Alternative Characterization of Ergodicity for Doubly Stochastic Chains

Behrouz Touri and Angelia Nedić

***Abstract*— In this paper we discuss the ergodicity of stochastic and doubly stochastic chains. We define absolute infinite flow property and show that this property is necessary for ergodicity of any stochastic chain. The proof is constructive and makes use of a rotational transformation, which we introduce and study. We then focus on doubly stochastic chains for which we prove that the absolute infinite flow property and ergodicity are equivalent. The proof of this result makes use of a special decomposition of a doubly stochastic matrix, as given by Birkhoff-von Neumann theorem. Finally, we show that a backward product of doubly stochastic matrices is convergent up to a permutation sequence and, as a result, the set of accumulation points of such a product is finite.**

I. Introduction

Ergodicity of a chain of stochastic matrices is one of the fundamental concepts in the study of time homogeneous and time inhomogeneous Markov chains. For finite state Markov chains, ergodicity reduces to the convergence of a forward product of a chain of stochastic matrices to a rank one matrix. On the other hand, ergodicity of backward products of stochastic matrices has many applications within distributed computation [10], [11], decentralized control [1], distributed optimization [10], [5], [6], and modeling of social opinion dynamics [3].

In this paper, we study the ergodicity of a backward product of matrices associated with a stochastic chain. We introduce the concept of absolute infinite flow property, which is a more restrictive property than the infinite flow property, as introduced in [8], [9]. We show that this property is indeed stronger than infinite flow by providing a concrete example of a chain that has infinite flow property but not absolute infinite flow property. We further establish that, even-though it is more restrictive, *absolute infinite flow property* is still necessary for ergodicity of any stochastic chain. This result is a non-trivial extension of our earlier results in [8], [9]. The proof technique relies on the development and study of *rotational transformation*, which we introduce as a tool for the analysis of ergodicity.

We then focus on backward product of matrices associated with a doubly stochastic chain and investigate a sufficient condition for ergodicity. We show that, in this case, *absolute infinite flow property* is also sufficient for ergodicity. The proof of this result makes use of properties of rotational transformation of the chain and the special representation of a doubly stochastic matrix as a convex combination of permutation matrices (as given by Birkhoff-von Neumann theorem).

Finally, using the properties of the rotational transformation and our results established in [7], we show that the backward product of any doubly stochastic chain is convergent up to a permutation sequence. As a consequence, we prove that the set of accumulation points of any such a product is finite, a result that does not hold for stochastic chains in general.

The paper is structured as follows. In Section II we provide our basic notation and introduce the notion of ergodicity. In Section III we introduce absolute infinite property and discuss how it compares with infinite flow property. We introduce a rotational transformation of a chain and study its properties in Section IV. In Section V we show that absolute infinite flow property is necessary for ergodicity of stochastic chains. We prove that this property is also sufficient for ergodicity of doubly stochastic chains in Section VI. In Section VII, we consider products of doubly stochastic chains that do not necessarily have absolute infinite flow property. We show that such products are always convergent up to a permutation sequence. We conclude in Section VIII.

II. Notation and Basic Terminology

We use subscripts for indexing elements of vectors and matrices. We write $x \geq 0$ or $x > 0$ if, respectively, $x_i \geq 0$ or $x_i > 0$ for all i. We use e to denote the vector with all entries equal to one. We use x^T and A^T, respectively, for the transpose of a vector x and a matrix A. A vector a is stochastic if $a \geq 0$ and $\sum_i a_i = 1$. A matrix W is stochastic if all of its rows are stochastic vectors, and it is doubly stochastic if its rows and columns are stochastic vectors. We use $[m]$ to denote the set $\{1, \ldots, m\}$. We denote a proper subset of $[m]$ by $S \subset [m]$, and we use $|S|$ to denote the cardinality of S and $\bar{S}$ to denote the set complement of S. A set $S \subset [m]$ such that $S \neq \emptyset$ is a nontrivial subset of $[m]$. *Throughout the paper we work with nontrivial subsets S of $[m]$.*

For a given $m \times m$ matrix W and a subset $S \subset [m]$, we let

$$W_S = \sum_{i \in S, j \in \bar{S}} (W_{ij} + W_{ji}).$$

A matrix P is a permutation matrix if it has only one entry equal to 1 in each row and column. For a permutation matrix P and a subset $S \subset [m]$, the set $T \subset [m]$ is the image of S under P if

$$\sum_{j \in T} e_j = \sum_{i \in S} P e_i,$$

Department of Industrial and Enterprise Systems Engineering, University of Illinois, Urbana, IL 61801, Email: {touri1,angelia} @illinois.edu. This research is supported by the National Science Foundation under CAREER grant CMMI 07-42538.

We will often write $T = P(S)$ to indicate that T is the image of the set S under the permutation P.

Throughout the paper we work with $m \times m$ stochastic matrices, unless clearly stated otherwise. We will often refer to a sequence of stochastic matrices as *stochastic chain*. We will study the ergodicity of such chains by considering the backward product of the matrices defining the chain. The backward product of a stochastic chain $\{A(k)\}$ is the infinite product given by: $\cdots A(2)A(1)A(0)$. Often, we use truncated backward product over a window of time, namely $A(k-1)A(k-2)\cdots A(s+1)A(s)$ for $k > s$. We find it convenient to denote such a truncated product by $A(k, s)$, i.e., we let

$$A(k : s) = A(k-1)\cdots A(s) \qquad \text{for } k > s \geq 0.$$

We use the following definition of ergodicity.

Definition 1: We say that a stochastic chain $\{A(k)\}$ is ergodic if for any $t_0 \geq 0$, we have

$$\lim_{k\to\infty} A(k : t_0) = ev^T(t_0)$$

for some stochastic vector $v(t_0) \in \mathbb{R}^m$.

In other words, $\{A(k)\}$ is ergodic if, for any $t_0 \geq 0$, the backward product $A(k : t_0)$ converges to a rank one matrix (which must be stochastic) as k approaches infinity.

III. Absolute Infinite Flow Property

In [8], [9], we have introduced the concept of infinite flow property and have shown that this property is necessary for ergodicity. In this section we provide a more restrictive property, namely absolute infinite flow property, and we discuss its relation with the infinite flow property.

We start our discussion by recalling the definition of the infinite flow property, as it appeared in [8], [9].

Definition 2: We say that a chain $\{A(k)\}$ has infinite flow property if

$$\sum_{k=0}^{\infty} A_S(k) = \infty \qquad \text{for all } S \subset [m].$$

The infinite flow property is a necessary condition for ergodicity of any stochastic chain (as shown in [8], [9]). However, in general, this property is not sufficient for ergodicity. For example, consider the chain $\{A(k)\}$ with

$$A(k) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \text{for } k \geq 0. \tag{1}$$

This chain is not ergodic while it has infinite flow property. Our work in [8], [9] revolves around exploring some additional properties of stochastic chains that together with infinite flow property imply ergodicity.

In this paper, we provide a more stringent property stronger than infinite flow, by strengthening the requirement of the infinite flow property. In particular, observe that in order to have infinite flow property, a chain $\{A(k)\}$ has to satisfy $\sum_{k=0}^{\infty} \sum_{i\in S, j\in\bar{S}} (A_{ij}(k) + A_{ji}(k)) = \infty$ for every $S \subset [m]$. Note that the set S is fixed and we sum, over time k, all entries $A_{ij}(k)$ and $A_{ji}(k)$ for i and j crossing from S to $\bar{S}$.

One may consider the version of the infinite flow property where the set S is also allowed to be time-varying. To do so, at first, we need to select well-behaved sequences $\{S(k)\}$ of sets $S(k) \subset [m]$, which brings us to the following definition.

Definition 3: We say that a sequence $\{S(k)\}$ of subsets of $[m]$ is regular if $|S(k)| = |S(0)|$ for any $k \geq 0$, i.e., the cardinality of $S(k)$ does not change with time.
As a simple example of a regular sequence, consider $m = 2$ and let $S(k) = \{1\}$ when k is even and $S(k) = \{2\}$ when k is odd.

We note that every regular sequence can be viewed as the image of a set $S = S(0)$ under a permutation chain and vice versa. To see this, let $\{P(k)\}$ be a permutation chain and let $S(0) = S$ for some $S \subset [m]$. Then, recursively define $S(k+1) = P(k)(S(k))$ for all $k \geq 0$. Since $\{P(k)\}$ is a permutation chain, it follows that $|S(k)| = |S(0)|$ for all $k \geq 0$. Therefore, any sequence $\{S(k)\}$ constructed in this way is regular. On the other hand, for any given regular sequence $\{S(k)\}$, let $P(k)$ be any permutation that maps the indices in $S(k)$ to the indices in $S(k+1)$ (note that such a permutation exists since $|S(k)| = |S(k+1)|$). In this case, the regular sequence $\{S(k)\}$ can be viewed as the image of the initial set $S(0)$ under the permutation chain $\{P(k)\}$.

Now, for two subsets $S, S' \subset [m]$ with the same cardinality and for a matrix A, let

$$A_{S'S} = \sum_{i\in S', j\in\bar{S}} A_{ij} + \sum_{i\in\bar{S}', j\in S} A_{ij}. \tag{2}$$

Note that for the case of $S' = S$, we have $A_{S'S} = A_S$.

The relation in (2) serves as a basis for the version of infinite flow property with time-varying sets. In particular, we define the flow over a regular set sequence $\{S(k)\}$ as follows.

Definition 4: For a chain $\{A(k)\}$ and a regular sequence $\{S(k)\}$, the flow of $\{A(k)\}$ over $\{S(k)\}$ is given by

$$F(\{A(k)\}; \{S(k)\}) = \sum_{k=0}^{\infty} A_{S(k+1)S(k)}(k).$$

We next define the extension of infinite flow property.

Definition 5: We say that a chain $\{A(k)\}$ has absolute infinite flow property if $F(\{A(k)\}; \{S(k)\}) = \infty$ for any regular sequence $\{S(k)\}$.

Note that, if in the definition of absolute infinite flow property, we restrict our attention to constant regular sequences, i.e., sequences $\{S(k)\}$ with $S(k) = S$ for all $k \geq 0$ and some $S \subset [m]$, then the definition reduces to the definition of infinite flow property. Also, notice that some chains may have infinite flow property but not absolute infinite flow property. As a concrete example, consider the chain $\{A(k)\}$ defined in Eq. (1) As discussed earlier, this chain has infinite flow property. However, it does not have absolute infinite flow property. This can be seen by considering the sequence $\{S(k)\}$ with $S(k)$ given by $S(2k) = \{1\}$ and $S(2k+1) = \{2\}$ for $k \geq 0$, which results in $F(\{A(k)\}; \{S(k)\}) = 0$. Hence, such a chain does not have absolute infinite flow property.

At the first look it may occur that absolute infinite flow property is very restrictive but, as will be shown subsequently, it is not more restrictive than ergodicity. In other words, it turns out that this property is a necessary condition for ergodicity of any stochastic chain. To show this result, we introduce a transformation of stochastic chains with respect to permutation chains and establish some properties of the transformation, as seen in the next section.

IV. Rotational Transformation

In this section we define and study a transformation of a stochastic chain. This transformation is defined with respect to a sequence $\{P(k)\}$ of permutation matrices. We refer to $\{P(k)\}$ as *permutation chain* and we reserve the notation $\{P(k)\}$ for such chains.

Definition 6: For a stochastic chain $\{A(k)\}$ and a permutation chain $\{P(k)\}$, let the chain $\{B(k)\}$ be defined by

$$B(k) = P^T(0 : k+1)A(k)P(0 : k) \quad \text{for } k \geq 0, \tag{3}$$

with $P(0 : 0) = I$. We say that the chain $\{B(k)\}$ is a rotational transformation of $\{A(k)\}$ with respect to $\{P(k)\}$.

First, note that each permutation chain induces a rotational transformation of a given stochastic chain $\{A(k)\}$. Since the product of stochastic matrices is a stochastic matrix, a rotational transformation is a stochastic chain.

Rotational transformation preserves ergodicity properties of the original stochastic chain $\{A(k)\}$. This and some other properties of a rotational transformation are discussed in the following lemma.

Lemma 1: Let $\{B(k)\}$ be the rotational transformation of a stochastic chain $\{A(k)\}$ with respect to a permutation chain $\{P(k)\}$. Then, the following holds:

(a) For any $t_0 \geq 0$ and $k > t_0$, we have
$$B(k : t_0) = P^T(k : 0)A(k : t_0)P(t_0 : 0). \tag{4}$$

(b) The chain $\{A(k)\}$ is ergodic if and only if $\{B(k)\}$ has infinite flow property.

(c) Let $\{S(k)\}$ be the trajectory of a set $S \subset [m]$ under a permutation chain $\{P(k)\}$, i.e., $S(k+1) = P(k)(S(k))$ for $k \geq 0$ with $S(0) = S$. Then, we have $A_{S(k+1)S(k)}(k) = B_S(k)$ for all $k \geq 0$.

Proof:

(a) Let $t_0 \geq 0$ be a fixed starting time. The proof follows by induction on $k > t_0$. For $k = t_0+1$, by the definition of the rotational transformation, we have
$$\begin{aligned} B(t_0+1 : t_0) = B(t_0) &= P^T(t_0+1 : 0)A(t_0)P(t_0 : 0) \\ &= P^T(k : 0)A(k : t_0)P(t_0 : 0) \end{aligned}$$
and hence, the claim is true for $k = t_0+1$. Now, suppose that the claim holds for some $k > t_0$. Then, by the induction hypothesis, we have
$$\begin{aligned} B(k+1 : t_0) &= B(k)B(k : t_0) \\ &= B(k)\left(P^T(k : 0)A(k : t_0)P(t_0 : 0)\right) \\ &= \left(P^T(k+1 : 0)A(k)P(k : 0)\right) \\ &\quad \left(P^T(k : 0)A(k : t_0)P(t_0 : 0)\right) \\ &= P^T(k+1 : 0)A(k+1 : t_0)P(t_0 : 0), \end{aligned}$$
where the last equality holds since for a permutation matrix P we have $P^TP = I$. Therefore, the claim holds for all $k > t_0$. Note that the choice of t_0 was arbitrary and, hence, the claim holds for any $k > t_0 \geq 0$.

(b) Suppose that $\{A(k)\}$ is an ergodic chain. Then, by the definition of ergodicity, we have $\lim_{k\to\infty} A(k : t_0) = ev^T(t_0)$ for a stochastic vector $v(t_0)$. Therefore, it follows that for any $\epsilon > 0$, there exists sufficiently large integer $N > 0$ such that for all $k \geq N_\epsilon$, we have $\|A_i(k : t_0) - v^T(t_0)\| \leq \epsilon$ for all $i \in [m]$, where $A_i(k : t_0)$ denotes the ith row of $A(k : t_0)$. Now, notice that $P^T(k : 0)A(k : t_0)$ is a matrix that is obtained by permuting the rows of $A(k : t_0)$. Therefore, for all $k \geq N_\epsilon$, we have $\|[P^T(k : 0)A(k : t_0)]_i - v^T(t_0)\| \leq \epsilon$ for all $i \in [m]$, where $[P^T(k : 0)A(k : t_0)]_i$ is the ith row of $P^T(k : 0)A(k : t_0)$. Thus, we have
$$\begin{aligned} &\|[P^T(k : 0)A(k : t_0)P(t_0 : 0)]_i - v^T(t_0)P(t_0 : 0)\| \\ &\qquad = \|[P^T(k : 0)A(k : t_0)]_i - v^T(t_0)\| \\ &\qquad \leq \epsilon, \end{aligned}$$
where we use $\|x^TP\| = \|x\|$ for any $x \in \mathbb{R}^m$ and any permutation matrix. The preceding relation shows that $\|B_i(k) - v^T(t_0)P(t_0 : 0)\| \leq \epsilon$ for all $i \in [m]$ and $k \geq N_\epsilon$. Note that for a fixed t_0, the vector $v^T(t_0)P(t_0 : 0)$ is stochastic. Since the preceding argument holds for any t_0 and any ϵ, it follows that $\{B(k)\}$ is ergodic. Using a similar argument one can prove that ergodicity of $\{B(k)\}$ implies ergodicity of $\{A(k)\}$.

(c) Let $\{S(k)\}$ be the trajectory of a set $S \subset [m]$ under a permutation chain $\{P(k)\}$. Then, using $B_{ij}(k) = e_i^TB(k)e_j$, we have
$$\begin{aligned} B_{ij}(k) &= e_i^TB(k)e_j \\ &= e_i^TP^T(k+1 : 0)A(k)P(k : 0)e_j \\ &= e_{i(k+1)}^TA(k)e_{j(k)} \\ &= A_{i(k+1)j(k)}(k), \end{aligned} \tag{5}$$
where $\{i(k)\}$ and $\{j(k)\}$ are the trajectories, respectively, of the singleton sets $\{i\} \subset [m]$ and $\{j\} \subset [m]$ under the permutation chain $\{P(k)\}$. Therefore by Eq. (5), we have
$$\begin{aligned} B_S(k) &= \sum_{i\in S, j\in \bar{S}} B_{ij}(k) + \sum_{i\in \bar{S}, j\in S} B_{ij}(k) \\ &= \sum_{i\in S, j\in \bar{S}} A_{i(k+1)j(k)}(k) + \sum_{i\in \bar{S}, j\in S} A_{i(k+1)j(k)}(k) \\ &= \sum_{i\in S(k+1), j\in \bar{S}(k)} A_{ij}(k) + \sum_{i\in \bar{S}(k+1), j\in S(k)} A_{ij}(k) \\ &= A_{S(k+1)S(k)}(k). \end{aligned}$$

■

Lemma 1 shows that certain properties, such as ergodicity, are preserved by rotational transformation. Using Lemma 1, one can show that a rotational transformation also preserves absolute infinite flow property (but not infinite flow property).

In the following two sections, through the use of the properties of rotational transformation, we explore connections between absolute infinite flow property and ergodicity.

V. Necessity of Absolute Infinite Flow for Ergodicity

In this section, we show the necessity of absolute infinite flow property for ergodicity of any stochastic chain. Before stating the main result let us restate a weaker version of this result, as appeared in [8], [9].

Theorem 1: Infinite flow property is necessary for ergodicity of stochastic chains.

Using the properties of rotational transformation and Theorem 1, the necessity of absolute infinite flow property for ergodicity follows.

Theorem 2: Absolute infinite flow property is a necessary condition for ergodicity of any stochastic chain $\{A(k)\}$.

Proof: Let $\{A(k)\}$ be an ergodic chain. Consider an arbitrary regular set sequence $\{S(k)\}$ and suppose that it is a trajectory of some $S \subset [m]$ under a permutation chain $\{P(k)\}$. Let $\{B(k)\}$ be the rotational transformation of $\{A(k)\}$ with respect to the permutation chain $\{P(k)\}$. Then, by Lemma 1-b, the chain $\{B(k)\}$ is also ergodic. By Theorem 1, the chain $\{B(k)\}$ must have infinite flow property, i.e., $\sum_{k=0}^{\infty} B_S(k) = \infty$ for all k. According to Lemma 1-c, there holds $B_S(k) = A_{S(k+1)S(k)}(k)$ and, therefore, it follows that $F(\{A(k)\};\{S(k)\}) = \infty$. Since $\{S(k)\}$ is an arbitrary regular sequence, we conclude that $\{A(k)\}$ has absolute infinite flow property. ■

As indicated by Theorem 2 ergodicity of a chain $\{A(k)\}$ implies absolute infinite flow property of $\{A(k)\}$. However, absolute infinite flow property is not sufficient to ensure ergodicity. As an example, consider the static chain $\{A(k)\}$ defined by

$$A(k) = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{for all } k \geq 0.$$

This chain is not ergodic since both e and e_1 are fixed points dynamics $x(k+1) = A(k)x(k)$ for $k \geq 0$ with an arbitrary $x(0) \in \mathbb{R}^3$. On the other hand, it can be seen that for any $S, S' \subset [m]$ with $|S| = |S'|$, we have $A_{S'S}(k) \geq \frac{1}{3}$, implying that the chain $\{A(k)\}$ has absolute infinite flow property.

VI. Ergodicity of Doubly Stochastic Chains

Although the converse result of Theorem 2 is not true for an arbitrary stochastic chain, it is true for doubly stochastic chains, as shown in this section. The result is based on properties of rotational transformation (Lemma 1) and the special representation of a doubly stochastic matrix in terms of permutation matrices.

Let A be an arbitrary doubly stochastic matrix. Then, by the Birkhoff-von Neumann theorem ([4], page 527), the matrix A can be written as a convex combination of permutation matrices. Specifically, there holds

$$A = \sum_{\xi=1}^{m!} \beta_\xi P^{(\xi)}, \tag{6}$$

where $\beta_\xi \geq 0$ with $\sum_{i=1}^{m!} \beta_\xi = 1$. Since scalars β_ξ are non-negative and sum up to one, there exists some $\xi^* \in \{1, \ldots, m!\}$ such that $\beta_{\xi^*} \geq \gamma$ for some $\gamma \geq \frac{1}{m!}$. Therefore, we can rewrite Eq. (6) as

$$\begin{aligned} A &= \gamma P^{(\xi^*)} + \left((\beta_{\xi^*} - \gamma) P^{(\xi^*)} + \sum_{\xi=1}^{m!} \beta_\xi P^{(\xi)} \right) \\ &= \gamma P^{(\xi^*)} + (1-\gamma)\bar{A}, \end{aligned}$$

where $\bar{A} = \frac{1}{1-\gamma}\left((\beta_{\xi^*} - \gamma) P^{(\xi^*)} + \sum_{\xi=1}^{m!} \beta_\xi P^{(\xi)} \right)$. Note that the matrix $\bar{A}$ is a convex combination of permutation matrices. Hence, by the Birkhoff-von Neumann theorem, $\bar{A}$ is doubly stochastic. All in all, we conclude that any doubly stochastic matrix A can be written as:

$$A = \gamma P + (1-\gamma)\bar{A},$$

where P is a permutation matrix, $\bar{A}$ is a doubly stochastic matrix, and γ is a scalar with $\gamma \geq \frac{1}{m!}$. Therefore, the following lemma holds.

Lemma 2: For any doubly stochastic chain $\{A(k)\}$, there exist a permutation chain $\{P(k)\}$ and a scalar $\gamma > 0$ such that

$$A(k) = \gamma P(k) + (1-\gamma)\bar{A}(k) \qquad \text{for all } k \geq 0. \tag{7}$$

For convenience, let us refer to $\{P(k)\}$ as a *permutation component* of the chain $\{A(k)\}$ and γ as a *mixing parameter* for $\{A(k)\}$.

Lemma 2 plays a crucial role in the establishment of the converse result of Theorem 2 for doubly stochastic chains. The idea is to use rotational transformation of a given doubly stochastic chain with respect to one if its permutation components and, then, to show ergodicity of the resulting chain. The advantage of using rotational transformation will become apparent from the following result.

Lemma 3: Let $\{A(k)\}$ be a doubly stochastic chain with a permutation component $\{P(k)\}$ and a mixing parameter $\gamma > 0$. Let $\{B(k)\}$ be the rotational transformation of $\{A(k)\}$ with respect to $\{P(k)\}$. Then $B_{ii}(k) \geq \gamma$ for all $i \in [m]$ and $k \geq 0$.

Proof: For all $k \geq 0$, we have

$$A(k) = \gamma P(k) + (1-\gamma)\bar{A}(k).$$

Therefore, we have

$$\begin{aligned} B(k) &= P^T(k+1:0)A(k)P(k:0) \\ &= \gamma P^T(k+1:0)P(k)P(k:0) \\ &\quad + (1-\gamma)P^T(k+1:0)\bar{A}(k)P(k:0) \\ &= \gamma I + (1-\gamma)\bar{B}(k), \end{aligned}$$

where $\bar{B}(k) = P^T(k+1:0)\bar{A}(k)P(k:0)$. Therefore, the result follows. ■

The final step toward the proof of the main result of this section is the following result which is a direct consequence of Theorem 5 in [8].

Theorem 3: Let $\{B(k)\}$ be a doubly chain with $B_{ii}(k) \geq \gamma > 0$ for all $i \in [m]$ and $k \geq 0$. Then, $\{B(k)\}$ is ergodic if and only if $\{B(k)\}$ has infinite flow property.

By putting together Theorem 3, Lemma 2, and Lemma 3, we now prove that absolute infinite flow property is sufficient for ergodicity of doubly stochastic chains. Specifically, the converse of Theorem 2 holds for doubly stochastic chains.

Theorem 4: A doubly stochastic chain is ergodic if and only if it has absolute infinite flow property.

Proof: The necessity of absolute infinite flow property is a consequence of Theorem 2. We prove the sufficiency. So let $\{A(k)\}$ be a doubly stochastic chain with absolute infinite flow property. By Lemma 2, any doubly stochastic chain has a permutation component $\{P(k)\}$ with mixing parameter $\gamma > 0$. Now, let $\{B(k)\}$ be the rotational transformation of $\{A(k)\}$ with respect to $\{P(k)\}$. Since $\{A(k)\}$ has absolute infinite flow property, by Lemma 1-c it follows that $\{B(k)\}$ has infinite flow property. On the other hand, by Lemma 3, we have $B_{ii}(k) \geq \gamma$ f or all $i \in [m]$ and $k \geq 0$. Therefore, by Theorem 3 the chain $\{B(k)\}$ is ergodic. This in turn, by Lemma 1-a implies ergodicity of $\{A(k)\}$. ■

Theorem 4 shows that for doubly stochastic chains ergodicity and absolute infinite flow property are two *equivalent* concepts.

An immediate corollary to Theorem 4 is an equivalent formulation of strong ergodicity for an inhomogeneous finite state Markov chain with doubly stochastic probability transition matrices. Generally, for a Markov chain the concept of ergodicity has two different aspects, namely *weak ergodicity* and *strong ergodicity*. A Markov chain with transition probability matrices $\{Q(k)\}$ is said to be *weakly ergodic* if $\lim_{k\to\infty} \|Q_i(t_0 : k) - Q_j(t_0 : k)\| = 0$ for all $i, j \in [m]$ and all $t_0 \geq 0$ [2]. If a Markov chain is weakly ergodic and the limit $\lim_{k\to\infty} Q(t_0 : k)$ exists for all $t_0 \geq 0$, it is said to be *strongly ergodic* (see [2]).

Using the transpose of the forward product $Q(t_0 : k)$ for $k > t_0$, we arrive at the backward product $Q^T(k : t_0) = Q^T(k) \cdots Q^T(t_0)$. Since the transpose of any doubly stochastic matrix is a doubly stochastic matrix, the resulting backward product is the product of some *doubly stochastic matrices*. Also note that for $S, S' \subset [m]$ with $|S| = |S'|$, we have $A_{S'S} = A^T_{SS'}$ which follows from the definition of $A_{S'S}$ in Eq. (2). Therefore, the following corollary immediately follows from Theorem 4.

Corollary 1: An inhomogeneous finite state Markov chain $\{x(k)\}$ with doubly stochastic transition probability matrices $\{Q(k)\}$ is strongly ergodic if and only if $\{Q^T(k)\}$ has absolute infinite flow property, or in other words:

$$\sum_{k=0}^{\infty} Q_{S(k)S(k+1)}(k) = \infty,$$

for any regular sequence $\{S(k)\}$ of sets $S(k) \subset [m]$.

VII. Infinite Flow Graph

In this section, we consider doubly stochastic chains that do not have infinite flow property and we show that their backward products always have finitely many accumulation points.

Let $\{A(k)\}$ be a doubly stochastic chain. In [7], we showed that if $A_{ii}(k) \geq \gamma > 0$ for all $i \in [m]$ and $k \geq 0$, then the limit $\lim_{k\to\infty} A(t : k)$ always exists. We established this result by considering the infinite flow graph of a given chain $\{A(k)\}$, as defined bellow.

Definition 7: The infinite flow graph of a stochastic chain $\{A(k)\}$ is the graph $G^\infty = ([m], \mathcal{E}^\infty)$ with $\mathcal{E}^\infty = \{\{i, j\} \mid \sum_{k=0}^{\infty} (A_{ij}(k) + A_{ji}(k)) = \infty, i \neq j \in [m]\}$.

In fact, the infinite flow graph of a chain $\{A(k)\}$ is a graph on m vertices with edges that carry infinite flow over the time. Note that a chain $\{A(k)\}$ has infinite flow property if and only if its infinite flow graph is connected. However, in general, the infinite flow graph may not be connected and, hence, it may have several connected components. The following result is a direct consequence of Theorem 5 in [7].

Theorem 5: Let $\{A(k)\}$ be a doubly stochastic chain with $A_{ii}(k) \geq \gamma > 0$ for all $i \in [m]$ and $k \geq 0$. Let G^∞ be the infinite flow graph of $\{A(k)\}$. Then, the limit $\Phi(t_0) = \lim_{k\to\infty} A(k : t_0)$ exists for all $t_0 \geq 0$. Moreover, $\Phi_i(t_0) = \Phi_j(t_0)$ for all $t_0 \geq 0$ if and only if i and j belong to the same connected component of G^∞.

Note that Theorem 3 is a special case of Theorem 5 when G^∞ is a connected graph.

For the main result of this section, let us define the following property.

Definition 8: We say that a chain $\{H(k)\}$ is convergent up to a permutation sequence, if there exists a permutation sequence $\{Q(k)\}$ such that $\lim_{k\to\infty} Q(k)H(k)$ exists.

Note that $\lim_{k\to\infty} Q(k)H(k)$ exists if and only if a permutation of rows of $H(k)$ is convergent which is essentially equivalent to convergence of the *set of vectors* consisting of the rows of $H(k)$, as k goes to infinity.

Now, using Theorem 5 and the properties of the rotational transformation we can prove the following result.

Theorem 6: Let $\{A(k)\}$ be a doubly stochastic chain. Then, as $k \to \infty$, the product $H(k) = A(k : t_0)$ is convergent up to a permutation sequence for all $t_0 \geq 0$.

Proof: By Lemma 2, $\{A(k)\}$ has a permutation component $\{P(k)\}$ with a mixing coefficient $\gamma > 0$. Let $\{B(k)\}$ be the rotational transformation of $\{A(k)\}$ with respect to $\{P(k)\}$. Then, by Lemma 3 we have $B_{ii}(k) \geq \gamma$ for all $i \in [m]$ and $k \geq 0$. Therefore, by Theorem 5 the product $B(k : t_0)$ converges as k goes to infinity, for all $t_0 \geq 0$. On the other hand, by Lemma 1-a, we have $B(k : t_0) = P^T(k : 0)A(k : t_0)P(t_0 : 0)$, or equivalently

$$B(k : t_0)P^T(t_0 : 0) = P^T(k : 0)A(k : t_0).$$

Since $B(k : t_0)P^T(t_0 : 0)$ is a matrix given by a fixed permutation of the columns of $B(k : t_0)$, it follows that the limit

$$\lim_{k\to\infty} B(k : t_0)P^T(t_0 : 0) = \lim_{k\to\infty} P^T(k : 0)A(k : t_0)$$

exists for all $t_0 \geq 0$. The matrix $Q(k) = P^T(k : 0)$ is a permutation matrix for all $k \geq 0$ and, hence, the product $A(k : t_0)$ is convergent up to a permutation sequence for all $t_0 \geq 0$. ■

A direct consequence of Theorem 6 is the following corollary.

Corollary 2: Let $\{A(k)\}$ be a doubly stochastic chain. Then, the set $\{A(k:0) \mid k \geq 0\}$ has at most $m!$ accumulation points.

Proof: By Theorem 6, the product $A(k:0)$ is convergent up to a permutation sequence, i.e., $B = \lim_{k\to\infty} Q(k)A(k:0)$ exists for a permutation sequence $\{Q(k)\}$. Now, consider an accumulation point $\bar{A}$ of $\{A(k:0) \mid k \geq 0\}$, as $k \to \infty$. Thus, we have $\lim_{r\to\infty} A(k_r:0) = \bar{A}$ for a subsequence $\{k_r\}$ and

$$B = \lim_{k\to\infty} Q(k)A(k:0) = \lim_{r\to\infty} Q(k_r)\bar{A}.$$

From this, we conclude that $\bar{A}$ is a permutation of the rows of B. Since there are $m!$ permutation matrices in $\mathbb{R}^{m\times m}$, it follows that the set $\{A(k:0) \mid k \geq 0\}$ has at most $m!$ accumulation points. ∎

Corollary 2 does not hold in general for an arbitrary stochastic chain. As an example, consider the chain $\{A(k)\}$ given by

$$A(k) = \begin{bmatrix} 1 & 0 & 0 \\ q_k & 0 & 1-q_k \\ 0 & 0 & 1 \end{bmatrix},$$

where the sequence $\{q_k\}$ is chosen from rational numbers in $[0,1]$ such that each rational number in $[0,1]$ appears infinitely often in this sequence (since the rational numbers can be enumerated, such a sequence exists). It can be shown that in this case the set of accumulation points of $\{A(k:0) \mid k \geq 0\}$ is the set of stochastic matrices of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ c & 0 & 1-c \\ 0 & 0 & 1 \end{bmatrix} \quad \text{where } c \in [0,1].$$

Obviously, the set of such matrices is not finite.

VIII. Conclusion

In this paper, we studied the backward product of stochastic and doubly stochastic matrices. We introduced the concept of absolute infinite flow property and showed that this property is necessary condition for ergodicity of a stochastic chain. Moreover, we proved that this property is also sufficient for ergodicity of a doubly stochastic chain. This result provides an alternative formulation of ergodicity for doubly stochastic chains. Finally, by considering arbitrary doubly stochastic chains, we showed that the set of accumulation points of the backward product of any doubly stochastic chain is finite.

References

[1] F. Bullo, J. Cortés, and S. Martínez, *Distributed control of robotic networks*, Applied Mathematics Series, Princeton University Press, 2009, Electronically available at http://coordinationbook.info.

[2] J. Hajnal, *The ergodic properties of non-homogeneous finite markov chains*, vol. 52, Proceedings of the Cambridge Philosophical Society, no. 1, 1956, pp. 67–77.

[3] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence models, analysis, and simulation*, Journal of Artificial Societies and Social Simulation **5** (2002).

[4] Roger A. Horn and Charles R. Johnson, *Matrix analysis*, Cambridge University Press, 1985.

[5] A. Nedić and A. Ozdaglar, *On the rate of convergence of distributed subgradient methods for multi-agent optimization*, Proceedings of IEEE CDC, 2007, pp. 4711–4716.

[6] ———, *Distributed subgradient methods for multi-agent optimization*, IEEE Transactions on Automatic Control **54** (2009), no. 1, 48–61.

[7] B. Touri and A. Nedić, *On approximations and ergodicity classes in random chains*, Under review, 2010.

[8] B. Touri and A. Nedić, *On ergodicity, infinite flow and consensus in random models*, `http://arxiv.com/abs/1001.1890`, 2010.

[9] ———, *When infinite flow is sufficient for ergodicity*, Proceedings of the 49th IEEE Conference on Decision and Control, 2010, pp. 7479–7486.

[10] J.N. Tsitsiklis, *Problems in decentralized decision making and computation*, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, MIT, 1984.

[11] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Transactions on Automatic Control **31** (1986), no. 9, 803–812.