Abstract—The distributed averaging problem is a consensus problem whose objective is to devise a protocol which will enable all the members of a group of autonomous agents to compute the average of the initial values of their individual consensus variables in a distributed manner. Periodic gossiping is a deterministic method for solving the distributed averaging problem by stipulating that each pair of agents which are allowed to gossip, do so repeatedly in accordance with a pre-specified periodic schedule. Agent pairs which are allowed to gossip correspond to edges on a given connected, undirected graph. In general, the rate at which the agents’ consensus variables converge to the desired average value depends on the order in which the gossips occur over a period. The main contributions of this paper are first to characterize classes of periodic gossip sequences which have the same convergence rate and second to prove that if the graph of allowable gossips is a tree with each edge restricted to gossiping once per period, the convergence rate is quite surprisingly, fixed and invariant over all possible periodic gossip sequences the graph allows. To arrive at these results, use is made of a new and unusual graph theoretic concept, namely the transfer function of a node of an undirected graph. Among all the trees with the same number of edges, optimal tree structures which yield the fastest convergence rate can then be sought.

Index Terms—Gossiping, Multi-Agent Systems, Consensus, Edge-Coloring

I. INTRODUCTION

Gossiping is a particular way of achieving average consensus, i.e. deriving from a collection of networked agents’ initial values of a certain variable what the average value is. With a conventional gossiping algorithm (as opposed to one allowing multi-gossiping, see below), at any instant of time commonly just one pair of agents interacts, i.e. exchanges and averages their values. For analysis purposes, typically an underlying graphical structure is assumed for a gossiping algorithm, and the graph must be connected. Vertices correspond to agents of course, and the edges of the graph join those vertex pairs corresponding to agent pairs which are permitted to gossip. (Variants involving a jointly connected property over a time interval are sometimes invoked, to deal with time-varying topologies.) One can conceive of synchronous or asynchronous selection of edges, and random or deterministic selection of edges. Random gossiping algorithms are studied in, e.g. [1]–[4]. There are nevertheless clear reasons which may lead one on occasion to prefer deterministic gossiping to random gossiping. Here are three:

1) With random gossiping, there is only a convergence rate on averaged quantities, including mean square error. There are in fact particular random sequences with arbitrarily slow convergence rate which occur with positive probability; for these sequences, the actual squared error is greater than the mean squared error. In contrast, when using deterministic gossiping sequences, a convergence rate may be explicitly and often easily computed (and of course applies to all sequences).

2) Deterministic gossiping offers the opportunity for further convergence rate improvement using multi-gossiping; in multi-gossiping, several pairs of agents, with no agent in more than one pair, can gossip at a given time instant. (There is a random but centralized and synchronous algorithm for multi-gossiping mentioned in [1], but it does not seem to have enjoyed much attention, perhaps due to the requirements of centralization and synchronization.)

3) Distributed optimization algorithms [5] often rely on deterministic distributed averaging algorithms, of which deterministic gossip algorithms are a special case.

It is also relevant to ask what are the distinguishing features of distributed gossiping which differentiate it from other types of distributed averaging algorithms.

On the one hand, especially if there is an underlying clock, there is the almost certain likelihood that convergence will be slower, although the possibility of multigossiping when the gossiping is synchronous can mitigate the convergence loss. As against the speed disadvantage, there can however be advantages, starting with the simplicity of the information that is exchanged, and the way it is processed. If one uses the methods for distributed averaging set out in e.g. [6], [7], linear iterations with doubly stochastic matrices are used, requiring each agent to know an upper bound on the number of neighbors of each of its neighbors. A second possibility is to use double linear iterations (push-sum), e.g. [8]–[12]. Knowledge concerning the number of neighbors of neighbors is no longer a requirement, but the algorithms are more complicated than linear iterations, and require more data to be communicated between agents.

Yet another and old variant on averaging achieves computation in finite time: values are flooded through the network in finite time and each node computes an average at the end of the flooding process. A moment’s reflection illustrates a key difference (but not the only one of course) between this type
of algorithm and the ones mentioned above, which exhibit exponential convergence; the latter algorithms are robust to noisy values, or to limited short-term loss of communication between pairs of agents, and this in part accounts for their popularity. The flooding algorithm in contrast has nowhere near the same level of robustness.

Periodic multi-gossiping for message passing is an old and well-understood idea in computer science, see e.g. [13], [14]: multi-gossiping (but not periodic multi-gossiping) for averaging was also mentioned in [1]. Deterministic periodic multi-gossiping is dealt with in more detail below.

A criticism of deterministic gossiping algorithms is that they have some overhead for implementation, in that before gossiping commences, a protocol must be supplied to all agents to schedule when they gossip. However, such protocols can be very simple, and once gossiping commences, the gossiping process can then run in a very straightforward way [15]–[18] – no centralized activity (leaving aside synchronisation) is required. It seems indeed hard to rank logically two algorithms where one is more complicated to initialise but easier to run than the other.

In this work, we largely focus on gossiping algorithms where there is a (deterministic) periodic protocol causing the gossip labeled by each edge to occur exactly once in the underlying period. We do however briefly also consider situations where some edges gossip more than once in a single period; the situation is roughly analogous to having unequal probabilities of gossiping of allowed pairs in a random gossiping framework. Unsurprisingly, examples show one can achieve a speed-up, but we lack insights at this point that suggest how to systematically exploit the phenomenon.

Assuming convergence to a consensus average does occur, it is reasonable to suppose that the convergence rate associated with deterministic periodic gossiping will be dependent on the order of the gossips within one period. It is easy to construct examples which verify this. Nevertheless it is not always the case. We establish here a result which we found very surprising, originally having observed it experimentally. Consider a graph in which no edge gossips more than once per period. Then any two successive gossips within the period can be interchanged without affecting the convergence rate if either the nodes on which the edges are incident are all distinct, or, when there is a common node for the two edges, that the edges are not within any cycle of the graph. As an immediate corollary, if the graph is a tree, the convergence rate is independent of the order of the gossip sequence. Building on the main result, we can also establish conditions applicable to the case where within one period, one or more edges can gossip more than once.

The discovery of the invariance of convergence rate with respect to the ordering of gossips over a period for tree graphs (given no repeats within the period) was originally announced in our paper [15], and later in [16]. Neither paper contained a detailed proof of the claim. A complete proof of the invariance property for tree graphs was derived in an unpublished technical report which was the basis of this paper; that report did not treat the case of general graphs addressed here. The work in this paper inspired the work in [18], which derived the result in an alternative way through a detailed calculation of specific entries of matrices arising in the gossip process, including products of matrices used to characterise single gossips. In contrast to this paper, none of the cited works treated graphs more general than tree graphs, and it is quite unclear how [18] could be extended beyond tree graphs.

In the course of establishing the results, we introduce a form of transfer function associated with a node and periodic gossiping sequence in a graph. This concept appears to be of interest in its own right, and is also used as a device for obtaining the characteristic polynomial of a gossiping matrix associated with a graph formed by joining two or more simpler graphs at a common node.

The next section is devoted to preliminaries and some definitions. Section III states the main result for general graphs, and in a corollary, the application of the main result to trees. Section IV provides a proof of the main result. Section V considers various examples of tree and star graphs, with illustrations of convergence rate. There is also examination of the question of what trees for a fixed number of nodes will give the fastest convergence rate for periodic gossiping with no repeats in a period. The section also makes some observations on the distinction between random and periodic gossiping. Section VII contains concluding remarks.

II. Definitions and Background

Let $\mathbb{F} = (\mathcal{V}, \mathcal{E})$ be an undirected simple graph with $n$ nodes in the node set $\mathcal{V}$ and $m$ edges in the edge set $\mathcal{E}$. The vertices correspond to agents which store a variable of interest, and the edges define the agent pairs between which a gossip potentially occurs. A pair of agents with labels $i$ and $j$ are said to gossip at time $t$ if both $x_i(t + 1)$ and $x_j(t + 1)$ are set equal to the average of $x_i(t)$ and $x_j(t)$. Thus $\mathbb{F}$ is the allowed gossip graph. Let $\mathbf{E} = e_1, e_2, \ldots, e_m$ with $m = |\mathcal{E}|$ be an ordering of the edges of $\mathbb{F}$, with each edge appearing just once. We call such a sequence non-redundant. Call the sequence $\mathbf{E}$ complete if $\mathbb{F}$ is connected, and call it minimally complete if it is complete and $\mathbb{F}$ is a tree.

Over one period of gossiping, the gossips defined by $\mathbf{E}$ occur in the defined order. Over an infinite interval, periodic repetition is assumed to occur, giving rise to an infinite periodic gossip sequence.

Any individual gossip whether or not in a periodic gossip sequence can be described by a specially structured doubly stochastic matrix (call it a single-gossip matrix). If edge $i$ defines a gossip, and is incident on nodes $j, l$ the associated single-gossip matrix $S_i$ has entries of $1/2$ in the $jj, jl, lj, ll$ entries, 1’s elsewhere on the main diagonal and zeros elsewhere. It can be easily checked that if two edges are not incident, the corresponding two single-gossip matrices commute. If $x(k)$ is the vector of stored values at the different nodes at time $k$ and edge $i$ then gossips, there holds $x(k + 1) = S_i x(k)$. If in all there are $m$ edges in a periodic gossiping sequence, the composition of the $m$ successive gossips corresponding to a product of such single-gossip matrices is again a doubly stochastic matrix, call it $T$. Where we wish to emphasise that $T$ is indeed a product, we shall term it a composite gossip
matrix. The term gossip matrix may denote a single-gossip or composite gossip matrix. Define $\rho$ to be the magnitude of the eigenvalue of second largest magnitude of the matrix $T$ (the largest eigenvalue is 1). If $\rho < 1$ (and conditions for this are reviewed below) then the convergence rate on a per period basis is $\rho$ (in the sense that $T^k$ approaches a limit as $k \to \infty$ as fast as $\rho^k$ tends to zero). Since there are $m$ gossip processes making up one period, then the convergence rate on a per gossip basis, which is the usual basis for describing convergence rates, is $\rho^{1/m}$.

Denote by $T_E$ the composite gossip matrix defined by the sequence $E$. Then

$$T_E = S_m S_m^{-1} \ldots S_1$$  \hspace{1cm} (1)

Call $T_E$ a (minimally) complete gossip matrix just when $E$ is (minimally) complete. If the first gossip within a single period occurs at time $k$ and $x(k)$ denotes the vector of gossip variables, then after one period, the new vector of gossip variables is $x(k+m) = T_E x(k)$. Now given a periodic gossip sequence $E$, the completeness property, equivalently connectedness of the associated graph $F$, is necessary and sufficient to ensure that $\rho(T_E) < 1$ and convergence of the gossip process occurs, see e.g. [19]. Indeed it is well known, see [19], that because $\rho(T_E) < 1$, then $T_E^k \to \frac{1}{\rho} I$ as $k \to \infty$. (Here, as usual, $I$ denotes a vector of all 1’s).

Below, we shall be working with composite gossip matrices formed from a subgraph of a given graph. On occasion, this subgraph may have no edges, though it will have (a nonzero number of) vertices. In this case, we simply take $T = I$.

An important tool in the following is provided by the transfer function of a gossip matrix, especially a complete gossip matrix, which is defined in conjunction with identifying a particular node of the graph; for convenience we take this to be the last numbered node, since node reordering will deal with the case where a different node is used. Suppose $T$ is an $n \times n$ gossip matrix, and written in partitioned form as

$$T = \begin{bmatrix} A & b \\ c & d \end{bmatrix}$$  \hspace{1cm} (2)

with $d$ a scalar. Then

$$w(z) = d + c(zI - A)^{-1}b$$  \hspace{1cm} (3)

is defined to be the transfer function associated with $T$. We will always assume that $w$ is given in reduced form, i.e. pole-zero cancellations are effected. To distinguish transfer functions associated with different gossip matrices, we shall usually write $w(T)$, suppressing the $z$ dependence. Note that if $T = I$, then $w(z) = 1$.

**Remark 1:** An interpretation of $w(z)$ can be easily given. Consider a signal flow graph [20] in which the nodes correspond to nodes 1 through $n-1$ of $F$, with two further nodes, an input node, label it $n_i$ and an output node, labelled $n_o$, that can together be thought of as corresponding to node $n$ of $F$. Let $a_{ij}, b_i, c_i$ denote respectively the $ij$ entry of $A$, the $i$-th entry of the column vector $b$ and the $i$-th entry of the row vector $c$. In the signal flow graph, there is a directed edge from node $j$ to node $i$ which is labelled with the gain $a_{ij} z_j^{-1}$ for $1 \leq i, j < n$, a directed edge from node $n_i$ to node $i$ with the gain $b_i z_i^{-1}$, a directed edge from node $i$ to node $n_o$ with gain $c_i$ for $1 \leq i < n$, and a directed edge from node $n_i$ to node $n_o$ with gain $d$. The transfer function $w(z)$ is the gain from $n_i$ to $n_o$. This is rather like a loop gain, obtained by constructing a signal flow graph corresponding to the gossip algorithm, and but with an opening of the loop at node $n$.

Denote by $ip(T)$ (shorthand for the ‘internal polynomial of’ $T$) the characteristic polynomial $|zI - A|$. It is straightforward to see that

$$|zI - T| = ip(T)(z - w(T))$$  \hspace{1cm} (4)

Our key interest in this paper is in considering variations of the gossip ordering within a period. To this end, let $\pi : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\}$ denote a permutation map. Let $E_{\pi} = e_{\pi(1)}, e_{\pi(2)}, \ldots, e_{\pi(m)}$ denote the edge sequence after acting on $E$ with the permutation $\pi$. Then

$$T_{E_{\pi}} = S_{\pi(m)} S_{\pi(m-1)} \ldots S_{\pi(1)}$$  \hspace{1cm} (5)

**III. MAIN RESULTS**

The first main result we will prove in the next section is as follows.

**Theorem 1:** Let $F = (V, E)$ be an arbitrary simple, connected, undirected graph with $n$ nodes and $m$ edges. Let $E$ be an ordering of the edges in $E$, defining a non-redundant complete gossip sequence of length $m$. Consider the group $\Pi(E)$ (under composition) of permutations $\pi : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\}$ generated by permutations of the following types:

1) $\pi$ is the identity permutation
2) $\pi$ preserves the cyclic order of the edges
3) $\pi$ is a transposition permutation which for some $i < m$ interchanges $i$ and $i+1$ provided that in $F$, edges $e_i, e_{i+1}$ are not incident on the same node
4) $\pi$ is a transposition permutation which for some $i < m$ interchanges $i$ and $i+1$ provided that neither $e_i$ nor $e_{i+1}$ is contained in any cycle of $F$.

Let $T_E$ and $T_{E_{\pi}}$ be as defined in (1) and (5). Then for any $\pi$ in $\Pi(E)$, these two matrices have the same eigenvalues.

That the eigenvalues of $T_E$ are left invariant by the first three transformations is straightforward to demonstrate, as we now show. The first claim is obvious. For the second claim, recall that it is well known that for two square matrices $S_1, S_2$ of the same size, the eigenvalues of $S_2 S_1$ are identical with those of $S_1 S_2$. From this it follows at once that the eigenvalues of

![Fig. 1.](image-url)
which occurring in rows and columns corresponding to the nodes on the identity matrix in precisely four entries, which are of value 1/2, for the third claim, observe that \( S_i, S_{i+1} \) each differ from the identity matrix in precisely four entries, which are of value 1/2, occurring in rows and columns corresponding to the nodes on which \( e_i, e_{i+1} \) are incident. Because all four nodes are distinct, \( S_i \) and \( S_{i+1} \) commute. Thus this permutation not only leaves the eigenvalues invariant, but also the product matrix itself.

It is the eigenvalue invariance property under the fourth permutation type which is nontrivial to verify. Verification is provided in Section IV. Note that in the light of the third condition, we only need to consider the situation where \( e_i, e_{i+1} \) are incident on a common node. Figure 1 illustrates (a) a case in which the transposition of \( e_1 \) and \( e_2 \) belongs to the fourth permutation type; and (b) a case in which the transposition of \( e_1 \) and \( e_2 \) belongs to neither the third nor fourth permutation types.

A. The result for trees

We now indicate the most important consequence of the above theorem, achieved by restricting the underlying graph to be a tree.

**Theorem 2:** Adopt the same hypothesis as for Theorem 1 and suppose additionally that \( E \) is a minimally complete sequence, or equivalently that \( F \) is a tree. Then for any permutation \( \pi : \{1, 2, \ldots, m\} \rightarrow \{1, 2, \ldots, m\} \) the complete gossip matrices \( T_E, T_{E_\pi} \) have the same eigenvalues.

The proof is immediate: it is well known that every permutation of a finite set is expressible as a composition of transpositions of adjacent elements, see e.g. [21]; since \( F \) is a tree, no edge is in a cycle, and so under the fourth category, all transpositions of adjacent elements, and therefore all permutations, are allowed.

B. Speed-up of convergence by multi-gossiping

Lastly in this section, we remark on a consequence of Theorem 2 which allows speed up for gossiping over trees. When two successive gossiping steps (involving two edges) do not involve any agent in common, the two edges are said to be non-adjacent, and then the gossips can be performed simultaneously [1].

To state this more formally in relation to the graph \( F = (V, E) \), let us define a *multi-gossiping sequence* as a cyclic sequence \( \{E_1, E_2, \ldots, E_i, E_1, \ldots\} \) of subsets \( E_i \) of individual gossiping edges, such that (a) \( E_i \cap E_j = \emptyset \) (\( i \neq j \)); (b) \( \bigcup E_i = E \); (c) no two edges in the same \( E_i \) have a common node. Then at any clock pulse, a set of edges \( E_i \) become active and gossiping is performed over these edges simultaneously.

An optimal multi-gossiping schedule for a graph may be characterized as one where there is the least number of simultaneous gossips required in one period; such a schedule can be identified by solving an edge-coloring problem on \( F = (V, E) \). In graph theory [22], edge coloring is an assignment of different colors to the edges of a graph \( F \), such that no two edges having the same color are incident on the same node. The minimum required number of colors for a graph is called the chromatic index, denoted by \( \chi'(F) \). Various results exist involving the chromatic index, but the most relevant one is Vizing’s Theorem: the chromatic index of any graph equals either the maximum node degree, \( \delta(F) \), or \( \delta(F) + 1 \). There are also algorithms for solving the edge-coloring problem in a distributed way, see e.g. [23], [24].

While determining the exact chromatic index of a general graph is NP-complete, it is easy to determine two possible values which differ only by one, so the acceleration of the convergence rate for multi-gossiping over single gossiping is almost guaranteed by an order of \( \delta(F) \). For some common families of graphs, the chromatic index is easily obtainable, see e.g. [25]. In particular, for a path graph, the chromatic index is 2 (which is \( \delta \)), for a ring graph it is 2 or 3 according as the number of edges is even or odd, and for a tree graph, the chromatic index is \( \delta \).

C. Allowing redundant gossip sequences

To this point, we have postulated that in one period, the gossiping sequence is defined by an ordering \( E \) of the edges in \( E \) which is non-redundant, i.e. no repetitions occur. We now extend Theorem 1 to relax this constraint.

**Theorem 3:** Let \( F = (V, E) \) be an arbitrary simple, connected, undirected graph with \( n \) nodes and \( m \) edges. Let \( E \) be an ordering of the edges in \( E \), such that each edge occurs at least once, and suppose there are \( m' > m \) elements in \( E \). Consider the group \( \Pi(E) \) (under composition) of permutations \( \pi : \{1, 2, \ldots, m'\} \rightarrow \{1, 2, \ldots, m'\} \) acting on ordered sets of \( m' \) elements and generated by the following permutations types:

1) \( \pi \) is the identity permutation
2) \( \pi \) preserves cyclic ordering
3) \( \pi \) is a transposition permutation which interchanges two adjacent elements of the ordering provided that in \( F \), the corresponding edges are not incident on the same node
4) \( \pi \) is a transposition permutation which interchanges two adjacent elements of the ordering provided that in \( F \), each corresponding edge is not contained in any cycle of \( F \), nor is it a repeated edge, i.e. it does not correspond to more than one element of \( E \).

Let \( T_E \) and \( T_{E*} \) be as defined in (1) and (5), save that \( m \) is replaced by \( m' \). Then for any \( \pi \) in \( \Pi(E) \), these two matrices have the same eigenvalues.

**Proof:** Define a new graph \( F^* = (V, E^*) \), where \( E^* \) is obtained from \( E \) by adjoining to it extra edges between any node pair where the corresponding edge appears more than once in the ordering defined by \( E \). If an edge, \( e_j \), say, gives rise to \( q \) entries in \( E \), then \( q - 1 \) extra edges are inserted between the nodes on which it is incident. Call these additional edges \( e_{j2}, e_{j3}, \ldots, e_{jq} \). An ordering \( E^* \) is defined for \( F^* \) by identifying the \( q \) entries of \( E \) corresponding to \( e_j \) and assigning \( q - 1 \) of them arbitrarily to correspond to \( e_{j2}, e_{j3}, \ldots, e_{jq} \).

It is evident that \( E^* \) is a non-redundant and complete gossip sequence of length \( m' \) for \( F^* \). As such, Theorem 1 applies. It is then straightforward to conclude that condition 4 of this theorem applied to \( F \) corresponds to condition 4 of Theorem 1 applying to \( F^* \).
IV. Establishing the Main Result

Our aim is now to establish Theorem 1, and in particular the eigenvalue invariance property under the fourth condition. The general approach is inductive. An overview of the approach is as follows. Consider a graph in which there is a node of degree at least 2 with at least two incident edges which are not in any cycle. Decompose, in a manner to be described and guided by the existence of this node, the overall graph into a union of spanning subgraphs (i.e., subgraphs with the same node set as the original graph) with disjoint edge sets. Relate the complete gossip matrix for the original graph to gossip matrices for the spanning subgraphs, and likewise relate the transfer functions and the characteristic polynomials. Then prove the desired result that the permutations of the fourth type listed in the theorem statement do not change the eigenvalues of the composite gossip matrix over one period for the original graph, assuming by induction that it is valid for the spanning subgraphs.

We first recall some further terminology which is useful for this section. Let \( G \) be a simple, not necessarily connected graph with node set \( \{1, 2, \ldots, n\} \). By an isolated node of \( G \) is meant any node with degree 0; all other nodes are non-isolated. Obviously, if \( i \) is an isolated node of \( G \), then in any gossip matrix, \( G \) say, formed from \( G \), whether it is single or composite the \( i \)th row of \( G \) has a one in its \( i \)th column and zeros elsewhere.

A. Background Lemmas

We first state two straightforward lemmas to be used in the sequel; their simple proofs are omitted.

**Lemma 1:** Let \( T \) be an \( n \times n \) matrix and let \( \pi : \{1, 2, \ldots, n\} \rightarrow \{1, 2, \ldots, n\} \) define a permutation which leaves \( n \) unchanged, with \( Q \) the associated permutation matrix. Then the last row of \( QT \) is the same as that of \( T \) and there holds
\[
    w(QTQ^\top) = w(T) \quad \text{and} \quad ip(QTQ^\top) = ip(T) \quad (6)
\]

**Lemma 2:** Let \( G = (V, E) \) be a simple graph with \( n \) nodes, and suppose that for some integer \( r \geq 1 \), there exist spanning subgraphs \( G_1 = (V, E_1), G_2 = (V, E_2), \ldots, G_r = (V, E_r) \) of \( G \) such that
(a) \( \cup G_i = G \)
(b) For all \( 1 \leq i < j \leq r \), subgraphs \( G_i \) and \( G_j \) have no non-isolated node in common.

Let \( G \) be a composite gossip matrix for \( G \) associated with a non-redundant sequence \( E \) and suppose it can be expressed in terms of composite gossip matrices \( G_i \) associated with non-redundant subgraphs of each \( G_i \) as follows:
\[
    G = G_{\pi(1)}G_{\pi(2)} \cdots G_{\pi(r)} \quad (8)
\]

for some particular permutation \( \pi : \{1, 2, \ldots, r\} \rightarrow \{1, 2, \ldots, r\} \). Suppose that the nodes are ordered, and node \( v \) is the last node in the sequence. Then there holds:
\[
    w(G) = \prod_{i=1}^{r} w(G_i)
\]

and
\[
    ip(G) = \frac{1}{(z - 1)(r - 1)(n - 1)} ip(G_1)ip(G_2) \cdots ip(G_r) \quad (9)
\]

and
\[
    |zI - G| = |zI - G_{\pi(1)}G_{\pi(2)} \cdots G_{\pi(r)}| \quad (10)
\]

for an arbitrary permutation \( \sigma : \{1, 2, \ldots, r\} \rightarrow \{1, 2, \ldots, r\} \).

**Proof:** Note an immediate implication of the hypothesis (b’), that no edge can be in two different edge sets \( E_i \). Hence the \( E_i \) form a partition of \( E \), so that in particular, \( E_i \cap E_j = \emptyset, 1 \leq i < j \leq r \), a fact we will use below.

We will establish the formulae for \( w(G) \) and \( ip(G) \) by induction on \( r \). For the basis step, suppose that \( r = 2 \).

Let \( G_{\pi(1)} \) have \( k + 1 \) non-isolated nodes. We will consider two cases. Suppose first that \( k + 1 = n \), implying all nodes imply that the edge sets \( E_i \) of the \( G_i \) form a partition of the edge set \( E \) of \( G \) and if for any node \( u \) of \( G \), there is an incident edge, call it \( e \), then \( e \) and all other edges incident at \( u \) must lie in one and the same subgraph, \( G_i \) say (depending on \( u \)), with \( u \) isolated in \( G_j \) for all \( j \neq i \). Hence any connected component of \( G \) in fact must lie entirely in one \( G_i \). Lemma 2 is then an immediate consequence of Claim 3 of Theorem 1: if any two successive single-gossip matrices in a non-redundant gossip sequence for \( G \) correspond to edges from different \( G_i \), then they may be interchanged, since they cannot share a common node. In this way, all the individual gossips corresponding to edges in the one \( G_i \) can be sequentially performed, and the sequence of gossips associated with \( G_i \) can be done before or after the sequence associated with \( G_j \) for any \( i \neq j \).

The next lemma is of independent interest. When a graph \( G \) has the structure defined in the hypothesis of the lemma and when the gossip sequence order is, in a sense made clear in the lemma statement, consistent with the graph structure, the transfer function and internal polynomial can be related to those of the spanning subgraphs. Following the lemma, we shall specialize the structure to which it is applicable.

**Lemma 3:** Let \( G = (V, E) \) be a simple not necessarily connected graph with \( n \) nodes, and suppose that for some \( r \), there exist spanning subgraphs \( G_1 = (V, E_1), G_2 = (V, E_2), \ldots, G_r = (V, E_r) \) of \( G \) such that
(a) \( \cup G_i = G \)
(b) For all \( 1 \leq i < j \leq r \), the spanning subgraphs \( G_i \) and \( G_j \) have no non-isolated node in common with the exception of a single non-isolated node \( v \), which is non-isolated in at least two of the \( G_i \).

Let \( G \) be a composite gossip matrix for \( G \) associated with a non-redundant sequence \( E \) and suppose it can be expressed in terms of composite gossip matrices \( G_i \) associated with non-redundant sequences in each \( G_i \) as follows:
\[
    G = G_{\pi(1)}G_{\pi(2)} \cdots G_{\pi(r)} \quad (8)
\]

for some particular permutation \( \pi : \{1, 2, \ldots, r\} \rightarrow \{1, 2, \ldots, r\} \). Suppose that the nodes are ordered, and node \( v \) is the last node in the sequence. Then there holds:
\[
    w(G) = \prod_{i=1}^{r} w(G_i)
\]

and
\[
    ip(G) = \frac{1}{(z - 1)(r - 1)(n - 1)} ip(G_1)ip(G_2) \cdots ip(G_r) \quad (9)
\]

and
\[
    |zI - G| = |zI - G_{\pi(1)}G_{\pi(2)} \cdots G_{\pi(r)}| \quad (10)
\]

for an arbitrary permutation \( \sigma : \{1, 2, \ldots, r\} \rightarrow \{1, 2, \ldots, r\} \).
of \( \mathcal{G}_{\pi(1)} \) are non-isolated. Condition (b') implies that with \( r = 2 \), both \( \mathcal{G}_1 \) and \( \mathcal{G}_2 \) must have at least one edge. Hence \( \mathcal{G}_{\pi(2)} \) has at least two nonisolated nodes, which are evidently nonisolated for \( \mathcal{G}_{\pi(1)} \) as well. This is a contradiction and so this case cannot arise.

For the second case, still with \( r = 2 \), suppose \( k+1 < n \), and denote the set of labels of the non-isolated and isolated nodes of \( \mathcal{G}_{\pi(1)} \) by \( \{i_1, i_2, \ldots, i_k, n\} \) and \( \{i_{k+1}, \ldots, i_{n-1}\} \) respectively. Recall \( v \) is chosen to be the node corresponding to index \( n \). Since \( r = 2 \), \( n \) indexes the common non-isolated node to the two subgraphs. Let \( Q \) denote the permutation matrix representing the permutation \( \{1, 2, \ldots, n\} \rightarrow \{i_1, i_2, \ldots, i_{n-1}, n\} \); note for use below that \( QT \) has the same last row as \( T \) since the permutation leaves \( n \) invariant and note that \( Q \) is not a representation of \( \pi \). It follows that \( QG_{\pi(2)}Q^T \) takes the following form for some \( A_2, b_2, c_2, d_2 \),

\[
QG_{\pi(2)}Q^T = \begin{bmatrix} I_{k \times k} & 0 & 0 \\ 0 & A_2 & b_2 \\ 0 & c_2 & d_2 \end{bmatrix}
\]

with \( d_2 \) scalar. Similarly, we can argue that for some \( A_1, b_1, c_1, d_1 \),

\[
QG_{\pi(1)}Q^T = \begin{bmatrix} A_1 & 0 & b_1 \\ 0 & I_{(n-k-1) \times (n-k-1)} & 0 \\ c_1 & 0 & d_1 \end{bmatrix}
\]

Evidently,

\[
QG_{\pi(1)}G_{\pi(2)}Q^T = QG_{\pi(1)}Q^T QG_{\pi(2)}Q^T = \begin{bmatrix} A_1 & b_1c_2 & b_1d_2 \\ 0 & A_2 & b_2 \\ c_1 & d_1c_2 & d_1d_2 \end{bmatrix}
\]

It is then straightforward algebra to verify that

\[
w(QG_{\pi(1)}G_{\pi(2)}Q^T) = w(QG_{\pi(1)}Q^T)w(QG_{\pi(2)}Q^T)
\]

and it is immediate that

\[
wp(QG_{\pi(1)}G_{\pi(2)}Q^T) = |zI - A_1||zI - A_2|
\]

\[
wp(QG_{\pi(1)}Q^T) = (z - 1)^{(n-k-1)}|zI - A_1|
\]

\[
wp(QG_{\pi(2)}Q^T) = (z - 1)^k|zI - A_2|
\]

Now recall that \( QT \) for any \( T \) has the same last row as \( T \). By Lemma 1, it follows from these equations that

\[
w(G_{\pi(1)}G_{\pi(2)}) = w(G_{\pi(1)})w(G_{\pi(2)})
\]

\[
wp(G_{\pi(1)}G_{\pi(2)}) = \frac{1}{(z - 1)^{(n-1)}}wp(G_{\pi(1)})wp(G_{\pi(2)})
\]

This establishes the basis step of the induction. We observe the fact, used below, that had \( v \) been an isolated node of \( \mathcal{G}_{\pi(2)} \), the same equations (16) would result; in the derivation, one would simply have \( b_2 = 0, c_2 = 0, d_2 = 1 \).

The inductive step is more straightforward: assume the result holds for \( r = 1, 2, \ldots, i \), and suppose now \( r = i + 1 \). Then by hypothesis \( G \) can be written as a product \( G_{\pi(1)}G_{\pi(2)} \ldots G_{\pi(i+1)} = [G_{\pi(1)}G_{\pi(2)} \ldots G_{\pi(i)}]G_{\pi(i+1)} \) of composite gossip matrices, for some permutation \( \pi \). Define the union graph \( \mathcal{G}_{\pi(i+1)} := \bigcup_{k=1}^{i+1} \mathcal{G}_{\pi(k)} \). By hypothesis, any nonisolated node of the union graph other than the node labelled \( m \) is an isolated node of \( \mathcal{G}_{\pi(i+1)} \). Also, \( \mathcal{G}_{\Pi(i)} \cup \mathcal{G}_{\pi(i+1)} = \mathcal{G} \) and \( \mathcal{E}_{i+1} \) is disjoint from \( \bigcup_{k=1}^{i} \mathcal{E}_{\pi(k)} \). Accordingly, the basis step of the induction yields (whether or not \( m \) labels a nonisolated node of \( \mathcal{G}_{\pi(i+1)} \))

\[
w(G_{\Pi(i+1)}) = w(G_{\Pi(i)}G_{\pi(i+1)}) = w(G_{\Pi(i)})w(G_{\pi(i+1)})
\]

\[
wp(G_{\Pi(i)}G_{\pi(i+1)}) = \frac{1}{(z - 1)^{(n-1)}}wp(G_{\Pi(i)})wp(G_{\pi(i+1)})
\]

and then invoking the standard induction assumption,

\[
w(G_{\Pi(i)}) = \prod_{j=1}^{i} w(G_{\pi(j)})
\]

\[
wp(G_{\Pi(i)}) = \frac{1}{(z - 1)^{(n-1)}} \prod_{j=1}^{i} wp(G_{\pi(j)})
\]

it is immediate that

\[
w(G_{\Pi(i+1)}) = \prod_{j=1}^{i+1} w(G_{\pi(j)})
\]

\[
wp(G_{\Pi(i+1)}) = \frac{1}{(z - 1)^{(n-1)}} \prod_{j=1}^{i+1} wp(G_{\pi(j)})
\]

This completes the induction.

Finally, to establish the sequence independence property (10) for \( |zI - G| \), recall (8). By what has just been proved, for any permutation \( \sigma \), the gossip matrix \( \hat{G} = G_{\sigma(1)}G_{\sigma(2)} \ldots G_{\sigma(r)} \) has the same transfer function, viz. \( \prod_{i=1}^{r} w(G_{i}) \) and same internal polynomial, viz \( \prod_{i=1}^{r} wp(G_{i}) \). Hence \( \hat{G} \) and \( G \) have the same characteristic polynomial by (4). This proves the result. \( \blacksquare \)

**Remark 2:** There is a helpful way to think about the last lemma. Consider \( r \) graphs, \( \mathbb{H}_i, i = 1, 2, \ldots, r \) with disjoint node sets of size \( n_i \). Identify one node in each graph and connect the \( r \) graphs at this node, to form a new graph \( G \) with \( n = \sum_{i=1}^{r} n_i - (r - 1) \) nodes. [Many graphs can be grown in this fashion, and all trees can be grown by a sequence of such operations.] Then the transfer function and internal polynomial for \( G \) can be found in terms of those of the \( \mathbb{H}_i \). To see this, define first spanning subgraphs \( \mathcal{G}_i \) of \( G \) by adding to \( \mathbb{H}_i \) the set of nodes of \( \cup_{j \neq i} \mathbb{H}_j \), except for the node which becomes the common node at which all the \( \mathbb{H}_i \) are joined when forming \( G \). Clearly, gossip matrices for the \( \mathbb{H}_i \) induce gossip matrices for the \( \mathcal{G}_i \); up to node reordering, a gossip matrix for \( \mathcal{G}_i \) is the direct sum of an identity matrix \( I_{n_i - n} \) and a gossip matrix for \( \mathbb{H}_i \). In obvious notation, \( w(H_i) = w(G_i), \) \( wp(H_i) = (z - 1)^{n_i - n} wp(G_i) \). So the lemma explains how to obtain the transfer function and internal polynomial for a graph in terms of the corresponding quantities for subgraphs (not spanning subgraphs) obtained by separating the given graph at a node.

**B. Proof of Main Theorem**

We now have the machinery to complete the proof of the Theorem 1.

Permutations of the first three types listed in the theorem statement are assumed to have been verified as providing the eigenvalue result. Given that in particular the eigenvalue
property has been verified for the third type, it is enough
to consider the fourth type when \( e_i, e_{i+1} \) have a common
node. Without loss of generality, suppose that the nodes are so
ordered that this is the last node, with index \( n \). Recalling that
\( S_j \) is the single-gossip matrix corresponding to edge \( e_j \), let \( \pi \)
denote the permutation of the edges which simply interchanges
edges \( i \) and \( i + 1 \), so that, as defined in (1),
\[
T_E = S_m S_{m-1} \ldots S_{i+1} S_i \ldots S_1
\]
while also
\[
T_{E_n} = S_m S_{m-1} \ldots S_i S_{i+1} \ldots S_1 \tag{17}
\]

We are required to show that \( |zI - T_E| = |zI - T_{E_n}| \). As

illustrated in Figure 2, let \( e_i \) join nodes \( u, n \) and \( e_{i+1} \) join \( v, n \).
Let \( G \) denote the subgraph of \( F \) when \( e_i, e_{i+1} \) are
deleted from \( F \). Define nonempty subgraphs \( G_i, i = 1, 2, 3 \) of \( G \)
(which will then be subgraphs also of \( F \)) as follows: \( G_1 \) contains the
connected component of \( G \) containing node \( u \); \( G_2 \) contains the
connected component of \( G \) containing node \( v \); \( G_3 \) contains the
other edges of \( G \) (there will be none if \( e_i \) and \( e_{i+1} \) are the only
edges incident on \( n \)). Note that if \( e_i \) or \( e_{i+1} \) were in a cycle,
there could not be three such nonempty subgraphs. The definition
results in \( G, G_1, G_2 \) and \( G_3 \) fulfilling the conditions of Lemma 2
with \( r = 3 \). By cyclic permutation, we have
\[
|zI - T_E| = |zI - MS_{i+1} S_i| \tag{18}
\]
for some \( M \) which is is associated with a non-redundant edge
sequence for \( G \). By Lemma 2, there are corresponding gossip
matrices \( G_i \) associated with the \( G_i \) such that \( M = G_3 G_2 G_1 \),
and so
\[
|zI - T_E| = |zI - G_3 G_2 G_1 S_{i+1} S_i| \tag{19}
\]
Moreover, the edge set of \( G_1 \) is incident on no nodes in
common with the pair associated with the gossip \( S_{i+1} \), so
that \( G_1 \) and \( S_{i+1} \) commute. Hence
\[
|zI - T_E| = |zI - G_3 G_2 S_{i+1} G_1 S_i| \tag{20}
\]

Now define spanning graphs \( H_j, j = 1, 2 \) of \( F \) with \( H_j \)
having edge set as the union of the edge set of \( G_j \) and \( e_{i+j-1} \).
Notice that \( F = H_1 \cup H_2 \cup G_3 \), that \( H_1, H_2, G_3 \) have a single
non-isolated node in common, viz. \( n \), and that the edge sets
of the three graphs are disjoint. Accordingly, Lemma 3 is
applicable, with \( r = 3 \).

Now observe that \( G_1 S_j \) and \( G_2 S_{i+1} \) are composite gossip
matrices associated with non-redundant edge sequences of \( H_1 \)
and \( H_2 \). Hence by Lemma 3, equation (20) yields
\[
|zI - T_E| = |zI - G_3 G_1 S_j G_2 S_{i+1}|
\]

Because \( G_2 \) commutes with \( S_j \) and \( G_1 \), there then holds
\[
|zI - T_E| = |zI - G_3 G_1 S_j S_{i+1}|	ag{21}
\]
\[
= |zI - MS_{i+1} S_j| = |zI - T_{E_n}|
\]

with the last step following by cyclic permutation again. This
completes the proof of the main theorem.

\textbf{Remark 3:} In [26], it is pointed out that replacement of the
usual gossip equations for every edge \( j \in E \) by
\[
x_j(k + 1) = w x_j(k) + (1 - w) x_i(k)
\]
\[
x_i(k + 1) = (1 - w) x_j(k) + w x_i(k)
\]
with \( w \in (0, 1) \) still leads to convergence to an average, but
the convergence rate can be faster for some values of \( w \). The
main results of this paper and their proofs still hold with this
replacement, there being nothing special about \( w = \frac{1}{2} \).

\section{Examples}

In this section, we present some examples of trees, including
path and star graphs, and identify optimal tree topologies,
i.e. those offering fastest convergence. Some comparison with
random gossiping is also made.

\subsection{Path and star graphs}

Consider a path graph with \( m \) edges. A rather lengthy
induction argument will show that the transfer function ob-
tained when a leaf node is taken as the last vertex is
expressible in terms of Chebyshev polynomials, and the
eigenvalues of the characteristic polynomial of the composite
gossip matrix \( T_m \) over one period are \( 1, \cos^2 \frac{2\pi}{m+1}, j = 1, 2, \ldots, \lfloor m/2 \rfloor, 0, \ldots, 0 \).
Then \( \rho(T_m) \) is \( \cos^2 \frac{2\pi}{m+1} \). This
means that without multigossiping, the convergence rate is
\( \rho^\frac{1}{2} \), and for multigossiping it is \( \rho^\frac{1}{2} \).
For \( m = 10 \), there results .9918 and .9595 respectively.

For a star graph with \( m \) edges, with \( T_m \) the composite
gossip matrix over one period, it is straightforward to conclude
using the transfer function concept that
\[
|zI - T_m| = z(z - \frac{1}{2})^m - (\frac{z}{2})^m
\]
No analytic formula is known for the zeros. For \( m = 10 \), the
divergence rate for gossiping results as \( \rho^{1/10}(T_m) = .9759 \).
Multigossiping is not possible.

\subsection{Optimal Tree Graphs}

A table listing the number of nonisomorphic trees with
vertex count up to 40 is available in [27], and a web page
associated with Brendan McKay [28] describes a program
‘Trees’ which lists all (unlabelled) trees according to diameter,
and vertex count up to 22. Using this data, an exhaustive search of optimal trees giving fastest gossip or multigossip rates was conducted up to \( m = 21 \). For \( m = 10 \), the rates were \( .9259 \) and \( .8706 \) respectively. For \( m = 21 \), the figures were \( .9804 \) and \( .9532 \). The optimum shapes tended to involve a path of two or three nodes, each of which was the node of a star. To clarify the terminology, which is peculiar to this paper: a double star graph comprises two connected vertices, and from each of the vertices a number of rays spread out. A triple star graph comprises three vertices forming a path subgraph, and from each of the vertices a number of rays spread out. Optimal double stars tend to be balanced in the sense that the degrees of the central vertices are roughly equal, but this is not so for triple star graphs. We are not sure whether the triple star will become more balanced when \( n \) tends to be large. As far as we know there is no existing algorithm to list all unlabelled trees for large \( n \), which limits the possibility of extensive search for optimal trees even in simulations. The optimum graphs are not the same for gossiping and multigossiping however.

C. Comparison with Random Gossiping

Any comparison with random gossiping is an apples and oranges comparison, as noted in the introduction. Random gossiping convergence rates are average and the optimum graph is that with the greatest algebraic connectivity, which is always a star graph. The second largest magnitude eigenvalue of the average gossip matrix is \( 1 - \frac{1}{2m} \) in an \( m \)-edge graph. For \( m = 10 \), this is \( .9500 \).

VI. CONCLUSIONS

The main results of this paper are concerned with deterministic periodic gossiping. We have shown that for a general graph, a complete gossip matrix obtained by multiplying together single-gossip matrices corresponding to each edge of the graph has a spectrum that remains unaltered if time-adjacent gossips are interchanged, under the proviso that the associated edges share no common node, or, if they do share a common node, that they are not in any cycle of the graph.

There remain some problems whose solutions would be of interest, but they do seem challenging. First, is there a way of applying the order invariance results to graphs which contain cycles, that would yield some insight as to the range of possible convergence rates which can be obtained, and how they might be optimized? Equivalently, the latter task is to identify for a non-tree graph what complete periodic gossip sequence will result in the fastest convergence; such a sequence would not necessarily have to include all edges in the graph, so long as those edges which were included defined a connected spanning subgraph. A second issue with periodic gossiping is to allow the possibility that over one period some edges gossip more than once, and to draw conclusions on what sort of improvement can be achieved in convergence rates (this is equivalent to allowing graphs which are not simple, and having each edge gossip once in a period); ideally, one would like to understand which edges in the graph should be repeated and which should not.

REFERENCES