Vector-Valued Gossip over w-Holonomic Networks

E. Bayram*, M.-A. Belabbas*, T. Başar*

Abstract

We study the weighted average consensus problem for a gossip network of agents with vector-valued states. For a given matrix-weighted graph, the gossip process is described by a sequence of pairs of adjacent agents communicating and updating their states based on the edge matrix weight. Our key contribution is providing conditions for the convergence of this non-homogeneous Markov process as well as the characterization of its limit set. To this end, we introduce the notion of "w-holonomy" of a set of stochastic matrices, which enables the characterization of sequences of gossiping pairs resulting in reaching a desired consensus in a decentralized manner. Stated otherwise, our result characterizes the limiting behavior of infinite products of (non-commuting, possibly with absorbing states) stochastic matrices.

Keywords: Consensus; Gossiping; Non-Homogeneous Markov Processes; Holonomy; Convergence of Matrix Products; Permutation Groups

1 Introduction

Consensus entails reaching an agreement between a set of agents [1]. Many applications of distributed control systems require agents to reach a consensus for a given quantity; for example consensus to the average value of their respective initial states. An extension of average value consensus is the weighted average consensus, in which each agent contributes to the agreed-upon consensus value based on its assigned weight; see the literature review below for more details. In this paper, we study the weighted average consensus problem for a gossiping network of agents with vector-valued states. Specifically, given a matrix-weighted communication graph, we study the process whereby at each time step, two adjacent agents in the network communicate and update their states based on the matrix weight of the edge adjoining them. These two agents are called a gossiping pair and the overall process is called a **weighted gossip process** [2, 3]. It is akin to a non-homogeneous Markov process, and the study of its convergence thus reduces to the study of convergence of an infinite product of row stochastic matrices taken from a finite set. It is well known that this is a hard problem for which no general solution is known.This is due in part to the fact that, save for particular cases such as a set of commuting matrices, the order in which stochastic matrices appear in the

*E. Bayram, M.-A. Belabbas and T. Başar are with Coordinated Science Laboratory, University of Illinois, Urbana-Champaign, Email: {ebayram2,belabbas,basar1}@illinois.edu

infinite product clearly affects the limit set; in fact, this set can be a continuum (see, e.g. [4] for examples) or it can be finite [5].

We adopt here a vantage point on the consensus problem similar to the one of [5], where the notion of holonomy of a set of stochastic matrices was introduced. There, the authors used the term holonomy to indicate a change of a certain left eigenvector (referred to as *weight vector* below) corresponding to eigenvalue 1 of the product of stochastic matrices along any cycle in the graph. For each cycle in the graph, one can associate a holonomy group (see [6] for the precise definition of a holonomy group). This group characterizes how the eigenvector changes as gossiping occurs along the cycle. In [5], the authors consider gossip processes for agents with scalar states and impose that the entries of a gossip matrix be strictly positive. Together, these restrictions imply that the holonomy group for a cycle, if it exists, can only be the trivial group.

Our work here extends this earlier work in two fundamental ways. First, we allow vector-valued states for the agents. Second, and more importantly, we allow zero entries in the gossip matrices. Said otherwise, we allow for update matrices that have absorbing states (i.e., have a standard unit vector as a row). These extensions together make possible the existence of a non-trivial, finite holonomy group in a gossip process, whose investigation will be one of the main concerns of this paper.

More generally, the hereby adopted set-up raises the following questions: (1) How to understand the appearance of non-trivial holonomy groups? I.e., situations where the weight vector changes after completing one loop in the gossip graph, but then returns to its initial value after completing this loop a finite number of times? (2) Can we still guarantee the convergence of the weighted gossip process to a limit or a finite limit set by following a sequence of gossiping pairs in a decentralized manner? (3) How does the potential presence of absorbing states in gossip updates impact the consensus weight vector? These three questions are fully addressed in this paper.

To understand the phenomenon described in the first question, we introduce a concept which we call *w-holonomy* of a set of stochastic matrices. This concept helps us describe the set of stochastic matrices that possess finite orbit sets when acting on some vectors. Such matrices are the ones enabling the appearance of holonomy groups in gossip processes.

For the second question, we introduce the so-called derived graph of the communication graph G for the weight vector w, which we denote by $D_G(w)$. Infinite exhaustive closed walks in the derived graph will correspond to *allowable sequences* of updates in the gossip process. These updates can be implemented in a decentralized manner, and yield a process which converges to a finite limit set.

For the third question, the presence of zeros and ones in the update matrices can significantly impact the consensus weights in our analysis. In particular, they can lead to some agents not contributing to the consensus value average and to the appearance of permutation matrices as update matrices. In fact, even when none of the update matrices are permutation matrices, their product within the gossip iteration can result in a permutation matrix (as will be illustrated later). This fact greatly complicates the analysis, and is at the root of the existence of finite limit sets.

The paper is organized as follows: We provide a brief review of the relevant literature on distributed control and weighted average consensus in the following paragraph. We then describe the notations and conventions used in the paper at the end of this section. In Section 2, we provide a precise formulation of the problem solved in this paper. The notion of holonomy and the main results of the paper are presented in Section 3. The proof of the main theorem is provided in Section 4 along with some auxiliary results. A summary of the results of the paper and outlook for future research are provided in Section 5.

Literature Review In recent decades, there has been an increase in the applications of multi-agent systems and distributed control. These applications aim to achieve consensus among agents, as seen in works like [7, 8, 9, 10, 11, 12]. Many of these systems involve agents with multiple states, highlighting the importance of addressing weighted average consensus.

The field of weighted average consensus has seen diverse perspectives and contributions over the years, such as [1, 9, 13, 14, 15, 16, 17] and [18]. Research has tackled challenges like time delays and asynchronous information spread [19, 20, 21], as well as changing network topologies due to link failures or reconfiguration [20, 22, 23]. Moreover, works presented by [24, 25] have focused on continuous-time consensus problems. Various techniques have been used to solve consensus problems, including Lyapunov function-based methods [14, 26], and approaches inspired by ergodicity theory [27, 28, 29]. Furthermore, research efforts have addressed the constant network topology driven by the gossip process [4, 5, 30, 31]. Our work falls within the scope of this latter category of research.

Notations and conventions. We denote by $G = (V, E)$ an undirected graph, with $V = \{v_1, \dots, v_{|V|}\}$ the node set and $E = \{e_1, \dots, e_{|E|}\}$ the edge set. The edge linking nodes v_i and v_j is denoted by (v_i, v_j), a self-arc or self loop is denoted by (v_i, v_i). We call G *simple* if it has no self-loops.

Given a sequence of edges $\gamma = e_1 \cdots e_k$ in E, a node $v \in V$ is called **covered** by γ if it is incident to an edge in γ. Given a sequence $\gamma = e_1 e_2 \cdots$, we say that γ' is a string of γ if it is a contiguous subsequence, i.e., $\gamma' = e_k e_{k+1} \cdots e_l$ for some $k \geq 1$ and $l \geq k$. Let $\gamma = e_1 \cdots e_k$ be a finite sequence and e_{k+1} be an edge of G. The sequence $e_1 \cdots e_k e_{k+1}$ obtained by adding e_{k+1} to the end of γ is denoted by $\gamma \vee e_{k+1}$.

For a given simple undirected graph G as above, we denote by $\overrightarrow{G} = (V, \overrightarrow{E})$ a directed graph on the same node set and with a "bidirectionalized" edge set; precisely, $\overrightarrow{E}$ is defined as follows: we assign to every edge (v_i, v_j) of G, $i \neq j$, two directed edges $v_i v_j$ and $v_j v_i$.

We denote a walk in $\overrightarrow{G}$ either by the succession of edges or the succession of nodes visited. We say that $\gamma = v_{i_1} \cdots v_{i_k}$ is a walk in the directed graph $\overrightarrow{G}$ if $v_{i_l} v_{i_{l+1}}$, for $\ell = 1, \cdots, k-1$, is an edge of $\overrightarrow{G}$. We refer to v_{i_1} and v_{i_k} as the starting- and ending-nodes of γ, respectively. We define $\gamma^{-1} := v_{i_k} v_{i_{k-1}} \cdots v_{i_1}$. Let $\gamma' = v_{i_l} v_{i_{l+1}} \cdots v_{i_m}$ be another walk in $\vec{G}$. We denote by $\gamma \vee \gamma' = v_{i_1} \cdots v_{i_k}, v_{i_l} \cdots v_{i_m}$ the concatenation of the two walks.

If each edge e in G is labeled with some quantity, the graph G is called a *weighted graph*. If each edge e in G is labeled with some matrix A_e, then the graph G is called a *matrix-weighted graph*.

We say that $p \in \mathbb{R}^n$ is a probability vector if $p_i \geq 0$ and $\sum_{i=1}^n p_i = 1$. The set of probability vectors in $\mathbb{R}^n$ is the $(n-1)$-simplex, which is denoted by Δ^{n-1}. Its interior with respect to the standard Euclidean topology in $\mathbb{R}^n$ is denoted by $\text{int}\,\Delta^{n-1}$. Then, if $p \in \text{int}\,\Delta^{n-1}$, all entries of p are positive.

On the space of $n \times m$ real matrices, we define the following semi-norm for a given $A \in \mathbb{R}^{n\times m}$,

$$\|A\|_S := \max_{1\le j\le m} \max_{1\le i_1,i_2\le n} |a_{i_1 j} - a_{i_2 j}|.$$

It should be clear that the semi-norm of A is zero if and only if all rows of A are equal.

We let $\mathbb{1}$ be a vector of all ones, whose dimension will be clear from the context.

The *support of a matrix* $A = [a_{ij}]$, denoted by $supp(A)$, is the set of indices ij such that $a_{ij} \neq 0$. We denote by $\min A$ the smallest non zero entry of A: $\min A = \min_{ij\in supp(A)} a_{ij}$.

A matrix A with order n is *reducible* if there exist a permutation matrix P such that,

$$P^\top AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \tag{1}$$

where A_{11} and A_{12} are nonempty square matrices. If A is not reducible, then A is called *irreducible*. For convenience, we denote similarity through the permutation matrix P by $\sim_P$.

The spectral radius of a matrix A is the maximum of the modulus of the elements of its spectrum, denoted by $\rho(A)$. A circle on $\mathbb{C}$ with radius $\rho(A)$ is called *spectral circle* of the matrix A. A nonnegative irreducible matrix A having $h > 1$ eigenvalues on its spectral circle is called *imprimitive*, and h is referred to as the index of imprimitivity. If there is only one eigenvalue on the spectral circle of A, then the matrix A is *primitive*.

The *period of the i^{th} entry* of a nonnegative matrix A is defined as $\omega_A(i) := gcd\{m : [A^m]_{ii} > 0, m \in \mathbb{N}\}$. If A is irreducible, then $\omega_A(i) = \omega_A(j), \forall i, j$ [32]. This common value is called the *period* of the matrix A, denoted by ω^A.

2 Preliminaries

Let $G = (V, E)$ be an undirected simple graph on n nodes. Each node represents an agent, and each agents' state is a vector in $\mathbb{R}^m$. We denote the state vector of the agent i at time t by $x^i(t) = \begin{bmatrix} x^i_1(t) & x^i_2(t) & \dots & x^i_m(t) \end{bmatrix}^\top$. The state of the system is the concatenation of the agents' states

$$x(t) = [(x^1(t))^\top (x^2(t))^\top \dots (x^n(t))^\top]^\top \in \mathbb{R}^{nm}.$$

To an edge $(v_i, v_j) \in E$, we associate a $2m \times 2m$ row stochastic matrix $\tilde{A}_{ij} = \{a_{kl}\}$. We refer to $\tilde{A}_{ij}$ as a **pre-local stochastic matrix** for agents i and j. It describes the local information exchange when these two agents interact as part of the gossip process.

The stochastic process we analyze here is described by sequences of edges $\gamma = e_{i_1} \cdots e_{i_t} \cdots$ in G with the convention that if $e_{i_t} = (v_i, v_j)$, then agents i and j update their states according to

$$\begin{bmatrix} x^i(t+1) \\ x^j(t+1) \end{bmatrix} = \tilde{A}_{ij} \begin{bmatrix} x^i(t) \\ x^j(t) \end{bmatrix} \tag{2}$$

while the other agents' states remain constant

$$x^k(t+1) = x^k(t) \text{ for all } k \neq i, j. \tag{3}$$

The update equation for $x(t)$ is given by the **local stochastic matrix** A_{ij}. It is an nm-dimensional stochastic matrix such that the rows/columns corresponding to the states of agent i and agent j is the submatrix $\tilde{A}_{ij}$ (2) and the rows/columns corresponding to the other agents is the identity matrix (3). Then, the gossip process on edge $e_{i_t} = (v_i, v_j)$ at time t is given by

$$x(t+1) = A_{ij}x(t). \tag{4}$$

For example, the local stochastic matrix A_{12}, which is associated with the edge (v_1, v_2), is given by

$$A_{12} = \begin{bmatrix} \tilde{A}_{12} & 0_{2m\times(n-2)m} \\ 0_{(n-2)m\times 2m} & I_{(n-2)m\times(n-2)m} \end{bmatrix} \tag{5}$$

We assume here that $A_{ij} = A_{ji}$. Hence, when dealing with sequences of edges in $\vec{G}$, we associate A_{ij} with both v_iv_j and v_jv_i. This makes the graph $\vec{G}$ a directed matrix-weighted graph.

It is important to note that we allow a pre-local stochastic matrix to have zeros in any row. Consequently, it is possible to construct a valid local stochastic matrix by having a single non-zero element in any given row while setting all other elements in that row to zero (e.g. having a standard unit vector as a row). Moreover, this leads to the possibility that a valid stochastic matrix can either be a complete permutation matrix or include a permutation block. We can elaborate on the concept of a permutation block using the following definition. Given an index set $\pi \subseteq \{1, \dots, nm\}$, a stochastic matrix A has *a permutation block for* π if the submatrix of A with rows/columns indexed by π is a permutation matrix in $S_{|\pi|}$. Let π_A be the largest index set among the sets indexing permutation submatrices in A; we refer to it as the *maximal permutation index of* A.

To illustrate, consider the following instances of pre-local stochastic matrices, which are associated with the edges e_1, e_2 and e_3, respectively, in the graph $\vec{G}$:

$$\tilde{A}_{e_1} := \begin{bmatrix} 0 & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 1 & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}, \tilde{A}_{e_2} := \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ b_{53} & b_{54} & b_{55} & b_{56} \\ b_{63} & b_{64} & b_{65} & b_{66} \end{bmatrix}, \tilde{A}_{e_3} := \begin{bmatrix} c_{11} & c_{12} & c_{15} & c_{16} \\ 1 & 0 & 0 & 0 \\ c_{51} & c_{52} & c_{55} & c_{56} \\ c_{61} & c_{62} & c_{65} & c_{66} \end{bmatrix} \tag{6}$$

with a_{ij}, b_{ij} and c_{ij} real numbers in the open interval $(0, 1)$. Note that none of the pre-local matrices provided in (6) has a permutation block for any index set; in contrast, they have rows that have a single nonzero entry.

For a finite sequence $\gamma = e_1 \cdots e_k$ of edges in G and for a given pair of integers $0 \le s \le t \le k$, we define the transition matrix $P_\gamma(t:s)$ for $t \ge s+1$ as follows:

$$P_\gamma(t:s) := A_{e_t}A_{e_{t-1}} \cdots A_{e_{s+1}} \tag{7}$$

We set $P_\gamma(t:s) = I$ for $t \le s$. This allows us to write the following update for the state vector x at s :

$$x(t) = P_\gamma(t:s)x(s) \tag{8}$$

When clear from the context, we will simply write P_γ for $P_\gamma(t:s)$.

A pointed cycle in $\vec{G}$ is a walk $v_{i_1}v_{i_2} \cdots v_{i_k}v_{i_1}$, where v_{i_1} is called the *basepoint* of the cycle. Let $\vec{\mathcal{C}}$ be the set of all pointed cycles in $\vec{G}$. We define an equivalence relation on the set $\vec{\mathcal{C}}$ by saying that $C_i, C_j \in \vec{\mathcal{C}}$ are equivalent if they visit the same vertices in the same cyclic order. The set of equivalence classes of pointed cycles are referred to as *cycles*. By abuse of notation, we also denote by $\vec{\mathcal{C}}$ the set of cycles.

To each pointed cycle $C \in \vec{\mathcal{C}}$, we assign a transition matrix P_C as in (7); when we want to emphasize the basepoint, we write $P_{C,i}$ if the basepoint is v_i.

Given a cycle C in $\vec{G}$, we can reduce the dimension of vectors in Δ^{nm-1} and stochastic matrices $A \in \mathbb{R}^{nm \times nm}$ by removing rows and/or columns corresponding to nodes that are *not* covered by C. For example, if $C = v_1 v_2 v_1$, then we let $\bar{A} \in \mathbb{R}^{2m \times 2m}$ be the principal submatrix of A obtained by keeping the first $2m$ rows and columns; similarly, $\bar{w} \in \mathbb{R}^{2m}$ is the subvector of w obtained by keeping the first $2m$ entries. The cycle C, and hence the dimension of the operation $\bar{\cdot}$, will always be clear from the context.

Within the state transition matrix definition (7), we can elaborate on the concept of permutation block. Consider a pointed cycle C_1 to be a walk such that $C_1 := e_1 e_2 e_3$. Assume that the associated pre-local stochastic matrices with the edges in C_1 are provided in (6). According to (7), we then have the following:

$$\bar{P}_{C,1} = \bar{A}_{e_1} \bar{A}_{e_2} \bar{A}_{e_3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ d_{21} & d_{22} & d_{23} & 0 & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & 0 & d_{35} & d_{36} \\ d_{41} & d_{42} & d_{43} & 0 & d_{45} & d_{46} \\ d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} \\ d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} \end{bmatrix} \tag{9}$$

with d_{ij} real numbers in the open interval $(0, 1)$. The matrix $\bar{P}_{C,1}$ has a permutation block for the index set $\pi = \{1\}$. It is worth noting that, in contrast, none of the pre-local stochastic matrices provided in (6), which are associated with the edges in C_1 have a permutation block for any index set. This observation underlines that even though none of the pre-local stochastic matrices for the edges in C_1 conform to the criteria of a permutation block, the corresponding state transition matrix for the pointed cycle C_1 distinctly features such a permutation block.

3 Main Result

In this section, we introduce the main concepts and present the main result of this paper. As already mentioned, an important concept is the one of *holonomy*. In differential geometry, holonomy deals with the variation of some quantity (e.g., a vector) along the loops in a given space. If there is no variation in this quantity after completing a loop, the process is defined as *holonomic*. Otherwise, it is said to be *non-holonomic*. In our convention, if there exists a finite $k > 1$ such that the quantity does change after completing the loop once but comes back to the initial value after completing the loop k-times, the process which evolves the said quantity is called *finitely non-holonomic*. For our purpose, the quantity is a left weight vector, the process evolving the quantity is the gossip process, and the space is the graph $\vec{G}$.

Holonomy in the network. We need the following lemma to introduce the notion of holonomy.

Lemma 1. *Let C_i and C_j be two pointed cycles in a cycle C. If there exists a weight vector w such that $w = w(P_{C,i})^k$ holds for some positive k, then there exists a weight vector w' such that $w' = w'(P_{C,j})^k$ holds.*

Proof. Let $C = v_i v_1 \cdots v_j v_i$ be a cycle in $\vec{G}$. Consider the pointed cycles $C_i = v_i v_1 \cdots v_2 v_j v_i$ and $C_j = v_j v_i v_1 \cdots v_2 v_j$ in the cycle C. According to the statement, it holds that:

$$w = w(P_{C,i})^k \text{ where } P_{C,i} = A_{v_j v_i} A_{v_2 v_j} \cdots A_{v_i v_1}. \tag{10}$$

By multiplying (10) by the matrix $A_{v_j v_i}$ from the right, we get

$$\begin{aligned} wA_{v_j v_i} &= w(A_{v_j v_i} A_{v_2 v_j} \cdots A_{v_i v_1})^k A_{v_j v_i} \\ &= wA_{v_j v_i}(A_{v_2 v_j} \cdots A_{v_i v_1} A_{v_j v_i})^k = wA_{v_j v_i}(P_{C,j})^k \end{aligned}$$

It is clear that the product $wA_{v_j v_i}$ is a weight vector. This shows that there exists a weight vector w' such that $w' = w'(P_{C,j})^k$ and it is equal to $wA_{v_j v_i}$. □

Thanks to Lemma 1, we can introduce the following definition:

Definition 3.1 (Holonomic Stochastic Matrices)**.** *Let C be a cycle in $\vec{G}$ of length greater than 2 and $w^\top \in \operatorname{int} \Delta^{nm-1}$ be a weight vector. The w-*order *of C is defined as*

$$\operatorname{ord}_w C := \min\{k \geq 1 : \bar{w} = \bar{w}(\bar{P}_C)^k\},$$

*and $\operatorname{ord}_w C = 0$ if the set is empty. The local stochastic matrices A_e, $e \in C$, are said to be **w-holonomic for** C if there exists a weight vector w such that $\operatorname{ord}_w C$ is finite and non-zero.*

Note that the definition of holonomy is independent from the basepoint of C in $\vec{G}$. We observe that if $\bar{w} = \bar{w}(\bar{P}_C)^k$ holds for some positive integer k, then $\bar{w} = \bar{w}(\bar{P}_C)^{nk}$ holds for all $n \in \mathbb{N}$, thus making the set of integers k for which $\bar{w} = \bar{w}(\bar{P}_C)^k$ holds of infinite cardinality.

The local stochastic matrices A_e are *w-holonomic for G* if there exists a *common* weight vector w so that the A_e's are w-holonomic for all $C \in \vec{\mathcal{C}}$ of length greater than 2.

It is easy to see that the w-order of a given cycle C can vary as a function of w. For our purpose, we need to consider the w's that yield the largest w-order and thus define the **order of a cycle** as

$$\operatorname{ord} C := \sup_{w^\top \in \operatorname{int} \Delta^{nm-1}} \operatorname{ord}_w C. \tag{11}$$

Now, we can define the holonomy of a gossip process as follows.

- If $\operatorname{ord} C = 0$, then the process on C is *non-holonomic.*
- If $\operatorname{ord} C = 1$, then the process on C is *holonomic.*
- If $\operatorname{ord} C > 1$, then the process on C is *finitely non-holonomic.*

One can assign a (holonomy) group to the process on the cycle C if $\operatorname{ord} C \geq 1$. If $\operatorname{ord} C = 1$, the cycle C is said to have **trivial holonomy** since the corresponding group has only identity operation (trivial group). If $\operatorname{ord} C > 1$, the cycle C is said to have **non-trivial holonomy** since the corresponding group is a cyclic group with an order greater than one.

We denote **the orbit set of a weight vector** w **around a cycle** C as $\mathcal{O}_w^C$, that is,

$$\mathcal{O}_w^C := \{w_C^{(a)} \in \mathbb{R}^{nm} | w_C^{(a)} = w(P_C)^a \text{ for } a \in \mathbb{N}\}. \quad (12)$$

For the sake of simplicity, if $a = 0$, we denote the weight vector $w_C^{(0)}$ as $w_C := w$.

Graph topology. In an undirected graph G, two nodes v_i and v_j are called connected if the graph G contains a path from v_i to v_j. We need the following notion.

Definition 3.2 (Bridge)**.** *Let $G = (V, E)$ be an undirected graph. Let $\mathcal{S}_G$ be the set of pairs of nodes that are connected in G. Let $\tilde{G}_e$ be the undirected graph obtained by removing the edge e from G, that is, $\tilde{G}_e = (V, E \setminus \{e\})$. If $|\mathcal{S}_G|$ is strictly greater than $|\mathcal{S}_{\tilde{G}_e}|$, the edge e is called a* ***bridge*** *(cut-edge) of G.*

A graph G without a bridge is called **bridgeless**. We record the following result characterizing cut-edges.

Proposition 1. *An edge e in a connected graph G is a bridge if and only if no cycles of G contain both vertices adjacent to e.*

See [33, Theorem 3.3] for a proof of Proposition 1. Paraphrasing, the statement says that every node in a connected, simple, bridgeless graph G is covered by at least one cycle.

Derived graphs. We now focus on describing the allowed sequences of updates which yields the consensus at the limit for the gossip process. These will be defined via paths in what we call the *derived graph* of G by w, denoted by $D_G(w)$. We present this graph as a *geometric* graph, with nodes embedded in $\mathbb{R}^{nm}$.

Definition 3.3 (Derived Graph $D_G(\cdot)$)**.** *Let $G = (V, E)$ be a* matrix-weighted *graph on n nodes, with weights $A_e \in \mathbb{R}^{nm \times nm}$. For a weight vector $w^\top \in \operatorname{int} \Delta^{nm-1}$, the derived graph of G generated by w, denoted by $D_G(w) = (N_w, \vec{E}_w)$, is a directed matrix-weighted graph, possibly with multi-edges and self-loops, with $N_w = \bigcup_{C \in \vec{\mathcal{C}}} \mathcal{O}_w^C$. For $w_i, w_j \in \mathcal{O}_w^C$, there exists an edge $w_i w_j \in \vec{E}_w$ if $w_i = w_j P_C$ for a (pointed) cycle C; in this case, the edge weight is P_C.*

Remark 1. *Note that the derived graph being a geometric graph, ensures that elements of distinct orbit sets with the same coordinates correspond to a unique vertex in the derived graph.*

Let $e \in \vec{E}_w$ have weight P_C with $C = v_i v_{i+1} \cdots v_k v_i$. We set $\psi(e) = v_i v_{i+1} \cdots v_k v_i$. We extend the domain of ψ to the set of paths in $D_G(w)$ according to

$$\psi(\gamma \vee e) = \psi(e) \vee \psi(\gamma)$$

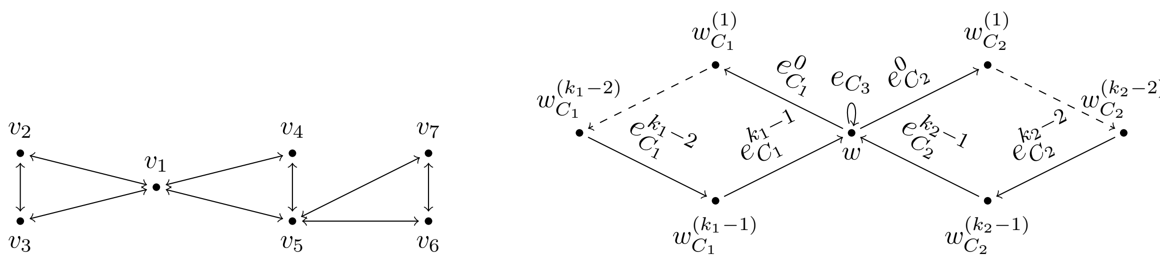
for any walk γ in $D_G(w)$. Note the order reversal in the above equation, a change that is essential for maintaining coherence. The gossip process evolves as the left multiplication of local stochastic matrices while the paths in the derived graph $D_G(w)$ correspond to the right multiplication of matrices with row vectors.

We provide an example for the derivation of $D_G(w)$.

Example 1. Consider a simple, connected, bridgeless graph with matrix weights $A_e \in \mathbb{R}^{nm\times nm}$ as depicted in Figure 1a. Consider the pointed cycles $C_1 = v_1v_2v_3v_1$, $C_2 = v_4v_1v_5v_4$ and $C_3 = v_5v_7v_6v_5$ in G. Assume that the set of local stochastic matrices $A_e, e \in E$ is w-holonomic for G and the corresponding orbit sets of the weight vector w around each of these cycles are

$$\begin{aligned}\mathcal{O}_w^{C_1} &= \{w, w_{C_1}^{(1)}, \cdots, w_{C_1}^{(k_1-2)}, w_{C_1}^{(k_1-1)}\}\\ \mathcal{O}_w^{C_2} &= \{w, w_{C_2}^{(1)}, \cdots, w_{C_2}^{(k_2-2)}, w_{C_2}^{(k_2-1)}\}\\ \mathcal{O}_w^{C_3} &= \{w\}\end{aligned}$$

where the w-order of the cycles C_1, C_2 and C_3 are k_1, k_2 and 1, respectively.



(a) The graph $\vec{G}$ (b) The graph $D_G(w)$

Figure 1: The graph $\vec{G}$ and $D_G(w)$

We denote an edge e in the graph $D_G(w)$ as $e_{C_i}^k$ if its weight is the matrix P_{C_i} and its starting node is $w_{C_i}^{(k)}$. By abuse of notation, we denote a self-loop e in the graph $D_G(w)$ as e_{C_i} if its weight is the matrix P_{C_i}. Consider the path $\gamma := e_{C_1}^{k_1-1} \vee e_{C_3} \vee e_{C_2}^0$ in the graph $D_G(w)$. “Travelling” over the path γ in $D_G(w)$ translates into the matrix product

$$w_{C_1}^{(k_1-1)} P_{C_1} P_{C_3} P_{C_2} (= w_{C_2}^{(1)}).$$

□

Now, we can state our main theorem.

Theorem 1. *Let $G = (V, E)$ be a simple, connected, bridgeless graph on n nodes with matrix-valued edge weights A_e, $e \in E$. Let $w \in \mathrm{int}\Delta^{nm-1}$ be such that the set of local stochastic matrices $\{A_e \in \mathbb{R}^{nm\times nm}, e \in E\}$ is w-holonomic for G. Then, for any infinite exhaustive closed walk γ in $D_G(w)$, we have that:*

(i) The limit set of $P_{\psi(\gamma)}$ is a finite set $\mathcal{L}$.

(ii) There exists a relabeling of the nodes such that each element of $\mathcal{L}$ can be expressed as

$$P_{\psi(\gamma)} = \begin{bmatrix} \tilde{P}_{\psi(\gamma)} & 0 \\ 0 & M_{\psi(\gamma)} \end{bmatrix}$$

where $\tilde{P}_{\psi(\gamma)}$ is a permutation matrix with rows/columns indexed by the set $\cap_{C \in \vec{\mathcal{C}}} \pi_{P_C}$ and $M_{\psi(\gamma)}$ is a block diagonal matrix.

(iii) The blocks $M^{ii}_{\psi(\gamma)}$ of $M_{\psi(\gamma)}$ are rank-one matrices.

We will provide an explicit description of the limit set $\mathcal{L}$ and the block $M^{ii}_{\psi(\gamma)}$ as a function of w in the proof of Theorem 1. The descriptions of the blocks $M^{ii}_{\psi(\gamma)}$ guarantee that they have no zero entries. It follows that if the transition matrix P_C for a cycle C has a 1 in a row (e.g. having transpose of a standard unit vector in $\mathbb{R}^{|nm|}$ as a row), then the 1 in the row either shows up in the permutation block $\tilde{P}_{\psi(\gamma)}$ for a matrix in the limit set or disappears by converging a value as a function of w in the block $M_{\psi(\gamma)}$.

Note that Theorem 1 holds for a broader class of allowable sequences but we only consider what the derived graph has generated.

4 Analysis and Proofs of Theorems

In this section, we analyze the cycles that have nonzero order and then prove Theorem 1. The major contribution of this section is to develop a necessary condition for a cycle to have nonzero order and to provide a proof of Theorem 1. We start with an analysis of the spectrum of the product of local stochastic matrices and its relationship with holonomy.

4.1 Cycles with Nonzero Order

Let $S^1 := \{z \in \mathbb{C} : |z| = 1\}$ and let $S^{1^*} := S^1 \backslash \{1\}$. The following lemma provides a necessary condition for a cycle to have non-trivial holonomy in G.

Lemma 2. *If a cycle C has non-trivial holonomy, then the matrix $\bar{P}_C$ has at least one eigenvalue in the set S^{1^*}.*

Proof. Let $\sigma(\bar{P}_C)$ be the spectrum of the matrix $\bar{P}_C$. If C has non-trivial holonomy, then there exists $\bar{w}$ such that $\bar{w} = \bar{w}(\bar{P}_C)^k$, for some $k > 1$ and $\bar{w} \neq \bar{w}(\bar{P}_C)$. The former condition implies that $\exists \lambda \in \sigma(\bar{P}_C)$ such that $\lambda^k = 1$ or, equivalently, $\lambda \in S^1$ and the latter implies that $\lambda \neq 1$, from which the result follows. □

Motivated by the previous Lemma, we now seek to describe stochastic matrices whose spectra have non-empty intersections with S^{1^*}. Note that the spectral radius of a stochastic matrix is 1 and its spectral circle is S^1.

Let A be a reducible matrix. We know that $A \sim_P B$ where B is of the form:

$$B = \begin{bmatrix} B_{11} & \cdots & B_{1r} & B_{1,r+1} & \cdots & B_{1m} \\ & \ddots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & & B_{rr} & B_{r,r+1} & \cdots & B_{rm} \\ & & & B_{r+1,r+1} & & \mathbf{0} \\ & \mathbf{0} & & & \ddots & \\ & & & \mathbf{0} & & B_{mm} \end{bmatrix} \quad (13)$$

where each principal block B_{ii} is either irreducible or the zero matrix [34]; we say that the form of B given in (13) is a *canonical form for reducible matrices*. In the literature, the subset of states corresponding to B_{kk} for $1 \leq k \leq r$ is called the k^{th} *transient class* of matrix B and the subset of states corresponding to $B_{r+j,r+j}$ for $j \geq 1$ is called the j^{th} *ergodic class* of matrix B.

Lemma 3. *Let C be a cycle with nonzero order. Then, there exists a permutation matrix P such that $\bar{P}_C \sim_P B$ where B is as in (13) with $r = 0$ and $m \geq 1$.*

Proof. If the matrix $\bar{P}_C$ is irreducible, then the result trivially holds for $P = I$. If the matrix $\bar{P}_C$ is reducible, by definition, there exists a permutation matrix P such that $\bar{P}_C \sim_P B$ where B as in (13). The principal submatrices B_{ii} for $i \leq r$ act solely on the transient states of $\bar{P}_C$. Therefore, $\rho(B_{ii}) < 1$ for $i \leq r$. This then implies that the matrices B_{ii}, $i \leq r$ are convergent. Since $\operatorname{ord} C > 0$, the matrix B cannot have convergent submatrices on the diagonal, from which the result follows. □

Paraphrasing, the statement says that the matrix $\bar{P}_C$ is permutationally similar to a matrix B such that it is block-diagonal with principal blocks being irreducible matrices. Note that a permutation matrix P is an irreducible stochatic matrix. Isolating the permutation part of $\bar{P}_C$, we can further write, up to relabeling, that

$$\bar{P}_C \sim_P \begin{bmatrix} B_C^{00} & 0 \\ 0 & M_C \end{bmatrix} = \begin{array}{c} \begin{array}{cccc} \pi_0^C & \pi_1^C & \cdots & \pi_m^C \end{array} \\ \begin{bmatrix} B_C^{00} & & & \mathbf{0} \\ & B_C^{11} & & \\ & & \ddots & \\ \mathbf{0} & & & B_C^{mm} \end{bmatrix} \end{array} \begin{array}{c} \pi_0^C \\ \pi_1^C \\ \vdots \\ \pi_m^C \end{array} . \quad (14)$$

where the principal submatrix B_C^{00} is a permutation matrix. For convenience, we denote the irreducible block $B_{r+j,r+j}$ in (13) by B_C^{jj}.

Let π_j^C be the set of indices labeling the rows/columns of the corresponding block matrix B_C^{jj} up to relabeling through the matrix P. It then follows that we have a partition of the index set $\{1, \cdots, nm\}$ induced by C, denoted by π^C, precisely $\pi^C := \{\{\pi_i^C\}_{i=0}^m\}$. As a matter of convention, we use the notation π_0^C interchangeably with π_{P_C} to refer to the largest index set among the sets that index the permutation submatrices in P_C. We will revisit the concept of the partition of the index set induced by a cycle in the next sub-section. We now aim to better understand irreducible principal blocks of the matrix M_C. Hence, without loss of generality, we can assume that $\bar{P}_C$ is irreducible (e.g., only the B_{11} block is nontrivial).

More is known about irreducible stochastic matrices. Let A be an irreducible stochastic matrix; then there exists a unique vector p satisfying

$$Ap = p,\ p > 0 \text{ and } \|p\|_1 = 1 \tag{15}$$

which is called the *Perron vector*. The Perron-Frobenius Theorem says that an irreducible *primitive* stochastic matrix A converges to a scrambling matrix whose rows are equal to the Perron vector of $A^\top$ [34], which is called the row Perron vector of the matrix A. To be more precise, $\lim_{k\to\infty} A^k = \mathbb{1}q^\top$ where $A^\top q^\top = q^\top$.

We now recall an extension of the Perron-Frobenious Theorem.

Lemma 4 (Frobenius Form)**.** *For each imprimitive matrix A with index of imprimivity $h > 1$, there exists a permutation matrix P such that $A \sim_P F$ where F is of the form:*

$$F = \begin{bmatrix} \boldsymbol{0} & A_{12} & 0 & \cdots & 0 \\ 0 & \boldsymbol{0} & A_{23} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{0} & A_{h-1,h} \\ A_{h1} & 0 & \cdots & 0 & \boldsymbol{0} \end{bmatrix}, \tag{16}$$

where the zero blocks, denoted by $\boldsymbol{0}$, on the main diagonal are square.

See [34] for a proof of Lemma 4. This is known as *Frobenius Form* for an irreducible imprimitive matrix. It is easy to see that the matrix F^h is a block diagonal matrix with blocks that are primitive.

4.2 Permutation Blocks

We now aim to gain a better understanding of permutation matrices. We demonstrate that the block matrix B_C^{00} is similar to a matrix that exhibits a permutation block structure after relabeling in (14). The following proposition establishes that the matrix B_C^{00} itself possesses a permutation block structure.

Proposition 2. *If A is a stochastic matrix which is conjugate to a permutation matrix, then A is a permutation matrix.*

To prove Proposition 2, we use the following lemma.

Lemma 5. *A matrix A is a stochastic matrix with A^{-1} also a stochastic matrix if and only if A is a permutation matrix.*

See [35] for a proof of Lemma 5. As an easy corollary, we have the following lemma:

Lemma 6. *A matrix A is a stochastic matrix with all its eigenvalues on the unit circle if and only if A is a permutation matrix.*

Using Lemma 5, we now prove Proposition 2:

Proof of Proposition 2. Let $A = PSP^{-1}$ where $S \in S_n$, and $P \in \mathbb{GL}(n)$. We have that $A^k = PS^kP^{-1}$. Since S is a permutation matrix, there exists a $k > 1$ so that $S^k = I$. Putting the previous two equalities together, we get

$$A^k = PIP^{-1} = PP^{-1} = I.$$

Hence, we have $A^{-1} = A^{k-1}$. Furthermore, by Lemma 5, the matrices A^{-1} and A are permutation matrices. ∎

4.3 Non-Permutation Blocks

Proposition 3. *For an exhaustive closed walk γ in the derived graph $D_G(w)$, there exists a permutation matrix P such that*

$$P_{\psi(\gamma)} \sim_P \begin{array}{c} \begin{array}{cccc} \pi_0^{\psi(\gamma)} & \pi_1^{\psi(\gamma)} & \cdots & \pi_l^{\psi(\gamma)} \end{array} \\ \left[\begin{array}{cccc} B_{\psi(\gamma)}^{00} & & & \boldsymbol{0} \\ & B_{\psi(\gamma)}^{11} & & \\ & & \ddots & \\ \boldsymbol{0} & & & B_{\psi(\gamma)}^{ll} \end{array}\right] \end{array} \begin{array}{c} \pi_0^{\psi(\gamma)} \\ \pi_1^{\psi(\gamma)} \\ \vdots \\ \pi_l^{\psi(\gamma)} \end{array} \tag{17}$$

where $B_{\psi(\gamma)}^{ii}$ are irreducible matrices for $i \geq 1$, the rows/columns of which are indexed by the set $\pi_i^{\psi(\gamma)}$ for $i \geq 1$ and the rows/columns of the permutation matrix $B_{\psi(\gamma)}^{00}$ are indexed by the set $\cap_{C \in \psi(\gamma)} \pi_0^C$.

To prove Proposition 3, we need to follow lemmas and definitions. First, we know that B_C^{00} is a permutation matrix for a cycle C with nonzero order due to Lemma 2. We need the following lemma to show that $B_{\gamma(\psi)}^{00}$ is also a permutation matrix:

Lemma 7. *If $C = AB$ where $C \in S_n$ and A,B are stochastic matrices order n, then $B \in S_n$ and $A \in S_n$*

Proof. We have $|\det(C)| = 1$ since it is a permutation matrix. Thanks to the homomorphism of the determinant, we have the following:

$$\begin{aligned} \det(C) &= \det(A)\det(B) \\ 1 &= |\det(A)||\det(B)| \end{aligned}$$

This shows that $|\det(A)|$ and $|\det(B)|$ are 1. It implies that $\sigma(A)$ and $\sigma(B)$ lie on the unit circle. Lemma 6 then implies that $A \in S_n$ and $B \in S_n$. □

We need the following lemma to study properties of irreducible block matrices in the submatrix $M_{\psi(\gamma)}$.

Lemma 8. *For a nonzero $\operatorname{ord}_w C$, there exists a permutation matrix P such that the matrix $(M_C)^{\operatorname{ord}_w C} \sim_P B$ where B is a block diagonal matrix with blocks that are primitive.*

Proof. From Lemma 3 and Definition 3.1, we know that the following holds:

$$\underbrace{[w_p, w_M]}_{\bar{w}}(\bar{P}_C)^{\operatorname{ord}_w C} = [w_p, w_M]\begin{bmatrix} B_C^{00} & 0 \\ 0 & M_C \end{bmatrix}^{\operatorname{ord}_w C} = [w_p(B_C^{00})^{\operatorname{ord}_w C}, \underbrace{w_M(M_C)^{\operatorname{ord}_w C}}_{*}] = \underbrace{[w_p, w_M]}_{\bar{w}} \tag{18}$$

up to relabeling. From the definition of Perron vector and $(*)$, we know that $\operatorname{ord}_w C$ is a multiple of the index of imprimivity of all block matrices B_C^{jj} for $j \geq 1$, from which the result follows. □

For convenience, we denote the matrix $(M_C)^{\mathrm{ord}_w C}$ by $(M_C)^w$. To prove Proposition 3, we need to consider the graph of a matrix. Let A be a matrix. Let $\mathbb{G}_A$ be a directed graph such that the transpose of its adjacency matrix is equal to the matrix obtained by replacing non-zero entries of the matrix A by one. The directed graph $\mathbb{G}_A$ is called *the graph of* A.

For an irreducible matrix, the period of the matrix is equal to the index of imprimivity of the matrix [36]. Then, one can easily prove the following corollary to Lemma 8:

Lemma 9. *The graph of the matrix* $(M_C)^w$ *is the union of the graphs of the submatrices* $(B_C^{jj})^w$*, each of which is strongly connected with self-arc at every node for* $j \geq 1$.

We need to introduce the composition of the graph of matrices. Let $\mathbb{G}_A$ and $\mathbb{G}_B$ be two directed graphs with the same node set V. The composition of $\mathbb{G}_A$ with $\mathbb{G}_B$, denoted by $\mathbb{G}_B \circ \mathbb{G}_A$, is a digraph with the node set V and the edge set defined as follows: $v_i v_j$ is an edge of $\mathbb{G}_B \circ \mathbb{G}_A$ whenever there is a node v_k such that $v_i v_k$ is an edge of $\mathbb{G}_A$ and $v_k v_j$ is an edge of $\mathbb{G}_B$ [5]. We have the following Lemma based on the composition definition:

Lemma 10. *For any sequence of stochastic matrices* $A_1, A_2, \cdots, A_k$ *which are all of the same size, we have that* $\mathbb{G}_{A_k \cdots A_2 A_1} = \mathbb{G}_{A_k} \circ \cdots \circ \mathbb{G}_{A_2} \circ \mathbb{G}_{A_1}$

See [22, Lem. 5] for a proof of Lemma 10. One can easily prove the following lemma:

Lemma 11. *If the graphs* $\mathbb{G}_A$ *and* $\mathbb{G}_B$ *have self-arcs at every node, then the union of the arc sets of* $\mathbb{G}_A$ *and* $\mathbb{G}_B$ *is a subset of the arc set of the graph* $\mathbb{G}_{AB}$

The condition of the Lemma says that A and B have all non-zero diagonal entries. Using the preceding lemmas, we now prove Proposition 3:

Proof of Proposition 3. By construction, a cycle in $D_G(w)$ has no chord. Hence, it is guaranteed that any exhaustive walk γ in $D_G(w)$ is a concatenation of all the cycles and the self-loops in $D_G(w)$. Let C_a and C_b be cycles in G with w-orders k_a and k_b, respectively. It is sufficient to check the particular case for γ. Assume that $\gamma := e^1_{C_a} e^2_{C_a} \cdots e^{k_a-1}_{C_a} e^1_{C_b} \cdots e^{k_b}_{C_b} \cdots e^{k_a-1}_{C_a}$. We map the walk γ to a sequence of edges in G via $\psi(.)$:

$$\psi(\gamma) = \underbrace{C_a \cdots C_a}_{|\mathcal{O}_w^{C_a}|} \cdots \underbrace{C_b \cdots C_b}_{|\mathcal{O}_w^{C_b}|} \underbrace{C_a \cdots C_a}_{|\mathcal{O}_w^{C_a}|}.$$

The matrix P_C is the product of the local stochastic matrices associated with $e \in C$. For convenience, we denote the matrix $P_C^{|\mathcal{O}_w^C|}$ by P_C^w. Then, we have the following:

$$P_{\psi(\gamma)} = P^w_{C_a} \cdots P^w_{C_b} P^w_{C_a}. \tag{19}$$

Let π^{C_a} and π^{C_b} be the partition of the index sets induced by the cycles C_a and C_b, respectively. Due to Lemma 8 and (19), if the partitions π^{C_a} and π^{C_b} are the same, the result trivially holds and $\pi^{\psi(\gamma)} = \pi^{C_a}$. It remains to prove the result if the partitions π^{C_a} and π^{C_b} are different.

First consider the permutation block in each matrix $P^w_{C_a}$ and $P^w_{C_b}$. Assume that for some index i, $i \in \pi_0^{C_a}$, but $i \notin \pi_0^{C_b}$, the permutation index set for the matrix $P^w_{C_a} P^w_{C_b}$ does not contain the index i due to Lemma 7. Owing to the above, the corresponding matrix $P_{\psi(\gamma)}$ only has permutation matrices corresponding to columns/rows associated with the states indexed by the intersection of the permutation index sets of the visited cycles by $\psi(\gamma)$, which is $\pi_0^{\psi(\gamma)} = \cap_{C\in\psi(\gamma)}\pi_0^C$. It proves the first part of the proposition. It remains to show the product of a set of irreducible blocks indexed by some elements in π^{C_a} and π^{C_b} produces a new irreducible block indexed by an element in $\pi^{\psi(\gamma)}$ under some conditions.

We first assume that the entries of the weight vector w are distinct. Then, it is easy to see that $\operatorname{ord}_w C$ is a multiple of the order of B_C^{00}. This implies that $(B_C^{00})^w$ is the identity, the graph of which consists of self-arcs at every node. Consider the graphs $\mathbb{G}_{P_{C_b}^w}$ and $\mathbb{G}_{P_{C_a}^w}$, Lemma 11 implies that they have self-arcs at every node. Due to Lemma 9 and Lemma 10, the graph $\mathbb{G}_{P_{C_b}^w P_{C_a}^w}$ has a strongly connected component for an index set $\pi_k^{C_aC_b} := \pi_i^{C_a} \cup \pi_j^{C_b}$ if $\pi_i^{C_a} \cap \pi_j^{C_b} \neq \emptyset$ for some k, i, j. Owing to the above, one can easily find the elements of partition $\pi^{\psi(\gamma)}$ of the index set induced by $\psi(\gamma)$ by induction and each element $\pi_k^{\psi(\gamma)}$ labels a strongly connected component in $\mathbb{G}_{P_{\psi(\gamma)}}$. It then follows that each block $B_{\psi(\gamma)}^{kk}$, which is indexed by the set $\pi_k^{\psi(\gamma)}$, is an irreducible block [22, Thm. 6.2.44] for $k > 1$.

We denote by $w_{\pi_0^C}$ the entries of the weight vector w which are indexed by the set π_0^C for a cycle C. Now, consider the case when some of the entries of the weight vector w are equal. On the one hand, if the corresponding indexes to these entries are not in the set π_0^C, then one can easily show that $\operatorname{ord}_w C$ is a multiple of the order of permutation matrix B_C^{00}. Therefore, $(B_C^{00})^w$ is identity. The remaining part of the proof will then be similar to the above. On the other hand, consider the case when the corresponding indices to these entries are in the set π_0^C. By definition of the orbit set (12), the matrix $(P_C)^w$ is in the stabilizer of weight vector w. But, it does not imply that $\operatorname{ord}_w C$ is a multiple of the order of the permutation B_C^{00}. For example, if all entries of the vector $w_{\pi_0^C}$ are equal, then any permutation matrix with the proper dimension will be in the stabilizer of the vector $w_{\pi_0^C}$. This implies that the matrix $(B_C^{00})^w$ is not necessarily the identity matrix, but it includes an identity block labeled by the indexes that correspond to the distinct weights in $w_{\pi_0^C}$. This observation leads us to the conclusion that the remaining part of the proof follows a similar pattern to the one discussed earlier. With this, we conclude the proof. $\square$

To elaborate on Proposition 3, consider two different exhaustive closed walks γ_1 and γ_2 in $D_w(G)$. Note that the partitions induced by $\psi(\gamma_1)$ and $\psi(\gamma_2)$ are the same, but the corresponding block matrices $B_{\psi(\gamma_1)}^{ii}$ and $B_{\psi(\gamma_2)}^{ii}$ are not necessarily the same. Indeed, the partition of the index set induced by an exhaustive walk can be found by intersection and union operation on the index sets, which are associative. This implies that the appearance order of the cycles in the exhaustive walk does not affect the corresponding index sets. We then conclude that the block structures of the matrices $P_{\psi(\gamma_1)}$ and $P_{\psi(\gamma_2)}$ are the same up to relabeling. We can thus denote the elements of the corresponding partition by π^G (e.g. $\pi^G = \pi^{\psi(\gamma)}$). On the other hand, the order in which cycles are traversed affects the order in which local stochastic matrices (which are not necessarily commutative in our work) are multiplied, and thus the block matrices are not necessarily equal.

For this section, we continue with Example 1.

Example 1 (Cont.). Assume that the number of states for each node m is 3 in the graph shown in Figure 1a; then $x(t) \in \mathbb{R}^{21}$. Consider the cycles C_1 and C_2. Assume that we have the following partition of the index set:

$$\pi^{C_1} = \{\underbrace{\{2,4,7,10,11,\cdots,20,21\}}_{\pi_0^{C_1}}, \underbrace{\{1,3,5\}}_{\pi_1^{C_1}}, \underbrace{\{6,8,9\}}_{\pi_2^{C_1}}\} \tag{20}$$

$$\pi^{C_2} = \{\underbrace{\{4,5,6,7,8,10,11,16,17,\cdots,20,21\}}_{\pi_0^{C_2}}, \underbrace{\{12,13\}}_{\pi_1^{C_2}}, \underbrace{\{1,2\}}_{\pi_2^{C_2}}, \underbrace{\{14,15\}}_{\pi_3^{C_2}}\}. \tag{21}$$

Consider the partition of the index set induced by C_2C_1. We observe that the sets $\pi_1^{C_1}$ and $\pi_2^{C_2}$ have a nonempty intersection. Hence, the set $\pi_1^{C_2C_1} := \pi_1^{C_1} \cup \pi_2^{C_2}$ is an element of the partition $\pi^{C_2C_1}$. On the other hand, the index sets $\pi_0^{C_1}$ and $\pi_0^{C_2}$ labels the maximal permutation block in the corresponding matrix. From Proposition 3, the set $\pi_0^{C_2C_1} := \pi_0^{C_1} \cap \pi_0^{C_2}$ is an element of the partition of the index set $\pi^{C_2C_1}$. Then, we obtain the following:

$$\pi^{C_2C_1} := \{\underbrace{\{4,7,10,11,16,17,\cdots,20,21\}}_{\pi_0^{C_2C_1}}, \underbrace{\{1,2,3,5\}}_{\pi_1^{C_2C_1}}, \underbrace{\{6,8,9\}}_{\pi_2^{C_2C_1}}, \underbrace{\{12,13\}}_{\pi_3^{C_2C_1}}, \underbrace{\{14,15\}}_{\pi_4^{C_2C_1}}\} \tag{22}$$

□

4.4 Proof of Theorem 1

For a cycle C in G with non-zero order, we define,

$$\epsilon_C := \min_{1 \le j \le m} (\min B_C^{jj}) \tag{23}$$

where B_C^{jj} is the irreducible block in the matrix M_C (14). We then set:

$$\epsilon := \min_{C \in G} \epsilon_C. \tag{24}$$

The coefficient of ergodicity of a stochastic matrix $A \in \mathbb{R}^{n \times n}$ is [37]

$$\mu(A) := \frac{1}{2} \max_{i,j} \sum^{n} |a_{ik} - a_{jk}|. \tag{25}$$

It is clear that $\mu(A) \le 1$ for any stochastic matrix A. The following inequality holds for any two stochastic matrices B and C [38],

$$\|BC\|_S \le \mu(B) \|C\|_S. \tag{26}$$

For any scrambling matrix A, we have the following inequality:

$$\mu(A) \le 1 - \min(A). \tag{27}$$

We now need the following lemma:

Lemma 12. *The product of any set of $l \ge \lfloor \frac{n}{2} \rfloor$ irreducible $n \times n$ stochastic matrices with positive diagonal entries is a scrambling matrix.*

See [5, Lem. 5] for a proof of Lemma 12.

Definition 4.1 (Spanning Sequence)**.** *Let $G = (V, E)$ be a simple, undirected graph. A finite sequence of edges of G is spanning if it covers a spanning tree of G. An infinite sequence of edges is spanning if it has infinitely many disjoint finite strings that are spanning.*

Lemma 13. *Let G be a bridgeless,simple, connected graph. For an exhaustive walk γ in $D_G(w)$ where the set of local stochastic matrices $\{A_e \in \mathbb{R}^{nm\times nm}, e \in E\}$ is w-holonomic for G, the sequence of edges $\psi(\gamma)$ is a spanning sequence of edges in G.*

Proof. By Definition 3.1, the set $\mathcal{O}_w^C$ is non-empty for any cycle C. Hence, there exists an edge, say e, in $D_G(w)$ which has the matrix-weight P_C. The exhaustive walk γ visits all edges in $D_G(w)$. Without loss of generality, assume that $\gamma = \gamma_a \vee e \vee \gamma_b$ where γ_a and γ_b are walks in $D_G(w)$; we have that $\psi(\gamma) = \psi(\gamma_b) \vee C \vee \psi(\gamma_a)$. This shows that the edges in any cycle C in G are visited. Because G is bridgeless, every edge in G is covered by at least one cycle. This implies that every edge in G is visited by the sequence of edges $\psi(\gamma)$, which concludes the proof. ☐

With the auxiliary results above, we are now in a position to prove Theorem 1.

Proof of Theorem 1. Let $w \in \operatorname{int}\Delta^{nm-1}$ be a weight vector. Recall that the set of local stochastic matrices $\{A_e \in \mathbb{R}^{nm\times nm}, e \in E\}$ is assumed to be w-holonomic for G. Let γ be an infinite exhaustive closed walk in $D_G(w)$.

Proposition 3 shows that the block of $P_{\psi(\gamma)}$ indexed by $\pi_0^{\psi(\gamma)} := \cap_{C\in\psi(\gamma)}\pi_0^C$ is a permutation matrix. Due to Lemma 13, all cycles in G are covered by $\psi(\gamma)$, which implies that $\pi_0^{\psi(\gamma)} = \cap_{C\in\vec{\mathcal{C}}}\pi_0^C = \pi_0^G$. Consequently, by relabeling the states so that the submatrix indexed by π_0^G is in the upper-left corner of $P_{\psi(\gamma)}$, we have proven the first part of assertion (ii) in Theorem 1.

It remains to characterize the limit set $\mathcal{L}$. Let Γ be the set of all finite exhaustive walks in $D_G(w)$. Let $\mathcal{S}_\Gamma$ be the set of permutation matrices $B^{00}_{\psi(\gamma_i)}$, for all $\gamma_i \in \Gamma$. Owing to the above paragraph, each $B^{00}_{\psi(\gamma_i)}$ is of the same dimension $|\pi_0^G|$. The set $\mathcal{S}_\Gamma \subseteq S_{|\pi_0^G|}$ is obviously finite; let $\mathcal{K}$ be the subgroup generated by the elements of $\mathcal{S}_\Gamma$. We can write an infinite exhaustive closed walk γ as the concatenation of finite exhaustive closed walks γ_i, $i \geq 1$, in $D_G(w)$. We then have $\tilde{P}_{\psi(\gamma)} = \cdots B^{00}_{\psi(\gamma_{i+1})}B^{00}_{\psi(\gamma_i)}B^{00}_{\psi(\gamma_{i-1})}\cdots$ which shows that $\tilde{P}_{\psi(\gamma)} \in \mathcal{K}$.

We now show that, given a weight vector w, the block $M_{\psi(\gamma)}$ in Theorem 1 is *uniquely* given. This implies that there exists an injection between the limit set $\mathcal{L} \ni P_{\psi(\gamma)}$ of the process and the set $\mathcal{K}$, which proves that the limit set is finite. To proceed, recall that Proposition 3 states that the matrix $M_{\psi(\gamma)}$ consists of principal block matrices that are irreducible with dimensions $|\pi_i^G|$ for $i \geq 1$ and are denoted by $M^{ii}_{\psi(\gamma)}$. Let $l_G := \max_i(|\pi_i^G|)$. Let $0 := t_0 < t_1 < t_2 \cdots$ be a monotonically increasing sequence such that every string $\psi(\gamma(t_k : t_{k+1}))$ for $k \geq 0$, has $\lfloor \frac{l_G}{2} \rfloor$ exhaustive walks in $D_G(w)$. Since $\|M_{\psi(\gamma)}\|_S$ is non-increasing by (26), it has a limit for $|\gamma| \to \infty$. We now show that $\|M^{ii}_{\psi(\gamma)}\|_S$ is 0 for $i \geq 1$.

Lemma 12 implies that $M^{ii}_{\psi(\gamma(t_k:t_{k+1}))}$ is a scrambling matrix. Plugging the lower bound (24) into (27), we have the following inequality $\mu(M^{ii}_{\psi(\gamma(t_k:t_{k+1}))}) \leq (1-\epsilon)$. Then, we can use this inequality in (26) for each block matrix $M^{ii}_{\psi(\gamma(t_k:t_{k+1}))}$ to obtain:

$$\begin{aligned}\lim_{k\to\infty} \left\|M^{ii}_{\psi(\gamma(t_k:0))}\right\|_S &\le \lim_{k\to\infty} (1-\epsilon) \left\|M^{ii}_{\psi(\gamma(t_{k-1}:0))}\right\|_S \\ &\le (1-\epsilon)^k = 0\end{aligned}$$

which implies that $\lim_{k\to\infty} \left\|M^{ii}_{\psi(\gamma(t_k:0))}\right\|_S = 0$. We conclude using [37] that $M^{ii}_{\psi(\gamma(t_k:0))}$ converges to a rank-one matrix, say $M^{ii}_{\psi(\gamma)} = \mathbb{1}p_i^\top$ for some vector $p_i \in \mathbb{R}^{|\pi_i^G|}$. This establishes asymptotic convergence in the assertion (iii) in Theorem 1.

It remains to provide an explicit characterization of the vector $p_i^\top$ such that $M^{ii}_{\psi(\gamma)} = \mathbb{1}p_i^\top$ as a function of the weight vector w. We denote by $\tilde{w}$ the entries of the weight vector w which are indexed by the set $\cup_{i\ge 1}\pi_i^G$. By construction of the derived graph $D_G(w)$, the closed walk γ induces the mapping:

$$\tilde{w} \mapsto \tilde{w} M_{\psi(\gamma)} (= \tilde{w}).$$

Then, we have the following up to relabeling,

$$\tilde{w} = \tilde{w} \begin{bmatrix} M^{11}_{\psi(\gamma)} & & \\ & M^{22}_{\psi(\gamma)} & \\ & & \ddots \end{bmatrix} = \tilde{w} \begin{bmatrix} \mathbb{1}p_1^\top & & \\ & \mathbb{1}p_2^\top & \\ & & \ddots \end{bmatrix}. \tag{28}$$

From the block structure (28), we have that,

$$p_i := \frac{[\tilde{w}]_{j\in\pi_i^G}}{\alpha_i} \text{ where } \alpha_i := \sum_{j\in\pi_i^G} \tilde{w}_j.$$

Since w is a weight vector by definition, we know that $\alpha_i \in (0,1]$. This then shows that p_i has no zero entry. □

5 Summary and outlook

In this paper, we have investigated the weighted average consensus problem for a gossiping network of agents with vector states. We have introduced the concept of w-holonomy for a set of stochastic matrices, which helped us to investigate the existence of non-trivial, finite holonomy groups in the gossip process. The allowable sequences of updates in the gossip process were obtained as closed walks in the so-called derived graph $D_G(w)$, in that any infinite exhaustive closed walks in $D_G(w)$ could be mapped to an allowable sequence of updates for the gossip process. Such sequences could be implemented in a decentralized manner, and we have shown that the corresponding infinite product of stochastic matrices converges to a finite limit set, whose elements we have explicitly characterized.

Our results have established a unified framework that connects the methodologies presented in [5] and [4]. Indeed, on the one hand, we have extended the framework of [5] by allowing gossip processes that display non-trivial holonomy groups, which results in a finite limit set for the process (as stated in Theorem 1). This is in contrast to [5, Thm. 1], where the limit set is a singleton. As a drawback of the existence of non-trivial holonomy groups, our results require following an allowable sequence of updates in the gossip process, whereas the order of the gossiping pairs does not matter in [5, Thm. 1].

On the other hand, in [4], allowable sequences have been found through a derived graph, (see [4, Definition 2.3]) nodes of which correspond to three nodes in the gossip graph that communicate simultaneously. This approach, motivated by the design of secure protocols, results in allowable sequences consisting of an infinite concatenation of triangles within the communication graph. In our work, we have developed a different perspective. Each node in our derived graph corresponds to an element of the orbit sets of the weight vector around a cycle. Consequently, our allowable sequences consist of the concatenation of cycles in the communication graph. It is worth noting that our approach requires a bridgeless communication graph, while the methodology presented by [4] requires a triangulated Laman graph.

The present work can be extended in several directions. Among others, we focus here on algorithm design, which comprises two parts. The first is to characterize the set of local stochastic matrices that yield a predefined consensus weight for agents with vector-valued states. It is relatively straightforward to use the results developed in this paper to develop a method that yields the desired local stochastic matrices given that the communication graph is *bridgeless*. Removing this topological constraint requires further research. This leads us to the second aspect, which is to remove the requirement of bridgeless graphs. To understand what it entails, we first note that the requirement can be traced back to the definition of w-holonomy involving cycles in G. Thus, one approach to remove the requirement is to modify the definition of w-holonomy to consider paths rather than cycles in the communication graph.

Namely, for a path ζ, we would redefine the orbit set in (12) as $\mathcal{O}_w^\zeta := \{w_\zeta^{(a)} \in \mathbb{R}^{nm} | w_\zeta^{(a)} = w(P_\zeta)^a \text{ for } a \in \mathbb{N}\}$ and state that if the set $\mathcal{O}_w^\zeta$ has finite and non-zero cardinality, then set of local stochastic matrices $\{A_e \in \mathbb{R}^{nm\times nm}, \forall e \in \zeta\}$ will be w-holonomic for ζ. With this modification, we can still employ the derived graph approach to characterize the allowable sequences. However, this introduces a significant challenge: the allowable sequences of updates cannot be followed in a decentralized manner (at least in an obvious manner). The ability of gossip processes to operate without a central authority is however crucial. This highlights the need to further understand the connection between topological constraints on G, and the development of decentralized update rules.

References

[1] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical association*, vol. 69, no. 345, pp. 118–121, 1974.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE transactions on information theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[3] F. Bénézit, V. Blondel, P. Thiran, J. Tsitsiklis, and M. Vetterli, "Weighted gossip: Distributed averaging using non-doubly stochastic matrices," in *2010 ieee international symposium on information theory*, pp. 1753–1757, IEEE, 2010.

[4] M.-A. Belabbas and X. Chen, “Triangulated laman graphs, local stochastic matrices, and limits of their products,” *Linear Algebra and its Applications*, vol. 619, pp. 176–209, 2021.

[5] X. Chen, M.-A. Belabbas, and J. Liu, “Gossip over holonomic graphs,” *Automatica*, vol. 136, p. 110088, 2022.

[6] W. Kuhnel, *Differential geometry : curves - surfaces - manifolds*. Providence, R.I: American Mathematical Society, 2002.

[7] Z. Zhu, J. Zhu, J. Liu, and Y. Liu, “Federated bandit: A gossiping approach,” *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 5, no. 1, 2021.

[8] R. Olfati-Saber, “Distributed kalman filtering for sensor networks,” in *2007 46th IEEE Conference on Decision and Control*, pp. 5492–5498, IEEE, 2007.

[9] M. Zhu and S. Martinez, “Discrete-time dynamic average consensus,” *Automatica*, vol. 46, no. 2, pp. 322–329, 2010.

[10] X. Chen, M.-A. Belabbas, and T. Başar, “Cluster consensus with point group symmetries,” *SIAM Journal on Control and Optimization*, vol. 55, no. 6, pp. 3869–3889, 2017.

[11] S. Sundaram and B. Gharesifard, “Distributed optimization under adversarial nodes,” *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1063–1076, 2018.

[12] B. Touri and B. Gharesifard, “Continuous-time distributed convex optimization on time-varying directed networks,” in *2015 54th IEEE Conference on Decision and Control (CDC)*, pp. 724–729, IEEE, 2015.

[13] V. D. Blondel and A. Olshevsky, “How to decide consensus? a combinatorial necessary and sufficient condition and a proof that consensus is decidable but np-hard,” *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 2707–2726, 2014.

[14] A. Nedić and J. Liu, “On convergence rate of weighted-averaging dynamics for consensus problems,” *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 766–781, 2016.

[15] J. N. Tsitsiklis, *Problems in decentralized decision making and computation*. PhD thesis, Massachusetts Institute of Technology, 1984.

[16] A. Kashyap, T. Başar, and R. Srikant, “Quantized consensus,” *Automatica*, vol. 43, no. 7, pp. 1192–1203, 2007.

[17] S. R. Etesami and T. Başar, “Convergence time for unbiased quantized consensus over static and dynamic networks,” *IEEE Transactions on Automatic Control*, vol. 61, no. 2, pp. 443–455, 2015.

[18] X. Chen, M.-A. Belabbas, and T. Başar, “Distributed averaging with linear objective maps,” *Automatica*, vol. 70, pp. 179–188, 2016.

[19] X. Li and J. Zhu, "Event-triggered weighted average consensus in networks of dynamic agents with time-varying delay," *IFAC-PapersOnLine*, vol. 54, no. 18, pp. 127–132, 2021.

[20] M. Cao, A. S. Morse, and B. Anderson, "Reaching a consensus in a dynamically changing environment: Convergence rates, measurement delays, and asynchronous events," *SIAM Journal on Control and Optimization*, vol. 47, no. 2, pp. 601–623, 2008.

[21] L. Fang, P. J. Antsaklis, and A. Tzimas, "Asynchronous consensus protocols: Preliminary results, simulations and open questions," in *Proceedings of the 44th IEEE Conference on Decision and Control*, pp. 2194–2199, IEEE, 2005.

[22] M. Cao, A. S. Morse, and B. Anderson, "Reaching a consensus in a dynamically changing environment: A graphical approach," *SIAM Journal on Control and Optimization*, vol. 47, no. 2, pp. 575–600, 2008.

[23] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on automatic control*, vol. 50, no. 5, pp. 655–661, 2005.

[24] J. M. Hendrickx and J. Tsitsiklis, "A new condition for convergence in continuous-time consensus seeking systems," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 5070–5075, IEEE, 2011.

[25] J. M. Hendrickx and J. Tsitsiklis, "Convergence of type-symmetric and cut-balanced consensus seeking systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 1, pp. 214–218, 2012.

[26] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 634–649, 2008.

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

[28] B. Touri and A. Nedic, "On ergodicity, infinite flow, and consensus in random models," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1593–1605, 2010.

[29] A. Aghajan and B. Touri, "Ergodicity of continuous-time distributed averaging dynamics: A spanning directed rooted tree approach," *IEEE Transactions on Automatic Control*, vol. 67, no. 2, pp. 918–925, 2021.

[30] J. Liu, S. Mou, A. S. Morse, B. Anderson, and C. Yu, "Deterministic gossiping," *Proceedings of the IEEE*, vol. 99, no. 9, pp. 1505–1524, 2011.

[31] F. He, A. Morse, J. Liu, and S. Mou, "Periodic gossiping," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 8718–8723, 2011.

[32] B. P. Kitchens, *Symbolic dynamics: one-sided, two-sided and countable state Markov shifts*. Springer, 1997.

[33] R. Merris, *Graph theory*. John Wiley & Sons, 2011.

[34] C. D. Meyer, *Matrix analysis and applied linear algebra*. OT ; 71, Philadelphia: Society for Industrial and Applied Mathematics, 2000.

[35] J. Ding and N. H. Rhee, "Teaching tip: When a matrix and its inverse are stochastic," *The College Mathematics Journal*, vol. 44, no. 2, pp. 108–109, 2013.

[36] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.

[37] E. Seneta, *Non-negative matrices and Markov chains*. Springer, 2006.

[38] J. Hajnal and M. S. Bartlett, "Weak ergodicity in non-homogeneous markov chains," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 54, pp. 233–246, Cambridge University Press, 1958.