

Discrete-Time Polar Opinion Dynamics with Heterogeneous Individuals

Ji Liu, Mengbin Ye, Brian D.O. Anderson, Tamer Başar, Angelia Nedić

Abstract—This paper considers a discrete-time opinion dynamics model in which each individual’s susceptibility to being influenced by others is dependent on her current opinion. We first propose a general opinion dynamics model based on the DeGroot model, with a general function to describe the functional dependence of each individual’s susceptibility to her own opinion, and characterize the set of all equilibria and stability of nontrivial equilibria. We then consider two classes of functions in which the individual’s susceptibility depends on the *polarity* of her opinion (i.e., how extreme her opinion is), and provide motivating social examples. First, we consider stubborn positives, who have reduced susceptibility if their opinions are at one end of the interval and increased susceptibility if their opinions are at the opposite end. Second, we consider stubborn extremists, who are less susceptible when they hold opinions at either end of the opinion interval. For each susceptibility model, we establish limiting behavior for different initial conditions. Networks consisting of individuals with both types of susceptibility functions are also considered.

I. INTRODUCTION

The problem of opinion dynamics, which considers how an individual’s opinion forms and evolves through interactions with others in a social network, has been widely studied in the social sciences for decades. The classical discrete-time DeGroot model, in which each individual updates her opinion by taking a convex combination of the opinions of her neighbors at each time step, is perhaps one of the most well known models [1]. This model is closely related to discrete-time linear consensus algorithms, which have been heavily studied in multi-agent coordination literature [2]–[4]. Since the time the DeGroot model was proposed, numerous other models have been introduced, in both continuous- and discrete-time settings. These various models, which describe the opinion formation process in the context of different social processes, all attempt to understand the formation and evolution of opinions in social networks of all sizes, and

The proofs of some theorems in this paper are omitted due to space limitations and will be given in an expanded version of the paper.

This work was supported by the Australian Research Council (ARC) under the ARC grant DP-160104500 and by Data61-CSIRO, Australia.

J. Liu is with the Department of Electrical and Computer Engineering at Stony Brook University (ji.liu@stonybrook.edu). M. Ye is with the Research School of Engineering at Australian National University, and the Faculty of Science and Engineering, University of Groningen (m.ye@rug.nl). B.D.O. Anderson is with the Research School of Engineering at Australian National University (Brian.Anderson@anu.edu.au), Hangzhou Dianzi University, Hangzhou, China, and Data61-CSIRO (formerly NICTA Ltd.) in Canberra, A.C.T., Australia. T. Başar is with the Department of Electrical and Computer Engineering at University of Illinois at Urbana-Champaign (basarl@illinois.edu). A. Nedić is with the School of Electrical, Computer and Energy Engineering at Arizona State University (angelia.nedich@asu.edu).

explain observed social phenomena such as polarization or attitude extremity [5]–[8], and subculture formation [9], [10].

There are many variants of the DeGroot model for opinion dynamics. The Altafini model, which assumes that the interactions between individuals can be cooperative or antagonistic, has been studied as a discrete-time process in [11]–[13], and the continuous-time counterpart has been considered in [13]–[15]. A notable conclusion is that the model links the limiting opinion behavior with the structural balance of the graph representing the social network. Some other models primarily focus on linking the limiting opinion behavior with a social process for an individual. For example, the Hegselmann-Krause model shows that the social process of homophily is linked to the fact that opinions in social network eventually form clusters [16]–[18]. It was shown in [9] that an individual’s desire to strive for uniqueness could generate subcultures which continuously form and vanish over time. On the other hand, an individual conforming to a social norm could generate pluralistic ignorance [10]. Finally, some models attempt to link final opinion behavior to a combination of social processes and the underlying network structure. The Friedkin-Johnsen model [19], [20] considers individual susceptibility to influence and shows that opinions reach a persistent diversity under general graph structures. The DeGroot-Friedkin model [21], [22] studied an individual’s ability to reflect on her impact in the opinion formation process, and showed that her self-confidence depended on the graph structure.

A key aspect of the DeGroot model is the interpersonal influence, which describes the amount of influence each individual’s neighbors have in determining that individual’s new opinion. Some of the results consider arbitrary, time-varying interpersonal influence, e.g. [12], [23]. However, many of the aforementioned models consider influence determined by a social process, e.g. homophily [18], social distancing [9], [10], conformity [10], desire for uniqueness [9], biased assimilation [24], or reflected self-appraisal [21], [22]. Because the social process is often dependent on the states, i.e. opinions (which change with time), necessarily the interpersonal influences are state-dependent, and thus time-varying. A recent paper [25] introduced a continuous-time model, called “polar opinion dynamics”. The term “polar” relates to the fact that the level of influence depends on how extreme, i.e. polar, an individual’s opinion is. For fixed network topologies, a sufficient condition for consensus is established for general susceptibility functions. Then, three different cognitive processes to drive the influence change are captured by three special susceptibility functions. Each case is separately analyzed, under the assumption that all

individuals' functions are homogeneous.

In this paper, we study a discrete-time opinion dynamics model where an individual's susceptibility to influence is dependent on her current opinion, bearing in mind that discrete-time models may be more appropriate to describe opinion dynamics, at least from the viewpoint that individuals change their minds from time to time, instead of continuously. The first contribution of this paper is to propose a general model, and establish some general properties of the model, including the set of all equilibria and their stability.

The second main contribution of this paper is to go beyond [25] by considering general functions for two classes of susceptibility functions (with the ones considered in [25] as special examples) and by assuming the functions are heterogeneous among the individuals (since individuals may naturally have different susceptibility to interpersonal influence). Limiting behavior of the model with each type of susceptibility function under different initial conditions is established, including when consensus is achieved exponentially fast (only asymptotic convergence to consensus was established in [25]). In addition, social examples from existing literature are provided to motivate these susceptibility functions; [25] introduced, but did not provide thorough justification for two of the cognitive processes.

The third major contribution, which was not explored by [25], is to study networks which allow individuals with two different types of susceptibility functions to coexist. Moreover, we characterize all possible equilibria and the stability of some of them. In summary, this paper considers a discrete-time model with heterogeneous individuals within the same network in two senses: (1) all the individuals are within the same class of susceptibility functions, but can have different specific functional forms; (2) individuals have two different types susceptibility functions.

Notation: For any positive integer n , we use $[n]$ to denote the index set $\{1, 2, \dots, n\}$. We view vectors as column vectors and write x^\top to denote the transpose of a vector x . For a vector x , we use x_i to denote the i th entry of x . For any matrix $M \in \mathbb{R}^{n \times n}$, we use m_{ij} to denote its ij th entry. A nonnegative $n \times n$ matrix is called a stochastic matrix (or row-stochastic matrix) if its row sums are all equal to 1. We use $\mathbf{0}$ and $\mathbf{1}$ to denote the vectors whose entries all equal to 0 and 1, respectively, and I to denote the identity matrix, while the dimensions of the vectors and matrices are to be understood from the context. For a real square matrix M , we use $\rho(M)$ to denote its spectral radius. For any real number x , we use $|x|$ to denote the absolute value of x . For any two real vectors $a, b \in \mathbb{R}^n$, we write $a \geq b$ if $a_i \geq b_i$ for all $i \in [n]$, $a > b$ if $a \geq b$ and $a \neq b$, and $a \gg b$ if $a_i > b_i$ for all $i \in [n]$. For any two sets \mathcal{A} and \mathcal{B} , we use $\mathcal{A} \setminus \mathcal{B}$ to denote the set of elements in \mathcal{A} but not in \mathcal{B} . We will use the terms "individual" and "agent" interchangeably.

A. Preliminaries

The graph of an $n \times n$ matrix M with real-valued entries is an n -vertex directed graph defined so that (i, j) is an arc from vertex i to vertex j in the graph whenever the ji th entry

of M is nonzero. A directed graph \mathbb{G} is strongly connected if there is a directed path between each pair of its distinct vertices. A matrix is irreducible if and only if its graph is strongly connected.

Lemma 1: (Lemma 2.3 in [26]) Suppose that M is an irreducible nonnegative matrix. Then, $\rho(M)$ is a simple eigenvalue of M , and there exists a unique (up to scalar multiple) vector $x \gg \mathbf{0}$ such that $Mx = \rho(M)x$.

Lemma 2: (Lemma 2.6 in [26]) Suppose that A is an irreducible nonnegative matrix. If B is a principal square submatrix of A , then $\rho(B) < \rho(A)$.

Lemma 3: (Lemma 1 in [25] and the Neumann series in [27]) Suppose that M is an irreducible stochastic matrix with the expression

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

where M_{11} and M_{22} are square submatrices. Then, $I - M_{22}$ is invertible, and $(I - M_{22})^{-1}M_{21}$ is a stochastic matrix.

Note that Lemma 1 in [25] concluded that $(I - M_{22})^{-1}M_{21}\mathbf{1} = \mathbf{1}$, but did not conclude that $(I - M_{22})^{-1}M_{21}$ is stochastic. To show the stochasticity, first consider M_{22} . Since M is irreducible and stochastic, from Lemma 2, $\rho(M_{22}) < 1$. Then, the Neumann series yields $(I - M_{22})^{-1} = \sum_{k=0}^{\infty} M_{22}^k$ [27]. Since M_{21} is also nonnegative, it follows that $(I - M_{22})^{-1}M_{21}$ is nonnegative. Since $(I - M_{22})^{-1}M_{21}\mathbf{1} = \mathbf{1}$, we conclude the stochasticity property. This property will later be used to provide insight into equilibria of the model.

II. THE MODEL

In this section, we propose a general model for describing opinion dynamics where each individual's susceptibility to being influenced by others is affected by some social process, and give some results on the trajectories of the opinions. In the next section, we shall consider two specific models to describe two different social processes that might affect an individual's susceptibility.

Consider a social network of $n > 1$ agents, labeled 1 through n , discussing opinions on a given topic. Each agent i can only learn, and be influenced by, the opinions of certain other agents called the neighbors of agent i . Neighbor relationships among the n agents are described by a directed graph \mathbb{G} , called the neighbor graph. Agent j is a neighbor of agent i whenever (j, i) is an arc in \mathbb{G} . Thus, the directions of arcs indicate the directions of information flow (specifically opinion flow). For convenience, we assume that each agent is a neighbor of herself. Thus, \mathbb{G} has self-arcs at all n vertices. Each agent i has an opinion x_i , which is a real-valued quantity that may change due to interpersonal influence over the graph.

In the DeGroot model [1], each agent i 's opinion evolves at each discrete time $t \in \{0, 1, 2, \dots\}$ as

$$x_i(t+1) = \sum_{j \in \mathcal{N}_i} w_{ij} x_j(t), \quad i \in [n],$$

where \mathcal{N}_i denotes the set of neighbors of agent i including i herself, and w_{ij} are positive influence weights satisfying

$j \notin \mathcal{N}_i \Rightarrow w_{ij} = 0$ and $\sum_{j \in \mathcal{N}_i} w_{ij} = 1$, for all $i \in [n]$. To illustrate the interpersonal influence, we rewrite the model as

$$\begin{aligned} x_i(t+1) &= x_i(t) + \left(\sum_{j \in \mathcal{N}_i} w_{ij} x_j(t) - x_i(t) \right) \\ &= x_i(t) + \sum_{j \in \mathcal{N}_i} w_{ij} (x_j(t) - x_i(t)), \end{aligned}$$

with the second equality obtained by using $\sum_{j \in \mathcal{N}_i} w_{ij} = 1$. By defining

$$u_i(t) = \sum_{j \in \mathcal{N}_i} w_{ij} (x_j(t) - x_i(t)),$$

it is clear that $u_i(t)$ represents the influence of agent i 's neighbors, which generates a change in the opinion of agent i at time t . We now suppose that agent i may not fully accept the influence of her neighbors, and her openness to influence, or susceptibility, is captured by the real-valued function $f_i(x_i(t))$. We make the following assumption on f_i throughout the paper.

Assumption 1: The susceptibility function f_i takes on values in $[0, 1]$.

We consider the following model for opinion dynamics with susceptibility:

$$x_i(t+1) = x_i(t) + f_i(x_i(t)) \sum_{j \in \mathcal{N}_i} w_{ij} (x_j(t) - x_i(t)). \quad (1)$$

The model is an analog of the continuous-time model in [25]. In the case when $f_i(x_i(t)) = 1$, agent i fully accepts her neighbors' influence at time t , and the model reduces to the DeGroot model. In the case when $f_i(x_i(t)) = 0$, agent i will ignore her neighbors and not change her opinion at time t ; in such a case, the agent is sometimes called stubborn [19], [28]. It is worth emphasizing that an agent's susceptibility function depends on her current opinion, i.e. state. This is consistent with the many works discussed in the introduction, which considers social processes which are dependent on the individual's opinion and, in some instances, the opinions of her neighbors.

For convenience of analysis, (1) may be rearranged to be written as

$$\begin{aligned} x_i(t+1) &= x_i(t) + f_i(x_i(t)) \left(\sum_{j \in \mathcal{N}_i} w_{ij} x_j(t) - x_i(t) \right) \\ &= (1 - f_i(x_i(t))) x_i(t) + f_i(x_i(t)) \sum_{j \in \mathcal{N}_i} w_{ij} x_j(t). \end{aligned}$$

The above set of n equations can be combined in a state-space form. Toward this end, let $x(t)$ be the vector in \mathbb{R}^n whose i th entry equals $x_i(t)$, $F(x(t))$ be the $n \times n$ diagonal matrix whose i th diagonal entry equals $f_i(x_i(t))$ with $0 \leq f_i \leq 1$, and W be the $n \times n$ matrix whose ij th entry equals w_{ij} if $j \in \mathcal{N}_i$, and zero otherwise. Then, it follows that

$$\begin{aligned} x(t+1) &= (I - F(x(t)))x(t) + F(x(t))Wx(t) \\ &= S(x(t))x(t), \end{aligned} \quad (2)$$

where $S(x(t)) = I - F(x(t)) + F(x(t))W$. It is worth noting that $S(x(t))$ is a function of $x(t)$ since $F(x(t))$ is. Thus, (2) is a nonlinear system. From Assumption 1 and the fact that W is a stochastic matrix, it is easy to verify that $S(x(t))$ is a stochastic matrix for all times t . Thus, each $x_i(t+1)$ is a convex combination of all $x_i(t)$, $i \in [n]$.

In this paper, we assume that all the initial opinions $x_i(0)$, $i \in [n]$, lie in the interval $[-1, 1]$, where -1 and 1 represent the extreme positive and negative opinions, respectively. Such a scaling is typical in opinion dynamics problems where x_i may represent individual i 's attitude towards an idea, e.g. the legalization of recreational marijuana, with $x_i = 1$ maximally supporting and $x_i = -1$ maximally opposing. The following lemma, whose proof is simple and thus omitted, shows that $[-1, 1]$ is an invariant set of each agent's opinion dynamics given by (1), and that the most negative and positive opinions will never become more negative and more positive, respectively.

Lemma 4: Suppose that each agent i follows the update rule (1) and that $x_i(0) \in [-1, 1]$ for all $i \in [n]$. Then, $x_i(t) \in [-1, 1]$ for all $i \in [n]$ and time t . Moreover, $x_{\min}(t) = \min_i x_i(t)$ is nondecreasing and $x_{\max}(t) = \max_i x_i(t)$ is nonincreasing as t increases.

III. MAIN RESULTS

In this section, we provide several main results of model (1). We impose the following assumption on \mathbb{G} throughout the paper, which implies that W (defined above (2)) is irreducible.

Assumption 2: The neighbor graph \mathbb{G} is strongly connected.

The set of all equilibria of model (1) is characterized by the following theorem.

Theorem 1: Every equilibrium of system (1) has one of the following forms:

- 1) $x = a\mathbf{1}$, $a \in [-1, 1]$;
- 2) x such that $f_i(x_i) = 0$ for all $i \in [n]$;
- 3) x defined as follows: Let $\mathcal{I} = \{i_1, i_2, \dots, i_m\}$, with $1 \leq m < n$, be any nonempty proper subset of $[n]$. Let π be any permutation map for which $\pi(i_j) = j$ for all $j \in [m]$, and P be the corresponding permutation matrix. Write

$$PWP^\top = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix},$$

where W_{11} is an $m \times m$ matrix. The equilibrium is

$$x = P^\top \begin{bmatrix} y \\ z \end{bmatrix},$$

where $f_{i_j}(y_j) = 0$ for all $j \in [m]$, and

$$z = (I - W_{22})^{-1} W_{21} y.$$

Proof: Cases 1) and 2) are obvious from examination of (2). We thus only need to prove that case 3) contains all nontrivial equilibria. Suppose that x is a nontrivial equilibrium, i.e., it is not included in case 1) or 2). Then, there exists at least one index $i \in [n]$ such that $f_i(x_i) \neq 0$. If $f_i(x_i) \neq 0$

for all $i \in [n]$, then it follows from (1) that $Wx = x$. Since W is an irreducible stochastic matrix, from Lemma 1, x must be equal to $\alpha \mathbf{1}$ for some $\alpha \in \mathbb{R}$, which contradicts the hypothesis. Thus, there exists at least one index $j \in [n]$ such that $f_j(x_j) = 0$. Without loss of generality, assume that $f_j(x_j) = 0$ for $j \in \{i_1, i_2, \dots, i_m\}$ with $1 \leq m < n$. To be an equilibrium of system (1), it must be true that $z = W_{22}z + W_{21}y$. Since W is an irreducible stochastic matrix, so is PWP^\top . From Lemma 3, $I - W_{22}$ is invertible. Therefore, $z = (I - W_{22})^{-1}W_{21}y$. ■

Theorem 1 says that three types of equilibria exist for the model. Case 1) implies that a consensus among all the individuals' opinions has been reached. Case 2) implies that all the individuals are at a point where they are maximally stubborn (which includes consensus as a special case if that point is the same for all individuals). Case 3) includes an equilibrium of general disagreement in the network, in which individual $i \in \mathcal{I}$ is maximally closed to influence. From Lemma 3, $(I - W_{22})^{-1}W_{21}$ is a stochastic matrix, which implies that each entry of z lies in the convex hull of y_j , $j \in [m]$. Consensus of opinions is a special case of 3).

We call an equilibrium trivial if it is included in case 1) or 2), and nontrivial otherwise. It is worth emphasizing that the three cases may be disjoint or not, depending on when the susceptibility functions (which are in general different for each individual) are equal to zero. From Theorem 1 and its proof, if x is a nontrivial equilibrium, there exist $i, j \in [n]$ such that $f_i(x_i) = 0$ and $f_j(x_j) \neq 0$, with $i \in \mathcal{I}$ and $j \notin \mathcal{I}$.

Suppose that x is a nontrivial equilibrium of system (1). From Theorem 1 and its proof, x can be uniquely given in the following manner. First, let \mathcal{I} be a subset of n such that $f_i(x_i) = 0$ if $i \in \mathcal{I}$, and $f_i(x_i) \neq 0$ if $i \in [n] \setminus \mathcal{I}$. Then, all x_i , $i \in \mathcal{I}$, constitutes y in Theorem 1, and z can be computed via $(I - W_{22})^{-1}W_{21}y$. The stability of such a nontrivial equilibrium may be determined by checking its $n \times n$ Jacobian matrix J whose entries are given by

$$J_{ii} = 1 - f_i(x_i) + w_{ii}f_i'(x_i) + f_i'(x_i) \sum_{j=1}^n w_{ij}(x_j - x_i), \quad (3)$$

$$J_{ij} = w_{ij}f_i'(x_i), \quad i \neq j, \quad (4)$$

with the assumption that each susceptibility function f_i is differentiable, whose derivative is denoted by f_i' . A nontrivial equilibrium may be stable or unstable.

Without loss of generality, let $\mathcal{I} = [m]$ with $1 \leq m < n$. Fix $x_i = y_i$ such that $f_i(y_i) = 0$ for all $i \in [m]$. Now we consider the reduced system only consisting of variables x_i , $i \in \{m+1, \dots, n\}$, which is defined as follows:

$$\begin{aligned} x_i(t+1) &= x_i(t) + f_i(x_i(t)) \sum_{j=1}^m w_{ij}(y_j - x_i(t)) \\ &\quad + f_i(x_i(t)) \sum_{j=m+1}^n w_{ij}(x_j(t) - x_i(t)), \\ &i \in \{m+1, \dots, n\}. \end{aligned} \quad (5)$$

From the preceding discussion, system (5) has a unique equilibrium $(I - W_{22})^{-1}W_{21}y$. The following result states that the unique equilibrium is locally exponentially stable.

Theorem 2: System (5) has a unique locally exponentially stable equilibrium.

Proof: Since $f_i(x_i) \neq 0$ for all $i \in \{m+1, \dots, n\}$, it follows from (5) that

$$\sum_{j=1}^m w_{ij}(y_j - x_i) + \sum_{j=m+1}^n w_{ij}(x_j - x_i) = 0$$

at the equilibrium point. With this fact and from (3) and (4), the Jacobian matrix of system (5) at the unique equilibrium equals $I - F_2(x) + F_2(x)W_{22}$, where $F_2(x)$ is the diagonal matrix whose diagonal entries consist of $f_i'(x_i)$, $i \in \{m+1, \dots, n\}$. Note that all diagonal entries of $F_2(x)$ are positive. Let F_1 be any $m \times m$ diagonal matrix whose diagonal entries are all in $(0, 1]$. Then, it is easy to check that

$$V = \begin{bmatrix} I - F_1 & 0 \\ 0 & I - F_2(x) \end{bmatrix} + \begin{bmatrix} F_1 & 0 \\ 0 & F_2(x) \end{bmatrix} W$$

is a stochastic matrix. Since W is an irreducible and all diagonal entries of F_1 and $F_2(x)$ are positive, V is also irreducible. Note that

$$V = \begin{bmatrix} I - F_1 + F_1W_{11} & F_1W_{12} \\ F_2(x)W_{21} & I - F_2(x) + F_2(x)W_{22} \end{bmatrix}.$$

From Lemma 2, $\rho(I - F_2(x) + F_2(x)W_{22}) < \rho(V) = 1$, which implies that the unique equilibrium is locally exponentially stable. ■

We call a vector of opinions $x \in \mathbb{R}^n$ a consensus state if all its entries are equal; otherwise, we call it a non-consensus state. A nontrivial equilibrium must be a non-consensus state.

Theorem 3: If all n susceptibility functions do not have any zero point or have the same single zero point, then all equilibria of system (1) are consensus states, and the system will asymptotically reach a consensus.

Proof: Let $\mathcal{C} = \{a\mathbf{1} : a \in [-1, 1]\}$ denote the set of all consensus states, which are equilibria of the system. Consider the Lyapunov function

$$V(x(t)) = \max_{i \in [n]} x_i(t) - \min_{j \in [n]} x_j(t).$$

It is clear that among all $x \in [-1, 1]^n$, $V(x(t)) = 0$ if and only if $x \in \mathcal{C}$. If $x(t) \notin \mathcal{C}$, since the neighbor graph is strongly connected, there must exist a finite positive time step k , such that $V(x(t+k)) < V(x(t))$. By Lyapunov's stability theorem for discrete-time autonomous systems [29], the system will converge to a point in \mathcal{C} , which implies that the system will reach a consensus. ■

Next we will consider networks where all individuals have one of two classes of susceptibility functions, and study the behavior of the corresponding models. In conjunction, we provide motivating examples from sociology for the susceptibility functions taking on these general forms.

A. Stubborn Positives

We begin with the case where $f_i(x_i(t))$ satisfies the following assumption.

Assumption 3: The susceptibility function $f_i(x_i)$ is a smooth and monotone for $x_i \in [-1, 1]$, takes values in $[0, 1]$, and equals 0 only when $x_i = 1$.

An agent with such a susceptibility function is said to be stubborn positive. In [25], a special case $f_i(x_i) = \frac{1}{2}(1 - x_i)$ is considered for all $i \in [n]$. It is worth emphasizing that the assumption allows heterogeneous susceptibility functions, which is also true for stubborn extremists discussed later.

Motivation: Note that for stubborn positive agents, her susceptibility decreases as $x_i \rightarrow 1$, and increases as $x_i \rightarrow -1$. In other words, the closer an agent is to a “positive opinion” (respectively a “negative opinion”), the more stubborn or unwilling (respectively more open or susceptible) she is to changing her opinion. Our motivating example is a jury panel. The paper [30] conducted extensive surveys of criminal juries after trials were complete. A clear pattern was observed: a juror was more likely to be extremely stubborn when believing the defendant should be acquitted, than when believing the defendant should be convicted; asymmetric stubbornness arose from the fact that a false conviction carried an enormous amount of consequence for defendants in criminal cases, e.g. a prison sentence. In our context, a juror with $x_i = 1$ is maximally supportive of acquitting the defendant, while a juror with $x_i = -1$ is maximally opposing acquittal (and thus supportive of convicting).

The following theorem characterizes the limiting behavior of system (1) consisting of stubborn positives.

Theorem 4: Suppose that all n susceptibility functions satisfy Assumption 3.

- 1) If $x_i(0) < 1$ for all $i \in [n]$, then all $x_i(t)$ will reach a consensus exponentially fast at some value in the interval $[-1, 1)$.
- 2) If $x_i(0) = 1$ for at least one $i \in [n]$, then all $x_i(t)$ will reach a consensus asymptotically at value 1.

The theorem says that system (1) consisting of stubborn positives will reach a consensus for any initial condition, which has the same limiting behavior as the continuous-time model considered in [25].

B. Stubborn Extremists

Consider now the case in which each individual i has a susceptibility function satisfying the following assumption.

Assumption 4: The susceptibility function $f_i(x_i)$ is a smooth and monotone for x_i in the intervals $[-1, 0]$ and $[0, 1]$, takes values in $[0, 1]$, and equals 0 only when $x_i = \pm 1$.

This assumption says that $f_i(x_i)$ is a non-decreasing function of x_i in $[-1, 0]$, and a non-increasing function in $[0, 1]$. An agent who has such a susceptibility function is said to be a stubborn extremist. In [25], a special case $f_i(x_i) = 1 - x_i^2$ is considered for all $i \in [n]$, and motivation for this special function is also provided.

Motivation: Stubborn extremists become less susceptible to influence as $x_i(t)$ approaches the two extreme ends of the

opinion interval $[-1, 1]$. They are fully open to influence when $x_i(t) = 0$, taking a neutral opinion value. That is, stubborn extremists are individuals whose strength of conviction increases as they become more extreme in their position. Sociology and social psychology literature points to the fact that individuals become more resistant to change of opinion as they become more polarized, or extreme, in their stance [31]–[33]. One can see that this is particularly likely to arise during the discussion of topics for which there are two competing positions, e.g. iPhone versus Android.

The following theorem characterizes some limiting behavior of system (1) with stubborn extremist individuals.

Theorem 5: Suppose that all n susceptibility functions satisfy Assumption 4.

- 1) If $|x_i(0)| \neq 1$ for all $i \in [n]$, then all $x_i(t)$ will reach a consensus exponentially fast at some value in the interval $(-1, 1)$.
- 2) If $x_i(0) \neq -1$ for all $i \in [n]$ and there exists at least one $j \in [n]$ such that $x_j(0) = 1$, then all $x_i(t)$ will asymptotically reach a consensus at 1.
- 3) If $x_i(0) \neq 1$ for all $i \in [n]$ and there exists at least one $j \in [n]$ such that $x_j(0) = -1$, then all $x_i(t)$ will asymptotically reach a consensus at -1 .
- 4) If there exists at least one $i \in [n]$ such that $|x_i(0)| \neq 1$, and there exist $j, k \in [n]$ such that $x_j(0) = 1$ and $x_k(0) = -1$, then there exists a unique¹ nontrivial equilibrium which is unstable. Moreover, the reduced system, only consisting of those individuals whose initial opinions do not equal 1 or -1 , has a unique equilibrium which is locally exponentially stable.

Item 4) of Theorem 5 implies that for any initial condition where there exists $j, k \in [n]$, with $j \neq k$, such that $x_j(0) = 1$ and $x_k(0) = -1$, the unique nontrivial equilibrium is stable with respect to external disturbance only on those individuals whose initial opinions do not equal 1 or -1 . Extensive simulations suggest a conjecture that the system will asymptotically converge to this unique equilibrium for all initial conditions satisfying the hypothesis in item 4).

C. Heterogeneous Networks

In this subsection, we consider networks in which two types of susceptibility functions coexist. For such networks, all possible equilibria are fully categorized by Theorem 1. Here, we establish the limiting behavior of some of the equilibria. In particular, for certain sets of initial conditions, we detail the equilibria that arise and their convergence properties. For other sets of initial conditions, no complete characterization is yet available.

Theorem 6: Suppose that stubborn extremists and positives coexist in system (1).

- 1) If $|x_i(0)| \neq 1$ for all stubborn extremists and $x_j(0) < 1$ for all stubborn positives, then all $x_i(t)$ will reach a consensus exponentially fast at some value in the interval $(-1, 1)$.

¹By unique, we mean that the equilibrium is independent of the initial conditions of all p individuals satisfying $|x_p(0)| \neq 1$.

- 2) If $x_i(0) \neq -1$ for all stubborn extremists and there exists at least one $j \in [n]$ such that $x_j(0) = 1$, then all $x_i(t)$ will asymptotically reach a consensus at 1.
- 3) If $x_j(0) \neq 1$ for all $j \in [n]$ and there exists at least one stubborn extremist such that $x_i(0) = -1$, then all $x_i(t)$ will asymptotically reach a consensus at -1 .
- 4) If there exist $j, k \in [n]$ such that $x_j(0) = 1$ and $|x_k(0)| \neq 1$, and there exists a stubborn extremist such that $x_i(0) = -1$, then there exists a unique nontrivial equilibrium which is unstable. Moreover, the reduced system, only consisting of those stubborn positives, whose initial opinions do not equal 1, and those stubborn extremists, whose initial opinions do not equal 1 or -1 , has a unique equilibrium which is locally exponentially stable.

IV. CONCLUSIONS

In this paper, a discrete-time polar opinion dynamics model with general susceptibility functions has been studied. The set of all equilibria of the system and stability of nontrivial equilibria have been characterized. We have studied two types of susceptibility functions motivated by social examples, allowing heterogeneous susceptibility functions. We have also considered some heterogeneous networks in which two different types of susceptibility functions coexist.

For future work, we seek to study stubborn neutrals, who have reduced susceptibility when their opinions are in the middle of the spectrum, and stubborn conformists, who become more closed to influence when they have an opinion similar to the network average. We also aim to better understand the behavior of the system with time-varying neighbor graphs and when more than two different types of susceptibility functions coexist.

REFERENCES

- [1] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.
- [2] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Department of Electrical Engineering and Computer Science, MIT, 1984.
- [3] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, 2003.
- [4] R. Olfati-Saber and R. M. Murray, "Consensus seeking in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, 2004.
- [5] C. E. Osgood and P. H. Tannenbaum, "The principle of congruity in the prediction of attitude change," *Psychological Review*, vol. 62, no. 1, pp. 42–55, 1955.
- [6] A. G. Miller, J. W. McHoskey, C. M. Bane, and T. G. Dowd, "The attitude polarization phenomenon: role of response measure, attitude extremity, and behavioral consequences of reported attitude change," *Journal of Personality and Social Psychology*, vol. 64, no. 4, p. 561, 1993.
- [7] E. M. Pomerantz, S. Chaiken, and R. S. Tordesillas, "Attitude strength and resistance processes," *Journal of Personality and Social Psychology*, vol. 69, no. 3, pp. 408–419, 1995.
- [8] J. N. Bassili, "Meta-judgmental versus operative indexes of psychological attributes: the case of measures of attitude strength," *Journal of Personality and Social Psychology*, vol. 71, no. 4, p. 637, 1996.
- [9] M. Mäs, A. Flache, and J. A. Kitts, "Cultural Integration and Differentiation in Groups and Organizations," in *Perspectives on Culture and Agent-based Simulations*. Springer, 2014, pp. 71–90.

- [10] P. Duggins, "A Psychologically-Motivated Model of Opinion Change with Applications to American Politics," *Journal of Artificial Societies and Social Simulation*, vol. 20, no. 1, pp. 1–13, 2017.
- [11] W. Xia, M. Cao, and K. H. 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.
- [12] Z. Meng, G. Shi, K. H. Johansson, M. Cao, and Y. Hong, "Behaviors of networks with antagonistic interactions and switching topologies," *Automatica*, vol. 73, pp. 110–116, 2016.
- [13] J. Liu, X. Chen, T. Başar, and M.-A. Belabbas, "Exponential convergence of the discrete- and continuous-time Altafini models," *IEEE Trans. Autom. Control*, vol. 62, no. 12, pp. 6168–6182, 2017.
- [14] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 935–946, 2013.
- [15] A. V. Proskurnikov, A. Matveev, and M. Cao, "Opinion dynamics in social networks with hostile camps: consensus vs. polarization," *IEEE Transactions on Automatic Control*, vol. 61, no. 6, pp. 1524–1536, 2016.
- [16] 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, pp. 1–33, 2002.
- [17] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis, "On Krause's Multi-Agent Consensus Model with State-Dependent Connectivity," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2586–2597, 2009.
- [18] S. R. Etesami and T. Başar, "Game-theoretic analysis of the Hegselmann-Krause model for opinion dynamics in finite dimensions," *IEEE Trans. Autom. Control*, vol. 60, no. 7, pp. 1886–1897, 2015.
- [19] N. E. Friedkin, "The problem of social control and coordination of complex systems in sociology: a look at the community cleavage problem," *IEEE Control Systems Magazine*, vol. 35, no. 3, pp. 40–51, 2015.
- [20] N. E. Friedkin and E. C. Johnsen, *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. Cambridge University Press, 2011, vol. 33.
- [21] 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.
- [22] 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*, 2019. [Online]. Available: <https://doi.org/10.1109/TAC.2018.2805261>
- [23] L. Moreau, "Stability of multi-agent systems with time-dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, no. 2, pp. 169–182, 2005.
- [24] P. Dandekar, A. Goel, and D. T. Lee, "Biased assimilation, homophily, and the dynamics of polarization," *Proceedings of the National Academy of Sciences*, vol. 110, no. 15, pp. 5791–5796, 2013.
- [25] V. Amelkin, F. Bullo, and A. K. Singh, "Polar opinion dynamics in social networks," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 5650–5665, 2017.
- [26] R. Varga, *Matrix Iterative Analysis*. Springer-Verlag, 2000.
- [27] R. C. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.
- [28] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks with stubborn agents: equilibrium and convergence rate," *Automatica*, vol. 50, no. 12, pp. 3209–3215, 2014.
- [29] M. Vidyasagar, *Nonlinear systems analysis*. SIAM, 2002.
- [30] N. L. Waters and V. P. Hans, "A Jury of One: Opinion Formation, Conformity, and Dissent on Juries," *Journal of Empirical Legal Studies*, vol. 6, no. 3, pp. 513–540, 2009.
- [31] C. E. Osgood and P. H. Tannenbaum, "The Principle of Congruity in the Prediction of Attitude Change," *Psychological Review*, vol. 62, no. 1, pp. 42–55, 1955.
- [32] A. G. Miller, J. W. McHoskey, C. M. Bane, and T. G. Dowd, "The attitude polarization phenomenon: role of response measure, attitude extremity, and behavioral consequences of reported attitude change," *Journal of Personality and Social Psychology*, vol. 64, no. 4, p. 561, 1993.
- [33] E. M. Pomerantz, S. Chaiken, and R. S. Tordesillas, "Attitude Strength and Resistance Processes," *Journal of Personality and Social Psychology*, vol. 69, no. 3, pp. 408–419, 1995.