Northumbria Research Link

Consensus and Clustering of Expressed and Private Opinions in Dynamical Networks Against Attacks

Yilun Shang

***Abstract*—A continuous-time opinion dynamics with both expressed and private opinions on a given topic is introduced. An opinion consensus strategy is proposed to achieve resilience consensus against Byzantine attacks in dynamical networks. Necessary and sufficient criteria are established for guaranteeing consensus among normal nodes when the attacks are bounded in each neighborhood of normal nodes. A modification that allows opinion clustering, featuring non-global consensus, is presented. Numerical examples are worked out to illustrate the effectiveness of our theoretical results.**

***Index Terms*—Consensus, clustering, social network, opinion dynamics, expressed opinion, attack.**

I. Introduction

A vast literature has been developed regarding resilient or secure decision making against malicious attacks in social networks, or more generally, networked systems with interacting components [1]–[4]. Across e-commerce platforms, for example, fake online reviews from ostensible customers have caused significant damage on business [5], [6]. Security against cyber attacks has been a key issue in cyber-physical systems [7]. To uphold the performance of network, distributed algorithms have been designed to cope with the compromise of a group of malicious individuals and guarantee consensus of the opinions or states of normal individuals based on nearest-neighbor rules. In many realistic social networks, there is a discrepancy between an individual's private and expressed opinions on a given topic, meaning that single-opinion models are not accurate. For instance, a politician may falsify his or her view to garner votes. A common rationale behind such discrepancy arguably stems from social pressure exerted on an individual to conform, either deliberately or passively, to the group opinion [8]. It is linked to a variety of social phenomena from pluralistic ignorance (where a majority of individuals privately oppose a view but incorrectly assume that most others accept it and hence go along with it) to the spiral of silence [9], [10].

Here, we study a class of opinion dynamics model accommodating both a private opinion which evolves under social influence from the expressed opinions of its neighbors, and an expressed opinion which varies under a pressure to conform to the local environment. We propose a purely distributed opinion consensus strategy and establish sufficient and necessary conditions for consensus against Byzantine attacks, in which each malicious agent may have complete knowledge of the whole network and even collude with other malicious agents posing a serious threat to the group decision making process. Next, we examine a class of resilient opinion clustering problem in dynamical networks, where multiple private and expressed opinions co-exist in the final opinion configuration instead of a common consensus. A couple of numerical examples illustrating our theoretical results are provided.

Related work. There has been some related work on the resilient opinion dynamics in complex social networks. In [11], Sobkowicz introduced emotion effect into the Deffuant opinion model [12] to make the neutral opinions tend to be unstable while the extreme opinions resilient to change. This idea has been largely extended by Amelkin et al. [13] to a general model featuring polar opinion dynamics, where the susceptibility of each individual takes the form of a function of its present opinion. By leveraging cloud computing resources, Alcaraz [14] proposed an opinion consensus based approach to tackle Byzantine faults. Building upon compressive sensing theory, an algorithm identifying the susceptible individuals to stubborn individuals in online social networks is studied in Wai et al. [15]. Moreover, a hybrid resilient consensus model consisting of three types of individuals, i.e., averagers, copiers, and voters is proposed by the author [16]. These work, nevertheless, overlooked the difference between private and expressed opinions.

II. Problem formulation

Let $t \geq 0$ be the time. A directed time-varying network $G(t) = (V, E(t), A(t))$ of n nodes is considered, where $V = \{v_1, \cdots, v_n\}$, an edge $(v_i, v_j) \in E(t)$ if information flows from v_i to v_j at time t, and $A(t) = (a_{ij}(t))$ describes the non-negative adjacency matrix with $a_{ij}(t) = 0$ if $(v_j, v_i) \notin E(t)$. The neighborhood of v_i is denoted by $\mathcal{N}_i(t) = \{v_j \in V : (v_j, v_i) \in E(t)\}$. The node set V is divided into two subsets with $V = N \cup B$, where N represents the normal individuals while B represents the Byzantine attackers. The number and identities of Byzantine nodes are not available to the normal ones in the network, meaning that a normal individual knows neither how many of its neighbors are Byzantine nor whether a neighbor is Byzantine. Byzantine individuals are viewed as the worst case attackers since they may adopt arbitrary opinion update rules and potentially send different information to different neighbors [2], [3]. The dynamics of each normal individual $v_i \in N$ is described by the following continuous-

Y. Shang is with the Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK (e-mail: yilun.shang@northumbria.ac.uk).

Manuscript received November 29, 2018; revised xxxx xx, xxxx. This work was supported in part by the UoA Flexible Fund (No. 201920A1001) from Northumbria University.

time system:

$$\dot{x}_i(t) = u_i(t), \quad t \geq 0 \tag{1}$$
$$\dot{\tilde{x}}_i(t) = \tilde{u}_i(t), \quad t \geq 0 \tag{2}$$

where $x_i(t), \tilde{x}_i(t) \in \mathbb{R}$ represent the private and expressed opinions of v_i at time t, respectively, and $u_i(t)$ and $\tilde{u}_i(t)$ are control inputs to be specified later. The normal individuals in N is said to achieve opinion consensus in the presence of Byzantine ones in B if $\lim_{t\to\infty} x_i(t) - x_j(t) = 0$ and $\lim_{t\to\infty} \tilde{x}_i(t) - \tilde{x}_j(t) = 0$ for all $v_i, v_j \in N$ and all initial conditions $\{x_i(0)\}_{i=1}^n$ and $\{\tilde{x}_i(0)\}_{i=1}^n$. Different from single-opinion models, we here formally require both private and expressed opinions reach a consensus.

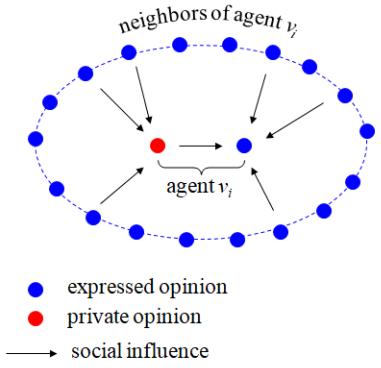

Fig. 1. Schematic illustration of opinion evolution for a normal individual v_i.

Next, we present our opinion consensus strategy for normal individuals as follows (c.f. Fig. 1). Fix an integer r. At time t, each $v_i \in N$ received the expressed opinions $\{\tilde{x}_j^i(t)\}$ of its neighbors, and creates a decreasingly ordered list for $\{\tilde{x}_j^i(t)\}_{v_j \in \mathcal{N}_i(t)}$. (Here, $\tilde{x}_j^i(t) \in \mathbb{R}$ represents the opinion sent from v_j to v_i at time t, and we assume $\tilde{x}_j^i(t) = \tilde{x}_j(t)$ for all $v_j \in N \cap \mathcal{N}_i(t)$, meaning that normal individuals always send their real expressed opinions to their neighbors.) The greatest r values that are greater than $x_i(t)$ are censored by deleting the corresponding incoming edges in $G(t)$ (if there are fewer than r greater values, all of them are erased). Similarly, the smallest values in the list undergo this censor process. We denote by $\mathcal{R}_i(t)$ the set of neighbors erased by v_i at time t. The opinion algorithm for $v_i \in N$ is proposed as:

$$\dot{x}_i(t) = \sum_{v_j \in \mathcal{N}_i(t) \setminus \mathcal{R}_i(t)} a_{ij}(t) f_{ij}(\tilde{x}_j^i(t), x_i(t)) \tag{3}$$

and

$$\dot{\tilde{x}}_i(t) = \lambda_i(x_i(t) - \tilde{x}_i(t)) + (1-\lambda_i) \cdot \sum_{v_j \in \mathcal{N}_i(t) \setminus \mathcal{R}_i(t)} b_{ij}(t) g_{ij}(\tilde{x}_j^i(t), \tilde{x}_i(t)), \tag{4}$$

where $\lambda_i \in [0,1]$, the entries of the adjacency matrix at any time t satisfy $a_{ij}(t) \leq \bar{a}$ for some constant $\bar{a} > 0$, and similarly, $0 \leq b_{ij}(t) \leq \bar{b}$ for some constant $\bar{b} > 0$. Here, Equation (3) describes the evolution of the private opinion of v_i, which is driven by the quantities $f_{ij}(\tilde{x}_j^i(t), x_i(t))$ exerted by its neighbors. For each neighbor $v_j \in \mathcal{N}_i(t) \setminus \mathcal{R}_i(t)$, f_{ij} explains the influence of the expressed opinions of v_j on the private opinion of v_i. We assume:

Assumption 1. The function $f_{ij} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is locally Lipschitz continuous satisfying (i) $f_{ij}(x,y) = 0$ if and only if $x = y$; and (ii) $(x-y) f_{ij}(x,y) > 0$ for $x \neq y$. The similar assumptions hold for g_{ij}.

Remark 1. The parameter λ_i characterizes the resilience to pressure to conform to the local environment encapsulated by g_{ij}. In cooperative control, a typical choice of f_{ij} and g_{ij} in (3) and (4) is $f_{ij}(x,y) = g_{ij}(x,y) = x - y$ [17]. The individual v_i is minimally resilient when $\lambda_i = 0$, meaning that its expressed opinion fully conforms to the group decision, e.g., the average group opinion by taking $b_{ij}(t) = (|\mathcal{N}_i(t)| - |\mathcal{R}_i(t)|)^{-1}$, where $|\cdot|$ represents the cardinality of a set. It is maximally resilient when $\lambda_i = 1$ as its expressed opinion is governed fully by the difference between its own private and expressed opinions. In this situation, the system (3) and (4) shares the same equilibrium, namely, $x_i(t) \to c_i$ as t tends to infinity, with the single-opinion dynamics delineated by $x_i(t) = \tilde{x}_(t)$.

Note that the above opinion consensus strategy is associated with a parameter r, which will be used to bound the number of Byzantine individuals in the neighborhood of any normal individual in the network (see Theorem 1 and Theorem 2 below). As such, an estimate of the upper bound of r suffices in our scheme. Some conservative choices for example can be $r = |B| = n - |N|$ or the maximum degree in the network; c.f. Remark 2 below.

It is also worth mentioning that the network $G(t)$ (and the matrix $A(t)$) is time-dependent and that the censor process further adds to the variation of the network topology. To facilitate the analysis, we assume the following.

Assumption 2. Let $\{\tau_k\}_{k=1}^{\infty}$ denote the time instants that the network topology $G(t)$ changes. There exists a constant $\tau > 0$ such that $\tau_{k+1} - \tau_k \geq \tau$ for any k.

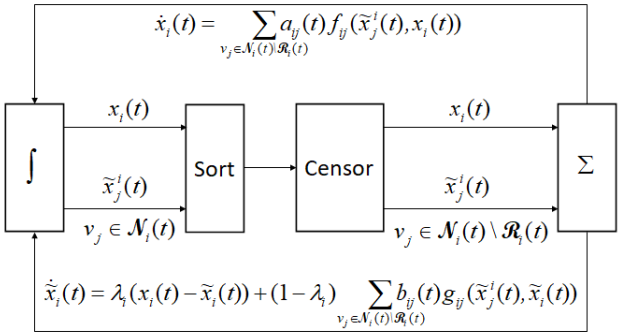

Fig. 2. Synchronous data flow for a normal individual v_i.

In Fig. 2 we present the flow chart of a normal individual $v_i \in N$. Our proposed opinion consensus strategy includes methods like sorting, censoring, and summation, in which the sorting subroutine is the worst part. By using the Quicksort algorithm, which has worst-case time complexity $O(|\mathcal{N}_i|^2)$ and worst-case space complexity $O(|\mathcal{N}_i|)$, our resilient consensus strategy also is worst-case quadratic in time and linear in space. Therefore, the complexity of the algorithm is low.

III. Consensus of expressed and private opinions

We now investigate the consensus evolution of expressed opinions and private opinions when at most r Byzantine individuals exit in each neighborhood of normal individuals. Our strategy is to first show the maxima and minima of opinions are bounded, and then establish the convergence of their difference as time goes on.

Define $M(t) = \max_{v_i \in N} x_i(t)$ and $m(t) = \min_{v_i \in N} x_i(t)$ as the maximum and minimum private opinions, respectively, for normal individuals at time t. Similarly, define $\tilde{M}(t) = \max_{v_i \in N} \tilde{x}_i(t)$ and $\tilde{m}(t) = \min_{v_i \in N} \tilde{x}_i(t)$ as the maximum and minimum expressed opinions, respectively, for normal individuals. Let $M^*(t) = \max\{M(t), \tilde{M}(t)\}$ and $m^*(t) = \min\{m(t), \tilde{m}(t)\}$. The following result shows that $[m^*(0), M^*(0)]$ is an invariant set, meaning that the final opinions for normal individuals remain in this interval, regardless of the network topology.

Theorem 1. Fix an integer r. Suppose each normal individual updates its expressed and private opinions according to the opinion consensus strategy with parameter r. Then for any network $G(t)$ containing at most r Byzantine individuals in the neighborhood of each normal individual, we have $x_i(t) \in [m^*(0), M^*(0)]$ and $\tilde{x}_i(t) \in [m^*(0), M^*(0)]$ for all $v_i \in N$ and all $t \geq 0$.

Proof. We first show $x_i(t), \tilde{x}_i(t) \leq M^*(0)$ for all $t \geq 0$ and argue by contradiction. If this is not the case, there must exist time t_0 such that there is a node $v_i \in N$ satisfying (a) $x_i(t) \leq M^*(0)$ for all $t \leq t_0$; $\tilde{x}_j(t) \leq M^*(0)$ for any $v_j \in N$ and $t \leq t_0$; $x_i(t_0) = M^*(0)$ and $\dot{x}_i(t_0) > 0$; or (b) $\tilde{x}_i(t) \leq M^*(0)$ for all $t \leq t_0$; $x_j(t) \leq M^*(0)$ for any $v_j \in N$ and $t \leq t_0$; $\tilde{x}_i(t_0) = M^*(0)$ and $\dot{\tilde{x}}_i(t_0) > 0$. In other words, either the private opinion of some v_i or the expressed opinion of some v_i firstly attains the upper bound $M^*(0)$.

Case (a). Thanks to (3), we obtain

$$0 < \dot{x}_i(t_0) = \sum_{v_j \in \mathcal{N}_i(t_0) \backslash \mathcal{R}_i(t_0)} a_{ij}(t_0) f_{ij}(\tilde{x}_j^i(t_0), x_i(t_0)). \quad (5)$$

For $v_j \in N \cap (\mathcal{N}_i(t_0) \backslash \mathcal{R}_i(t_0))$, we have $x_i(t_0) = M^*(0) \geq \tilde{x}_j(t_0) = \tilde{x}_j^i(t_0)$ by assumption. For $v_j \in B \cap (\mathcal{N}_i(t_0) \backslash \mathcal{R}_i(t_0))$, $\tilde{x}_j^i(t_0)$ can be written as a convex combination of $\{\tilde{x}_i(t_0)\}_{v_i \in N}$ since the number of Byzantine nodes in the neighborhood of a normal node is no more than r. Hence, $x_i(t_0) \geq \tilde{x}_j^i(t_0)$ still holds. By Assumption 1, $f_{ij}(\tilde{x}_j^i(t_0), x_i(t_0)) \leq 0$, which implies that the righthand side of (5) is non-positive, offering the desired contradiction.

Case (b). It follows from (4) that

$$0 < \dot{\tilde{x}}_i(t_0) = \lambda_i(x_i(t_0) - \tilde{x}_i(t_0)) + (1 - \lambda_i) \cdot \sum_{v_j \in \mathcal{N}_i(t_0) \backslash \mathcal{R}_i(t_0)} b_{ij}(t_0) g_{ij}(\tilde{x}_j^i(t_0), \tilde{x}_i(t_0)). \quad (6)$$

By assumption, we have $x_i(t_0) \leq M^*(0) \leq \tilde{x}_i(t_0)$ and $\lambda_i \in [0, 1]$. Since there are at most r Byzantine nodes in the neighborhood of a normal node, $\tilde{x}_i^i(t_0) \leq \tilde{x}_i(t_0)$ for any $v_j \in \mathcal{N}_i(t_0) \backslash \mathcal{R}_i(t_0)$ by invoking the censoring strategy. Therefore, $g_{ij}(\tilde{x}_j^i(t_0), \tilde{x}_i(t_0)) \leq 0$. The righthand side of (6) is a sum of non-positive terms and hence is non-positive. This leads to a contradiction and concludes the proof of $x_i(t), \tilde{x}_i(t) \leq M^*(0)$ for all $t \geq 0$. The proof for $x_i(t), \tilde{x}_i(t) \geq m^*(0)$ can be done analogously. □

Given $S \subseteq V$, if there exists a node $v_i \in S$ such that $|\mathcal{N}_i \backslash S| \geq r$, we say that S is r-reachable [1]. The idea behind a r-reachable set S is that if there exists one node in S that has sufficiently many neighbors outside of S, then S can be influenced by nodes in $V \backslash S$ when appropriate censoring strategies are applied. For any pair of nonempty, disjoint subsets of V, G is called r-robust if one of these sets is r-reachable. It is shown in [2] that G is 1-robust if and only if it has a directed spanning tree. Therefore, in a robust graph, at least one of any two mutually exclusive sets has good "expansion properties" amenable to the external information, which makes the overall consensus possible.

Define $G_N(t) = (N, E_N(t))$ to be the subgraph of $G(t) = (V, E(t))$ induced by the set of normal nodes, where $E_N(t)$ consists of all directed edges among the normal nodes at time t. Furthermore, define $\Theta(t) = M^*(t) - m^*(t) \geq 0$ for each time $t \geq 0$. With these preparation, we can show our sufficient and necessary criteria for resilient opinion consensus.

Theorem 2. Fix an integer r. Suppose each normal individual updates its expressed and private opinions according to the opinion consensus strategy with parameter r. If $G(t)$ is $(2r+1)$-robust, then opinion consensus for normal individuals can be achieved for $G(t)$ containing at most r Byzantine individuals in the neighborhood of each normal individual. Moreover, a necessary condition for reaching consensus is that $G_N(t)$ is $(r+1)$-robust.

Proof. We first show sufficiency part. Let i_0 and j_0 be indices such that $x_{i_0}^*(t) = \max_{v_i \in N}\{x_i(t), \tilde{x}_i(t)\}$ and $x_{j_0}^*(t) = \min_{v_i \in N}\{x_i(t), \tilde{x}_i(t)\}$, respectively. We take those with maximum derivatives if there are multiple such indices. Capitalizing the property of the Dini derivatives [18], we consider two cases.

(i) The maximum derivative is attained by a private opinion. The Dini derivative of $M^*(t)$, denoted as $D^+M^*(t)$, along the trajectory of (3) is give by

$$D^+M^*(t) = \dot{x}_{i_0}^*(t) = \sum_{v_j \in \mathcal{N}_{i_0}(t) \backslash \mathcal{R}_{i_0}(t)} a_{i_0 j}(t) f_{i_0 j}(\tilde{x}_j^{i_0}(t), x_{i_0}^*(t)), \quad (7)$$

where $D^+M^*(t) = \limsup_{h \to 0^+} \frac{1}{h}(M^*(t+h) - M^*(t))$. For $v_j \in N \cap (\mathcal{N}_{i_0}(t) \backslash \mathcal{R}_{i_0}(t))$, we have $x_{i_0}^*(t) \geq \tilde{x}_j^{i_0}(t)$ by definition. For $v_j \in B \cap (\mathcal{N}_{i_0}(t) \backslash \mathcal{R}_{i_0}(t))$, $x_{i_0}^*(t) \geq \tilde{x}_j^{i_0}(t)$ still holds as there are at most r Byzantine nodes in the neighborhood of a normal node. It follows from Assumption 1 and (7) that $D^+M^*(t) \leq 0$. Similarly, we can show that $D^+m^*(t) \geq 0$.

(ii) The maximum derivative is attained by an expressed opinion. In this case, the Dini derivative of $M^*(t)$ along the trajectory of (4) is give by

$$D^+M^*(t) = \dot{\tilde{x}}_{i_0}^*(t) = \lambda_{i_0}(x_{i_0}(t) - \tilde{x}_{i_0}^*(t)) + (1 - \lambda_{i_0}) \cdot \sum_{v_j \in \mathcal{N}_{i_0}(t) \backslash \mathcal{R}_{i_0}(t)} b_{i_0 j}(t) g_{i_0 j}(\tilde{x}_j^{i_0}(t), \tilde{x}_{i_0}^*(t)). \quad (8)$$

A similar argument shows that the righthand side of (8) is a sum of non-positive terms and hence $D^+M^*(t) \leq 0$. Similarly, we can show that $D^+m^*(t) \geq 0$.

Combining (i) and (ii), we obtain $D^+\Theta(t) = D^+M^*(t) - D^+m^*(t) \leq 0$ for all t. We claim that $\lim_{t\to\infty} D^+\Theta(t) = 0$. If this is true, then $M^*(t) \to \rho_M$ and $m^*(t) \to \rho_m$ for some constants ρ_M and ρ_m. Therefore, $\lim_{t\to\infty} x^*_{i_0}(t) = \rho_M$ and $\lim_{t\to\infty} x^*_{j_0}(t) = \rho_m$ by definition. Since $G(t)$ is $(2r+1)$-robust and our algorithm discard at most $2r$ neighbors in the neighborhood of each normal node, the resulting network is 1-robust. According to the previous comment, there exists a spanning tree in the network at any time. In the light of our algorithm, all normal nodes along the path starting from the root node to v_{i_0} has the private and expressed opinions ρ_M for sufficiently large t. Similarly, all normal nodes along the path starting from the root node to v_{j_0} has the private and expressed opinions ρ_m for large enough t. As the root takes both the maximum and the minimum opinions, we arrive at $\rho_M = \rho_m$. The sufficiency is proved.

It remains to show the claim $\lim_{t\to\infty} D^+\Theta(t) = 0$. Suppose that this does not occur. There must exist some $\varepsilon > 0$ for any $T > 0$ there is $t > T$ such that $D^+\Theta(t) \leq -2\varepsilon$. Hence, there is $\delta > 0$ and a sequence $\{t_l\}_{l=1}^{\infty}$ tending to infinity such that $D^+\Theta(t) \leq -2\varepsilon$ and $|t_{l+1} - t_l| > \delta$ for all l. For any time interval I with $\{\tau_k\}_{k=1}^{\infty} \cap I = \emptyset$, it follows from Assumption 1 that $\dot{x}_i(t), \dot{\tilde{x}}_i(t)$ are bounded for all $v_i \in N$, and hence $D^+\Theta(t)$ is uniformly continuous on I. There exists $\delta_0 > 0$ such that for any $t^{(1)}$ and $t^{(2)}$ satisfying $|t^{(1)} - t^{(2)}| < \delta_0$, $|D^+\Theta(t^{(1)}) - D^+\Theta(t^{(2)})| < \varepsilon$. Consequently, for any $t \in [t_l - \delta_0, t_l + \delta_0]$,

$$\begin{aligned} D^+\Theta(t) &= -|D^+\Theta(t_l) - (D^+\Theta(t_l) - D^+\Theta(t))| \\ &\leq -(|D^+\Theta(t_l)| - |D^+\Theta(t_l) - D^+\Theta(t)|) \\ &\leq -2\varepsilon + \varepsilon = -\varepsilon. \end{aligned} \tag{9}$$

On the other hand, if there exist some $\tau_k \in [t_l - \delta_0, t_l + \delta_0]$, by Assumption 2 there exists $\delta_1 \in (0, \tau)$ such that $D^+\Theta(t) \leq -\varepsilon$ for all $t \in [t_l - \delta_1, t_l + \delta_1]$. Combining the above discussion, we see that there exists some $\delta_2 > 0$ satisfying $\int_0^{\infty} D^+\Theta(t)dt \leq \lim_{N\to\infty} \sum_{l=1}^{N} \int_{t_l-\delta_2}^{t_l+\delta_2} D^+\Theta(t)dt \leq -2\lim_{N\to\infty} N\varepsilon\delta_2 = -\infty$, which contradicts the fact that $\Theta(t)$ is lower bounded, namely, $\Theta(t) \geq 0$ for all t. The claim is then proved.

Finally, we show the necessity. Suppose that $G_N(t)$ is not $r+1$-robust. There exist two nonempty and disjoint sets $S_1, S_2 \subseteq N$ which are not $r+1$ reachable. Every node in these two sets has no more than r normal neighbors not in the set. Fix $\rho_1 < \rho_2$. Let $x_i(0) = \tilde{x}_i(0) = \rho_1$ for all $v_i \in S_1$, and $x_i(0) = \tilde{x}_i(0) = \rho_2$ for all $v_i \in S_2$. For all the other nodes v_i in the network, set $x_i(0) = \tilde{x}_i(0) \in (\rho_1, \rho_2)$. Assume that all Byzantine nodes always send the expressed opinion ρ_1 to each node v_i in S_1, and the expressed opinion ρ_2 to each node v_i in S_2 at all time t. Through out opinion consensus strategy with parameter r, nodes in S_1 and S_2 will never adopt opinions not in their own sets. Accordingly, consensus cannot be achieved among normal nodes in N. The necessity is proved. □

Remark 2. In Theorem 2, the number r is a given parameter, which is assumed to be known for each normal node (some estimates are given in Section II). We here assume that r is known as it is related to the robustness condition of the network; see also [19] for a discrete time system. In practice, it would be possible to obtain an agreed r in a distributed manner via some consensus protocol coupled with the opinion consensus process. For example, if a normal node v_i has a feasible value r_i, it can adopt the max consensus strategy; e.g. [20]. Otherwise, v_i can increase r_i by 1.

Remark 3. It is worth mentioning that in the sufficiency in Theorem 2, we showed that both private and expressed opinions of normal nodes converge to the same consensus value, which is stronger than what is required in our definition of consensus in Section II. Persistence to the individuals' initial opinions has been identified to be a possible cause of non-vanishing discrepancy between expressed and private opinions in [21]. In our framework, such "stubbornness" to the initial opinions has been considered as a malicious behavior and is aimed to be conquered. Hence, we are able to show the strong result of vanishing discrepancy, which interestingly echoes recent data analysis on social networks [22].

Remark 4. Despite the fault tolerance mechanism considered in our proposed protocol, the consensus outcome critically relies on the requirement of normal agents' private and expressed opinions following a predetermined protocol (Eqs. (3) and (4)). In reality, it may be questionable as mass media and systemic bias may impinge on private opinions and distort expressed opinions. This limitation, however, tends to be an inherent part of any form of distributed consensus algorithm. If we for example adopt the proposed algorithm for a cryptocurrency to maintain the transactions of distributed ledger. An intruder can exploit this weakness and can open the doors for double spending problem. As such, our work may fit better in some well-trained environments such as jury deliberation and panel meetings.

IV. Clustering of expressed and private opinions

In this section, we consider the clustering, namely, the coexistence of different opinion clusters for both expressed and private opinions. This can be achieved by generalizing the scaled consensus method [3].

Definition 1. Given scalar number $\alpha_i \neq 0$ for every node $v_i \in V$, we say that the normal nodes in N achieve resilient opinion clustering with respect to $(\alpha_1, \cdots, \alpha_n)$ in the presence of Byzantine nodes in B if $\lim_{t\to\infty} \alpha_i x_i(t) - \alpha_j x_j(t) = 0$ and $\lim_{t\to\infty} \alpha_i \tilde{x}_i(t) - \alpha_j \tilde{x}_j(t) = 0$ for all $v_i, v_j \in N$ and all initial conditions $\{x_i(0)\}_{i=1}^n$ and $\{\tilde{x}_i(0)\}_{i=1}^n$.

From Definition 1, we easily reproduce the resilient opinion consensus defined in Section II by setting $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 1$. In general, we have $x_i/x_j \to \alpha_j/\alpha_i$ and $\tilde{x}_i/\tilde{x}_j \to \alpha_j/\alpha_i$ as t tends to infinity. The concept of opinion clustering is useful in social opinion networks [23]. For instance, an agent may simply reject whatever its competitors support and advocate whatever its competitors oppose in an antagonistic or competitive scenario, which can be appropriately modeled by setting $\alpha_i = 1$ while $\alpha_j = -1$ for a pair of rivals v_i and v_j.

To the end of clustering, the opinion consensus strategy presented in Section II can be modified as follows. Fix an

integer r. At time t, each $v_i \in N$ received the expressed opinions $\{\tilde{x}_j^i(t)\}$ of its neighbors, and creates a decreasingly ordered list for $\{\alpha_j \tilde{x}_j^i(t)\}_{v_j \in \mathcal{N}_i(t)}$. The greatest r values that are greater than $\alpha_i x_i(t)$ are censored by deleting the incoming edges in $G(t)$ (if there are fewer than r such values, all of them are erased). Similarly, the smallest values in the list undergo this censor process. Similarly, the opinion algorithm for $v_i \in N$ is proposed as:

$$\begin{aligned}&\dot{x}_i(t)\\ &= \mathrm{sgn}(\alpha_i) \sum_{v_j \in \mathcal{N}_i(t) \setminus \mathcal{R}_i(t)} a_{ij}(t) f_{ij}(\alpha_j \tilde{x}_j^i(t), \alpha_i x_i(t)) \end{aligned} \quad (10)$$

and

$$\begin{aligned}\dot{\tilde{x}}_i(t) =& \lambda_i(x_i(t) - \tilde{x}_i(t)) + (1-\lambda_i)\\ &\cdot \sum_{v_j \in \mathcal{N}_i(t) \setminus \mathcal{R}_i(t)} b_{ij}(t) g_{ij}\big(\frac{\alpha_j}{\alpha_i} \tilde{x}_j^i(t), \tilde{x}_i(t)\big), \end{aligned} \quad (11)$$

where $\mathrm{sgn}(\cdot)$ is the standard signum function, and all previous assumptions are applied here. The following corollary can be established with similar arguments as in Section III.

Corollary 3. Fix an integer r. Suppose each normal individual updates its expressed and private opinions according to the above opinion clustering strategy with parameter r. If $G(t)$ is $(2r+1)$-robust, then opinion clustering for normal individuals can be achieved for $G(t)$ containing at most r Byzantine individuals in the neighborhood of each normal individual. Moreover, a necessary condition for reaching clustering is that $G_N(t)$ is $(r+1)$-robust.

V. Numerical Examples

In this section, numerical simulations are presented to illustrate our theoretical results.

Example 1. Consider a network G having node set $V = N \cup B$ with normal $N = \{v_1, \cdots, v_5\}$ and Byzantine $B = \{v_6\}$. It is direct to check that G is 3-robust. The initial configuration is chosen as $x_1(0) = \tilde{x}_1(0) = -3$, $x_2(0) = \tilde{x}_2(0) = 5$, $x_3(0) = \tilde{x}_3(0) = -1$, $x_4(0) = \tilde{x}_4(0) = -2$, $x_5(0) = \tilde{x}_5(0) = 2$, $\tilde{x}_6(0) = 3$. The normal nodes follow (3) and (4) with $f_{ij}(x,y) = g_{ij}(x,y) = x - y$, $\lambda = 0.9$, $a_{ij} = b_{ij}$ being binary characterizing the adjacency of the nodes. The Byzantine node follows the dynamics $\dot{\tilde{x}}_6(t) = -\tilde{x}_6(t) - \frac{1}{2}\ln(t+1)$ featuring a monotonic decrease, which potentially drives the group decision towards the minus infinity if not appropriately dealt with. The opinion trajectories are shown in Fig. 3(a) and the discrepancies between private and expressed opinions, denoted by $\Delta_i(t) = |x_i(t) - \tilde{x}_i(t)|$, are shown in Fig. 3(b). Here, $\Delta_i(t)$ measures the discrepancy between expressed and private opinions of each normal node v_i. We observe from Fig. 3(a) that resilient opinion consensus is reached as one would expect despite the intervention of a malicious node v_6. Fig. 3(b) shows that the discrepancies $\Delta_i(t)$ $(i = 1, \cdots, 5)$ start at zero and finally die out leading to consensus in line with our theoretical result of Theorem 2. This further reveals that the difference between expressed and private opinions may arise even when there is no discrepancy between expressed and private opinions initially, which shadows some real-life phenomena.

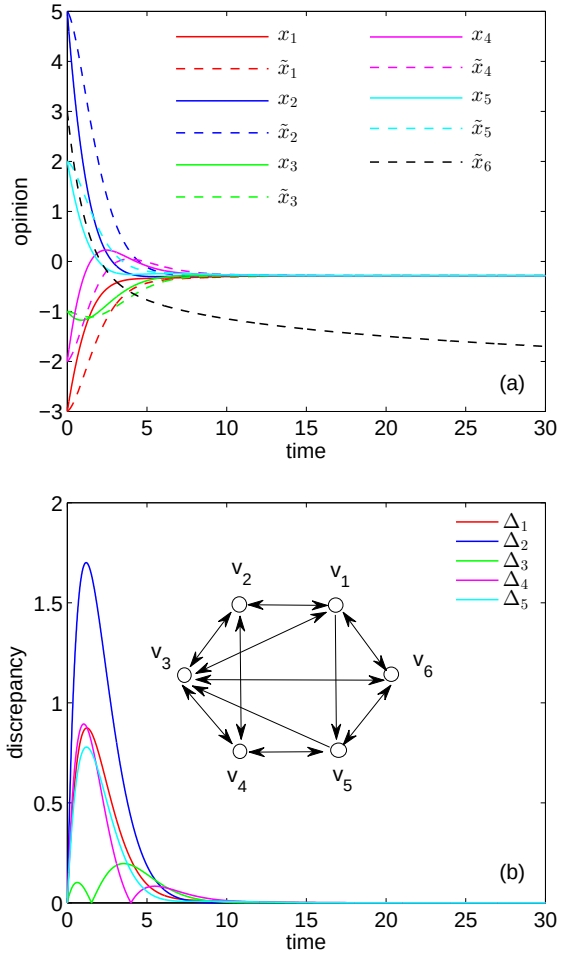

Fig. 3. (a) Opinion consensus trajectories and (b) discrepancy between private and expressed opinions over network G with a Byzantine node v_6 shown in the inset of (b).

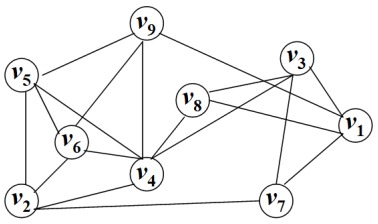

Fig. 4. Rich-core of Zachary's karate club with $N = \{v_2, v_3, \cdots, v_9\}$ and $B = \{v_1\}$.

Example 2. We consider a real-world social network, the Zachary's karate club [24], which describes the relationship between members of a university karate club in 1977. The rich-core of Zachary's karate club is identified in [25]; see Fig. 4. It is a undirected 3-robust graph characterizing the relationship between core members. We assume that the normal node set is $N = \{v_2, v_3, \cdots, v_9\}$ and the Byzantine node set is $B = \{v_1\}$. The initial private and expression opinions of all nodes are randomly taken in the interval $[0, 1]$. The Byzantine node v_1 has its own dynamics $\dot{\tilde{x}}_1(t) = -\tilde{x}_1(t) + \frac{1}{2}\sin(\frac{t}{5})$, which features a periodic fluctuation modeling repeated divergence.

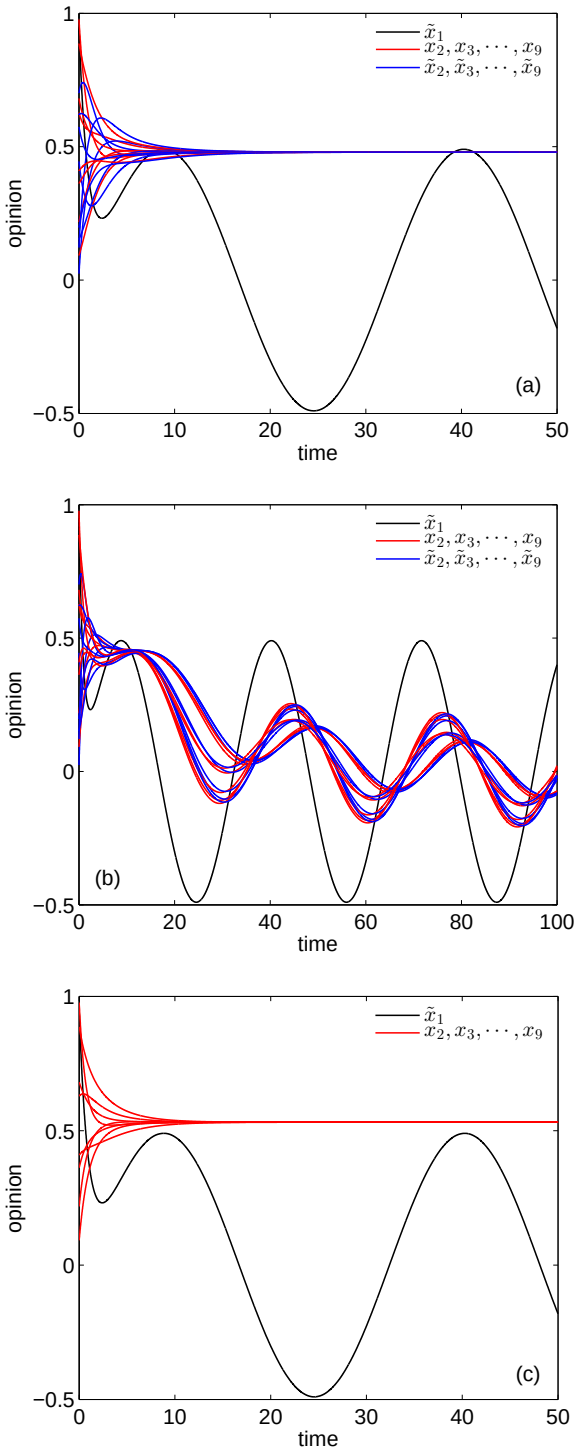

Fig. 5. Expressed and private opinion evolution for Example 2 over network depicted in Fig. 4 with (a) or without (b) the censoring strategy. The same parameters are used in panel (c) as in panel (a) except that the expressed dynamics are scraped. v_1 is the Byzantine node following its own dynamics.

As in Example 1, the Byzantine equation is chosen to illustrate the theoretical results as Byzantine nodes can determine their own strategy.

Recall that the identity of Byzantine v_1 is not known by these normal ones mimicking the real situation. We first compare (a) our censoring strategy, where all normal nodes follow (3), (4) with $f_{ij}(x,y) = g_{ij}(x,y) = x - y$, $\lambda = 0.8$, $a_{ij} = b_{ij} = 1$ if v_i and v_j are connected and 0 otherwise, with (b) a "placebo" strategy which takes the similar algorithm by replacing $\mathcal{N}_i(t)\backslash\mathcal{R}_i(t)$ with $\mathcal{N}_i(t)$ in (3) and (4). In other words, each normal individual will take stock of the situation by sorting its neighbors expressed opinions and removing the maximum and minimum extremes in case (a), while no extreme opinions are censored in case (b). The opinion trajectories for cases (a) and (b) are shown, respectively, in Fig. 5(a) and 5(b). As expected, all opinions converge in Fig. 5(a) in line with Theorem 2. From Fig. 5(b), we observe that with the intervention of the Byzantine node v_1, neither private nor expressed opinions could reach consensus, both of which display periodic-like fluctuations. This highlights the non-triviality of the censoring strategy.

Finally, in Fig. 5(c) we plot the opinion evolution where each normal agent v_i has only a single private opinion [3], [26], namely, $\tilde{x}_j^i(t) \equiv x_j^i(t)$ for $v_j \in \mathcal{N}_i(t)\backslash\mathcal{R}_i(t)$ in (3) and the dynamics (4) are scraped. All other parameters are the same as in case (a). Comparing with Fig. 5(a), we observe that the convergence becomes slightly faster due to the omission of expressed opinions. For example, the time to consensus in Fig. 5(a) is around $t = 20$ and the corresponding time in Fig. 5(c) is around $t = 15$. Also, the final consensus values are different between them as one would expect.

VI. Conclusion

In this paper, we studied consensus and clustering of both expressed and private opinions in directed complex networks against Byzantine individuals. The communication network $G(t)$ is modeled as a time-varying directed graph with positive dwell time. A general purely distributed censoring algorithm is proposed to rule the opinion dynamics of both private opinions and expressed opinions. Necessary and sufficient conditions for resilient consensus and clustering are established based upon the concept of network robustness. In particular, if $G(t)$ is $(2r+1)$-robust, the resilient consensus can be achieved under our proposed dynamics when $G(t)$ contains at most r Byzantine nodes in the neighborhood of each normal node. Analogous result is derived for resilient clustering via a modified opinion consensus strategy. Our results shed light on the discrepancy and evolution between expressed and private opinions in social networks.

In addition to the acquisition of r and correct expression of private opinions mentioned in the work, there are numerous avenues of interesting future research, such as a thorough study of the mechanism of expressed and private opinions, other types of malicious behaviors, and the effect of possible time delays in communication. As this work is theoretical by nature, it would also be appealing to explore cross-validation for the experimental parameters in real-life social networks.

Acknowledgment

The author is grateful to the anonymous reviewers and the editor for their constructive and insightful comments that helped improve the paper significantly.

REFERENCES

[1] H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "Resilient asymptotic consensus in robust networks," *IEEE J. Select. Areas Commun.*, vol. 31, pp. 766–781, 2013.

[2] Y. Wu and X. He, "Secure consensus control for multi-agent systems with attacks and communication delays," *IEEE/CAA J. Autom. Sinica*, vol. 4, pp. 136–142, 2017.

[3] Y. Shang,, "Resilient consensus of switched multi-agent systems," *Syst. Cont. Lett.*, vol. 122, pp. 12–18, 2018.

[4] B. Wu, X. Zhou, Q. Jin, F. Lin, and H, Leung, "Analyzing social roles based on a hierarchical model and data mining for collective decision-making support," *IEEE Syst. J.*, vol. 11, pp. 356–365, 2017.

[5] D. Plotkina, A. Munzel, and J. Pallud, "Illusions of truth-Experimental insights into human and algorithmic detections of fake online reviews," *J. Bus. Res.*, doi:10.1016/j.jbusres.2018.12.009.

[6] N. Kumar, D. Venugopal, L. Qiu, and S. Kumar, "Detecting review manipulation on online platforms with hierarchical supervised learning," *J. Manag. Inf. Syst.*, vol. 35, pp. 350–380, 2018.

[7] V. Venkataramanan, A. Hahn, and A. Srivastava, "CP-SAM: Cyber-physical security assessment metric for monitoring microgrid resiliency," *IEEE Trans. Smart Grid*, doi:10.1109/TSG.2019.2930241.

[8] A. Flache, M. Mäs, T. Feliciani, E. Chattoe-Brown, G. Deffuant, S. Huet, and J. Lorenz, "Models of social influence: towards the next frontiers," *J. Artif. Soc. & Soc. Simul.*, vol. 20, no. 4, pp. 1–31, 2017.

[9] S. V. Grootel, C. V. Laar, L. Meeussen, T. Schmader, and S. Sczesny, "Uncovering pluralistic ignorance to change men's communal self-descriptions, attitudes, and behavioral intentions," *Front. Psychol.*, vol. 9, art. 1344, 2018.

[10] R. Sokoloski, E. M. Markowitz, and D. Bidwell, "Public estimates of support for offshore wind energy: false consensus, pluralistic ignorance, and partisan effects," *Energy Pol.*, vol. 112, pp. 45–55, 2018.

[11] P. Sobkowicz, "Extremism without extremists: Deffuant model with emotions," *Front. Phys.*, vol. 3, art. no. 17, 2015.

[12] C. Antonopoulos and Y. Shang, "Opinion formation in multiplex networks with general initial distributions," *Sci. Rep.*, vol. 8, art. no. 2852, 2018.

[13] V. Amelkin, F. Bullo, and A. K. Singh, "Polar opinion dynamics in social networks," *IEEE Trans. Autom. Contr.*, vol. 62, pp. 5650–5665, 2017.

[14] C. Alcaraz, "Cloud-assisted dynamic resilience for cyber-physical control systems," *IEEE Wireless Commun.*, vol. 25, pp. 76–82, 2018.

[15] H.-T. Wai, A. E. Ozdaglar, and A. Scaglione, "Identifying susceptible agents in time varying opinion dynamics through compressive measurements," *IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Calgary, Canada, pp. 4114–4118, 2018.

[16] Y. Shang, "Hybrid consensus for averager-copier-voter networks with non-rational agents," *Chaos, Solitons & Fractals*, vol. 110, pp. 244–251, 2018.

[17] J. Qin, Q. Ma, Y. Shi, and L. Wang, "Recent advances in consensus of multi-agent systems: a brief survey," *IEEE Trans. Ind. Electron.*, vol. 64, pp. 4972–4983, 2017.

[18] G. D. Shi, K. H. Johansson, and Y. G. Hong, "Reaching an optimal consensus: dynamical systems that compute intersections of convex sets," *IEEE Trans. Autom. Contr.*, vol. 58, pp. 610–622, 2013.

[19] Y. Shang, "Resilient consensus for expressed and private opinions," *IEEE Trans. Cybern.*, doi:10.1109/TCYB.2019.2939929.

[20] R. Lucchese and D. Varagnolo, "Average consensus via max consensus," *IFAC PapersOnLine*, vol. 48, pp. 58–63, 2015.

[21] M. Ye, Y. Qin, A. Govaert, B. D. O. Anderson, and M. Cao, "An influence network model to study discrepancies in expressed and private opinions," *Automatica*, vol. 107, pp. 371–381, 2019.

[22] Y. Zhu, Z. Huang, Z. Wang, L. Luo, and S. Wu, "Influence and extension of the spiral of silence in social networks: a data-driven approach," in *Social Network Based Big Data Analysis and Applications*, Lecture Notes in Social Networks, Springer, Cham, pp. 143–164, 2018.

[23] H. Chen, H. Yin, X. Li, M. Wang, W. Chen, and T. Chen, "People opinion topic model: opinion based user clustering in social networks," *26th Int. Conf. World Wide Web*, Geneva, Switzerland, pp. 1353–1359, 2017.

[24] W. W. Zachary, "An information flow model for conflict and fission in small groups," *J. Anthropol. Res.*, vol. 33, 452–473, 1977.

[25] A. Ma and R. J. Mondragón, "Rich-cores in networks," *PLoS ONE*, vol. 10, art. no. e0119678, 2015.

[26] Y. Shang, "Resilient multiscale coordination control against adversarial nodes," *Energies*, vol. 11, art. no. 1844, 2018.

Yilun Shang received B.S. and Ph.D. degrees in mathematics from Shanghai Jiao Tong University, in 2005 and 2010, respectively. He was a Postdoctoral Fellow with University of Texas at San Antonio, Singapore University of Technology and Design, and Hebrew University of Jerusalem from 2010 to 2014. From 2014 to 2018, he was an Associate Professor with Tongji University. He was an short-stay International Visiting Fellow with University of Essex in 2017. Since 2018, he has been an Associate Professor with Northumbria University. He has been an Associate Editor with Cell Press since 2019.

Dr. Shang's research interests include complex systems, applied probability, algebra, combinatorics, network science, data science and algorithms. He was awarded the Pujiang Talent Program in 2015 by Science and Technology Commission of Shanghai Municipality. He was the recipient of 2016 Dimitrie Pompeiu Prize and received the Open Arms Grant from ICM2018. He is an academic editor for some international journals including Scientific Reports, IEEE Access, Frontiers in Physics, and European Journal of Pure and Applied Mathematics.