen and Griffiths, 2017) have developed methods to interpret vertical transmission between Bayesian agents as Markov chains, thus revealing innate cognitive constraints and structures as the outcome of such processes. For example, (Griffiths and Kalish, 2007) interpret transmission of language from parents to children as a Markov chain, which leads to the conclusion that, in the absence of other influences, the resulting observed distribution of languages reflects our prior biases about language and language structures.

However, cognition and memory are sustained by both communicative and cultural aspects (Candia et al., 2019; Zhou et al., 2020) and reflect social influences (Abrams et al., 2011; Roberts and Fedzechkina, 2018). This horizontal transmission is intrinsically bidirectional and introduces the possibility of long term consequences of social network structures for beliefs. Network theory has studied transmission over social networks (Boccaletti et al., 2006; Delvenne et al., 2015) for cases including diseases (Huang et al., 2019; Pastor-Satorras et al., 2015), information (Wang et al., 2013; Zhou et al., 2020; Zhan et al., 2019), opinions (Castellano, 2012; Quattrociocchi et al., 2014), and rumours (Moreno et al., 2004). However, in these models transmission is formalized as a property that can be caught or passed between agents. This is suitable for diseases and facts, but beliefs are more naturally represented as distributions over some latent space, as in probabilistic models of cognition used to model vertical transmission.

In this article, we combine both vertical and horizontal transmission to explore the long term evolution of beliefs. We provide a mathematical formulation to analyze the limiting distribution of beliefs in societies based on sociodynamic aspects and cognitive aspects of belief evolution. This limiting distribution tells us the long term belief distribution of each individual. Moreover this provides a framework to explore the long term belief evolution of groups and/or of the society as a whole. Integrating classical results of time homogeneous and inhomogeneous Markov chain theories, we provide conditions on the network structures—static and dynamic at random—that result in homogeneous/heterogeneous belief systems among individuals (or groups). Moreover, we provide rates of convergence of the models to their limiting behaviors, for both static and random cases. Prior studies show that individuals in a social network may tend to connect to individuals who share similar interests, and thus it is considered as an important evolutionary mechanism (Liu et al., 2018). We integrate this assortive dynamics in which networks are formed based on homophily and prove conditions under which societies will converge to heterogenous beliefs.

There has been extensive research on how belief diversity enhances the collective intelligence. A society that collectively has similar beliefs offers little chance for collective decision making to improve over any individuals. If individuals have different beliefs, collective accuracy can be enhanced. A simple example comes from “wisdom of crowds” effects in which the average of a group of people’s guesses is more accurate than most individuals (Galton, 1907), but many more examples exist in the decision making literature. Integration of multiple beliefs and, diversity in beliefs is thus required for underlying collective intelligence (Broomell and Budescu, 2009; Keuschnigg and Ganser, 2017; Novaes Tump et al., 2018). In this work we explore the network and belief structures that result in belief homogeneity vs heterogeneity under three scenarios: static networks, randomly changing networks and homophily-based networks. Thus the results can be used to explore conditions on optimal structures that improves collective accuracy and evolution.

Formulation of the problem

Our aim is to develop a model that one can use to analyze evolution of individual and societal beliefs through both vertical and horizontal transmission. Our approach builds on prior research in the cognitive science literature formalizing vertical transmission as a Markov Chain (Griffiths and Kalish, 2007; Kirby et al., 2007; Whalen and Griffiths, 2017), while integrating horizontal transmission from network theory. To integrate horizontal transmission, we formalize interactions among individuals in a society with a given, possibly dynamic, structure. As in prior work, individuals’ initial beliefs are assumed to be sampled from a given distribution. Individuals within a society will interact with subsets of other individuals as defined by an adjacency matrix defining network structure. Networks may take a variety of forms including unidirectional and bidirectional, static and
dynamic, and belief dependent. Each of these cases can be represented as a (collection of) adjacency matrix (matrices).

**Definition 1. Evolution of beliefs in social networks.** Consider a set of people \( P = \{a_i\}_{i=1}^n \) in the society and a set of concepts \( H = \{\beta_k\}_{k=1}^r \). Denote people’s priors on \( H \) by \( M = \{m_{jk}\}_{j,r} \), the network structure over which people may communicate at time \( t \) by \( P_t = (p_{ij})_{i,r} \), and the concept structure at time \( t \) by \( H_t = (h_{kl})_{s,r} \). All are row stochastic matrices. Let \( P_0 = H_0 = I \), where \( I \) is the identity matrix of corresponding order. Define

\[
Q_n(P_t,H_t,M) = \prod_{t=0}^{n-1} P_t M \prod_{t=0}^{n-1} H_t,
\]

where \( Q_n \) represents the society’s beliefs at time \( n \), and the long-run beliefs are analyzed as \( n \rightarrow \infty \). That is, at each time \( n \), \( Q_n \) is the product \( P_n \ldots P_0 M H_n \ldots H_0 \).

The model formulates the time evolution of people’s beliefs. \( m_{jk} \) represents the initial belief (prior) of the person \( a_j \) on the concept \( \beta_k \); \( p_{ij} \) denotes the weight that the \( i \)th person gives to the \( j \)th person’s information and \( h_{kl} \) denotes the degree to which concept \( \beta_l \) may be confused for \( \beta_k \). A variety of properties can be captured in the matrices \( P_t \) and \( H_t \).

Consider \( P_t \). Absence of direct transmission is formalized when \( p_{ij} = p_{ji} = 0 \). Bidirectional transmission is formalized by \( p_{ij} > 0 \) and \( p_{ji} > 0 \). Unidirectional transmission is formalized when either \( p_{ij} > 0 \) and \( p_{ji} = 0 \) or \( p_{ij} = 0 \) and \( p_{ji} > 0 \). Different network structures including random graphs, small-world and scale-free networks (Albert and Barabási, 2002) can be formalized through the construction of adjacencies. Dynamic network structures (Li et al., 2017) are formalized by introducing the subscripts \( P_t \) and \( H_t \) to indicate the network structure at time \( t \). Notice that \( P_t, M \) and \( H_t \) are stochastic matrices. Therefore for any \( n, Q_n(P_t,H_t,M) \) is also a stochastic matrix. This approach combines both the sociodynamic and the cognitive aspects of belief evolution which helps to evaluate society’s vertical and horizontal transmission simultaneously. **Table 1.**

Next, we illustrate the design of the structures and the model using some stylized examples.

**Example 2.** Consider a neighborhood with three people \( a_1, a_2, a_3 \) with the network structure at time \( t \) given

\[
P_t = \begin{pmatrix}
    a_1 & a_2 & a_3 \\
    1 & 0 & 0 \\
    2 & 3 & 1 \\
    2 & 4 & 4
\end{pmatrix}
\]

The network structure can be depicted in Figure 1, where the directed edge from \( a_i \) to \( a_j \) denotes \( p_{ij} \).

Here, \( a_1 \) does not believe what anyone else says but believes himself 100%, while \( a_2 \) only believes what others say. However, \( a_3 \) believes him self and others with certain percentages. In practice, this may model the communication between three listeners \( a_1, a_2, a_3 \), where \( a_1 \) may be a speaker, \( a_2 \) may model a new student in the class who is learning from his teacher \( a_1 \) and a peer \( a_3 \).

Similarly, the formulation of the concept structure can be viewed as a graph. Instead of pointing from speaker to listener, arrows point from a concept toward a concept it can replace (be confused with). Note that the concept structure is modeled as a distribution rather than single values. Each value corresponds to the degree (weight) to which one concept may be confused with another. Notice that the concept structure corresponds to the vertical component of the model. That is, the generational or cultural transmission of beliefs. Confusability of beliefs is a reasonable notion to denote the imperfect dynamics over generations due to changes in cultural traits and new found information over time leading to generation gaps.

**Example 3.** Consider a group of 5 people, each holding a belief on 5 distinct concepts. Suppose people have a prior belief distribution given by \( M = I_{5 \times 5} \). That is \( a_i \) believes only on the concept \( \beta_i \) for all \( i = 1, \ldots, 5 \). Let \( P = \)

| Notation | Definition |
|----------|------------|
| \( P = \{a_i\}_{i=1}^n \) | a set of people in the society, where \( a_i \) denotes the \( i \)th person |
| \( H = \{\beta_k\}_{k=1}^r \) | a set of concepts, where \( \beta_k \) denotes the \( k \)th concept |
| \( M = \{m_{jk}\}_{j,k} \) | a row stochastic matrix records a set of people’s priors over a set of concepts each row represents a person, each column represents a concept \( m_{jk} \) denotes the initial belief of person \( a_j \) on concept \( \beta_k \) |
| \( P_t = (p_{ij})_{i,r} \) | a row stochastic matrix denotes the network structure at time \( t \) \( p_{ij} \) denotes the weight that \( i \)th person gives to \( j \)th person’s information |
| \( H_t = (h_{kl})_{s,r} \) | a row stochastic matrix denotes the concept structure at time \( t \) \( h_{kl} \) the degree to which concept \( \beta_l \) may be confused for \( \beta_k \) |
| \( Q_n(P_t,H_t,M) \) | a row stochastic matrix modeled by \( \prod_{t=0}^{n-1} P_t M \prod_{t=0}^{n-1} H_t \) records the society’s beliefs at time \( n \) |
In the next section, we analyze the long term behavior of the model theoretically which sheds light on belief evolution and societal belief diversity. The above example illustrates the model for a static network structure and a static concept structure. However the structures could be dynamic, thus we will consider three phenomena: static structures, randomly changing structures and homophily-based dynamic structures. Analyzing this model will help us better understand the minimal conditions necessary for sustained belief heterogeneity, conditions on which the homogeneity is attained.

Note: Markov chains are widely used in many applications in predicting variation tendencies of random processes including modeling inter generational beliefs. Belief evolution can be studied as transmission chains where the beliefs evolve through time via horizontal and vertical transmission, which is mathematically parallel to analyzing Markov chains. So in our model both the network structure and the concept structure are considered as Markov chains and are represented by corresponding transition matrices.\footnote{Note: Markov chains are widely used in many applications in predicting variation tendencies of random processes including modeling inter generational beliefs. Belief evolution can be studied as transmission chains where the beliefs evolve through time via horizontal and vertical transmission, which is mathematically parallel to analyzing Markov chains. So in our model both the network structure and the concept structure are considered as Markov chains and are represented by corresponding transition matrices.}

**Analyzing the belief evolution in social networks**

In this section, we explore belief change in the long run, individually and societally. We analyze under what conditions a society will attain homogeneity of beliefs and whether the society will evolve into groups with distinct beliefs. Moreover, we explore how fast a society will converge to its final belief system. As discussed in Section 2, networks can be time invariant as well as time variant. Therefore we investigate the belief evolution for time homogeneous and time inhomogeneous cases separately.

**Belief evolution over stable social and belief networks**

First, we analyze the belief evolution when network and concept structures are time homogeneous. That is, we assume that $\forall t$, $P_t = P$ and $H_t = H$; $P$ and $H$ are fixed matrices. Then the operator $Q$ simplifies to

$$Q_n(P, H, M) = P^n M H^n.$$ 

For a square matrix $A$, $A^n$ denotes the multiplication of $A$ for $n$ times.

**Convergence and limiting distribution.** As transition matrices of Markov Chains, important distinctions about the network and concept structure are whether they are indecomposable/decomposable and reducible/irreducible. The limiting behavior depends on the structures as well as the states of the people and beliefs (transient/persistent). Therefore we define:

**Definition 4. Irreducibility:** A set $C$ of states is closed if no state outside $C$ can be reached from any state $j$ in $C$. A
Markov chain is irreducible if there exists no closed sets other than the set of all sets; otherwise, it is reducible.

**Definition 5. Indecomposability:** A Markov chain is indecomposable if it contains at most one closed set of states other than the set of all states. Otherwise it is decomposable.

**Definition 6. Transient/Recurrent states:** State $i$ is called transient if, given that we start in state $i$, there is a non-zero probability that we will never return to $i$. State $i$ is called recurrent (or persistent) if it is not transient.

**Definition 7. Stationary distribution:** Let $A$ be a transition matrix. A stationary distribution (steady state distribution) $\pi$ is a non-negative stochastic (row) vector, such that $\pi A = \pi$.

An irreducible and aperiodic (Definition A.2) Markov chain has a unique stationary distribution $\pi$ as $n \to \infty$ (Gravner, 2010). That is, the transition matrix converges to a matrix with same rows equals to $\pi$. Moreover, $\pi$ is the left eigenvector of the associated transition matrix that corresponds to the unit eigenvalue (which exists and is unique). If the Markov chain is indecomposable but reducible, the transient states vanish in the limit. Similarly, we can analyze the structure of the limit of decomposable, aperiodic chains using Propositions B.1 and B.2.

**Proposition 8**. Assume the network structure $P_{rs}$ and the concept structure $H_{rs}$ are aperiodic matrices. Let $M_{rs}$ be the initial belief distribution in the society. The society will stabilize in the long run. That is, $\lim_{n \to \infty} |Q_{n+1} - Q_n| \to 0$ as $n \to \infty$.

(i) If $H$ is indecomposable, then in the long run, the society stabilizes to a single belief distribution that does not depend on $P$ or $M$. That is, there exists a steady state distribution $\pi = \{\pi_1, \ldots, \pi_s\}$ such that for any $P$ and $M$

$$\lim_{n \to \infty} Q_n = \lim_{n \to \infty} P^n M H^n = \left( \begin{array}{ccc} \pi_1 & \cdots & \pi_s \\ \vdots & \ddots & \vdots \\ \pi_1 & \cdots & \pi_s \end{array} \right)_{(s \times s)}$$

(ii) If $H$ is decomposable and $P$ is indecomposable, then in the long run, the society stabilizes to single a belief distribution that depends on $M$ and $H$. That is, for any $M$ and $H$, there exists a steady state distribution $\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_s\}$ such that:

$$\lim_{n \to \infty} Q_n = \lim_{n \to \infty} P^n M H^n = \left( \begin{array}{ccc} \sigma_1 & \cdots & \sigma_s \\ \vdots & \ddots & \vdots \\ \sigma_1 & \cdots & \sigma_s \end{array} \right)_{(s \times s)}$$

(iii) If $H$ and $P$ are both decomposable. Then the society will not have a single belief distribution in the long run. That is, the rows of the matrix $\lim_{n \to \infty} Q_n$ are not all the same.

Notice that, if $H$ is indecomposable neither $P$ or $M$ have an effect on the limit of the model. That is, the concept structure dominates and controls the long run behavior, regardless of what the network structure is or what people initially believe. Moreover, as a Markov chain, transient beliefs (if there are any) vanish from the society. However if $H$ is decomposable, in addition to the concept structure, the network structure as well as the initial belief distribution affect the long run behavior. The homogeneity of beliefs among people in the society depends on the network structure. In particular, if the network structure is indecomposable, then there will be a unique belief distribution in the society regardless of the initial beliefs. In summary, if either the network structure or the concept structure is indecomposable, the society will converge to a unique belief distribution in the long run. However, if both are decomposable, there will not be a single belief distribution in the society; there will be heterogeneity among individuals. These scenarios are illustrated in Examples 9, 10, 11.

**Example 9.** Consider $P, M$ and $H$ given in Example 3. Here, $P$ is decomposable and has two closed communicating classes and $H$ is indecomposable. Then, $\lim_{n \to \infty} Q_n = \begin{pmatrix} 0.285 & 0.203 & 0.2 & 0.155 & 0.156 \\ 0.285 & 0.203 & 0.2 & 0.155 & 0.156 \\ 0.285 & 0.203 & 0.2 & 0.155 & 0.156 \\ 0.285 & 0.203 & 0.2 & 0.155 & 0.156 \\ 0.285 & 0.203 & 0.2 & 0.155 & 0.156 \end{pmatrix}$ In this example, even though there are people in the network who never talk to each other, everyone converges to the same beliefs in the long run. This is because the concept structure $H$, which is indecomposable, dominates.

**Example 10.** Consider $P = \begin{pmatrix} 0.3 & 0.7 & 0 \\ 0.6 & 0.4 & 0 \end{pmatrix}$, $M = \begin{pmatrix} 0.1 & 0.3 & 0.12 & 0.02 & 0.28 & 0.18 \\ 0.06 & 0.23 & 0.1 & 0.23 & 0.23 & 0.15 \\ 0.03 & 0.01 & 0.39 & 0.38 & 0.09 & 0.1 \end{pmatrix}$, and $H = \begin{pmatrix} 0.803 & 0.197 & 0 & 0 & 0 & 0 \\ 0.464 & 0.536 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.272 & 0.038 & 0.02 & 0.669 \\ 0 & 0 & 0.515 & 0.017 & 0.401 & 0.068 \\ 0 & 0 & 0.144 & 0.319 & 0.002 & 0.535 \\ 0 & 0 & 0.16 & 0.357 & 0.242 & 0.241 \end{pmatrix}$, where both $P$ and $H$ are decomposable. $P$ has two closed communicating classes: $\{\alpha_1, \alpha_2\}$ and $\{\alpha_3\}$. Then the
stationary distribution (rounded up to 3 decimal places) is
\[
\lim_{n \to \infty} Q_n =
\begin{pmatrix}
0.239 & 0.102 & 0.169 & 0.131 & 0.115 & 0.241 \\
0.239 & 0.102 & 0.169 & 0.131 & 0.115 & 0.241 \\
0.028 & 0.012 & 0.245 & 0.191 & 0.167 & 0.351
\end{pmatrix}
\]
We can see that in the limit, people’s beliefs are not the same. However, the beliefs of people who are in the same closed communicating class are the same. That is, \( \alpha_1 \) and \( \alpha_2 \) have the same beliefs while \( \alpha_3 \) has different beliefs.

**Example 11.** Consider \( P = \) \[
\begin{pmatrix}
0.268 & 0.422 & 0.31 \\
0.331 & 0.232 & 0.437 \\
0.094 & 0.364 & 0.542
\end{pmatrix}
\]
and \( H = \) \[
\begin{pmatrix}
0 & 0.3 & 0.7 & 0 \\
0 & 0.6 & 0.4 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]
where both \( P \) and \( H \) are decomposable. In \( P \), \( \{\alpha_1\} \) and \( \{\alpha_6\} \) are closed classes. \( \lim_{n \to \infty} Q_n = \) \[
\begin{pmatrix}
0.318 & 0.372 & 0.31 \\
0.052 & 0.061 & 0.887 \\
0.082 & 0.996 & 0.823 \\
0.074 & 0.887 & 0.839 \\
0.081 & 0.094 & 0.825 \\
0.027 & 0.032 & 0.941
\end{pmatrix}
\]
We can see that there will not be a unique belief distribution among people in the limit.

**How are the social/concept structures represented by these different matrices?** Social structures are typically highly structured. For example, some are strongly connected. That is, it is possible to communicate from any person by a chain of individuals to any other person in the network. This scenario can be represented by an indecomposable matrix. On the other hand there are social networks where the communication is unidirectional. For instance, media can be thought of as a unidirectional communication path in the sense that the news is broadcast, and no matter how loud one yells at the screen, the newscaster does not hear you; hence, the audience’s beliefs are transient. Also, some structures have a strong asymmetry between groups. Colonialism is such an example. These scenarios can be represented by different structures of decomposable matrices. Similar analogy can be made for concept structures based on the relatedness between beliefs.

**Example 12.** What if \( M \) is decomposable? Consider a situation where different groups of people have no common beliefs. For example, people in different countries may have different sets of languages (or dialects), with no common language between the countries. Assuming people learn languages by talking to others, and that \( H \) has some structure representing relatedness of the languages, we can explore the long term distribution of languages among people using our framework. Notice, here the prior matrix \( M \) is a block diagonal matrix. Let \( P^nM = P \). If \( H \) is indecomposable, the society will follow a same language distribution. However, if \( H \) is decomposable but \( P \) is indecomposable, then the society will stabilize to a same distribution of languages that depends on \( P; M \) and \( H \). If both \( H \) and \( P \) are decomposable, then the society will stabilize to a heterogenous distribution of languages.

**Rate of convergence.** One may ask how fast the individuals or the society attain their limiting beliefs. This provides insights to the rate of belief evolution. More precisely, what is the effect of the structure of \( P \) and \( H \) matrices on how fast the model converges to its stationary distribution. We provide a lower bound on the rate of convergence of the model that represents how quickly the sequence approaches its stationary distribution. (See definition B.3)

According to Proposition B.4, the convergence rate of an indecomposable Markov chain is governed by the second largest eigenvalue, which is less than 1. If the chain is decomposable, it has more than one closed communicating class. We can treat each class as an indecomposable chain and find each of its rate of convergence. The slowest of those rates will be considered as the convergence rate of the decomposable chain.

**Proposition 13.** Suppose the network structure \( P_{r,s} \), and the concept structure \( H_{r,s} \), are indecomposable and aperiodic. Let \( \lambda_P \) and \( \lambda_H \) denote the second largest eigenvalues of \( P \) and \( H \), respectively, and \( Q_n = P^nM H^n = (q_{ij})_{n \times n} \). Then there exists a positive constant \( C_0 \) such that for all \( i = 1, \ldots, r \) and \( j = 1, \ldots, s \)

\[
|q_{ij} - \pi_j| \leq C_0 (\lambda_P \lambda_H)^n
\]
where \( \pi = \{\pi_1, \ldots, \pi_N\} \) is the stationary distribution of \( Q \). Note that \( \lambda_P < 1, \lambda_H < 1 \), therefore \( C_0 (\lambda_P \lambda_H)^n \to 0 \) as \( n \to \infty \).

**Proposition 14.** Let \( R_P \) and \( R_H \) be the convergence rates of \( P \) and \( H \), respectively. Then the model converges to the stationary distribution with a rate of at least \( R = \min\{R_P, R_H\} \).

That is, the society will reach the steady state distribution only when both network and concept structures are stabilized.

**What if the social structure and the concept structure change over time?**
In this section, we consider time inhomogeneous models, where network and concept structures can change over time.
We provide conditions for the model convergence to homogeneous beliefs convergence in expectation, and a lower bound for the rate of convergence of the model.

**Convergence and limiting distribution.** For simplicity, $M$ is assumed to be indecomposable in the formulation of the problem. For time homogeneous case, Proposition 8 suggests that if either $P$ or $H$ is stochastic, indecomposable and aperiodic (SIA), homogeneity of beliefs is guaranteed. This can be generalized to time inhomogeneous case as following:

**Proposition 15.** Let $S_P = \{P_i\}_{i=1}^k$ be a set of social structure matrices, and $S_H = \{H_i\}_{i=1}^l$ be a set of concept structure matrices. At each time $t$, $P_t$ and $H_t$ are chosen from $S_P$ and $S_H$ respectively. Then $Q_n(P_t, H_t, M) = \prod_{i=0}^n P_t M \prod_{i=0}^n H_t$ converges to a rank one matrix as $n \rightarrow \infty$ if and only if every possible product of matrices in $S_P$ or $S_H$ converges to rank one matrix as $n \rightarrow \infty$.

Proposition 15 provides a condition that guarantees a homogeneous belief distribution in the society in the long run. In particular, if every product in the set of network structures and the set of concept structures is SIA, the society will stabilize to a unique belief distribution. However, note that each matrix in a set $S$ (a set of stochastic matrices) being SIA does not guarantee that every product is SIA, and as the order of the transition matrices increases (that is, as the number of states of the Markov chain increases) it is difficult to check if every product is SIA. Therefore, we now discuss conditions on the individual matrices from $S$ which guarantees that any product of matrices from $S$ is SIA.

**Definition 16.** For a square stochastic matrix $P$, let $\lambda(P) = 1 - \gamma$, where $\gamma$ is the ergodic coefficient of $P$. If $\lambda(P) < 1$, then $P$ is called a **scrambling** matrix.

**Proposition 17.** If every matrix in $S$ is stochastic and scrambling, then any product of matrices from $S$ converges to a rank one matrix.

This reduces the required amount of computations as it is relatively easy to check if a matrix is scrambling or not. Moreover, given a set $S$, we only need to check all matrices in $S$, rather than every possible product of matrices in $S$.

**Example 18.** Suppose there are two belief evolution systems, one with concept structure set $S_H = \{H_1, H_2\}$, the other one with $S_H = \{H_1, H_3\}$, where $H_1 = \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.6 & 0.4 & 0 \\ 0.8 & 0.2 \end{pmatrix}$, $H_2 = \begin{pmatrix} 0.6 & 0.1 & 0.3 \\ 0 & 1 & 0 \\ 0.7 & 0.3 \end{pmatrix}$, $H_3 = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0.3 & 0.2 & 0.5 \\ 0.4 & 0.5 & 0.1 \end{pmatrix}$.

Convention shows that $\gamma(H_1) = \sum_{i=0}^{\infty} \min(p_{ij}) = 0 + 0 = 0.4$, hence $\lambda(H_1) = 1 - 0.4 = 0.6 < 1$, $H_1$ is scrambling. Similarly, one have $\lambda(H_2) = 1$, $\lambda(H_3) = 0.7$. Hence $H_2$ is scrambling, and $H_3$ is not. Therefore according to Proposition 17, people in the second system must converge to the same belief. Whether the first system converges to the same belief is further depending on its social structure set $S_P$.

**Convergence in probability setting.** Proposition 15 provides a necessary and sufficient condition on when every product of stochastic matrices from $S = \{A_1, \ldots, A_l\}$ converges to a rank one matrix. In contrast to this absolute setting, we now consider the convergence in probability.

**Proposition 19.** Given a set of stochastic matrices $S = \{A_1, \ldots, A_l\}$ and a positive vector $w = (w_1, \ldots, w_l) \in \mathbb{R}^l$ with $w_i > 0$ and $\sum_{i=1}^l w_i = 1$, a product of matrices that are i.i.d. sampled from $S$ according to $w$ converges to a rank one matrix with probability one if and only if there exists a finite product $B = \prod_{i=1}^N B_i$ of matrices, where $B_i$ is from $S$ such that $B$ is scrambling.

**Remark 20.** We may replace ‘scrambling’ in Proposition 19 by ‘SIA’ as sufficiently large powers of an SIA matrix are scrambling and any product that has a scrambling matrix as a factor is SIA.

Although it is easy to check if a matrix is scrambling, to make sure whether a scrambling product $B$ exist in Proposition 19 could still be challenging. We now introduce an equivalent condition in form of graphs, which is straightforward to verify.

Associated with the finite state Markov chain of a transition matrix $A$, there is a directed graph $G_A$. For instance, let $A = \begin{pmatrix} 0.7 & 0.3 \\ 0 & 1 \end{pmatrix}$, then the corresponding graph is . Similarly, associated with a set of transition matrices $S = \{A_1, \ldots, A_l\}$ with a fixed collection of states, we may define a directed graph $G_S$, where $G_S$ has the same vertex set as any $G_{A_l}$ and the edge set contains $(i,j)$ if there exists a $k \in \{1, \ldots, l\}$ such that $G_{A_k}$ contains $(i,j)$. For instance, let $S = \{A_1, A_2, A_3\}$, where $A_1 = A$ as above, $A_2 = \begin{pmatrix} 0.3 & 0.2 & 0.5 \\ 0 & 1 & 0 \\ 0.8 & 0.2 \end{pmatrix}$, and $A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.6 & 0.4 \\ 0 & 0 & 1 \end{pmatrix}$, then the corresponding graph $G_S$ is:

A set of vertices are said to be
**strongly connected** if there exists a directed path between any pairs of vertices in the set. Each $G_S$ further induces a condensed graph $\hat{G}_S$ by combining vertices in each strongly connected set into a ‘super-vertex’. In our example $v_2$ is strongly connected to $v_3$. Hence we have $G_S$ as:

\[
\begin{array}{c}
\text{A state is defined to be} \quad \text{recurrent} \quad \text{if it is} \quad \text{contained in a leaf of} \quad \hat{G}_S, \quad \text{otherwise the state is} \quad \text{transient}. \quad \text{Thus} \quad \text{in our example,} \quad v_{2,3} \quad \text{are recurrent, and} \quad v_1 \quad \text{is transient.}
\end{array}
\]

Notice that if $\hat{G}_S$ is connected and has one leaf, this is equivalent to a finite product of matrices from $S$ which are scrambling. Hence, as a consequence of Proposition 19, we have,

**Corollary 21.** Given a time-inhomogeneous Markov chain with $S$ and $w$, a product of transition matrices that are i.i.d. sampled from $S$ according to $w$ converges to a rank one matrix with probability one if and only if $\hat{G}_S$ is connected and has one leaf.

The limit of the product of sampled transition matrices may not exist when there are more than one leaf of $\hat{G}_S$. Hence instead we now consider the expectation of such limit.

**Proposition 22.** Given a time-inhomogeneous Markov chain with $S$ and $w$, the expectation of the product of sampled transition matrices is equal to the limiting product of the expectation of the transition matrix, i.e.

\[
\mathbb{E}(\lim_{n \to \infty} \prod_{t=1}^{n} X_t^t) = \lim_{n \to \infty} \prod_{t=1}^{n} \mathbb{E}(X_t^t), \quad \text{where} \quad \mathbb{E}(X_t^t) = \sum_{k=1}^{N} w_k \cdot A_k.
\]

Based on the above analysis, we can now investigate long-term behavior when both network and concept structures are sampled from a collection of matrices $S_P = \{P_1, \ldots, P_k\}$ and $S_H = \{H_1, \ldots, H_l\}$ respectively. Let the condensed graphs corresponding to $S_P$ and $S_H$ be $G_P$ and $G_H$. According to Proposition 19 and Corollary 21, analogous to Proposition 8 for the time homogeneous case, the following holds:

When $G_H$ has only one leaf, or equivalent there is a finite product from $S_H$ is scrambling or SIA, then with probability one everyone in the network converges to the same posterior distribution over the hypothesis set $H$. In particular, the posterior distribution is supported only on the recurrent hypotheses, i.e. hypotheses in the leaf. When $G_H$ has more than one leaf, or equivalently there is a common indecomposable structure for every matrix in $S_H$, but $G_P$ has only one leaf, then everyone in the network still converges to the same posterior distribution over $H$ with probability one. Moreover, the shared posterior distribution is a mixture of isolated posterior distribution of recurrent people (people in the leaf vertices). Thus, the shared posterior distribution is completely determined by priors of recurrent people and their belief’s corresponding confusion parameters.

When both $G_H$ and $G_P$ have more than one leaf, people in different recurrent classes (people in different leaf vertices) converge to different posterior distributions. In general, posteriors of people in transient states is a mixture of posteriors for recurrent classes where the mixture weights differ over time (no limit exists).

In all cases, Proposition 22 suggests that the expectation of people’s posterior is:

\[
\lim_{n \to \infty} \mathbb{E}(P^n \cdot M \cdot (\mathbb{E}H)^n) = \mathbb{E}(P^n) \cdot M \cdot (\mathbb{E}H)^n.
\]

**Rate of convergence.** Next, we explore the rate of convergence of inhomogeneous Markov chains. We then discuss how to obtain the rate of convergence of the model when both $P_n$ and $H_n$ are time inhomogeneous.

**Proposition 23.** (Anthonisse and Tijms, 1977) Suppose that any product of matrices from $S$ converges to a rank one matrix. Then there exist an integer $v \geq 1$, for any sequence $\{A_1, \ldots, A_n\}$, $n \geq 1$ of matrices from $S$, such that for all $i, j = 1, \ldots, N,$

\[
|A_1 \ldots A_n| - \pi^i_j \leq (1 - \gamma)^{|x|}
\]

for all $n \geq 1$, where $\pi = (\pi_1, \ldots, \pi_N)$ is the stationary distribution, $\gamma = \min \{|(A_1 \ldots A_n)|A_i \in S, 1 \leq i \leq v\}$, and $[x]$ is the largest integer less than or equal to $x$.

In other words, the above proposition provides an upper bound for the rate at which the network structure (or concept structure) stabilizes, for any SIA product of matrices in $S_P$ (or $S_H$). Note that this depends on the ergodic coefficients of the matrices. Integer $v$ can always be taken less than or equal to $v^* = \frac{1}{2}(3^N - 2^{N+1} + 1)$ (Anthonisse and Tijms, 1977). Rate of convergence of an indecomposable Markov chain with transition matrices from $S$ is upper bounded by $(1 - \gamma)^{|S_H|}$. If the transition matrices are decomposable, we can perform similar analysis as discussed in section 3.1.2 by considering the convergence rate of each communicating class.

Now, we look at the convergence rate of the model when the network structure and the concept structure change over time. That is $P_n$ and $H_n$ are inhomogeneous. We assume that $P_n \in S_P$ and $H_n \in S_H$ where $S_P$ is a finite set of social structure matrices and $S_H$ is a finite set of concept structures. In other words, at each time step, people’s network structure takes the form of a stochastic matrix from the finite set $S_P$. Similarly, for $S_H$.

**Proposition 24.** Let $R_P$ and $R_H$ be the convergence rates of $P_n$ and $H_n$, respectively. Then the model converges to its stationary distribution with a rate of at least $R = \min\{R_P, R_H\}.$

**Collective Intelligence**
Belief evolution over dynamic, homophily-based networks

Results in the previous section assume that network structures are either static or change at random. However, network structures in society, especially in terms of who we communicate with, are affected by our beliefs (Liu et al., 2018; McPherson et al., 2001; Murase et al., 2019). For example, people may be more likely to talk with people whose beliefs are more similar to their own, either because of consistency of beliefs in a geographic region (Čepić and Tonković, 2020; Khanam et al., 2020), or through active selection of partners. Because beliefs change based on who one talks with, networks that are based on homophily may be dynamic. In this section, we analyze belief evolution for societies whose structures are governed by homophily.

Given people’s initial priors $M$ we create the network structure of people and cognitive structure of concepts based on belief similarity. Specifically, we construct the homophily network structure by linking people whose beliefs are sufficiently similar. Let $M$ be the matrix representing beliefs of individuals at time $n$. Further $S: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^+$ is a function that measures divergence between to vectors, where $S(v, u) = 0$ indicates $v = u$. Then for a given similarity threshold $\epsilon_\beta>0$, individuals $a_i$ and $a_j$ are linked, i.e. $p_{ij} > 0$, if $S(p_i, p_j) < \epsilon_\beta$ where $p_i$, $p_j$ are the row vectors in $M_n$ corresponding to $a_i$ and $a_j$; otherwise $p_{ij} = 0$.

Similarly, we construct the homophily concept structure by linking concepts that are held to similar degrees across people. In particular, let $M_c$ be the column normalization of $M$. Then for a given similarity threshold $\epsilon_\alpha>0$, concepts $\beta_k$ and $\beta_l$ are linked, i.e. $h_{kl} > 0$ if $S(h_k, h_l) < \epsilon_\alpha$ where $h_k$, $h_l$ are the column vectors of $M_n$ corresponding to $\beta_k$ and $\beta_l$; otherwise, $h_{kl} = 0$. In this section, we measure similarity of beliefs between pairs of people, and of degrees between pairs of concepts via Kullback-Leibler (KL) divergence. In addition to being a natural measure of divergence between beliefs, KL divergence is asymmetric, which means that our network and concept structures are not restricted to be symmetric.

**Definition 25.** [KL divergence] Let $A = (a_{ik})$ be a row stochastic matrix. Define KL divergence between two discrete probability distributions, $p$ and $q$, in $\mathbb{R}^K$,

$$KL(p, q) = \sum_{k \in K} p_k \log \left( \frac{p_k}{q_k} \right).$$

If two individuals or concepts are sufficiently similar, they will be linked. Next we calculate the strength of the links as a relative divergence. In particular, the strength of the link is related to their divergence relative to other linked individuals or concepts by the softmax function.

**Definition 26.** [Softmax function] For a given vector, $a = (a_1, ..., a_n) \in \mathbb{R}^n$ and a parameter $\beta$, the softmax function $\sigma$ of $a$ is, $\sigma(a) = e^{a\beta} / \sum_{j} e^{a_j\beta}$.

We define the weights of the links between individuals as follows: Let $S_{ij} = S(p_i, p_j)$ be the similarity of beliefs between individuals $a_i$ and $a_j$ and $S_t = \{S_1, S_2, ..., S_n\}$. We define $w_{ij} = \sigma(S_{ij})$, where $w_{ij}$ be the vector with weights of the probabilistic links from $a_i$ to $a_j$, $\forall j \in \{1, 2, ..., r\}$. Similarly we define the weights of the links for concept structure using Softmax function.

We now introduce the homophily-based model, which at each time step adapts its structure on $P_{n+1}$ and $H_{n+1}$ based on $M_n$,

$$Q_{n+1}(P_t, H_t, M) = \prod_{t=0}^{t=n+1} P_t \cdot M \cdot H_t = P_{n+1}H_{n+1}$$

where $M_n$ is the matrix representing beliefs of individuals, $j \in \{1, 2, ..., r\}$, in concepts, $k \in \{1, 2, ..., s\}$, at time $n$ and $P_{n+1}$ is the network structure matrix and $H_{n+1}$ is the concept structure derived from $M_n$ as described above.

One question we may ask is whether the dynamic nature of the homophily structures yield interesting changes in the asymptotic structure of the society. We have seen from previous results that as long as one of the network or concept structures is indecomposable, the long run behavior is that everyone converges to a single group with the same beliefs. We now prove a lower bound on the number of groups of beliefs for homophily-based dynamic structures, which shows the same does not hold.

**Definition 27.** Let $V = \{v_1, ..., v_k\} \subset \mathbb{R}^n$ be a set of vectors. Given $\epsilon > 0$, an $\epsilon$-KL cluster over $V$ is defined to be a subset $V \subset V$ such that: for any $v_i \in V$, $KL(Conv(V_{-i}), v_i) < \epsilon$ holds, and for any $v_j \notin V$, $KL(Conv(V), v_j) \geq \epsilon$ holds, where $V_{-i}$ represents omitting $v_i$ in $V$, and Conv($) represents the convex hull.

**Theorem 28.** Let $M$ be the initial belief matrix and $\epsilon_\rho$ be the threshold of network structure. Assume the concept structure $H$ is the identity. Then the number of groups in network
structure (number of communicating classes in $P_t$) is bounded below by the number of $\epsilon_p$-KL clusters over row vectors of $M$. Similarly, assume $P$ is identity, then the number of groups in concept structure (number of communicating classes in $H_t$) is bounded below by the number of $\epsilon_h$-KL clusters over column vectors of $M$.

We first describe an algorithm to construct $\epsilon_p$-KL clusters, the proof then follows along.

- **Step 0** Each row of $M$ (representing belief of a person) can be realized as a point $p_l^i \in \mathbb{R}^t$. View each point as a vertex (representing a person) to obtain $G_{P_t}$. Note: Let $V$ be the vertex set of a connected component of $G_{P_t}$, and $a_l \in V$ be a person belongs to this group. Then since $a_l$’s belief will be updated as a linear combination of concepts in $V$, the point $p_l^0$ in $\mathbb{R}^t$ representing $a_l$ can only move to a new point $p_l^1$ in the convex hull $\operatorname{Conv}(V)$ of $V$.

- **Step 1** For a pair of vertices $a_l$ and $a_l'$, add an edge $e_{ij}$ if $\text{KL}(p_l^0, p_l'^0) < \epsilon_p$ to obtain $G_{P_t}$. Here, three subgroups has emerged: groups $a_1$, $a_2$ and the isolated individual $a_3$. We can see that people in each subgroup has their unique belief distribution.

- **Step 2** For each pair of connected components $V_1$ and $V_2$ of $G_{P_t}$, if $B_{KL}(\operatorname{Conv}(V_1)) \cap \operatorname{Conv}(V_2) \neq \emptyset$ or $B_{KL}(\operatorname{Conv}(V_2)) \cap \operatorname{Conv}(V_1) \neq \emptyset$, then add an edge from a person $a_l$ in $V_1$ to $a_l'$ in $V_2$ to obtain $G_{P_t}$.

- **Repeat step 2** until connected components of $G_{P_t}$ stabilize, denote the converged graph by $G_{P_t}$.

It is clear from the above construction that each vertex set of a connected component of $G_{P_t}$ forms a $\epsilon_{P_t}$-KL cluster over rows of $M$. On the other hand, communication can only happen between people within the same connected component of $G_{P_t}$ for any choice of $t$. Indeed, if $a_l$ and $a_l'$ communicate time $t^*$, i.e. $\text{KL}(p_l^t, p_l'^t) < \epsilon$ or $\text{KL}(p_l^t, p_l'^t) < \epsilon$, the connected components containing $a_l$ and $a_l'$ will be connected for any $t > t^*$. Hence, the number of groups in network structure (number of communicating classes in $P_t$) is bounded below by the number $\epsilon_p$-KL clusters over row vectors of $M$. Thus, Theorem 28 holds. Even though we use KL divergence as a natural choice, above results hold for any divergence.

Even though $P_t$ and $H_t$ change over time in the homophily model, numerical simulations show that the model converges to its stationary distribution after few steps. Using this framework we can illustrate various behavior including isolated individuals, emergence of subgroups, and evolving into a society with a homogeneous belief system. Example 29 below shows how society evolves into sub communities with people in the same group having the same belief distribution, and that isolated individuals are also possible.

**Example 29.** Consider a network with $\epsilon_p = 0.3$, $\epsilon_h = 0.25$ and $M = \begin{pmatrix} 0.263 & 0.472 & 0.084 & 0.181 \\ 0.06 & 0.103 & 0.683 & 0.154 \\ 0.029 & 0.136 & 0.479 & 0.357 \\ 0.252 & 0.463 & 0.08 & 0.204 \end{pmatrix}$. Using the homophily-based framework developed above we get,

$$P_t = \begin{pmatrix} 0.5 & 0 & 0 & 0.5 \\ 0 & 0.532 & 0.468 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 \end{pmatrix} \text{. For all } t \geq 2,$$

and for all $t \geq 3$,

$$H_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0 \end{pmatrix} \text{ and } Q_t = \begin{pmatrix} 0.258 & 0.468 & 0.095 & 0.179 \\ 0.044 & 0.12 & 0.29 & 0.546 \\ 0.025 & 0.176 & 0.046 & 0.088 \\ 0.258 & 0.468 & 0.095 & 0.179 \end{pmatrix} \text{.}$$

All the matrices are rounded up to 3 decimal places. Notice that when $t = 1$, $H$ does not have any links. At $t = 5$, $P$ and $H$ have stabilized to stationary distribution $\lim_{n \to \infty} P_n$ and $\lim_{n \to \infty} H_n$, respectively. $Q_4$ shows the limiting belief distribution of the society, $\lim_{n \to \infty} Q_n$. Here, three subgroups has emerged: groups $a_1$, $a_3$ and $a_5$, and the isolated individual $a_2$. It is intuitive that when $\epsilon_p$ or $\epsilon_h$ is increased enough while the number of people and number of belief are fixed, the society will display a homogeneous belief distribution in the long run. Example 30 illustrates this scenario.

**Example 30.** Next we consider the same $M$ and $\epsilon_p$, as in 29 and let $\epsilon_h = 0.4$. Then the society stabilizes to a unique belief distribution. In particular, for all $t \geq 4$:

$$Q_t = \begin{pmatrix} 0.295 & 0.186 & 0.258 & 0.262 \\ 0.295 & 0.186 & 0.258 & 0.262 \\ 0.295 & 0.186 & 0.258 & 0.262 \\ 0.295 & 0.186 & 0.258 & 0.262 \end{pmatrix} \text{.}$$

Further examples below illustrates the evolution of homophily networks over time. Namely, Examples 31, 32 and 33 show how each person’s beliefs evolves with time, for different initial $M$ matrices and threshold values. In each example we consider three concepts $h_1$, $h_2$, $h_3$. The concept space is represented by an equilateral triangle with vertices $h_1 = (1, 0, 0)$, $h_2 = (0, 1, 0)$, $h_3 = (0, 0, 1)$. Each person’s belief at time $t$ is denoted by a point inside the triangle. All the points $x \in \mathbb{R}^3 : \text{KL}(p_t, x) < \epsilon_p$ for each person $a_t$ (referred as KL regions), are represented by the coloured regions, at every time step until the model converges. At any time step, if a person $a_t$ is in the KL region of another person $a_t'$, then $a_t$ creates a communication link with $a_t'$, represented by a line connecting the
two corresponding points. (These are unidirectional links, however we show them by a line). Observe that these regions evolve with time. 31 illustrates that the creation as well as destroying of network links are possible. Moreover, the changes in threshold parameters changes the limiting behavior. In 32 the society converge to two groups. However in 32 where we increase $\epsilon_h$ while keeping everything else the same, society converge to one group.

Example 31. We consider five people $i = 1, \ldots, 5$ and three concepts $h_1, h_2, h_3$. Initial $M$ is

\[
\begin{bmatrix}
0.348 & 0.039 & 0.6132 \\
0.321 & 0.609 & 0.07 \\
0.884 & 0.083 & 0.033 \\
0.082 & 0.185 & 0.733 \\
0.372 & 0.281 & 0.347
\end{bmatrix},
\]

$\epsilon_p = 0.3$ and $\epsilon_h = 0.2$. At each time step, each individual’s KL region changes, leading to creating new communication links or destroying existing ones. We observe that at $t = 5$, the society stabilizes and persons $i = 1$ and $2$ become isolated while the others converge to one subgroup (Figure 2).

Example 32. In this example, we consider four people $i = 1, \ldots, 4$ and three concepts. Initial $M$ is

\[
\begin{bmatrix}
0.489 & 0.104 & 0.407 \\
0.033 & 0.712 & 0.255 \\
0.543 & 0.182 & 0.275 \\
0.248 & 0.375 & 0.3776
\end{bmatrix}, \quad \epsilon_p = 0.3 \text{ and } \epsilon_h = 0.05.
\]

We observe that, at $t = 3$ society stabilizes into two subgroups. This example shows that, as time evolves existing links can be destroyed as well (Figure 3).

Example 33. Now we consider $M$ and $\epsilon_p$ as in Example 32 and let $\epsilon_h = 0.5$. We observe that, no links will be destroyed and at $t = 3$ society stabilizes to one stationary belief distribution (Figure 4). Observe that this example clearly illustrates the fact that the structure of the concept space affect

![Figure 2](image-url)

*Figure 2.* Each colored point inside a triangle represents the belief of a person $i$ at time $t$. The corresponding colored region represents the KL region of the person $i$. A line that connects two points denotes that the corresponding two people are communicating.

![Figure 3](image-url)

*Figure 3.* Each colored point inside a triangle represents the belief of a person $i$ at time $t$. The corresponding colored region represents the KL region of the person $i$. A line that connects two points denotes that the corresponding two people are communicating.
the long term dynamics of belief distribution in a society, just as the social structure.

Discussion

We presented a mathematical model of that allows for transmission of beliefs over a set of concepts both across people (horizontal) and across time (vertical). The model assumes structures over both individuals and concepts. Individuals’ beliefs about a particular concept can change either because they are connected to an individual with different beliefs or because of a change in beliefs about a related concept. We analyzed three cases: static social network and concept structures, social network and concept structures that change at random over time, and structures that vary dynamically based on homophily.

For static and randomly changing networks, we proved that if indecomposibility is satisfied by the initial (collection of) structures, then individuals in society will converge to a single group with the same beliefs. In the case of dynamically changing networks, we find a sufficient condition for heterogeneity to occur. We also provided lower bounds for the rate of convergence of the model for both static and changing networks. Our results align with previous studies showing rates of convergence slow with multidimensional transmission (Page et al., 2007). For network structures that dynamically change based on homophily, we find that the society could either converge to a homogeneous distribution or sub groups with same beliefs and or to isolated individuals, based on a threshold on divergence between people and between beliefs. We proved conditions under which the lower bound on the number of groups is greater than one, thus identifying sufficient conditions under which individuals within society will converge to more than one group characterized by different beliefs.

Prior analyses of horizontal transmission have investigated richer social network structures, but have not considered learners who maintain distributions of beliefs. This research has focused on rate of transmission as a function of the connectivity pattern in the graph. Transmission is assumed to occur by copying a random neighbor in the graph. For example, small world network structures (Albert and Barabási, 2002) yield rapid transmission to a large proportion of the network due to the short average minimal distance between individuals. Thus, it does not allow for the possibility of polarization.

Our findings differ from prior analyses of vertical transmission which consider static network structures and chains of individuals passing beliefs via random selection of data unidirectionally (Griffiths and Kalish, 2007) which show that convergence to a stationary distribution. This analysis holds for cases where individuals do not receive information from the world, and for cases where they receive data from both their predecessor and the world (Griffiths and Kalish, 2005). Across these cases, individuals in society, after long enough, all hold the same beliefs up to some variance that depends on the amount of data sampled from the world.

For example, (Whalen and Griffiths, 2017) considered vertical transmission of languages together with social structure. In their model, at each timepoint, a random learner was paired with a random neighbor and heard their language, updating their own language probabilistically based on their prior and that observation. The primary findings were that the distribution of languages over the society converged to the prior and that the degree to which neighbors in the graph spoke the same language depended on the social structure. Their study differed from ours in that they assumed each individual spoke only one language at a time, rather than maintaining a distribution and that individuals updated their language based on Bayesian inference. In contrast, we analyzed learners who maintained a distribution over beliefs and integrated information from prior timesteps with neighbors’ evidence based on information integration theory (Cohen et al., 1980; Frey and Kinnear, 1980). Most important, though, by allowing for both networks of individuals and concepts to adapt, we enable the potential emergence of heterogeneity in beliefs through homophily.

Figure 4. Each colored point inside a triangle represents the belief of a person at time . The corresponding colored region represents the KL region of the person . A line that connects two points denotes that the corresponding two people are communicating.
Our results suggest that homophily based networks, which dynamically change to connect people with similar beliefs, yield stable heterogeneity; however, simpler arrangements in which changes in network structure over time are not related to beliefs do not. An implication of this work is to focus attention on homophily as a critical component in shaping stable, long term differences in beliefs that define communities.

Evolution of beliefs is a type of collective learning in the absence of meaningful feedback on any ground truth (Almaatouq et al., 2020). Our model illustrates the importance of looking at vertical and horizontal transmission together: from a horizontal transmission perspective, any connected group of people converges to a society with a single belief distribution; while from a vertical transmission perspective, any connected structure leads to homogeneity convergence. When changes in horizontal structure accumulate over time because of homophily, we find stable heterogeneity. Collective intelligence requires differences in beliefs across individuals (Bednar et al., 2010; March, 1991) and is enabled by homophily. However, collective intelligence is endangered by extremes of homophily in which one only talks with those of like beliefs.

There remain a number of interesting open directions for future work including the death and birth of people and concepts, alternative models of transmission between neighbors, the possibility that people may obtain information from the environment, and the potential for dishonest actors who inject false information. Experimental or empirical work could attempt to calibrate our models to behavioral data which could produce more realistic models of the horizontal and vertical evolution of beliefs and potentially bound rates of convergence. Finally, our framework could be used to compare how the variation in the concept structure influences rates of convergence and possible to investigate the extent to which allocating concepts into disciplines impedes learning.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in part by DARPA grant HR00112020039, NSF MRI 1828528, and NSF Inspire 1549981 to PS.

ORCID iD

Pushpi Paranamana © https://orcid.org/0000-0003-0497-0904

Supplemental Material

Supplemental material for this article is available online.

Notes

1. We use Markov chain and corresponding transition matrix interchangeably, when there is no confusion.
2. All proofs are included in the Supplemental Material.
3. (Wolfowitz, 1963) provides an algorithm to determine if every product in a given set of matrices is SIA, in a bounded number of arithmetic operations.
4. Refer to Supplemental Material A.1 for a detailed graph theoretic interpretation of Markov chains.
5. See proof in Supplemental Material C.

References

Abrams DM, Yaple HA and Wiener RJ (2011) Dynamics of social group competition: modeling the decline of religious affiliation. Physical Review Letters 107(8):088701.
Albert R and Barabási A-L, (2002) Statistical mechanics of complex networks. Reviews of modern physics 74(1):47–97.
Almaatouq A, Noriega-Campero A, Alothabib A, et al. (2020) Adaptive social networks promote the wisdom of crowds. Proceedings of the National Academy of Sciences 117(21):11379–11386.
Anthonisse JM and Tijms H (1977) Exponential convergence of products of stochastic matrices. Journal of Mathematical Analysis and Applications 59(2):360–364.
Bednar J, Bramson A, Jones-Rooy A, et al. (2010) Emergent cultural signatures and persistent diversity: A model of conformity and consistency. Rationality and Society 22(4):407–444.
Boccaletti S, Latora V, Moreno Y, et al. (2006) Complex networks: Structure and dynamics. Physics Reports 424(4–5):175–308.
Broomell SB and Budescu DV (2009) Why are experts correlated? decomposing correlations between judges. Psychometrika 74(3):531–553.
Candia C, Jara-Figueroa C, Rodriguez-Sickert C, et al. (2019) The universal decay of collective memory and attention. Nature Human Behaviour 3(1):82–91.
Castellano C (2012) Social influence and the dynamics of opinions: the approach of statistical physics. Managerial and Decision Economics 33(5–6):311–321.
Cepić D and Tonković Ž (2020) How social ties transcend class boundaries? network variability as tool for exploring occupational homophily. Social Networks 62:33–42.
Chater N, Tenenbaum JB and Yuille A (2006) Probabilistic Models of Cognition: Conceptual Foundations. Trends Cogn Sci 10:287–291.
Chevalier P-Y, Gusev VV, Jungers RM, et al. (2017) Sets of Stochastic Matrices with Converging Products: Bounds and Complexity. arXiv preprint arXiv:1712.02614.
Cohen JB, Miniard PW and Dickson PR (1980) Information Integration: An Information Processing Perspective. Atlanta, GA: ACR North American Advances.

Delvenne J-C, Lambiotte R and Rocha LE (2015) Diffusion on networked systems is a question of time or structure. Nature Communications 6(1):1–10.

Feller W (1957) An Introduction to Probability Theory and its Applications: 2d Ed. Hoboken, NJ: Wiley.

Frey CJ and Kinnear TC (1980) An Introduction to Probability Theory and its Applications Human Behaviour. Davis, CA: Janko Gravner. Chapter, 13:151–160.

Gravner J (2010) Lecture notes for introductory probability. Davis, CA: Janko Gravner.

Grifiths TL and Kalish ML (2005) A bayesian view of language evolution by iterated learning. In Proceedings of the Annual Meeting of the Cognitive Science Society, 27. California: Escholarship.

Grifiths TL and Kalish ML (2007) Language evolution by iterated learning with bayesian agents. Cognitive Science 31(3):441–480.

Huang C-Y, Wen T-H, Fu Y-H, et al. (2019) Epirank: Modeling homophily in social networks. Scientific Reports 9(1):1–15.

Keuschmig M and Ganser C (2017) Crowd wisdom relies on agents’ ability in small groups with a voting aggregation rule. Management Science 63(3):818–828.

Khanam KZ, Srivastava G and Mago V (2020) The Homophily Principle in Social Network Analysis. arXiv preprint arXiv: 2008.10383.

Kirby S, Dowman M and Griffiths TL (2007) Innateness and culture in the evolution of language. Proceedings of the National Academy of Sciences 104(12):5241–5245.

Li A, Cornelius SP, Liu Y-Y, et al. (2017) The fundamental advantages of temporal networks. Science 358(6366):1042–1046.

Liu Y, Li L, Wang H, et al. (2018) The competition of homophily and popularity in growing and evolving social networks. Scientific Reports 8(1):1–15.

March JG (1991) Exploration and exploitation in organizational learning. Organization Science 2(1):71–87.

McPherson M, Smith-Lovin L and Cook JM (2001) Birds of a feather: Homophily in social networks. Annual Review of Sociology 27(1):415–444.

Moreno Y, Nekovee M and Pacheco AF (2004) Dynamics of rumor spreading in complex networks. Physical Review E 69(6):066130.

Murase Y, Jo H-H, Török J, et al. (2019) Structural transition in social networks: The role of homophily. Scientific Reports 9(1):1–8.

Newman ME (2003) The structure and function of complex networks. SIAM Review 45(2):167–256.

Novaes Tump A, Wolf M, Krause J, et al. (2018) Individuals fail to reap the collective benefits of diversity because of over-reliance on personal information. Journal of the Royal Society Interface 15(142):20180155.

Page SE, Sander LM and Schneider-Mizell CM (2007) Conformity and dissonance in generalized voter models. Journal of Statistical Physics 128(6):1279–1287.

Pastor-Satorras R, Castellano C, Van Mieghem P, et al. (2015) Epidemic processes in complex networks. Reviews of modern physics 87(3):925–979.

Pishro-Nik H (2016) Introduction to Probability, Statistics, and Random Processes. Blue Bell, PA: Kappa Books.

Quattrociocchi W, Caldarelli G and Scala A (2014) Opinion dynamics on interacting networks: media competition and social influence. Scientific Reports 4:4938.

Roberts G and Fedzechkina M (2018) Social biases modulate the loss of redundant forms in the cultural evolution of language. Cognition 171:194–201.

Rosenthal JS (1995) Minorization conditions and convergence rates for markov chain monte carlo. Journal of the American Statistical Association 90(430):558–566.

Schatzman M and Schatzman M (2002) Innateness and the influence. Trends in Cognitive Sciences 21(7):522–530.

Wang F, Wang H, Xu K, et al. (2013) Characterizing information diffusion in online social networks with linear diffusive model. In 2013 IEEE 33rd International Conference on Distributed Computing Systems, pages 307–316. Manhattan, NY: IEEE.

Whalen A and Griffiths TL (2017) Adding population structure to models of language evolution by iterated learning. Journal of Mathematical Psychology 76:1–6.

Wolfowitz J (1963) Products of indecomposable, aperiodic, stochastic matrices. Proceedings of the American Mathematical Society 14(5):733–737.

Zhan X-X, Hanjalic A and Wang H (2019) Information diffusion backbones in temporal networks. Scientific Reports 9(1):1–12.

Zhou B, Pei S, Muchnik L, et al. (2020) Realistic modelling of information spread using peer-to-peer diffusion patterns. Nature Human Behaviour 4, 1–10.