Mathematical Models of Self-Appraisal in Social Networks*

Brian D.O. ANDERSON · Mengbin YE

Abstract In social networks where individuals discuss opinions on a sequence of topics, the self-confidence an individual exercises in relation to one topic, as measured by the weighting given to their own opinion as against the opinion of all others, can vary in the light of the self-appraisal by the individual of their contribution to the previous topic. This observation gives rise to a type of model termed a DeGroot-Friedkin model. This paper reviews a number of results concerning this model. These include the asymptotic behavior of the self-confidence (as the number of topics goes to infinity), the possible emergence of an autocrat or small cohort of leaders, the effect of changes in the weighting given to opinions of others (in the light for example of their perceived expertise in relation to a particular topic under discussion), and the inclusion in the model of individual behavioral characteristics such as humility, arrogance, etc. Such behavioral characteristics create new opportunities for autocrats to emerge.

Keywords Social influence network, DeGroot, Degroot-Friedkin, Opinion dynamics, self-appraisal.

1 Introduction

The subject of opinion formation and evolution in a group of individuals interacting in a social network, conveniently termed opinion dynamics, was originally considered as an activity pursued within the framework of social science, rather than engineering, starting perhaps with the French-DeGroot or simply DeGroot model [1–3]. However, with the increasing dependence on dynamic models for opinion formation, and indeed the parallels between DeGroot-like models and models of physical systems with multiple coordinating autonomous agents, which may be tasked with aligning velocities, or positions, or angular velocities through limited information interchange (see e.g. [4] for a typical contribution), tools of control theory have started to be applied, especially those dealing with issues like asymptotic behavior and stability. The DeGroot model centers around the concept of *social influence*; this influence drives an individual to resolve any difference in opinion with their neighbor in the network. Mathematically, an

Brian D.O. ANDERSON. Email: brian.anderson@anu.edu.au
School of Engineering, Australian National University, Acton 2601, A.C.T., Australia.
Mengbin YE. Email: mengbin.ye@curtin.edu.au
School of Electrical Engineering, Computing and Mathematical Sciences
Curtin University, Perth 6102, W.A., Australia.

*Ye is supported by the Western Australian Government, under the Premier's Science Fellowship Program.

individual's opinion evolves using the weighted average of the opinions of their neighbors, where the influence of each neighbor is captured by an interpersonal weight.

In this paper, we seek to survey a number of examples of the application of control theory tools to a class of opinion dynamics problems, involving selective adjustment of interpersonal weights. Underpinning the relevant models is the notion that when groups of people meet to discuss a sequence of topics, the role that one or another individual plays in a discussion can change over time. One individual or a small number of individuals may come to dominate the discussion, with the outcome heavily depending on their initial views on the topic of discussion. Apart then from the dynamics associated with discussing an individual topic (typically but not always resulting in consensus among the individuals in the group), the nature of the dynamics evolves as one proceeds through a progression of topics, due to the group members' perceptions of the contribution they and other members of the group are making. Put another way, the parameter values governing the interactions vary from topic to topic, depending on the nature of the discussion outcome with previous topics.

Following a significant literature stream, we focus on the mechanism of *reflected self-appraisal*, or simply self-appraisal, [5–7] as the driver of interpersonal weights adjustment. We pay particular attention to a model termed the DeGroot-Friedkin model, which (based on the DeGroot model) arose in [8], with [9] providing a survey and additional material on its asymptotic behavior. Here, an individual updates their self-confidence (a term with technical meaning as given later) at the end of the discussion of a particular topic, in a manner dependent on their contribution to the outcome. (The discussion during the topic is modeled using the usual DeGroot model). Broadly speaking, and as described with more formal language and in detail subsequently in this paper, this evaluation means that if they have contributed substantially to the outcome, their self-confidence will increase, and as a result, during the discussion on the next topic, the individual will accord their opinion greater weighting in relation to the opinions of the remaining individuals. In turn, this change in weighting affects the amount of contribution by each individual in the subsequent topics, illustrating a form of feedback loop in a dynamical social system. The DeGroot-Friedkin model itself provides a systematic way to change the relevant weighting.

Besides surveying the construction of the model, and the asymptotic results provided by [9], in this paper we review the outcomes of a series of our own investigations, [10–17]. In the first instance, we focus on the convergence results.

Depending on the graph structure capturing the interactions between the members of the group, two types of broad outcome are possible. There is an important result which says that, roughly speaking, asymptotically either one individual becomes completely dominant (or autocratic) with other individuals' relevance to discussion outcomes trending to zero; or alternatively, this does not happen, and all individuals contribute something towards the discussion. These are two fundamentally very different scenarios.

In relation to the first, we consider how the autocracy might be disturbed by adjustments in the graph, through addition of further group members, or the establishing of coalitions between group members other than the autocrat. We also consider how a collective autocracy, involving say two individuals, might be created.

In relation to the second scenario, where there is no autocrat, we first establish that the set of weights appearing in the DeGroot equations for each topic, that evolves due to an updating at the end of discussions on each topic, asymptotically approaches the relevant limits exponentially fast in the number of topics. The control theoretic tools for demonstrating this important result are somewhat unusual, being Lefschetz-Hopf theory and nonlinear contraction analysis. The exponentially fast aspect of the convergence suggests the possibility of examining time-varying situations, such as might arise with substitution of group members, or because when topics change, different weights might better reflect specialist knowledge of group members in relation to the new topic as compared to the previous topic. A key conclusion is that even in time-varying networks, an exponential type of convergence applies: in particular, initial conditions are forgotten exponentially fast. We also consider the situation where one or more group members displays certain behavioral characteristics, such as arrogance or humility (where the individual overrates or underrates their ability to contribute) and emotionality or nonreactivity (where the individual reacts particularly positively/negatively to a perceived success or failure, or is nonresponsive to a perceived success or failure). The previous dichotomy of an autocrat emerging or not emerging depending purely on the graphical structure no longer applies, and inclusion of emotional or arrogant individuals can give rise to new possibilities leading to autocracy.

The paper is structured as follows. In the next section, following a review of notation and graph theory, we recall the DeGroot model. The following Section 3 reviews formation of the DeGroot-Friedkin model. We spend some time discussing the asymptotic properties of the model in Section 4, which necessarily distinguishes the two broad cases (convergence to an autocracy, and convergence but not to an autocracy). Section 5 considers the introduction of variations to an autocracy which will disturb it. Section 6 deals with the important issue of time-variation in the weightings given by an individual to neighbors, arising from the nature of the topic, or even adjustment of a group membership. Periodic time-variation is also considered. In Section 7, we discuss the incorporation of specific behavioral characteristics in each individual's self-appraisal process. Section 8 reviews briefly various other works on self-appraisal and Section 9 contains concluding remarks.

2 Background, including DeGroot model review

2.1 Notation, linear algebra and graph theory preliminaries

We record some notation to be used in this paper. Column vectors of all ones and all zeros are denoted by $\mathbf{1_n}$ and $\mathbf{0_n}$ respectively, the vector size being n. The unit vector with 1 in the i-th position will be denoted by e_i. For vectors $x, y \in \mathbb{R}^n$, $x \geq y$ denotes $x_i \geq y_i$ for $i = 1, 2, \ldots, n$, $x > y$ denotes $x \geq y, x \neq y$ and $x \gg y$ denotes $x_i > y_i$ for all i. The n-simplex is $\Delta_n = \{x \in \mathbb{R}^n : x \geq \mathbf{0_n}, \mathbf{1_n}^\top x = 1\}$. Define $\widetilde{\Delta}_n = \Delta_n \setminus \{e_1, e_2, \ldots, e_n\}$ and $\text{int}(\Delta_n) = \{x \in \mathbb{R}^n : x \gg \mathbf{0}_n, \mathbf{1_n}^\top x = 1\}$

A square matrix with all entries nonnegative is termed row stochastic when the sum of the entries in each row is 1. A row stochastic matrix has at least one eigenvalue equal to 1, and no other eigenvalue has magnitude greater than 1. Associated with the eigenvalue 1 is an eigenvector $\mathbf{1}_n$ and also a left eigenvector of nonnegative entries. A square matrix A is termed irreducible if there does not exist a permutation matrix P such that $P^\top AP$ is block triangular. An irreducible row stochastic matrix has a simple eigenvalue at 1. Of course, $\mathbf{1}_n$ is still an eigenvector, and the associated left eigenvector has all positive entries. We term the left eigenvector, call it $\zeta^\top$, normalized when $\zeta^\top \mathbf{1}_n = 1$, and it is of course unique. If in addition to being irreducible, a row stochastic matrix is primitive, i.e. 1 is the only eigenvalue of magnitude 1, there holds $\lim_{n\to\infty} A^n = \mathbf{1}_n\zeta^\top$. A sufficient condition for an irreducible row stochastic A to be primitive is that $a_{ii} > 0$ for some i. Irreducibility of A also corresponds to the property that for all i, j, $(A^k)_{ij} > 0$ for some natural number k depending on i, j; primitivity of A corresponds to k being independent of i, j. For these and other properties of row stochastic matrices, see [18], [19] or [20, Section 2].

To model interactions between individuals in a social network, it is helpful to use a directed graph. A directed graph (digraph) $\mathcal{G}$ is represented by a triple $(\mathcal{V}, \mathcal{E}, A)$. The first of the three members of this triple, $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$, is a vertex set, one vertex corresponding to each individual; the second member of the triple is $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, being a set of directed edges, with the edge from v_i to v_j denoted by $e_{ij} = (v_i, v_j)$; note that e_{ij} and e_{ji} are distinct. There are allowed to be self-loops, i.e. edges starting and ending at the same vertex. Not every vertex pair gives rise to an ordered edge; in a social network context, the presence of the ordered edge e_{ij} connotes that individual j learns and takes into account the opinion of individual i in determining their opinion. Individual i has two neighbor sets, the incoming neighbor set defined by $\mathcal{N}_i^+ = \{v_j \in \mathcal{V} : e_{ji} \in \mathcal{E}\}$ and the outgoing neighbor set $\mathcal{N}_i^- = \{v_j \in \mathcal{V} : e_{ij} \in \mathcal{E}\}$. The third component of the triple is the matrix A, which has a positive entry a_{ij} if and only if $e_{ji} \in \mathcal{E}$, and it measures the weight individual i assigns* to the opinion of individual j. Usually in a social network model, the weights of the edges coming into a vertex sum to 1, implying the matrix A is row stochastic.

A directed path in a graph is a sequence of edges $(v_{i_1}, v_{i_2}), (v_{i_2}, v_{i_3}), \ldots (v_{i_{p-1}}, v_{i_p})$ that starts at v_{i_1} and terminates at v_{i_p}, with no vertex except possibly for the first and last allowed to appear twice. Existence of such a directed path connotes that the opinion of i_1 is taken into account by i_2, that of i_2 is taken into account by i_3 and so on, through to i_p; ultimately then, i_p is influenced, albeit indirectly, by i_1 even if there is no edge (v_{i_1}, v_{i_p}). A cycle is a closed path, i.e. one where the first and last vertex coincide; the length of a cycle is the number of edges in it. A digraph is termed *strongly connected* if and only if every vertex has a path to every other vertex. A digraph is strongly connected if and only the weight matrix A is irreducible [21]. A strongly connected digraph is termed *aperiodic* if the greatest common divisor of all cycle lengths is 1. A strongly connected digraph is aperiodic if and only if the weight matrix A is primitive, [19].

*In some literature, the direction of the edge is reversed so that e_{ij} corresponds to a_{ij}; results and analysis in this paper are unaffected given the appropriate adjustment: an edge e_{ij} connotes that individual i learns of and considers the opinion of individual j.

2.2 The DeGroot model

In social network modelling using the DeGroot model, each individual i holds a scalar real opinion variable y_i which takes values at discrete instants of time. Conventionally, opinion variables are taken as lying in $[0,1]$ or sometimes $[-1,1]$. We will adopt the $[0,1]$ assumption. This formulation of y_i is in fact a highly general one, applicable to many different scenarios. For instance, y_i may represent individual i's attitude or orientation toward something, e.g. electric cars as a mode of personal transport, with $y_i = 0$ and $y_i = 1$ representing maximal rejection and maximal acceptance, respectively, and values between 0 and 1 reflecting partial rejection/acceptance. Alternatively, y_i may represent the belief in a statement, e.g. "electric cars are the future of personal transport", with $y_i = 0$ and $y_i = 1$ representing complete disbelief and complete belief of the statement, respectively.

With A a row stochastic matrix, the DeGroot model [3] posits that at each time instant $k \in \mathbb{Z}^+$, opinions are updated according to the rule

$$y_i(k+1) = \sum_{j \in \mathcal{N}_i^+} a_{ij} y_j(k) \tag{1}$$

or with $y(k)$ denoting the column vector $[y_1(k)\ y_2(k)\ \dots\ y_n(k)]^\top$, one can write

$$y(k+1) = Ay(k). \tag{2}$$

Usually A is irreducible and primitive (equivalently, the underlying graph is strongly connected and aperiodic), and then with $\zeta^\top$ the normalized left eigenvector of A corresponding to eigenvalue of 1, the limiting behavior is

$$\lim_{n\to\infty} y(n) = \lim_{n\to\infty} A^n y(0) = \mathbf{1}_n \zeta^\top y(0) \tag{3}$$

and the convergence is exponentially fast. In the limit, all entries of y tend to the same value, i.e. consensus of opinions is achieved. Unsurprisingly, the actual value is dependent on the initial value of y, with the weight given to $y_i(0)$ being ζ_i; the n weights ζ_i it will be recalled sum to 1. Note that the DeGroot model assumes a synchronous updating scheme, where all individuals update according to Eq. (1) at each time instant. Variations of the model have been studied that assume asynchronous updating, whereby a single individual or a subset of individuals update at each time instant, e.g. [22, 23].

For completeness, we note that the condition that A is irreducible and primitive is actually *not* necessary for consensus to be achieved. Consider for example the following system where $\alpha \in (0,1)$:

$$y_1(k+1) = y_1(k) \tag{4}$$

$$y_2(k+1) = (1-\alpha)y_1(k) + \alpha y_2(k) \tag{5}$$

It is clear that A is reducible. Nevertheless, consensus is easily seen to be achieved at the initial value of $y_1(0)$. This sort of behavior can occur if a single node v_i has $a_{ii} = 1$, and there is a path from v_i to every other node $v_j \in \mathcal{V}$. This possibility will frequently be specifically excluded in the later material. For the interested reader, necessary and sufficient conditions for consensus to be reached in the DeGroot model, both in graph theoretic and algebraic terms, are given in [20, Theorem 12 and Lemma 9].

2.3 Relative interaction matrix

In the sequel, it proves useful to consider a small modification of the basic DeGroot model framework. We consider a nonnegative weighting matrix C, termed a relative interaction matrix, in which $c_{ii} = 0$ for all i. In other words, the associated graph has no self-loops, and the updated value $y_i(k+1)$ for each i does not directly depend on $y_i(k)$ but only on the opinions associated with certain individuals other than i. The quantity $c_{ij} \geq 0$ is the fractional weight given by individual i to the opinion of agent j as compared with the opinions of all of its neighbors (again, excluding i itself). The fractions add to 1, so that the stochastic matrix constraint $\sum_{j=1}^{n} c_{ij} = \sum_{j\in\mathcal{N}_i^+} c_{ij} = 1$ is met.

If we have any row stochastic matrix A, perhaps associated with a DeGroot model, we can associate a relative interaction matrix C with it in the following way. First, in words, the new opinion at time $k+1$ of individual i is a certain combination of its own opinion at time k and the opinions at time k of all other agents which lie in the set $\mathcal{N}_i^+$. Thus if a_{ii} is the weight given to its own opinion, one can express this idea in an equation such as the following:

$$y_i(k+1) = a_{ii}y_i(k) + (1-a_{ii})(\text{representative opinion of neighbors}) \tag{6}$$

Note that the two weights, a_{ii} and $(1-a_{ii})$ add to 1. Thought of along these lines, it is evident that a_{ii} is a measure of the *self-confidence* of an individual. The closer a_{ii} is to 1, the more the individual relies on their preexisting opinion and the less weight they give to the opinion of others; likewise the closer it is to zero, the less weight they are giving to their own preexisting opinion and the more they become simply reactive to the opinion of others.

What should go inside the parentheses as the 'representative opinion of the neighbors'? If a_{ii} were zero, the choice $\sum_{j\in\mathcal{N}_i^+} c_{ij}y_j(k)$ would be logical, where the c_{ij} are entries of the relative interaction matrix. It remains logical even if a_{ii} is not zero. If we use X to denote $\text{diag}[a_{11}, a_{22}, \ldots, a_{nn}]$, we are therefore proposing that the update matrix be

$$A = X + (I-X)C \tag{7}$$

Since X is a matrix whose diagonal entries are the self-confidences of the different individuals, it is apparent that A can be formed from the set of those self-confidences and a relative interaction matrix. Conversely, given A, and provided that no $a_{ii} = 1$, (which would mean that $(I-X)$ is singular) we can obtain the self-confidence values from the diagonal, and determine the associated relative interaction matrix C by

$$C = (I-X)^{-1}(A-X) \tag{8}$$

It is useful also to consider the properties of irreducibility and primitivity for the two matrices A and C. Note that irreducibility of A means that $a_{ii} < 1$ for all i, but the converse does not hold. Clearly then, if A is irreducible, X and C can both be formed. Further, irreducibility of A implies irreducibility of C, a fact which can be seen from the pattern of zeros in the two matrices or seen from the underlying graphs; the graphs differ solely by the presence or absence of self-loops, which have no relevance to the existence of paths connecting vertex pairs. If instead of starting with A to form X and C, we start with X and C to form A, an irreducible C will yield an irreducible A if and only no diagonal entry of X is 1, because $a_{ii} = 1$ for some i means A is reducible.

In contrast to the irreducibility property, primitivity of A does not imply primitivity of C (though the converse is true). Recall though that the presence of any nonzero diagonal entry in A is enough to ensure that A is primitive (assuming it is already irreducible).

We sum up these simple calculations as follows:

Lemma 2.1 *Consider a digraph $\mathcal{G}$, with row stochastic weighting matrix A in which $a_{ii} < 1\ \forall i$. Then a unique vector of self-confidences can be formed, as $x = [a_{11}, a_{22}, \dots, a_{nn}]$ with an associated diagonal matrix $X = \mathrm{diag}[a_{11}, a_{22}, \dots, a_{nn}]$, together with a relative interaction matrix C defined by Eq. (8), which is a row stochastic matrix with zeros on the main diagonal and with all the nondiagonal entries in each row the same positive multiple of the corresponding nondiagonal entries in the corresponding row of A. Further, A, X and C satisfy Eq. (7). Conversely, given a relative interaction matrix C, i.e. a row stochastic matrix with zeros on the diagonal, and a vector of self-confidences x with $\mathbf{0}_n \leq x \leq \mathbf{1}_n$, the matrix A defined by Eq. (7) is a row stochastic matrix. If A is irreducible, so is C. If C is irreducible and $x_i < 1$ for all i, so is A.*

3 The DeGroot-Friedkin model

The scenario for the DeGroot-Friedkin model envisages discussion of a sequence of topics, indexed by a set $\mathcal{S} = \{1, 2, 3, \dots\}$. For each topic, a DeGroot model is used to achieve consensus. This will be achieved by requiring that each of the matrices $A(s)$ is primitive; in fact, this will be assured by arranging they are irreducible, and have at least one nonzero diagonal entry. While during the consideration of any one topic, the weighting matrix $A(s)$ remains constant, the weighting matrix for the successive topics varies, but with one aspect in common, viz. the relative interaction matrix; this invariance of the relative interaction matrix is a core property of the DeGroot-Friedkin model. The way the diagonal entries of $A(s)$ vary with s is set out later in this section; the fact that they do vary lies at the core of the distinction between the simple DeGroot model and its DeGroot-Friedkin development.

By insisting that the relative interaction matrix is irreducible, we ensure then that any corresponding $A(s)$ with no diagonal entry equal to 1 has the same property. However, it is not immediately clear what we should do to ensure that every $A(s)$ is in addition primitive, which is required to secure consensus in the treatment of topic s. One might reasonably make a primitivity assumption about the initial $A(0)$, but probably not all $A(s)$; our strategy in fact will be to prove the primitivity property for $A(s), s \geq 1$, by showing that there is always a positive diagonal entry. Likewise, we will make an assumption that no diagonal entry of $A(0)$

is equal to 1, and show that this property also propagates to $A(s)$ for $s \geq 1$.

These observations motivate the first assumption we make concerning the DeGroot-Friedkin model:

Assumption 1 In the discussion of a sequence of topics by n individuals, the relative interaction matrix C is the same for all topics, and is an irreducible row stochastic matrix with zero entries on its diagonal, while the associated weighting matrix $A(0)$ has one or more diagonal entries positive but no diagonal entry equal to 1.

We now turn to examine the basis on which $A(s)$ changes from topic to topic. The DeGroot-Friedkin formulation postulates that there are two time scales operative. There is the fast time scale by which the opinions on each single topic are updated until a consensus is achieved, and there is the slow time scale by which topics are updated. In practice, for a fixed topic, we could if necessary assume that discussion ceases when in practical terms consensus is achieved (and since consensus is approached exponentially fast, this means after four or five times the dominant time constant governing the opinion convergence).

Let us consider the discussion of topic $s \in \mathcal{S}$. Suppose the associated self-confidence of individual i is denoted as $x_i(s)$. The associated vector of all self-confidences is $x(s)$ and the associated diagonal matrix is $X(s)$. This means that the DeGroot model used for the discussion of topic s uses a weighting matrix

$$A(s) = X(s) + (I - X(s))C \tag{9}$$

We will show below that $A(s)$ never has a diagonal entry equal to 1. Then Assumption 1 and Lemma 2.1 assure that $A(s)$ is irreducible; let us further assume it is in fact primitive (a fact which still needs to be justified, and will also be demonstrated below). Consequently, a consensus is achieved. Let $\zeta(s)^\top$ denote the normalized left eigenvector of $A(s)$. Then the common value in the limit as $k \to \infty$ of the entries of the opinion vector $y(k; s)$ will be $\zeta(s)^\top y(0; s)$. The initial opinion of individual i is $y_i(0; s)$ and this opinion is reflected in the final consensus opinion by weighting it by the quantity $\zeta_i(s)$, which is positive if $A(s)$ is irreducible. Note that these weights add up 1; so $\zeta_i(s)$ is not just a weight but indeed the fraction of influence assigned to the opinion of individual i in coming to the consensus opinion. We will use the term *social power* to describe the quantity $\zeta_i(s)$. It is a quantity associated with an individual, but depends at the same time on the whole social network structure including the weighting values (though not the initial conditions). Different networks thus give different social power to each individual when consensus is reached.

The fundamental premise in the DeGroot-Friedkin model is to postulate that the individual's self-confidence in addressing topic $(s + 1)$ depends on their social power in respect of topic s. That is:

Assumption 2 With $x_i(s)$ and $\zeta_i(s)$ denoting the self-confidence and social power of individual i in addressing topic s, there holds for some function ϕ_i a relation of the form

$$x_i(s + 1) = \phi_i(\zeta_i(s)) \tag{10}$$

If this thinking is to make sense, one would expect that ϕ_i must be constrained in some way, and that is indeed the case. Later in this paper, we shall consider families of such functions. But the basic DeGroot-Friedkin model studied in [9] makes a very simple choice indeed:

Assumption 3 With $x_i(s)$ and $\zeta_i(s)$ denoting the self-confidence and social power of individual i in addressing topic s, there holds

$$x_i(s+1) = \zeta_i(s) \tag{11}$$

The vector of self-confidences evolves every time a topic changes. How does this occur? Assume that Assumption 1 holds. Starting with topic 0, and assumed values of self-confidence $x(0)$ and a relative interaction matrix C, a consensus will be reached since $A(0)$ is primitive. The consensus value is determined by a vector, viz $\zeta(0)^\top$, and its inner product with the vector of initial opinions $y(0;0)$. The vector $\zeta(0)^\top$ is the normalized left eigenvector of $A(0)$, corresponding to the unity eigenvalue (and it is unique since $A(0)$ is irreducible). Following Assumption 3 the i-th entry of $\zeta(0)$ is taken as $x_i(1) = a_{ii}(1)$, i.e. $x(1) = \zeta(0)$, and $A(1)$ is assembled using $X(1)$ and the same relative interaction matrix C as for $A(0)$, in accordance with Assumption 1 and the equation (9). Since $\zeta(0)$ is a positive vector with entries summing to 1, or equivalently $x(1)$ is positive and its entries sum to 1, the diagonal entries of $A(1)$ are necessarily all positive and less than 1, ensuring $A(1)$ is irreducible and primitive. Again a consensus is established but now for topic 1, and the weights determining the consensus value are the entries of the normalized left eigenvector $\zeta(1)$ of $A(1)$. This eigenvector also gives a positive vector $x(2) = \zeta(1)$, and so the process iterates on.

Observe that the update process is such that the irreducibility and primitivity properties of $A(0)$ propagate, i.e. $A(s)$ is necessarily primitive for all finite s and has no diagonal entry equal to 1. (Note that the case of $s \to \infty$ will be the subject of later examination, and it is affected by the underlying topology.)

We sum this up in a formal statement.

Lemma 3.1 *Consider a sequence of DeGroot models of n interacting individuals with Assumptions 1 and 3 holding. Then for all finite s, the matrix $A(s)$ defined by Eq. (9) has positive entries on the main diagonal, with no entry equal to 1, and is primitive.*

Remark 3.2 The lemma is silent on the question of the limit as $s \to \infty$, and indeed we will identify below a class of C for which $A(s)$ in the limit loses irreducibility (and thus primitivity).

In considering every topic, there is an initial opinion vector, call it $y(0;s)$. An immediate and important observation is that the evolution of the $x(s)$ sequence occurs independently of these particular initial opinion vectors $y(0;s)$, i.e. the same sequence $x(s)$ results irrespective of the $y(0;s)$. Therefore, it makes sense to study the evolution of the vector of self-confidences when topics change, irrespective of the initial conditions applying to opinion vectors in considering each topic.

One can notice that Eq. (11) (and more generally Eq. (10)) captures a form of feedback, but in a dynamical social system as opposed to an electrical or mechanical system, as might appear more traditionally in control theory. More precisely, the social power $\zeta_i(s)$ determines the self-confidence for the next $x_i(s+1)$ which in turn determines $\zeta_i(s+1)$. In the context of the problem, Assumption 2 captures reflected self-appraisal and Assumption 3 a particular form of reflected self-appraisal; each individual i separately appraises their own state in the social network, (specifically the level of contribution to consensus), and this occurs via reflection at the end of topic discussion leading to a consensus.

It is straightforward to obtain a recursion for the $x(s)$ vector, as argued in [9]. We record the result here and provide an abbreviated proof below.

Theorem 3.3 *Consider a sequence of DeGroot models of n interacting individuals with Assumptions 1 and 3 holding. Let γ denote the normalized left eigenvector of the relative interaction matrix C corresponding to eigenvalue 1. Then one can write*

$$x(s+1) = F(x(s)) \tag{12}$$

where the nonlinear map $F : \Delta_n \to \Delta_n$ is defined as

$$F(x(s)) = \begin{cases} e_i, & \text{if } x(s) = e_i \\ \alpha(x(s)) \begin{bmatrix} \frac{\gamma_1}{1-x_1(s)} \\ \cdots \\ \frac{\gamma_n}{1-x_n(s)} \end{bmatrix} & \text{otherwise} \end{cases} \tag{13}$$

with

$$\alpha(s) = \frac{1}{\sum_{i=1}^{n} \frac{\gamma_i}{1-x_i(s)}} \tag{14}$$

and in fact F is invariant on $\text{int}(\Delta_n)$.

Remark 3.4 Before sketching a proof of the result, we comment on several properties of F. First, it maps corner points of Δ_n into corner points of Δ_n. Second, although it is invariant on $\text{int}(\Delta_n)$, one cannot assume that any limit of a sequence $x(1), x(2), \ldots$ with all $x(s) \in \text{int}(\Delta_n)$ lies in the same set; the limit may be on the boundary of the set, i.e. it is possible to have for example $x(s) \to e_1$ with all $x(s) \in \text{int}(\Delta_n)$. Third, the map F is also well-defined for any $x(s)$ with $\mathbf{1}_n > x(s) > \mathbf{0}_n$. This means that when $s = 0$, $F(x(0))$ is indeed well-defined (despite the absence of a restriction in Assumption 1 that $x(0) \in \Delta_n$). Further, even if $x(0) \notin \Delta_n$, there holds $x(1) = F(x(0)) \in \text{int}(\Delta_n)$, provided there is no i satisfying $x_i(0) = 1$.

Proof The vector $\zeta(s)^\top$ is the normalized left eigenvector of $X(s) + (I - X(s))C$ corresponding to eigenvalue 1, i.e.

$$\zeta(s)^\top X(s) + \zeta(s)^\top [I - X(s)]C = \zeta(s)^\top$$

or

$$\zeta(s)^\top [I - X(s)]C = \zeta(s)^\top [I - X(s)]$$

This equation says that either $\zeta(s)^\top[I - X(s)]$ is a left eigenvector of C corresponding to eigenvalue 1 or it is the null vector. In case $x_i(s) = 1$ for some i (which means that $x_j(s) = 0$ for $j \neq i$, since the entries of $x_i(s)$ must sum to 1), the row vector $\zeta(s)[I - X(s)]$ will have a zero i-th entry. Since all entries of any eigenvector of C corresponding to the unity eigenvalue must be of the same sign by irreducibility (see Assumption 1), this implies that the whole vector is zero. Since $x_j(s) = 0$ for $j \neq i$, it follows immediately that $\zeta_j(s) = 0$ for $j \neq i$, and as entries of $\zeta(s)$ sum to 1, there results $\zeta_i(s) = 1$, i.e. $\zeta(s) = e_i$. Since $x(s+1) = \zeta(s)$, this verifies the first expression for $F(x(s))$.

If $x_i \neq 1$ for every i, it follows that $\zeta(s)^\top[I - X(s)]$ is a scaled version of $\gamma^\top$ and then the second expression for $F(x(s))$ is almost immediate.

The detailed formula for F makes clear that F is not only invariant on Δ_n but also on the interior of Δ_n. □

It is natural to ask whether there is experimental evidence for the validity of the DeGroot-Friedkin model. Indeed, there are a number of studies that have been conducted by Noah Friedkin and his collaborators, see [24, 25].

There are a number of properties of the map F, and much of this paper will be spent exploring them. Here is one, established in [13]. Note that continuity was established in [9]:

Corollary 3.5 *With the same hypothesis as the theorem above, the mapping F is C^∞.*

Of course, the whole issue is to prove that F is well behaved at the vertices of the unit simplex Δ_n but this is not difficult.

Of great interest are the asymptotics of Eq. (12). It turns out we must pursue two different paths, associated with the nature of the underlying graph linked to the relative interaction matrix.

4 A dichotomy: Asymptotic convergence properties of star graphs and non-star graphs

Before delving into the details, we make three preliminary observations concerning what can be proved regarding the asymptotic behavior of the DeGroot-Friedkin model. First, limit cycles are never encountered; in fact there is always convergence to a fixed point. Second, there are unstable fixed points: the corners of the simplex Δ_n, viz. the vectors e_i, are all fixed points, as is evident from the expression for F. The corner fixed points are either all unstable equilibria, or are all unstable equilibria but one.

Third, and perhaps surprisingly, the asymptotic behavior of the DeGroot-Friedkin model for three or more individuals is different for a (strongly connected) star graph than with any other graph. A strongly connected graph is said to be a star graph if there is a 'central' vertex that is a neighbor (both incoming and outgoing) of every other vertex, while every other vertex has just the central vertex as a neighbor (both incoming and outgoing). See Fig. 1. (The graph is best regarded as corresponding to the relative interaction matrix. Self-loops would be shown in the corresponding graph for $A(s)$). Roughly speaking, convergence occurs from all initial conditions other than the corners e_i of Δ_n to that corner of Δ_n corresponding to the central node of the star graph. For a non-star graph, convergence occurs to a unique point in $\mathrm{int}(\Delta_n)$ from all initial conditions other than the corners e_i of Δ_n.

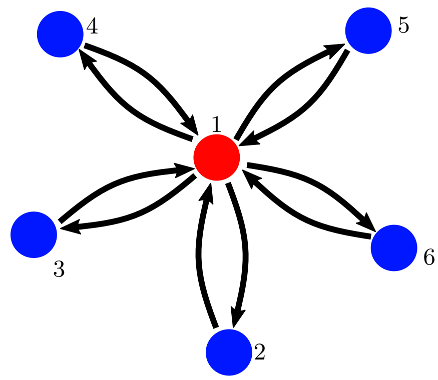


Figure 1: A strongly connected star graph, with central vertex v_1.

As it turns out, the star graph property is reflected in a property of the vector γ associated with the relative interaction matrix C which is helpful to record.

Lemma 4.1 *[9, Lemma 2.3] With $n \geq 3$ vertices, the graph associated with the relative interaction matrix C is a star graph if and only if the normalized left eigenvector $\gamma^\top$ of C corresponding to eigenvalue 1 has one entry (corresponding to the central vertex) equal to 1/2. For a non-star graph there holds $\|\gamma\|_\infty < 1/2$.*

The key to establishing this result, see [9], is to observe that if (for convenience) the central vertex is the first vertex v_1, the relative interaction matrix must have the form

$$C = \begin{bmatrix} 0 & c_{12} & c_{13} & \dots & c_{1n} \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} \tag{15}$$

with $\sum_{j=2}^n c_{1j} = 1$. The first entry of the eigenvector equation $\gamma^\top C = \gamma^\top$ states that $\sum_{j=2}^n \gamma_j = \gamma_1$ whence $\gamma_1 = 1/2$ since the γ_i sum to 1. Conversely, if $\gamma_1 = 1/2$, there must hold $\sum_{j=2}^n \gamma_j = \gamma_1$; taking this with the eigenvector constraint (and noting also that the relative interaction property of C implies $c_{11} = 0$ and $c_{i1} \leq 1$ for $i > 1$) yields the first column constraint in the above equation.

4.1 Convergence of the self-confidence vector for star graphs

The key convergence result, [9, Theorem 4.1], with an embellishment eliminating the possibility of exponentially fast convergence for star graphs in [14, Corollary 2] and [26], is:

Theorem 4.2 *Consider a sequence of DeGroot models of $n \geq 3$ interacting individuals with Assumptions 1 and 3 holding. Suppose that the underlying digraph is a star graph, with central node v_1. Then from any initial $x(0) \in \widetilde{\Delta}_n$, the sequence $x(s), s = 1, 2, \dots$ converges* *asymptotically to* e_1*, but not exponentially fast. Moreover,* $x_1(s)$ *converges monotonically to 1, while* $\sum_{i=2}^{n} x_i(s)$ *converges monotonically to 0.*

The key ideas of the proof are as follows. Through algebraic manipulation, one can establish by contradiction that there is no equilibrium $x^* \in \mathrm{int}(\Delta_n)$. Again through algebraic manipulation, one can establish that $x_1(s+1) > x_1(s)$ for $s \geq 1$. Since $x_1(s)$ is bounded above by 1, the sequence converges. Using $1 - x_1$ as a Lyapunov function for $x \in \Delta_n$, and appealing to the LaSalle Invariance Principle, one establishes that the limit is in fact 1. Since $\sum_i x_i(s) = 1$, this means that the limit of the $x_i(s)$ for $i \neq 1$ is 0. Monotonicity of $x_1(s)$ follows from the proved decrease along trajectories of the Lyapunov function, as also does the monotonicity of $\sum_{i=2}^{n} x_i(s) = 1 - x_1(s)$.

Establishing an absence of exponential convergence is achieved in [14] by looking at the Jacobian of F. It is not hard to check that the Jacobian at e_i for any i has a single eigenvalue at $(1-\gamma_i)/\gamma_i$ and all other eigenvalues are at 0. This means that if $\gamma_1 = 1/2$, i.e. the graph is a star graph with central node 1, the Jacobian at e_1 has a unity eigenvalue, implying exponential convergence cannot be concluded. For e_i with $i \neq 1$, or in the case of a non-star graph, $\gamma_i < 1/2$ by Lemma 4.1, which implies the Jacobian at e_i has an eigenvalue greater than 1, implying instability. This establishes the properties of the fixed points e_i, as noted at the start of this section.

In the next section, we provide a more detailed examination of the consequences of autocracy or, equivalently, having a star graph structure.

4.2 Convergence of the self-confidence vector for non-star graphs

The main result is as follows. In its original form, it can be found in [9, Theorem 4.1], but without a proof of exponential convergence. Exponential convergence is established in [14] and by a different approach in [13].

Theorem 4.3 *Consider a sequence of DeGroot models of* $n \geq 3$ *interacting individuals with Assumptions 1 and 3 holding. Suppose that the underlying digraph corresponding to the relative interaction matrix is not a star graph. Then for all initial conditions* $x(0) \in \widetilde{\Delta}_n$*, i.e.* $x(0) \neq e_i$ *for any* i*, the vector of self-confidences* $x(s)$ *converges exponentially fast to a point* $x^* \in \mathrm{int}(\Delta_n)$*, and is the unique fixed point in* $\widetilde{\Delta}_n$ *satisfying* $x^* = F(x^*)$*.*

The conclusions of the theorem evidently imply that the equilibrium point is independent of the initial condition.

The exponential convergence rate is far more than an academically pleasant conclusion. It is generally well-known that if a collection of systems parametrized by some parameter is exponentially stable uniformly for all values of the parameter (which is possibly confined to a bounded set), then one can expect that slow variation of the parameter will not destroy the exponential convergence property (although the actual convergence rate may be affected). As it turns out here, far stronger results are possible using analysis of the Jacobian matrix, as we illustrate later in the paper, for problems in which the relative interaction matrix can vary from topic to topic.

There are however some more detailed consequences of the above theorem which we record now.

Theorem 4.4 *Adopt the same hypotheses as Theorem 4.3. Then*

1. $x_i^* < x_j^*$ *if and only if* $\gamma_i < \gamma_j$*, and* $x_i^* = x_j^*$ *if and only if* $\gamma_i = \gamma_j$
2. $x_i^* < \gamma_i/(1-\gamma_i)$ *for all* i.
3. *Suppose that* $\|\gamma\|_\infty < 1/3$. *Set* $\kappa = 2\max_i\left(\frac{\gamma_i}{1-\gamma_i}\right)$. *Then there exists a finite* s_1 *such that for all* $s \geq s_1$,
$$\|x^* - x(s+1)\|_1 \leq \kappa\|x^* - x(s)\|_1 \tag{16}$$
4. *If* C *is doubly stochastic (row and column sums are both 1), then* $x_i^* = 1/n$ *for all* i.
5. *Define* $\alpha^* = [\sum_{i=1}^n \frac{\gamma_i}{1-x_i^*}]^{-1}$. *If* $\alpha^* > 0.5$*, then* $\gamma_i > \overline{\gamma}$ *implies* $x_i^* > \gamma_i$*,* $\gamma_i = \overline{\gamma}$ *implies* $x_i^* = \gamma_i$*, and conversely in both cases. If* $\alpha^* \leq 0.5$*, there exists just one individual with self-confidence* x_i^* *exceeding* γ_i.

The first, fourth and fifth results can be found in [9] and the second and third in [14]. A different form of convergence rate bound than that of the third result can be found in [27], involving the value of the entries of x^*.

4.3 Reflections on the two convergence results

Aside from the fact that the proofs are different and the speeds of convergence are different, there are important differing interpretations of the two convergence results when recalling that the DeGroot-Friedkin model captures self-appraisal during discussion on a sequence of topics. In both cases, although we have described convergence as a result dealing with the self-confidence vector, Assumption 3 ensures that simultaneously, we are talking about convergence of the social power by virtue of Eq. (11). In terms of the self-confidence convergence, the result for star graphs is saying that the equation

$$y_1(k+1;s) = a_{11}(s)y_1(k;s) + (1-a_{11}(s))\sum_{j\neq 1} c_{1j}y_j(k;s) \tag{17}$$

is one in which $a_{11}(s) \to 1$ as $s \to \infty$, implying that the central node v_1 accords less and less weighting to the other nodes, and in the limit (which is not actually reached in finite time), there holds $y_1(k;s+1) = y_1(k;s)$. At the same time, for $j \neq 1$, there holds

$$y_j(k+1;s) = a_{jj}(s)y_j(k;s) + (1-a_{jj}(s))y_1(k;s) \tag{18}$$

This says that as $s \to \infty$ and $a_{jj}(s) \to 0$, node v_j accords less and less weight to its own prior opinion, and progressively more weight to the opinion of the central node. In the limit (though not achieved for finite s), $y_j(k;s) = y_1(k;s)$, implying that the node v_j opinion simply mirrors that of the central node.

By way of side remark, we note that even though one starts with $A(0)$ irreducible and the irreducibility condition propagates through all $A(s)$ with $s \geq 1$, it fails in the limit because $\lim_{s\to\infty} a_{11}(s) = 1$.

In terms of the social power, recall that the consensus value for the s-th topic is $\zeta(s)^\top y(0; s)$. Hence the convergence of $\zeta(s)$ to the unit vector e_1 means that in the limit as $s \to \infty$, the consensus value achieved for all nodes depends solely on the initial opinion $y_1(0; s)$ held by the central node v_1, and is unchanged from that value. All nodes are required to follow the value of the central node's opinion. One can also note that for all finite s, $\zeta(s)$ is a positive vector, but the positivity property fails in the limit as $s \to \infty$.

In summary then, given a star graph (corresponding to a relative interaction matrix) an autocracy develops, with all the social power being held by the individual corresponding to the central vertex. Apparently, even though 'consultations' occur on topics (two-way opinion exchanges between the central vertex individual and the other individuals, but not between pairs of other individuals), when they are all one-on-one (the implication of having a star graph), the centrally located individual turns into an autocrat in the limit.

In contrast to the situation for star graphs, no possibility of autocracy arises for a non-star graph. The limiting value of the self-confidence vector is always an interior point of the region $\widetilde{\Delta}_n$, and obviously then, no node acquires all the social power, i.e. a value of 1 for its entry of the corresponding social power vector, which is $\zeta(s)$ at topic s. The limit of the sequence $A(s)$ will be irreducible and primitive, with diagonal entries in $(0, 1)$. And while there is no simple explicit formula for the limiting vector, the ordering of entries is precisely determined, in that it reflects the ordering of the entries of γ, which is the normalized left eigenvector of the relative interaction matrix. In case C is doubly stochastic, one has $\gamma = (1/n)\mathbf{1}_n$, and all x_i^* take the same value. This is the antithesis of autocracy, and might be termed total democracy.

4.4 Commentary on the proof of convergence for non-star graphs

In this subsection, we summarize a number of aspects regarding the approaches to proving convergence of the iterative algorithm in Eq. (12) for the self-confidence vector in the case of non-star graphs.

By way of a preliminary observation, the continuity of F, and the fact that Δ_n is a convex and compact set invariant under F, guarantee by Brouwer's Fixed Point Theorem, see e.g. [28], that there is at least one fixed point of the mapping. By itself however, this is not helpful in identifying whether an interior fixed point exists, given that the corner points are all fixed points. To establish that there is at least one interior fixed point requires establishing that there is a convex and compact set excluding the boundary points (corners, edges, faces, etc.) of Δ_n, and further establishing that this set is in fact invariant for F. This can indeed be done, with the set a sort of 'shrunk' version of Δ_n, the shrinkage being arbitrarily small but nonzero. We shall refer to this set as $\widehat{\Delta}_n$ and for sufficiently small $\delta > 0$, it is given by

$$\widehat{\Delta}_n = \{x \in \mathbb{R}^n : \delta \leq x_i \leq 1 - \delta \ \forall i = 1, \ldots, n, \mathbf{1}_n^\top x = 1\} \tag{19}$$

To prove Theorem 4.3, and using the fact that at least one fixed point is guaranteed to exist by Brouwer's Theorem in $\widehat{\Delta}_n$ for sufficiently small δ, reference [9] then uses a series of algebraic inequality calculations to show that the interior fixed point is unique. Certain properties of the trajectories are established, and together with a proposed Lyapunov function (whose functional form requires the coordinates of the equilibrium point) and application of the LaSalle Invariance principle, asymptotic stability is concluded. No conclusion regarding exponential stability is given.

A different approach yielding exponential stability is to be found in [13] and [16]. This approach depends on Lefschetz-Hopf theory [28], and the following powerful, but little known consequence of the main result of that theory:

Theorem 4.5 *Consider a smooth map $F : X \to X$ where X is a compact oriented and convex manifold or a convex triangulable space of arbitrary dimension. Suppose that the eigenvalues of the Jacobian dF_x have magnitude less than 1 for all fixed points of F. Then F has a unique fixed point, and in a local neighborhood about the fixed point, the iteration $x(k+1) = F(x(k))$ converges exponentially fast.*

For the self-confidence update map F of interest to us, it is possible to prove that at any interior point of Δ_n, any fixed point in $\widehat{\Delta}_n$ necessarily gives a Jacobian dF_x whose eigenvalues are less than 1 in magnitude, see [13] for details. This immediately implies the fixed point is unique, with a local exponential stability property. What still remains unproved when relying on Lefschetz-Hopf ideas and not the Lyapunov approach described above, is that convergence to the unique fixed point in $\widehat{\Delta}_n$ occurs from any initial condition apart from the corners e_i in Δ_n. Note also that there is no general proof that the Jacobian of F at an arbitrary point in $\widehat{\Delta}_n$ has eigenvalues less than 1 in magnitude; indeed examples reveal this is not necessarily so. If it were so, then application of Banach's fixed-point theorem would allow to conclude convergence from any initial condition [29]. By way of an aside, the Jacobian of $F(x(s))$ can of course be expressed in terms of $x(s)$. With minor manipulation it turns out that one can write

$$
\begin{aligned}
\frac{\partial F_i}{\partial x_i}(x) &= x_i \frac{1 - F_i(x)}{1 - x_i} \\
\frac{\partial F_i}{\partial x_j}(x) &= -\frac{F_i(x)F_j(x)}{1 - x_j}, \quad j \neq i
\end{aligned}
\tag{20}
$$

The relevance of these expressions becomes clear at a fixed point of F, when x_i and $F_i(x)$ coincide.

A different approach again, yielding almost global exponential stability, is to be found in [14]. Jacobians again play a role. The key is to work with what are termed 'virtual dynamics', where the word 'virtual' connotes infinitesimal displacement [30]. Thus the following equation is studied:

$$
\delta x(s+1) = \frac{\partial F(x(s))}{\partial x(s)} \delta x(s) \tag{21}
$$

and then a particular nonsingular transformation $\delta z(s) = \Theta((x(s), s)\delta x(s)$ is used. The particular $\Theta(x(s), s)$ is constructed to be bounded and uniformly nonsingular. The transformed system, of the form

$$\delta z(s+1) = \widehat{F}(s)\delta z(s),$$

is then studied. It turns out that the 1-norm of $\widehat{F}(s) \in \mathbb{R}^{n\times n}$ is uniformly less than 1, which implies any pair of neighboring trajectories of the original system $x(s+1) = F(x(s))$ converge to one another exponentially fast. Since the trajectories are all confined to a compact region, one can argue that this means all trajectories must tend to a single trajectory. Since a trajectory beginning at a fixed point is a trajectory which stays at the fixed point, it follows that all trajectories approach the fixed point known to exist via Brouwer's fixed point theorem, and it is necessarily unique. Exponentially fast convergence is a further consequence. The introduction of the set $\widehat{\Delta}_n$ plays the same technical role here as earlier, and serves to exclude the corner points of Δ_n which are themselves fixed points, though unstable.

It would seem that this last approach is the most powerful, in terms of efficiently delivering the desired result of almost global exponential convergence to a unique equilibrium point. Furthermore, it underpins our later consideration of networks with varying relative interaction matrices. There would appear to be no way the methods of [9] or [13, 16] could be straightforwardly applied to time-varying problems. One potential disadvantage however is the need to determine the transformation matrix $\Theta(x(s), s)$; the problem is analogous to finding a Lyapunov function to demonstrate stability of some given system, there being no universal and standard way in either case.

5 Exploration of star graphs: attacking autocracy and embracing collective leadership

Motivated by general social disapproval of autocracies, we will first examine how additional nodes and/or edges might be connected to a star topology to change the equilibrium self-confidence vector, and in particular, upset the autocracy property. Then we shall examine the consequence of having a *leadership group*, where a small number of nodes which in rough terms are centrally located can acquire most of the social power[†]. Most of the results can be found in [10]; the remainder are in [11].

With the notation x^* being used to denote an equilibrium of $x(s+1) = F(x(s))$, we make the following definitions:

Definition 5.1 (Autocratic network/autocrat) A social network is said to be autocratic or have an autocratic configuration, with node v_i being the autocrat, if $x^* = e_i$.

Definition 5.2 (Domination of one individual by another/Social dominance/leadership) Individual i, associated with node v_i, is said to dominate individual j, associated with node v_j, if $x^*_i > x^*_j$, and to be the socially dominant individual or leader in a social network if $x^*_i > x^*_j$ for all $j \neq i$.

[†]To this point, we have preserved the distinction between self-confidence and social power as concepts, even if the values associated with the respective variables are the same. In the remainder of this section, we will blur the distinction.

For would-be autocrats, social dominance of the network is the next best thing.

Given a graph $\mathcal{G}$ with star topology, with centre node say v_1, let us call the other nodes *subject nodes*, in the sense they are subjects of the centre autocrat node. Now there are multiple ways where simple modifications of the graph structure might undermine the autocratic status of v_1. (As we know, there cannot be an autocrat node in steady state if the graph structure is not a star graph, and so the modifications we suggest are consistent with that fact.)

Here are some possibilities, where $\mathcal{G}$ is modified in some way to become a non-star graph $\overline{\mathcal{G}}$. They are illustrated in Fig. 2.

1. Single attack. Suppose that $n \geq 4$. Suppose that $\mathcal{G}$ has a star topology with v_1 the central node (autocrat) and $(n-2)$ subject nodes. Introduce an attacker node v_n attached to subject node v_{n-1} by adding two edges, $e_{n-1,n}$ and $e_{n,n-1}$, in opposite directions between the two nodes to form $\overline{\mathcal{G}}$.

2. Coordinated double attack. Suppose that $n \geq 5$. Suppose that $\mathcal{G}$ has a star topology with v_1 the central node (autocrat) and $(n-3)$ subject nodes. Two attacker nodes v_{n-1} and v_n are attached to subject node v_{n-2} by adding two oppositely directed edges between each attacker and v_{n-2}, with edges $e_{n-2,n-1}, e_{n-1,n-2}, e_{n-2,n}, e_{n,n-2}$, to form $\overline{\mathcal{G}}$.

3. Uncoordinated double attack. As for Case 2, except that one attacker node v_{n-1} attaches to subject node v_{n-3} and the other attacker node v_n attaches to subject node v_{n-2}, each attachment involving two oppositely directed edges, $e_{n-3,n-1}, e_{n-1,n-3}$ and $e_{n-2,n}, e_{n,n-2}$.

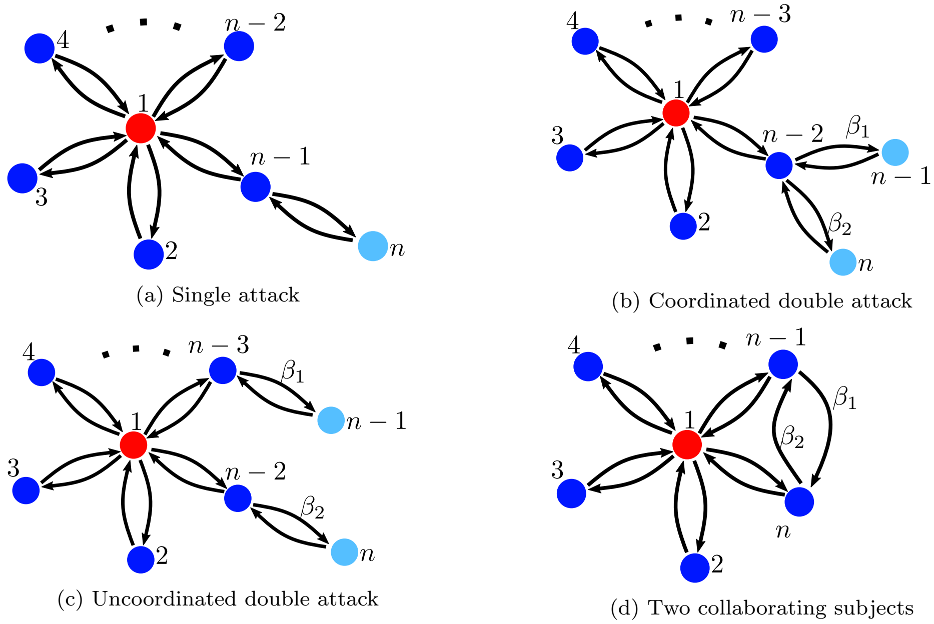
4. Two collaborating subjects. Suppose that $n \geq 4$. Suppose that $\mathcal{G}$ has a star topology with v_1 the central node (autocrat) and $(n-1)$ subject nodes. There are no attacker nodes but subject nodes v_{n-1} and v_n become neighbors through two oppositely directed edges, $e_{n,n-1}$ and $e_{n-1,n}$, between them to form $\overline{\mathcal{G}}$.

The conclusions are as follows. Note that in every case, after adjustment there will necessarily hold $x_1^* < 1$ since there is no longer a star graph; also, in those cases where a new nonzero entry appears in a row of C, the remainder of that row is scaled to preserve the row stochastic property. The first key question is then under what circumstances the central node retains social dominance. Since a single attack is a specialized coordinated double attack, we consider the coordinated double attack possibility first.

Coordinated double attack conclusions. (Refer to Figure 2b). Suppose that $\beta_1 = c_{n-2,n-1} \in (0,1)$ and $\beta_2 = c_{n-2,n} \in (0,1)$ (these being entries of the relative interaction matrix C) and consider these parameters to be adjustable. Note that necessarily $\beta_1 + \beta_2 < 1$ since $c_{n-2,1} > 0$ and each row of the C matrix sums to 1.

1. Node v_1 remains dominant over any subject node where there is no attacker connected: $x_i^* < x_1^*$ for $i \neq 1, n-2, n-1, n$.

2. Node v_1 also remains dominant over nodes v_{n-2}, v_{n-1}, v_n, i.e. it is the socially dominant node, if and only if $\beta_1 + \beta_2 < 1 - c_{1,n-2}$. If $\beta_1 + \beta_2$ exceeds this threshold, then node



(a) Single attack

(b) Coordinated double attack

(c) Uncoordinated double attack

(d) Two collaborating subjects

Figure 2: An illustration of the different topology variations used to modify a star graph with central node v_1 (red). The parameters β_1 and β_2 are the weights in the relative interaction matrix that selected subject nodes (blue) accord to the newly introduced attacker nodes (cyan) in 2b and 2c. In 2d, β_1 and β_2 are the relative interaction weights the two subjects accord to each other.

v_{n-2} is socially dominant in the network, i.e. leadership is transferred from node v_1, but of course, node v_{n-2} is not an autocrat.

3. Separately, x_n^* and x_{n-1}^* are greater than x_1^* if $\beta_1 > (1-\beta_1)/(1+c_{1,n-2})$ and $\beta_1 > (1-\beta_2)/(1+c_{1,n-2})$, respectively.

4. Node v_{n-1} is dominant over node v_n if and only if $\beta_2 > \beta_1$.

The variations to cope with a single attack are trivial, and likewise for using three attack nodes, rather than two.

The second conclusion, and in particular the formula for the threshold, reveals another important fact: other things being equal, if the aim is to dislodge the social leadership status of the central node (as well as dislodging its autocrat status) the attacks should be focused through the subject node which is the most trusted by the central node, i.e. the node j for which $c_{1,j}$ is greatest. This is the choice which gives the lowest threshold for $\beta_1 + \beta_2$. The choice is also logical in everyday terms: to change the leader with external intervention one should concentrate all support on the leader's most trusted subject.

Uncoordinated double attack conclusions. (Refer to Figure 2c). As before, $\beta_1 \in (0,1), \beta_2 \in (0,1)$.

1. Node 1 remains dominant over any subject node where there is no attacker connected: $x_i^* < x_1^*$ for $i \neq 1, n-3, n-2, n-1, n$.

2. Node v_1 will be dominated by v_{n-3} or v_{n-2}, respectively, if and only if β_1 or β_2 exceed the separate thresholds of $1 - c_{1,n-3}$ and $1 - c_{1,n-2}$, respectively

3. The inequality $\beta_1 > 1/(1 + c_{1,n-3})$ (which implies $\beta_1 > 1 - c_{1,n-3}$) assures that node v_1 is dominated by an attacker node v_{n-1} and $\beta_2 > 1/(1 + c_{1,n-2})$ (which implies $\beta_2 > 1 - c_{1,n-2}$) assures that v_1 is dominated by v_n. A further inequality determines which of v_{n-3} and v_{n-2} dominates the other: in fact, $(1 - \beta_2)/(1 - \beta_1) > c_{1,n-2}/c_{1,n-3}$ implies $x_{n-3}^* > x_{n-2}^*$.

Notice that in contrast to the conclusions for a coordinated attack which revolve round the level of the sum $\beta_1 + \beta_2$, the conclusions here involve β_1 and β_2 separately. Closer examination in fact reveals that, in accordance with intuition, a coordinated attack on the social dominance of the central node is more effective than an uncoordinated double attack, i.e. in the sense of requiring a lower level of $\beta_1 + \beta_2$ to destroy that social dominance.

Two collaborating subjects conclusions. (Refer to Figure 2d). There are no attacker nodes; two subjects instead achieve dissent from the leader by collaborating, and this involves two parameters $\beta_1 = c_{n-1,n}$ and $\beta_2 = c_{n,n-1}$.

1. Node v_1 remains dominant over nodes other than the dissenting nodes: $x_i^* < x_1^*, = 2, 3, \ldots, n-2$.

2. There are threshold values for β_1 and β_2 which when exceeded guarantee that $x_n^* > x_1^*$ or $x_{n-1}^* > x_1^*$, i.e. v_1 loses social dominance. For example, $x_n^* > x_1^*$ if and only if $\beta_1 > (1 - c_{1,n})/(c_{1,n-1} + \beta_2)$ with $\beta_1 \in (0, 1)$ and there exists such a β_1 only if $\beta_2 > \sum_{i=2}^{n-2} c_{1,i}$.

3. Node v_{n-1} dominates v_n, i.e. $x_{n-1}^* > x_n^*$, if and only if $c_{1,n-1} + c_{1,n}\beta_2 > c_{1,n} + c_{1,n-1}\beta_1$.

When the fine structure of the inequalities for securing social dominance of node v_{n-1} or v_n is examined, one finds that the two dissenting nodes must adopt a cooperative strategy: each node must trust the other sufficiently (i.e. both β_1 and β_2 have to be suitably large).

Now we will consider a different form of variation to a star graph, one in fact where a single leader is replaced by a leadership group, For convenience we will consider a leadership group of two individuals only. More precisely, suppose that $\mathcal{G}_1$ and $\mathcal{G}_2$ are star graphs with node sets $\mathcal{V}_1 = \{1, \ldots, n\}$ and $\mathcal{V}_2 = \{n+1, \ldots, n+m\}$, where $n \geq 3,\ m \geq 3$. A graph $\overline{\mathcal{G}}$ is formed by merging $\mathcal{G}_1$ and $\mathcal{G}_2$ through insertion of edges $e_{1,n+1}$ and $e_{n+1,1}$ linking the two leaders. The nodes v_1, v_{n+1} are regarded as the leadership group and other nodes are subjects.

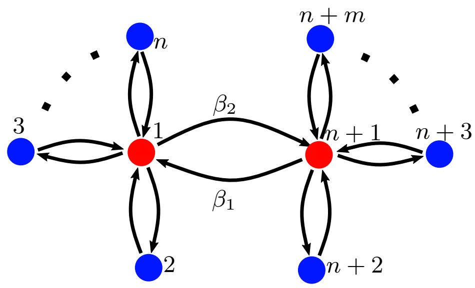


Figure 3: Two star graphs, with central nodes v_1 and v_{n+1} (red), joined through the central nodes to form a leadership group, with relative interaction matrix weights β_1 and β_2.

The graphs $\mathcal{G}_1, \mathcal{G}_2$ have relative interaction matrices C_1, C_2. The relative interaction matrix $\overline{C}$ associated with $\overline{\mathcal{G}}$ is almost simply a block diagonal matrix with C_1 and C_2 the block diagonal entries; however, after the addition of two edges, adjustments must be made to the entries to ensure that $\overline{C}$ is row stochastic.

Leadership group conclusions (Refer to Figure 3).

1. For all $\beta_i \in (0,1)$, node v_1 continues to dominate nodes $v_2, \ldots, v_n$ and node v_{n+1} continues to dominate nodes $v_{n+2}, \ldots, v_{n+m}$
2. $x_1^* > x_{n+1}^*$ iff $\beta_2 > \beta_1$
3. Node v_1 dominates node v_{n+k} for $k \in \{1, \ldots, m\}$, i.e. $x_1^* > x_{n+k}^*$, if and only if $c_{n+1,k}(\beta_2/\beta_1) < 1$. Node v_{n+1} dominates node v_i for $i \in \{2, \ldots, n\}$ iff $c_{1,i}\beta_1/\beta_2 < 1$.

Statement 1 indicates that neither leader loses its dominance over those nodes over which, before merger, it had social dominance. Statement 2 deals with which of the two leaders is dominant over the other. The interesting result is statement 3, indicating when one of the leaders might acquire dominance over all subject nodes in $\overline{\mathcal{G}}$. A sufficient condition for each leader to dominate all the subject nodes is that the leaders are fully cooperative, in the sense that $\beta_1 = \beta_2$. The actual value of the β_i is, perhaps surprisingly, irrelevant. Even with some difference in the two values, this dominance will continue to hold.

6 Allowing changing topologies in the social network

The social networks we have been considering to this point can be viewed as having changing topologies, in that in a topic series, there are changes in the self-confidence of each agent, or equivalently changes linked to the diagonals of the weighting matrices $A(s)$ used in the opinion formation process (with corresponding changes implemented to off-diagonal elements). The relative interaction matrix C has however been constant.

In this section, we consider results dealing with a changing relative interaction matrix. A contribution dealing with periodic changes was considered in [12], but the main contribution is to be found in [14].

There are at least two distinct scenarios which motivate the consideration of time-varying interaction matrices, and we term them *issue-driven* and *individual-driven*.

Issue-driven variation: Consider a regular government cabinet meeting to discuss issues including e.g. economic growth, foreign policy, environment. Each minister will have their own portfolio, and presumably be more knowledgeable in relation to that portfolio than others, but nevertheless they will participate in the discussion of almost all issues. Such issues give rise to the specific topics that appear in the set $\mathcal{S}$, discussed in each cabinet agenda. It is natural then for the weights c_{ij} to change with topic: the relative trust given to the minister of environment by the minister's neighbors may be high when discussing an environmental topic, but not a foreign policy topic.

Individual-driven variation: Considering the cabinet analogy further, one can imagine changes in the c_{ij} matrix arising because of changes in the personal relationships of ministers, or replacement of individuals within the cabinet. It could be too that individual i might decide, after some experience, that individual j is not worth listening too, and set c_{ij} to zero. This might be a product of an extreme form of homophily, which is a classical social psychological theory that people tend to form relationships with those others who hold similar beliefs, values, interests, etc. [31].

These remarks then motivate the study of the following replacement of Eq. (9):

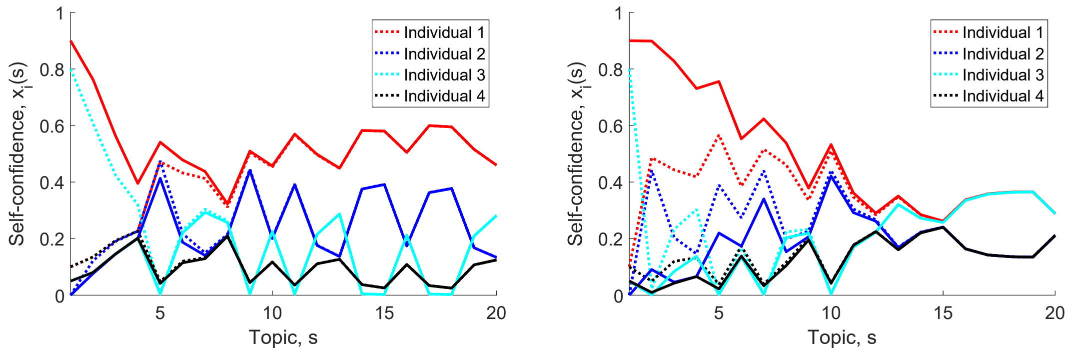
$$A(s) = X(s) + (I - X(s))C(s) \tag{22}$$

where we emphasise in the notation the dynamic nature of $C(s)$. Of course, $X(s) = \mathrm{diag}(x(s))$ evolves as previously.

Issue-driven variations may well be periodic in nature, a cabinet for example having a standardized list of issues for each meeting (though the particular topics dealing with one issue may vary from meeting to meeting). The paper [12] adopted such an assumption, and with partial proofs and simulations argued that a limiting regime for $x(s)$ would be reached, in which the self-confidence vector varied periodically; convergence to a periodic trajectory occurred exponentially fast, and initial conditions were forgotten, i.e. the limiting trajectory was unique.

In the discussion in Section 4.4 of approaches to establishing the asymptotic behavior of the vector of self-confidences for a non-star graph with a constant relative interaction matrix C applying to the discussion of every topic, we noted one approach using the notion of virtual dynamics. In this approach, a sort of linearization procedure established that two neighboring trajectories necessarily converged exponentially fast. Using the compactness of $\widehat{\Delta}_n$ it followed that any trajectories beginning in $\widehat{\Delta}_n$ then had to converge to one another. In the case where a fixed point exists, the argument established almost global convergence to the fixed point. When the $C(s)$ matrix is issue-dependent[‡], essentially the same calculations continue to apply; hence for the time-varying case, a limiting trajectory independent of initial conditions has to exist and is approached exponentially fast.

[‡]We make a mild assumption that $C(s+1)$ is independent of $x(s')$ for all $s' \leq s$.



(a) Switching topology sequence 1

(b) Switching topology sequence 2

Figure 4: An illustration of the convergence result for changing topologies. A network with $n = 4$ individuals with 3 different possibly relative interaction matrices C_1, C_2, C_3, with the dotted and solid lines denoting two different initial self-confidence vectors $x(0)$. (a) For a particular topology sequence $C(s)$, $s = 0, 1, 2, \ldots$ with $C(s) \in \{C_1, C_2, C_3\}$, the trajectories from two different initial condition vectors converge exponentially fast to a unique limiting trajectory determined only by the sequence $C(s)$. (b) When a different sequence $C(s)$ is used, the exponential convergence property is preserved, but the unique limiting trajectory is different from (a).

We can prove from this observation that if the $C(s)$ sequence is periodic, the limiting trajectory is periodic (and is approached exponentially fast). Crucially, it is independent of initial conditions. The argument is as follows. Let $x(s; x_0, s_0), s = s_0, s_0 + 1, \ldots$ denote successive values of the social power vector given an initial condition $x(s_0) = x_0$. For a periodic system with period K, evidently

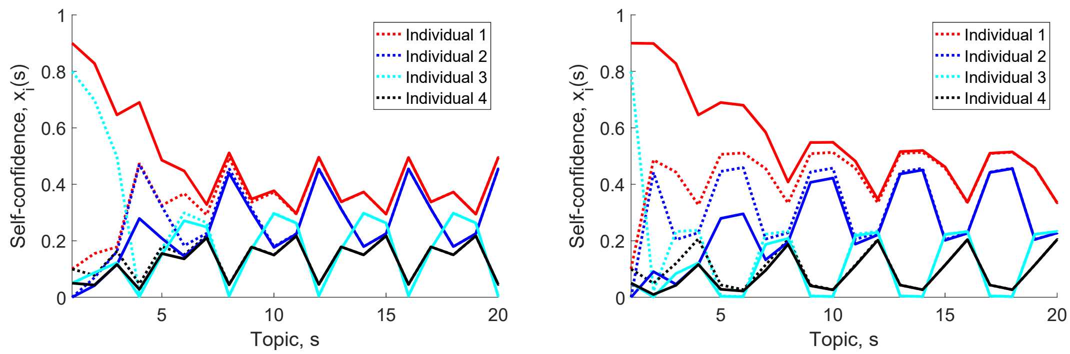
$$x(s + K; x_0, s_0 + K) = x(s; x_0, s_0) \tag{23}$$

Now the convergence property guarantees that $\lim_{s_0 \to -\infty} x(s; x_0, s_0) = \widehat{x}(s)$ is independent of x_0. So taking the limit as $s_0 \to -\infty$ in the above equation yields $\widehat{x}(s + K) = \widehat{x}(s)$, i.e. the limiting trajectory is periodic.

Figures 4 and 5 illustrate these ideas, for the nonperiodic case and the periodic case, respectively. The precise simulation parameters, and the `MATLAB` code, are available in the following repository: `https://github.com/mengbin-ye/degroot-friedkin-simulations`.

7 Incorporating behaviors

The developments up to now have all rested on the critical Assumption 3, which is an intuitive specialisation of Assumption 2. The self-confidence $x_i(s + 1)$ which an individual brings to topic $s + 1$ is identical with the social power $\zeta_i(s)$ they exhibited in the treatment of topic s. But individuals react differently to success or failure. An arrogant individual may adopt a higher self-confidence for topic $s + 1$ than their social power for topic s and a humble



(a) Periodic topology sequence 1 (b) Periodic topology sequence 2

Figure 5: An illustration of the convergence result for changing topologies. A network with $n = 4$ individuals with 3 different possibly relative interaction matrices C_1, C_2, C_3 (the same as in Fig. 4), with the dotted and solid lines denoting two different initial self-confidence vectors $x(0)$. (a) For the periodic sequence $C_2, C_1, C_3, C_1, C_2, C_1, C_3, C_1, \ldots$, the trajectories from two different initial condition vectors converge exponentially fast to a unique limiting periodic trajectory determined only by the sequence $C(s)$. (b) When a different periodic sequence of $C_3, C_2, C_1, C_3, C_3, C_2, C_1, C_3, \ldots$ is used, the exponential convergence property is preserved, but the unique limiting periodic trajectory is different from (a).

individual might do the reverse. We can also identify at least two other types of behavior leading to self-confidence and social power being unequal. An individual may be emotional (in which case they over-react to an outcome, boosting their self-confidence when their social power is high, or decreasing their self-confidence when their social power is low.) An individual may be nonreactive, maintaining an almost constant level of self-confidence over a wide range of social powers.

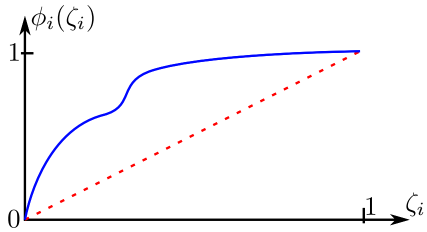
In this section, we describe results presented in [15] where such behavioral characteristics are taken into account.

Accordingly, we replace Assumption 3 by Assumption 2. In addition, we require:

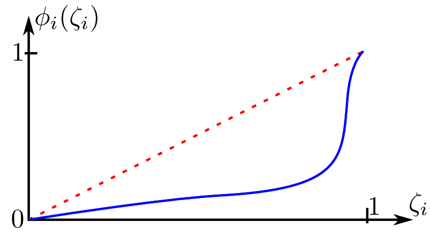
Assumption 4 For every $i \in \{1, \ldots, n\}$, $\phi_i(\zeta_i) : [0, 1] \to [0, 1]$ is a smooth monotonically increasing function satisfying $\phi_i = 0 \iff \zeta_i = 0$ and $\phi_i = 1 \iff \zeta_i = 1$.

An individual i for which $\phi_i(\zeta_i) = \zeta_i$ is termed a *well-adjusted* individual. Figure 6 in contrast illustrates the shape of example ϕ_i corresponding to the distortions introduced by the four behavioral characteristics identified above, viz. arrogant, humble, emotional and nonreactive. The specifics of the ϕ_i are given in [15], but are highly general and admit a broad range of ϕ_i functions for each behavioral characteristic.

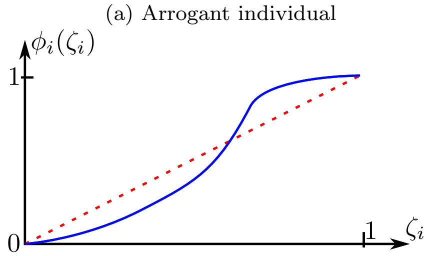
With Assumptions 2 and 4 replacing Assumption 3, Lemma 3.1 continues to hold, the argument being very little changed from the preceding argument. One difference is that the entries of $x(s)$, while obeying $\mathbf{0}_n \ll x(s) \ll \mathbf{1}_n$, no longer sum to 1, and it is easier to cast the update equation using ζ as opposed to x; this means that the relevant mapping continues to



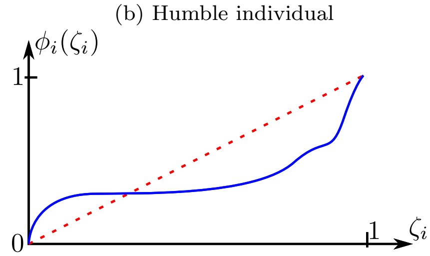
(a) Arrogant individual



(b) Humble individual



(c) Emotional individual



(d) Nonreactive individual

Figure 6: Different behaviors of individual i during the self-appraisal process, with example ϕ_i functions. The red dotted line illustrates the well-adjusted individual with $\phi_i(\zeta_i) = \zeta_i$ for all $\zeta_i \in [0, 1]$, which is the original DeGroot-Friedkin model as detailed in Assumption 3. (6a) An arrogant individual has $\phi_i(\zeta_i) > \zeta_i$ for all $\zeta_i \in (0, 1)$. (6b) A humble individual has $\phi_i(\zeta_i) < \zeta_i$ for all $\zeta_i \in (0, 1)$. (6c) An emotional individual behaves as a humble and arrogant individual for low and high values of ζ_i respectively. (6a) A nonreactive individual behaves as an arrogant and humble individual for low and high values of ζ_i respectively.

be invariant on the unit simplex.

Let Φ denote the mapping from the vector $\zeta(s)$ to the vector $x(s+1)$, so that $\Phi(\zeta(s)) = x(s+1)$. The following Theorem follows easily:

Theorem 7.1 *Consider a sequence of DeGroot models of $n \geq 3$ interacting individuals with Assumptions 1, 2 and 4 holding. Then*

$$\zeta(s+1) = F(x(s+1)) = F(\Phi(\zeta(s))) \tag{24}$$

where

$$F(x) = \begin{cases} e_i & \text{if } x_i = 1 \text{ for any } i \\ \alpha(x) \begin{bmatrix} \frac{\gamma_1}{1-x_1} \\ \vdots \\ \frac{\gamma_n}{1-x_n} \end{bmatrix} & \text{otherwise} \end{cases} \tag{25}$$

where $\alpha(x) = 1/\sum_{i=1}^{n} \frac{\gamma_i}{1-x_i}$. The map $G = F \circ \Phi : \Delta_n \to \Delta_n$ is smooth on Δ_n, and $\zeta(s) \in \Delta_n$, $x_i(s) \in [0, 1]$ for all i. The map G is also invariant on $\mathrm{int}(\Delta_n)$.

As foreshadowed, we are now considering a dynamical system for the social power vector, $\zeta(s+1) = G(\zeta(s))$. This is because Φ is surjective according to Assumption 4, and therefore establishing the dynamics of $\zeta(s)$ allows one to establish the dynamics of $x(s)$, but not vice versa.

It is evident that if $x(s') = e_i$ for some i and s', or indeed $x_i(s') = 1$ with possibly $x_j(s') > 0, j \neq i$, then for all $s > s'$, there will hold $x_i(s) = \zeta_i(s) = 1$, with individual i holding all the social power, and all other individuals having no self-confidence and no social power. On the other hand, Assumption 1 ensures that we consider initial conditions in which $x_i(0) < 1$ for all i with $x_i(0) > 0$ for some i, which ensures $\zeta_j(0) > 0$ for all j, and thus $\mathbf{0}_n \ll x(s) \ll \mathbf{1}_n$ for all $s \geq 1$. So we cannot obtain $x(s') = e_i$ under the stated assumptions for finite s'. Of course, in the limit, matters are different. It is then possible to have $\lim_{s\to\infty} x(s) = e_i$.

The first major issues to consider are whether there is a fixed point for the mapping G, as there was for F, and whether any such point is unique.

Here are the key results from [15]:

Theorem 7.2 *Adopt the same hypotheses as for Theorem 7.1, and suppose that the underlying graph is a non-star graph. The following conclusions hold:*

1. *If every individual is either well adjusted, humble, or nonreactive, then there is a fixed point of the mapping G in $\widetilde{\Delta}_n$, i.e. an equilibrium value of the social power vector.*

2. *If every individual is either well adjusted or humble, then the fixed point of G in $\widetilde{\Delta}_n$ is unique, and in fact lies in* $\text{int}(\Delta_n)$*, and is approached exponentially fast from any initial condition for which $x_i(0) < 1$ for all i and $\zeta_j(0) > 0$ for some j.*

3. *If individual v_i is has an emotional ϕ_i function, then the fixed point e_i is locally exponentially stable if and only if $\gamma_i^{-1}(1-\gamma_i)\phi_i' < 1$, where ϕ_i' is the derivative of ϕ_i with respect to ζ_i evaluated at $\zeta_i = 1$.*

4. *Networks of emotional individuals may have multiple attractive equilibria, including one in* $\text{int}(\Delta_n)$.

The sorts of conclusions in the last two points involving emotional individuals also apply to arrogant individuals. Evidently, either of these behavioral characteristics can create a major upset in the type of results that can occur, compared to the original model with no special behavior, under Assumption 3. Crucially, a star graph is no longer a prerequisite for the emergence of an autocrat. Emotional behavior can also lead to the emergence of an autocrat, provided the network begins with $\zeta(0)$ close to the corner of Δ_n associated with an emotional individual and the inequality on statement 3 is satisfied.

The result of Statement 3 is not proved in [32], but the proof is straightforwardly obtained by deriving the necessary and sufficient conditions for the eigenvalues of the Jacobian to be less than 1 in magnitude at the corner of Δ_n associated with an emotional individual. The expression for the Jacobian itself is given in [15]. Finally, we remark that a comprehensive convergence analysis is still missing for networks of nonreactive, emotional or arrogant individuals (or a network containing a mixture of such individuals).

8 Other works on self-appraisal dynamics

Besides the work covered in the previous sections, a number of other results have been developed for self-appraisal dynamics originating from the ideas of Noah E. Friedkin in [8]. For the purposes of brevity, we do not go into extensive details here.

One of the key assumptions imposed was on the irreducibility of C, Assumption 1, which implies the associated underlying graph is strongly connected. Reducible C matrices are considered comprehensively in [33]. In another direction, the work [26] explores a number of results for the model without special behaviors of the type considered in the previous section, but in which several significant expansions of the earlier model and the results thereon are considered: i) $C(s)$ matrices vary with topic sequence but the variations around a nominal constant C are sufficiently small for all s, ii) for constant C, convergence rates are identified as a function of the fixed point x^* for non-star graphs, iii) individuals retain memory of previous relative interaction matrices when $C(s)$ varies over the topic sequence, and iv) individuals have memory of their previous self-confidence in past topics.

Moving further from the original model (that adopts Assumptions 1 and 3), a single-time scale for both the opinion dynamics and the self-appraisal update is considered in [34, 35] which relaxes the fast-slow time-scale separation discussed below Assumption 1. In the real world, individuals often fail to come to a complete consensus of opinions (but may still become closer in opinion over time), and this can be captured in an extension of the DeGroot model, called the Friedkin-Johnsen model [36] which posits that individuals remain somewhat stubbornly attached to their initial opinion. Replacement of the DeGroot opinion dynamics with Friedkin-Johnsen dynamics is explored in [37, 38] although there is still lacking a comprehensive convergence result for all topologies and model parameters. Continuous-time versions of the DeGroot-Friedkin model are considered in [26, 39].

So far, the above works have all considered the matrix C with all entries nonnegative, which can be viewed as representing positive interactions between individuals which are friendly or cooperative. Negative weights on edges representing interactions have come into focus [40, 41], representing enemy/competitive/antagonistic relationships. The DeGroot-Friedkin model with negative interactions has been explored in [42, 43]. For strongly connected networks, it has been shown that negative interactions in opinion updating lead to either i) all opinions converging to a consensus at the 0 value (in which case there is no easily definable social power measure), or ii) an equivalent DeGroot dynamics with all positive interactions under a simple coordinate transform [44, 45] (in which case the above results already provide a comprehensive treatment).

9 Reflections and future work

Much of this paper has been devoted to detailed examination of an existing model, the DeGroot-Friedkin model, for opinion dynamics which has the great advantage of being backed by experimental data. The most significant conclusions beyond those of the major work of [9] deal with the nature of the convergence, nonexponential but asymptotic for star graphs, and exponential for non-star graphs. They can be regarded as resting also on the experimental data.

We have also treated two significant extensions of this base scenario. One deals with changing relative interaction matrices, including periodically changing matrices. While not backed by experimental data (as distinct from simulations), the general conclusions appear to accord with what intuition suggests. In broad terms, initial conditions are forgotten exponentially fast. Limiting behavior is determined just by the sequence of topologies (including the associated weights embedded in the sequence of relative interaction matrices). The other significant extension is to attempt to include behavioral characteristics of individuals in the model. That such characteristics exist, resulting in individuals over- or underestimating their contribution to a discussion outcome, is something that most people would endorse. It requires however a significant leap to endorse the particular way these are captured in the models for such behavior we have advanced. Nevertheless, one of the outcomes of such models, to the effect that an arrogant individual can end up dominating a group, is an outcome which accords with many people's experience.

As far as future work is concerned, any experimental verification of the ideas summarized here but going beyond [9] would be welcome. In terms of theoretical work, it would seem that investigations of the incorporation of self-appraisal going beyond those reported in the previous section would be welcome in relation to Friedkin-Johnson models, incorporating bias, and in adding self-appraisal to the treatment of simultaneous examination of multiple differing but dependent issues [46–48]. Another possibility again for theoretical work would be to consider those extensions of the DeGroot model such as the Hegselmann-Krause bounded confidence model, [49], [50], [51] (the key feature of such models is that an individual limits the extent to which they take account of a very different opinion), and extend such models to accommodate self-appraisal.

Acknowledgements **While Professor Lei Guo did not directly contribute to the ideas of this paper, we wish to acknowledge his profound influence on the development of control science and applications to complex systems, and the inspiration to those of both older and younger generations including ourselves. We also thank the editor and his colleagues for their initiative in putting together this special issue, and the referees for the helpful comments on our manuscript.**

References

[1] J. R. P. French Jr, "A Formal Theory of Social Power," *Psychological Review*, vol. 63, no. 3, pp. 181–194, 1956.

[2] F. Harary, *Studies in social power.* Ann Arbor, University of Michigan Press, 1959, ch. A criterion for unanimity in French's theory of social power, pp. 198–182.

[3] M. H. DeGroot, "Reaching a Consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[4] W. Ren and R. W. Beard, "Consensus Seeking in Multiagent Systems Under Dynamically Chang-

ing Interaction Topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[5] C. H. Cooley, *Human Nature and the Social Order*. Transaction Publishers, 1992.

[6] J. S. Shrauger and T. J. Schoeneman, "Symbolic interactionist view of self-concept: Through the looking glass darkly." *Psychological Bulletin*, vol. 86, no. 3, p. 549, 1979.

[7] V. Gecas and M. L. Schwalbe, "Beyond the Looking-Glass Self: Social Structure and Efficacy-Based Self-Esteem," *Social Psychology Quarterly*, pp. 77–88, 1983.

[8] N. E. Friedkin, "A Formal Theory of Reflected Appraisals in the Evolution of Power," *Administrative Science Quarterly*, vol. 56, no. 4, pp. 501–529, 2011.

[9] P. Jia, A. MirTabatabaei, N. E. Friedkin, and F. Bullo, "Opinion Dynamics and the Evolution of Social Power in Influence Networks," *SIAM Review*, vol. 57, no. 3, pp. 367–397, 2015.

[10] M. Ye, J. Liu, B. D. O. Anderson, C. Yu, and T. Başar, "Modification of Social Dominance in Social Networks by Selective Adjustment of Interpersonal Weights," in *Proceedings of IEEE 56th Annual Conference on Decision and Control, Melbourne, Australia*, 2017, pp. 2906–2911.

[11] ——, "Modification of Social Dominance in Social Networks by Selective Adjustment of Interpersonal Weights," 2017, arXiv:1703.03166 [cs:SI]. [Online]. Available: https://arxiv.org/abs/1703.03166

[12] ——, "On the Analysis of the DeGroot-Friedkin Model with Dynamic Relative Interaction Matrices," in *Proceedings of the 20th IFAC World Congress, Toulouse, France*, 2017, pp. 12 408–12 413.

[13] B. D. O. Anderson and M. Ye, "Nonlinear Mapping Convergence and Application to Social Networks," in *European Control Conference, Limassol, Cyprus*, Jun. 2018, pp. 557–562.

[14] M. Ye, J. Liu, B. D. O. Anderson, C. Yu, and T. Başar, "Evolution of Social Power in Social Networks with Dynamic Topology," *IEEE Transaction on Automatic Control*, vol. 63, no. 11, pp. 3793–3808, Nov. 2018.

[15] M. Ye and B. D. O. Anderson, "Modelling of Individual Behaviour in the DeGroot–Friedkin Self-Appraisal Dynamics on Social Networks," in *European Control Conference, Naples, Italy*, Jun. 2019, pp. 2011–2017.

[16] B. D. O. Anderson and M. Ye, "Extrema without convexity and stability without Lyapunov," *Communications in Information and Systems*, vol. 20, no. 3, pp. 253–281, 2020, mathematical Engineering: A special issue at the occasion of the 85th birthday of Prof. Thomas Kailath.

[17] M. Ye, *Opinion Dynamics and the Evolution of Social Power in Social Networks*. Springer, Cham, 2019.

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, New York, 2012.

[19] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, ser. Computer Science and Applied Mathematics. Academic Press: London, 1979.

[20] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part I," *Annual Reviews in Control*, vol. 43, pp. 65–79, 2017.

[21] C. D. Godsil and G. Royle, *Algebraic Graph Theory*. Springer: New York, 2001, vol. 207.

[22] C. Ravazzi, P. Frasca, R. Tempo, and H. Ishii, "Ergodic Randomized Algorithms and Dynamics Over Networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 1, pp. 78–87, 2015.

[23] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part II," *Annual Reviews in Control*, vol. 45, pp. 166–190, 2018.

[24] N. E. Friedkin, P. Jia, and F. Bullo, "A Theory of the Evolution of Social Power: Natural Trajectories of Interpersonal Influence Systems along Issue Sequences," *Sociological Science*, vol. 3, pp.

444–472, 2016.

[25] N. E. Friedkin and F. Bullo, "How truth wins in opinion dynamics along issue sequences," *Proceedings of the National Academy of Sciences*, vol. 114, no. 43, pp. 11 380–11 385, 2017.

[26] G. Chen, X. Duan, N. E. Friedkin, and F. Bullo, "Social Power Dynamics over Switching and Stochastic Influence Networks," *IEEE Transactions on Automatic Control*, vol. 64, no. 2, pp. 582–597, 2019.

[27] Y. Chen, W. Xia, M. Cao, and J. Lü, "Random asynchronous iterations in distributed coordination algorithms," *Automatica*, vol. 109, p. 108505, Nov. 2019.

[28] V. Guillemin and A. Pollack, *Differential Topology*. American Mathematical Soc., 2010, vol. 370.

[29] M. A. Khamsi and W. A. Kirk, *An Introduction to Metric Spaces and Fixed Point Theory*. John Wiley & Sons, 2011, vol. 53.

[30] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.

[31] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001.

[32] M. Ye and B. D. O. Anderson, "Effects of Behaviour in Self-Appraisal Dynamics on Social Networks," submitted for publication, 2018.

[33] P. Jia, N. E. Friedkin, and F. Bullo, "Opinion Dynamics and Social Power Evolution over Reducible Influence Networks," *SIAM Journal on Control and Optimization*, vol. 55, no. 2, pp. 1280–1301, 2017.

[34] ——, "Opinion Dynamics and Social Power Evolution: A Single-Timescale Model," *IEEE Transactions on Control of Networked Systems*, 2019, to appear.

[35] Z. Xu, J. Liu, and T. Başar, "On a Modified DeGroot-Friedkin Model of Opinion Dynamics," in *American Control Conference (ACC), Chicago, USA*, July 2015, pp. 1047–1052.

[36] N. E. Friedkin and E. C. Johnsen, "Social Influence and Opinions," *Journal of Mathematical Sociology*, vol. 15, no. 3-4, pp. 193–206, 1990.

[37] A. Mirtabatabaei, P. Jia, N. E. Friedkin, and F. Bullo, "On the reflected appraisals dynamics of influence networks with stubborn agents," in *2014 American Control Conference*, June 2014, pp. 3978–3983.

[38] Y. Tian, P. Jia, A. Mirtabatabaei, L. Wang, N. E. Friedkin, and F. Bullo, "Social Power Evolution in Influence Networks with Stubborn Individuals," *IEEE Transactions on Automatic Control*, 2021, to appear.

[39] X. Chen, J. Liu, M.-A. Belabbas, Z. Xu, and T. Başar, "Distributed Evaluation and Convergence of Self-Appraisals in Social Networks," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 291–304, 2017.

[40] C. Altafini, "Consensus Problems on Networks with Antagonistic Interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, 2013.

[41] W. Xia, M. Cao, and K. Johansson, "Structural Balance and Opinion Separation in Trust–Mistrust Social Networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 46–56, 2016.

[42] D. Xue, S. Hirche, and M. Cao, "Evolution of social power over influence networks containing antagonistic interactions," *Information Sciences*, vol. 540, pp. 449–468, 2020.

[43] S. Bhowmick and S. Panja, "Evolution of self-confidence in opinion dynamics over signed network," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 2975–2981, 2020.

[44] J. Liu, X. Chen, T. Başar, and M.-A. Belabbas, "Exponential Convergence of the Discrete- and Continuous-Time Altafini Models," *IEEE Transaction on Automatic Control*, vol. 62, no. 12, pp. 6168–6182, 2017.

[45] J. M. Hendrickx, "A lifting approach to models of opinion dynamics with antagonisms," in *IEEE 53rd Annual Conference on Decision and Control (CDC)*. IEEE, 2014, pp. 2118–2123.

[46] M. Ye, M. H. Trinh, Y.-H. Lim, B. D. O. Anderson, and H.-S. Ahn, "Continuous-time Opinion Dynamics on Multiple Interdependent Topics," *Automatica*, vol. 115, no. 5, p. 108884, May 2020.

[47] N. E. Friedkin, A. V. Proskurnikov, R. Tempo, and S. E. Parsegov, "Network science on belief system dynamics under logic constraints," *Science*, vol. 354, no. 6310, pp. 321–326, 2016.

[48] M. Ye, J. Liu, L. Wang, B. D. O. Anderson, and M. Cao, "Consensus and Disagreement of Heterogeneous Belief Systems in Influence Networks," *IEEE Transactions on Automatic Control*, vol. 65, no. 11, pp. 4679–4694, Nov. 2020.

[49] 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.

[50] S. R. Etesami and T. Başar, "Game-Theoretic Analysis of the Hegselmann-Krause Model for Opinion Dynamics in Finite Dimensions," *IEEE Transactions on Automatic Control*, vol. 60, no. 7, pp. 1886–1897, 2015.

[51] J. Lorenz, "Continuous Opinion Dynamics Under Bounded Confidence: A Survey," *International Journal of Modern Physics C*, vol. 18, no. 12, pp. 1819–1838, 2007.