The Role of Active Leaders in Opinion Formation on Social Networks

F. Baumann¹, I. M. Sokolov¹,², and M. Tyloo³,⁴

¹Institute of Physics and ²IRIS Adlershof, Humboldt-Universität zu Berlin, Newtonstraße 15, 12481 Berlin, Germany.
³Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.
⁴School of Engineering, University of Applied Sciences of Western Switzerland HES-SO, CH-1951 Sion, Switzerland.

(Dated: October 7, 2019)

The intentional polarization of opinions and controlled changes of a consensus represent potentially harmful processes for any liberal society. Within the framework of a simple model for constructive opinion exchange, we analytically study the role of active leaders on the overall opinion dynamics on social networks. Similar to zealots with rigid opinions, active leaders are assumed to follow a prescribed and individual agenda. We characterize both the ability of single active leading agents to shift a prevailing consensus as well as the emerging social polarization in the case of two antagonistic opinion leaders. To this end we define measures for the coherence of opinions and their polarizability and relate those to graph based quantities derived from a modified Laplacian of the system. We apply the formalism to synthetic and empirical networks and find that while polarizability is decreased for small-worldness and favored in the case of a pronounced community structure the opposite is true for the coherence of opinions during a consensus change.

I. INTRODUCTION

Communication infrastructures represent highly networked systems. This holds true in the case of face-to-face [1] interactions but also and especially for large online social media platforms, which dominate the information exchange among millions of people on a daily basis. Although embedded in a connected topology, the potential impact of individual agents differs widely across the system. Naturally, this leads to the notion of influential agents, which play a crucial role due to their particular position in the network.

In this work we focus on the role of active leaders with respect to opinion formation processes in social systems. We aim to identify nodes in the social network which are particularly suitable to be occupied by opinion leaders to maximize their influence on the system. Such influential agents have mostly been identified and studied in the context of rumor and epidemic spreading [2-3]. Less attention has been paid to the role of active leader placement in the context of opinion formation on social networks. Here we investigate potentially alarming developments of opinion interference such as the controlled change of an existing consensus or the intentional polarization of opinions. Real-world correspondences to our theoretical considerations might include phenomena such as the targeted disruption of a scientifically established consensus, e.g. on climate change [5] or antagonistically reporting news agencies responsible for a measurably increased opinion polarization around certain political issues [6]. To identify the most relevant agents in the system, we utilize an adaption of the simplest model for constructive opinion dynamics [7, 8]. Here, the ensemble of agents is partitioned into passive followers and active leaders which interact on a fixed network topology. We assume that while followers exclusively average their neighbors opinions, active leaders are subject to an additional signal. In the light of opinion exchange such signals can be interpreted twofold. Individual preferences or biases might render agents stubbornly focused on a certain viewpoint [9, 10], conceptionally similar to zealots in social networks [11]. Or, in contrast to any internal preferences, the leader-follower setup yields an exemplary social setting in which a fixed subset of agents gets influenced from outside the system. Either way, leaders introduce heterogeneity to the dynamics and trigger the emergence of rich dynamical behaviors and relevant stationary opinion states.

We tackle both the linear dynamics of opinion exchange and the underlying network topology in a unified framework. We study at which position in the network active leaders need to be placed to drive the system away from consensus in two different relevant ways: First, we analytically characterize the ability of single active leaders to change the overall population consensus to a new desired value. By defining a measure for network coherence we identify agents who are able to alter the global opinion while being closely followed by the remaining agents during the transition from one consensus to another. Furthermore, we study opinion polarization as a relevant stationary state. It occurs upon the introduction of two antagonistic active opinion leaders, modeling single activists or news agencies [12] following prescribed opposed agendas. We find that the strength of the developing polarization crucially depends on positions of active agents in the network. As in the case of single leaders we define useful graph based measures to classify social networks with respect to polarizability. Applying our formalism to different synthetic and empirical networks we find that community structure and small-worldness strongly affect the coherence as well as the polarizability of opinions in interacting social networks.

The paper is organized as follows. In Sec. 1 the model for leader-follower dynamics is defined. Section 2 discusses properties of a modified Laplacian of the system from which a new distance measure is deduced, similarly defined as the resistance distance of an electrical net-
work. Section IV investigates single leader dynamics with respect to network coherence. In Sec. V we discuss properties of the system in the case of two antagonistic active leaders. While we present efficient methods to evaluate the arising opinion polarization in the system our formalism is able to directly associate followers to one of the two opposed active leaders. Sec. VI illustrates our theory numerically on synthetic and empirical networks. The work is concluded in Sec. VII.

II. LEADER-FOLLOWER DYNAMICS

We consider a system of interacting agents, which has previously been studied in [13]. Each agent $i$ is characterized by a continuous opinion variable $x_i \in \mathbb{R}$ and can be assigned to one of the two subgroups of leaders ($V_l$) or followers ($V_f$). The dynamics of opinion exchange are governed by the following set of differential equations,

\begin{align}
\dot{x}_i &= -\sum_j b_{ij} (x_i - x_j) - \kappa (x_i - P_i), \quad i \in V_l, \tag{1a} \\
\dot{x}_i &= -\sum_j b_{ij} (x_i - x_j), \quad i \in V_f, \tag{1b}
\end{align}

where $b_{ij}$ denotes the coupling strength between agents $i$ and $j$. We assume a symmetric connectivity structure and solely constructive dynamics, i.e. $b_{ij} = b_{ji}$ and $b_{ij} \geq 0$. Due to the averaging mechanism of Eqs. (1), follower agents $i$ simply aim to minimize the difference between their own opinions $x_i$ and those of their neighbours $j$ specified by the adjacency matrix $b_{ij}$. By contrast, active leaders are subject to an additional constraint encoded by the last term in Eq. (1a): with rate $\kappa$ they aim to converge to a well-defined constraint opinion value $P_i$. In our model constraint opinions can be individually chosen for each active leader which generalizes previous work where only a single leading opinion $P$ was assumed [13]. Fig. 1 shows example dynamics of our model system with a single (left panel) and two opposing active leaders (right panel). One observes that, after a transient, in the single leader case all opinions converge to the constraint opinion, $P_a$, of the active single leader. Instead, for two active leaders, the final opinions are distributed over an interval of finite width. In the next sections, we analytically describe both the transient behavior of the agents and the final state of opinions in terms of simple network based quantities. Defining the vector $\mathbf{x} = (x_1, \ldots, x_n)$, with $n = n_l + n_f$ being the total number of nodes, Eqs. (1) can be rewritten in a vectorial form as,

\[ \dot{\mathbf{x}} = -(\mathbf{L} + \mathbf{K}) \mathbf{x} + \mathbf{P}^c, \tag{2} \]

where $\mathbf{K}$ is a diagonal matrix with vanishing components for follower agents and a global $\kappa$ value for leading agents. Similarly, $\mathbf{P}^c$ has vanishing components for followers and $P^c_i = \kappa P_i$ for active leaders and $\mathbf{L}$ denotes the Laplacian of the coupling graph, defined as

\[ \mathbf{L}_{ij} = \begin{cases} -b_{ij}, & i \neq j, \\
\sum_k b_{ik}, & i = j. \end{cases} \tag{3} \]

As we consider undirected connectivity graphs both $\mathbf{L}$ and $\mathbf{L}^c$ are symmetric matrices. In the following we denote eigenvectors and eigenvalues of $\mathbf{L}^c$, as $\mathbf{u}_k^c$ and $\lambda_k^c$, respectively and discuss important properties of this modified Laplacian matrix which will be used to derive our theoretical results.

III. LEADER DYNAMICS AND MODIFIED RESISTANCE DISTANCES

Exploiting the linearity of Eqs. (1) we can fully express the transient dynamics as well as the final state of the system in terms of $\mathbf{u}_k^c$ and $\lambda_k^c$. Based on the modified Laplacian $\mathbf{L}^c$, we define a new distance metric $\Omega_{ij}^{(k,1)}$ which is similar to resistance distance as introduced in [13]. This metric, which takes into account both Eqs. (1) and the interaction topology $\mathbf{B}$, will be related to important dynamical characteristics of the system. Thus, we first explore important properties of $\mathbf{L}^c$ and present definitions of node and network descriptors.
A. Properties of $L^\kappa$

Depending on the number of active leaders the matrix $L^\kappa$ and its associated eigenvalues have important properties. Elementwise, the modified Laplacian $L^\kappa$ reads,

$$L^\kappa_{ij} = \mathbb{I}_{ij} + \kappa \sum_{a \in V_l} \delta_{ij}\delta_{ia},$$  \hspace{1cm} (4)

where $\mathbb{I}$ satisfies $\sum_{j} L_{kj} = 0 \forall k$. The matrix elements of the inverse of $L^\kappa$ along rows and columns corresponding to leader nodes satisfy accordingly,

$$\sum_{a \in V_l} [L^\kappa]^{-1}_{ia} = 1/\kappa.$$  \hspace{1cm} (5)

which holds true for $n_l \geq 1$ (see proof in Appendix A). Note, that $[L^\kappa]^{-1}_{ij}$ denotes the $ij$-th element of the inverse of $L^\kappa$. Particularly, in the considered case of two active leaders ($V_l = \{a, b\}$) Eq. (5) leads to the following relation between the diagonal elements corresponding to leader nodes,

$$[L^\kappa]^{-1}_{aa} = [L^\kappa]^{-1}_{bb}.$$  \hspace{1cm} (6)

In the case of $n_l > 2$, Eq. (6) can easily be generalized. For an arbitrary number of active leaders, the eigenvectors $u^\kappa_a$ satisfy

$$\sum_{i} u^\kappa_{a,i} = \kappa \sum_{a \in V_l} \frac{u^\kappa_{a,a}}{\lambda^\kappa_a},$$  \hspace{1cm} (7)

which combined with Eq. (5), gives

$$\sum_{a,b \in V_l} [L^\kappa]^{-2}_{ab} = \sum_{a,b \in V_l} \sum_{\alpha} \frac{u^\kappa_{\alpha,a} u^\kappa_{\alpha,b}}{\lambda^\kappa_\alpha} = n/\kappa^2.$$  \hspace{1cm} (8)

The above properties will be used to derive expression for the final state of opinions.

B. Modified Resistance Distances (MRDs)

The resistance distance is a graph theoretic metric originally defined by the Laplacian of an interaction network as,

$$\Omega_{ij} = \mathbb{L}_{ii}^{\dagger} + \mathbb{L}_{jj}^{\dagger} - \mathbb{L}_{ij}^{\dagger} - \mathbb{L}_{ji}^{\dagger},$$  \hspace{1cm} (9)

where $\mathbb{L}^{\dagger}$ is the pseudoinverse of $\mathbb{L}$. As $L^\kappa$ is nonsingular, a distance metric similar to $\Omega_{ij}$ can thus simply be defined using the inverse of $L^\kappa$, i.e.

$$\Omega_{ij}^{(\kappa,p)}(V_l) = [L^\kappa]^{-1}_{ii} + [L^\kappa]^{-1}_{jj} - [L^\kappa]^{-1}_{ij} - [L^\kappa]^{-1}_{ji},$$  \hspace{1cm} (10)

where $\Omega_{ij}^{(\kappa,p)}(V_l)$ depends on the set of leaders $V_l$ as well as the parameter $\kappa$. The modified resistance distance (MRD) and its higher order generalizations can be expressed using eigenvectors and eigenvalues of $L^\kappa$ as,

$$\Omega_{ij}^{(\kappa,p)}(V_l) = \sum_{\alpha} \frac{(u^\kappa_{\alpha,a} - u^\kappa_{\alpha,b})^2}{\lambda^\kappa_\alpha},$$  \hspace{1cm} (12)

The superscript $p$ denotes the generalization of Eq. (10) to the $p^{th}$ power of $[L^\kappa]^{-1}$ \[16\,18\]. Using the modified resistance distance, we introduce its corresponding closeness centrality,

$$C_p(i, V_l) = \left( n^{-1} \sum_{j} \Omega_{ij}^{(\kappa,p)}(V_l) \right)^{-1},$$  \hspace{1cm} (13)

quantifying the average MRD of node $i$ to every other node in the network \[17\,19\]. More precisely, if $C_p(i, V_l)$ is large (small), then node $i$ is central (peripheral) according the distance defined in Eq. (11). Note, that if Eq. (13) is re-expressed using Eq. (11), properties of the eigenvectors $u^\kappa_a$ which depend on $n_l$ have to be taken into account.

IV. OPINION COHERENCE FOR SINGLE ACTIVE LEADERS

In this section we analytically quantify the coherence of the system during the consensus change induced by a single active leading agent. We assume that one specified leader $i = a$, aims to change the global consensus to a new value $P_a$. Depending on the leader’s position in the network the remaining agents will closely follow its opinion trajectory or exhibit a more spread out behavior during the transition, as depicted in Fig. 1. A related problem was previously studied for a noisy system of followers and leaders in \[8\]. In order to rank the agents according to the resulting network coherence we define the measure

$$\mathcal{C} = \sum_{i} \int_{t_0=0}^{\infty} |x_a(t) - x_i(t)| dt,$$  \hspace{1cm} (14)

which assesses the system’s behavior for the crossover between two consensuses starting from $x_i(t) = 0, \forall i$ up until $x_a(t) = P_a = x_i(t), \forall i$. Note, that a large value of $\mathcal{C}$ corresponds to a low network coherence. In the case of a single leader the opinion of the $i$-th agent reads,

$$x_i(t) = \kappa P_a \sum_{\alpha} \frac{u^\kappa_{\alpha,a}(1 - e^{-\lambda^\kappa_\alpha t})}{\lambda^\kappa_\alpha} u^\kappa_{\alpha,i}.$$  \hspace{1cm} (15)

Taking $t \to \infty$ and using Eq. (5) yields,

$$x_i(t \to \infty) = \kappa P_a \sum_{\alpha} \frac{u^\kappa_{\alpha,a} u^\kappa_{\alpha,i}}{\lambda^\kappa_\alpha} = P_a,$$  \hspace{1cm} (16)
which ensures a complete consensus in the long term limit for the single leader case. Integrating Eq. (15), \( C \) can be expressed as,

\[
C = \kappa P_a \sum_i \sum_\alpha \frac{u_{\alpha,a}^\kappa - u_{\alpha,b}^\kappa}{\lambda_\alpha^2},
\]

\[
= \kappa P_a \left( n C_2^{-1}(a, \{a\}) + \sum_\alpha \left[ \frac{u_{\alpha,a}^\kappa}{\lambda_\alpha^2} - \lambda_\alpha^{-2} \right] \right).
\]

(17)

Hence, in order to drive all agents to the desired value while maintaining a high network coherence the single active leader should be selected such that s/he exhibits a high centrality \( C_2 \), as defined in Eq. (13). The second term in Eq. (17) indicates that, in order to minimize the spreading of opinions, the leader node should be selected such that \( u_{\alpha,a} \) is maximized/minimized on eigenvectors corresponding to large/small \( \lambda_\alpha^\kappa \). Finally the last term in Eq. (17) corresponds to the sum of the square of the timescales of the system, namely the eigenvalues of \( L^\kappa \). To assess the overall coherence of a network, we define \( \bar{C} \) where Eq. (17) is averaged over all possible leader nodes, i.e.

\[
\bar{C} = \langle C \rangle_{\{a\}}.
\]

(18)

This averaged coherence measure quantifies the overall drivability of a network of agents. Equation (18) is numerically investigated on different interaction networks in Sec. VI.

V. TWO OPPOSING LEADERS

In the context of political discussions on social networks, opinion polarization has been quantified empirically [20]. We define an opinion state as polarized, as soon as it deviates from situations of perfect consensus, i.e. \( x_i = c, \forall i \). In our model such heterogeneous opinion distributions can only be stable over time if multiple active leaders with different constraint opinions \( P_i \neq P_j \) are introduced [see Eq. (16)]. Hence, in the following, we consider the prototypical case leading to polarization of two opposed active leaders \( (i = a, b) \) and assume all other agents to be followers. In order to single out implications of differently strong constraint opinions and focus solely on the effects of the social network and the positionings of active leaders therein, we assume leaders to be balanced, i.e. \( P_a = P_b = -P_b \), cf. Fig. [1]. Starting from an initial consensus, i.e. \( x_i(t_0) = 0 \), the time evolution of the system is formally solved by,

\[
x(t) = \kappa P \sum_\alpha \frac{u_{\alpha,a}^\kappa - u_{\alpha,b}^\kappa}{\lambda_\alpha^\kappa} (1 - e^{-\lambda_\alpha t}) u_{\alpha}^\kappa
\]

(19)

which yields the starting point of the following considerations. We will analytically quantify the polarizability of arbitrary social networks and efficiently solve the leader association problem of stationary opinion states. With regard to empirical networks, the developed methods can be applied in order to answer urgent questions of democratic opinion formation and the influence of highly active or stubborn individuals on the opinion formation processes.

A. Leader association

Given a fixed pair of opposed leaders the question arises which followers will be associated to which active leader. We show that this problem can be reduced to the evaluation of the corresponding MRDs, which represents a very efficient tool especially for large social networks. While the population is initialized at \( x_i(t_0) = 0, \forall i \), the introduction of two opposed active leading agents with \( P_a > 0 > P_b \) generally shifts the opinion of each follower either to the positive or negative side, defining its resulting opinion leader association as \( \text{sgn}(x_i) \pm 1 \). According to Eq. (19) the final opinion of an arbitrary follower \( i \) can be expressed using Eq. (6) as

\[
x_i^\infty = \kappa P \sum_\alpha \frac{u_{\alpha,a}^\kappa - u_{\alpha,b}^\kappa}{\lambda_\alpha^\kappa} u_{\alpha,i}^\kappa
\]

\[
= \frac{\kappa P}{2} \left( \Omega_{\{a,b\}}(\{a,b\}) - \Omega_{\{a,b\}}(\{a,b\}) \right).
\]

(20)

Interestingly, Eq. (20) suggests that an agent \( i \) will be associated to the closest of the two leaders with respect to the \( \Omega_{ij}^{(\kappa,1)} \). Fig. [2] illustrates the leader association exemplarily for four different types of synthetically generated networks. Fig. [2] shows the final association for a Watts-Strogatz (WS) graph [21] obtained for a low edge rewiring probability. Even though the MRDs between leader \( i = b \) (red cross) and the right half part of the cycle have been reduced by the rewiring procedure, its distance to the left half part of the cycle has merely been changed. Consequently, the final state of opinions splits into two associative parts of roughly equal size. Generally, the average MRD for node \( i, (\Omega_{ij}^{(\kappa,1)})_j \), is largely reduced for high degree nodes, which corresponds to a high centrality as defined in Eq. (13). Therefore, agents associated with hubs will generally attract more followers in contrast to agents sitting on low degree nodes at the periphery of the network. Such a case is depicted in Fig. [3] for a scale free Barabási-Albert network. In Fig. [2], we consider the case of a network generated by the stochastic block model (SBM) [22] where each of the two leaders is placed within one of two separate communities. The high internal connectedness within each community heavily reduces the MRDs with respect to in-versus out-community node pairings. This leads most probably to a situation in which all nodes of a community are associated to the leader which is situated within that community. For a regular network as in Fig. [2], a leader position within the lattice away from the boundaries, favors small MRDs to most other agents in the system and
defined as $D_L = x_a^\infty - x_b^\infty$, reduces from Eq. \[20\] to

$$D_L([a, b]) = \kappa P \sum_{\alpha} \frac{(u^{a}_{\alpha, a} - u^{b}_{\alpha, b})^2}{\lambda^{a}_{\alpha}} = \kappa P \Omega^{(\alpha, 1)}_{ab}. \tag{21}$$

Using $D_L$ we quantify how strongly those opinions diverge. Moreover we derive expressions for the mean $\mu_x = \frac{1}{n} \sum_i x_i$ and the variance $\sigma^2_x = \frac{1}{n} \sum_i (x_i - \mu_x)^2$ of the resulting opinion distributions, i.e.

$$\mu_x([a, b]) = \frac{\kappa P}{2} \left[ C^{-1}_1(b) - C^{-1}_1(a) \right] \tag{22}$$

$$\sigma^2_x([a, b]) = \left( \frac{\kappa P}{2} \right)^2 \left( \frac{4\Omega^{(\alpha, 2)}_{ab}}{n} - \left[ C^{-1}_1(b) - C^{-1}_1(a) \right]^2 \right), \tag{23}$$

respectively, where $\Omega^{(\alpha, 2)}_{ab}$ and $C_1$ were defined in Sec. III B. Note, that the first term in Eq. \[23\] can be also expressed by first order modified resistance distances between leaders $a$ and $b$ using $\Omega^{(\alpha, 2)}_{ab} = \frac{1}{2} \sum_i \left( \Omega^{(\alpha, 1)}_{b\alpha} - \Omega^{(\alpha, 1)}_{a\alpha} \right)^2$. Interestingly, all three introduced polarization descriptors can be formulated in terms of first order modified resistance distances $\Omega^{(\alpha, 1)}_{ij}$.

Those expressions, are all defined up to a fixed pair of leaders $([a, b])$. Hence, to obtain global indices which characterize the network’s, polarizability Eqs. \[21\], \[22\], \[23\] are averaged over all possible sets of leading pairs, yielding

$$\tilde{D}_L = \langle D_L \rangle_{[a, b]} \tag{24a}$$

$$\tilde{\mu}_x = \langle |\mu_x| \rangle_{[a, b]} \tag{24b}$$

$$\tilde{\sigma}^2_x = \langle \sigma^2_x \rangle_{[a, b]} \tag{24c}$$

For large empirical networks Eqs. \[24\] become computationally exhaustive to evaluate. Our presented graph-based approach, however, still outperforms the explicit solution of the original system of Eqs. [1] in time as a means to quantify the different aspects of opinion polarization.

\section{VI. NUMERICAL RESULTS}

We investigate numerically the global network coherence measure $\tilde{C}$ and the different quantifiers for opinion polarization, $\tilde{D}_L$, $\tilde{\mu}_x$, $\tilde{\sigma}^2_x$, as introduced in Sec. IV and Sec. V B respectively. Using WS and SBM networks, we assess how those quantities relate to relevant network features, such as small-worldness or the peculiarity of community structure. Those network properties have been shown to play important roles in various dynamical processes on networks such as epidemic and rumor spreading [24] or oscillator synchronization [29]. Therefore we expect them also to influence aspects of opinion dynamical processes such as polarization and network coherence. As an interesting application, we additionally discuss the

Figure 2. \textbf{Final leader association of follower nodes on different network topologies.} Initialized as perfect consensus ($x_i(t_0) = 0$) the system of agents gets polarized by the introduction of two opposed leaders $a$ and $b$ with $P_a = -P_b = 10$ ($\kappa = 1$) depicted as blue and red cross, respectively. Given this fixed pair of two opposed leaders the evaluation of Eq. \[20\] allows to efficiently determine the association of each follower to one of the two leaders, cf. Eq. \[20\]. The final political leaning of each follower $\text{sgn}(x_i^\infty) = \pm 1$ is colored in blue/red, respectively, to visualize the association to the respective leader $a$ (blue) or $b$ (red) on different example networks: Watts-Stogatz network (a), scale-free Barabási-Albert network (b), stochastic block model (c) and a regular grid network (d).

B. Opinion polarization

Beyond follower-leader associations it is relevant to relate the placement of two antagonistic active leaders to the resulting opinion formation process. Hence, we study the polarization of opinions by different accessible quantities that can be derived from the modified laplacian of the system to characterize the stationary opinion states. We will focus on the resulting opinion distance between the most extreme opinions, the final mean opinion as well as the opinion variance. By that we characterize not only the heterogeneity of opinions but also quantify how strongly the system of agents tends to favor on average one of the two opposed views upon the introduction of two antagonistic active leaders.

In the polarizing scenario, the two leading agents always take on the most extreme opinions in the system. The final opinion distance between the two leading agents therefore attracts the larger number of followers.
The lower panel (black line) in Fig. 3a shows the network coherence quantifier $\hat{C}$ defined in Sec. IV for WS networks as a function of $p_r$. The coherence measure steadily decreases with the rewiring probability, with a strong decay in the small-world region around $p_r = 0.1$. While rewiring, the MRDs in the networks are reduced such that all nodes become more central [cf. Eq. (13)] and thus yielding a smaller coherence measure. For the SBM network, in Fig. 3b, $\hat{C}$ increases with $p_{\text{intra}}$. For high values of $p_{\text{intra}}$, two highly connected distinct communities emerge with few links connecting them. Hence, MRDs are increased between the communities while they are reduced within the communities. The control of the whole population becomes less efficient when only few links connect the two communities, as the leader naturally resides in one of them. Finally we consider an empirical network of face-to-face interactions of high school students depicted in Fig. 3c. The performed randomization procedure blurs the network’s socially grown structure, which is predominantly characterized by a community structure representing three different school classes. One observes that upon randomization, $\hat{C}$ is reduced. As shown with WS and SBM, networks with high average distance and distinct communities seem to be harder to drive to the desired opinion with high coherence, i.e. low values of $\hat{C}$.

Similar conclusions seem to hold true for opinion polarization in the two leader case. In the three panels of Fig. 3, $D_L$ and $\bar{\sigma}^2$ behave similarly as the network coherence discussed previously. As a matter of fact, strong community structures and absence of small-worldness, namely high clustering coefficient and average geodesic distance, tend to promote polarization phenomena. However, when considering the absolute mean opinions, one
observes with SBM (Fig. 3b) and the empirical data set (Fig.3c) that strong community structures seems to keep a balance in the opinions as μx is the smallest for high values of pntra for SBM and low number of rewiring for the empirical network. Moreover, for SBM, when pntra becomes smaller than 0.5, the interaction network tends to become a bipartite graph. Therefore, each node in one of the community is close to the other community, reducing MRDs in the whole networks and yielding small ˜μx as well as coherence and polarization quantifiers. A similar behavior is observed for WS networks in Fig. 3a. For pr = 0, the network is regular and therefore, all nodes are equivalent explaining that ˜μx is vanishing. When a few edges are rewired from the initial regular cycle, symmetries are broken and some communities with reduced MRDs are built exhibiting a large ˜μx. Edges rewired afterwards alterate the community structure, and eventually lead to an overall random network with one large community. In summary, adding long range or inter community coupling seems to prevent polarization among the population while gathering opinions further away from initial consensus, x = 0.

VII. CONCLUSIONS

In this work we considered a simple model for constructive opinion dynamics to study the role of active opinion leaders on the opinion formation in social networks. The dynamical system was reformulated in terms of a modified Laplacian, which incorporates the follower and leader dynamics as well as the underlying interaction topology. Within that framework the impact of a single leader on the coherence of the opinions as well their opinion polarization is studied in terms of quantities derived from a modified form of the resistance distance. In the case of opinion polarization in a networked social system we focus on the final leader distance, the opinion variance as well as the mean absolute opinion of the stationary opinion distribution. We present exact expressions to asses those characteristics of the expected opinion landscape. Simultaneously, our framework solves the leader association problem in a system of two opposing active leaders. The numerical evaluation of the derived analytical expressions for the global network measures of coherence and polarization is computationally quite extensive. This is due to the fact that it requires to compute the inverse of the modified Laplacian Lκ for each possible set of leader nodes, which becomes especially exhaustive for |Vl| > 1. However, we found numerically that for large networks, small values of κ and a low number of active leading agents (|Vl| < N), the resulting modified resistance distances as defined Ωij(k,p) do not deviate significantly from the values of Ωij as defined in Eq. (6). In such cases the explicit evaluation of Lκ for each possible set of leaders becomes redundant and yields a dramatically accelerated evaluation of the derived analytical expressions. Our work opens up the possibility to quantify the polarizability of large social systems if the network topology is known. Most of the presented work can easily be generalized to the case of directed networks, i.e. asymmetric coupling schemes within the respective social systems. This could correspond to scenarios in which leading agents are not subject to opinion averaging and would therefore not be influenced by the followers within the ensemble. Research along these lines is currently in progress.

VIII. ACKNOWLEDGMENT

This work was developed within the scope of the IRTG 1740/TRP 2015/50122-0 and funded by the DFG/FAPESP and the Swiss National Science Foundation under grant No. 200020_182050. We thank Philipp Lorenz-Spreen and Philippe Jacquod for interesting discussions.

Appendix A: Proofs of properties of Lκ

1. No vanishing eigenvalues

We denote uα, λα respectively the eigenvectors and eigenvalues of the Laplacian L. Eigenvalues satisfies λ1 = 0 < λ2 < ... < λn. The matrix containing leaders dynamics is given by Kij = \sum α∈Vl κδiαδjα. At first order in perturbation, we have,

\[ u'_{\alpha,i} = u_{\alpha,i} + \sum_{\beta \neq \alpha} \kappa \sum_{a \in Vl} u_{\alpha,a} u_{\beta,a} \frac{\lambda_{\alpha} - \lambda_{\beta}}{\lambda_{\alpha}} u_{\alpha,i} , \quad (A1) \]

\[ \lambda_\alpha' = \lambda_\alpha + \kappa \sum_{a \in Vl} u_{\alpha,a}^2 , \quad (A2) \]

where the second equation shows that Lκ has no vanishing eigenvalue. The eigenvalues of the perturbed Laplacian that are the most influenced by adding K are those with the largest components u_{α,a}, a \in Vl.

2. Proof Eq. (5)

Multiplying matrix Lκ by its inverse reads element-wise,

\[ \sum_j L_{ij}^{k}[L^k]^{-1}_{jk} = \sum_j \left( L_{ij} + \kappa \sum_{a \in Vl} \delta_{ai} \delta_{ij} \right)[L^k]^{-1}_{jk} \quad (A3) \]

\[ = \delta_{ik} , \quad (A4) \]

from which, summing over i, we have that,

\[ \sum_{a \in Vl} [L^k]^{-1}_{ja} = 1/\kappa . \quad (A5) \]
Both matrices $L^\kappa$ and its inverse can be written using eigenvectors and eigenvalues as,

$$L^\kappa_{ij} = \sum_\alpha \lambda_\alpha^\kappa u_{\alpha,i}^\kappa u_{\alpha,j}^\kappa,$$

(A6)

$$L^{-1} = \sum_\alpha \frac{u_{\alpha,i}^\kappa u_{\alpha,j}^\kappa}{\lambda_\alpha^\kappa}.$$  

(A7)

And thus Eq. (5) implies,

$$\sum_{\alpha \in V_i} \sum_{\alpha \in V_j} \frac{u_{\alpha,k}^\kappa u_{\alpha,a}^\kappa}{\lambda_\alpha^\kappa} = 1/\kappa.$$  

(A8)

3. Proof of Eq. (6)

In the case of two leaders $\{a, b\}$, we have for the element $(a, b)$ of the product of $L^\kappa$ by its inverse,

$$\sum_j L^\kappa_{aj} L^{-1}_{jb} = \sum_j L^\kappa_{aj} L^{-1}_{jb} + nL^\kappa_{ab}^{-1} = 0.$$  

(A9)

Exploiting the symmetry of matrices $L^\kappa$ and $L$ together with Eq. (A5) we get,

$$L^{-1}_{aa} = L^{-1}_{bb},$$

(A10)

$$\sum_j L^\kappa_{aj} L^{-1}_{jb} = 0.$$  

(A11)

With Eq. (A6) we have the relation between eigenvectors and eigenvalues,

$$\sum_\alpha \frac{u_{\alpha,a}^2}{\lambda_\alpha^\kappa} = \sum_\alpha \frac{u_{\alpha,b}^2}{\lambda_\alpha^\kappa}.$$  

(A12)

4. Proof of Eq. (7)

Multiplying $L^\kappa$ by eigenvector $u_{\alpha}^\kappa$ reads elementwise,

$$\sum_j L^\kappa_{ij} u_{\alpha,j}^\kappa = \sum_j \left(L_{ij} + \delta_{ij} \sum_\alpha \delta_{\alpha,j} \right) u_{\alpha,j}^\kappa,$$

(A13)

$$= \lambda_\alpha^\kappa u_{\alpha,i}^\kappa,$$

(A14)

then summing over $i$,

$$\kappa \sum_j \sum_\alpha \delta_{\alpha,j} u_{\alpha,j}^\kappa = \lambda_\alpha^\kappa \sum_i u_{\alpha,i}^\kappa.$$  

(A15)

Which finally yields,

$$\kappa \sum_{\alpha \in V_i} \frac{u_{\alpha,a}^\kappa}{\lambda_\alpha^\kappa} = \sum_i u_{\alpha,i}^\kappa.$$  

(A16)

5. Comparison between $\Omega_{ij}$ and $\Omega_{ij}^{(\kappa,1)}$

Figure 4 depicts scatter plots of the resistance distances based on the regular graph Laplacian $L$ and the MRDs computed from $L^\kappa$ for a WS-graph of $N = 200$ nodes of. For small $\kappa = 0.1$ and only $n_1 = 2$ active leaders $\Omega_{ij}$ and $\Omega_{ij}^{(\kappa,1)}$ are perfectly correlated, cf. Fig. 4a. This drastically changes in the case of higher $\kappa$ values and a larger set of active leaders. Figure 4b depicts this exemplarily for $\kappa = 2$ and $n_1 = 30$.

![Figure 4](image)

6. Network models and empirical data

Watts-Strogatz (WS) model. Introduced in [21] the model constructs networks in the following way: In the first step a ring of $n$ vertices, where each node is connected to its $k$ nearest neighbors is constructed. In a second step, each edge is re-wired with probability $p_r$. Typically, for increasing values of $p_r$ the networks go from a regular graphs via small-world networks to completely random topologies.

Stochastic block model (SBM): The basic concept of the SBM was originally introduced in the social sciences [22]. The principle is that nodes of a network are organized into building blocks which are then connected depending on their association to a certain group. In the simple case considered in this work we assume two similar communities $\text{Com}_1$ and $\text{Com}_2$ and introduce $p_{\text{intra}}$ stating the probability for an arbitrary connection within the community. To be able to change the community structure of the network but fix the average number of links in the network the connection probability between two arbitrary nodes in two different communities is defined as $p_{\text{inter}} = 1 - p_{\text{intra}}$.

Empirical friendship network. The data-set contains information about friendships within in a US high-school [20]. To construct the network each student was asked twice about his friends. The original network is directed and weighted to account for multiple namings.
of a single student by a friend. For our purposes we symmetrize the interaction topology and dismiss weights such that the post-processed adjacency matrix contains a one between node $i$ and node $j$ if one of the two students named the other as a friend. To randomize the original network topology we perform an increasing number of double edge swaps, which fix the degrees of the nodes but randomize it connectivity structure.

[1] M. Génois, C. L. Vestergaard, J. Fournet, A. Panisson, I. Bonmarin, and A. Barrat, *Network Science* **3**, 326 (2015).
[2] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, *Nature physics* **6**, 888 (2010).
[3] S. Ahajjam and H. Badir, *Scientific reports* **8**, 11932 (2018).
[4] F. Iannelli, M. S. Mariani, and I. M. Sokolov, *Physical Review E* **98**, 062302 (2018).
[5] S. Lewandowsky, J. Cook, N. Fay, and G. E. Gignac, *Memory & cognition*, 1 (2019).
[6] M. P. Fiorina and S. J. Abrams, *Annu. Rev. Polit. Sci.* **11**, 563 (2008).
[7] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, *Advances in Complex Systems* **3**, 87 (2000).
[8] S. Patterson and B. Bamieh, in *49th IEEE Conference on Decision and Control (CDC)* (IEEE, 2010) pp. 2692–2697.
[9] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione, *ACM Transactions on Economics and Computation (TEAC)* **1**, 19 (2013).
[10] S. Galam, *Mind & Society* **15**, 117 (2016).
[11] S. M. Shekatkar, arXiv preprint [arXiv:1902.07084] (2019).
[12] D. Bhat and S. Redner, arXiv preprint [arXiv:1907.13103] (2019).
[13] S. Patterson, Y. Yi, and Z. Zhang, *IEEE Transactions on Control of Network Systems* (2018).
[14] D. J. Klein and M. Randić, *Journal of mathematical chemistry* **12**, 81 (1993).
[15] W. Xiao and I. Gutman, *Theoretical Chemistry Accounts* **110**, 284 (2003).
[16] M. Tyloo, T. Coletta, and P. Jacquod, *Physical review letters* **120**, 084101 (2018).
[17] M. Tyloo, L. Pagnier, and P. Jacquod, arXiv preprint [arXiv:1810.09694] (2018).
[18] M. Tyloo and P. Jacquod, *Phys. Rev. E* **100**, 032303 (2019).
[19] P. Boldi and S. Vigna, *Internet Mathematics* **10**, 222 (2014).
[20] C. A. Bail, L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky, *Proceedings of the National Academy of Sciences* **115**, 9216 (2018).
[21] D. J. Watts and S. H. Strogatz, *nature* **393**, 440 (1998).
[22] P. W. Holland, K. B. Laskey, and S. Leinhardt, *Social networks* **5**, 109 (1983).
[23] A.-L. Barabási and R. Albert, *science* **286**, 509 (1999).
[24] C. Moore and M. E. J. Newman, *Phys. Rev. E* **61**, 5678 (2000).
[25] M. Barahona and L. M. Pecora, *Physical review letters* **89**, 054101 (2002).
[26] R. A. Rossi and N. K. Ahmed, in *AAAI* (2015).