

## Eigenvalue Invariance of Inhomogeneous Matrix Products in Distributed Algorithms

Shaoshuai Mou    Brian. D. O. Anderson

**Abstract**—This paper establishes a general theorem concerning the eigenvalue invariance of certain inhomogeneous matrix products with respect to changes of individual multiplicands' orderings. Instead of detailed entries, it is the zero-nonzero structure that matters in determining such eigenvalue invariance. The theorem is then applied in analyzing the convergence rate of a distributed algorithm for solving linear equations over networks modelled by undirected graphs.

### I. INTRODUCTION

Intensive research efforts have recently been devoted to studying distributed control of networked systems, the goal of which is to achieve global objectives through local coordination among connected subsystems [1]–[3]. Networked systems operating in discrete time may often lead to a mathematical model involving inhomogeneous products of matrices. Such inhomogeneity can result from, for example, change of network structure [4], [5], cyclic consideration (measurement or control) of subsystems, or communications among subsystems in the overall networked system [6], [7], or alternation of different operations, such as seeking local optimization involving one subsystem, followed by deliberate incorporation with another subsystem, then reversion to the local optimization using distributed optimization [8]–[10].

A common style of interaction between subsystems joined in a network involves gossiping, a process where, typically in a random sequence, individual pairs of systems exchange information, often with a view to establishing something like a common average of initially stored values. Deterministic periodic gossiping is also one distributed way of driving all nodes in a network to achieve a consensus that can be the global average of all nodes' initial state values, though on practical grounds random gossiping [11] may be preferred. Nevertheless, periodic gossiping [12] brings with it a deterministic convergence, which is determined by the second largest eigenvalue of an inhomogeneous product of matrices [13] (Such convergence could be further speeded up by a computable amount by allowing multi-gossiping corresponding to disjoint edges [14]). The convergence rate has been shown to be invariant with respect to the gossip orders for tree networks due to an invariant property of eigenvalues of a class of inhomogeneous products [15]–[17].

In a quite different direction, by a combination of the consensus idea with projections to linear constraints on nodes'

states, distributed algorithms have recently been developed for solving linear equations [18]–[22]. It is possible to use periodic gossiping as the tool for achieving consensus in these distributed algorithms, and it then transpires that a class of inhomogeneous products of matrices becomes relevant in considering the convergence. These products are inherently different from (and indeed structurally more complicated than) those arising in ordinary periodic gossiping aimed at achieving an average of initially stored variables. Surprisingly we have observed that the property of eigenvalue invariance still holds as long as the underlying network is a tree. Proof of this observation is nontrivial since previous results in [15]–[17] based on the details of matrix entries are not directly applicable. This motivates us to establish a general theorem of linear algebra about the invariance of eigenvalues of certain inhomogeneous matrix products with respect to changes of ordering of the individual multiplicands.

Accordingly in this paper, we seek to first explain the general theorem, and then show how it applies to the case of a decentralized discrete-time algorithm for solving  $Ax = b$ . The approach to proving the main linear algebra result is motivated by the proofs we found for a result on achieving consensus using periodic gossiping [15]–[17]. The proof in this paper involves delicately interrelating properties of graphs with properties of block matrices; however it does not draw on the detailed structure of the individual blocks. Any such details are in fact a distractor, though they become relevant in considering particular applications of the theorem, and are needed in computing convergence rates for infinite products.

The next section of the paper establishes the eigenvalue invariance property for inhomogeneous products. It is followed by sections treating the distributed solution of  $Ax = b$ . Simulations and conclusions complete the paper.

### II. EIGENVALUE INVARIANCE OF A CLASS OF INHOMOGENEOUS MATRIX PRODUCTS

Consider an  $n$ -node- $\bar{n}$ -edge *simple* graph  $\mathbb{E}$ , i.e. undirected with no self-loops or multiple edges. An  $n \times n$  block matrix  $Q \in \mathbb{R}^{nm \times nm}$  is called a *primitive matrix* for edge  $(i, j)$  in  $\mathbb{E}$  if each block in its diagonal is  $I_m$  except the  $i$ th and the  $j$ th, and each off-diagonal block is  $0_m$  except the  $ij$ th and the  $ji$ th. Note that in a primitive matrix for edge  $(i, j)$ , we do not put any restrictions on its  $i$ th,  $ij$ th,  $ji$ th and  $j$ th block entry. As will be shown later, the detailed contents of these block entries have no impact to the eigenvalue invariance property that we will establish.

Let  $\mathbf{E} = e_1, e_2, \dots, e_{\bar{n}}$  denote an ordered edge sequence. If  $e_i \neq e_j$  for any  $i \neq j$  in  $\mathbb{E}$  and  $\{e_1, e_2, \dots, e_{\bar{n}}\}$  is the edge set of  $\mathbb{E}$ , the ordered edge sequence  $\mathbf{E}$  is called *complete*.

S. Mou is with the School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47906. [mous@purdue.edu](mailto:mous@purdue.edu); his work is supported by a funding from Northrop Grumman Corporation. B. D. O. Anderson is with Hangzhou Dianzi University, Hangzhou, China, The Australian National University and Data-61 CSIRO (formerly NICTA), Canberra, ACT 2600 Australia, [Brian.Anderson@anu.edu.au](mailto:Brian.Anderson@anu.edu.au); his work is supported by Data-61, CSIRO, and by the Australian Research Council's Discovery Projects DP-130103610 and DP-160104500.

Correspondingly, the matrix product  $E = Q_{\bar{n}}Q_{\bar{n}-1}\cdots Q_1$  is called a *complete matrix* of  $\mathbb{E}$ , where  $Q_h$  is a primitive matrix for the edge  $e_h$ ,  $h = 1, 2, \dots, \bar{n}$ . Let  $\mathcal{G}(\mathbb{G})$  denote the set of complete matrices of any simple graph  $\mathbb{G}$ . By an *isolated node* of  $\mathbb{G}$  is meant any node in  $\mathbb{G}$  with degree zero; any node of  $\mathbb{G}$  with positive degree is *non-isolated*. If  $G \in \mathcal{G}(\mathbb{G})$  and  $i$  is an isolated node, then both the  $i$ th block row and the  $i$ th block column of  $G$  have  $I_m$  in the  $i$ th block entry and zeros elsewhere. By a *spanning subgraph* of  $\mathbb{G}$  is meant a graph with the same set of nodes as  $\mathbb{G}$  but a subset of edges of  $\mathbb{G}$  (Even an empty subset is allowed here).

A major result of this paper is the following theorem, which presents the invariance of characteristic polynomials of complete matrices under certain permutations.

*Theorem 1:* Let  $E = Q_{\bar{n}}Q_{\bar{n}-1}\cdots Q_1$  be a complete matrix of an  $n$ -node- $\bar{n}$ -edge simple graph  $\mathbb{E}$ , where  $Q_h$  is a primitive matrix for the edge  $e_h$  in  $\mathbb{E}$ ,  $h = 1, 2, \dots, \bar{n}$ . Let  $E_\pi = Q_{\pi(\bar{n})}Q_{\pi(\bar{n}-1)}\cdots Q_{\pi(1)}$ , where  $\pi : \{1, 2, \dots, \bar{n}\} \rightarrow \{1, 2, \dots, \bar{n}\}$  denotes a permutation map. Let  $cp(\cdot)$  denote the characteristic polynomial of a matrix. Then

$$cp(E) = cp(E_\pi)$$

if  $\pi$  satisfies one of the following conditions:

- 1)  $\pi$  is the identity map.
- 2)  $\pi$  is a cyclic permutation of  $\{1, 2, \dots, \bar{n}\}$
- 3)  $\pi$  interchanges  $h$  and  $h+1$  provided that  $e_h$  and  $e_{h+1}$  are disjoint, that is, not incident on the same node;
- 4)  $\pi$  interchanges  $h$  and  $h+1$  provided that  $e_h$  and  $e_{h+1}$  **are incident on the same node but neither edge is contained in any cycle of  $\mathbb{E}$ .**

To prove Theorem 1, we need the following two lemmas with proofs to be provided in the Appendix.

*Lemma 1:* Suppose that  $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_r$  are spanning subgraphs of a simple graph  $\mathbb{G}$  such that the union of  $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_r$  is  $\mathbb{G}$ ; and for any  $1 \leq i, j \leq r$ ,  $\mathbb{G}_i$  and  $\mathbb{G}_j$  have no non-isolated node in common. Then for any  $G \in \mathcal{G}(\mathbb{G})$ , there exists  $G_i \in \mathcal{G}(\mathbb{G}_i)$ ,  $i = 1, 2, \dots, r$  such that

$$G = G_{\pi(1)}G_{\pi(2)}\cdots G_{\pi(r)}$$

for any permutation map  $\pi : \{1, 2, \dots, r\} \rightarrow \{1, 2, \dots, r\}$ .

*Lemma 2:* Suppose that  $\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_r$  are spanning subgraphs of an  $n$ -node simple graph  $\mathbb{G}$ . Suppose node  $n$  is the only common non-isolated node of all  $\mathbb{G}_i$ ,  $i = 1, 2, \dots, r$ , and for any  $1 \leq i, j \leq r$ ,  $\mathbb{G}_i$  and  $\mathbb{G}_j$  have no other common non-isolated node except node  $n$ . Then for any permutation map  $\pi : \{1, 2, \dots, r\} \rightarrow \{1, 2, \dots, r\}$  and any  $G_i \in \mathcal{G}(\mathbb{G}_i)$ ,  $i = 1, 2, \dots, r$ , one has

$$cp(G_{\pi(r)}G_{\pi(r-1)}\cdots G_{\pi(1)}) = cp(G_rG_{r-1}\cdots G_1). \quad (1)$$

**Proof of Theorem 1:** It is easy to verify that Theorem 1 is true under Condition 1) or 2), and accordingly we omit the details of the proof. Next note that if two edges have no node in common, their corresponding primitive matrices commute, which implies Theorem 1 holds under condition 3). All we still need to do is to prove the theorem is true under condition 4) by showing  $cp(E) = cp(E_\pi)$ , where

$$E = Q_{\bar{n}}\cdots Q_{h+2}Q_{h+1}Q_hQ_{h-1}\cdots Q_1$$

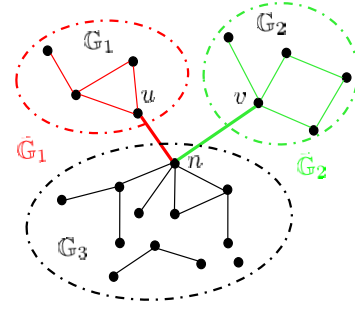


Fig. 1. Graphs that have the same set of nodes indicated by black dots but different edges:  $\mathbb{G}_1$  (with edges in the red circle),  $\mathbb{G}_2$  (with edges in the green circle),  $\mathbb{G}_3$  (with edges in the black circle),  $\tilde{\mathbb{G}}_1$  (with edges in red color), and  $\tilde{\mathbb{G}}_2$  (with edges in green color);

and

$$E_\pi = Q_{\bar{n}}\cdots Q_{h+2}Q_hQ_{h+1}Q_{h-1}\cdots Q_1.$$

Without losing any generality, and with inessential node renumbering if necessary we suppose  $Q_h$  and  $Q_{h+1}$  correspond to the edges  $(u, n)$  and  $(v, n)$ , respectively, of  $\mathbb{E}$ . Let  $\mathbb{G}$  denote a spanning subgraph of  $\mathbb{E}$  which results from deleting  $(u, n)$  and  $(v, n)$  from  $\mathbb{E}$ . As indicated in Fig.1, we let  $\mathbb{G}_1$  be the spanning subgraph of  $\mathbb{E}$  with edges indicated in red circle, which are edges of the connected component in  $\mathbb{G}$  containing node  $u$ ; let  $\mathbb{G}_2$  denote the spanning subgraph of  $\mathbb{E}$  with edges indicated in the green circle, which are those of the connected component in  $\mathbb{G}$  containing node  $v$ ; let  $\mathbb{G}_3$  denote the spanning subgraph of  $\mathbb{E}$  with edges indicated in the black circle. Here,  $\mathbb{G}_3$  is obtained by deleting  $(u, n)$ ,  $(v, n)$  and all edges belonging to  $\mathbb{G}_1$  and  $\mathbb{G}_2$ . Then  $\mathbb{G}$  is the union of  $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3$ . Note that  $\mathbb{G}_1, \mathbb{G}_2$  are necessarily connected while  $\mathbb{G}_3$  may not be. Let  $\tilde{\mathbb{G}}_1, \tilde{\mathbb{G}}_2$  denote the spanning subgraphs of  $\mathbb{E}$ , with edge set of  $\mathbb{G}_1$  plus  $(u, n)$  and edge set of  $\mathbb{G}_2$  plus  $(v, n)$ , respectively. Then  $\mathbb{E}$  is the union of  $\tilde{\mathbb{G}}_1, \tilde{\mathbb{G}}_2, \mathbb{G}_3$ .

By cyclic permutations to the matrix product  $E$ , one can construct a matrix product  $M$ , where

$$M = Q_{h-1}\cdots Q_1Q_{\bar{n}}Q_{\bar{n}-1}\cdots Q_{h+2} \quad (2)$$

Then

$$cp(E) = cp(Q_{h+1}Q_hM).$$

Note that  $M \in \mathcal{G}(\mathbb{G})$ ,  $\mathbb{G}$  is the union of  $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3$ , and  $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3$  are spanning subgraphs of  $\mathbb{E}$ . Since  $(u, n)$  and  $(v, n)$  are not contained in any cycle in  $\mathbb{E}$ , then any two of  $\{\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3\}$  do not have a non-isolated node in common. By Lemma 1, we have

$$M = G_2G_1G_3 \quad (3)$$

where  $G_1 \in \mathcal{G}(\mathbb{G}_1)$ ,  $G_2 \in \mathcal{G}(\mathbb{G}_2)$ ,  $G_3 \in \mathcal{G}(\mathbb{G}_3)$ . Then

$$cp(E) = cp(Q_{h+1}Q_hG_2G_1G_3). \quad (4)$$

Note that  $(u, n)$  and any edge of  $\mathbb{G}_2$  are disjoint. Thus  $Q_h$  commutes with  $G_2$ . It follows that

$$cp(E) = cp(Q_{h+1}G_2Q_hG_1G_3). \quad (5)$$

Observe that  $Q_h G_1 \in \mathcal{G}(\bar{\mathbb{G}}_1)$ ,  $Q_{h+1} G_2 \in \mathcal{G}(\bar{\mathbb{G}}_2)$ , and any two of  $\bar{\mathbb{G}}_1, \bar{\mathbb{G}}_2, \bar{\mathbb{G}}_3$  only have  $n$  as the non-isolated node in common. Thus by Lemma 2, we have

$$cp(E) = cp(Q_h G_1 Q_{h+1} G_2 G_3). \quad (6)$$

Note that  $Q_{h+1}$  commutes with  $G_1$  since  $(v, n)$  and any edge of  $\bar{\mathbb{G}}_1$  are disjoint; and  $G_2$  commutes with  $G_1$  since any edge in  $\bar{\mathbb{G}}_1$  and any one in  $\bar{\mathbb{G}}_2$  are disjoint. Hence, we have

$$cp(E) = cp(Q_h Q_{h+1} G_2 G_1 G_3) \quad (7)$$

which by (3) leads to

$$cp(E) = cp(Q_h Q_{h+1} M) \quad (8)$$

with  $M$  as defined in (2). Again by cyclic permutation to  $Q_h Q_{h+1} M$  at the right-hand side of (8), one has

$$cp(E) = cp(Q_{\bar{n}} \cdots Q_{h+2} Q_h Q_{h+1} Q_{h-1} \cdots Q_1) = cp(E_\pi).$$

This completes the proof. ■

### III. A DISTRIBUTED ALGORITHM FOR SOLVING LINEAR EQUATIONS

The purpose of this section is to illustrate that a distributed, periodic, iterative algorithm for solving linear equations which rests on a gossiping framework can be obtained. Moreover, when an undirected tree graph underpins the algorithm, a very similar result to our previous gossiping results [15]–[17] can be obtained on the independence of the rate of convergence to the ordering of the individual steps in one period. Additionally, assuming a periodic algorithm is used, its rate of convergence can be obtained. Knowing a deterministic rate of convergence is of course considered advantageous, sometimes even *de rigueur*, in any discussion of iterative linear algebra algorithms which do not converge in a finite number of steps. It would clearly be harder to ‘sell’ an algorithm using probabilistic selection of individual steps to a potential user, given the inability to make any statement guaranteeing the rate of convergence in any one instance.

#### A. Outline of the algorithm

Consider the equation  $A\bar{x} = b$ ,  $A \in \mathbb{R}^{\bar{m} \times m}$ ,  $b \in \mathbb{R}^{\bar{m}}$ , and assume that a unique solution exists. Of course, this means that  $\bar{m} \geq m$  and  $A$  is with full column rank. We arbitrarily partition the matrix  $A$  into a number of  $n$  blocks, say  $A = [A_1 \ A_2 \ \cdots \ A_n]^\top$  where  $A_i^\top$  has  $m_i$  rows and  $m_1 + m_2 + \cdots + m_n = \bar{m}$ . The vector  $b$  is correspondingly partitioned, with  $b_i$  denoting now the  $i$ -th subvector of dimension  $m_i$ . A straightforward partition is by using blocks with one row. Consider also a connected and simple graph  $\mathbb{E}$  with  $n$  nodes and  $\bar{n}$  edges. Associate with node  $i$  an  $m$ -vector  $x_i$ , initialized so that there holds  $A_i^\top x_i(0) = b_i$ . Suppose that a single step of the algorithm can be associated with a particular edge of the graph, say edge  $h$  adjacent to nodes  $i$  and  $j$ . In this step, a *combined projection-gossip operation* is undertaken. The gossip component is

$$\begin{aligned} x_{ig}(k+1) &= \frac{1}{2}[x_i(k) + x_j(k)] \\ x_{jg}(k+1) &= \frac{1}{2}[x_i(k) + x_j(k)] \end{aligned}$$

and the projection component serves to project  $x_{ig}(k+1), x_{jg}(k+1)$  onto, respectively, the two affine spaces  $A_i^\top x_i = b_i, A_j^\top x_j = b_j$ . Let  $P_i \in \mathbb{R}^{m \times m}$  denote the projection matrix to the kernel of  $A_i$ . One has the following combined projection-gossip operation for updating states of node  $i$  and  $j$  given by

$$\begin{aligned} x_i(k+1) &= x_i(k) + P_i[x_{ig}(k+1) - x_i(k)] \\ &= (I_m - \frac{1}{2}P_i)x_i(k) + \frac{1}{2}P_i x_j(k) \end{aligned} \quad (9)$$

$$\begin{aligned} x_j(k+1) &= x_j(k) + P_j[x_{jg}(k+1) - x_j(k)] \\ &= \frac{1}{2}P_j x_i(k) + (I_m - \frac{1}{2}P_j)x_j(k); \end{aligned} \quad (10)$$

and leaving all other nodes’ states unchanged, namely

$$x_l(k+1) = x_l(k), \quad l = 1, 2, \dots, n, l \neq i, l \neq j \quad (11)$$

We have just described a single step of the algorithm corresponding to one edge in  $\mathbb{E}$ . These steps could be put together in the order defined by a periodic, ordered, infinite sequence  $\mathbf{E}, \mathbf{E}, \mathbf{E}, \dots$ , in which  $\mathbf{E} = e_1, e_2, \dots, e_{\bar{n}}$  is an ordered sequence of all the edges in  $\mathbb{E}$ , with each edge appearing just once. Then we are led to a periodic algorithm and over one period, the combined projection-gossip steps occur in the order defined by  $\mathbf{E}$ .

#### B. Convergence of the algorithm

In this section, we show that the algorithm in question converges by using cyclic projections. We claim:

**Theorem 2:** Given a linear equation  $A\bar{x} = b$ , with  $A \in \mathbb{R}^{\bar{m} \times m}$  and  $b \in \mathbb{R}^{\bar{m}}$ , which has a unique solution. Then the algorithm just described above on a connected and undirected graph  $\mathbb{E}$  with  $n$  nodes and  $\bar{n}$  edges results in  $x_i(k)$  for all  $i$  approaching a common limit, viz.  $\bar{x}$ , as  $k \rightarrow \infty$

**Proof of Theorem 2:** Let

$$x(k) = [x_1(k)^\top, x_2(k)^\top, \dots, x_n(k)^\top]^\top \in \mathbb{R}^{mn}.$$

For any edge  $h$  that is incident on nodes  $i, j$ , we define an operation  $\mathcal{S}_h$  on  $x(k)$  which replaces the subvectors  $x_i(k)$  and  $x_j(k)$  by their average, leaving the other subvectors unchanged. Notice that this operation is actually a projection. Also, define an operation  $\mathcal{P}_h$  on  $x(k)$  which replaces the subvectors  $x_i(k), x_j(k)$  by their projections on to the affine spaces  $A_i^\top x_i = b_i, A_j^\top x_j = b_j$ , again leaving the other subvectors unchanged. Notice that this operation is also a projection. Finally notice that one step of the algorithm, corresponding to a combined projection-gossip operation associated with edge  $h$ , consists of application of the composite operator  $\mathcal{P}_h \mathcal{S}_h$  to  $x(k)$ . It is evident that over one period of the algorithm, which has  $\bar{n}$  steps in it, corresponding to edges  $e_1, e_2, \dots, e_{\bar{n}}$ , there results

$$x(k + \bar{n}) = \mathcal{P}_{\bar{n}} \mathcal{S}_{\bar{n}} \mathcal{P}_{\bar{n}-1} \mathcal{S}_{\bar{n}-1} \dots \mathcal{P}_1 \mathcal{S}_1 x(k) \quad (12)$$

It is standard, see [23], that the vector resulting from repeated iterations of such cyclic projections, converges exponentially fast to a point which lies in the intersection of all the projected sets, i.e. in the set

$$\{\cap \mathcal{P}_h \mathbb{R}^{mn}\} \cap \{\cap \mathcal{S}_h \mathbb{R}^{mn}\}$$

The connectedness of  $\mathbb{E}$  ensures that  $\cap \mathcal{S}_h \mathbb{R}^{mn} = \{x = [x_1^\top \ x_2^\top \ \dots \ x_n^\top]^\top \in \mathbb{R}^{mn}\}$  with  $x_1 = x_2 = \dots = x_n$ . Call the common value  $x_*$ . Then the requirement that  $x = [x_*^\top \ x_*^\top \ \dots \ x_*^\top]^\top \in \mathcal{P}_h \mathbb{R}^{mn}$  for all  $h$  implies that  $A_i^\top x_* = b_i$  for all  $i$ , i.e.  $x_* = \bar{x}$ . ■

### C. Invariance of the Convergence Rate

We will now argue that the main result Theorem 1 can effectively be extended to capture algorithms of the form just described. To this end, recall that the equations (9)-(11) capture what happens for a projection-gossip operation on edge  $h$  at time  $k$ , which leads to the corresponding update in  $x(k)$  given by

$$x(k+1) = Q_h x(k) \quad (13)$$

Here  $Q_h$  denotes an  $n \times n$  block matrix with entries of size  $m \times m$  and whose entries in blocks  $ii, ij, ji, jj$  are  $I_m - \frac{1}{2}P_i, \frac{1}{2}P_i, \frac{1}{2}P_j, I_m - \frac{1}{2}P_j$ ; the other entries of  $Q_h$  are  $0_m$  off the diagonal and  $I_m$  on the diagonal. Thus  $Q_h$  is a primitive matrix corresponding to the edge  $h$  in  $\mathbb{E}$ . For the periodic algorithm just described above with edge sequence  $\mathbf{E} = e_1, e_2, \dots, e_{\bar{n}}$  in one period, one could define

$$T_{\mathbf{E}} = Q_{\bar{n}} Q_{\bar{n}-1} \dots Q_1. \quad (14)$$

Then the convergence rate of the periodic algorithm is determined by the second largest eigenvalue of  $T_{\mathbf{E}}$  by [23]. For an arbitrary permutation  $\pi : \{1, 2, \dots, \bar{n}\} \rightarrow \{1, 2, \dots, \bar{n}\}$ , one defines

$$T_{\mathbf{E}_\pi} = Q_{\pi(\bar{n})} Q_{\pi(\bar{n}-1)} \dots Q_{\pi(1)} \quad (15)$$

Notice that  $T_{\mathbf{E}}$  and  $T_{\mathbf{E}_\pi}$  are all complete matrices. The eigenvalues of  $T_{\mathbf{E}}$  and  $T_{\mathbf{E}_\pi}$  are the same if the permutation is of certain types. More specifically, with these definitions and Theorem 1, we have:

*Theorem 3:* Consider the algorithm of the type described above for solving the equation  $A\bar{x} = b$  that is known to have a unique solution. Suppose the associated connected and simple graph  $\mathbb{E}$  has  $n$  nodes and  $\bar{n}$  edges. Let  $\mathbf{E}$  be an ordered edge sequence of length  $\bar{n}$  that defines a complete matrix  $T_{\mathbf{E}}$  as in (14). Consider the group  $\Pi(\mathbf{E})$  (under composition) of permutations  $\pi : \{1, 2, \dots, \bar{n}\} \rightarrow \{1, 2, \dots, \bar{n}\}$  generated by permutations of the four types listed in the statement of Theorem 1. Let  $T_{\mathbf{E}_\pi}$  be as defined in (15). Then for any  $\pi$  in  $\Pi(\mathbf{E})$ ,  $T_{\mathbf{E}}$  and  $T_{\mathbf{E}_\pi}$  have the same eigenvalues.

*Remark 1:* Obviously if the underlying graph  $\mathbb{E}$  is a tree, the convergence rate of the periodic algorithm described above will be the same irrespective of the order of the combined projection-gossip steps. This invariance property implies that the convergence rate of such periodic algorithms will be purely determined by the underlying tree topology, which will motivate a control engineer to figure out the optimal convergence rate by investigating the underlying tree topologies.

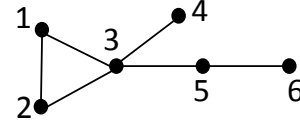


Fig. 2. A six-node simple graph  $\mathbb{E}$

## IV. SIMULATIONS

In this section, we will perform numerical simulations on a six-node simple graph as indicated in Fig. 2 for the validation of Theorem 1- Theorem 3. Let

$$A = \begin{bmatrix} 0.4249 & 0.0741 & 0.1425 & 0.9394 & 0.0684 & 0.9616 \\ 0.1192 & 0.3939 & 0.2681 & 0.2212 & 0.4363 & 0.7624 \\ 0.4951 & 0.0034 & 0.1749 & 0.4827 & 0.1739 & 0.0073 \\ 0.7064 & 0.2207 & 0.1386 & 0.3760 & 0.0261 & 0.6800 \\ 0.2436 & 0.0013 & 0.5989 & 0.5238 & 0.9547 & 0.7060 \\ 0.7851 & 0.1892 & 0.9011 & 0.2649 & 0.4306 & 0.6451 \end{bmatrix}$$

and

$$b = [0.5523 \quad 0.2181 \quad 0.7724 \quad 0.2280 \quad 0.3709 \quad 0.8909]^\top$$

whose entries are picked uniformly from the interval  $[0, 1]$ . The linear equation  $A\bar{x} = b$  has a unique solution

$$\bar{x} = [-0.1106 \quad 1.3378 \quad 1.3245 \quad 1.3965 \quad -0.4357 \quad -1.0094]^\top$$

Suppose each node  $i$  knows the  $i$ th row of  $A$  and  $b$ ,  $i = 1, 2, \dots, 6$ , and updates its state vector  $x_i(k)$  according to the periodic algorithm for solving linear equations described in the paper, in which the sequence in one period is

$$\mathbf{E} = (1, 2), (2, 3), (3, 1), (3, 4), (3, 5), (5, 6).$$

Let  $V(k) = \frac{1}{2} \sum_{i=1}^6 |x_i(k) - \bar{x}|^2$  measure the closeness between all  $x_i(t)$  and the unique solution  $x^*$ . Simulations in MATLAB suggest that  $V(k)$  converges to 0 exponentially fast as indicated in Fig. 3. Then all  $x_i(t)$  converge to  $\bar{x}$  exponentially fast, which agrees with Theorem 2.

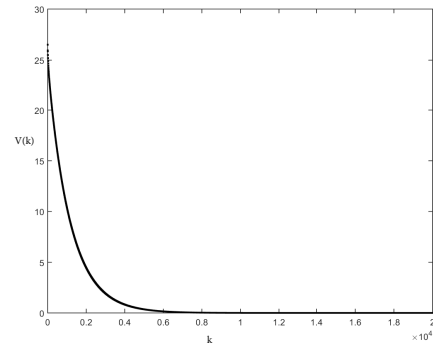


Fig. 3. Simulations show the exponential convergence of  $V(k)$  to 0.

Let  $T_{\mathbf{E}}$  denote the complete matrix defined by the edge sequence  $\mathbf{E}$ , that is,  $T_{\mathbf{E}} = Q_{56} Q_{35} Q_{43} Q_{13} Q_{23} Q_{12}$ . The convergence rate of  $x_i(t)$  to  $\bar{x}$  is determined the second largest eigenvalue of  $T_{\mathbf{E}}$ . Let  $\bar{\mathbf{E}}$  denote the complete matrix defined by the edge sequence

$$\bar{\mathbf{E}} = (1, 2), (2, 3), (3, 1), (3, 5), (3, 4), (5, 6).$$

Simulations in MATLAB suggest  $cp(T_{\mathbb{E}}) = cp(T_{\bar{\mathbb{E}}})$ . This is because the edge sequence  $\bar{\mathbb{E}}$  could be achieved by permute edges (3, 5), (3, 4) in  $\mathbb{E}$ , where (3, 5) and (3, 4) has one common adjacent node but neither of them is in any cycle in the graph  $\mathbb{E}$ .

## V. CONCLUSION

In this paper we have established a general theorem about the eigenvalues invariance of certain inhomogeneous matrix products with respect to changes of individual multiplicands' orderings. Each matrix in the product corresponding to an edge in a simple graph is a block matrix with no restriction to the details of blocks. We then present a periodic algorithm for solving linear equations, and employ the general theorem to achieve invariance of the convergence rate of the proposed distributed algorithm.

## APPENDIX

**Proof of Lemma 1:** Since for any  $1 \leq i, j \leq r$ ,  $\mathbb{G}_i$  and  $\mathbb{G}_j$  have no non-isolated node in common, any edge in  $\mathbb{G}_i$  and any edge in  $\mathbb{G}_j$  are not adjacent to the same node. Correspondingly, a primitive matrix corresponding to any edge in  $\mathbb{G}_i$  commutes with any primitive matrix corresponding to any edge in  $\mathbb{G}_j$  for any  $1 \leq i, j \leq r$ . In other words, two primitive matrices commute when they belong to any two different spanning subgraphs from  $\{\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_m\}$ . Thus for any  $G \in \mathcal{G}(\mathbb{G})$ , there **exists** a permutation map  $\pi : \{1, 2, \dots, r\} \rightarrow \{1, 2, \dots, r\}$  such that

$$G = G_{\pi(1)} G_{\pi(2)} \cdots G_{\pi(r)}$$

where  $G_i \in \mathcal{G}(\mathbb{G}_i)$ . Note further that  $G_i \in \mathcal{G}(\mathbb{G}_i)$  commutes with  $G_j \in \mathcal{G}(\mathbb{G}_j)$  when  $i \neq j$ . Thus  $\pi$  could be **any** permutation map. ■

Before proving Lemma 2, we will show how to apply the classical systems and control concept of a matrix transfer function to describe certain properties of graphs, and to explain how the transfer function associated with a union of graphs under certain conditions can be related to the transfer functions of the individual graphs. For any real  $(p+1)m \times (p+1)m$  matrix  $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$  with  $D \in \mathbb{R}^{m \times m}$ , we let  $\tau(M)$  denote its associated *transfer function*, i.e.  $\tau(M)$  is the reduced rational function

$$C(sI - A)^{-1}B + D$$

Note that

$$\begin{aligned} sI - M &= \begin{bmatrix} sI - A & -B \\ -C & sI - D \end{bmatrix} \\ &= \begin{bmatrix} sI - A & 0 \\ -C & I \end{bmatrix} \begin{bmatrix} I & -(sI - A)^{-1}B \\ 0 & sI - \tau(M) \end{bmatrix} \end{aligned}$$

Then it follows easily that

$$cp(M) = ip(M) \det(s - \tau(M)) \quad (16)$$

with  $ip(M) = cp(A)$  that we call the *internal polynomial* of  $M$ . Moreover, for any square matrix  $M \in \mathbb{R}^{(p+1)m \times (p+1)m}$

and any permutation matrix  $T$  which leaves unchanged the last block row of  $TM$ , one has

$$\tau(TMT^\top) = \tau(M) \quad ip(TMT^\top) = ip(M) \quad (17)$$

This results from the fact that the Markov parameter sequences of  $\tau(TMT^\top)$  and  $\tau(M)$  are the same.

The proof of Lemma 2 will depend on the use of the following lemma:

**Lemma 3:** Let  $\mathbb{K}$  and  $\mathbb{L}$  be spanning subgraphs of a  $n$  node graph  $\mathbb{G}$ . Suppose that node  $n$  is a non-isolated node in both  $\mathbb{K}$  and  $\mathbb{L}$ . All non-isolated nodes of  $\mathbb{K}$  except  $n$  are isolated nodes of  $\mathbb{L}$ . Then for all  $K \in \mathcal{G}(\mathbb{K})$  and  $L \in \mathcal{G}(\mathbb{L})$

$$\tau(KL) = \tau(K)\tau(L) \quad (18)$$

$$ip(KL) = \frac{1}{(s-1)^{(n-1)m}} ip(K)ip(L) \quad (19)$$

**Proof of Lemma 3:** Let  $q+1$  denote the number of non-isolated nodes of  $\mathbb{K}$ . Suppose first that  $q+1 = n$ . Since all non-isolated nodes of  $\mathbb{K}$  are isolated nodes of  $\mathbb{L}$ , this means that  $\mathbb{L}$  must have at least  $n-1$  isolated nodes. But this can be true only if all  $n$  nodes of  $\mathbb{L}$  are isolated. Therefore the only matrix in  $\mathcal{G}(\mathbb{L})$  is the  $nm \times nm$  identity. Thus (18) is true. Moreover  $ip(I_{nm \times nm}) = (s-1)^{(n-1)m}$ , so (19) must be true.

Now suppose that  $q+1 < n$ . Let  $\{i_1, i_2, \dots, i_n\}$  be any permutation of the set  $\{1, 2, \dots, n\}$  for which  $\{i_1, i_2, \dots, i_q, n\}$  is the set of labels of the non-isolated nodes of  $\mathbb{K}$  and write  $T$  for that permutation matrix which represents the permutation  $\{1, 2, \dots, n-1, n\} \rightarrow \{i_1, i_2, \dots, i_{n-1}, n\}$ . By assumption, the nodes with labels in  $\{i_1, i_2, \dots, i_q\}$  must all be isolated in  $\mathbb{L}$ . Then for  $j \in \{i_1, i_2, \dots, i_q\}$  the  $j$ th block row and the  $j$ th block column of  $L$  must have  $I_m$  in the  $j$ th block entry and zero blocks elsewhere. This means that the matrix  $TLL^\top$  must be of the following form

$$TLL^\top = \begin{bmatrix} I_{qm \times qm} & 0 & 0 \\ 0 & A & B \\ 0 & C & D \end{bmatrix}, \quad (20)$$

where  $D \in \mathbb{R}^m$ . Note that  $\{i_{q+1}, \dots, i_{n-1}\}$  are all isolated nodes of  $\mathbb{K}$ . Then for  $j \in \{i_{q+1}, \dots, i_{n-1}\}$ , one has the  $j$ th block row and the  $j$ th block column of  $K$  must have  $I_m$  in its  $j$ th block entry and zeros elsewhere. This means that the matrix  $TKT^\top$  must be of the form

$$TKT^\top = \begin{bmatrix} \bar{A} & 0 & \bar{B} \\ 0 & I_{(n-q-1)m \times (n-q-1)m} & 0 \\ \bar{C} & 0 & \bar{D} \end{bmatrix}, \quad (21)$$

where  $\bar{D} \in \mathbb{R}^m$ . Note that the definition of  $T$  implies that it leaves unchanged the last block rows of  $TK$  and  $TL$ , implying that (17) holds. In view of this, all that is necessary to complete the proof is to show that

$$\tau(TKLT^\top) = \tau(TKT^\top)\tau(TLL^\top) \quad (22)$$

$$ip(TKLT^\top) = \frac{1}{(s-1)^{(n-1)m}} ip(TKT^\top)ip(TLL^\top) \quad (23)$$

Here we exploit the fact that  $TKLT^\top = TKT^\top LT^\top$ . Note that

$$TKLT^\top = \begin{bmatrix} \bar{A} & \bar{B}C & \bar{B}D \\ 0 & A & B \\ \bar{C} & \bar{D}C & \bar{D}D \end{bmatrix} \quad (24)$$

because of (20) and (21). Clearly

$$ip(TKLT^\top) = cp(\bar{A})cp(A) \quad (25)$$

But  $ip(TLT^\top) = (s-1)^{qm}cp(A)$  because of (20) and  $ip(TKT^\top) = (s-1)^{(n-q-1)m}cp(\bar{A})$  because of (21). From these two relations and (25) it follows that (23) is true.

From (24), one has

$$\tau(TKLT^\top) = [\bar{C} \quad \bar{D}C] M^{-1} \begin{bmatrix} \bar{B}D \\ B \end{bmatrix} + \bar{D}D \quad (26)$$

where

$$M = \begin{bmatrix} sI - \bar{A} & -\bar{B}C \\ 0 & sI - A \end{bmatrix}$$

Note that

$$M^{-1} = \begin{bmatrix} (sI - \bar{A})^{-1} & (sI - \bar{A})^{-1}\bar{B}C(sI - A)^{-1} \\ 0 & (sI - A)^{-1} \end{bmatrix}$$

It follows that

$$\begin{aligned} \tau(TKLT^\top) &= \bar{C}(sI - \bar{A})^{-1}\bar{B}D \\ &+ [\bar{C}(sI - \bar{A})^{-1}\bar{B}][C(sI - A)^{-1}B] \\ &+ \bar{D}C(sI - A)^{-1}B + \bar{D}D \end{aligned}$$

which can be rewritten as

$$(\bar{C}(sI - \bar{A})^{-1}\bar{B} + \bar{D})(C(sI - A)^{-1}B + D)$$

Thus

$$\tau(TKLT^\top) = \tau(TKT^\top)\tau(TLT^\top)$$

Thus (22) is true. We complete the proof. ■

**Proof of Lemma 2:** Note that any two spanning subgraphs in  $\{\mathbb{G}_1, \mathbb{G}_2, \dots, \mathbb{G}_r\}$  only have  $n$  as the non-isolated node in common. Then for any  $i \in \{1, 2, \dots, r-1\}$ ,  $n$  is the only common non-isolated node of the union graph  $\bigcup_{q=1}^i \mathbb{G}_{\pi(q)}$  and the graph  $\mathbb{G}_{\pi(i+1)}$ . Moreover, all non-isolated nodes of the union graph  $\bigcup_{q=1}^i \mathbb{G}_{\pi(q)}$  are isolated nodes in  $\mathbb{G}_{\pi(i+1)}$ . It follows from Lemma 3 that for any  $i \in \{1, 2, \dots, r-1\}$

$$\begin{aligned} \tau(G_{\pi(i+1)} \cdots G_{\pi(2)} G_{\pi(1)}) &= \tau(G_{\pi(i+1)})\tau(G_{\pi(i)} \cdots G_{\pi(1)}) \\ ip(G_{\pi(i+1)} \cdots G_{\pi(2)} G_{\pi(1)}) &= \frac{1}{(s-1)^{(n-1)m}} ip(G_{\pi(i+1)}) \\ &\quad \cdot ip(G_{\pi(i)} \cdots G_{\pi(1)}) \end{aligned}$$

These equations clearly imply that

$$\begin{aligned} \tau(G_{\pi(r)} G_{\pi(r-1)} \cdots G_{\pi(1)}) &= \tau(G_{\pi(r)})\tau(G_{\pi(r-1)}) \cdots \tau(G_{\pi(1)}) \\ ip(G_{\pi(r)} G_{\pi(r-1)} \cdots G_{\pi(1)}) &= \frac{1}{(s-1)^{(r-1)(n-1)m}} ip(G_{\pi(r)}) \\ &\quad \cdot ip(G_{\pi(r-1)}) \cdots ip(G_{\pi(1)}) \end{aligned}$$

Together with (16), these equations imply (1) and prove Lemma 2. ■

## REFERENCES

- [1] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.
- [2] B. D. O. Anderson, C. Yu, B. Fidan, and J. M. Hendrickx. Rigid graph control architectures for autonomous formations. *IEEE Control Systems*, 28(6):48–63, 2008.
- [3] Z. Sun, S. Mou, B. D. O. Anderson, and M. Cao. Exponential stability for formation control systems with generalized controllers: A unified approach. *Systems & Control Letters*, 93:50–57, 2016.
- [4] M. Cao, A. S. Morse, and B. D. O. Anderson. Reaching a consensus in a dynamically changing environment: a graphical approach. *SIAM Journal on Control and Optimization*, 47(2):575–600, 2008.
- [5] X. Chen, M. A. Belabbas, and T. Basar. Controllability of formations over directed time-varying graphs. *IEEE Transactions on Control of Network Systems*, pages 4362–4367, 2015.
- [6] K. B. Mahieddine and P. M. Dew. A periodic symmetrically-initiated load balancing algorithm for distributed systems. *ACM SIGOPS Operating Systems Review*, 28(1):66–79, 1994.
- [7] R. Labahn, S. T. Hedetniemi, and R. Laskar. Periodic gossiping on trees. *Discrete Applied Mathematics*, 53(1):235–245, 1994.
- [8] A. Nedic and A. Ozdaglar. Distributed sub-gradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61, 2009.
- [9] J. C. Duchi, A. Agarwal, and M. J. Wainwright. Dual averaging for distributed optimization: Convergence analysis and network scaling. *IEEE Transactions on Automatic Control*, 57(3):592–606, 2012.
- [10] D. Jakovetic, J. M. F. Moura, and J. Xavier. Fast distributed gradient methods. *IEEE Transactions on Automatic Control*, 59(5):1131–1146, 2014.
- [11] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transaction on Information Theory*, pages 2508–2530, 2006.
- [12] S. Mou, C. Yu, B. D. O. Anderson, and A. S. Morse. Deterministic gossiping with a periodic protocol. *Proc. of the 49th Conference on Decision and Control*, pages 5787–5791, 2010.
- [13] J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, and C. Yu. Deterministic gossiping. *Proceedings of the IEEE*, 99(9):1505–1524, 2011.
- [14] A. L. Liestman and D. Richards. Network communication in edge-colored graphs: Gossiping. *IEEE Transactions on Parallel and Distributed Systems*, 4(4):438–445, 1993.
- [15] B. D. O. Anderson, C. Yu, and A. S. Morse. Convergence of periodic gossiping algorithms. *Perspectives in Mathematical System Theory, Control, and Signal Processing. Springer Lecture Notes in Control and Information Sciences*, pages 127–138, 2010.
- [16] F. He, S. Mou, J. Liu, and A. S. Morse. Convergence rate on periodic gossiping. *Information Sciences*, 364:111–125, 2016.
- [17] C. Yu, B. D. O. Anderson, S. Mou, J. Liu, and F. He. Distributed averaging using a periodic gossiping scheme. *IEEE Transactions on Automatic Control*, 2017. Accepted.
- [18] S. Mou, J. Liu, and A. S. Morse. A distributed algorithm for solving a linear algebraic equation. *IEEE Transactions on Automatic Control*, 60(11):2863–2878, 2015.
- [19] S. Mou, Z. Lin, L. Wang, D. Fullmer, and A. S. Morse. A distributed algorithm for efficiently solving linear equations and its applications (special issue jcw). *Systems & Control Letters*, 91:21–27, 2016.
- [20] B. D. O. Anderson, S. Mou, U. R. Helmke, and A. S. Morse. Decentralized gradient algorithm for solution of a linear equation. *Numerical Algebra, Control and Optimization*, 6(3):319–328, 2016.
- [21] G. Shi and B.D. O.Anderson. Distributed network flows solving linear algebraic equations. *Proceedings of American Control Conference*, pages 2864–2869, 2016.
- [22] X. Wang, S. Mou, and D. Sun. Improvement of a distributed algorithm for solving linear equations. *IEEE Transactions on Industrial Electronics*, 64(4):3113–3117, 2017.
- [23] F. Deutsch and H. Hundal. The rate of convergence for the method of alternating projections.II. *Journal of Mathematical Analysis and Applications*, 205:381–405, 1997.