Role of social environment and social clustering in spread of opinions in coevolving networks

Nishant Malik[a)] and Peter J. Mucha
Department of Mathematics, CB 3250, University of North Carolina, Chapel Hill,
North Carolina 27599, USA

Taking a pragmatic approach to the processes involved in the phenomena of collective opinion formation, we investigate two specific modifications to the coevolving network voter model of opinion formation studied by Holme and Newman [Phys. Rev. E **74**, 056108 (2006)]. First, we replace the rewiring probability parameter by a distribution of probability of accepting or rejecting opinions between individuals, accounting for heterogeneity and asymmetric influences in relationships between individuals. Second, we modify the rewiring step by a path-length-based preference for rewiring that reinforces local clustering. We have investigated the influences of these modifications on the outcomes of simulations of this model. We found that varying the shape of the distribution of probability of accepting or rejecting opinions can lead to the emergence of two qualitatively distinct final states, one having several isolated connected components each in internal consensus, allowing for the existence of diverse opinions, and the other having a single dominant connected component with each node within that dominant component having the same opinion. Furthermore, more importantly, we found that the initial clustering in the network can also induce similar transitions. Our investigation also indicates that these transitions are governed by a weak and complex dependence on system size. We found that the networks in the final states of the model have rich structural properties including the small world property for some parameter regimes. © 2013 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4833995]

As the study of networks is applied to an ever broadening variety of phenomena, it is important to study the properties of networks, dynamical processes coupled across networks, and the interplay between the two where the coupled dynamics affect the network topology. A minimal mathematical model that has been used to model the social phenomena of collective opinion formation is the coevolving voter model.[1,16–26] We introduce two additional attributes to the multi-opinion coevolving voter model, in order to describe processes and networks that are closer to real-world situations within a still relatively simple model. Our model includes a "social environment," modeling the inherent heterogeneity and asymmetry in relationships within a social group. We also include a path-length-based preference for rewiring that reinforces social clustering. Our inclusion of this second attribute has been influenced by the fact that clustering is a ubiquitous feature of networks and has not been incorporated as a dynamic entity in most coevolving voter models. We explore the consequences of these two additional attributes within the coevolving voter model, comparing and contrasting the behaviors of this only slightly more complicated model with those of the minimal coevolving voter model. Our results highlight the important role of clustering, with possible consequences for future applications of coevolving voter models.

I. INTRODUCTION

It has been widely reported in the media that online social networks like Facebook, Twitter, Blackberry messenger, etc. played a key role in recent events in the world political sphere such as the Arab spring and London riots of 2011.[2–6] Meanwhile, there has also been increased interest in the quantitative and analytical analysis of the mechanisms and dynamics of the spread of social contagions such as rumors and opinions on complex networks.[6–15] In such studies, individuals in the society are represented by nodes with edges indicating relationships between them, and then techniques from social network analysis and from statistical and nonlinear science are employed to analyze plausible models of the dynamics of spread of social contagions on a network.[1,16–26]

We propose a variation of the simplest coevolving network voter model of opinion formation, studied by Holme and Newman.[1] In the model of Holme and Newman, an edge between individuals with different opinions is either rewired to connect two nodes having the same opinion or the opinion of an individual is changed to agree with the opinion of one of its neighbors. The selection from these two options is based on a parameter named the rewiring probability. We add two more simple mechanisms to this model, inspired by a pragmatic approach to the modeling of asymmetric influences and tendencies to local clustering in the phenomenon of collective opinion formation. We then observe a broader array of model behaviors induced by these modifications to the coevolving voter model. For convenience of the present

a)Electronic mail: nmalik@email.unc.edu

exposition, we refer to these additional mechanisms as: (1) *Social environment* and (2) *Social clustering*. Below we describe their meaning and significance in the processes of opinion formation.

Acceptance or rejection of somebody else's opinion by an individual depends on a multitude of factors including the strength of relationship between the concerned individuals. A prevailing *social environment* (as defined, e.g., in Ref. 27) not only alters relationships between individuals but can also affect their opinions on different issues in a fundamental way. A highly divisive society may be an outcome of inflexibilities in relationships that exist between individuals who resist accepting or sharing each others' opinions, choices, or views. Moreover, these inflexibilities themselves could be due to the prevailing "negative" social environment. In contrast, a "positive" social environment leading to flexible relationships between individuals leads to less resistance to accept each others' opinions, choices, or views. In modern times, media and advertising also play a significant role in altering the social environment and in constructing consent around certain opinions or choices.[28]

We propose to incorporate the effect of the social environment on the model of opinion formation on coevolving networks by a distribution of probability of accepting or rejecting opinions between individuals. The distribution for social environment replaces the constant rewiring probability that has been used before in other studies on voter model with coevolving networks.[1,16–18] Such description of the social environment becomes more plausible if we note the fact that relationships among individuals in a social group are inherently heterogeneous and asymmetric. For simplicity, we have assumed that the social environment remains the same over the temporal evolution of the model.

Another important aspect that has not yet been sufficiently analyzed in the models of opinion formation on coevolving networks has been the role of local clustering of edges in the network and other similar preferences for new links to be formed between nodes that are already near each other in the network. Indeed, in most models studied to date, the dynamics involved have been assumed to be independent of the any distances in the network beyond the immediate nature of whether two nodes are already connected. In the present model we have attempted to explore the complex consequences of a simple introduction of network-distance and clustering effects into the model. Specifically, we replace the random rewiring step of other models with a step that prefers rewiring to nodes/actors who are already close within the network and who have higher probability of accepting new opinions. This process reinforces local clustering in the evolving network so that clustering coefficients do not vanish in the large-network limit (as in other previous models). Clustering is a fundamental property of most network representations of social contexts, i.e., friends of friends have a higher likelihood (relative to the rest of the network) of also being friends.[13,14,29] However, rewiring rules for coevolving network models that do not reinforce clustering (as in, e.g., Refs. 1 and 16) can randomize away any initial clustering, greatly simplifying the associated opinion dynamics at the cost of dynamically generating networks that are unlike real social networks.

The explicit incorporation of model processes for *social environment* and *social clustering* provides a simple simulation for the coupled effects of opinions with clustering and *homophily*, the tendency of individuals to connect with individuals having similar characteristics.[30]

II. DESCRIPTION OF THE MODEL

Let $G(N,E)$ be a network of N nodes and E edges with a predefined topology. Let $\{O_i\}$ represent a set of O number of opinions initially uniformly distributed over the N nodes of $G(N,E)$. Let p_{ij} be the probability of some node j accepting an opinion from node i. The distribution $P(p_{ij})$ describes the *social environment* determining the values of p_{ij}, the probability of the jth node accepting the opinion of the ith node. If an edge exists between node i and j then we say $E_{ij} = 1$. An edge connecting two nodes with different opinions (i.e., $E_{ij} = 1$ with $O_i \neq O_j$) is called a *discordant edge*. The total number of discordant edges in G is represented by E_- and $E = E_+ + E_-$, where E_+ represents the number of harmonious edges (i.e., edges connecting nodes with the same opinion).

Algorithm 1: A voter model on a coevolving network with clustering and heterogeneous levels of influence.

1:     Generate a graph G of given topology
2:     Generate a given distribution for p_{ij}, i.e., $P(p_{ij})$
3:     Populate nodes with O number of uniformly distributed opinions $\{O_i\}$
4:     Calculate E_-
5:     **while** $E_- \neq 0$ **do**
6:        Randomly choose a discordant edge E_{ij} (with equal probabilities of which end of the edge is labeled i and j)
7:        Generate a uniform random number ξ between 0 and 1
8:        **if** $\xi < p_{ij}$ **then**
9:           $O_j \leftarrow O_i$
10:       Calculate E_-
11:       **else**:
12:         Remove the link between i and j, i.e., set $E_{ij} = 0$
13:         Find the set $\mathcal{N}' = \{j\}_{j \neq i} \cap \{k\}$
    ▷ Here $\{j\}_{j\neq i}$ is a set containing all nodes such that each element satisfies $p_{ij} \geq \xi$ and set $\{k\}$ contains all nodes not directly connected to node i that are within graph-theoretic distance d from note i, with d the minimum possible distance such that $\mathcal{N}' \neq \emptyset$. If no such distance exists, proceed
14:        **if** $\mathcal{N}' \neq \emptyset$ **then**
15:           Connect i randomly to any node $l \in \mathcal{N}'$
16:           $O_l \leftarrow O_i$
17:        **else**:
18:           Connect i randomly to any node j s.t. $O_j = O_i$
19:        **end if**
20:       Calculate E_-
21:       **end if**
22:   **end while**

Different individuals have different probabilities of acceptance of others' opinions, which is here taken to be independent of the existence of a link between the individuals. Several factors ranging from socio-cultural affinity to the prevailing political and economic situation can influence these probabilities differently for different individuals. To account for such variability we have used a distribution $P(p_{ij})$ for rewiring probabilities rather than a constant. We call $P(p_{ij})$ the *social environment function*, accounting for the heterogeneous and asymmetric relationships among individuals. For the purposes of simply exploring a variety of settings, we have considered two different kinds of power laws for the *social environment*. We set $P(p_{ij}) = p_{ij}^{\alpha}$ to represent a *flexible* social environment, i.e., many individuals are able to accept others' opinions readily. Alternatively, we consider $P(p_{ij}) = 1 - p_{ij}^{\alpha}$ to represent an *inflexible* social environment where individuals do not accept others' opinion readily and, hence, more churning happens in the network dynamics [see Fig. 1(b)]. While there has been some empirical evidence to suggest that election results in multi-party democracies have power law distributions of votes among candidates from different parties[31–33] (but see also Refs. 34 and 35), our use of a power law distribution in the present context is driven only by its simplicity for simulation and for the presentation of a qualitatively diverse set of social environments as a one-parameter family. Other distributions such as the exponential, beta, and extreme value distributions should also suffice to reproduce similar features. Any distribution which can qualitatively describe "flexible" and "inflexible" regimes would be sufficient here for the intended purpose here though of course some of the quantitative results would vary.

Steps 13–16 in Algorithm 1 ensure that rewiring connections are mostly made according to *social clustering*, i.e., a node has higher probability of connecting to a person who is either a friend of a friend or, if no such connections are available, connecting to a person at the shortest possible distance identified in the network. The set $\mathcal{N}'$ in Algorithm 1 consists of nodes/individuals who are close to the node i both in terms of path length between them in the network and also in that they have higher probabilities of accepting the opinion
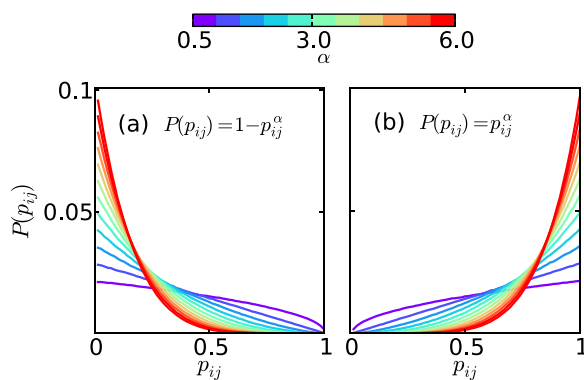
of node i. Hence, we call the nodes within the set N' to be socially close to the node i. If node i is not able to find an individual satisfying the required constraints, it connects uniformly at random to somebody else holding the same opinion to avoid complete social isolation.

Here, we aim to study the role of clustering of the network in altering the opinion space and network properties of the final end state. In so doing, our emphasis will be on transitions that occur in the network structure (notably, sizes and clustering of connected components) rather then just the space of opinions. We refer to the ratio of the number of opinions to the number of nodes, O/N, as *diversity*. We have fixed the average degree $\langle k \rangle = 4$ and number of opinions $O = 100$ for the simulations, if not mentioned otherwise. We have additionally investigated other values of numbers of $\langle k \rangle$ and O to confirm the robust nature of the qualitative properties that we describe. By the definition of the dynamics, the number of edges is conserved: at any time t, $E(t) = \langle k \rangle \frac{N}{2}$. Letting the coevolution of the network and opinions start at $t = t_\circ$ with initial number of discordant edges $E_-(t_\circ)$, the dynamics stop at the earliest $t = t_f$ such that $E_-(t_f) = 0$. That is, the final state of this model has no discordant edges remaining.

There are several levels of additional complexity that might be considered, and other plausible choices could be made to provide new insights into the coevolving dynamics of opinions and networks. But most such choices come at the price of making the model analytically harder to track. Indeed, even the very limited analytical tractability of graph fission in a two-opinion coevolving voter model presented in Ref. 16 is undoubtedly aided by the rewiring rule considered there randomizing away all non-trivial clustering. In light of the significant complications introduced by the distance-influenced rewiring rule considered here, we have attempted to computationally analyze this model in a thorough manner.

A. Basic features of the model

In this section we give a brief introduction to the basic features of this model. First, we obtain two qualitatively distinct final states as we vary the social environment from flexible to inflexible. For a flexible social environment with $P(p_{ij}) = p_{ij}^{\alpha}$ and $\alpha = 6.0$, we observe formation of a single large connected component of size comparable to the initial network with each node in the component having the same opinion. We call this kind of final state a *hegemonic consensus* because of the emergence of one single dominating opinion. For an inflexible social environment, simulated by setting $P(p_{ij}) = 1 - p_{ij}^{\alpha}$ with $\alpha = 6.0$, we observe that the initially connected network disintegrates into a large number of small connected components where every node in a given component holds the same opinion, i.e., each component is in a state of internal consensus. We will refer to this kind of final state as a *segregated consensus* as this feature is qualitatively similar to the segregation of individuals in a society. A lattice based classical model of this social phenomena was given by Thomas Schelling,[20] where he showed segregation of two groups of populations ("red" and "white") who move over a check board following some simple rules. Several

FIG. 1. Different types of *social environment function* $P(p_{ij})$, where p_{ij} is the probability of the jth node accepting the opinion of the ith node: (a) "inflexible," $P(p_{ij}) = 1 - p_{ij}^{\alpha}$, so that more links will have lower probabilities of accepting opinions; (b) "flexible," $P(p_{ij}) = p_{ij}^{\alpha}$, so that more links will have higher probabilities of accepting opinions.
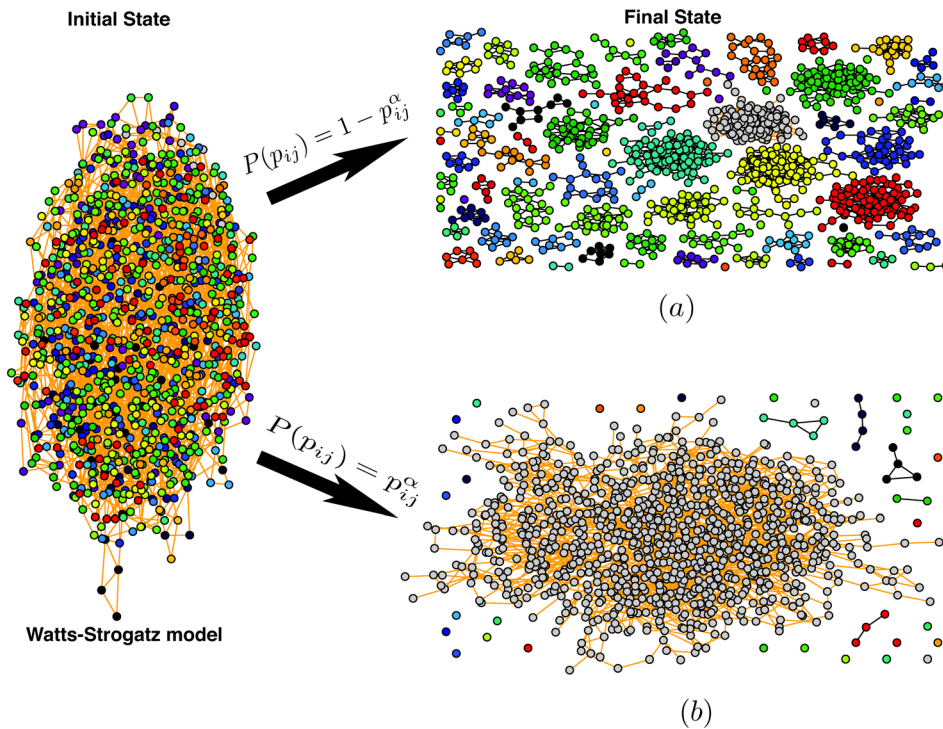
FIG. 2. A visual representation of the formation of qualitatively distinct consensus states for two different social environments. Both systems start with an initial Watts-Strogatz network (with $N = 1000$, $\langle k \rangle = 4$, and $\mathcal{C} = 0.1$). (a) Setting $P(p_{ij}) = 1 - p_{ij}^{\alpha}$ and $\alpha = 6$ creates an "inflexible" social environment. We observe disintegration of the network into small connected components with each component in internal consensus, i.e., *segregated consensus* occurs in the network. (b) Setting $P(p_{ij}) = p_{ij}^{\alpha}$ and $\alpha = 6$ creates a "flexible" social environment. We observe a dominant connected component in the final consensus, with size comparable to the initial network, while a large number of the initial opinions go extinct. We refer this kind of final state as a *hegemonic consensus*.

analytical and simulation results have been obtained following Schelling's model on networks as well as on coevolving networks but not in the context of modeling processes involved in collective opinion formation.[36–38]

A visualization for these observations for $N = 1000$ nodes with $O = 100$ is shown in Fig. 2. The drastic transition between the *hegemonic consensus* and *segregated consensus* in the final states of the systems seems to occur somewhere between the extreme flexible to inflexible social environments. Holme and Newman[1] observe some transitions qualitatively similar to this distinction between hegemonic and segregated consensus by changing their constant rewiring probability parameter. Intuitively, it is not surprising that changing the distribution of the social environment induces a transition similar to that studied by Holme and Newman,[1] insofar as the change in the distribution changes the overall average level of rewiring. Nevertheless, a priori we have no reason to expect that change in the form of the distribution of probabilities of accepting or rejecting of others' opinions should have similar effects as the changes to the single rewiring probability parameter employed by Holme and Newman.[1] Also, the detailed structural properties of the network in the *hegemonic consensus* and *segregated consensus* in the final states observed here are expected to be much richer, as shown and discussed below in some detail.

In Fig. 3 we observe the effect of varying the social environment, where s_i is the size (fraction of nodes in the network) of the ith component in the final consensus state, with i indexing the rank of the component sizes ($i = 1$ being the largest component). A further analysis of the phase transition involved in emergence of these two distinct states in this system has been attempted in detail in Sec. III, as one of the two central themes of this paper.

The giant consensus community occurring in the Holme and Newman model[1] would appear to be structurally similar
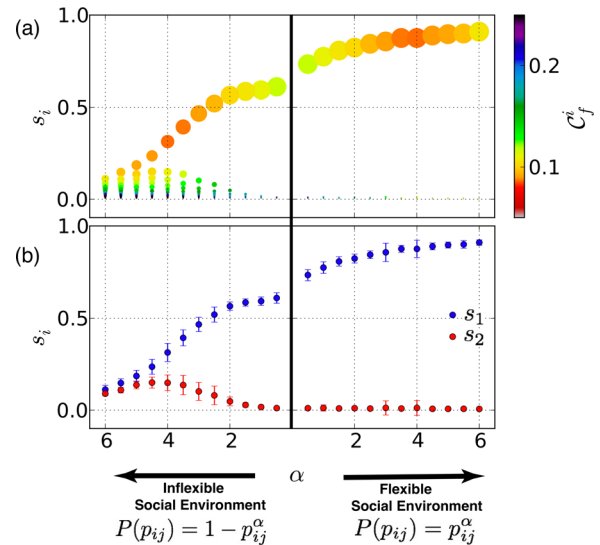

FIG. 3. The effect of different social environments on a network of $N = 1000$ nodes with $O = 100$ opinions initially present. The starting network is an Erdős-Rényi random network (i.e., clustering $\sim 1/N$). (a) The distribution of component sizes s_i, the fraction of nodes in the ith (ranked by size) connected component in the final consensus state, is plotted as a function of social environment (with marker sizes proportional s_i). Colors indicate $\mathcal{C}_f^i$, the clustering coefficient of the ith component. The thick bold line in the middle separates the two types of social environment function: on the right we consider flexible social environments, with a single large connected component with size increasing with increasing α; on the left we consider inflexible social environments, observing a decrease of the size of the largest connected component with increasing α, finally leading to its disintegration into several components of comparable sizes. (b) The sizes of the two largest connected components, s_1 and s_2, is plotted versus social environment. Simulations were conducted on 100 realizations of the network and initial opinion distribution, with the plotted component sizes estimated as the means over these realizations. Error bars give the standard deviation of these sizes across realizations.

to networks obtained under a configuration model with the observed final state degree distribution. In contrast, as observed in Fig. 4(a) the largest connected component in the *hegemonic consensus* has small world properties in that the average path lengths are comparable to independently distributed random networks but have high clustering coefficients. The dominant component also consists of nodes with higher numbers of connections, as is apparent from the change in cumulative degree distribution shown in Fig. 4(b). Because these features are qualitatively closer to organized political or religious movements, which usually have a hierarchy of leadership and high clustering, we have pointedly not referred to this structure as a *mob*. We have not observed variation in *diversity O/N* to bring about any significant change to the above discussed basic properties, while varying O from 2 to 100.
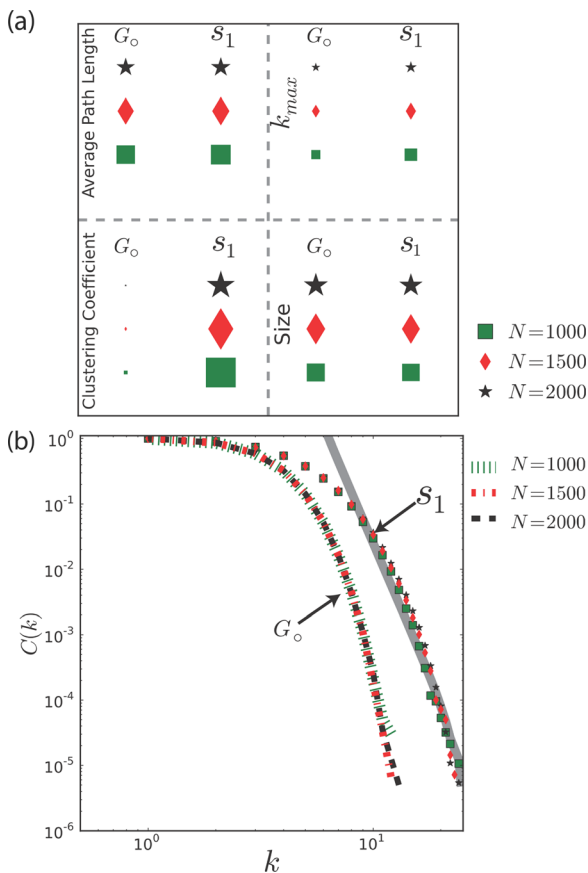

FIG. 4. Properties of the largest connected component (*hegemonic consensus*) for the final state reached for a flexible environment, $P(p_{ij}) = p_{ij}$, with $\alpha = 6.0$ from an initially Erdős-Rényi network. (a) Comparisons of average path length, maximum degree, clustering coefficient and size for different network sizes, with different markers representing different network sizes (see the legend). The initial network is G_0, with initial clustering close to zero, while s_1 here denotes both the largest connected component in the final state and its size (as a fraction of the nodes in the network). We observe that s_1 has a significantly higher clustering coefficient (0.2) whereas it has path length comparable to the initial Erdős-Rényi network G_0, implying that s_1 has small world features. Also, s_1 typically has higher k_{max} (maximum degree) while its size remains comparable to G_0. (b) The cumulative degree distribution $C(k)$ of the initial network G_0 (dashed lines) is compared with that for s_1 (markers), further showing that s_1 has nodes with higher degrees. In its tail, the cumulative degree distribution of s_1 appears to approximately follow a power law as shown by solid grey line of exponent -8 though the steepness of this line does not preclude other distributions in the tail.

Another crucial aspect to consider in this model is the role of initial network topology in transitions between *hegemonic consensus* and *segregated consensus* as the two distinct final states. Does the variation of the initial clustering coefficient change the final state? This question has not been considered in the previous studies of voter models on coevolving networks, as the previously introduced models have not treated clustering as a consequence of those models, even though clustering is one of the essential characteristics of social networks.[13,14,29] In the model considered here, the formation of a *hegemonic consensus* state apparently does not take place in networks with high initial clustering coefficient. To understand this feature we investigate the evolution of clustering in this model.

We specify the *clustering coefficient* of the network, $\mathcal{C}$, as three times the ratio of the number of loops of length three to the number of connected triples of nodes, also known as *transitivity*.[39] Symbols $\mathcal{C}_\circ$ and $\mathcal{C}_f$ are used here for the initial clustering at the start of the simulation and the final clustering at the end of the simulation, respectively. In the Watts-Strogatz model, for example, the maximum possible initial clustering $\mathcal{C}^{max}$ corresponding to the ring topology is $\mathcal{C}^{max} = \frac{3(\langle k \rangle - 2)}{4(\langle k \rangle - 1)}$. Therefore, with $\langle k \rangle = 4$, we would have $\mathcal{C}^{max} = 0.5$ (see, e.g., Ref. 40). The $\mathcal{C}^{max}$ value is also an upper bound for the observed $\mathcal{C}_f$. In Fig. 5 we plot the evolution of different variables of the model from a single simulation of $N = 1000$ nodes as discordant edges are removed. The social environment of this simulation was set to be flexible, $P(p_{ij}) = p_{ij}^\alpha$, with $\alpha = 6.0$, in a parameter regime where we expect formation of a *hegemonic consensus* state for initially unclustered networks. When we set $\mathcal{C}_\circ = 0$, the opinion space does undergo a transition as expected, and we see one opinion dominating [see Fig. 5(a)]. Correspondingly, there is no transition in the size of the largest connected component [see red dotted line in Fig. 5(c)]. For the black curve in Fig. 5, we have set $\mathcal{C}_\circ = \mathcal{C}^{max}$, and we observe a counter intuitive and unexpected transition viz. that the largest connected component starts to disintegrate and become smaller in size [see Fig. 5(c)] while in opinion space we do not observe emergence of a single dominant opinion [see Fig. 5(b)]. We also observe in the lowest panel of Fig. 5 that $\mathcal{C}$ saturates to $\mathcal{C}_f$ before reaching the consensus. This is a special feature of this model and provides an opportunity to study the evolution of a clustered network topology with opinion formation. For the case $P(p_{ij}) = p_{ij}^\alpha$ with $\alpha = 6$, $\mathcal{C}_f$ appears to be well approximated by a linear function of $\mathcal{C}_\circ$.

We also see in panel (d) of Fig. 5 that the system starts to slow down, in terms of the time necessary to remove the next discordant edge, right before the consensus states emerges. That is, more iterations are required to decrease the number of discordant edges, possibly indicative of some form of critical slowing of the system as *segregation* is reached. This feature is not as apparent in the red dotted curve, implying that processes involved in formation of *hegemonic consensus* might not involve critical slowing of the system. To identify the region of the parameter space where initial clustering $\mathcal{C}_\circ$ plays a dominant role in determining the final state, we have plotted the values of s_1 in Fig. 6 across the parameter space of α and $\mathcal{C}_\circ$. We observe in Fig. 6 that
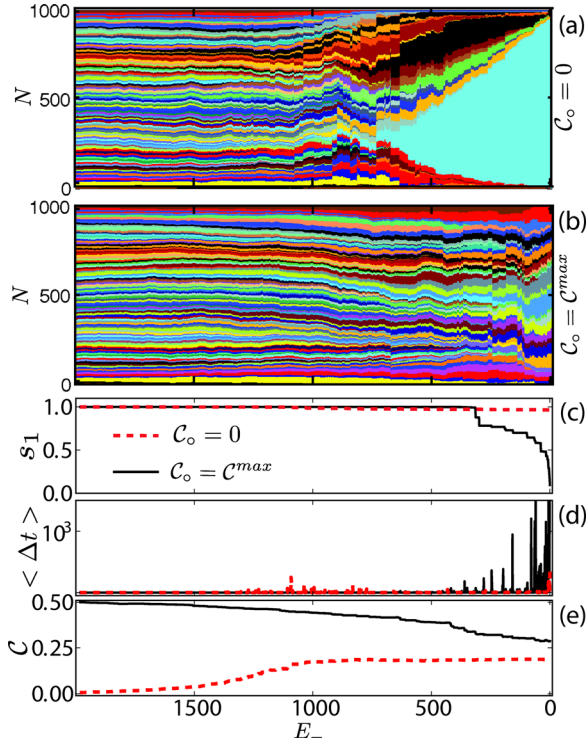
FIG. 5. The evolution of system variables with decreasing number of discordant edges. Each variable is plotted at the last time step when that number of discordant edges, E_-, was present in the system. The black line and panel (b) correspond to simulations starting at the highest possible clustering coefficient $\mathcal{C}^{max}$ whereas the red dotted line and panel (a) correspond to simulations starting at the negligible clustering coefficient obtained with a random network of independent edges. In (a) and (b), each color corresponds to one of the opinions, with width indicating the number of nodes holding that opinion. The wide width of cyan at the end in (a) represents the formation of a *hegemonic consensus* (one large connected component of size comparable to the initial network). We do not observe a similar transition in (b) even though the only difference in this simulation is the large initial clustering coefficient. In (c), we plot s_1, the size of the largest connected component. Observe the abrupt drop of the black line in s_1, indicating the disintegration of the network into smaller components (i.e., *segregated consensus*). In contrast, we do not observe any such transition for the red dotted line, corresponding to the formation of *hegemonic consensus*. In (d), we show $\langle \Delta t \rangle$, the average number of iterations of the system between last-observed times for each E_-, with a substantial increase for the black curve near the end. In (e), $\mathcal{C}$ is the corresponding evolution of the clustering coefficient.

higher values of initial clustering in the flexible social environment regime can induce fragmentation. In Fig. 7 we show the disintegration of the network into smaller components as we increase the initial clustering coefficient from 0 to $\mathcal{C}^{max}$. Now that we have briefly illustrated some of the features in the evolution of clustering in the model, we present a more systematic analysis of this transition below.

III. PHASE TRANSITIONS

A. Role of social environment in transitions

As discussed above this model shows transition to two distinct final states. For *flexible* social environments we have observed that as α is increased, the largest connected component's size approaches that of the whole network, $s_1 \rightarrow 1$ (see Fig. 3). In contrast, for the *inflexible* social environments, we have disintegration of the network into several smaller sized connected components. As we move from
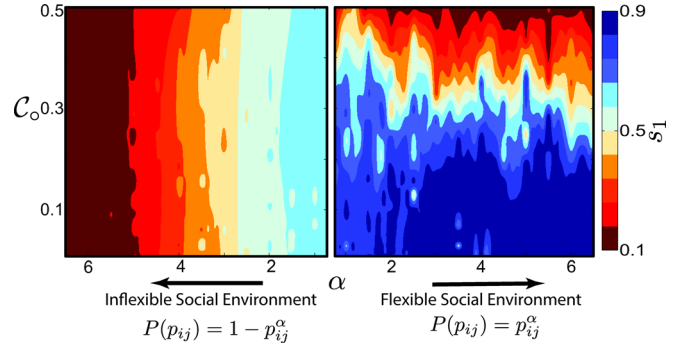

FIG. 6. A phase diagram for s_1 (size of the largest connected component) varying α and $\mathcal{C}_\circ$ in both the inflexible and the flexible social environments. Colors represent the values of s_1 (see the color bar). The left panel belongs to the inflexible social environment regime whereas the right panel belongs to the flexible social environment regime. Observe the disintegration of the largest connected components in the flexible social environment regime (the right panel) for higher values of initial clustering $\mathcal{C}_\circ$ (lower values of s_1 in shades of red). For lower values of $\mathcal{C}_\circ$ we do not observe any such disintegration (higher values of s_1 in shades of blue). In the inflexible regime (the left panel) we observe that values of α dominate the final outcome of the simulation. A network of $N = 1000$ nodes and with $O = 100$ initial opinions was employed for each α and $\mathcal{C}_\circ$. For visualization, data were interpolated onto a regular grid by a combination of natural neighbor and spline interpolation.
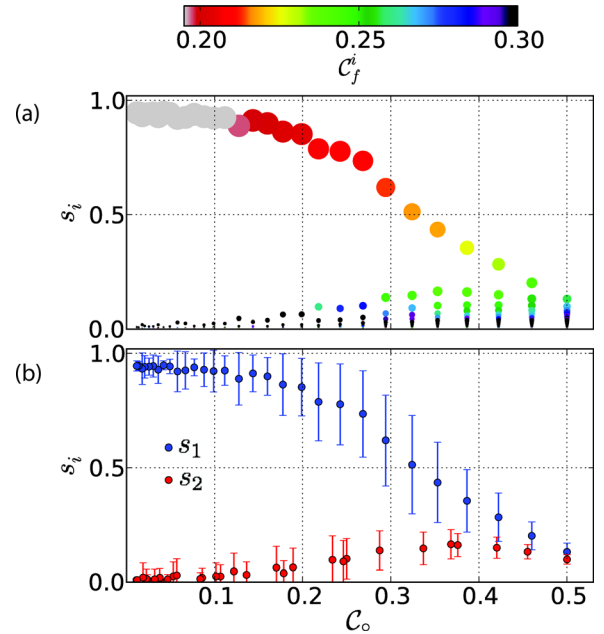

FIG. 7. The effect of different initial clustering $\mathcal{C}_\circ$ on a network of $N = 1000$ nodes with $O = 100$ initial opinions for flexible social environment $P(p_{ij}) = p_{ij}^\alpha$ with $\alpha = 6$. Large initial clustering $\mathcal{C}_\circ$ leads to final states with *segregated consensus*, contrary to the expected *hegemonic consensus* for initially unclustered networks in the same flexible social environment. (a) The distribution of component sizes s_i, the fraction of nodes in the ith (ranked by size) connected component in the final consensus state, is plotted as a function of the initial clustering coefficient, $\mathcal{C}_\circ$ (with marker sizes proportional s_i). Colors indicate $\mathcal{C}_f^i$, the clustering coefficient of the ith component in the final state. (b) The sizes of the two largest connected components, s_1 and s_2, are plotted versus $\mathcal{C}_\circ$. Simulations were conducted on 100 realizations of the network and initial opinion distribution, with the plotted component sizes estimated as the means over these realizations. Error bars give the standard deviation of these sizes across realizations.

inflexible to flexible social environments, fewer and fewer initial opinions survive, with the most extreme case being where only one dominant opinion survives with formation of a *hegemonic consensus*. Here we will attempt to infer from numerical simulations whether these transitions have a finite size effect.[41] The complexities involved in this model makes analytical analysis hard, but it is possible to obtain a variety of details using numerical simulations.

From Fig. 3 we observe that somewhere when the parameters of the model are in the *inflexible* social environment regime there is emergence of smaller sized connected components. Hence, we will focus on those transitions which occur within the parameter setting of the *inflexible* social environment, i.e., $P(p_{ij}) = 1 - p_{ij}^\alpha$. For all of the simulations below, we have used an Erdős-Rényi random network as the initial network at the start of the simulation. In Fig. 8(a) we observe multiple transitions in the system when α is varied from 0.5 to 6.0 for inflexible social environments. The first transition is visible in the size of s_1, where a weak dependence on the size of the system seems to emerge [see inset Fig. 8(a)]. The data from different system sizes appear to collapse onto a single curve when a small factor $N^{-0.05}$ is multiplied to α. That is, it appears that this transition value of α has dependence on the size of the system proportional to $N^{0.05}$ [see vertical lines in the inset of Fig. 8(a)], and this transition point would move to infinity in the thermodynamic limit.

A second transition for finite systems appears to occur near $\alpha = 4.25$ where the best fit to the data points changes from a polynomial fit to power law fit [see Figs. 8(b) and 8(c) and Figs. 8(e) and 8(f)]. All fitting reported here has been obtained using a least squares routine provided in SciPy's optimize package, which uses minpack's lmdif and lmder algorithms.[42] This transition is more apparent in Fig. 8(d) for the size of the second largest connected component, s_2. In Fig. 8(g), we have plotted the Shannon entropy over the sizes of the 10 largest components, $H = \sum_{i=1}^{i=10} s_i \ln(s_i)$. Considering only 10 largest components for this calculation is a close approximation to the total Shannon entropy of the size distribution in most cases, given the rapid decrease in the tail of the size distribution. In this figure, the transition near $\alpha = 4.25$ is visibly very much apparent as H tends to

saturate and then start to decrease. The polynomial fit in Fig. 8(a) has the following form:

$$
\begin{aligned}
s_1 &= a\alpha^2 + b\alpha + c && \text{if } \alpha < 4.25, \\
s_1 &\sim f(N)\alpha^{-2.4 \pm 0.02} && \text{if } \alpha \geq 4.25,
\end{aligned}
\tag{1}
$$

where $a \approx -0.029$, $b \approx N^{0.052 \pm 0.001} - 1.4$, $c \approx N^{-0.36}\log(N)$, and $f(N)$ is a function dependent on N. A similar analysis for s_2 also yields a polynomial fit

$$
\begin{aligned}
s_2 &= a\alpha^{-2.1} + b\alpha^{2.1} + c && \text{if } \alpha < 4.25, \\
s_2 &\sim f(N)\alpha^{1.42 \pm 0.12} && \text{if } \alpha \geq 4.25,
\end{aligned}
\tag{2}
$$

where $a \sim N^{0.0027} - 1.02$, $b \approx -2.68^{-6}N - 1.54$, and $c \sim N^{1.75}$, and again $f(N)$ is a function dependent on N. This analysis brings out a highly complex dependence of s_1 and s_2 on system size for the transition occurring near $\alpha = 4.25$. However, as indicated by the errors to the polynomial fit and power law fits in Figs. 8(b) and 8(c) and 8(e) and 8(f), a polynomial fit becomes systematically less erroneous as N is increased. It remains possible that for large N these multiple transitions might coalesce into a single continuous transition.

B. Role of network structure in transitions

Social networks are generally known to have higher clustering.[43] The initial definition of global and local clustering was in the context of social ties.[13,14,29,44] In previously studied coevolving voter models with random rewiring the clustering tends to decay away to that of independently distributed edges ($\sim 1/N$) as the system evolves with time.[1,16–18] In contrast, because the present model reinforces clustering, we observe non-trivial clustering is sustained throughout the dynamics, never dropping to near zero (see Fig. 5).

Such a model provides an opportunity to explore the influence of the clustering coefficient on transitions between the formation of a *hegemonic consensus* and *segregated consensus*. We are here mainly interested in knowing whether $C_\circ$, the initial clustering, can affect the formation of the *hegemonic consensus*. We know from the discussion above that if we set $P(p_{ij}) = p_{ij}^\alpha$ (flexible social environment) with $\alpha = 6$
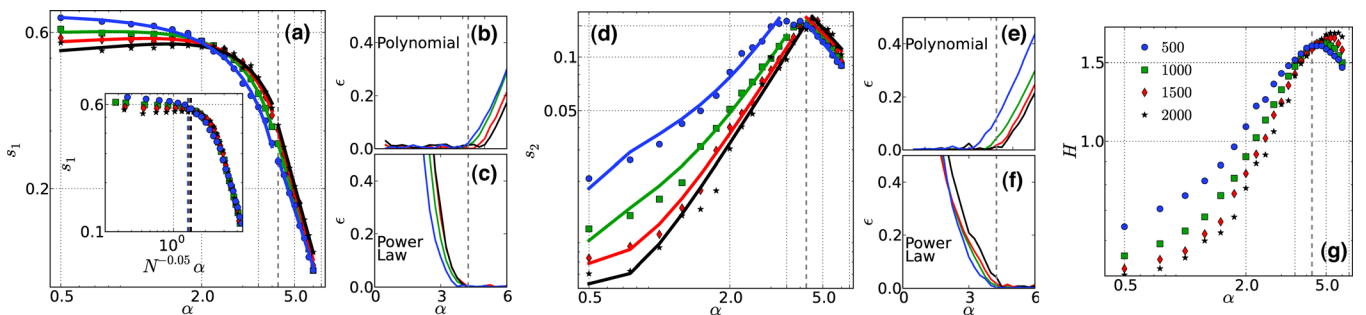

FIG. 8. Variation of s_1 (size of largest connected component) and s_2 (size of second largest connected component) with α for *inflexible* social environments, $P(p_{ij}) = 1 - p_{ij}^\alpha$. Different shapes and colors of the markers represent networks of different sizes [see legend in (g)]. In (a) we observe multiple transitions in s_1 collapse onto a similar curve (inset) for rescaling α by $N^{0.05}$. A second transition is observed near $\alpha = 4.25$ [dashed grey vertical lines in (b) and (c)], where a best fit to the data changes from a polynomial to power law, as indicated by the values of ϵ, the errors between the fitted function and the s_1 data points. This second transition appears to be collocated with a transition in s_2 appearing in (d), with an abrupt decreasing of s_2 after $\alpha = 4.25$ (dashed grey vertical line). Similar to (b) and (c), the best fit to the s_2 data changes from a polynomial to power law [see (e) and (f)]. In (g), we plot the Shanon entropy H of the 10 largest connected components versus α, observing that H tends to saturate near $\alpha = 4.25$ (dashed grey vertical line) and decreases for higher α.

and an initial random network of independently distributed edges (or network with negligible clustering coefficient), we will get a *hegemonic consensus* as the final state, with the size of the largest connected component $s_1 \sim 1$. If we instead vary the initial clustering $\mathcal{C}_\circ$ of the system, employing a Watts-Strogatz model for the initial network, we observe (in the inset of Fig. 9) that with increasing initial clustering the largest connected component tends to disintegrate into smaller sizes. For even higher $\mathcal{C}_\circ$, rather then having only one dominant connected component of size $s_1 \sim 1$, we get multiple smaller similarly-sized connected components, i.e., *segregated consensus* occurs in place of *hegemonic consensus*. So, even in the case of a highly flexible social environment ($P(p_{ij}) = p_{ij}^\alpha$, $\alpha = 6$), we can still get disintegration and no single dominant opinion if the initial clustering of the network is high enough.

To get an estimate on the critical values $\mathcal{C}_\circ$ where we could start observing the disintegration in the consensus state, we further analyze the results obtained in Fig. 9. We observe that if we multiply $\mathcal{C}_\circ$ by a factor of $\log(N)$, the data collapse onto one curve (see Fig. 9), implying that transition seems to be occurring at $\mathcal{C}_\circ \sim 1/\log(N)$. If we plot the transition scales $\frac{1}{\log(N)}$ (as done in the inset of Fig. 9 by means of vertical lines), we observe drop off in the values of s_1 starts near these scales. The form of the function fitted to the data in Fig. 9 is as follows:

$$
s_1 \sim
\begin{cases}
1 & \text{if } \mathcal{C}_\circ \leq \dfrac{1}{\log(N)} \\
a\mathcal{C}_\circ^\gamma \exp(-\lambda\mathcal{C}_\circ) & \text{if } \mathcal{C}_\circ > \dfrac{1}{\log(N)},
\end{cases}
$$

where, $\lambda \sim N^{-0.37 \pm 0.018}$, $a \sim N^{-0.95 \pm 0.07}$, and $\gamma \sim N^{-0.13 \pm 0.012}$. Though this functional form has a complex dependence on system size, the critical values $\mathcal{C}_\circ$ appear to be varying as $\frac{1}{\log(N)}$. Hence, this transition would exist in a finite network and the critical value of $\mathcal{C}_\circ$ would tend to zero in the thermodynamic limit.
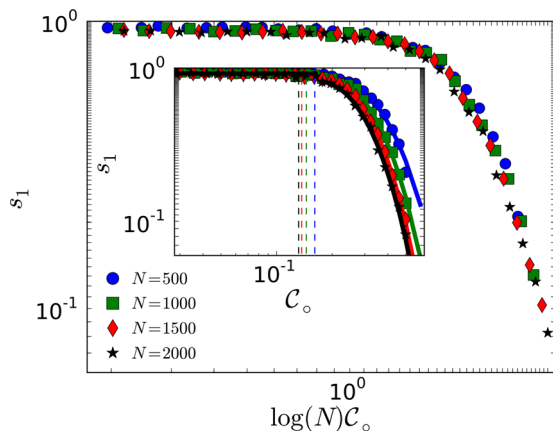

FIG. 9. Variation in the size of the largest connected component s_1 with initial clustering coefficient for flexible social environment, $P(p_{ij}) = p_{ij}^\alpha$, with $\alpha = 6$. When $\log(N)$ is multiplied to $\mathcal{C}_\circ$ the data for different system sizes appears to collapse onto a single curve. The inset curve shows the fits to data without scaling, with vertical lines $\frac{1}{\log(N)}$ indicating the scales of the transition points.
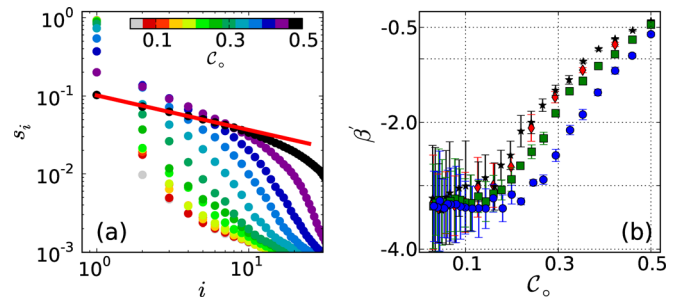

FIG. 10. The sizes of different connected components in the consensus state for networks of $N = 1500$ nodes. (a) Sizes of connected components v. their ordered (by decreasing size) indices. As initial clustering of the network $\mathcal{C}_\circ$ (color bar) is increased, there is emergence of smaller components of comparable sizes. (b) The values of the exponents, β', of the slopes fitted to the sizes of components in the final consensus state v. indices at each value of $\mathcal{C}_\circ$ [the thick red line in (a) is an example for $\mathcal{C}_\circ = 0.5$]. In (b), observe the decrease in the slope and error bars for higher initial clusterings, indicating the formation of several components of comparable sizes.

A further analysis of the connected components formed in *segregated consensus* shows that their sizes are approximately power law distributed. In Fig. 10(b) we have plotted the slope of the line fitted to the sizes of connected components and in Fig. 10(a) there is an illustration of the same for $N = 1500$ nodes. As $\mathcal{C}_\circ$ increases, the slope becomes smaller and the error bar to the fit is reduced, indicating that sizes of the connected components are becoming comparable as $\mathcal{C}_\circ$ is increased, i.e., similar sized contrarian social groups or *cults* are formed. Importantly, we also note from Fig. 7 that these similar sized components generally have very high clustering.

IV. CONCLUSIONS

We have considered a model for opinion formation on coevolving networks with two additional attributes: social environment, modeled by a distribution of susceptibilities to opinion change, and a path-length-based preference for rewiring that reinforces social clustering. The social clustering component intrinsically links the topological evolution of the network with the processes involved in collective opinion formation and vice versa.

We observed that two qualitatively distinct final states can emerge in this model. In *hegemonic consensus*, a dominating large connected component with each node within the component having the same opinion. Importantly, this dominating large connected component also maintains nontrivial local clustering. Such clustering contrasts with the properties of previously studied models, as random rewiring in those models leads to non-clustered random networks as the final consensus state. The other outcome that emerges under parameter settings of inflexible social environments is the disintegration of the network and formation of small isolated components consisting of nodes holding the same opinion. As a feature qualitatively similar to the *segregation* of individuals in a society, we have named this final state a *segregated consensus*.

A fundamentally key aspect we have studied using the features of this model is the role of clustering in the coevolving network/opinion process since the clustering of the network is continually reinforced by the preference to rewire to

nodes at smaller path length in this model. We observed that if the initial network has clustering above a critical value, then even in a flexible social environment we get *segregated consensus* as the final state. This is contrary to what happens with a network having initially negligible clustering (as in a random network of independently distributed edges). Injection of this additional attribute to the model makes the dynamics of this system richer and more relevant to social networks, but at the price of making any analytical study much more difficult than for other models, such as discussed in Refs. 16–23.

One can observe similar features in the process of opinion formation in society, for example, *hegemonic consensus* may be analogous to situations in multi-party democratic elections where one party wins by a landslide. In contrast, some hung elections may be similar to a *segregated consensus*.[45] A similar situation can also occur when choices are made on a product among many available brands, with monopoly of one brand over the product being a *hegemonic consensus* and *segregated consensus* being when there is more even competition over a product between different brands.[46]

Further analysis of the transitions in numerical simulations of different sizes has indicated complex and weak dependence on system size. In particular, it is possible that the multiple transitions induced by variations in social environment might coalesce into a single continuous transition for large systems. Meanwhile, the scaling of the transition induced by clustering in the initial network indicates that it may only exist for a finite system. Importantly, because this latter transition occurs for initial clustering $\sim 1/\log(N)$ (cf., independently distributed edges giving clustering $\sim 1/N$), we note that one should be careful making any claims about the applicability of coevolving network models that lack reinforcement of clustering to real-world network situations that have non-trivial transitivity.

ACKNOWLEDGMENTS

[1]P. Holme and M. Newman, Phys. Rev. E **74**, 056108 (2006).

[2]R. Srinivasan, "London, Egypt and the nature of social media," The Washington Post, August 11, 2011, see http://articles.washingtonpost.com/2011-08-11/national/35271787_1_social-media-social-networks-mobile-communication.

[3]J. Ball and P. Lewis, "Riots database of 2.5m tweets reveals complex picture of interaction," The Guardian, August 24, 2011, see http://www.guardian.co.uk/uk/2011/aug/24/riots-database-twitter-interaction.

[4]C. Huang, "Facebook and Twitter key to Arab spring uprisings: Report," The National, June 6, 2011, see http://www.thenational.ae/news/uae-news/facebook-and-twitter-key-to-arab-spring-uprisings-report.

[5]P. N. Howard, A. Duffy, D. Freelon, M. Hussain, W. Mari, and M. Mazaid, "Opening closed regimes: What was the role of social media during the Arab spring?," Seattle, PIPTI, 2011, see http://pitpi.org/index.php/2011/09/11/opening-closed-regimes-what-was-the-role-of-social-media-during-the-arab-spring/.

[6]P. Ball, *Why Society is a Complex Matter* (Springer, 2012).

[7]C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81**, 591 (2009).

[8]J. Onnela and F. Reed-Tsochas, Proc. Natl. Acad. Sci. U.S.A. **107**, 18375 (2010).

[9]D. Centola, Science **329**, 1194 (2010).

[10]D. Kempe, J. Kleinberg, S. Oren, and A. Slivkins, in *Proceedings of European Commission* (2013).

[11]R. M. Bond, C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler, Nature **489**, 295 (2012).

[12]Y. Moreno, M. Nekovee, and A. F. Pacheco, Phys. Rev. E **69**, 066130 (2004).

[13]S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge University Press, 1994).

[14]F. Vega-Redondo, *Complex Social Networks* (Cambridge University Press, 2007).

[15]V. Sood and S. Redner, Phys. Rev. Lett. **94**, 178701 (2005).

[16]R. Durrett, J. P. Gleeson, A. L. Lloyd, P. J. Mucha, F. Shi, D. Sivakoff, J. E. S. Socolar, and C. Varghes, Proc. Natl. Acad. Sci. U.S.A. **109**, 3682 (2012).

[17]F. Vazquez, V. Eguíluz, and M. S. Miguel, Phys. Rev. Lett. **100**, 108702 (2008).

[18]D. Kimura and Y. Hayakawa, Phys. Rev. E **78**, 016103 (2008).

[19]W. Weidlich, *Sociodynamics: A Systematic Approach to Modelling the Social Sciences* (Harwood, Academic Amsterdam, 2000).

[20]T. C. Schelling, *Micromotives and Macrobehaviour* (W.W. Norton, New York, 1978).

[21]A. M. Timpanaro and C. P. C. Prado, Phys. Rev. E **80**, 021119 (2009).

[22]G. Zschaler, G. A. Böhme, M. SeiBinger, C. Huepe, and T. Gross, Phys. Rev. E **85**, 046107 (2012).

[23]G. A. Böhme and T. Gross, Phys. Rev. E **85**, 066117 (2012).

[24]J. P. Gleeson, D. Cellai, J.-P. Onnela, M. A. Porter, and F. Reed-Tsochas, e-print arXiv:1305.7440 (2013).

[25]S. Redner, Eur. Phys. J. B **4**, 131 (1998).

[26]D. Volovik and S. Redner, J. Stat. Mech. **2012**, P04003 (2012).

[27]E. Barnett and M. Casper, Am. J. Public Health **91**, 465 (2001).

[28]E. S. Herman and N. Chomsky, *Manufacturing Consent: The Political Economy of the Mass Media* (Pantheon Books, 1988).

[29]D. J. Watts and S. Strogatz, Nature **393**, 440 (1998).

[30]M. McPherson, L. Smith-Lovin, and J. M. Cook, Annu. Rev. Sociol. **27**, 415 (2001).

[31]R. N. C. Filho, M. P. Almeida, J. S. Andrade, Jr., and J. E. Moreira, Phys. Rev. E **60**, 1067 (1999).

[32]A. Maulana and H. Situngkir, *Power Laws in Elections: A Survey* (Bandung Fe Institute, 2011), see http://cogprints.org/6934/.

[33]J. Farmer and J. Geanakoplos, "Power laws in economics and elsewhere," Technical Report, Santa Fe Institute, 2006.

[34]A. Clauset, C. R. Shalizi, and M. E. J. Newman, SIAM Rev. **51**(4), 661 (2009).

[35]M. Stumpf and M. Porter, Science **335**, 665 (2012).

[36]A. D. Henry, P. Prałat, and C.-Q. Zhang, Proc. Natl. Acad. Sci. **108**, 8605 (2011).

[37]M. Fossett, J. Math. Sociol. **30**, 185 (2006).

[38]M. Pollicott and H. Weiss, Adv. Appl. Math. **27**(1), 17–40 (2001).

[39]S. N. Dorogovtsev, *Lectures on Complex Networks* (Oxford University Press, Oxford, 2010).

[40]M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).

[41]R. Toral and C. J. Tessone, Commun. Comput. Phys. **2**, 177 (2007).

[42]E. Jones, T. Oliphant, P. Peterson *et al.*, SciPy: Open source scientific tools for Python, 2001, see http://www.scipy.org/.

[43]S. Wasserman, *Social Network Analysis: Methods and Applications* (Cambridge University Press, Cambridge, 1994).

[44]P. W. Holland and S. Leinhardt, Comp. Group Studies **2**, 107 (1971).

[45]N. Schofield and I. Sened, *Multiparty Democracy: Elections and Legislative Politics* (Cambridge University Press, 2006).

[46]T. J. Webster, *Managerial Economics: Theory and Practices* (Academic Press, 2003).