Éminence Grise Coalitions: On the Shaping of Public Opinion

Sadegh Bolouki, *Member, IEEE*, Roland P. Malhamé, *Member, IEEE*, Milad Siami, *Student Member, IEEE*, and Nader Motee, *Senior Member, IEEE*

***Abstract*—We consider an opinion network of multiple individuals with dynamics evolving via a general time-varying continuous time consensus algorithm. In such a network, a subset of individuals forms an *éminence grise coalition* (EGC) if the individuals in that subset are capable of leading the entire network to agreeing on any desired opinion through a cooperative choice of their own initial opinions. In this endeavor, the coalition members are assumed to have access to full profile of the coupling graph of the network as well as the initial opinions of all other individuals. We establish the existence of a minimum size EGC and develop a nontrivial set of tight upper and lower bounds on that size. Thus, even when the coupling graph does not guarantee convergence to a global or multiple consensus, a generally restricted coalition of individuals can steer public opinion toward a desired consensus, provided they can cooperatively adjust their own initial opinions. Geometric insights into the structure of EGCs are also given.**

***Index Terms*—Eminence grise, ergodicity, networked control systems, opinion dynamics, rank of stochastic chains.**

I. Introduction

Distributed averaging algorithms have been widely used in the past few decades to describe dynamics of a network. There has been a growing interest from various research communities as to whether a global agreement, also known as consensus, is asymptotically achieved within the network. In biology, the notion of consensus arises when investigating the emergent behavior of bird flocks, fish schools, etc. [1], [2]. Applications in robotics and control relate to the coordination and cooperation of mobile agents and sensors [3], [4]. In sociology, averaging dynamics can shed light on the dynamics of opinion formation [5]. Consensus algorithms have also been studied within the computer science [6] and management science communities [7].

A class of distributed averaging algorithms is characterized in general by an *exogenous* coupling chain of opinion update matrices, which behave like intensity matrices of a Markov chain. If the Markov chain, which couples the update dynamics, is *ergodic*, then a global agreement is guaranteed irrespective of the initial opinions [8]. Classifying ergodic chains or its equivalent problems have been the subject of a large body of literature of which, to the best of the authors' knowledge, [9]–[11] appear to provide the largest class of ergodic chains.

In contrast to the ergodicity problem, which is concerned with an "unconditional" consensus, this work is motivated by the following questions: What if the coupling chain is not ergodic? How can a global agreement be achieved in a network with absolutely no assumption on the coupling chain? In other words, for a network with a general time-varying coupling chain, what can be said about particular (nontrivial) choices of initial opinions leading to a possible global agreement? Geometric insights on the nature of the "march" toward agreement allow one to realize that such choices of initial opinion vectors form a vector space, the dimension of which is related to the characteristics of the coupling chain. The fact that such initial opinion vectors form a vector space suggests the existence of a possibly small subgroup of individuals in the network who are capable of steering the whole group to eventually agree on any desired value *only* by collectively adjusting their own initial opinions.

A. Related Work

On the subject of steering the public opinion, a remarkable work has been performed by Lorenz and Urbig [12], where the bounded confidence model introduced in [13] is considered. For relatively small confidence bounds, the authors investigated the possibility of enforcing consensus via modification of coupling weights. Furthermore, *controllability* of a linear system via a preferably small subset of agents, referred to as the *leaders*, has been widely discussed in the literature, for example, [14] and [15], which addressed the time-invariant case. A general framework addressing controllability was introduced by Liu *et al.* [16]. For networks with a fixed topology, Monshizadeh *et al.* [17] characterize controllable sets of the network via the so-called zero forcing sets. A generalization to networks with switching topologies was carried out in [18].

B. Contributions

This paper deals with networks of individuals with scalar opinions whose dynamics are described by a distributed averaging algorithm in continuous time. An *Éminence Grise Coalition* (EGC) of a network is defined as a subset of individuals who can, for any arbitrary value $x^* \in \mathbb{R}$ and any distribution of the rest

Manuscript received September 24, 2014; revised April 20, 2015; accepted September 10, 2015. Date of publication September 25, 2015; date of current version June 16, 2017. This work was supported by NSF-ECCS-1202517, AFOSR-YIP FA9550-13-1-0158, ONR N00014-13-1-0636, and NSERC 6820-2011. Recommended by Associate Editor F. Fagnani.

S. Bolouki, M. Siami, and N. Motee are with the Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, PA 18015 USA (e-mail: bolouki@lehigh.edu; siami@lehigh.edu; motee@lehigh.edu).

R. P. Malhamé is with GERAD and the Department of Electrical Engineering, Polytechnique Montréal, Montreal, QC H3T 1J4 Canada (e-mail: roland.malhame@polymtl.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCNS.2015.2482218

of the individuals' initial opinions, set their own initial opinions such that a global agreement on x^* is asymptotically achieved. While it is trivial to establish the existence of at least one largest EGC, namely, the universal coalition of individuals, the main point of interest in this work is to characterize the size of a minimal EGC. Contributions of this work are listed as follows.

The notion of *rank* is defined for the coupling chain of a general network and shown to be equal to the size of a minimal EGC of the network. Therefore, the size of a minimal EGC of a network shall be referred to as the rank of its coupling chain. A geometric framework is developed to interpret the rank of the coupling chain and immediately results in a tight upper bound for the rank (Section III). Futhermore, it is shown that the ranks of two coupling chains are identical if the two chains are l_1-approximations of each other (Section IV). Tight lower bounds on the rank are established based on the existing notions in the literature, namely, the so-called infinite flow graph and the unbounded interactions graph of the chain (Section V). Employing the sharp bounds obtained on the rank, we determine its exact value for time-invariant chains (Section VI) as well as a large class of time-varying chains, the so-called Class $\mathcal{P}^*$ (Section VII). Finally, *full-rank* chains, namely, chains with rank equal to the network size, are precisely characterized, leading to another upper bound on the rank (Section VIII).

II. Notions and Terminology

Consider N individuals in a network and let $\mathcal{V} = \{1, \ldots, N\}$ be the set of individuals. Assume that t stands for the continuous time index. Let a time-varying chain $\{A(t)\}_{t\geq 0}$ of square real-valued matrices of size N satisfy the following two assumptions.

Assumption 1: Each matrix $A(t)$, $t \geq 0$, has zero row sum and non-negative offdiagonal elements.

Assumption 2: Each element $a_{ij}(t)$ of $A(t)$, $i, j \in \mathcal{V}$, $t \geq 0$ is measurable and uniformly bounded.

All continuous time chains considered in this paper are assumed to satisfy Assumptions 1-2. Let the dynamics of the network be described by the following distributed algorithm:

$$\dot{x}(t) = A(t)x(t), \qquad t \geq 0 \tag{1}$$

where $x(t) \in \mathbb{R}^N$ is the vector of opinions. Thus, $x_i(t)$ is the scalar opinion of individual i at time t. Chain $\{A(t)\}$ is referred to as the *coupling chain* of the network.

Assumption 2 implies that for an arbitrary initial condition, (1) has a unique solution in the Caratheodory sense.[1] Furthermore, there exists a unique[1] *state transition matrix* $\Phi(t, \tau)$, $t \geq \tau \geq 0$, associated with chain $\{A(t)\}$ for which

$$x(t) = \Phi(t, \tau)x(\tau) \qquad \forall\, t \geq \tau \geq 0. \tag{2}$$

Note that the state transition matrix, which is always invertible,[1] is the unique solution of

$$\Phi(t, \tau) = I + \int_{\tau}^{t} A(t')\Phi(t', \tau)\, dt', \qquad t \geq \tau \geq 0 \tag{3}$$

where I is the identity matrix.

[1]For instance, see [19, Subsections II.4-III.1].

We use the following notation throughout this paper: $\Phi_i(t, \tau)$ and $\Phi_{i,j}(t, \tau)$, $1 \leq i$, $j \leq N$, denote the ith column and the (i, j)th element of $\Phi(t, \tau)$, respectively. Moreover, the transposition of a matrix is indicated by the matrix followed by prime $(')$. We emphasize that $\Phi'_i(t, \tau)$ refers to the ith column of $\Phi'(t, \tau)$ (prime acts first). For an arbitrary vector $v \in \mathbb{R}^N$, and $1 \leq i \leq N$, v_i denotes the ith element of v. The vectors (and matrices) of all zeros and all ones of any size are denoted by $\mathbf{0}$ and $\mathbf{1}$, respectively. For an arbitrary subset $\mathcal{S} \subset \mathcal{V}$, $\mathcal{V} \setminus \mathcal{S}$ denotes the complement of $\mathcal{S}$ in $\mathcal{V}$.

Remark 1: According to Assumption 1, $A(t)$ can be viewed as the evolution of the intensity matrix of a time inhomogeneous Markov chain. Consequently, $\Phi_{i,j}(t, \tau)$, $t \geq \tau$, for a fixed τ can be viewed as a transition probability in a backward propagating inhomogeneous Markov chain. In particular, for every $t_2 \geq t_1 \geq \tau \geq 0$, we have

$$\Phi_{i,j}(t_2, \tau) = \sum_k \Phi_{i,k}(t_2, t_1)\Phi_{k,j}(t_1, \tau)$$

with the following conditions:

$$\begin{aligned} \Phi_{i,j}(t, \tau) &\geq 0 \\ \sum_{j\in\mathcal{V}} \Phi_{i,j}(t, \tau) &= 1 \\ \Phi_{i,j}(\tau, \tau) &= \delta_{ij} \end{aligned}$$

where δ_{ij} is the Kronecker symbol.

A. Éminence Grise Coalition

Definition 1 (EGC): For an opinion network with dynamics (1), a subgroup of individuals $\mathcal{S} \subset \mathcal{V}$ is said to be an *Éminence Grise Coalition* (EGC), if for any arbitrary $x^* \in \mathbb{R}$ and any initialization of opinions of individuals in $\mathcal{V} \setminus \mathcal{S}$, there exists an initialization of opinions of individuals in $\mathcal{S}$ such that $\lim_{t\to\infty} x(t) = x^*.\mathbf{1}$.

From another point of view, an EGC of a network is a subgroup of individuals who are capable of leading the whole group toward a global agreement on any desired ultimate opinion only by properly initializing their own opinions, with the assumption that they are cognizant of the coupling chain of the network and initial opinions of the remaining individuals.

Example 1: Consider a network with dynamics (1) whose coupling chain is defined by

$$A(t) = \begin{cases} \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} & \text{if } t \in [2k-2, 2k-1),\ k \in \mathbb{N} \\ \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} & \text{if } t \in [2k-1, 2k),\ k \in \mathbb{N}. \end{cases}$$

Note that Assumptions 1-2 are satisfied. Taking advantage of the periodicity of $\{A(t)\}$ and calculating matrix exponentials, one derives

$$\lim_{t\to\infty} \Phi(t, 0) = \mathbf{1}. \left[\frac{1}{e+1}\ 0\ \frac{e}{e+1}\right] \tag{4}$$

where e is the Euler's number. We show that $\{1\}$ is an EGC of the network. Let x^* be the desired value of agreement and

$x_i(0)$, $i \in \{2,3\}$ be the arbitrary initial opinion of individual i. If individual 1 sets: $x_1(0) = (e+1)x^* - ex_3(0)$, taking (2) and (4) into account, an agreement on x^* would asymptotically be established. Similarly, $\{3\}$ is an EGC of the network as well. However, $\{2\}$ is not an EGC. Finally, any subset of individuals with a size greater than 1 is also an EGC since it includes at least one of the individuals 1 and 3.

Lemma 1: In an opinion network with dynamics (1), a subset $\mathcal{S} \subset \mathcal{V}$ is an EGC if and only if for any initialization of opinions of individuals in $\mathcal{V} \setminus \mathcal{S}$, there exists an initialization of opinions of individuals in $\mathcal{S}$ such that $\lim_{t\to\infty} x(t) = \mathbf{0}$.

Proof: The *only if* part is obvious by setting $x^* = 0$ in Definition 1. Conversely, assume for any initialization of individuals in $\mathcal{V} \setminus \mathcal{S}$, there exists an initialization of individuals in $\mathcal{S}$ such that all opinions asymptotically converge to zero. We show in the following that $\mathcal{S}$ is an EGC. Let $x^* \in \mathbb{R}$ is the arbitrary desired value of agreement, and $x_i(0) = \hat{x}_i$, $i \in \mathcal{V} \setminus \mathcal{S}$ are arbitrary but fixed. For a moment, let us assume that the opinion of each individual i, $i \in \mathcal{V} \setminus \mathcal{S}$ is initialized at $\hat{x}_i - x^*$. By the assumption $x_i(0)$, $i \in \mathcal{S}$, could be set, say on $\hat{x}_i$, in such a way that all opinions would asymptotically converge to zero. In other words, the following initialization:

$$x_i(0) = \begin{cases} \hat{x}_i - x^*, & \text{if } i \in \mathcal{V} \setminus \mathcal{S} \\ \hat{x}_i, & \text{if } i \in \mathcal{S} \end{cases}$$

would result in $\lim_{t\to\infty} x(t) = \mathbf{0}$. Now, the following initialization, which is essentially a translation of the previous initialization by x^*:

$$x_i(0) = \begin{cases} \hat{x}_i, & \text{if } i \in \mathcal{V} \setminus \mathcal{S} \\ \hat{x}_i + x^*, & \text{if } i \in \mathcal{S} \end{cases}$$

will lead to an agreement on x^* noticing that translations are preserved in consensus dynamics (1). ■

Our primary objective in this work is characterizing the size of a minimal EGC.

B. Rank of a Chain

In this subsection, we define several operators for chains of matrices. **Bold** style is used for chain operators to distinguish them from matrix operators. Let $\{A(t)\}$ be a chain of matrices and $\Phi(t,\tau)$, $t \geq \tau \geq 0$ be its associated state transition matrix.

Definition 2 (Null Space): The *null space* of chain $\{A(t)\}$ at time $\tau \geq 0$, denoted by $\mathbf{null}_\tau(A)$, is defined by

$$\mathbf{null}_\tau(A) \triangleq \left\{ v \in \mathbb{R}^N \,|\, \lim_{t\to\infty} (\Phi(t,\tau)v) = \mathbf{0} \right\}. \quad (5)$$

One can easily show that $\mathbf{null}_\tau(A)$ is a subspace of $\mathbb{R}^N$ for any $\tau \geq 0$.

Lemma 2: The dimension of subspace $\mathbf{null}_\tau(A)$ is independent of τ.

Proof: Let $\tau_2 > \tau_1 \geq 0$ be two arbitrary time instants. Define linear operator $\phi_{\tau_2,\tau_1} : \mathbb{R}^N \to \mathbb{R}^N$ by

$$\phi_{\tau_2,\tau_1}(v) \triangleq \Phi(\tau_2,\tau_1)v, \ \forall v \in \mathbb{R}^N. \quad (6)$$

Since $\Phi(\tau_2,\tau_1)$ is invertible, the operator ϕ_{τ_2,τ_1} creates a one-to-one correspondence between the two vector spaces $\mathbf{null}_{\tau_1}(A)$ and $\mathbf{null}_{\tau_2}(A)$. Thus, the two vector spaces are of equal dimensions. ■

Definition 3 (Nullity and Rank): The dimension of $\mathbf{null}_\tau(A)$, $\tau \geq 0$, which is independent of τ, is called the *nullity* of chain $\{A(t)\}$ and is denoted by $\mathbf{nullity}(A)$. Moreover, the *rank* of chain $\{A(t)\}$ is defined by

$$\mathbf{rank}(A) \triangleq N - \mathbf{nullity}(A). \quad (7)$$

To further clarify the notions, we recall the chain of Example 1. For periodic chain $\{A(t)\}$ of Example 1, using (4), one concludes $\mathbf{null}_0(A) = \{(\alpha, \beta, -\alpha/e) \,|\, \alpha, \beta \in \mathbb{R}\}$, which is a subspace of dimension 2. Therefore, $\mathbf{nullity}(A) = 2$ and, consequently, $\mathbf{rank}(A) = 1$. Thus, for the network of Example 1, $\mathbf{rank}(A)$ is equal to the size of a minimal EGC. We state in the following theorem, that this is, in fact, true in general.

Theorem 1: The size of a minimal EGC in a network with dynamics (1) is $\mathbf{rank}(A)$.

Proof: Let $r \triangleq \mathbf{rank}(A)$ and h be the size of a minimal EGC. It is shown in the following that $h \leq r$ and $r \leq h$.

($h \leq r$): We show that there is an EGC of size r. From Lemma 1, it is sufficient to show that there exists a subset $\mathcal{S} \subset \mathcal{V}$ of size r with the property that for any initialization of the opinions of individuals in $\mathcal{V} \setminus \mathcal{S}$, there exists an initialization of the opinions of individuals in $\mathcal{S}$ such that all opinions asymptotically converge to 0. Note that $\mathbf{null}_0(A)$ is a vector space with dimension $\mathbf{nullity}(A) = N - r$. Let $\{\beta_1, \ldots, \beta_{N-r}\}$ be a basis of $\mathbf{null}_0(A)$. Notice that the column-rank of matrix $[\beta_1 \mid \cdots \mid \beta_{N-r}]$ is $N-r$, and so is its row-rank. Thus, it has $N-r$ linearly independent rows. Assume that $i_1, \ldots, i_{N-r}$ are the indices of $N-r$, those independent rows, and $\alpha_{i_1}, \ldots, \alpha_{i_{N-r}}$ denote those rows, respectively. We wish to show that $\mathcal{S}$ defined by

$$\mathcal{S} = \mathcal{V} \setminus \{i_1, \ldots, i_{N-r}\}$$

forms an EGC. Let

$$\left[x_{i_1}(0) \ \ldots \ x_{i_{N-r}}(0)\right]'$$

be an *arbitrary* vector representing initial opinions of individuals $i_1, \ldots, i_{N-r}$. Define $v \in \mathbb{R}^N$, to *potentially* represent the initial opinion vector of all individuals, by

$$v \triangleq [\beta_1 \mid \cdots \mid \beta_{N-r}] \begin{bmatrix} \alpha_{i_1} \\ \vdots \\ \alpha_{i_{N-r}} \end{bmatrix}^{-1} \begin{bmatrix} x_{i_1}(0) \\ \vdots \\ x_{i_{N-r}}(0) \end{bmatrix}. \quad (8)$$

Note that the inverse matrix in (8) exists since $\alpha_{i_1}, \ldots, \alpha_{i_{N-1}}$ were assumed linearly independent. Since $\{\beta_1, \ldots, \beta_{N-r}\}$ is a basis of $\mathbf{null}_0(A)$, we have $v \in \mathbf{null}_0(A)$, meaning that if v were the initial opinion vector, all opinions would eventually converge to 0. On the other hand, $v_{i_s} = x_{i_s}(0)$ for every s, $1 \leq s \leq N-r$, since

$$\begin{bmatrix} v_{i_1} \\ \vdots \\ v_{i_{N-r}} \end{bmatrix} = \begin{bmatrix} \alpha_{i_1} \\ \vdots \\ \alpha_{i_{N-r}} \end{bmatrix} \begin{bmatrix} \alpha_{i_1} \\ \vdots \\ \alpha_{i_{N-r}} \end{bmatrix}^{-1} \begin{bmatrix} x_{i_1}(0) \\ \vdots \\ x_{i_{N-r}}(0) \end{bmatrix} = \begin{bmatrix} x_{i_1}(0) \\ \vdots \\ x_{i_{N-r}}(0) \end{bmatrix}.$$

Thus, this part of the proof is now complete since $x_{i_s}(0)$, $1 \leq s \leq N-r$, was chosen arbitrarily.

$(r \leq h)$: Since there is an EGC of size h, there are $N - h$ individuals such that for any values of their initial opinions, there is an initial opinion vector which results in the asymptotic convergence of all opinions to 0 or, equivalently, an initial opinion vector that belongs to $\mathbf{null}_0(A)$. Thus, the dimension of the vector space $\mathbf{null}_0(A)$ is greater than or equal to $N - h$, that is, $N - r \geq N - h$. ■

Remark 2: Another point of interest regarding the issue of consensus that we will not further discuss in this work is that of the nature of the set of initial opinion vectors leading to consensus in the network with dynamics (1); more precisely

$$\left\{x(0) | \exists x^* \in \mathbb{R} : \lim_{t\to\infty} x(t) = x^*.\mathbf{1}\right\}. \tag{9}$$

We note that set (9) is a subspace of dimension $\mathbf{nullity}(A) + 1$ since it is the subspace spanned by $\mathbf{null}_0(A)$ and $\mathbf{1}$.

Keeping Theorem 1 in mind, we focus on the notion of rank in the rest of this paper.

C. Ergodicity and Class-Ergodicity

Definition 4 (Ergodicity): Chain $\{A(t)\}$ is said to be *ergodic* if for every $\tau \geq 0$, its associated state transition matrix $\Phi(t, \tau)$ converges to a matrix with equal rows as $t \to \infty$.

From [8], the ergodicity of $\{A(t)\}$ is equivalent to the occurrence of consensus in (1) irrespective of the initial conditions.

Definition 5 (Class-Ergodicity): Chain $\{A(t)\}$ is *class-ergodic* if for any $\tau \geq 0$, $\lim_{t\to\infty} \Phi(t, \tau)$ exists but has possibly distinct rows.

We note that chain $\{A(t)\}$ is class-ergodic if and only if multiple consensus occurs in (1) irrespective of the values of the initial opinions. (See [20] and [21].) According to [22], two individuals $i, j \in \mathcal{V}$ are said to be *mutually weakly ergodic* if for every $\tau \geq 0$

$$\lim_{t\to\infty} \left\| \Phi'_i(t, \tau) - \Phi'_j(t, \tau) \right\| = 0. \tag{10}$$

The relation of being mutually weakly ergodic is an equivalence relation on $\mathcal{V}$. The equivalence classes of this relation are referred to as *ergodicity classes* in this paper. Indeed, these equivalence classes form a partitioning of $\mathcal{V}$ and while, in some cases, they may simply be singletons, they can always be defined for any $\{A(t)\}$. If chain $\{A(t)\}$ is class-ergodic, that is, $\lim_{t\to\infty} \Phi'_i(t, \tau)$ exists for every $i \in \mathcal{V}$ and $\tau \geq 0$, then $i, j \in \mathcal{V}$ are in the same ergodicity class if $\lim_{t\to\infty} \Phi'_i(t, \tau) = \lim_{t\to\infty} \Phi'_j(t, \tau)$, for every $\tau \geq 0$. We refer to the ergodicity classes of a class-ergodic chain as *ergodic classes*.

III. Geometric Interpretation of the Rank

In this section, inspired by [23], we employ a geometric approach to analyze the asymptotic properties of a chain of matrices. This approach helps to geometrically interpret the rank of a time-varying chain and identify an upper bound for it.

For chain $\{A(t)\}$, define $C_{t,\tau}$, $t \geq \tau \geq 0$ as the convex hull of points in $\mathbb{R}^N$, corresponding to the columns of the transpose of associated state transition matrix $\Phi(t, \tau)$, i.e.,

$$C_{t,\tau} \triangleq \left\{ \sum_{i=1}^{N} \alpha_i \Phi'_i(t, \tau) \,\middle|\, \alpha_i \geq 0 \forall i \in \mathcal{V} \text{ and } \sum_{i=1}^{N} \alpha_i = 1 \right\}.$$

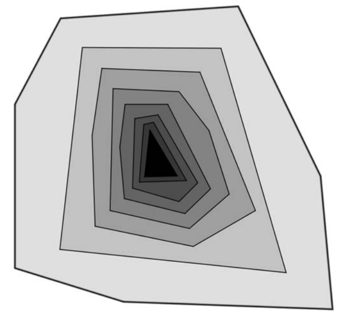


Fig. 1. Example of nested polygons converging to a triangle.

Note that since $\Phi(t, \tau)$ is invertible, $\Phi'_i(t, \tau)$'s, $i \in \mathcal{V}$, are linearly independent and none of them lies in the convex hull of the rest. Thus, $\Phi'_i(t, \tau)$'s form the N vertices of polytope $C_{t,\tau}$. From [23, Prop. 5.1], we know that for every $t_2 \geq t_1 \geq \tau$, we have $C_{t_2,\tau} \subset C_{t_1,\tau}$. It means that polytopes $C_{t,\tau}$, for an arbitrary fixed τ, form a decreasing sequence of polytopes in $\mathbb{R}^N$. An example of these nested polytopes projected on a 2-D subspace of $\mathbb{R}^N$ is depicted in Fig. 1.

For any $\tau \geq 0$, $\lim_{t\to\infty} C_{t,\tau}$ exists and is a polytope in $\mathbb{R}^N$ due to the existence of a uniform upper bound, namely, N on the number of vertices of the nested polytopes. Let C_τ denote the limiting polytope and c_τ be the number of its vertices.

Lemma 3: c_τ, $\tau \geq 0$, is independent of τ.

Proof: It suffices to show that $c_{\tau_1} = c_{\tau_2}$ for any two arbitrary time instants $\tau_2 \geq \tau_1 \geq 0$. We first claim that $c_{\tau_1} \leq c_{\tau_2}$. To prove our claim, we define linear operator $\phi'_{\tau_2,\tau_1} : \mathbb{R}^N \to \mathbb{R}^N$ by

$$\phi'_{\tau_2,\tau_1}(v) \triangleq \Phi'(\tau_2, \tau_1)v, \ \forall v \in \mathbb{R}^N. \tag{11}$$

As a property of the state transition matrix, for any $t \geq \tau_2$

$$\Phi'(t, \tau_1) = \Phi'(\tau_2, \tau_1)\Phi'(t, \tau_2). \tag{12}$$

Our claim $c_{\tau_1} \leq c_{\tau_2}$ shall be proved by taking the three following steps. From (11) and (12), we first show that for any $t \geq \tau_2$, C_{t,τ_1} is the image of C_{t,τ_2} under ϕ'_{τ_2,τ_1}. Then, we obtain that C_{τ_1} is the image of C_{τ_2} under ϕ'_{τ_2,τ_1}. Finally, we conclude our claim.

Any vector in convex hull C_{t,τ_2} can be written as a convex combination of the vertices of C_{t,τ_2} each of which is a column of $\Phi'(t, \tau_2)$. Therefore, for any $v \in C_{t,\tau_2}$, there exists a stochastic vector $u \in \mathbb{R}^N$ such that $v = \Phi'(t, \tau_2)u$. From (12), we now have

$$\begin{aligned}\phi'_{\tau_2,\tau_1}(v) &= \Phi'(\tau_2, \tau_1)v = \Phi'(\tau_2, \tau_1)\Phi'(t, \tau_2) \\ &= \Phi'(t, \tau_1)u \in C_{t,\tau_1}.\end{aligned}$$

Thus, the image of any vector in C_{t,τ_2} under ϕ'_{τ_2,τ_1} lies in C_{t,τ_1}. Similarly, for any $\bar{v} \in C_{t,\tau_1}$, there exists a stochastic vector $\bar{u}$ such that $\bar{v} = \Phi'(t, \tau_1)\bar{u}$. Using (12), we have

$$\bar{v} = \Phi'(t, \tau_1)\bar{u} = \Phi'(\tau_2, \tau_1)\Phi'(t, \tau_2)\bar{u} = \phi'_{\tau_2,\tau_1}\left(\Phi'(t, \tau_2)\bar{u}\right)$$

where $\Phi'(t, \tau_2)\bar{u} \in C_{t,\tau_2}$ since $\bar{u}$ is stochastic. Thus, any vector in C_{t,τ_1} is the image of a vector in C_{t,τ_2} under ϕ'_{τ_2,τ_1}. Hence, C_{t,τ_2} is mapped to C_{t,τ_1} under ϕ'_{τ_2,τ_1}. Now, by taking t to infinity, from the continuity of linear operator ϕ'_{τ_2,τ_1}, we conclude that C_{τ_2} is mapped to C_{τ_1} under ϕ'_{τ_2,τ_1}. Assume now that $v_1, \ldots, v_{c_{\tau_2}}$ are the c_{τ_2} vertices of C_{τ_2}. Since C_{τ_1} is the image of C_{τ_2} under ϕ'_{τ_2,τ_1}, for an arbitrary $\bar{v} \in C_{\tau_1}$, there exists $v \in C_{\tau_2}$

such that $\bar{v} = \phi'_{\tau_2,\tau_1}(v)$. Since $v \in C_{\tau_2}$, there exists a stochastic vector u such that

$$v = \begin{bmatrix} v_1 & \cdots & v_{c_{\tau_2}} \end{bmatrix} u. \tag{13}$$

Mapping both sides of (13) under operator ϕ'_{τ_2,τ_1}, we obtain

$$\bar{v} = \left[\phi'_{\tau_2,\tau_1}(v_1) \ \cdots \ \phi'_{\tau_2,\tau_1}\left(v_{c_{\tau_2}}\right)\right] u$$

meaning that vector $\bar{v}$ in C_{τ_1} can be expressed as a convex combination of c_{τ_2} vectors

$$\phi'_{\tau_2,\tau_1}(v_1), \ldots, \phi'_{\tau_2,\tau_1}\left(v_{c_{\tau_2}}\right). \tag{14}$$

Taking into account that the choice of $\bar{v} \in C_{\tau_1}$ was arbitrary, C_{τ_1} is a subset of the convex hull of the vectors of (14). On the other hand, all vectors of (14) lie in C_{τ_1} since each is the image of a vector in C_{τ_2}. Thus, C_{τ_1} is the convex hull of the vectors of (14). Consequently, the number of vertices of C_{τ_1} are within the vectors of (14), which results in $c_{\tau_1} \le c_{\tau_2}$. To complete the proof, it is sufficient to show that $c_{\tau_2} \le c_{\tau_1}$. To this aim, one defines linear operator $\phi'^{-1}_{\tau_2,\tau_1}$, which is the inverse of ϕ'_{τ_2,τ_1}, and proceeds with the exact same arguments as before. Notice that the inverse operator exists due to the invertibility of matrix $\Phi'(\tau_2, \tau_1)$. ■

Remark 3: The argument in the proof of Lemma 3 also implies that for every $\tau_2 \ge \tau_1$, the vertices of C_{τ_2} are mapped to the vertices of C_{τ_1} under ϕ'_{τ_2,τ_1}.

Let integer c be the constant value of c_τ, $\tau \ge 0$. We will show later in this section that c is equal to $\mathbf{rank}(A)$. To prove this, we first state the following two lemmas.

Lemma 4: $\mathbf{rank}(A)$ is equal to the dimension of the vector space generated by the vectors corresponding to the vertices of C_τ, for every $\tau \ge 0$.

Proof: It suffices to prove Lemma 4 for $\tau = 0$. Let $v_1, \ldots, v_c \in \mathbb{R}^N$ be the c vertices of C_0. It is easy to see that for any $u \in \mathbb{R}^N$

$$u \in \mathbf{null}_0(A) \Longleftrightarrow v_i' u = 0, \ \forall\, i, \ 1 \le i \le c.$$

It implies that the dimension of the vector space generated by $v_1, \ldots, v_c$ is $N - \mathbf{nullity}(A)$, which proves the lemma. ■

Lemma 5: For every $\tau \ge 0$, the vectors corresponding to the vertices of C_τ are linearly independent.

Sketch of Proof: In view of Lemmas 3 and 4, if one proves Lemma 5 for some $t \ge 0$, it is immediately implied for any $t \ge 0$. We shall show that there exists:

- a sufficiently large time T;
- agent sequences $\{i_t\}$, $1 \le i \le c$;
- subsets S^i of $\mathcal{V}$, $1 \le i \le c$

that satisfy the following properties:

i) $\Phi'_{i_t}(t, T)$ converges to an exclusive vertex of C_τ, say u_i, as t grows large;
ii) S^i's are nonempty and pairwise disjoint;
iii) $\sum_{j \notin S^i} \Phi_{i_t,j}(t, T)$ is sufficiently small for any $t \ge T$.

From Property iii), we then conclude that $\sum_{j \notin S^i} (u_i)_j$ is sufficiently small as well. This is, in fact, equivalent to $\sum_{j \in S^i} (u_i)_j$ being sufficiently close to 1 (since u_i is stochastic). Finally, from the disjointness of S^i's, we conclude that u_i's are linearly independent.

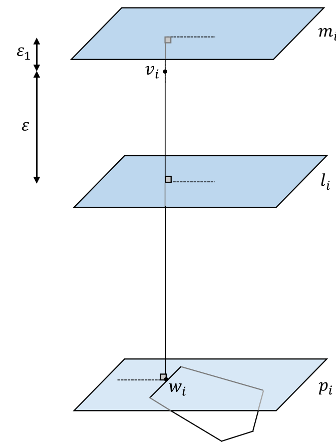


Fig. 2. Affine hyperplanes p_i, l_i, m_i are orthogonal to $v_i w_i$.

Proof: Let $v_1, \ldots, v_c$ be the c vertices of C_0. We note that vector v_i, $1 \le i \le c$, must lie outside the convex hull of v_j's, $j \ne i$, since otherwise it would not qualify as a vertex. For every i, $1 \le i \le c$, let w_i in the convex hull of v_j's, $j \ne i$, be such that it has the minimum distance from v_i. Notice that w_i exists since the convex hull of v_j's, $j \ne i$, is a compact subset of $\mathbb{R}^N$ as it is both bounded and closed. Define the following positive numbers:

$$\epsilon \triangleq \frac{1}{4} \min \{\|v_i - w_i\| \mid 1 \le i \le c\} \tag{15}$$

$$\epsilon_1 \triangleq \frac{\epsilon}{(2N)}. \tag{16}$$

For every i, $1 \le i \le c$, we define three affine hyperplanes in $\mathbb{R}^N$ as follows. As depicted in Fig. 2, let p_i be the affine hyperplane in $\mathbb{R}^N$ crossing w_i and orthogonal to vector $v_i - w_i$, i.e.,

$$p_i = \left\{ y \in \mathbb{R}^N \mid (v_i - w_i)'(y - w_i) = 0 \right\}.$$

Affine hyperplane p_i splits $\mathbb{R}^N$ into two half-spaces

$$p_i^- = \left\{ y \in \mathbb{R}^N \mid (v_i - w_i)'(y - w_i) < 0 \right\} \tag{17}$$

$$p_i^+ = \left\{ y \in \mathbb{R}^N \mid (v_i - w_i)'(y - w_i) > 0 \right\}. \tag{18}$$

Notice that $v_i \in p_i^+$ and the convex hull of v_j's, $j \ne i$, completely lies in $p_i \cup p_i^-$. Let l_i be the unique affine hyperplane which is parallel to p_i, distant ϵ from v_i, and crossing segment $v_i w_i$ (the convex hull of v_i and w_i). Formally, we have

$$l_i = \Big\{ y \in \mathbb{R}^N | (v_i - w_i)' \times \left[y - \left(\frac{\|v_i - w_i\| - \epsilon}{\|v_i - w_i\|} v_i + \frac{\epsilon}{\|v_i - w_i\|} w_i \right) \right] = 0 \Big\}.$$

Similar to (17) and (18), half-spaces l_i^- and l_i^+ are defined. Notice that $v_i \in l_i^+$ and the convex hull of v_j's, $j \ne i$, lies in l_i^-. Finally, let m_i be the affine hyperplane parallel to p_i, distant ϵ_1 from v_i, and not crossing segment $v_i w_i$

$$m_i = \Big\{ y \in \mathbb{R}^N | (v_i - w_i)' \times \left[y - \left(\frac{\|v_i - w_i\| + \epsilon_1}{\|v_i - w_i\|} v_i + \frac{-\epsilon_1}{\|v_i - w_i\|} w_i \right) \right] = 0 \Big\}$$

and define half-spaces m_i^- and m_i^+ similar to (17) and (18). Notice that C_0, including its vertex v_i, completely lies in m_i^-.

Since C_0 is the limit of $C_{t,0}$ as t goes to infinity, there must exist a sufficiently large time $T \geq 0$, such that for $t \geq T$, every point in $C_{t,0}$ lies within an ϵ_1-distance of C_0. Define for every i, $1 \leq i \leq c$

$$S^i = \left\{ j \in \mathcal{V} | \Phi'_j(T,0) \in l_i^+ \cap m_i^- \right\}.$$

Set $l_i^+ \cap m_i^-$ can be viewed as a strip in $\mathbb{R}^N$ which is margined by the two affine hyperplanes l_i and m_i. Note that by the assumption, T is so large that every point in $C_{T,0}$, including $\Phi'_j(T,0)$ for any j, lies within an ϵ_1-distance of C_0. Therefore, we must have $\Phi'_j(T,0) \in m_i^-$ for every j. In other words, $\Phi'_j(T,0)$ either lies in $l_i^+ \cap m_i^-$ or $l_i \cup l_i^-$. The latter cannot occur for every j since otherwise $C_{T,0}$ would completely lie in l_i^-, which would be a contradiction as $C_{T,0}$ must contain $v_i \notin l_i^-$. Thus, $\Phi'_j(T,0)$ lies in $l_i^+ \cap m_i^-$ for some j, meaning that S^i is nonempty. We claim that S^i's, $1 \leq i \leq c$, are pairwise disjoint sets. Let $i, j \in \mathcal{V}$ be arbitrary but fixed. By the defining property of T, for every $k \in S^i \cap S^j$, $\Phi'_k(T,0)$ must be within an ϵ_1-distance of C_0. Noticing that

$$\left\{ \Phi'_k(T,0) \mid k \in S^i \cap S^j \right\} \subset l_i^+ \cap l_j^+$$

to prove the claim, it suffices to show that any point in $l_i^+ \cap l_j^+$ cannot be within ϵ_1-distance of C_0. Consider an arbitrary point in C_0 and let $\sum_{k=1}^c \beta_k v_k$ be its representation as the convex combination of $v_1, \ldots, v_c$. Obviously, at least one of β_i and β_j does not exceed 1/2. If $\beta_i \leq 1/2$, then any point in l_i^+ has a distance greater than $(1/2)\|v_i - w_i\| - \epsilon$ from $\sum_{k=1}^c \beta_k v_k$, and from (15), we know that

$$\frac{1}{2}\|v_i - w_i\| - \epsilon \geq \epsilon > \epsilon_1.$$

Similarly, if $\beta_j \leq 1/2$, any point in l_j^+ has a distance greater than ϵ_1 from $\sum_{k=1}^c \beta_k v_k$. Thus, either way, we conclude that any point in $l_i^+ \cap l_j^+$ cannot be within ϵ_1-distance of $\sum_{k=1}^c \beta_k v_k$, which proves the claim. With polytope C_0 being the limit of shrinking convex hulls $C_{t,0}$'s, it follows that for $i = 1, \ldots, c$, there exist sequences $\{i_t\}$ of individuals such that $\Phi'_{i_t}(t,0)$ converges to v_i. Therefore, after some finite time, we have the following inequality:

$$\left\| \Phi'_{i_t}(t,0) - v_i \right\| < \epsilon_1. \tag{19}$$

Without loss of generality, we can assume that inequality (19) holds for any $t \geq T$ (otherwise, we would proceed by replacing T with T', $T' > T$, such that inequality (19) holds for every $t \geq T'$). For any $t \geq T$, we now have

$$\begin{aligned} \Phi'_{i_t}(t,0) &= \Phi'(T,0)\Phi'_{i_t}(t,T) \\ &= \sum_{j \in \mathcal{V}} \Phi_{i_t,j}(t,T)\Phi'_j(T,0) \\ &= \sum_{j \notin S^i} \Phi_{i_t,j}(t,T)\Phi'_j(T,0) \\ &\quad + \sum_{j \in S^i} \Phi_{i_t,j}(t,T)\Phi'_j(T,0). \end{aligned} \tag{20}$$

We now claim that for every i, $1 \leq i \leq c$, and any $t \geq T$, the following two inequalities hold:

$$\sum_{j \notin S^i} \Phi_{i_t,j}(t,T) < \frac{2}{(2N+1)} \tag{21}$$

$$\sum_{j \in S^i} \Phi_{i_t,j}(t,T) > 1 - 2/(2N+1). \tag{22}$$

To prove (21) and (22), we use (20) to find a lower bound for the distance from $\Phi'_{i_t}(t,0)$, $t \geq T$, to affine hyperplane m_i as drawn in Fig. 2. Remember that if $j \in S^i$, then $\Phi'_j(T,0)$ lies in $l_i^+ \cap m_i^-$, while if $j \notin S^i$, then $\Phi'_j(T,0)$ lies in $l_i \cup l_i^-$. For a fixed i, $1 \leq i \leq c$, let $\eta \triangleq \sum_{j \notin S^i} \Phi_{i_t,j}(t,T)$. With matrix $\Phi(t,T)$ being row-stochastic, it follows that $\sum_{j \in S^i} \Phi_{i_t,j}(t, T) = 1 - \eta$. Using (20), we now conclude that $\eta(\epsilon_1 + \epsilon) + (1-\eta)0$ is a lower bound for the distance from $\Phi'_{i_t}(t,0)$, $t \geq T$, to affine hyperplane m_i. On the other hand, this distance is upper bounded by $2\epsilon_1$ according to (19). Thus

$$\eta(\epsilon_1 + \epsilon) + (1-\eta)0 < 2\epsilon_1. \tag{23}$$

Therefore, remembering $\epsilon = 2N\epsilon_1$, inequality (23) implies that $\eta < 2/(2N+1)$, which proves (21) and (22). From Remark 3, every vertex v_i of C_0 is the image of a vertex of C_T, say u_i, under $\phi'_{T,0}$

$$v_i = \phi'_{T,0}(u_i). \tag{24}$$

We also have

$$\Phi'_{i_t}(t,0) = \Phi'(T,0)\Phi'_{i_t}(t,T) = \phi'_{T,0}\left(\Phi'_{i_t}(t,T)\right). \tag{25}$$

Recalling $\lim_{t\to\infty} \Phi'_{i_t}(t,0) = v_i$ from (24) and (25), and the continuity of the inverse of operator $\phi'_{T,0}$, we obtain $\lim_{t\to\infty} \Phi'_{i_t}(t,T) = u_i$. Considering (22) again, and taking limits as $t \to \infty$, we conclude

$$\sum_{j \in S^i} (u_i)_j \geq 1 - 2/(2N+1) \tag{26}$$

and consequently

$$\sum_{j \notin S^i} (u_i)_j \leq \frac{2}{(2N+1)}. \tag{27}$$

In the following text, we show that $u_1, \ldots, u_c$, that is, the vertices of C_T, are linearly independent. Assume that $\lambda_1, \ldots, \lambda_c \in \mathbb{R}$ are such that

$$\sum_{i=1}^c \lambda_i u_i = 0. \tag{28}$$

Moreover, let k, $1 \leq k \leq c$ be such that

$$|\lambda_k| = \max_{1 \leq i \leq c} \{|\lambda_i|\} \triangleq \lambda.$$

If $\lambda > 0$, noting that (26) and (27) hold for any vertex u_i of C_T including u_k, and that the S^i's are disjoint sets of individuals, from (28), we have

$$\begin{aligned}0 &= \left|\sum_{j\in S^k}\sum_{i=1}^{c}\lambda_i(u_i)_j\right| \\ &= \left|\sum_{j\in S^k}\lambda_k(u_k)_j + \sum_{j\in S^k}\sum_{i\neq k}\lambda_i(u_i)_j\right| \\ &\geq |\lambda_k|.\left|\sum_{j\in S^k}(u_k)_j\right| - \sum_{i\neq k}\left(|\lambda_i|.\sum_{j\in S^k}(u_i)_j\right) \\ &\geq |\lambda_k|.\left|\sum_{j\in S^k}(u_k)_j\right| - \sum_{i\neq k}\left(|\lambda_i|.\sum_{j\notin S^i}(u_i)_j\right) \\ &\geq \lambda\left(1-2/(2N+1)\right) - \lambda(c-1).2/(2N+1) \\ &= \lambda\left(2(N-c)+1\right)/(2N+1) > 0 \end{aligned} \tag{29}$$

which is a contradiction. We note that the second inequality in (29) is a result of S^i and S^k being disjoint sets, that is, $S^k \in \mathcal{V}\setminus S^i$. Thus, we must have $\lambda = 0$, which means $\lambda_i = 0$, $\forall i$, $1 \leq i \leq c$. This proves the lemma. ■

Theorem 2: The size of a minimal EGC of a network with dynamics (1) is equal to c, that is, the constant value of c_τ, $\tau \geq 0$, where c_τ is the number of vertices of limiting polytope C_τ.

Proof: Theorem 2 is an immediate result of Theorem 1 and Lemmas 4 and 5. ■

Theorem 3: The size of a minimal EGC of a network with dynamics (1) is less than or equal to the number of ergodicity classes of $\{A(t)\}$.

Proof: We show that the number of ergodicity classes is at least c. Recall limiting polytope C_0 with vertices $v_1,\ldots,v_c$ from earlier in the section. Remember from the proof of Lemma 5 that for $i = 1,\ldots,c$, there exist sequences $\{i_t\}$ of individuals such that $\Phi'_{i_t}(t,0)$ converges to v_i. Let

$$\epsilon_2 = \frac{1}{3}\min\left\{\|v_i - v_j\|\,|i,j\in\mathcal{V}, i\neq j\right\}. \tag{30}$$

In view of (10), for any fixed $\tau \geq 0$, there exists a sufficiently large time T such that for any $t \geq T$ and i, j is in the same ergodicity class

$$\left\|\Phi'_i(t,\tau) - \Phi'_j(t,\tau)\right\| < \epsilon_2. \tag{31}$$

Moreover, there is $T' > 0$ such that for any $t \geq T'$ and $i \in \mathcal{V}$

$$\left\|\Phi'_{i_t}(t,0) - v_i\right\| < \epsilon_2. \tag{32}$$

Thus, for every $t \geq T'$ and $i \neq j$, $1 \leq i, j \leq c$, we must have

$$\begin{aligned}3\epsilon_2 &\leq \|v_i - v_j\| \\ &\leq \left\|v_i - \Phi'_{i_t}(t,0)\right\| + \left\|\Phi'_{i_t}(t,0) - \Phi'_{j_t}(t,0)\right\| \\ &\quad + \left\|\Phi'_{j_t}(t,0) - v_j\right\| \\ &< \epsilon_2 + \left\|\Phi'_{i_t}(t,0) - \Phi'_{j_t}(t,0)\right\| + \epsilon_2. \end{aligned} \tag{33}$$

Note that the first, second, and third inequalities of (33) are implied from (30), the triangle inequality, and (32), respectively. From (33)

$$\left\|\Phi'_{i_t}(t,0) - \Phi'_{j_t}(t,0)\right\| > \epsilon_2 \qquad \forall\, t \geq T'. \tag{34}$$

Taking (31) into account, from (34), we conclude that i_t and j_t cannot be in the same ergodicity class for all $t \geq \max\{T, T'\}$. Thus, there are at least c distinct ergodicity classes, which completes the proof. ■

Remark 4: In case $\{A(t)\}$, the coupling chain of a network with dynamics (1), is class-ergodic, the occurrence of multiple consensus in the network is guaranteed, and the number of ergodic classes becomes equal to the number of consensus clusters. Yet, this number may be larger than the size of a minimal EGC of the network. In other words, there may exist an EGC in which some of the consensus clusters has no representative. As an illustrative example, consider system (1) of three individuals with a fixed coupling chain

$$A(t) = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{3} & -1 & \frac{2}{3} \\ 0 & 0 & 0 \end{bmatrix}, \qquad \forall\, t \geq 0.$$

We then have

$$\lim_{t\to\infty} x(t) = \begin{bmatrix} x_1(0) \\ \frac{(x_1(0)+2x_3(0))}{3} \\ x_3(0) \end{bmatrix}.$$

Also notice that for the corresponding state transition matrix, we have

$$\lim_{t\to\infty}\Phi(t,\tau) = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix}, \qquad \forall\, \tau \geq 0.$$

Therefore, each individual forms a consensus cluster, that is, there are three consensus clusters. However, subgroup {1,3} with size two is an EGC of the network since, irrespective of the initial opinion of individual 2, an agreement on value x^* is achieved if individuals 1 and 3 initialize their opinions at x^*.

IV. Approximation of Chains

In this section, we introduce the continuous time version of l_1-approximation of chains, initially introduced in [22] for discrete time chains, and show that the rank of a chain is invariant under any l_1-approximation.

Definition 6 (l_1-Approximation): Chain $\{A(t)\}$ is said to be an l_1-approximation of chain $\{B(t)\}$ if

$$\int_0^\infty \|A(t) - B(t)\|_\infty\, dt < \infty \tag{35}$$

where $\|.\|_\infty$ refers to the maximum absolute row sum.

It is not difficult to show that l_1-approximation is an equivalence relation in the set of chains that satisfy Assumptions 1 and 2. The importance of the l_1-approximation notion in this work comes from the following proposition.

Proposition 1: Rank of a chain is invariant under l_1-approximation.

In the remainder of this section, we prove Proposition 1. We first state a lemma which roughly implies that if two chains are close, their state transition matrices are close as well.

Lemma 6: For any two chains $\{A(t)\}$ and $\{B(t)\}$, and any $t \geq \tau \geq 0$, the following holds:

$$\|\Phi_A(t,\tau) - \Phi_B(t,\tau)\|_\infty \leq \int_\tau^t \|A(t') - B(t')\|_\infty\, dt' \tag{36}$$

where $\Phi_A(t,\tau)$ and $\Phi_B(t,\tau)$ are state transition matrices associated with $\{A(t)\}$ and $\{B(t)\}$, respectively.

Proof: For simplification, we write $\|.\|$ for $\|.\|_\infty$ in this proof. Inequality (36) is obviously satisfied if $t=\tau$. Let $t>\tau\geq 0$ be arbitrary but fixed. From Assumption 2, a_{ij} and b_{ij} are bounded on interval $[\tau,t]$ for every $i,j\in\mathcal{V}$. Define

$$M \triangleq \sup_{\tau\leq t'\leq t}\{\|A(t')\|,\|B(t')\|\}. \tag{37}$$

For any $\tau_0,t_0\in[\tau,t]$, $\tau_0\leq t_0$ and according to (3), we have

$$\begin{aligned}\|\Phi_A(t_0,\tau_0)-I\| &= \left\|\int_{\tau_0}^{t_0} A(t')\Phi_A(t',\tau_0)dt'\right\| \\ &\leq \int_{\tau_0}^{t_0}\|A(t')\|\,\|\Phi_A(t',\tau_0)\|\,dt' \\ &= \int_{\tau_0}^{t_0}\|A(t')\|\,dt' \leq M(t_0-\tau_0).\end{aligned} \tag{38}$$

From (38), since t_0 is arbitrary, we conclude

$$\begin{aligned}\sup_{\tau_0\leq t'\leq t_0}\|\Phi_A(t',\tau_0)-I\| &\leq \sup_{\tau_0\leq t'\leq t_0} M(t'-\tau_0) \\ &= M(t_0-\tau_0).\end{aligned} \tag{39}$$

Similarly

$$\sup_{\tau_0\leq t'\leq t_0}\|\Phi_B(t',\tau_0)-I\| \leq M(t_0-\tau_0). \tag{40}$$

Let $\alpha>0$ be arbitrary. If $t_0-\tau_0<(\alpha/2M^2)$, one further writes (41), shown at the bottom of the page. Note that to write the second inequality of (41), we used $\|\Phi_A\|=1$, together with (37), while the third inequality is implied from (39) and (40). Thus, if we define

$$f(t_0,\tau_0)\triangleq\frac{\|\Phi_A(t_0,\tau_0)-\Phi_B(t_0,\tau_0)\|-\int_{\tau_0}^{t_0}\|A(t')-B(t')\|\,dt'}{t_0-\tau_0}$$

then, for any $t_0-\tau_0<\alpha/(2M^2)$, we have

$$f(t_0,\tau_0)<\alpha. \tag{42}$$

On the other hand, one can prove

$$f(t,\tau)\leq\frac{1}{2}f\left(t,\frac{t+\tau}{2}\right)+\frac{1}{2}f\left(\frac{t+\tau}{2},\tau\right) \tag{43}$$

by taking into account

$$\begin{aligned}&\|\Phi_A(t,\tau)-\Phi_B(t,\tau)\| \\ &=\left\|\Phi_A\left(t,\frac{t+\tau}{2}\right)\Phi_A\left(\frac{t+\tau}{2},\tau\right) - \Phi_B\left(t,\frac{t+\tau}{2}\right)\Phi_B\left(\frac{t+\tau}{2},\tau\right)\right\| \\ &=\left\|\left[\Phi_A\left(t,\frac{t+\tau}{2}\right)-\Phi_B\left(t,\frac{t+\tau}{2}\right)\right]\Phi_A\left(\frac{t+\tau}{2},\tau\right) + \Phi_B\left(t,\frac{t+\tau}{2}\right)\left[\Phi_A\left(\frac{t+\tau}{2},\tau\right)-\Phi_B\left(\frac{t+\tau}{2},\tau\right)\right]\right\| \\ &\leq\left\|\Phi_A\left(t,\frac{t+\tau}{2}\right)-\Phi_B\left(t,\frac{t+\tau}{2}\right)\right\|\left\|\Phi_A\left(\frac{t+\tau}{2},\tau\right)\right\| + \left\|\Phi_B\left(t,\frac{t+\tau}{2}\right)\right\|\left\|\Phi_A\left(\frac{t+\tau}{2},\tau\right)-\Phi_B\left(\frac{t+\tau}{2},\tau\right)\right\| \\ &=\left\|\Phi_A\left(t,\frac{t+\tau}{2}\right)-\Phi_B\left(t,\frac{t+\tau}{2}\right)\right\| + \left\|\Phi_A\left(\frac{t+\tau}{2},\tau\right)-\Phi_B\left(\frac{t+\tau}{2},\tau\right)\right\|\end{aligned}$$

$$\begin{aligned}\frac{\|\Phi_A(t_0,\tau_0)-\Phi_B(t_0,\tau_0)\|}{t_0-\tau_0} &= \frac{\left\|\int_{\tau_0}^{t_0}[A(t')\Phi_A(t',\tau_0)-B(t')\Phi_B(t',\tau_0)]\,dt'\right\|}{t_0-\tau_0} \\ &= \frac{\left\|\int_{\tau_0}^{t_0}[(A(t')-B(t'))\,\Phi_A(t',\tau_0)+B(t')\,(\Phi_A(t',\tau_0)-I)-B(t')\,(\Phi_B(t',\tau_0)-I)]\,dt'\right\|}{t_0-\tau_0} \\ &\leq \frac{\int_{\tau_0}^{t_0}\|A(t')-B(t')\|\,\|\Phi_A(t',\tau_0)\|\,dt'}{t_0-\tau_0} + \frac{\int_{\tau_0}^{t_0}\|B(t')\|\,\|\Phi_A(t',\tau_0)-I\|\,dt'}{t_0-\tau_0} \\ &\quad + \frac{\int_{\tau_0}^{t_0}\|B(t')\|\,\|\Phi_B(t',\tau_0)-I\|\,dt'}{t_0-\tau_0} \\ &\leq \frac{\int_{\tau_0}^{t_0}\|A(t')-B(t')\|\,dt'}{t_0-\tau_0} + M\left(\sup_{\tau_0\leq t'\leq t_0}\|\Phi_A(t',\tau_0)-I\|\right) + M\left(\sup_{\tau_0\leq t'\leq t_0}\|\Phi_B(t',\tau_0)-I\|\right) \\ &\leq \frac{\int_{\tau_0}^{t_0}\|A(t')-B(t')\|\,dt'}{t_0-\tau_0} + M.M(t_0-\tau_0)+M.M(t_0-\tau_0) \\ &< \frac{\int_{\tau_0}^{t_0}\|A(t')-B(t')\|\,dt'}{t_0-\tau_0} + \alpha\end{aligned} \tag{41}$$

and

$$\int_{\tau}^{t} \|A(t') - B(t')\| \, dt' = \int_{\frac{t+\tau}{2}}^{t} \|A(t') - B(t')\| \, dt' + \int_{\tau}^{\frac{t+\tau}{2}} \|A(t') - B(t')\| \, dt'.$$

We claim that $f(t,\tau) < \alpha$. Inequality (43) means that if $f(t,(t+\tau)/2)$ and $f((t+\tau)/2,\tau)$ are both less than α, then so is $f(t,\tau)$. Thus, it suffices to prove the claim for $f(t,(t+\tau)/2)$ and $f((t+\tau)/2,\tau)$. Similarly, we split each interval $(t,(t+\tau)/2)$ and $((t+\tau)/2,\tau)$ into half and continue the process until we obtain intervals of length less than $\alpha/(2M^2)$. Then, using (42), our claim is proved. Since $\alpha > 0$ was chosen arbitrarily, $f(t,\tau) \leq 0$. This proves the lemma. ■

We now prove Proposition 1. Let chain $\{B(t)\}$ be an l_1-approximation of chain $\{A(t)\}$. We only show that $\mathbf{rank}(B) \geq \mathbf{rank}(A)$ which proves Proposition 1 due to the symmetry of l_1-approximation. Let $\mathbf{rank}(A) = c$. Assume that $v_1(\tau), \ldots, v_c(\tau)$ are the c vertices of C_τ formed based on the state transition matrix of $\{A(t)\}$. Let

$$S_\tau \triangleq \{v_1(\tau), \ldots, v_c(\tau)\}$$
$$v(\tau) \triangleq [v_1(\tau) \mid \cdots \mid v_c(\tau)] \,.$$

Define *linear independency index* ζ as follows to quantify the extent by which the vertices of C_τ are linearly independent:

$$\zeta(\tau) \triangleq \min\left\{\|v(\tau)\lambda\|_1 \mid \lambda \in \mathbb{R}^n, \|\lambda\|_1 = 1\right\} \quad (44)$$

where $\|.\|_1$ refers to the 1-norm. Note that existence of a minimum in (44) is guaranteed by continuity and the compactness of the unit ball. Furthermore, $\zeta(\tau) > 0$ for any τ since zero does not belong to the set on the right-hand side of (44) due to the linear independency of the vertices of C_τ. We now claim that $\zeta(\tau)$ is nondecreasing by τ. Assume that $\tau_2 \geq \tau_1$. According to Remark 3, for some permutation matrix P, we have $v(\tau_1) = P\Phi'(\tau_2,\tau_1)v(\tau_2)$. Thus, $\forall \lambda \in \mathbb{R}^N$

$$\begin{aligned} \|v(\tau_1)\lambda\|_1 &= \|P\Phi'(\tau_2,\tau_1)v(\tau_2)\lambda\|_1 \\ &\leq \|P\|_1 \, \|\Phi'(\tau_2,\tau_1)\|_1 \, \|v(\tau_2)\lambda\|_1 \leq \|v(\tau_2)\lambda\|_1 \end{aligned}$$

which proves our claim. Hence, for all $\tau \geq 0$

$$\zeta(\tau) \geq \zeta(0) \triangleq \zeta_0 > 0. \quad (45)$$

From the definition of l_1-approximation (35), we conclude

$$\lim_{\tau\to\infty} \int_{\tau}^{\infty} \|A(t) - B(t)\|_\infty = 0.$$

Thus, from Lemma 6, for any $\epsilon > 0$, there exists $T > 0$ such that for every $t \geq \tau \geq T$, we have

$$\|\Phi_A(t,\tau) - \Phi_B(t,\tau)\|_\infty < \epsilon$$

or equivalently

$$\|\Phi'_A(t,\tau) - \Phi'_B(t,\tau)\|_1 < \epsilon. \quad (46)$$

Let $\mathbf{rank}(B) \triangleq c'$. Assume for some fixed $\tau \geq T$, that $u_1(\tau), \ldots, u_{c'}(\tau)$ are the c' vertices of $C_\tau(B)$ which is the limiting polytope generated by the state transition matrix of $\{B(t)\}$ and

$$u(\tau) \triangleq [u_1(\tau) \mid \cdots \mid u_{c'}(\tau)] \,.$$

If $t \to \infty$ in (46), limiting polytopes $C_\tau(A)$ and $C_\tau(B)$ lie within ϵ-distance of each other with respect to 1-norm, i.e.,

$$\sup_{z_1 \in C_\tau} \left\{ \inf_{z_2 \in C_\tau(B)} \{\|z_1 - z_2\|_1\} \right\} \leq \epsilon.$$

In particular, every vertex $v_i(\tau)$ of C_τ is within an ϵ-distance of some point in $C_\tau(B)$. Thus, since any point in $C_\tau(B)$ can be expressed as a convex combination of the vertices of $C_\tau(B)$, for some $\gamma_i \in \mathbb{R}^{c'}$, $\|\gamma_i\|_1 = 1$, we must have

$$\|v_i(\tau) - u(\tau)\gamma_i\|_1 \leq \epsilon.$$

Let $\gamma = [\gamma_1 \mid \cdots |\gamma_c]$. Now, for any $\lambda \in \mathbb{R}^c$, where $\|\lambda\|_1 = 1$

$$\begin{aligned} \|v(\tau)\lambda\|_1 &= \|[v(\tau) - u(\tau)\gamma + u(\tau)\gamma]\,\lambda\|_1 \\ &\leq \|[v(\tau) - u(\tau)\gamma]\,\lambda\|_1 + \|u(\tau)\gamma\lambda\|_1 \\ &\leq \|v(\tau) - u(\tau)\gamma\|_1 \, \|\lambda\|_1 + \|u(\tau)\gamma\lambda\|_1 \\ &\leq \epsilon + \|u(\tau)\gamma\lambda\|_1 \,. \end{aligned} \quad (47)$$

Assume on the contrary, that $c' < c$. Then, since γ is a $c' \times c$ matrix, it cannot be full-row rank. Thus, there exists $\lambda_0 \in \mathbb{R}^c$, $\|\lambda_0\|_1 = 1$, such that $\gamma\lambda_0 = \mathbf{0}$. From (47), we now obtain

$$\|v(\tau)\lambda_0\|_1 \leq \epsilon. \quad (48)$$

Recalling linearly independency index ζ from (44), one concludes from (48) that $\zeta(\tau) \leq \epsilon$. Remembering that ϵ is an arbitrarily positive number, one could choose $\epsilon < \zeta_0$, which would contradict (45). Thus, $c' \geq c$, and the proof of Proposition 1 is now complete.

V. Lower Bounds on the Rank of Chains

In this section, we obtain lower bounds on the size of a minimal EGC in a network with dynamics (1), that is, on $\mathbf{rank}(A)$. We recall the following definition from [9], [24].

Definition 7 (Unbounded Interactions Graph): The unbounded interactions graph of a chain $\{A(t)\}$, $\mathcal{H}_1(\mathcal{V},\mathcal{E}_1)$, is a fixed directed graph such that for every distinct nodes $i, j \in \mathcal{V}$, $(i,j) \in \mathcal{E}_1$ if

$$\int_{0}^{\infty} a_{ji}(t)dt = \infty.$$

Thus, a link is drawn from i to j if the total influence of individual i on individual j is unbounded over the infinite time interval.

Definition 8 (Set-Root): A subset $\mathcal{S}' \subset \mathcal{V}$ is called a *set-root* of $\mathcal{H}_1(\mathcal{V},\mathcal{E}_1)$ if for every node $i \in \mathcal{V}$, we have $i \in \mathcal{S}'$ or there exists $j \in \mathcal{S}'$ such that i is reachable from j.

Theorem 4: Let $\mathcal{H}_1(\mathcal{V},\mathcal{E}_1)$ be the unbounded interaction graph associated with chain $\{A(t)\}$. Then, the size of a minimal EGC of a network with dynamics (1) is lower bounded by the size of the smallest set-root of $\mathcal{H}_1(\mathcal{V},\mathcal{E}_1)$.

Proof: Define a chain $\{B(t)\}$ from chain $\{A(t)\}$ as follows: For every $i \neq j \in \mathcal{V}$ and $t \geq 0$, set

$$b_{ij}(t) = \begin{cases} a_{ij}(t) & \text{if } (j,i) \in \mathcal{E}_1 \\ 0 & \text{if } (j,i) \notin \mathcal{E}_1 \end{cases}$$

and $b_{ii}(t) = -\sum_{j \neq i} b_{ij}(t)$ for every $i \in \mathcal{V}$ and $t \geq 0$. Consider an opinion network with coupling chain $\{B(t)\}$

$$\dot{y}(t) = B(t)y(t), \qquad t \geq 0 \tag{49}$$

where $y(t) \in \mathbb{R}^N$ is the opinion vector. Since $\{B(t)\}$ is an l_1-approximation of $\{A(t)\}$, from Proposition 1, the two chains share the same rank. Consequently, minimal EGC's of networks with dynamics (1) and (49) are of the same size. Moreover, the two chains share a common unbounded interactions graph. Therefore, it suffices to prove Theorem 4 for network with dynamics (49). To this aim, we show that every EGC of the network with dynamics (49) is also a set-root of $\mathcal{H}_1$. Assume on the contrary, that $\mathcal{S} \subset \mathcal{V}$ is an EGC which is not a set-root of $\mathcal{H}_1$. Define

$$n(\mathcal{S}) \triangleq \mathcal{S} \cup \{i \mid i \in \mathcal{V}, \exists j \in \mathcal{S} : i \text{ is reachable from } j \text{ in } \mathcal{H}_1\}.$$

Since $\mathcal{S}$ is not a set-root, $n(\mathcal{S}) \subsetneq \mathcal{V}$. It is easy to see that there is no link from $n(\mathcal{S})$ to $\mathcal{V} \setminus n(\mathcal{S})$ in $\mathcal{H}_1$. According to the way that chain $\{B(t)\}$ was constructed, this means that $n(\mathcal{S})$ has zero influence on $\mathcal{V} \setminus n(\mathcal{S})$ at any time instant. Thus, since $\mathcal{S} \subset n(\mathcal{S})$, individuals in $\mathcal{S}$ cannot, in general, lead individuals in $\mathcal{V} \setminus n(\mathcal{S})$ to agreeing on an arbitrary value x^*. For instance, given a desired consensus value x^*, if the opinions of individuals in $\mathcal{V} \setminus n(\mathcal{S})$ are all initialized at value $x^* + 1$, they will never change and, consequently, they will never converge to x^*. Thus, $\mathcal{S}$ is not an EGC, which completes the proof. ■

An important special case of Theorem 4 is described as follows. Let us first define the continuous time counterpart of the *infinite flow graph* of a chain according to [25].

Definition 9 (Infinite Flow Graph): The infinite flow graph of a chain $\{A(t)\}$, $\mathcal{H}_2(\mathcal{V}, \mathcal{E}_2)$, is a fixed undirected graph such that for every distinct nodes $i, j \in \mathcal{V}$, $\{i, j\} \in \mathcal{E}_2$ if

$$\int_0^\infty (a_{ij}(t) + a_{ji}(t))\, dt = \infty.$$

As a special case of Theorem 4, one obtains the following:

Corollary 1: Rank of a chain is lower bounded by the number of connected components of its infinite flow graph.

VI. Rank of Time-Invariant (TI) Chains

Let $\{A(t)\}$ be a TI chain, that is, $A(t) = \hat{A}, \forall t \geq 0$, where $\hat{A}$ is a fixed matrix with the property that each of its rows adds up to zero and its offdiagonal elemetns are non-negative. Assume that $\mathrm{rank}(\hat{A})$ and $\mathrm{nullity}(\hat{A})$ represent the rank and the nullity of $\hat{A}$. Notice that roman style is used for matrix operators as opposed to the chain operators in order to avoid any ambiguity. For state transition matrix $\Phi(t, \tau)$ associated with TI chain $\{\hat{A}\}$, we have

$$\Phi(t, \tau) = e^{\hat{A}(t-\tau)}, \qquad t \geq \tau \geq 0.$$

Note that $\hat{A}$ is marginally stable and has all eigenvalues negative but one equal to zero with algebraic multiplicity $\mathrm{nullity}(\hat{A})$. Thus, $\lim_{t-\tau\to\infty} \Phi(t, \tau)$ exists and the limit has eigenvalue zero

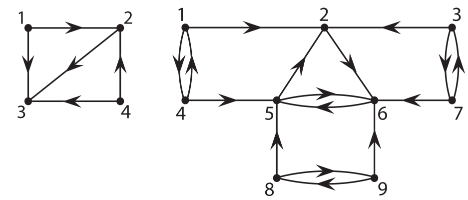


Fig. 3. Unweighted graphs associated with two TI chains. {1, 4} (left) and {1, 3, 8} (right) are the smallest set-roots.

with algebraic multiplicity $\mathrm{rank}(\hat{A})$ and eigenvalue one with algebraic multiplicity $\mathrm{nullity}(\hat{A})$. Hence

$$\mathbf{rank}(A) = \mathrm{nullity}(\hat{A}).$$

Employing a graph-theoretic approach, $\mathrm{nullity}(\hat{A})$ is the size of a minimal set-root of the *weighted directed* graph whose Laplacian is $\hat{A}$ [26, Cor. 1]. Since an unweighted version of the graph described serves as the unbounded interactions graph associated with TI chain $\{A(t)\}$, $A(t) = \hat{A}$, $\forall t \geq 0$, the following corollary is implied (see Fig. 3).

Corollary 2: For a TI chain $\{A(t)\}$, the lower bound provided in Theorem 4 is achieved. More specifically, the size of a minimal EGC of a network with dynamics (1) is equal to the size of the smallest set-root of the unbounded interactions graph associated with $\{A(t)\}$.

Remember that any TI chain $\{A(t)\}$ is class-ergodic and the numbers of ergodic classes provides an upper bound for the size of a minimal EGC according to Theorem 3. For example, for the coupling graphs depicted in Fig. 3, the number of ergodic classes are 4 (left) and 6 (right). The graph interpretation of the notion of rank explains the following two properties:

1) For any TI chain $\{A(t)\}$ and $\alpha > 0$

$$\mathbf{rank}\left(\{\alpha A(t)\}\right) = \mathbf{rank}\left(\{A(t)\}\right).$$

2) For any two TI chains $\{A(t)\}$ and $\{B(t)\}$

$$\mathbf{rank}(\{A(t)+B(t)\}) \leq \min\{\mathbf{rank}(\{A(t)\}), \mathbf{rank}(\{B(t)\})\}.$$

Remark 5: While Statement 1) seems to hold for any time-varying chain $\{A(t)\}$ as well, there exist time-varying chains $\{A(t)\}$ and $\{B(t)\}$ that do not satisfy Statement 2). This means that more interactions between individuals may surprisingly increase the size of the minimal EGC of a network. The following is an example; let:

$$A(t) = \begin{cases} \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \text{if } t \in [2^{2k} - 1, 2^{2k}),\ k \in \mathbb{N} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} & \text{if } t \in [2^{2k}, 2^{2k+1} - 1),\ k \in \mathbb{N} \end{cases}$$

and $A(t) = \mathbf{0}$ elsewhere. Also let

$$B(t) = \begin{cases} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix} & \text{if } t \in [2^{2k+1} - 1, 2^{2k+1}),\ k \in \mathbb{N} \\ \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \text{if } t \in [2^{2k+1}, 2^{2k+2} - 1),\ k \in \mathbb{N} \end{cases}$$

and $B(t) = \mathbf{0}$ elsewhere. Note that at every time instant, either $A(t)$ or $B(t)$ is $\mathbf{0}$. It is easy to see that $\{A(t)\}$ and $\{B(t)\}$ are ergodic chains. More specifically, for every $\tau \geq 0$, we have

$$\lim_{t\to\infty} \Phi_A(t,\tau) = [0\ 0\ 1][1\ 1\ 1]'$$
$$\lim_{t\to\infty} \Phi_B(t,\tau) = [1\ 0\ 0][1\ 1\ 1]'.$$

Therefore, $\mathbf{rank}(A) = \mathbf{rank}(B) = 1$. However, it can be shown that $\mathbf{rank}(\{A(t) + B(t)\}) = 2$. More precisely, $\{1, 3\}$ forms a minimal EGC of the network with coupling chain $\{A(t) + B(t)\}$.

VII. Rank of Chains in Class $\mathcal{P}^*$

From the fundamental work of Kolmogorov [27], for the state transition matrix $\Phi(t,\tau)$, $t \geq \tau \geq 0$, associated with an arbitrary chain $\{A(t)\}$, there exists a sequence of stochastic vectors $\{\pi(t)\}$, called an *absolute probability sequence*, such that for every $t \geq \tau \geq 0$

$$\pi'(\tau) = \pi'(t)\Phi(t,\tau). \tag{50}$$

Remember that by a stochastic vector, we mean a vector whose elements are all non-negative and add up to 1. For the sake of completeness, we state Kolmogorov's *non-constructive* proof for the existence of an absolute probability sequence as follows. View $\Phi_{ij}(t,\tau)$ in (50) as transition probabilities of a Markov chain in *reverse* time. In other words, view sequence $\{\pi(t)\}$ as a backward sequence of stochastic vectors that starts at $t = \infty$ and ends at $t = 0$. Note that we still do not know that if such a backward sequence satisfying (50) exists at all. However, we do know that for any fixed arbitrary $t_0 \geq 0$, there exists a backward sequence of stochastic vectors initialized at t_0 satisfying (50) for every t, τ such that $t_0 \geq t \geq \tau \geq 0$. Thus, let $\{\pi^{(t_0)}(\tau)\}_{t_0 \geq \tau \geq 0}$ represent such a backward sequence. Indeed, the choice of initial probabilities vector $\pi^{(t_0)}(t_0)$ is arbitrary as long as it is a stochastic vector. Now, given sequential compactness and using a diagonal argument, we can progressively extract based on the infinite vector sequences associated with $t_0 = 1, 2, 3, \ldots$, a gradually thinning infinite subsequence of time indices $\lambda_1, \lambda_2, \lambda_3, \ldots$, where $\lim_{n\to\infty} \lambda_n = \infty$, such that for any $i = 1, \ldots, N$ and $\tau \geq 0$, $\lim_{n\to\infty} \pi_i^{(\lambda_n)}(\tau)$ exists. Defining $\pi_i(\tau)$ as this limit value for every i, an absolute probability sequence is obtained.

We now state a continuous time version of [10, Def. 3].

Definition 10 (Class $\mathcal{P}^*$, Continuous Time Version): A chain $\{A(t)\}$ is said to be in Class $\mathcal{P}^*$ if its associated state transition matrix $\Phi(t,\tau)$, $t \geq \tau \geq 0$, admits an absolute probability sequence $\{\pi(t)\}$ for which there exists a constant $p^* > 0$ such that $\pi(t) > p^*$ for any $t \geq 0$.

It is possible to characterize chains of Class $\mathcal{P}^*$ more concretely. To do so, we state the following two lemmas.

Lemma 7: For every $\tau \geq 0$ and $j \in \mathcal{V}$

$$\pi_j(\tau) \leq \inf_{t\geq\tau} \left\{ \sum_{i\in\mathcal{V}} \Phi_{i,j}(t,\tau) \right\}.$$

Proof: Obvious, since for every $t \geq \tau$

$$\pi_j(\tau) = \pi'(t)\Phi_j(t,\tau) = \sum_{i\in\mathcal{V}} \pi_i(t)\Phi_{i,j}(t,\tau) \leq \sum_{i\in\mathcal{V}} \Phi_{i,j}(t,\tau).$$

■

Lemma 8: A chain $\{A(t)\}$ is in Class $\mathcal{P}^*$ if and only if for its state transition matrix $\Phi(t,\tau)$, $t \geq \tau \geq 0$, and any $j \in \mathcal{V}$

$$\inf_{t,\tau} \left\{ \sum_{i\in\mathcal{V}} \Phi_{i,j}(t,\tau) \mid t \geq \tau \geq 0 \right\} > 0. \tag{51}$$

Proof: The *only if* part is an immediate result of Lemma 7. To prove the *if* part, let $p^* > 0$ be $1/N$ times the value of the left-hand side of inequality (51). We now take advantage of the way an absolute probability sequence can be obtained in [27] as explained in the beginning of this section. Remembering that the choice of $\pi^{(t_0)}(t_0)$ is arbitrary, set $\pi_i^{(t_0)}(t_0) = 1/N$ for any $i \in \mathcal{V}$ and $t_0 \geq 0$. Subsequently, $\pi_i^{(\lambda_n)}(\lambda_n) = 1/N$ for every n. Notice also by definition that $(\pi^{(\lambda_n)}(\tau))' = (\pi^{(\lambda_n)}(\lambda_n))'$ $\Phi(\lambda_n,\tau)$ for every n and $0 \leq \tau \leq \lambda_n$. Thus

$$\pi_i^{(\lambda_n)}(\tau) = \frac{1}{N}\sum_{j=1}^{N} \Phi_{ji}(\lambda_n,\tau)$$

and consequently

$$\pi_i(\tau) = \lim_{n\to\infty} \pi_i^{(\lambda_n)}(\tau) = \lim_{n\to\infty} \frac{1}{N}\sum_{j=1}^{N} \Phi_{ji}(\lambda_n,\tau)$$
$$\geq \frac{1}{N}\inf_{t,\tau} \left\{ \sum_{i\in\mathcal{V}} \Phi_{i,j}(t,\tau) \mid t \geq \tau \geq 0 \right\} = p^*.$$

■

Lemma 8 roughly implies that the coupling chain of a system is in Class $\mathcal{P}^*$, if and only if the opinion of any individual at any time continues to have influence on the formation of individuals' opinions at all future times. According to [11, Theor. 6]), every chain $\{A(t)\}$ in Class $\mathcal{P}^*$ is class-ergodic, while the number of ergodic classes is equal to the number of connected components of the infinite flow graph of $\{A(t)\}$. Therefore, if chain $\{A(t)\}$ is in Class $\mathcal{P}^*$, the upper bound provided in Theorem 3 for the size of a minimal EGC of a network with dynamics (1) is equal to the lower bound provided in Corollary 1, leading to the following corollary.

Corollary 3: If $\{A(t)\}$ is in Class $\mathcal{P}^*$, the size of a minimal EGC of a network with dynamics (1) is the number of connected components of the infinite flow graph associated with $\{A(t)\}$.

VIII. Full-Rank Chains

Theorem 5: A chain $\{A(t)\}$ is *full-rank*, that is, $\mathbf{rank}(A) = N$ if and only if $\{A(t)\}$ is an l_1-approximation of the neutral chain, that is, the chain of matrix $\mathbf{0}$.

Proof: The sufficiency is immediately implied using Proposition 1 and taking into account that the neutral chain is full-rank. To prove the necessity, assume that $\mathbf{rank}(A) = N$. Recalling the geometric framework developed in Section III, limiting polytope C_0 must have N vertices. Letting $v_1, \ldots, v_N$ be the N vertices of C_0, for a permutation σ over $\{1, \ldots, N\}$, we must have

$$\lim_{t\to\infty} \Phi'(t,0) = \left[v_{\sigma(1)}|\cdots|v_{\sigma(N)}\right] \tag{52}$$

since each column of $\Phi'(t,0)$ is a continuous function of t such that its distance from $\{v_1, \ldots, v_N\}$ vanishes as t grows large. From the Jacobi-Liouville formula [28] and using (52), one concludes

$$\begin{aligned}\lim_{t\to\infty} \exp\left[\int_0^t \mathrm{tr}\left(A(\tau)\right) d\tau\right] &= \lim_{t\to\infty} \det\left(\Phi'(t,0)\right)\\ &= \det\left(\lim_{t\to\infty}\Phi'(t,0)\right)\\ &= \det\left(\left[v_{\sigma(1)}|\cdots|v_{\sigma(N)}\right]\right) \end{aligned} \tag{53}$$

where $\mathrm{tr}(.)$ and $\det(.)$ denote *trace* and *determinant* operators. respectively. Since $v_1, \ldots, v_N$ are linearly independent, from (53), one concludes that $\int_0^t \mathrm{tr}(A(\tau))d\tau$, which is always negative and decreasing, remains bounded as t grows large. Taking into account that diagonal elements of $A(\tau)$ are its only negative elements, one further obtains $\int_0^\infty \|A(t)\|_\infty dt < \infty$, which completes the proof. ■

Assume that the infinite flow graph of chain $\{A(t)\}$, that is, $\mathcal{H}_2(\mathcal{V}, \mathcal{E}_2)$, has h_2-connected components. Form chain $\{B(t)\}$, which is an l_1-approximation of $\{A(t)\}$, by eliminating all interactions between distinct connected components. Since the subchain corresponding to each connected component is full-rank if and only if it contains a single node, the following proposition follows from Proposition 1, that provides an upper bound for $\mathbf{rank}(A)$.

Proposition 2: Let $\{A(t)\}$ be a time-varying chain with infinite flow graph $\mathcal{H}_2$. Then, $\mathbf{rank}(A) \le N - h_2'$, where h_2' is the number of connected components of $\mathcal{H}_2$ containing two or more nodes.

IX. Conclusion

We considered a network of multiple individuals with opinions updated via a general time-varying continuous time linear algorithm. The notion of EGC, an acronym associated with ÉGC, in the network was defined as follows. An EGC is a subgroup of individuals who can cooperatively manage to create a global consensus on any desired opinion in the network only by adequately setting their initial opinions assuming that they are cognizant of the coupling chain of the network as well as the rest of individuals initial opinions. The size of a minimal EGC can be treated as a characteristic of the coupling chain of the network. We then introduced an extension of the notion of rank, from an individual matrix-related notion to one related to a Markov chain in continuous time. A key result is that the rank of the coupling chain of a network is also the size of its minimal EGC. Geometrically and associated with the chain, one can define a monotone decreasing sequence of convex hulls (polytopes) generated by an underlying sequence of vertices. The rank of the chain is the number of linearly independent vertices in the sequence of polytopes.

A collection of upper and lower bounds on the rank was also established that helped determine the rank for both time invariant chains (possibly not in Class $\mathcal{P}^*$), as well as for Class $\mathcal{P}^*$ chains in the time inhomogeneous case.

From a practical standpoint, this work establishes the rather intuitive result that the less "natural" dissension that exists in an opinion network, the easier it is to steer the network toward global agreement. In cases where an "average" amount of natural dissonance exists, then the theory points to the need to minimally "infiltrate" identifiable dissenting clusters and work from the inside so to speak to steer the global opinion to a consensus. Success in doing so hinges on an ability to enlist key individuals' cooperation given that they must act as a "grand coalition" of key individuals. This, in turn, opens the door to games over opinion networks where key individuals might choose to break up into smaller coalitions and work toward conflicting goals. This will be the subject of future work. Finally, developing simple algorithms to identify key individuals in an opinion network is another direction for future research.

Acknowledgment

The authors would like to thank the anonymous reviewers for their comments and insights.

References

[1] G. Flierl, D. Grünbaum, S. Levins, and D. Olson, "From individuals to aggregations: The interplay between behavior and physics," *J. Theor. Biol.*, vol. 196, no. 4, pp. 397–454, 1999.
[2] F. Cucker and S. Smale, "Emergent behavior in flocks," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 852–862, May 2007.
[3] J. N. Tsitsiklis *et al.*, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. 31, no. 9, pp. 803–812, Sep. 1986.
[4] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. AC-48, no. 6, pp. 988–1001, Jun. 2003.
[5] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis, simulation," *J. Artif. Soc. Social Simul.*, vol. 5, no. 3, pp. 1–33, 2002.
[6] N. A. Lynch, *Distributed Algorithms*. San Mateo, CA, USA: Morgan Kaufmann, 1996.
[7] M. H. DeGroot, "Reaching a consensus," *J. Amer. Stat. Assoc.*, vol. 69, no. 345, pp. 118–121, 1974.
[8] S. Chatterjee and E. Seneta, "Towards consensus: Some convergence theorems on repeated averaging," *J. Appl. Probability*, vol. 14, no. 1, pp. 89–97, 1977.
[9] J. M. Hendrickx and J. N. Tsitsiklis, "Convergence of type-symmetric and cut-balanced consensus seeking systems," *IEEE Trans. Autom. Control*, vol. 58, no. 1, pp. 214–218, Jan. 2013.
[10] B. Touri and A. Nedić, "Product of random stochastic matrices," *IEEE Trans. Autom. Control*, vol. 59, no. 2, pp. 437–448, Feb. 2014.
[11] S. Bolouki and R. P. Malhamé, "Consensus algorithms and the decomposition-separation theorem," in *Proc. IEEE 52nd Annu. CDC.*, 2013, pp. 1490–1495.
[12] J. Lorenz and D. Urbig, "About the power to enforce and prevent consensus by manipulating communication rules," *Adv. Complex Syst.*, vol. 10, no. 2, pp. 251–269, 2007.
[13] G. Weisbuch, G. Deffuant, F. Amblard, and J.-P. Nadal, "Meet, discuss, segregate!" *Complexity*, vol. 7, no. 3, pp. 55–63, 2002.
[14] H. G. Tanner, "On the controllability of nearest neighbor interconnections," in *Proc. IEEE 43rd Annu. Conf. Dec. Control*, 2004, pp. 2467–2472.
[15] A. Rahmani, M. Ji, M. Mesbahi, and M. Egerstedt, "Controllability of multi-agent systems from a graph-theoretic perspective," *SIAM J. Control Optimiz.*, vol. 48, no. 1, pp. 162–186, 2009.

[16] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.

[17] N. Monshizadeh, S. Zhang, and M. K. Camlibel, "Zero forcing sets and controllability of dynamical systems defined on graphs," *IEEE Trans. Autom. Control*, vol. 59, no. 9, pp. 2562–2567, Sep. 2014.

[18] W. Ni, X. Wang, and C. Xiong, "Consensus controllability, observability and robust design for leader-following linear multi-agent systems," *Automatica*, vol. 49, no. 7, pp. 2199–2205, 2013.

[19] J. L. Daleckii and M. G. Krein, *Stability of Solutions of Differential Equations in Banach Space. Translations of Mathematical Monographs vol. 43*. Providence, RI, USA: American Mathematical Soc., 1974.

[20] S. Bolouki and R. P. Malhamé, "Consensus algorithms and the decomposition-separation theorem," in *Proc. IEEE 52nd Annu. Conf. Dec. Control*, 2013, pp. 1490–1495.

[21] B. Touri and A. Nedić, "On backward product of stochastic matrices," *Automatica*, vol. 48, no. 8, pp. 1477–1488, 2012.

[22] B. Touri and A. Nedić, "On approximations and ergodicity classes in random chains," *IEEE Trans. Autom. Control*, vol. 57, no. 11, pp. 2718–2730, Nov. 2012.

[23] J. Shen, "A geometric approach to ergodic non-homogeneous markov chains," in *Lecture Notes in Pure and Applied Mathemathics*. New York, NY, USA: Marcel Dekker, 2000, pp. 341–366.

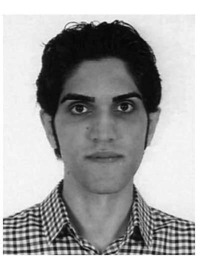
[24] S. Bolouki and R. Malhamé, "Ergodicity and class-ergodicity of balanced asymmetric stochastic chains," in *Proc. Eur. Control Conf.*, 2013, pp. 221–226.

[25] B. Touri and A. Nedić, "On ergodicity, infinite flow, consensus in random models," *IEEE Trans. Autom. Control*, vol. 56, no. 7, pp. 1593–1605, Jul. 2011.

[26] P. Chebotarev and R. Agaev, "Forest matrices around the Laplacian matrix," *Linear Algebra Appl.*, vol. 356, no. 1–3, pp. 253–274, 2002.

[27] A. Kolmogoroff, "On the theory of markov chains," in *Selected Works of A. N. Kolmogorov*, vol. 26. New York, NY, USA: Springer, 1992, pp. 182–187.

[28] C. G. J. Jacobi, Gesammelte Werke, vol. 1. Berlin, Germany: Chelsea, 1969.



Sadegh Bolouki (M'14) received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2008 and the Ph.D. degree in electrical engineering from École Polytechnique de Montréal, Montréal, QC, Canada, in 2014.

From 2014 to 2015, he was a Research Scholar with the Department of Mechanical Engineering and Mechanics at Lehigh University, Bethlehem, PA, USA. Since 2015, he has been a Postdoctoral Scholar at the Coordinated Science Lab at the University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA. His research interests include the areas of decentralized and distributed control, opinion dynamics, consensus, and game theory.

Roland P. Malhamé (S'82–M'92) received the B.Sc. degree in electrical engineering from the American University of Beirut, Beirut, Lebanon, in 1976, the M.Sc. degree in electrical engineering from the University of Houston, Houston, TX, USA, in 1978, and the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 1983.

After a single year, he was with the University of Quebec, Montreal, QC, Canada, and CAE Electronics Ltd., Montreal. In 1985, he joined École Polytechnique de Montréal, where he is Professor of Electrical Engineering. In 1994, 2004, and 2012, he was on sabbatical leave, respectively, with Laboratoire des Signaux & Systèms, Gif-sur-Yvette, France, and University of Rome Tor Vergata, Roma, Italy. His interest in statistical mechanics-inspired approaches to the analysis and control of large-scale systems led him to contributions in the area of aggregate electric load modeling, and to the early developments of the theory of mean field games. His current research interests are in collective decentralized decision-making schemes, and the development of mean-field-based control algorithms in the area of smart grids. From 2005 to 2011, he headed GERAD, the Group for Research on Decision Analysis. He is an Associate Editor of the *International Transactions on Operations Research*.

Milad Siami (S'12) received the B.Sc. degree in electrical engineering and in pure mathematics and the M.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2009 and 2011, respectively, the M.Sc. degree in mechanical engineering from Lehigh University, Bethlehem, PA, USA, in 2014, where he is currently pursuing the Ph.D. degree in mechanical engineering and mechanics.

From 2009 to 2010, he was a Research Student in the Department of Mechanical and Environmental Informatics at the Tokyo Institute of Technology, Tokyo, Japan. His research interests include distributed control systems, distributed optimization, and applications of fractional calculus in engineering.

Nader Motee (S'99–M'08–SM'13) received the B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2000, and the M.Sc. and Ph.D. degrees in electrical and systems engineering from the University of Pennsylvania, Philadelphia, PA, USA, in 2006 and 2007, respectively.

From 2008 to 2011, he was a Postdoctoral Scholar with the Control and Dynamical Systems Department at the California Institute of Technology, Pasadena, CA, USA. Since 2011, he has been the P.C. Rossin Assistant Professor in the Department of Mechanical Engineering and Mechanics at Lehigh University, Bethlehem, PA. His current research area is distributed dynamical and control systems with a particular focus on issues related to sparsity, performance, and robustness.

Prof. Motee is a past recipient of several awards, including the 2008 AACC Hugo Schuck best paper award, the 2007 ACC best student paper award, the 2008 Joseph and Rosaline Wolf best thesis award, a 2013 Air Force Office of Scientific Research (AFOSR) Young Investigator Award, and a 2015 National Science Foundation Faculty Early Career Development (CAREER) Award.