

Research Article

Paths to Polarization: How Extreme Views, Miscommunication, and Random Chance Drive Opinion Dynamics

Matthew A. Turner  and **Paul E. Smaldino**

Cognitive and Information Sciences, University of California, Merced, USA

Correspondence should be addressed to Matthew A. Turner; mturner8@ucmerced.edu

Received 15 May 2018; Accepted 13 September 2018; Published 1 November 2018

Guest Editor: Martin Hilbert

Copyright © 2018 Matthew A. Turner and Paul E. Smaldino. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Understanding the social conditions that tend to increase or decrease polarization is important for many reasons. We study a network-structured agent-based model of opinion dynamics, extending a model previously introduced by Flache and Macy (2011), who found that polarization appeared to increase with the introduction of long-range ties but decrease with the number of salient opinions, which they called the population's "cultural complexity." We find the following. First, polarization is strongly path dependent and sensitive to stochastic variation. Second, polarization depends strongly on the initial distribution of opinions in the population. In the absence of extremists, polarization may be mitigated. Third, noisy communication can drive a population toward more extreme opinions and even cause acute polarization. Finally, the apparent reduction in polarization under increased "cultural complexity" arises via a particular property of the polarization measurement, under which a population containing a wider diversity of extreme views is deemed less polarized. This work has implications for understanding the population dynamics of beliefs, opinions, and polarization as well as broader implications for the analysis of agent-based models of social phenomena.

1. Introduction

Diversity of opinions in a community is often difficult to maintain. Iterative exposure, norm enforcement, and psychological biases for conformity can drive consensus within a group [1–6]. On the other hand, in-group bias, outgroup aversion, and the tendency to further differentiate ourselves from those deemed different may lead to the emergence of strong intergroup differences [7–15]. Such differences can lead to polarization in opinions under certain conditions. Understanding the social conditions that tend to increase or decrease polarization is important for many reasons. Primary among these is that a functioning democratic society depends on clear communication among the citizenry, which is impeded by the mismatch in norms, the differential interpretation of facts, and the dehumanization that polarization can engender (see Pew Research Center [16] for a current analysis of these dynamics in the United States). The maintenance of social differences in the form of cliques and clubs may be inevitable, but cooperation depends on transcending differences.

We take a network theoretic approach to studying the conditions for polarization in an agent-based model of opinion dynamics. Empirical research on the population dynamics of opinions is challenging and must be supplemented by formal modeling [17]. Models reduce complex systems to ones that are tractable using mathematical or computational analysis and allow for the exploration of replicate and counterfactual scenarios. Of course, the conclusions we draw from our models depend essentially on the assumptions of those models, and so caution must be taken when using model results to make inferences about empirical phenomenon. For example, Smaldino and Schank [18] analyzed models of human mate choice and showed that very different individual decision rules could be fit to almost any empirical outcome by modulating assumptions about the population structure that had been ignored in prior analyses. When considering an important phenomena such as polarization, similar caution must be exercised, as we will demonstrate.

Our analysis extends the work of Flache and Macy [19], who used a network-structured model of opinions and biased

influence (hereafter the FM model) to study polarization. Network ties in this model exist between individuals as an indicator of social influence. Like several other models of opinions and beliefs, they operationalized the well-known phenomena of *biased assimilation* [9, 11], the tendency for an individual to become more similar to those to whom they are similar and to become more distinct from those with whom they already differ. Some empirical studies support the assumption of both positive and negative biased assimilation [20, 21], e.g. Other empirical studies failed to find evidence of negative biased assimilation at work where computational studies suggested it would be [22, 23], e.g. Of course, if further empirical research turns out to invalidate that assumption, then our model conclusions must also be reexamined, as with any theoretical model [24]. Flache and Macy found that, when compared with a highly clustered population structure, the addition of long-range ties could dramatically increase polarization. When individuals were clustered into relatively isolated groups, they tended to converge to local consensus while maintaining diversity in the population at large. However, the addition of long-range ties increased exposure to substantially different opinions. Whether by attractive or repulsive forces, these long-range ties tended to drive opinions more toward their extreme values, resulting in increased polarization. Another important result was that the extent of “cultural complexity”—the number of orthogonal traits that are important to individuals in assessing their similarities and differences with others—mitigated polarization. When the number of traits was large, polarization was reduced. DellaPosta, Shi, and Macy [25] used a variant of the FM model to explain data from the General Social Survey indicating that arbitrary traits tend to become associated with polarized identity groups, leading to often-puzzling stereotypes such as “latte-drinking liberals” and “bird-hunting conservatives.”

If we take the results of Flache and Macy [19] at face value, two possible recommendations for the reduction of polarization readily emerge. First, we might try to reduce the number of long-range ties in our social network. This is made difficult due to the pervasive influence of internet social media [26, 27]. Second, we might attempt to broaden the number of domains in the public discussion, so that points of agreement are easier to discover. This is also challenging, due to the increasingly fractured media landscape in which niche interests are increasing and common knowledge diminishing [28]. However, challenging is not the same thing as impossible. We must ask then how seriously should we take these recommendations? Might there be other solutions available?

To address these questions, we perform new analyses of the FM model and reveal several additional factors influencing polarization. First, polarization is almost always a probabilistic occurrence. Even when parameter exploration appears to reveal regularities in polarization, specific outcomes are strongly path dependent. Indeed, there is often a wide range of possible outcomes even given identically repeatable starting conditions, due to stochasticity in the dynamics of interactions. This result highlights potential limits of our ability to make reliable predictions about polarization in any particular social system. Complex systems are

often stochastic, and something that increases or decreases average polarization in a simulation is not guaranteed to do so in reality. Second, resultant polarization depends strongly on the initial distribution of opinions in the population. In the absence of extremists, polarization may be mitigated. This highlights the well-known danger of extremists and suggests new routes to avoiding polarization. More broadly, we show that too much diversity of extreme opinions makes polarization more likely. Third, noisy communication can drive a population toward more extreme opinions and even cause acute polarization. Cooperation and consensus building depend on individuals finding common ground, which can be jeopardized even in the presence of unbiased error [29]. Finally, we show that the apparent reduction in polarization under increased “cultural complexity” arises via a particular property of the polarization measurement, under which a population containing a wider diversity of extreme views is deemed less polarized. Although this may often be a reasonable assumption, it highlights the need for caution in our measurement of complex social phenomena.

2. Model

2.1. Modeling Individuals and Their Opinions. Our model is an extension of one presented by Flache and Macy [19] and shares many general features with other models of opinion dynamics in structured populations [7–9, 12, 25, 30–32]. The population is modeled as a network of individuals (or agents), each of whom is defined by a vector of opinions. The size of this vector, K , is called the “cultural complexity,” and may be more descriptively explained as the number of opinions that are important to individuals in assessing their similarities and differences with others. Opinions can represent political views, religious or moral values, artistic tastes, or myriad other beliefs. The opinion of agent i on issue k ($1 \leq k \leq K$), s_{ik} , is operationalized as a real number implicitly bounded in $[-1, 1]$ by smoothing (equation (3)). In Flache and Macy’s original analysis, all opinions were initialized as random draws from the uniform distribution $U(-1, 1)$. In order to study the importance of initially extreme opinions, each initial opinion is here drawn instead from $U(-S, S)$, where $0 < S \leq 1$.

2.2. Modeling Social Influence. The aggregation of the K opinions held by an agent determines its coordinates in opinion space. We adopt the FM model’s measure of distance between agents i and j as follows:

$$d_{ij} = \frac{1}{K} \sum_{k=1}^K |s_{jk,t} - s_{ik,t}|. \quad (1)$$

Distance thus defined measures the average absolute difference across opinion coordinates. Agents are nodes in a network, with an edge between agents reflecting a relationship and an opportunity for the agents to influence one another. The magnitude and direction of that influence is characterized by the *weight* of each edge. Weights are determined by the relative opinions of the two agents, as measured by their distance, and so can change dynamically. Positive weights represent positive influence, in which agents become

closer in their opinions, while negative weights represent the tendency toward differentiation. For descriptive convenience, if two agents are connected with a positive weight, they could be considered “friends” and if the weight is negative they could be considered “enemies.” In reality, no assumptions about such clear social roles are necessary. The weight of an edge between agents i and j is given by

$$w_{ij,t+1} = 1 - d_{ij,t}. \quad (2)$$

So, if the opinions of agents i and j are separated by $d_{ij} < 1$, the agents are friends and will harmonize their opinions. If $d_{ij} > 1$, the agents are enemies and will drive each other’s opinions to more extreme levels. This weighting rule embodies the psychological phenomena of *biased assimilation*, in which similar individuals grow more similar and dissimilar individuals grow further apart after interacting [11]. This is a common assumption in models of social influence [9, 19, 33]. It should be noted that while the empirical evidence for biased assimilation is quite strong and spans almost four decades, it is less clear how coherence on various opinions or beliefs affects influence on orthogonal opinions or beliefs. The assumption in this model is that it is only average distance in opinions that matters.

At time $t + 1$, agents update their opinions by adding the average influence from all neighbor agents. For each opinion k , agent i uses the following update rule:

$$s_{ik,t+1} = s_{ik,t} + \Delta s_{ik,t} (1 - \text{sgn}(s_{ik,t})s_{ik,t}), \quad (3)$$

where

$$\Delta s_{ik,t} = \frac{1}{2N_i} \sum_{j \neq i} w_{ij,t} (s_{jk,t} - s_{ik,t}) + \varepsilon. \quad (4)$$

Here, N_i is the number of agents with which agent i shares an edge, and ε is a noise term that reflects errors in the communication of opinions. This term is in each instance drawn at random from a normal distribution with a mean of zero and a standard deviation of σ . We conceptualize updating to be the result of agents sensing the communicated opinions of neighbors. Furthermore, we conceptualize this σ as representing noise either in an agent sensing the opinions of other agents, noise in agents communicating their opinions, or both. In their original study, Flache and Macy [19] considered only scenarios without noise ($\sigma = 0$). Time in the model progressed in discrete time steps. At each time step, each agent’s opinions were updated asynchronously in random order to avoid well-known artefacts that often accompany simultaneous agent updating.

It is worth noting a few immediate consequences of these update equations. First, agents with extreme opinions in dimension k will tend to make smaller changes to those opinions because of the smoothing factor $(1 - \text{sgn}(s_{ik,t})s_{ik,t})$. In other words, extreme opinions will be harder to change. Second, there are two opposing factors that modulate the magnitude of influence between two agents. On the one hand, edge weight is maximal when agents’ opinions are very

similar. On the other hand, $\Delta s_{ik,t}$ (which Flache and Macy refer to as the “raw” state change) increases the more agents’ opinions differ, presumably because larger distances provide larger room for change, with a mathematical form drawn from psychological models of reinforcement learning [34, 35]. Influence will therefore be maximal for agents who are an intermediate distance apart in opinion space. To facilitate an intuitive understanding of dyadic interactions, we illustrate the strength of influence on agent opinions in $K = 2$ opinion space in Figure 1. We see that an agent with opinions at the origin of opinion space has only a moderate, attractive influence on other agent opinions in the opinion space. Agents at the corners of opinion space are barely influenced by a central opinion vector. When we consider the influence of an agent opinion nearer to the corner, at $\vec{s}_i = (0.9, 0.9)$, we see that there is a clear line where relationships switch from friend to enemy ($s_{j2} = s_{j1} - 0.2$). Due to the comingling of effects described above, there is a varied and nonmonotonic landscape of influence.

2.3. Measuring Polarization. There are a multitude of measures for polarization [36], and no single measure is widely agreed upon. We follow Flache and Macy [19] and define polarization at time t to be the variance of all distances between agents

$$P_t = \text{var}(d_{ij,t}). \quad (5)$$

This metric has the advantage of simple interpretation. If half of all agents are in one corner of opinion space and the other half of agents are in the opposite corner, then the population is maximally polarized. As agent opinions spread to other corners and to other regions of opinion space, polarization will decrease. One disadvantage is that more general patterns of clustering, as would be detected using various machine learning clustering algorithms, will go undetected. In the final subsection of our results, we illustrate another limitation of this metric. Nonetheless, we generally find that it is a useful and suitable operationalization for the concept of polarization.

2.4. Network Structure. Our network structures are taken from Flache and Macy’s [19] experiment 2. We begin with the connected caveman network structure introduced by Watts [37]. Specifically, we consider a network of $N = 100$ agents, grouped into 20 fully connected clusters (caves) of five agents each. These caves are arranged on a circle, and for each cave, one edge is selected at random and rewired to connect to a random agent in the cave immediately to the right of the focal cave. This network has the appearance of tight-knit communities with weak ties to neighboring communities. The connected caveman network is highly clustered, meaning that if two agents are both neighbors of another single agent, there is a high probability that those two agents are also neighbors. However, relative path length is considerably greater in a connected caveman graph than for a totally random graph.

To assess the influence of adding long-range ties, we then consider a network for which 20 additional edges are added between randomly selected pairs of agents from across the entire network (Figure 2). Long-range ties are added at $t = 2000$ to give the local communities (caves)

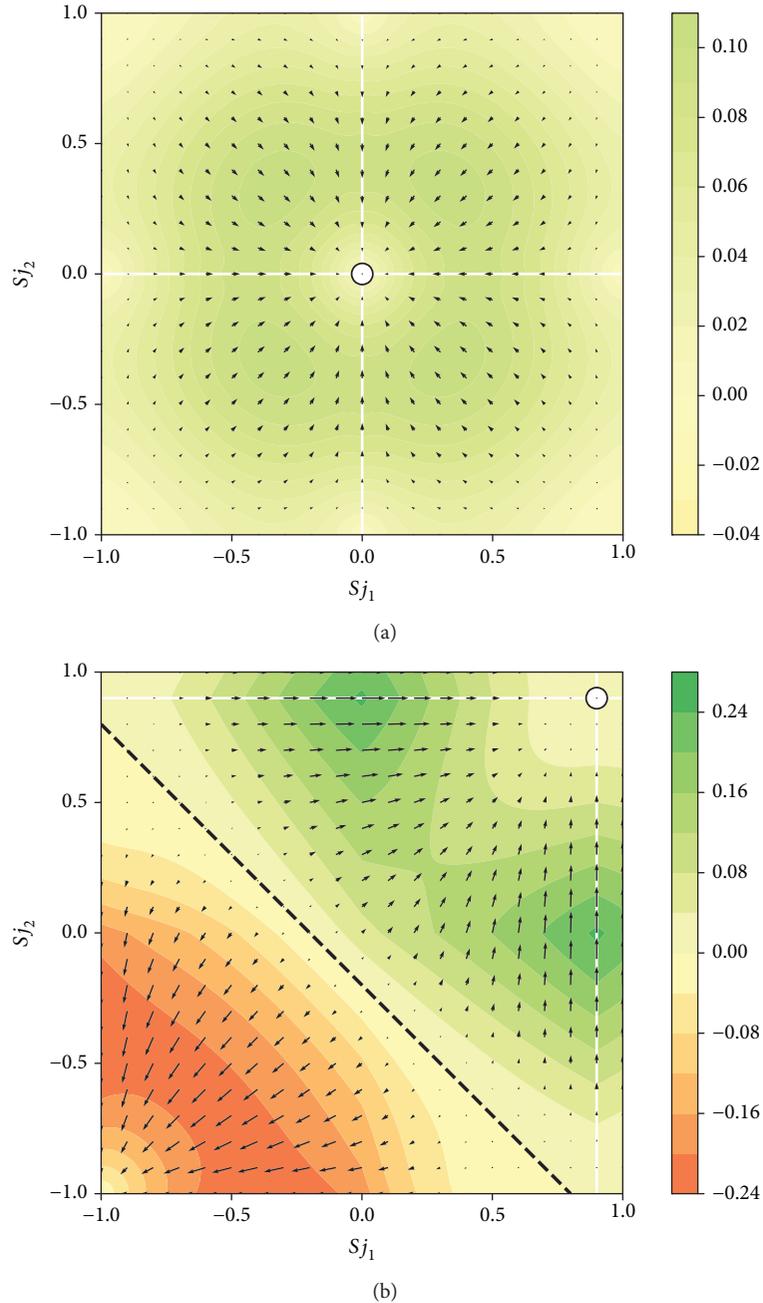


FIGURE 1: Influence by one agent on another changes depending on the location of each agent. This illustrates the influence exerted by a central agent (white circle) on another agent at different locations in opinion space.

time to yield enclaves of conformity that differ slightly from their neighboring enclaves, following Flache and Macy [19]. The long-range ties reduce the average path length of the network while retaining high clustering, yielding networks with “small-world” properties Watts [37].

Finally, as a way to control for the effect of simply adding additional ties, we also consider the connected caveman network with *short-range* ties. In this case, a randomly selected agent from each cave (who is not already connected to another cave) is connected to a random agent in the cave immediately to the right of the focal cave. Unless stated otherwise, all of our analyses were restricted to the connected caveman network

with long-range ties, as this was the network structure found by Flache and Macy [19] to maximize polarization.

2.5. Computational Experiments. Below, we present the results of our computational experiments. For all parameter combinations, we ran 100 simulations of the model, with data collected after 10^4 time steps. This was always sufficient time for the system to settle down into a relatively stable pattern (true equilibria were not always reached due to the stochasticity inherent in the model). By calculating the difference in polarization on the final time step for all simulations and finding all to be sufficiently small, we confirmed that 10^4 time

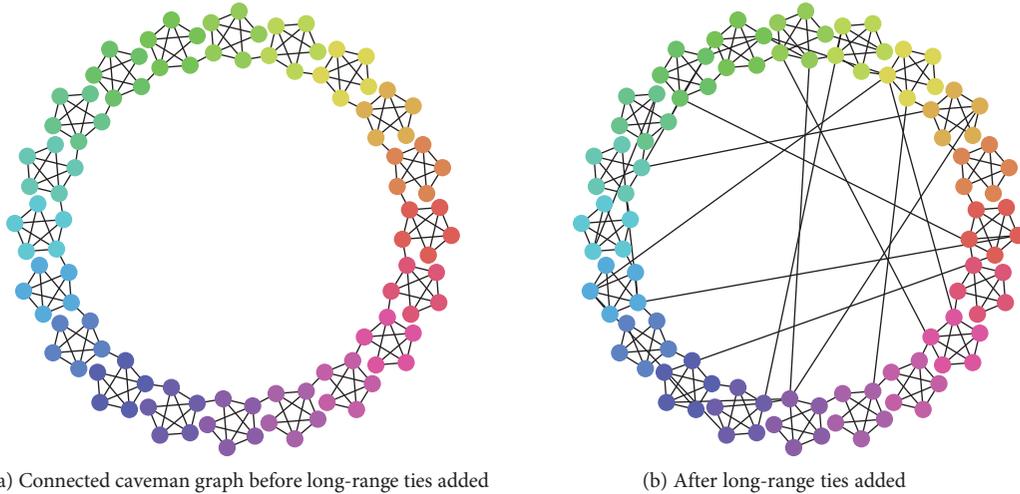


FIGURE 2: Connected caveman network with and without twenty long-range ties. Colors represent cave membership.

steps was sufficient to achieve stable behavior across all simulations. We first replicate the major result of Flache and Macy [19] that polarization increases with the addition of long-range ties but decreases with increasing cultural complexity, K . We then perform three sets of experiments:

- (1) Quantifying variation. We take a closer look at the variation among simulation runs and explore path dependence on the road to polarization
- (2) Reducing extremism. We investigate values of $S < 1$, in which the initial distribution of opinions is less extreme
- (3) Adding noise. We investigate values of $\sigma > 0$, in which communication about opinions is noisy and influence is therefore more stochastic

Unless stated otherwise, all simulations used a connected caveman network with random long-range ties, $S = 1$ and $\sigma = 0$. Model and analysis code is available on GitHub at <https://github.com/mt-digital/polarization>.

3. Results

In their original analysis of the FM model, Flache and Macy [19] found two main causes of polarization. First, random long-range ties decreased the average path length of the network and increased the average polarization of the system across trials. Second, average polarization across trials decreased with increasing cultural complexity, K . We replicated these results, as illustrated in Figure 3. The remainder of this section is dedicated to novel results. The first three subsections show results of new analyses of the original FM model. The final subsection shows our analysis of the FM model modified to include communication noise.

3.1. Polarization Is Probabilistic and Path Dependent. Averages do not carry information about variation between trials. Here, we explore that variation. Figure 4 shows the polarization for each of the individual trials averaged in Figure 3. We

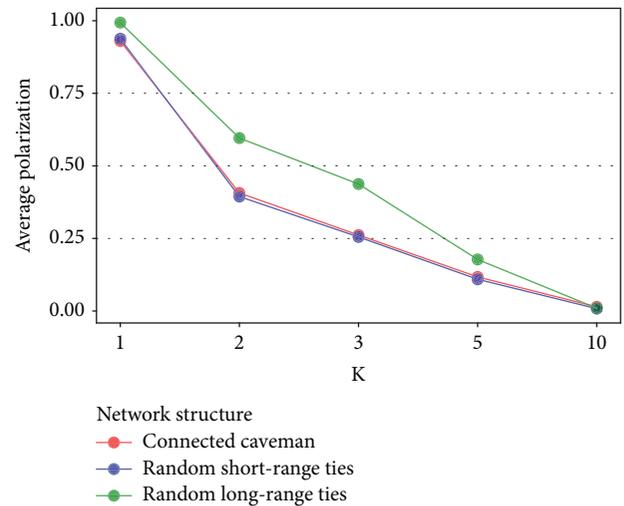


FIGURE 3: Reproduction of Figure 12(b) of Flache and Macy [19]. Average polarization decreases with K . However, as shown in subsequent figures, this does not mean trials with high polarization never obtain for large K . Average taken over 100 trials.

see a lot of variation around those averages, and that although polarization was low in all cases for large K , there are still individual trials for which polarization was high across all three network structures.

In addition to the demonstrated influence of the overall network structure, three possible sources of variation in system polarization are (1) the initial distribution of agent opinions, (2) the initial distribution of how agent opinions are clustered on the network, and (3) the update path—the order in which weights or agent opinions are updated. We performed additional analyses to investigate the contributions from each of these three factors, focusing on the initial distribution of agent opinions. We studied the nonrandom connected caveman network so as to keep network structure constant across trials, and for simplicity, we restricted this analysis to $K = 2$. Due to the nature of our polarization measure, at initialization, the system will have some nonzero

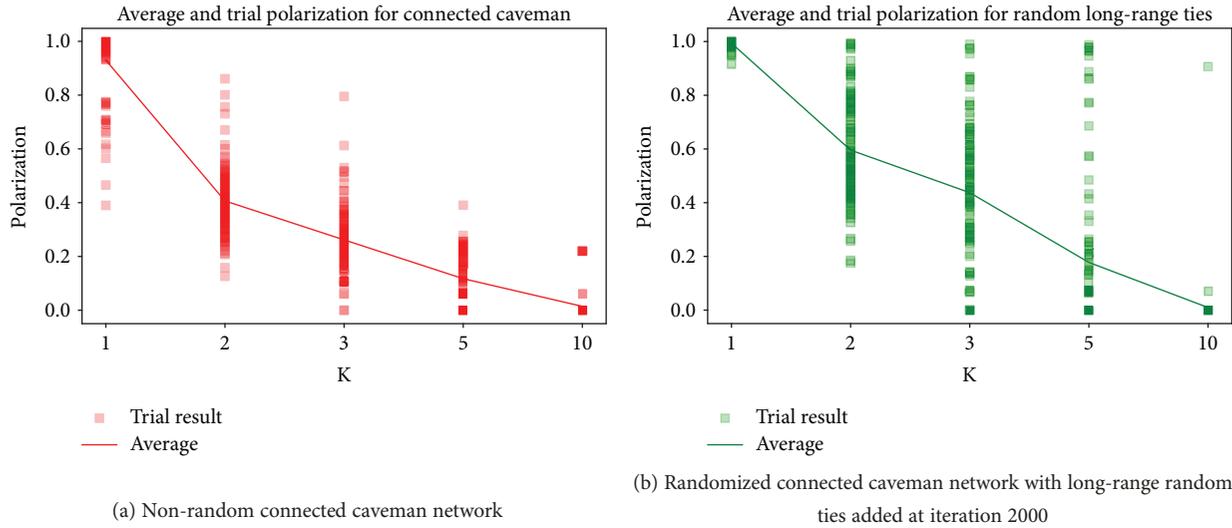


FIGURE 4: Results of individual model runs under different network conditions. The averages of these are shown in Figure 3. Even in the nonrandom connected caveman structure, there is variation in the final polarization for different values of K . Highly polarized final states may obtain even for large K . 100 trials are shown for each network condition. Solid lines indicate the average across all trials.

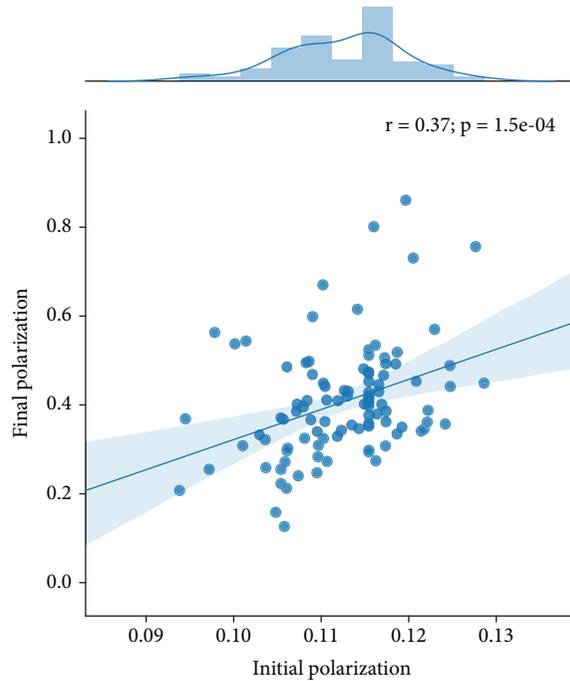


FIGURE 5: Regression of final polarization against initial polarization for $K = 2$ in the nonrandom connected caveman network configuration. Final polarizations are same as in the $K = 2$ column of Figure 4(a). 100 trials are shown. The top histogram shows the distribution of initial polarization across trials. The right histogram shows the distribution of final polarization across trials.

degree of polarization, which will vary depending on the random draws of agents' initial opinions. Over 100 trials, we compare the initial polarization of the system to the final polarization. We found a significant, if relatively small, correlation between the initial and final polarization of agent opinions, $r^2 = .137$ (Figure 5). This means that the level of initial polarization accounts for only about 14% of the variation in final polarizations. It seems, then, that initial clustering of agent opinions and the stochasticity of the update path

account for a large portion of the variability. In order to delineate the contributions of these two remaining factors to the overall variability in polarization, we considered the previously discussed simulations and ran 100 replicate trials with the initial conditions taken from the trials with the lowest and highest initial polarization. In other words, for each of two conditions, we ran replicate simulations with the exact same starting conditions between trials. Any variation in outcomes must therefore be due to stochasticity in the update

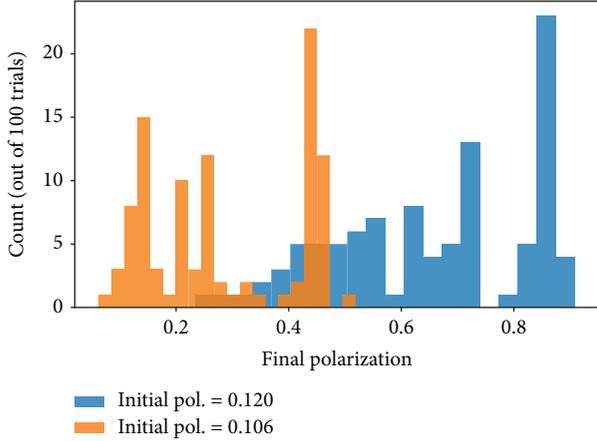


FIGURE 6: Distribution of final polarizations at $t = 10^4$ starting from initial conditions of either maximum or minimum polarization taken from the connected caveman trials with $K = 2$.

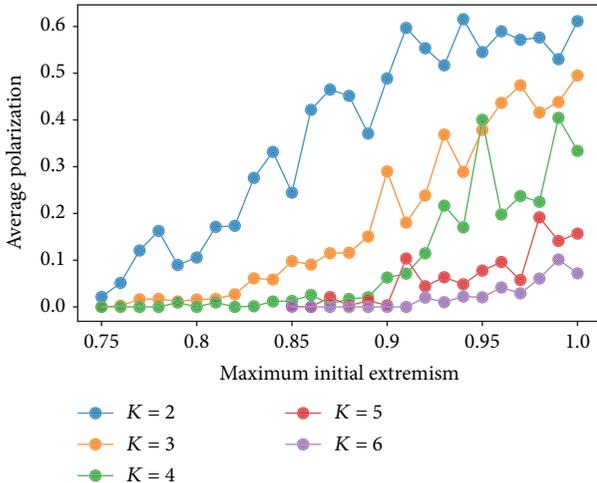


FIGURE 7: Average final polarization for different cultural complexities over maximum initial opinion magnitude, S . Averages are roughly zero for $S < 0.75$ for all cultural complexities.

paths. For example, if two opposing extremists influence a disjoint set of moderates disproportionately often, polarization will increase. The results are shown in Figure 6. Final polarization was clearly biased by the initial polarization (average final polarization across trials was 0.66 for the larger initial polarization and 0.290 for the smaller initial polarization), but showed considerable variability. In other words, a large proportion of the variation between trials was due to stochasticity not in the initial configuration of the population, but to stochasticity in the transient dynamics of agent interactions.

3.2. The Absence of Initially Extreme Opinions Reduces Polarization. Next, we extend our analysis of initial conditions further, by studying the breadth of opinions initially present in the population. Specifically, initial opinions were drawn from the uniform distribution $U(-S, S)$. Figures 7

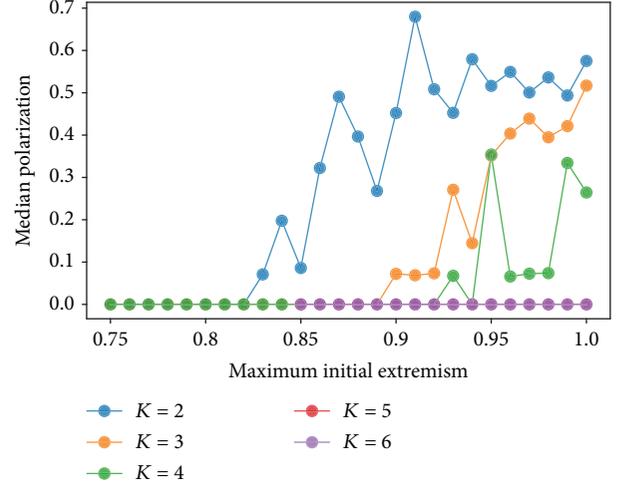


FIGURE 8: Median final polarization for different cultural complexities over maximum initial opinion magnitude, S . Median polarizations for $K = 5$ and $K = 6$ are both flat at zero; $K = 5$ data is obscured by $K = 6$.

and 8 show the mean and median polarization of the population as function of S , for $K = 2, \dots, 6$. In general, the average final polarization decreased with smaller S for all values of K . The lines are not perfectly smooth due to the large variation in outcomes described in the previous section (see Figure 9).

We again examined the within-condition variation in final polarization (Figure 9). Even when the average polarization was very small, we nevertheless saw instances of strongly polarized outcomes for $S < 1$ across all values of K . For small values of S , much more polarization occurred with small K . This further highlights the fact that initial conditions, in conjunction with the cultural complexity, bias the system towards larger or smaller levels of polarization, but do not eliminate the possibility of either conformity or extreme polarization.

3.3. The Meaning of Polarization in High-Dimensional Opinion Space. Clearly extreme positions are important in the FM model. Extremists are more stubborn (and therefore more influential) than centrists due to smoothing. Our analysis indicates that under a wide range of conditions, all opinions are likely to end up at extreme values. Indeed, the only stable states of the model are complete consensus, which can be at any point in opinion space in the absence of noise, or for all opinions to be at extreme values. This brings us back to a key result of the FM model, which is that increased cultural complexity, K , decreases polarization. Recall that polarization is measured as the variance among distances between agent opinions. To what extent is this decrease in polarization with increased cultural complexity driven by the fact that, for larger K , there are simply more “corners” (extreme opinion values) for agent opinions to settle on?

We investigated this question by comparing polarization emerging from the dynamics of the FM model with polarization that occurs when agents are artificially placed on a random vertex of the K -dimensional opinion hypercube. We found the polarization for this combinatorial condition is

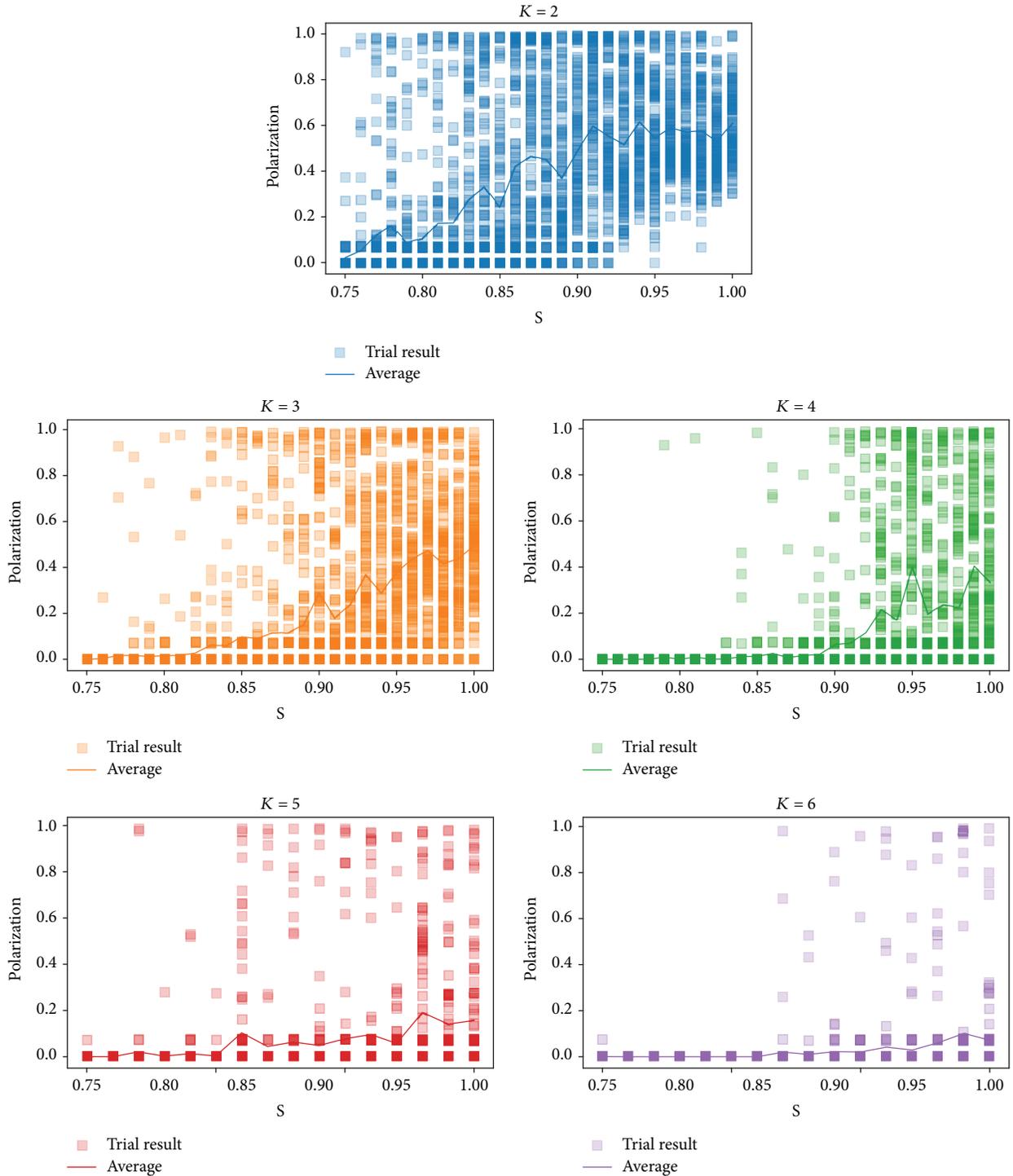


FIGURE 9: Final polarization of individual trial runs and averages from Figure 7.

$P_c \approx 1/K$ via Monte Carlo sampling with 100 agents and 1000 trials for each $K \in \{1, \dots, 12\}$. In the Appendix, we derive a formal proof that $P_c = 1/K$ exactly in the limit as $N \rightarrow \infty$.

When we compare the combinatorial result to the FM model results, we find that observed decrease in polarization with increased K follows the combinatorial results very closely (Figure 10). The connected caveman condition results

in a lower polarization, on average, than P_c for all K that we tested. The random long-range condition results in an average polarization roughly equal to P_c for $K=1$, higher average polarization than P_c from $K=2$ to $K=4$, and lower polarization for $K \geq 5$. The source of this jump from above-combinatorial to below-combinatorial is not clear, but is an interesting avenue for future work.

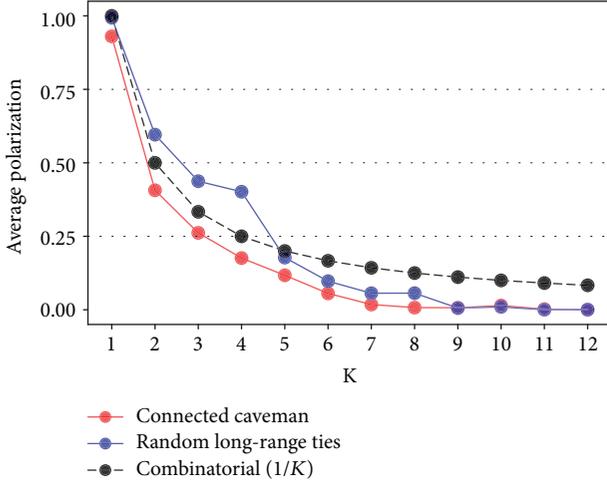


FIGURE 10: Polarization resulting from FM model simulations under connected caveman and random long-range tie conditions, compared with polarization resulting from agents arbitrarily choosing a corner of opinion space at random. Monte Carlo simulations revealed that polarization goes as $1/K$ if agents simply pick a corner at random. Random long-range and connected caveman data points are averaged from 100 trials with 10^4 iterations. Combinatorial condition data points are the average over 1000 trials and 10^4 iterations. Standard deviation around combinatorial trial averages was less than 10^{-2} .

3.4. Noisy Communication Increases Polarization, Particularly in the Absence of Initially Extreme Opinions. Up to this point, we have assumed that agents accurately express their own opinions and accurately receive information concerning the opinions of others. As this assumption is unlikely to fully hold in most cases of human interaction, it is important to assess the model’s robustness to noisy communication. To do this, we introduced random error into the opinion update equation, so that every cultural feature communication channel, for every connected dyad, was modulated by a noise term, ϵ , drawn from a normal distribution with mean 0 and standard deviation σ . Let us call σ the “noise level.” We varied the noise level from 0 to 0.2 in increments of 0.02. For each of these noise levels, we also varied S from 0.5 to 1.0 in steps of 0.05 for a total of 121 parameter pairs for each $K \in \{2, 3, 4, 5\}$. Note that we did not explicitly bound opinion components in the presence of noise. This led to us discarding 19 of the 60,500 runs due to runaway opinions that diverged to infinity, and this was only for the highest noise levels used. These discarded runs had noise levels of .18 or .2. Most parameter settings had only one discarded run if any, with one parameter setting having three discarded runs, lowering the number of samples to 97 from 100 for that parameter setting ($K=5$, $S=0.95$, and noise level=0.2). This lack of smoothing had no effect on nondivergent model runs polarization outcomes, as polarization was less than or equal to 1.0 for all.

These experiments reveal an interesting pattern of results. A sufficiently large amount of noise produced high levels of polarization for low values of S , which never produced polarization in the absence of noise.

Indeed, there appears to be a phase transition point for σ under low S , below which the system collapses to complete conformity and above which we see high levels of polarization (Figure 11). Across the values of K we tested, this threshold appeared to be around $\sigma = 0.8$, below which we never saw any polarization for low S (Figure 12). As S increases, however, the system behavior becomes less sensitive to noise, appearing to be completely insensitive to noise close to $S = 1$.

Even though polarization is rare at moderate noise levels, extremism is not. A noise level of over 0.1 was required to reliably drive the system to polarization in our simulations, but lower noise levels led to consensus around an extreme location in opinion space rather than at a most centrist position. We infer this because the average agent distance from center increases to the maximum, 1.0, with noise levels of only 0.6 (Figure 13). Thus, we obtain the interesting result that even small amounts of communication noise can move the population to extremist positions.

Figures 11 and 13 also illustrate a curious interaction between noise level, σ , and initial extremism, S . For smaller S , we observe clear phase transitions from centrist conformity to extremist conformity to polarization. For larger S , the populations responses are less clearly delineated. To help explain, we present illustrations of the spatiotemporal dynamics of the model for exemplar trials. Consider first a case of very low initial extremism, $S = 0.5$ (Figure 14). In the absence of noise, the system collapses around the center of opinion space at $t = 200$ and by $t = 3000$ has reached full consensus (Figure 14(a)). At the other extreme, under high levels of noise, $\sigma = 0.2$, agents reach a near consensus by $t = 1000$ and remain there until $t = 2000$, when random long-range ties are added. At this point, agents are exposed to individuals with very slightly different sets of opinions, and those differences are amplified by the noise, leading to repulsion. This is sufficient to jolt the system away from conformity and into opposing camps moving towards opposing corners (Figure 14(c)).

For $\sigma = 0.08$, we found most simulations end in extreme consensus. That is, all opinions were at the extremes (± 1) rather than closer to zero, but these opinions were universally shared so that final polarization was zero. One such trial is shown in Figure 14(b). This occurs because noise is sufficient to move the population toward the extremes (from which it is difficult to return to center), but agents remain sufficiently clustered so that all forces remain attractive rather than repulsive.

When initial opinions are drawn from the full range of possibilities ($S = 1$), the system always achieves some degree of polarization. Because noise only serves to increase the likelihood of extreme opinions, this condition is unaffected by noise. Typical cases are shown in Figure 15. The behavior for $t \leq 2000$ is similar in all three cases: each cave reaches a local consensus, and the network of caves reaches a stable configuration. Some of the caves find consensus values at the corners. When random ties are added, the stable configuration is broken, and agents are pulled towards one of the four corners, where some caves have already been stably established. The caves in the corners do not move. Recall that

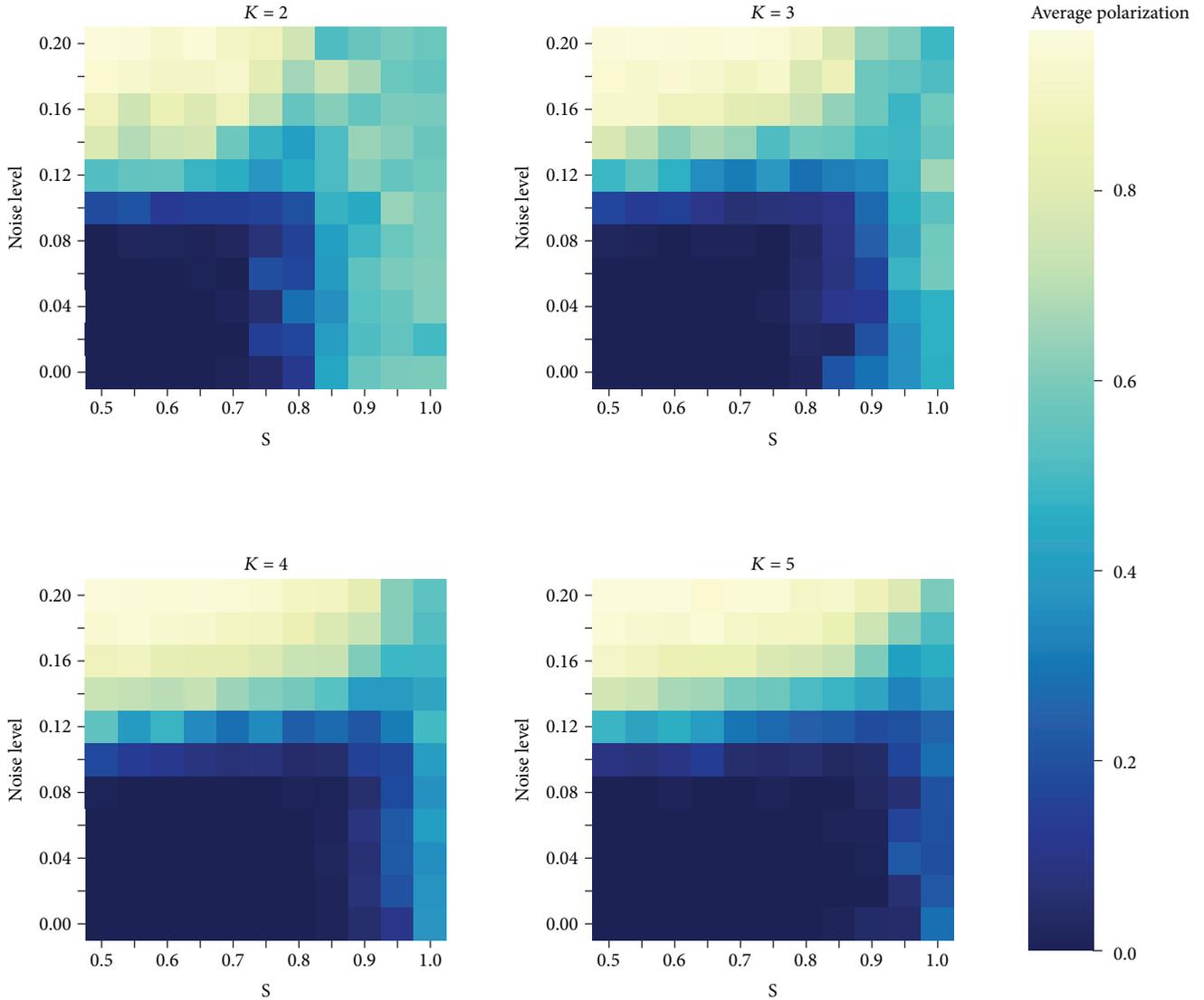


FIGURE 11: Final average polarization varies with both the width of the uniform distribution of initial opinion magnitudes and the noise level in the opinion updates. The value in each square of the heat map is the average of 100 trials.

a key assumption of the FM model is that extremist opinions influence centrist opinions more than centrists influence extremists. The noise is not strong enough to move extremists from extreme positions. In other words, in the presence of extreme opinions, network structure, not noise, dominates the dynamics. We extend the intuition to higher dimensions of opinions space using parallel coordinate plots, visualizing time series of opinion dynamics for $K = 4$ (Figure 16).

4. Discussion

Humans are the quintessential cultural species. Our instinct to learn from others is a key reason for our domination of the planet [38, 39]. An under-appreciated component of cultural learning concerns exacerbating differences and rejecting opinions when individuals are not likely to share one's current norms and beliefs. When those differences occur within a community, they can lead to discord. Many of us live in multicultural societies requiring cooperation and common

ground, and so it is natural to ask when do we expect polarization and is there anything we can do about it. Any suggestions based on our modeling efforts here should of course be compared with empirical studies. Hopefully, these results stimulate further empirical work to understand when and why polarization emerges in real-world situations. One such opportunity for future work is to connect our findings to the political science literature on polarization [40], especially in relation to communication. If agents had different roles, such as elite agents (politicians and media) and common agents, we could model the effects of ideologically biased news in political polarization [28, 41]. Our results show that in the presence of sufficiently large communication noise and small-world networks, a situation we are arguably in today, a state of polarization is the only stable state (Figures 14–16). It is interesting to consider this in light of one recent analysis suggesting that the United States Constitution was designed not just to accommodate polarization but to foster it for the sake of stability [42].

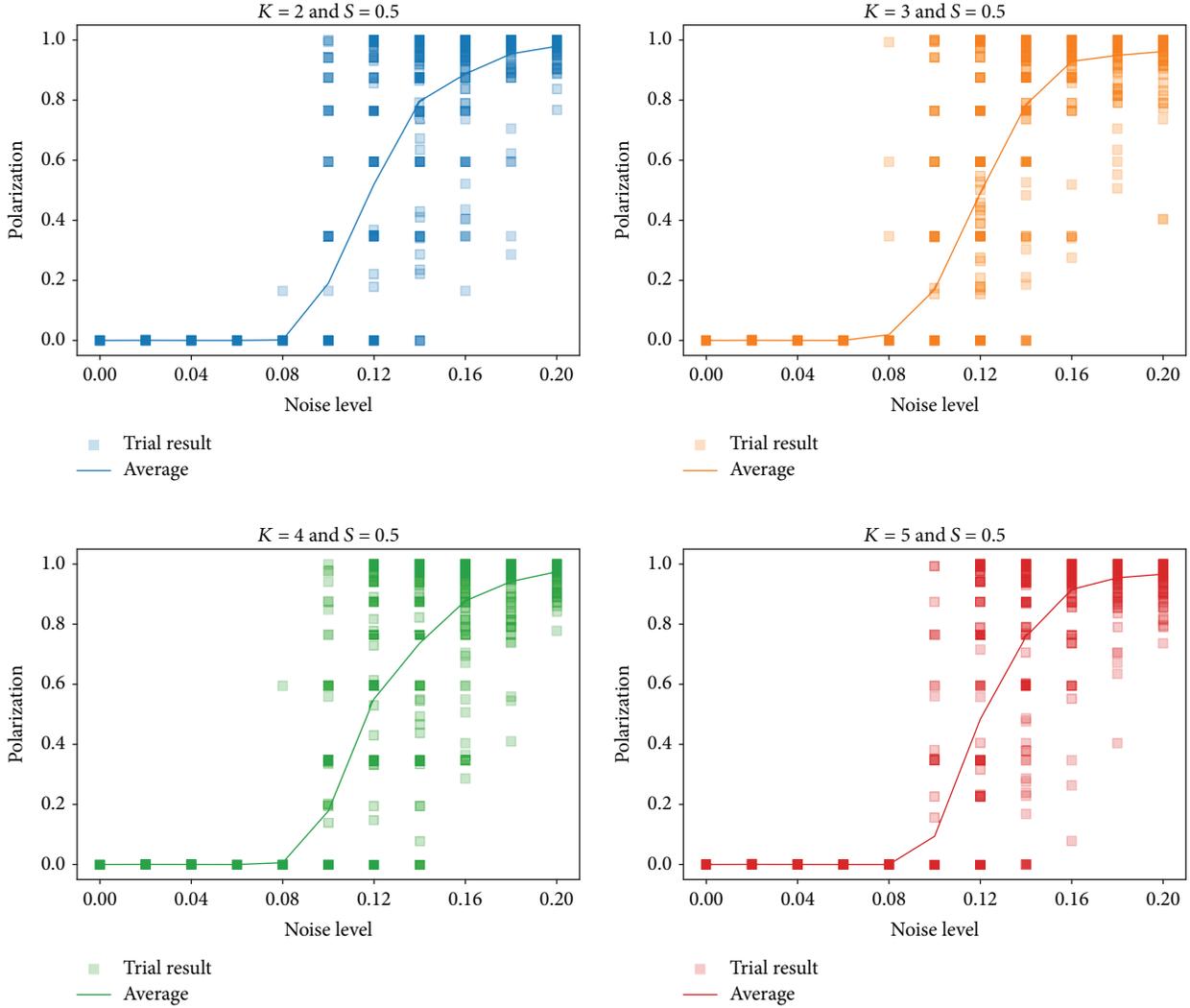


FIGURE 12: Final polarization of individual trial runs and averages from Figure 11 for $S = 0.5$ as a function of noise level, σ . As the noise level is increased, the system is increasingly biased towards larger final polarization outcomes.

We have highlighted the stochastic nature of the system being modeled. A key conclusion is that empirical results of opinions on social networks may, when taken on a case-by-case basis, exhibit trends that bear little resemblance to those predicted by the model. This is not necessarily an invalidation of the model, but merely a consequence of the variability inherent in complex systems. That said, given enough data, key trends should emerge. We have confirmed Flache and Macy’s [19] result that long-range ties increase polarization. As such, we might emphasize the importance of local communities being allowed to reach their own consensus. We have shown that decreasing initial extremism can reduce polarization, as one might expect. Achieving consensus in a community relies heavily on the absence of opinions at the extremes. However, this result is quite sensitive to noise in communication. A little bit of noise can shift consensus from centrist or ambivalent positions to more extreme views, while more noise can lead to polarization. Even if polarization is to be avoided, what about the intermediate case of “extreme consensus”? While it may be natural to view extreme

opinions as undesirable, an alternative perspective is that they represent a more stable system of cultural coherence. Note that these findings contradict computational and mathematical studies of the bounded confidence model under the influence of noise, where sufficient noise breaks polarization and leads to disordered opinion spreading [43–45]. This is because in the bounded confidence model, agents that are too far from one another do not interact. In the FM model, connected agents always interact, and the further apart they are in opinion space, the more strongly they repel one another in opinion space.

We confirmed Flache and Macy’s [19] result that increased “cultural complexity”—the number of opinions that are important to individuals in assessing their similarities and differences with others—decreased overall polarization. We also showed that this result stems directly from an increase in the number of permutations of extreme opinions individuals can hold when there are more items on which one can hold opinions. This might be viewed as a flaw in the metric of polarization used here. Alternatively, we believe it is

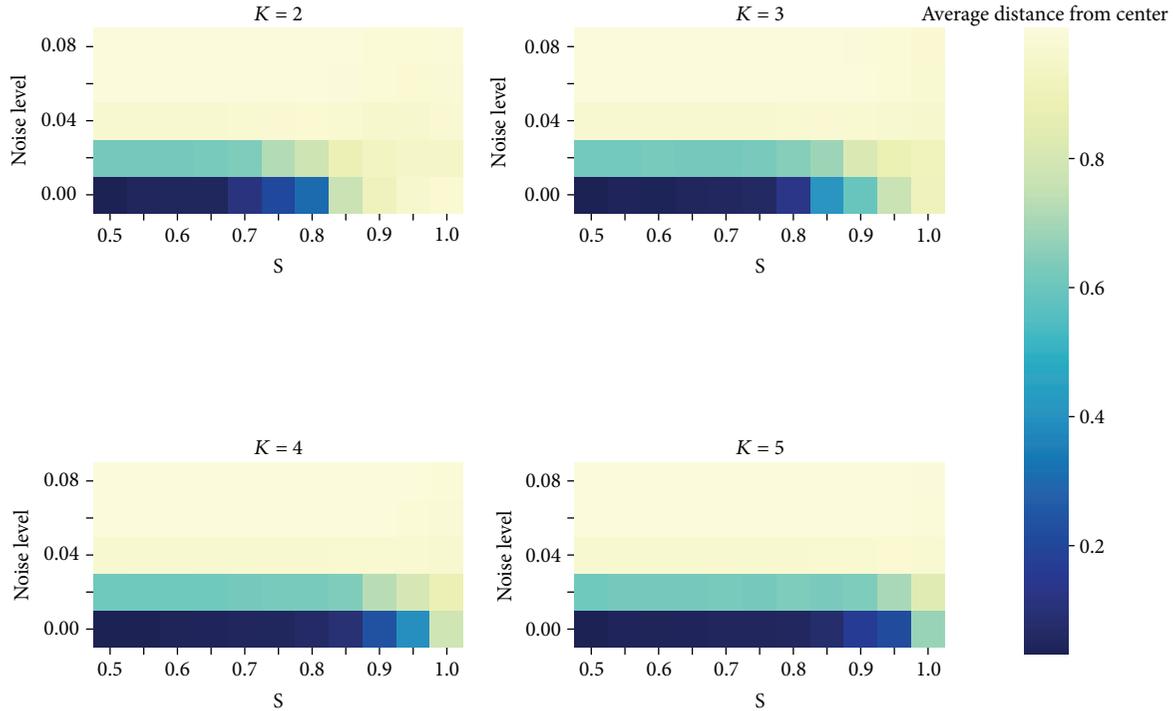


FIGURE 13: Noisy communication causes extremism without polarization before it causes extremism with polarization. For all K pictured, the average distance from center increases with moderate levels of noise, even though polarization has not increased, as shown in Figure 11. The value in each square of the heat map is the average of 100 trials.

reasonable to posit that a community with a wider diversity of views should be considered less polarized than a community with only a few suites of clustered opinions. In any case, this finding highlights the importance of a thorough understanding of one’s distance measure when dealing with multidimensional opinions. Our analysis may in fact cast doubt on the interpretation by Flache and Macy [19] that cultural complexity decreases opinion polarization, if one also rejects the interpretation that adding arbitrary traits on which actors are indifferent should reduce their opinion distance.

As noted, the model we have studied is a simplified abstraction and does not include many details that are important to the empirical reality of opinion dynamics. In general, theoretical modeling work should start simple and gradually add heterogeneity as the simpler versions of the system in question become fully described. Future work should explore these sources of heterogeneity. First, we did not distinguish between private opinions and public productions representing those opinions [32]. Our operationalization of communication noise could be interpreted as a modulation of private opinion, but communication noise could also be interpreted as misunderstanding of perfectly reproduced, publicly voiced opinions. People often communicate public opinions that differ from their private opinions when incentives for the parties involved are not aligned [46–48]. Second, we ignored the structural influence of explicit identity groups. It could be argued that clustering of agent opinions implicitly defines an identity group. For example, [25] measured network autocorrelation to explain

why people’s preferences cluster together. This data-driven approach was offered as an attempt to explain arbitrary opinion clustering, as indicated by the paper’s title, “Why do liberals drink lattes?”. Nevertheless, explicit identity with groups and roles influences human behavior far beyond homophilic clustering [49–51]. Third, we ignored individual differences in how individuals influence and are influenced. Some people may be stubborn while others are easily swayed. Some prestigious or charismatic individuals may have outsized influence while others are ineffective at communicating their opinions. Relatedly, individuals may also vary in their confidence in their opinions, which will influence the extent of their mutability and persuasion. The assumption that as agents become more extreme, their opinions become more stubborn, as formalized in (3), may not always hold. Indeed, our work highlights the need for additional empirical work on how individuals alter their opinions as a function of how extreme those opinions are. Finally, the social networks used in our model are simplistic in both dynamics and structure. Ties in many real-world networks change with greater frequency than we modeled, providing new opportunities for social influence. Moreover, interactions and opinions are contextual. Individuals are embedded in multilayered social networks, in which the dynamics of opinions may be considerably more nuanced than indicated by our relatively static, single-layer network [30, 52].

In our study of the FM model, we have found rich behaviors and theoretical lessons for understanding opinion dynamics. This work highlights the potential for complexity even in a very simple model of individual behavior, because

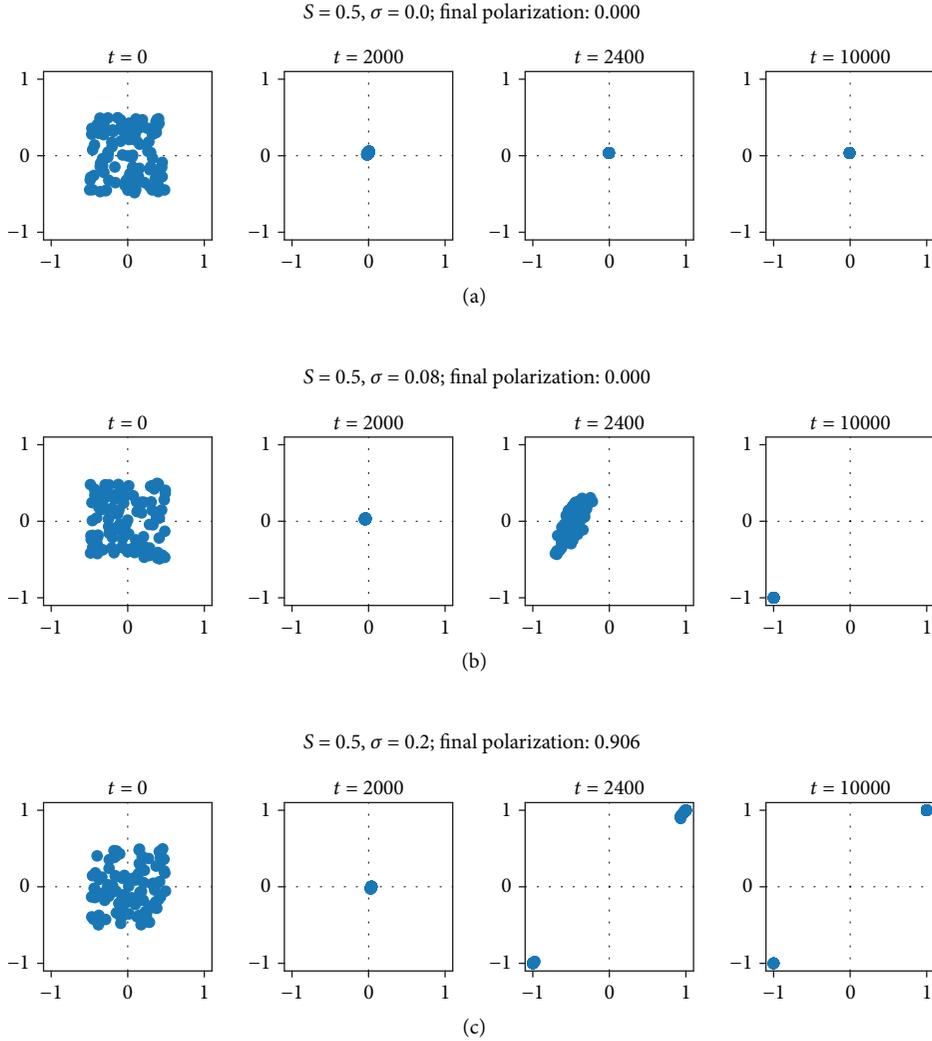


FIGURE 14: Exemplar spatiotemporal dynamics of agent opinion coordinates with $K = 2$ and $S = 0.5$ for $\sigma \in \{0.0, 0.08, 0.2\}$. There are three regimes. In the first, without noise, every simulation ends in centrist consensus (top row). In the presence of noise with $\sigma = 0.08$, agents find extremist consensus; in this trial, agents found consensus around the point $(-1, -1)$. The third regime is the high polarization regime at the highest level of communication noise we tested, $\sigma = 0.2$. In this regime, agents split into opposing camps, led by first-mover extremists.

network structure provides for path dependent effects and can be further influenced by initial conditions and noise. Our analytic approach highlights the value of systematic investigation of a model’s explicit and tacit assumptions.

Appendix

Proof that Polarization Scales with $1/K$

We hypothesized that the decrease in polarization with increasing K observed in simulations of the FM model was driven by an increase in the number of permutations of binary vectors of length K , in which each element was -1 or 1 . We supported this hypothesis in the main text with simulations in which agents were randomly initialized at such extreme positions in opinion space. Here, we derive a formal proof that polarization in the FM model scales with $1/K$ if we assume that agents are randomly assigned a vector of “extreme” opinions, such that $\forall i, k, s_{ik} \in \{-1, 1\}$. To do this,

we exactly calculate the polarization of a population where each agent occupies one of the 2^K corners of opinion space with K cultural features.

Recall that polarization is defined as the variance in pairwise distances between all agents. We define the *combinatorial polarization*, $P_c(K)$, as the polarization that arises from randomly placing each agent at one of the 2^K corners of opinion space with K cultural features, which is a K -hypercube, denoted Q_K . “Corners” of opinion space are simply vertices in the graph of Q_K . We computed this value numerically for $K \in \{1, \dots, 12\}$ and found it tracks closely to $1/K$ (see Figure 10). Here, we demonstrate that $P_c = 1/K$ exactly in the large N limit. To calculate $P_c(K)$, we need three elements. First, we need to calculate the distance between pairs of agents at different corners of Q_K . Second, we must count the number of agent pairs separated by the distance from one corner to another. We do this by first counting the number of subcubes of dimension L or L -subcube. Then, we count the number of maximally separated pairs in a subcube of L

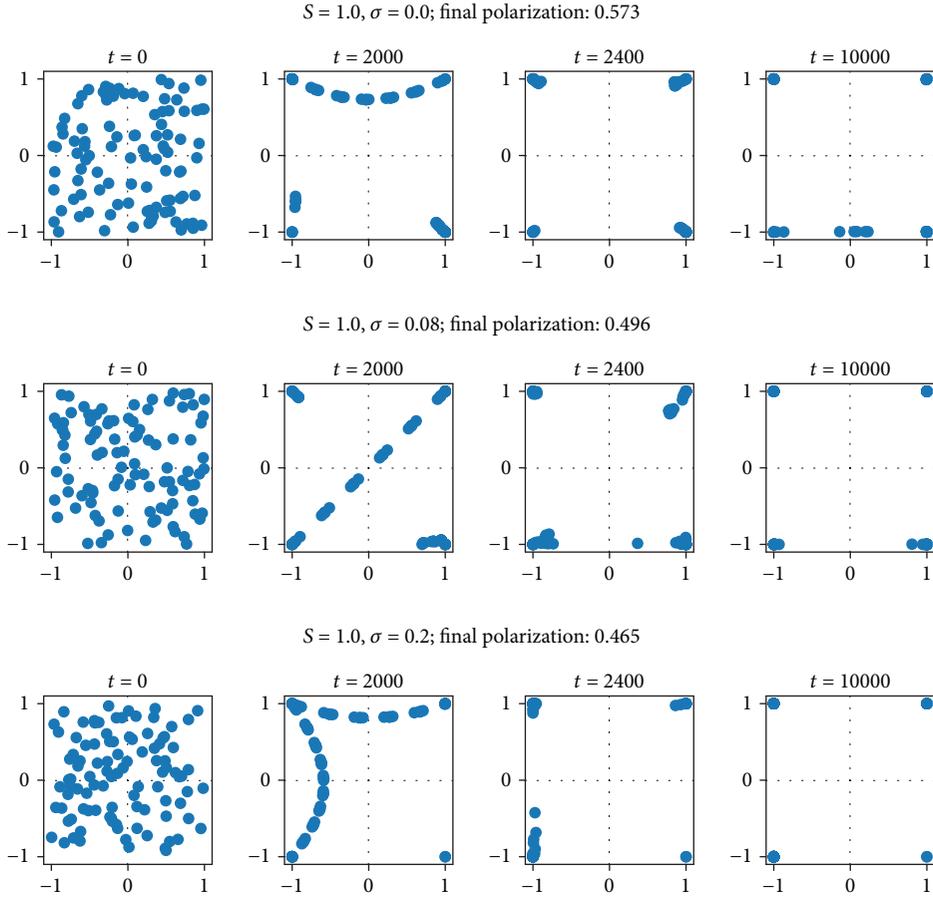


FIGURE 15: Exemplar spatiotemporal dynamics of agent opinion coordinates with $K = 2$ and $S = 1.0$ for $\sigma \in \{0.0, 0.08, 0.2\}$. Before the random long-range ties are added at $t = 2000$, extremists pull centrists to the extremes, but more centrist agent caves are balanced between more extreme caves. When long-range ties are added, the balance is broken and agents proceed to move to one of the extremes. Because at least some extremists held each of the corners, centrist agents do not move only to polar opposite corners, but in many cases to the nearest corner contained a neighboring (in the network sense) agent.

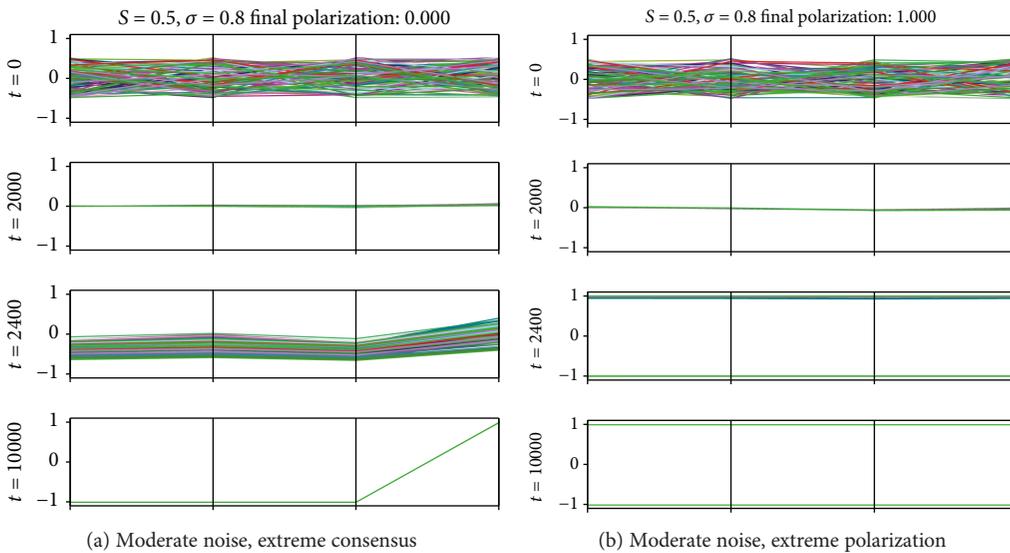


FIGURE 16: Exemplar parallel coordinate time series for $K = 4$ and $S = 0.5$. Here, the x -axis represents a single opinion coordinate, k_i , and the y -axis is the location of an agent for that coordinate. Each agent is represented by a line, colored by cave membership. With $\sigma = 0.1$, consensus emerges but at a corner of the opinion space.

-subcube. Finally, we calculate the distance of maximally separated, or *antipodal* pairs, of agents in an L -subcube. We can then calculate the expected value of pairwise distances, $\langle d \rangle$, and the expected square of pairwise distance, $\langle d^2 \rangle$, from which we will have the combinatorial polarization

$$P_c = \langle d^2 \rangle - \langle d \rangle^2. \quad (\text{A.1})$$

We will show that $P_c = 1/K$ by showing that $\langle d \rangle = 1$ and $\langle d^2 \rangle = K + 1/K$. Before we do that, we will derive functions to help us count the number of pairs separated by a particular distance and to calculate distances between vertices on subcubes $Q_L \subseteq Q_K$. First, we denote the total number of pairwise distances as $n = N(N-1)/2$, where N is the number of agents. The number of L -subcubes $Q_L \subseteq Q_K$ is

$$n_s(L, K) = 2^{K-L} \binom{K}{L}. \quad (\text{A.2})$$

This results from the fact that at all 2^K vertices of Q_K , $\binom{K}{L}$ subcubes can be created by choosing L nodes adjacent to the vertex. This gives us $2^K \binom{K}{L}$ subcubes. This overcounts since each generated subcube was generated once for each of its 2^L vertices. So, we must divide by a factor of 2^L , giving us the expression in (A.2). Within Q_L , the number of pairwise distances where agents occupy antipodal vertices is

$$n'_a(L, K) = 2^{L-1} \left(\frac{N}{2^K} \right)^2. \quad (\text{A.3})$$

There are 2^{L-1} pairs of antipodal vertices in Q_L . In the large N limit, agents are distributed in equal number to each vertex of Q_K . Then, the number of agents at a single vertex is $N/2^K$, so the number of pairwise distances between any two antipodal pairs is $(N/2^K)^2$. The total number of antipodal pairs across all Q_L is then

$$n_a(L, K) = n'_a(L, K) n_s(L, K). \quad (\text{A.4})$$

Finally, the distance between agent opinions \vec{s}_1 and \vec{s}_2 in antipodal vertices of Q_L is

$$d_a(L, K) = \frac{1}{K} \sum_{k=1}^K |s_{1k} - s_{2k}| = \frac{2L}{K}, \quad (\text{A.5})$$

since any antipodal vertices of Q_L share $K-L$ opinion coordinates, and the maximum magnitude of difference on a single opinion dimension is 2.

With these quantities, we can write the expected value of pairwise distance

$$\langle d \rangle = \frac{1}{n} \sum_{L=1}^K n_a(L, K) d_a(L, K). \quad (\text{A.6})$$

Simplifying and taking $N \rightarrow \infty$, this becomes

$$\langle d \rangle = \frac{(K-1)!}{2^{K-1}} \sum_{L=1}^K \frac{1}{(K-L)!(L-1)!}. \quad (\text{A.7})$$

Using the identity

$$\sum_{L=1}^K \frac{1}{(K-L)!(L-1)!} = \frac{2^{K-1}}{(K-1)!}, \quad (\text{A.8})$$

we find $\langle d \rangle = 1$. Calculating $\langle d^2 \rangle$ proceeds similarly, beginning with

$$\langle d^2 \rangle = \frac{1}{n} \sum_{L=1}^K n_a(L, K) d_a(L, K)^2. \quad (\text{A.9})$$

Simplifying and taking $N \rightarrow \infty$, this becomes

$$\langle d^2 \rangle = \frac{(K-1)!}{2^{K-2} K} \sum_{L=1}^K \frac{L}{(K-L)!(L-1)!}. \quad (\text{A.10})$$

With the identity

$$\sum_{L=1}^K \frac{L}{(K-L)!(L-1)!} = \frac{2^{K-2}(K+1)}{(K-1)!}, \quad (\text{A.11})$$

we find $\langle d^2 \rangle = K + 1/K$. So,

$$P_c = \langle d^2 \rangle - \langle d \rangle^2 = \frac{K+1}{K} - 1 = \frac{1}{K}. \quad (\text{A.12})$$

Data Availability

Data used for our analyses is available for download (14 GB) at http://mt.digital/static/data/polarization_v0.1-data.tar.

Conflicts of Interest

There are no conflicts of interest for either of the authors.

Acknowledgments

Computational experiments were performed on the MERCED computing cluster, which is supported by the National Science Foundation (Grant no. ACI-1429783).

References

- [1] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Advances in Complex Systems*, vol. 3, pp. 87–98, 2000.

- [2] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.
- [3] C. Efferson, R. Lalive, P. Richerson, R. McElreath, and M. Lubell, "Conformists and mavericks: the empirics of frequency-dependent cultural transmission," *Evolution and Human Behavior*, vol. 29, no. 1, pp. 56–64, 2008.
- [4] J. Henrich and R. Boyd, "The evolution of conformist transmission and the emergence of between-group differences," *Evolution & Human Behavior*, vol. 19, no. 4, pp. 215–241, 1998.
- [5] M. Muthukrishna, T. J. H. Morgan, and J. Henrich, "The when and who of social learning and conformist transmission," *Evolution & Human Behavior*, vol. 37, no. 1, pp. 10–20, 2016.
- [6] P. E. Smaldino and J. M. Epstein, "Social conformity despite individual preferences for distinctiveness," *Royal Society Open Science*, vol. 2, no. 3, article 140437, 2015.
- [7] R. Axelrod, "The dissemination of culture: a model with local convergence and global polarization," *Journal of Conflict Resolution*, vol. 41, no. 2, pp. 203–226, 1997.
- [8] K. M. Carley, "Group stability: a socio-cognitive approach," *Advances in Group Processes*, vol. 7, no. 1, p. 44, 1990.
- [9] P. Dandekar, A. Goel, and D. T. Lee, "Biased assimilation, homophily, and the dynamics of polarization," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 15, pp. 5791–5796, 2013.
- [10] K. Gray, D. G. Rand, E. Ert, K. Lewis, S. Hershman, and M. I. Norton, "The emergence of "Us and Them" in 80 lines of code: modeling group genesis in homogeneous populations," *Psychological Science*, vol. 25, no. 4, pp. 982–990, 2014.
- [11] C. G. Lord, L. Ross, and M. R. Lepper, "Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence," *Journal of Personality and Social Psychology*, vol. 37, no. 11, pp. 2098–2109, 1979.
- [12] N. Mark, "Beyond individual differences: social differentiation from first principles," *American Sociological Review*, vol. 63, no. 3, pp. 309–330, 1998.
- [13] R. McElreath, R. Boyd, and P. J. Richerson, "Shared norms and the evolution of ethnic markers," *Current Anthropology*, vol. 44, no. 1, pp. 122–130, 2003.
- [14] P. E. Smaldino, M. A. Janssen, V. Hillis, and J. Bednar, "Adoption as a social marker: innovation diffusion with outgroup aversion," *The Journal of Mathematical Sociology*, vol. 41, no. 1, pp. 26–45, 2017.
- [15] H. Tajfel, M. G. Billig, R. P. Bundy, and C. Flament, "Social categorization and intergroup behaviour," *European Journal of Social Psychology*, vol. 1, no. 2, pp. 149–178, 1971.
- [16] Pew Research Center, *The Partisan Divide on Political Values Grows Even Wider*, Pew Research Center, 2017.
- [17] A. Flache, M. Mäs, T. Feliciani et al., "Models of social influence: towards the next frontiers," *Journal of Artificial Societies and Social Simulation*, vol. 20, no. 4, p. 2, 2017.
- [18] P. E. Smaldino and J. C. Schank, "Human mate choice is a complex system," *Complexity*, vol. 17, no. 5, 22 pages, 2012.
- [19] A. Flache and M. W. Macy, "Small worlds and cultural polarization," *The Journal of Mathematical Sociology*, vol. 35, no. 1–3, pp. 146–176, 2011.
- [20] J. Adams and V. J. Roscigno, "White supremacists, oppositional culture and the world wide web," *Social Forces*, vol. 84, no. 2, pp. 759–778, 2005.
- [21] P. S. Hart and E. C. Nisbet, "Boomerang effects in science communication: how motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies," *Communication Research*, vol. 39, no. 6, pp. 701–723, 2012.
- [22] L. Boxell, M. Gentzkow, and J. M. Shapiro, "Greater Internet use is not associated with faster growth in political polarization among US demographic groups," *Proceedings of the National Academy of Sciences*, vol. 114, no. 40, pp. 10612–10617, 2017.
- [23] K. Takács, A. Flache, and M. Mäs, "Discrepancy and disliking do not induce negative opinion shifts," *PLoS One*, vol. 11, no. 6, article e0157948, 2016.
- [24] P. E. Smaldino, "Models are stupid, and we need more of them," in *Computational Models in Social Psychology*, R. R. Vallacher, A. Nowak, and S. J. Read, Eds., Psychology Press, 2017.
- [25] D. DellaPosta, Y. Shi, and M. Macy, "Why do liberals drink lattes?," *American Journal of Sociology*, vol. 120, no. 5, pp. 1473–1511, 2015.
- [26] Pew Research Center, *News Use Across Social Media Platforms 2016*, Pew Research Center, 2016.
- [27] Pew Research Center, *Anger Beat Love in Facebook Reactions to Lawmaker Posts after 2016 Election*, Pew Research Center, 2018.
- [28] Pew Research Center, *Political Polarization and Media Habits*, Internet Project, 2014.
- [29] H. Clark, *Using Language*, Cambridge University Press, Cambridge, UK, 5th edition, 1996.
- [30] F. Battiston, V. Nicosia, V. Latora, and M. S. Miguel, "Layered social influence promotes multiculturalism in the Axelrod model," *Scientific Reports*, vol. 7, no. 1, p. 1809, 2017.
- [31] N. P. Mark, "Culture and competition: homophily and distancing explanations for cultural niches," *American Sociological Review*, vol. 68, no. 3, pp. 319–345, 2003.
- [32] A. Nowak, J. Szamrej, and B. Latané, "From private attitude to public opinion: a dynamic theory of social impact," *Psychological Review*, vol. 97, no. 3, pp. 362–376, 1990.
- [33] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis, and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, 2002.
- [34] R. A. Rescorla and A. R. Wagner, "A theory of Pavlovian conditioning: variations on the effectiveness of reinforcement and non-reinforcement," in *Classical Conditioning II: Current Research and Theory*, pp. 64–99, Appleton-Century-Crofts, New York, NY, USA, 1972.
- [35] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA, USA, 1998.
- [36] A. Bramson, P. Grim, D. J. Singer et al., "Disambiguation of social polarization concepts and measures," *The Journal of Mathematical Sociology*, vol. 40, no. 2, pp. 80–111, 2016.
- [37] D. J. Watts, "Networks, dynamics, and the small-world phenomenon," *American Journal of Sociology*, vol. 105, no. 2, pp. 493–527, 1999.
- [38] J. Henrich, *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*, Princeton University Press, 2015.
- [39] K. N. Laland, *Darwin's Unfinished Symphony: How Culture Made the Human Mind*, Princeton University Press, 2017.
- [40] J. Sides and D. J. Hopkins, Eds., *Political Polarization in American Politics*, Bloomsbury, New York, NY, USA, 2015.
- [41] M. Prior, "Media and political polarization," *Annual Review of Political Science*, vol. 16, no. 1, pp. 101–127, 2013.

- [42] B. D. Wood and S. Jordan, "Polarization as the norm of the American system," in *Party Polarization in America*, pp. 300–314, Cambridge University Press, 2017.
- [43] A. Carro, R. Toral, and M. San Miguel, "The role of noise and initial conditions in the asymptotic solution of a bounded confidence, continuous-opinion model," *Journal of Statistical Physics*, vol. 151, no. 1-2, pp. 131–149, 2013.
- [44] T. Kurahashi-Nakamura, M. Mäs, and J. Lorenz, "Robust clustering in generalized bounded confidence models," *Journal of Artificial Societies and Social Simulation*, vol. 19, no. 4, 2016.
- [45] M. Pineda, R. Toral, and E. Hernández-García, "Noisy continuous-opinion dynamics," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2009, no. 8, 2009.
- [46] V. P. Crawford and J. Sobel, "Strategic information transmission," *Econometrica*, vol. 50, no. 6, pp. 1431–1451, 1982.
- [47] S. Pinker, M. A. Nowak, and J. J. Lee, "The logic of indirect speech," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 3, pp. 833–838, 2008.
- [48] P. E. Smaldino, T. J. Flamson, and R. McElreath, "The evolution of covert signaling," *Scientific Reports*, vol. 8, no. 1, p. 4905, 2018.
- [49] F. Barth, *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*, Little, Brown, 1969.
- [50] J. Berger and C. Heath, "Who drives divergence? Identity signaling, outgroup dissimilarity, and the abandonment of cultural tastes," *Journal of Personality and Social Psychology*, vol. 95, no. 3, pp. 593–607, 2008.
- [51] P. E. Smaldino, "Social identity and cooperation in cultural evolution," *Behavioural Processes*, 2017.
- [52] P. E. Smaldino, R. M. D'Souza, and Z. Maoz, "Resilience by structural entrenchment: dynamics of single-layer and multiplex networks following sudden changes to tie costs," *Network Science*, vol. 6, no. 2, pp. 157–175, 2018.

