Approximation and Limiting Behavior of Random Models

Behrouz Touri

Dept. of Industrial and Enterprise Systems Engineering
University of Illinois, Urbana, IL 61801
Email: touri1@illinois.edu

Angelia Nedić

Dept. of Industrial and Enterprise Systems Engineering
University of Illinois, Urbana, IL 61801
Email: angelia@illinois.edu

Abstract— In this paper, we investigate limiting behavior of linear dynamic systems driven by random stochastic matrices. We introduce and study the new concepts of partial ergodicity and ℓ_1-approximation of a given chain of stochastic matrices. We show that partial ergodicity of a chain is invariant under ℓ_1-approximations. We also introduce an infinite flow graph of a random chain and use the connectivity components of this graph to characterize the ergodicity classes of a chain. Finally, we provide a result showing that, under certain conditions, the ergodicity classes of an independent random chain and its expected counterpart are the same.

I. INTRODUCTION

Ergodicity of chains (products) of stochastic matrices is one of the central concepts in both deterministic and random settings, as it is closely related to the behavior of time-inhomogeneous linear dynamics driven by such matrices. Many of the works on this topic have focused on sufficient conditions for ergodicity and the rate of convergence in ergodic chains [1], [5], [9], [3], [6], [7]. However, not much is known about the behavior of random chains when ergodicity is not present. In this case, an array of questions arises about the limiting behavior including the existence and the characterization of the accumulation points. The main objective of this work is to study the limiting behavior of dynamics driven by random chains that are not ergodic.

The organization of the paper is as follows: in Section II, we introduce a natural generalization of ergodicity, the concept that we term *mutual ergodicity*, and we discuss some basic properties of the concept for a deterministic and random chains. Then, in Section III, we introduce the notion of ℓ_1-approximation of a chain and we show that ℓ_1-approximations do not change the ergodic properties of a given chain. We define the class $\mathcal{M}_2$ of random chains in Section IV, and we investigate ℓ_1-approximations of one sub-class in $\mathcal{M}_2$. Finally, in Section V we define *infinite flow graph* of a chain. Using this graph, and the tools developed in the preceding sections, we establish our main result of this paper. For a class of random chains, the result shows that the ergodic properties of a chain can be completely characterized by the connected component of the infinite flow graph associated with the chain. Furthermore, under certain conditions, we prove that the connected components of the

infinite flow graphs of a chain and its corresponding expected chain are the same.

Basic Notation and Terminology. We use subscripts to index the entries of vectors and matrices. We write $x \geq 0$ or $x > 0$ if $x_i \geq 0$ or $x_i > 0$ respectively holds for all i. We use $\ell_1(\mathbb{R})$ to denote the set of all scalar sequences $\{a_k\}$ such that $\sum_{k=0}^{\infty} |a_k| < \infty$. A sequence of vectors and matrices are ℓ_1-sequence if the sequence generated by each entry of the corresponding object is a sequence in $\ell_1(\mathbb{R})$. We use e_i to denote the vector in $\mathbb{R}^n$ whose ith entry is 1 and the other entries are 0. A vector a is stochastic if $a \geq 0$ and $\sum_i a_i = 1$. A matrix W is stochastic if all its rows are stochastic vectors, and it is doubly stochastic if its rows and columns are stochastic vectors. For a matrix W, the norm $\|W\|_p$ is the norm induced by the vector norm $\|x\|_p$ for $p \in [1, \infty]$. We use $[m]$ to denote the set $\{1, \ldots, m\}$. We denote a proper subset of $[m]$ by $S \subset [m]$ and its complement by $\bar{S}$. A set $S \subset [m]$ such that $S \neq \emptyset$ is a *nontrivial* subset of $[m]$. For an $m \times m$ matrix W, we use W^j to denote the jth column vector of W. We let $W_S = \sum_{i \in S, j \in \bar{S}} (W_{ij} + W_{ji})$ for $S \subset [m]$. We write $W = \text{diag}(W^{(1)}, \ldots, W^{(\tau)})$ to denote that the matrix W is a block diagonal matrix with its rth diagonal block being the matrix $W^{(r)}$. We use $G = (V, \mathcal{E})$ to denote a graph with the vertex set V and edge set $\mathcal{E}$. We write $\mathsf{E}[X]$ to denote the expected value of a random variable X. We often use *a.s.* to denote *almost surely*.

II. BASIC DEFINITIONS

Throughout this paper we work with $m \times m$ stochastic matrices. We denote the set of all such matrices by $\mathbb{S}^m$, and we let $\mathscr{F}_{\mathbb{S}^m}$ be the Borel sigma algebra on $\mathbb{S}^m$. Given a probability space $(\Omega, \mathscr{E}, \Pr(\cdot))$, a measurable function $W : (\Omega, \mathscr{E}, \Pr(\cdot)) \rightarrow \Pi_{k=0}^{\infty}(\mathbb{S}^m, \mathscr{F}_{\mathbb{S}^m})$ is referred to as a *random chain*, *random dynamic* or *random model*. We represent random chains by its coordinate map sequence $\{W(k)\}$. A random model $\{W(k)\}$ is an independent model if its coordinate sequence $\{W(k)\}$ is independent, and if, in addition, $W(k)$s are identically distributed, the model is *i.i.d.*

We discuss a generalization of the ergodicity notion for random (and deterministic) models. In particular, let $\{W(k)\}$ be a random chain. Given a starting time t_0 and a starting point $x(t_0) \in \mathbb{R}^m$, our goal is to investigate the limiting behavior of the following dynamic system:

$$x(t+1) = W(t)x(t) \qquad \text{for } t \geq t_0, \qquad (1)$$

where time $t_0 \geq 0$ is an initial time and the system is initiated at some starting state $x(t_0) \in \mathbb{R}^m$. In what follows, we often use the dynamic system (1) with a deterministic chain $\{A(t)\}$. In this case, we view $\{A(t)\}$ as a random model $\{W(t)\}$ with $W(t) = A(t)$ *a.s.* for $t \geq t_0$.

We next introduce some notions.

Definition 1: For the dynamic system in (1) driven by a deterministic chain $\{A(k)\}$, we have the following:

(a) The index $i \in [m]$ is an *ergodic index* for the chain $\{A(t)\}$ if the limit $\lim_{t \to \infty} x_i(t)$ exists for any starting time $t_0 \geq 0$ and any initial point $x(t_0) \in \mathbb{R}^m$. When each index $i \in [m]$ is ergodic, we say the chain is *partially-ergodic*.

(b) Two indices $i, j \in [m]$ with $i \neq j$, are *mutually ergodic indices* if i or j is an ergodic index and $\lim_{t \to \infty}(x_i(t) - x_j(t)) = 0$ for any initial time $t_0 \geq 0$ and initial point $x(t_0) \in \mathbb{R}^m$. We write $i \Leftrightarrow j$ when the indices i and j are mutually ergodic for the chain $\{A(k)\}$.

(c) The chain $\{A(t)\}$ is *ergodic chain* if all the indices $i \in [m]$ are mutually ergodic, i.e., $i \Leftrightarrow j$ for all $i, j \in [m]$.

For a random model $\{W(k)\}$, if any of the above properties hold almost surely, we say that the model has the corresponding property.

In view of Definition 1, when the chain $\{W(t)\}$ is partially-ergodic, the mutual ergodicity relation $\Leftrightarrow$ is an equivalence relation on the index set $[m]$, which induces the equivalency classes for the chain $\{W(t)\}$. We refer to these classes as *ergodic classes* for the chain. By Definition 1(c), we see that a chain is ergodic if and only if it has a single ergodic class.

When discussing the mutual ergodicity relation among the indices $i \in [m]$ with respect to different chains $\{W(k)\}$ and $\{U(k)\}$, we often write $\Leftrightarrow_W$ and $\Leftrightarrow_U$ to distinguish the equivalence relation among the indices for the chains $\{W(k)\}$ and $\{U(k)\}$, respectively.

Let $\Phi(k, s) = W(k) \cdots W(s)$ for $k \geq s \geq t_0$, and consider the dynamic system (1) started at $x(t_0) = e_\ell$ with $\ell \in [m]$. Then, $x(k) = \Phi(k-1, t_0)e_\ell$ and hence, $x_i(k) = \Phi_{i\ell}(k-1, 0)$ for all i. Therefore, if $\lim_{k \to \infty} x_i(k)$ exists almost surely, then $\lim_{k \to \infty} \Phi_{i\ell}(k-1, t_0)$ also exists almost surely. Particularly, if i is an ergodic index, the preceding assertion holds for any $\ell \in [m]$ and any initial time t_0, which implies that the ith row of $\Phi(k, t_0)$ converges almost surely for any $t_0 \geq 0$. The converse is also true, i.e., if the ith row of $\Phi(k, t_0)$ converges almost surely, $x_i(k)$ will converge almost surely. Similarly, the relation $i \Leftrightarrow j$ implies that

$$\lim_{k \to \infty}(\Phi_{i\ell}(k, t_0) - \Phi_{j\ell}(k, t_0)) = 0 \text{ for all } \ell \in [m] \text{ and } t_0 \geq 0.$$

Therefore, $i \Leftrightarrow j$ if and only if, for any $t_0 \geq 0$, the ith and jth row of $\Phi(k, t_0)$ converge *a.s.* to a common stochastic (random) vector $\pi(t_0)$.

For each $i \in [m]$, let E_i be the event that $\lim_{k \to \infty} x_i(k)$ exists for any initial time $t_0 \geq 0$ and $x(t_0) \in \mathbb{R}^m$. Also, for any $i, j \in [m]$ with $i \neq j$, let E_{ij} be the event that $\lim_{k \to \infty}(x_i(k) - x_j(k)) = 0$ for any $t_0 \geq 0$ and $x(t_0) \in \mathbb{R}^m$ happens on E_i. We have the following result.

Lemma 1: For an independent random chain $\{W(k)\}$, the events E_i and E_{ij} occur with probability either 0 or 1 for any $i, j \in [m]$.

Proof: We have $\Phi(k, t) = \Phi(k, t_0)\Phi(t_0, t)$ for $k \geq t_0 \geq t \geq 0$. Therefore, if $\lim_{k \to \infty} x_i(k)$ exists for starting time t_0 and any $x(t_0) \in \mathbb{R}^m$, then $\lim_{k \to \infty} x_i(k)$ exists for any starting time $t \leq t_0$ and $x(t) \in \mathbb{R}^m$. Also, if the ith row and jth row of $\Phi(k, t_0)$ converge to a common value, then the ith and jth row of $\Phi(k, t)$ converge to a common value. Thus, each E_i and E_{ij} depends on the tale of the chain $\{W(k)\}$. By Kolmogorov's 0-1 theorem ([2] page 61), the events E_i and E_{ij} occur with probability either 0 or 1. ∎

Note that it is possible that all events E_i and E_{ij} occur with probability 0. This is the case, for example, when the matrices $W(k)$ are chosen independently and uniformly from the set of $m \times m$ permutation matrices.

We have the following result, which is an extension of the result in Lemma 3 of [7] for ergodic independent chains.

Lemma 2: Let $\{W(k)\}$ be an independent random chain. If index i is ergodic for $\{W(k)\}$, then i is ergodic for the expected chain $\{\mathsf{E}[W(k)]\}$. Moreover, if $i \Leftrightarrow j$ for the chain $\{W(k)\}$, then $i \Leftrightarrow j$ for the expected chain $\{\mathsf{E}[W(k)]\}$.

Proof: If i is ergodic index for $\{W(k)\}$, then the ith row of $\Phi(k, s)$ converges *a.s.* for any $s \geq 0$, as $k \to \infty$. By the dominated convergence theorem, it follows that the ith row of $\mathsf{E}[\Phi(k, s)]$ also converges. Since the chain is independent, there holds $\mathsf{E}[\Phi(k, s)] = \mathsf{E}[W(k)] \cdots \mathsf{E}[W(s)]$. Hence, for any $s \geq 0$, the ith row of $\mathsf{E}[W(k)] \cdots \mathsf{E}[W(s)]$ converges as $k \to \infty$, implying that i is ergodic index for the chain $\{\mathsf{E}[W(k)]\}$. Similarly, we can show that $i \Leftrightarrow j$ in $\{W(k)\}$ implies $i \Leftrightarrow j$ in $\{\mathsf{E}[W(k)]\}$. ∎

III. APPROXIMATION OF CHAINS

In this section, we define ℓ_1-approximation for chains, and we show that two chains that approximate each other have the same ergodic properties. We define ℓ_1-approximation for deterministic chains, as follows.

Definition 2: A chain $\{B(k)\}$ is an ℓ_1-approximation of a chain $\{A(k)\}$ if $\{B_{ij}(k) - A_{ij}(k)\}$ is an ℓ_1-sequence for every $i, j \in [m]$, i.e.,

$$\sum_{k=0}^{\infty} |B_{ij}(k) - A_{ij}(k)| < \infty \quad \text{for all } i, j \in [m].$$

Note that the ℓ_1-approximation is an equivalence relation. Also, note that there are alternative (equivalent) definitions of ℓ_1-approximation. Since we deal with $m \times m$ matrices, we have $\sum_{k=0}^{\infty} |B_{ij}(k) - A_{ij}(k)| < \infty$ for all $i, j \in [m]$ if and only if $\sum_{k=0}^{\infty} \sum_{i,j \in [m]} |B_{ij}(k) - A_{ij}(k)| < \infty$, which yields an equivalent characterization of $\{B(k)\}$ being ℓ_1-approximation of $\{A(k)\}$. Also, since $\|\frac{1}{m}e\|_1 = 1$, we have $\frac{1}{m}\sum_{j \in [m]} |B_{ij}(k) - A_{ij}(k)| \leq \|B(k) - A(k)\|_1$ for any $i \in [m]$. Hence, $\sum_{ij} |B_{ij}(k) - A_{ij}(k)| \leq m\|B(k) - A(k)\|_1 \leq m \sum_{ij} |B_{ij}(k) - A_{ij}(k)|$. Thus, $\{B(k)\}$ is an ℓ_1-approximation of $\{A(k)\}$ if and only if $\sum_{k=0}^{\infty} \|B(k) - A(k)\|_1 < \infty$. Furthermore, since l_p-norms are equivalent

for finite dimensional matrices, *the chain $\{B(k)\}$ is an ℓ_1-approximation of $\{A(k)\}$ if and only if*

$$\sum_{k=0}^{\infty} \|B(k) - A(k)\|_p < \infty \quad \text{for any } p \in [1, \infty]. \quad (2)$$

Finally, note that for $M(k) = \max_{ij} |B_{ij}(k) - A_{ij}(k)|$, we have $\frac{1}{m}\|B(k) - A(k)\|_1 \leq M(k) \leq \|B(k) - A(k)\|_1$. Hence, $\{B(k)\}$ is an ℓ_1-approximation of $\{A(k)\}$ if and only if $\{M(k)\} \in \ell_1(\mathbb{R})$.

Now, let $\{W(k)\}$ and $\{U(k)\}$ be independent random models adapted to the same sigma-field. Because of the Kolmogorov's 0-1 law, $\sum_{k=0}^{\infty}|W_{ij}(k) - U_{ij}(k)| < \infty$ with either probability 0 or 1 for any $i, j \in [m]$. Hence, without any ambiguity, we can define two (adapted) independent random chains $\{W(k)\}$ and $\{U(k)\}$ to be an ℓ_1-approximations of each other if $\sum_{k=0}^{\infty}|W_{ij}(k) - U_{ij}(k)| < \infty$ almost surely for all $i, j \in [m]$. We will use the following result to provide an alternative characterization of ℓ_1-approximation of two random chains.

Lemma 3: Let $\{X(k)\}$ be a sequence of independent scalar random variables such that $X(k) \in [0, \alpha]$ almost surely for all k and for some $\alpha > 0$. Then, $\sum_{k=0}^{\infty} X(k) < \infty$ almost surely if and only if $\sum_{k=0}^{\infty} \mathsf{E}[X(k)] < \infty$.

Proof: If $\sum_{k=0}^{\infty} \mathsf{E}[X(k)] < \infty$, then by the monotone convergence theorem it follows that $\sum_{k=0}^{\infty} X(k) < \infty$. For the converse, suppose that $\sum_{k=0}^{\infty} X(k) < \infty$ almost surely. Then, by Kolmogorov's three series theorem ([2] page 63), it follows that $\sum_{k=0}^{\infty} \mathsf{E}\left[X(k)\mathbf{1}_{\{|X(k)|\leq\alpha\}}\right] < \infty$ where $\mathbf{1}_E$ is the indicator function of the event E. Since $X(k) \in [0, \alpha]$, we have $X(k)\mathbf{1}_{\{|X(k)|\leq\alpha\}} = X(k)$ and the result follows. ∎

For two stochastic matrices W and U, we have $0 \leq \max_{ij}|W_{ij} - U_{ij}| \leq 2$. From this and Lemma 3, we have the following alternative characterization of ℓ_1-approximation.

Corollary 1: Let $\{W(k)\}$ and $\{U(k)\}$ be two independent random models adapted to the same sigma-field. Then, the chain $\{W(k)\}$ is an ℓ_1-approximation of $\{U(k)\}$ if and only if $\sum_{k=0}^{\infty} \mathsf{E}[\max_{ij}|W_{ij}(k) - U_{ij}(k)|] < \infty$.

Our next result shows that the ergodicity of an index as well as the ergodic classes of a chain are invariant under ℓ_1-approximations of the chain. These properties, while important in their own right, provide us with some essential tools for later use.

Lemma 4: (Approximation lemma) Suppose that a deterministic chain $\{B(k)\}$ is an ℓ_1-approximation of a deterministic chain $\{A(k)\}$. Then, $\{A(k)\}$ and $\{B(k)\}$ have the same ergodic properties, i.e., i is an ergodic index for $\{A(k)\}$ if and only if i is ergodic index for $\{B(k)\}$, and $i \Leftrightarrow_A j$ if and only if $i \Leftrightarrow_B j$.

Proof: Recalling the discussion following Definition 1, it suffice to consider initial states $x(t_0) = e_\ell$ with $\ell \in [m]$ and arbitrary t_0. Thus, let $t_0 \geq 0$ and $\ell \in [m]$ be arbitrary. Let $\{x(k)\}$ be the sequence generated by (1) driven by $\{A(k)\}$, i.e., $x(k+1) = A(k)x(k)$ for $k \geq t_0$ with $x(t_0) = e_\ell$. Let $\epsilon > 0$ be arbitrary. Since $\{B(k)\}$ is an ℓ_1-approximation of $\{A(k)\}$ for the given ϵ, there exists $N_\epsilon \geq t_0$ such that $\sum_{k=N_\epsilon}^{\infty} \|B(k) - A(k)\|_\infty < \epsilon$ (see (2)).

Now, define $z(k+1) = B(k)z(k)$ for $k \geq N_\epsilon$ with initial condition $z(N_\epsilon) = x(N_\epsilon)$. By induction on k, we show that for all $k \geq N_\epsilon$,

$$\|x(k+1) - z(k+1)\|_\infty \leq \sum_{t=N_\epsilon}^{k} \|A(t) - B(t)\|_\infty. \quad (3)$$

To see this, note that for any $k \geq N_\epsilon$, we have

$$x(k+1) = A(k)x(k) = (A(k) - B(k))x(k) + B(k)x(k).$$

Since the matrices $A(k)$ are stochastic and $x(t_0) = e_\ell$, we have $0 \leq x_i(k) \leq 1$ for any $k \geq t_0$. Therefore,

$$\|(A(k) - B(k))x(k)\|_\infty \leq \|A(k) - B(k)\|_\infty. \quad (4)$$

Therefore, $\|x(N_\epsilon + 1) - B(N_\epsilon)x(N_\epsilon)\|_\infty \leq \|A(N_\epsilon) - B(N_\epsilon)\|_\infty$. Since $z(N_\epsilon) = x(N_\epsilon)$, by the definition of $z(k)$, we have $z(N_\epsilon + 1) = B(N_\epsilon)x(N_\epsilon)$, implying

$$\|x(N_\epsilon + 1) - z(N_\epsilon + 1)\|_\infty \leq \|A(N_\epsilon) - B(N_\epsilon)\|_\infty.$$

Thus, relation (3) is valid for $k = N_\epsilon$. Now, suppose that relation (3) holds for $k - 1$, i.e., $\|x(k) - z(k)\|_\infty \leq \sum_{t=N_\epsilon}^{k-1} \|A(t) - B(t)\|_\infty$. Using the definitions of $x(k)$ and $z(k)$, Eq. (4) and the triangle inequality, we have

$$\begin{aligned}
\|x(k+1) - z(k+1)\|_\infty &= \|A(k)x(k) - B(k)z(k)\|_\infty \\
&\leq \|(A(k) - B(k))x(k)\|_\infty + \|B(k)(x(k) - z(k))\|_\infty \\
&\leq \|A(k) - B(k)\|_\infty \|x(k)\|_\infty + \|B(k)\|_\infty \|x(k) - z(k)\|_\infty \\
&\leq \|A(k) - B(k)\|_\infty + \|x(k) - z(k)\|_\infty
\end{aligned}$$

where the last inequality follows by $\|x(k)\|_\infty \leq 1$ and $\|B(k)\|_\infty = 1$ (since $B(k)$ is a stochastic matrix). Now by using the induction hypothesis, we obtain $\|x(k+1) - z(k+1)\|_\infty \leq \sum_{t=N_\epsilon}^{k} \|A(t) - B(t)\|_\infty$, thus showing relation (3). Note that from relation (3) and the choice of N_ϵ, we have

$$\|x(k+1) - z(k+1)\|_\infty \leq \epsilon \quad \text{for all } k \geq N_\epsilon. \quad (5)$$

Now if i is an ergodic index for $\{B(k)\}$, then $\lim_{k\to\infty} z_i(k) = z_\infty$ exists. Hence, there exists $M \geq N_\epsilon$ such that $|z_i(k) - z_\infty| \leq \epsilon$. Therefore, for $k, k' \geq M$,

$$\begin{aligned}
|x_i(k) - x_i(k')| &= |(x_i(k) - z_i(k)) + (x_i(k') - z_i(k')) \\
&\quad + (z_i(k) - z_\infty) + (z_i(k') - z_\infty)| \leq 4\epsilon,
\end{aligned}$$

which holds by the triangle inequality and inequality (5). But this argument holds for any $\epsilon > 0$ and hence, $\{x_i(k)\}$ is a Cauchy sequence in $[0, 1]$ which proves that $\lim_{i\to\infty} x_i(k)$ exists. Here we did not use any assumption about the starting time being 0. Therefore, using the same argument, we obtain that i is an ergodic index for $\{A(k)\}$.

Suppose now that $i \Leftrightarrow_B j$, then i and j are ergodic indices for $\{B(k)\}$ and by the preceding discussion they are also ergodic indices for $\{B(k)\}$. To show $i \Leftrightarrow_A j$, it remains to prove that $x_i(k) - x_j(k) \to 0$. Since $i \Leftrightarrow_A j$, we have $\lim_{k\to\infty}(z_i(k) - z_j(k)) = 0$. Thus, by using inequality (5), we can see $\limsup_{k\to\infty}|x_i(k) - x_j(k)| \leq 2\epsilon$ for any $\epsilon > 0$. Therefore, $\limsup_{k\to\infty}|x_i(k) - x_j(k)| = 0$. The converse statement follows by the same argument. ∎

The result of Approximation Lemma 4 has immediate implications for random chains. Specifically, if a random

chain $\{W(k)\}$ is an ℓ_1-approximation of another random chain $\{U(k)\}$, then the result of Approximation Lemma 4 holds for almost all realizations of $\{W(k)\}$ and $\{U(k)\}$. Hence, ℓ_1-approximations preserve the ergodic properties of random chains almost surely.

The following result is an immediate consequence of Approximation Lemma 4.

Corollary 2: The class of deterministic ergodic chains is closed under ℓ_1-approximations. Also, the class of ergodic random chains is closed under ℓ_1-approximations.

IV. Approximation and Ergodicity on Class $\mathcal{M}_2$

As in [7], we say a chain $\{W(k)\}$ has a common steady state π in expectation if π is a stochastic vector such that $\pi^T \mathsf{E}[W(k)] = \pi^T$ for all k. In [7], we studied the ergodicity properties of such chains by using the function

$$V(x) = \sum_{i=1}^{m} \pi_i (x_i - \pi^T x)^2, \qquad (6)$$

and we have shown the following result ([7], Theorem 4).

Theorem 1: Let the random model $\{W(k)\}$ be independent with a common steady state π in expectation. Let the sequence $\{x(k)\}$ be given by (1) with $t_0 = 0$ and arbitrary $x(0) \in \mathbb{R}^m$. Then, we almost surely have for all $k \geq 0$,

$$\mathsf{E}[V(x(k+1)) \mid x(k)] \leq V(x(k)) - \sum_{i<j} H_{ij}(k) (x_i(k) - x_j(k))^2,$$

where $H(k) = \mathsf{E}[W^T(k)DW(k)]$ and $D = \mathrm{diag}(\pi)$.

Regarding this result, a natural question is: if $\{U(k)\}$ is an ℓ_1-approximation of the chain $\{W(k)\}$ for which Theorem 1 holds, will the result hold for the chain $\{U(k)\}$? The answer to the question is not obvious, since when $\{U(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$, the chain $\{U(k)\}$ need not have a common steady state in expectation. To answer the question, we define a class of random chains as follows.

Definition 3: Let $\mathcal{M}_2$ be the class of independent random models $\{W(k)\}$ with the following property: there exists a vector $u > 0$ such that for any $t_0 \geq 0$ and $x(t_0) \in \mathbb{R}^m$, the sequence $\{x(k)\}$ of the dynamic system (1) is such that

$$\sum_{k=t_0}^{\infty} \sum_{i<j} H_{ij}(k) \mathsf{E}[(x_i(k) - x_j(k))^2] < \infty \quad a.s.,$$

where $H(k) = \mathsf{E}[W^T(k)\mathrm{diag}(u)W(k)]$. The vector u is termed an *asymptotic distribution* for the chain $\{W(k)\}$.

Note that since the matrix $\mathsf{E}[W^T(k)\mathrm{diag}(u)W(k)]$ is a linear function of u and the expectation operation $\mathsf{E}[\cdot]$ is linear, if u is an asymptotic distribution for $\{W(k)\}$, then the vector $\lambda\pi$ for any $\lambda > 0$ is also an asymptotic distribution for $\{W(k)\}$. Thus, by a proper normalization of the vector u in Definition 3, we can obtain a stochastic vector as an asymptotic distribution for the chain.

By Definition 3, the class of independent random models that have some common steady state in expectation is a subclass of $\mathcal{M}_2$. We next show that ℓ_1-approximations of such models yield chains that stay in the class $\mathcal{M}_2$.

We now state a direct consequence of Theorem 1, which we need in the sequel.

Lemma 5: Let W be a random matrix and π be a stochastic vector such that $\mathsf{E}[\pi^T W] = \pi$. Then, for any vector $x \in \mathbb{R}^m$,

$$\mathsf{E}[V(Wx)] \leq V(x) - \sum_{i<j} H_{ij}(x_i - x_j)^2,$$

where $H = \mathsf{E}[W^T \mathrm{diag}(\pi)W]$.
Using this result, we establish the following theorem.

Theorem 2: Let $\{W(k)\}$ be an independent random chain with a common steady state π in expectation. Let $\{U(k)\}$ be an ℓ_1-approximation of $\{W(k)\}$. Then, $\{U(k)\} \in \mathcal{M}_2$ and π is an asymptotic distribution for $\{U(k)\}$.

Proof: Let $D = \mathrm{diag}(\pi)$. Let $\{z(k)\}$ be the sequence resulted from (1) with the chain $\{U(k)\}$, i.e., $z(k+1) = U(k)z(k)$ for $k \geq t_0$. Without loss of generality, let $t_0 = 0$ and $z(0) \in [0,1]^m$. We have $\mathsf{E}[V(z(k+1)) \mid z(k)] = \mathsf{E}[z^T(k+1)(D - \pi\pi^T)z(k+1) \mid z(k)]$. By letting $y = (U(k) - W(k))z(k)$, we write

$$\mathsf{E}[V(z(k+1)) \mid z(k)]$$
$$= \mathsf{E}\Big[(W(k)z(k) + y)^T (D - \pi\pi^T)(W(k)z(k) + y) \mid z(k)\Big]$$
$$\leq \mathsf{E}[V(W(k)z(k)) \mid z(k)]$$
$$\quad + 2\mathsf{E}\Big[y^T(D - \pi\pi^T)z(k+1) \mid z(k)\Big], \qquad (7)$$

which holds since $V(y) \geq 0$. Define $H(k) = \mathsf{E}[W^T(k)DW(k)]$ for $k \geq 0$. and note that by Lemma 5, we have $\mathsf{E}[V(W(k)z(k)) \mid z(k)] \leq V(z(k)) - \sum_{i<j} H_{ij}(k)(z_i(k) - z_j(k))^2$. Thus,

$$\mathsf{E}[V(z(k+1)) \mid z(k)] \leq V(z(k))$$
$$\quad - \sum_{i<j} H_{ij}(k)(z_i(k) - z_j(k))^2$$
$$\quad + 2\mathsf{E}\big[y^T(D - \pi\pi^T)z(k+1) \mid z(k)\big]. \qquad (8)$$

Since $U(k)$s are stochastic, $z(k) \in [0,1]^m$ for all k. Each entry of $D - \pi\pi^T$ is bounded by 1 in the absolute value. Furthermore, $|y_i| \leq \|W(k) - U(k)\|_\infty$ for all $i \in [m]$. Therefore, $y^T(D - \pi\pi^T)z(k+1) \leq m^2 \|W(k) - U(k)\|_\infty$ for all k. By the independency assumption, $U(k)$ and $W(k)$ are independent of $z(k)$. Thus, from Eqs. (7) and (8), it follows

$$\mathsf{E}[V(z(k+1)) \mid z(k)] \leq V(z(k))$$
$$- \sum_{i<j} H_{ij}(k)(z_i(k) - z_j(k))^2 + 2m^2 \mathsf{E}[\|W(k) - U(k)\|_\infty].$$

Since $\{U(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$, we have $\sum_{k=0}^{\infty} \mathsf{E}[\|W(k) - U(k)\|_\infty] < \infty$. By Robbins-Siegmund supermartingale result ([4], page 50), it follows

$$\sum_{k=0}^{\infty} \sum_{i<j} H_{ij}(k)(z_i(k) - z_j(k))^2 < \infty \quad a.s. \qquad (9)$$

To complete the proof, we show that the difference between the two sums $\sum_{k=0}^{\infty} \sum_{i<j} H_{ij}(k)(z_i(k) - z_j(k))^2$ and $\sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)(z_i(k) - z_j(k))^2$ is finite almost surely, where $L(k) = \mathsf{E}[U^T(k)DU(k)]$ for $k \geq 0$. Let $T(k) =$

$U(k) - W(k)$. Since $|T_{\ell i}(k)| = |U_{\ell i}(k) - W_{\ell i}(k)| \leq \|T(k)\|_\infty$, we have

$$L_{ij}(k) = \sum_{\ell=1}^{m} \mathsf{E}[\pi_\ell U_{\ell i}(k) U_{\ell j}(k)]$$

$$\leq \sum_{\ell=1}^{m} \mathsf{E}[\pi_\ell (W_{\ell i}(k) + T_{\ell i}(k))(W_{\ell j}(k) + T_{\ell j}(k))]$$

$$\leq H_{ij}(k) + \alpha \mathsf{E}[\|T(k)\|_\infty],$$

for some $\alpha > 0$ which holds since $\pi_\ell, W_{\ell i}(k), W_{\ell j}(k) \in [0,1]$. Since $(z_i(k) - z_j(k))^2 \leq 1$, we have

$$\sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)(z_i(k) - z_j(k))^2$$

$$\leq \sum_{k=0}^{\infty} \sum_{i<j} H_{ij}(k)(z_i(k) - z_j(k))^2 + m^2\alpha \sum_{k=0}^{\infty} \mathsf{E}[\|T(k)\|_\infty].$$

We note that $\sum_{k=0}^{\infty} \mathsf{E}[\|T(k)\|_\infty] < \infty$ since $\{U(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$. By using this and inequality (9), we obtain $\sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)(z_i(k) - z_j(k))^2 < \infty$ *a.s.*, thus showing that $\{U(k)\} \in \mathscr{M}_2$ and that π is an asymptotic distribution for $\{U(k)\}$. ∎

We now give some definitions, as introduced in [7].

Definition 4: ([7]) A random model $\{W(k)\}$ has *weak-feedback property* if there exists a scalar $\gamma > 0$ such that $\mathsf{E}\left[W^i(k)^T W^j(k)\right] \geq \gamma \mathsf{E}[W_{ij}(k) + W_{ji}(k)]$ for all $k \geq 0$ and $i, j \in [m]$. The scalar γ is termed a feedback constant for $\{W(k)\}$.

Definition 5: ([7]) A chain $\{W(k)\}$ has *infinite flow property* if $\sum_{k=0}^{\infty} W_S(k) = \infty$ for any $S \subset [m]$.

We next provide a generalization of the Infinite Flow Theorem established in [7].

Theorem 3: Let $\{W(k)\} \in \mathscr{M}_2$ have the weak feedback property. Then, the following statements are equivalent:

(a) The model is ergodic.
(b) The model has infinite flow property.
(c) The expected model has infinite flow property.
(d) The expected model is ergodic.

The proof of this result follows almost the same line of arguments as that of Theorem 3 in [8]. There one should notice that the proof can rest on the definition of the class $\mathscr{M}_2$ chains as opposed to models with a common steady state π in expectation.

V. INFINITE FLOW GRAPH

So far, we showed that the infinite flow property and ergodicity are equivalent under some conditions for chains $\{W(k)\}$ in the class $\mathscr{M}_2$. We now focus on the situation when the infinite flow property *does not hold* and we aim to answer the question "what happens with the dynamics in (1) when the infinite flow is not present?" We provide some insights into that question. For this, we let $\{W(k)\}$ be an independent random model and define the infinite flow graph associated with the model as follows.

Definition 6: For a given independent random model $\{W(k)\}$, the *infinite flow graph* of the model is the undirected graph $G^\infty = ([m], \mathcal{E}^\infty)$ where $\{i,j\} \in \mathcal{E}^\infty$ if and only if $\sum_{k=0}^{\infty}(W_{ij}(k) + W_{ji}(k)) = \infty$ almost surely. Since the model is independent, the event $\sum_{k=0}^{\infty}(W_{ij}(k) + W_{ji}(k)) = \infty$ is a tale event and hence, it happens with either probability zero or probability one. Thus, $\{i,j\} \notin \mathcal{E}^\infty$ if and only if $\sum_{k=0}^{\infty}(W_{ij}(k) + W_{ji}(k)) < \infty$ almost surely.

Note that the graph G^∞ is connected if and only if the model has infinite flow property. In [7], Theorem 1, we have shown that the infinite flow property is necessary for the ergodicity of any random chain. Here, we will give a stronger version of this result by the use of the following lemma.

Lemma 6: Let $\{A(k)\}$ be a deterministic chain, and let the indices $i, j \in [m]$ be such that $i \Leftrightarrow_A j$. Then, i and j belong to the same connected component of G^∞.

Proof: Let S be the vertex set corresponding to the connected component of G^∞ for which $i \in S$. To arrive at a contradiction, assume that $j \notin S$, so we have $j \in \bar{S}$ and $\sum_{k=0}^{\infty} A_S(k) < \infty$. Without loss of generality, we can assume that $S = \{1, \cdots, \ell\}$ for some $\ell < m$. Then, we construct a matrix $B(k)$ by modifying the entries of $A(k)$ for $i \in S$ and $j \notin S$. In particular, for $i \in S$ we let $B_{ij'} = A_{ij'}$ if $j' \in S$ with $j' \neq i$, $B_{ii} = A_{ii} + \sum_{j' \in \bar{S}} A_{ij'}$, and $B_{ij'} = 0$ for $j' \in \bar{S}$. Similarly, we define $B_{i'j'}$ for $i' \in \bar{S}$: $B_{i'j'} = A_{i'j'}$ if $j' \in \bar{S}$ with $j' \neq i'$, $B_{i'i'} = A_{i'i'} + \sum_{j' \in S} A_{i'j'}$, and $B_{i'j'} = 0$ for $j' \in S$. By this construction, we have $B(k) = \begin{bmatrix} B_1(k) & 0 \\ 0 & B_2(k) \end{bmatrix}$, where $B_1(k)$ and $B_2(k)$ are $\ell \times \ell$ and $(m-\ell) \times (m-\ell)$ stochastic matrices, respectively. Note that because $\sum_{k=0}^{\infty} A_S(k) < \infty$, the chain $\{B(k)\}$ is an ℓ_1-approximation of $\{A(k)\}$.

Now, let $z = \sum_{i=1}^{\ell} e_i$. Then, $B(k)z = z$ for any $k \geq 0$. Since $i \in S$ and $j \in \bar{S}$, we have $z_i = 1$ for all $i \in S$ and $z_j = 0$ for all $j \in \bar{S}$. Hence, $i \not\Leftrightarrow_B j$ and by Approximation Lemma 4 it follows $i \not\Leftrightarrow_A j$ - a contradiction. ∎

If in Lemma 6, the chain $\{A(k)\}$ is ergodic, then $i \Leftrightarrow_A j$ for any $i, j \in [m]$, implying that G^∞ is connected, which means that $\{A(k)\}$ has infinite flow property.

For convenience, in the rest of this paper we use the following notation. Given a random model $\{W(k)\}$, we let its infinite flow graph G^∞ have τ connected components, $\tau \geq 1$. We let $S_1, \ldots, S_\tau \subset [m]$ be the sets of vertices in the connected components of G^∞. Without loss of generality, we assume the sets S_r and their elements are ordered as follows: $S_1 = \{a_0 + 1, \ldots, a_1\}$, $S_2 = \{a_1 + 1, \ldots, a_2\}$, $\ldots, S_\tau = \{a_{\tau-1} + 1, \ldots, a_\tau\}$ for $a_0 = 0 < a_1 < \ldots < a_\tau = m$. Let m_r be the number of vertices in S_r for $1 \leq r \leq \tau$, for which we have $m_r = a_r - a_{r-1}$ and $\sum_{r=1}^{\tau} m_r = m$. Evidently, the sets S_r, $1 \leq r \leq \tau$ form a partition of the set $[m]$. In order to properly relate the elements in the sets $[m_r]$ and $[m]$, we define $i_r = i + a_{r-1}$ for $r = 1, \ldots, \tau$. Thus, $i_r \in [m]$ is the vertex in the set $[m]$ corresponding to $i \in [m_r]$ (the ith vertex of the rth connected component S_r).

Now, given a chain $\{W(k)\}$ and its infinite flow graph G^∞, we construct a special *block diagonal approximation* of $\{W(k)\}$ based on the partition $\{S_r, 1 \leq r \leq \tau\}$ of G^∞.

For any $r \in [\tau]$, we define a matrix $\tilde{W}^{(r)}(k)$ for each k based on the matrix $W(k)$ and the index set S_r, as follows: for $i, j \in [m_r]$ and any $k \geq 0$,

$$\tilde{W}_{ij}^{(r)}(k) = \begin{cases} W_{i_r j_r}(k) + \sum_{\ell \in \bar{S}_r} W_{i_r \ell}(k) & \text{if } j = i, \\ W_{i_r j_r}(k) & \text{if } j \neq i. \end{cases} \tag{10}$$

Thus, for each r, the matrices $\tilde{W}^{(r)}(k)$ are $m_r \times m_r$, and the model $\{\tilde{W}^{(r)}(k)\}$ is defined over $\mathbb{S}^{m_r}$. The *block diagonal approximation* of $\{W(k)\}$ is given by

$$\tilde{W}(k) = \text{diag}(\tilde{W}^{(1)}(k), \ldots, \tilde{W}^{(\tau)}(k))$$
$$= \begin{pmatrix} \tilde{W}^{(1)}(k) & 0 & \cdots & 0 \\ 0 & \tilde{W}^{(2)}(k) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{W}^{(\tau)}(k) \end{pmatrix}. \tag{11}$$

Based on the construction, the chain $\{\tilde{W}(k)\}$ has the same infinite flow graph as the chain $\{W(k)\}$. Furthermore, every sub-chain $\{\tilde{W}^{(r)}(k)\}$ has its infinite flow graph equal to the subgraph of G^∞ induced by the r-th connected component (with vertex set S_r) for $r \in [\tau]$. Therefore, the models $\{\tilde{W}^{(r)}(k)\}$, $r \in [\tau]$ have infinite flow property.

In fact, the block diagonal approximation of a model is an ℓ_1-approximation of the model, as seen in the following.

Lemma 7: Let $\{W(k)\}$ be an independent random model and $\{\tilde{W}(k)\}$ be its block diagonal approximation. Then, $\{\tilde{W}(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$.

Proof: First we show that $\tilde{W}(k)$ is a stochastic matrix for any $k \geq 0$. By the form of the matrix $\tilde{W}(k)$ in Eq. (10), it suffice to show that $\tilde{W}^{(r)}(k)$ is stochastic for $1 \leq r \leq \tau$. By the definition of $\tilde{W}^{(r)}(k)$, we have $\tilde{W}^{(r)}(k) \geq 0$. Also, for any $i \in [m_r]$, we have:

$$\sum_{j=1}^{m_r} \tilde{W}_{ij}^{(r)}(k) = \tilde{W}_{ii}^{(r)}(k) + \sum_{j \neq i, j \in [m_r]} \tilde{W}_{ij}^{(r)}(k)$$
$$= \left(W_{i_r i_r}(k) + \sum_{\ell \in \bar{S}_r} W_{i_r \ell}(k) \right) + \sum_{\ell \neq i_r, \ell \in S_r} W_{i_r \ell}(k)$$
$$= \sum_{\ell=1}^{m} W_{i_r \ell}(k) = 1. \tag{12}$$

Therefore, the matrices $\tilde{W}^{(r)}(k)$ are stochastic.

To show that $\{\tilde{W}(k)\}$ is an ℓ_1-approximation, we estimate the terms $|\tilde{W}_{ij}(k) - W_{ij}(k)|$ for $i, j \in [m]$. For this, we let $i \in S_r$ for $r \in [\tau]$. Then, we have three cases:
(1) If $j \notin S_r$, we have $\tilde{W}_{ij}(k) = 0$, implying that $|\tilde{W}_{ij}(k) - W_{ij}(k)| = W_{ij}(k)$.
(2) If $j \in S_r$ and $j \neq i$, by the definition of $\tilde{W}(k)$ we have $\tilde{W}_{ij}(k) = W_{ij}(k)$ and hence, $|\tilde{W}_{ij}(k) - W_{ij}(k)| = 0$.
(3) If $j = i$, we have

$$\tilde{W}_{ii}(k) = W_{ii}(k) + \sum_{j \notin S_r} W_{ij}(k).$$

Hence, $|\tilde{W}_{ij}(k) - W_{ij}(k)| = \sum_{j \notin S_r} W_{ij}(k)$.

From the preceding three cases, we have for any $r \in [\tau]$,

$$\sum_{j=1}^{m} |\tilde{W}_{ij}(k) - W_{ij}(k)| = 2 \sum_{j \notin S_r} W_{ij}(k) \quad \text{for all } i \in S_r.$$

By summing over all $i \in S_r$, we have for any $r \in [\tau]$,

$$\sum_{i \in S_r} \sum_{j=1}^{m} |\tilde{W}_{ij}(k) - W_{ij}(k)| = 2 \sum_{i \in S_r} \sum_{j \notin S_r} W_{ij}(k) \leq 2 W_{S_r}(k).$$

By summing these inequalities for $r = 1, \ldots, \tau$, we obtain

$$\sum_{i,j \in [m]} |\tilde{W}_{ij}(k) - W_{ij}(k)| = \sum_{r=1}^{\tau} \sum_{i \in S_r} \sum_{j=1}^{m} |\tilde{W}_{ij}(k) - W_{ij}(k)|$$
$$\leq 2 \sum_{r=1}^{\tau} W_{S_r}(k).$$

The sets $S_1, \ldots, S_\tau$ are the sets of vertices of the connected components in the infinite flow graph G^∞. By the definition of the infinite flow graph, we have $\sum_{k=0}^{\infty} \sum_{r=1}^{\tau} W_{S_r}(k) < \infty$ almost surely. Therefore, it follows that

$$\sum_{k=0}^{\infty} \sum_{i,j \in [m]} |\tilde{W}_{ij}(k) - W_{ij}(k) - | < \infty \quad a.s.,$$

proving that $\{\tilde{W}(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$. $\blacksquare$

By Theorem 3, we know that for $\mathcal{M}_2$ models with asymptotic distribution $\pi > 0$ and weak feedback property, the model is ergodic if and only if the infinite flow graph is connected. The following theorem characterizes the behavior of system (1) when the dynamics is driven by an independent chain $\{W(k)\}$ with a common steady state $\pi > 0$ in expectation and weak feedback property.

Theorem 4: Let $\{W(k)\}$ be an independent model with a common steady state $\pi > 0$ in expectation and weak feedback property. Then, the model is partially ergodic and $i \Leftrightarrow j$ if and only if i and j are in the same connected component of G^∞.

Proof: The "if" part follows from Lemma 6. To show the "only if" part, we use two ℓ_1-approximations successively. The first approximation is the block diagonal approximation, which gives us almost all of the properties we are looking for. However, it does not necessarily preserve weak-feedback property, so we construct another ℓ_1-approximation to cure this problem.

Let the infinite flow graph of $\{W(k)\}$ be G^∞ and suppose it has τ connected components. Let $S_1, \ldots, S_\tau$ be the vertex sets corresponding to the connected components of G^∞. Let $\pi_{\min} = \min_{i \in [m]} \pi_i > 0$.

Consider the diagonal approximation $\{\tilde{W}(k)\}$ of $\{W(k)\}$ with $\tilde{W}^{(r)}(k)$ be defined as in Eq. (10) for $r \in [\tau]$. Let $M(k) = \mathsf{E}\left[\max_{i,j \in [m]} |\tilde{W}_{ij}(k) - W_{ij}(k)|\right]$. Since $\{\tilde{W}(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$, by Lemma 1 $\sum_{k=0}^{\infty} M(k) < \infty$. Therefore $\lim_{k \to \infty} M(k) = 0$. Thus, there exists $N \geq 0$ such that $M(k) \leq \frac{\pi_{\min}}{8m}$ for any $k \geq N$.

Let $J^{(r)} = \pi^{(r)} e^{(m_r)T}$ where $\pi^{(r)} = (\pi_{a_{r-1}+1}, \ldots, \pi_{a_r})$, which is a sub-vector of π having the coordinates π_i with

$i \in S_r$. Let $U(k) = I$ for $k < N$ and for $k \geq N$,

$$U^{(r)}(k) = (1 - d(k))\tilde{W}^{(r)}(k) + d(k)J^{(r)},$$

where $d(k) = \frac{4m}{\pi_{\min}}M(k)$ for $k \geq 0$. Since $M(k) \leq \frac{\pi_{\min}}{8m}$ for $k \geq N$, we have $d(k) \in [0, \frac{1}{2}]$. Since convex combination of stochastic matrices is a stochastic matrix, it follows that each $U^{(r)}(k)$ is stochastic. Note that $\sum_{k=0}^{\infty} M(k) < \infty$ implies $\sum_{k=0}^{\infty} d(k) < \infty$ and hence, the model $\{U^{(r)}(k)\}_{k \geq N}$ is an ℓ_1-approximation of $\{\tilde{W}^{(r)}(k)\}_{k \geq N}$. But since, the entries of each matrix are in $[0, 1]$, changing finitely many matrices in a chain cannot change infinite flow properties. Therefore, $\{U^{(r)}(k)\}$ is an ℓ_1-approximation of $\{\tilde{W}^{(r)}(k)\}$ and the model $\{U(k)\}$ with matrices defined by $U(k) = \text{diag}(U^{(1)}(k), \ldots, U^{(\tau)}(k))$, $k \geq 0$ is an ℓ_1-approximation of $\{\tilde{W}(k)\}$. By Lemma 7, $\{\tilde{W}(k)\}$ is an ℓ_1-approximation of the original model $\{W(k)\}$ and therefore, $\{U(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$.

For $k < N$, $U(k) = I$ which has weak-feedback property with constant 1. So, let us fix $k \geq N$ and show weak feedback property for $U(k)$. Let $r \in [\tau]$ be arbitrary and for simplicity of notation, let $Q = U^{(r)}(k)$. For $i, j \in [m_r]$ recall that their corresponding indices in $[m]$ are given by $i_r = i + a_r - 1$, $j_r = j + a_r - 1$. Also, recall that W^s denotes the sth column vector of a matrix W. Using this, for any $i, j \in [m_r]$ with $i \neq j$ we have:

$$Q^{iT}Q^j = \left((1 - d(k))\tilde{W}^{(r)i_r}(k) + d(k)\pi^{(r)}\right)^T$$
$$\left((1 - d(k))\tilde{W}^{(r)j_r}(k) + d(k)\pi^{(r)}\right)$$
$$\geq (1 - d(k))^2(\tilde{W}^{(r)i_r}(k))^T\tilde{W}^{(r)j_r}(k)$$
$$+ (1 - d(k))d(k)\pi^{(r)T}(\tilde{W}^{(r)i_r}(k) + \tilde{W}^{(r)j_r}(k)). \quad (13)$$

Now, based on the definition of $\pi^{(r)}$, we have:

$$\pi^{(r)T}\tilde{W}^{(r)i_r}(k) = \pi^T W^{i_r}(k) + \pi^T(\tilde{W}^{i_r}(k) - W^{i_r}(k))$$
$$\geq \pi^T W^i(k) - \max_{i'j'}|\tilde{W}_{i'j'}(k) - W_{i'j'}(k)|,$$

which holds due to the stochasticity of π. Therefore,

$$\mathsf{E}\left[\pi^{(r)T}\tilde{W}^{(r)i_r}(k)\right] \geq \pi_i - M(k),$$

which follows from π being a common steady state in expectation of $\{W(k)\}$. Similarly, we have $\mathsf{E}\left[\pi^{(r)T}\tilde{W}^{(r)j_r}(k)\right] \geq \pi_j - M(k)$. Taking the expectation of the both sides in Eq. (13) and using the preceding inequalities, we obtain

$$\mathsf{E}[Q^{iT}Q^j] \geq (1 - d(k))^2\mathsf{E}\left[(\tilde{W}^{i_r}(k))^T\tilde{W}^{j_r}(k)\right]$$
$$+ (1 - d(k))d(k)(\pi_i + \pi_j - 2M(k))$$
$$\geq (1 - d(k))^2\mathsf{E}\left[(\tilde{W}^{i_r}(k))^T\tilde{W}^{j_r}(k)\right]$$
$$+ (1 - d(k))d(k)\frac{\pi_i + \pi_j}{2}, \quad (14)$$

which holds by $M(k) \leq \frac{\pi_{\min}}{8m} \leq \frac{\pi_i}{4}$ for $i \in [m]$ and $k \geq N$.

Since, $\tilde{W}_{\ell i_r}(k) \geq W_{\ell i_r}(k) - \max_{ij}|\tilde{W}_{ij}(k) - W_{ij}(k)|$ for all $\ell \in [m]$, we have

$$\mathsf{E}\left[(\tilde{W}^{i_r}(k))^T\tilde{W}^{j_r}(k)\right] \geq \mathsf{E}\left[(W^{i_r}(k))^T W^{j_r}(k)\right] - 2mM(k).$$

Therefore, using the above inequality in Eq. (14), we have:

$$\mathsf{E}[Q^{iT}Q^j] \geq (1 - d(k))^2\mathsf{E}\left[(W^{i_r}(k))^T W^{j_r}(k)\right]$$
$$- (1 - d(k))^2 2mM(k) + (1 - d(k))d(k)\frac{\pi_i + \pi_j}{2}.$$

But $0 \leq \frac{4m}{\pi_{\min}}M(k) = d(k) \leq 1$ and hence,

$$2mM(k) \leq d(x)\frac{\pi_i + \pi_j}{4}.$$

By combining the preceding two relations, we have

$$\mathsf{E}[Q^{iT}Q^j] \geq (1 - d(k))^2\mathsf{E}\left[(W^{i_r}(k))^T W^{j_r}(k)\right]$$
$$+ (1 - d(k))d(k)\frac{\pi_i + \pi_j}{4}$$
$$\geq (1 - d(k))^2\gamma\mathsf{E}[W_{i_r j_r}(k) + W_{j_r i_r}(k)]$$
$$+ (1 - d(k))d(k)\frac{\pi_i + \pi_j}{4}, \quad (15)$$

where the last inequality follows by weak feedback property of $\{W(k)\}$. Without loss of generality, we may assume $\gamma \leq \frac{1}{4}$ (otherwise we replace γ by $\frac{1}{4}$). Since $i_r, j_r \in S_r$ and $i_r \neq j_r$, by the construction of $\tilde{W}(k)$, we have $\tilde{W}_{i_r j_r}(k) = W_{i_r j_r}(k)$. Hence, $\mathsf{E}[Q_{ij} + Q_{ji}] = (1 - d(k))(\mathsf{E}[W_{i_r j_r}(k)] + \mathsf{E}[W_{j_r i_r}(k)]) + d(k)(\pi_i + \pi_j)$. By combining this with Eq. (15), we have

$$\mathsf{E}[Q^{iT}Q^j] \geq (1 - d(k))\gamma(\mathsf{E}[Q_{ij} + Q_{ji}] - d(k)(\pi_i + \pi_j))$$
$$+ (1 - d(k))d(k)\frac{\pi_i + \pi_j}{4}$$
$$= (1 - d(k))\gamma\mathsf{E}[Q_{ij} + Q_{ji}]$$
$$+ (1 - d(k))d(k)\left(-\gamma + \frac{1}{4}\right)(\pi_i + \pi_j)$$
$$\geq \frac{\gamma}{2}\mathsf{E}[Q_{ij} + Q_{ji}],$$

which follows from $d(k) \leq \frac{1}{2}$ and $\gamma \leq \frac{1}{4}$. Note that we set $Q = U^{(r)}(k)$ where $r \in [\tau]$ and $k \geq N$ was arbitrary. Hence, each of the decoupled random models $\{U^{(r)}(k)\}$ has weak-feedback property with feedback constant $\frac{\gamma}{2}$.

Let $x(0) \in \mathbb{R}^m$ and $\{x(k)\}$ be the chain resulted from $\{U(k)\}$ and dynamic system (1). By Theorem 2, it follows that $\{U(k)\} \in \mathcal{M}_2$ and $\pi > 0$ is an asymptotic distribution for $\{W(k)\}$. Hence, $\sum_{k=0}^{\infty}\sum_{i<j}L_{ij}(k)(x_i(k) - x_j(k))^2 < \infty$ almost surely, where $L(k) = \mathsf{E}[U^T(k)\text{diag}(\pi)U(k)]$. Hence, for any $r \in [\tau]$,

$$\sum_{k=0}^{\infty}\sum_{\substack{i<j \\ i,j \in S_r}}L_{ij}(k)(x_i(k) - x_j(k))^2 < \infty \quad a.s.$$

Due to the diagonal structure of $U(k)$, we have:
(a) $x(k) = (x^{(1)}(k), \cdots, x^{(\tau)}(k))$, where $x_i^{(r)}(0) = x_{i_r}(0)$, so $\{x^{(r)}(k)\}$ are the sequences of random vectors in $\mathbb{R}^{m_r}$ driven by the individual chains $\{U^{(r)}(k)\}$.

(b) For $i, j \in [m_r]$ and $r \in [\tau]$,

$$
\begin{aligned}
L_{i_r j_r}(k) &= \mathsf{E}\left[\sum_{\ell \in [m]} \pi_\ell U_{\ell i_r}(k) U_{\ell j_r}(k) \right] \\
&= \mathsf{E}\left[\sum_{\ell \in S_r} \pi_\ell U^{(r)}_{\ell i_r}(k) U^{(r)}_{\ell j_r}(k) \right] \\
&= \mathsf{E}\left[\sum_{\bar\ell \in [m_r]} \pi^{(r)}_{\bar\ell} U^{(r)}_{\bar\ell i}(k) U^{(r)}_{\bar\ell j}(k) \right]. \qquad (16)
\end{aligned}
$$

Therefore, by the above observations, the random model $\{U(k)\} \in \mathscr{M}_2$ acting on $\mathbb{R}^m$, decomposes into τ random models $\{U^{(1)}(k)\}, \ldots, \{U^{(\tau)}(k)\}$ each of $\mathscr{M}_2$ class and acting on $\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_\tau}$, respectively. Furthermore, the random model $\{\tilde W^{(r)}(k)\}$ has an asymptotic distribution $\pi^{(r)} > 0$. Also each model $\{\tilde W^{(r)}(k)\}$ has the infinite flow property and hence, their ℓ_1-approximations have the infinite flow property. Also, we showed that each random model $\{\tilde W^{(r)}(k)\}$ has weak feedback property. Hence, by Theorem 3, $\{U^{(r)}(k)\}$ is ergodic chain for any $r \in [\tau]$ which implies $i \Leftrightarrow j$ for any $i, j \in S_r$. Since, $U(k) = \mathrm{diag}(U^{(1)}(k), \ldots, U^{(\tau)}(k))$, hence, $i \Leftrightarrow j$ in $\{U(k)\}$. Therefore, by Approximation Lemma 4, $i \Leftrightarrow j$ in the original chain $\{W(k)\}$ if and only if $i \Leftrightarrow j$ in $\{U(k)\}$ which is true if and only if $i, j \in S_r$ for some $r \in [\tau]$. ∎

Using the above theorem, the following generalization of the Infinite Flow theorem [7] follows immediately.

Theorem 5: Let $\{W(k)\}$ be an independent chain with a common steady state $\pi > 0$ in expectation and weak feedback property. Then, the chain and the expected chain are partially ergodic. Furthermore, the following conditions are equivalent:

(a) $i \Leftrightarrow j$ in $\{W(k)\}$.
(b) i, j belong to the same connected component of G^∞.
(c) i, j belong to the same connected component of $\bar G^\infty$, where $\bar G^\infty$ is the infinite flow graph of $\{\mathsf{E}[W(k)]\}$.
(d) $i \Leftrightarrow j$ in $\{\mathsf{E}[W(k)]\}$.

Proof: By Theorem 4 $\{W(k)\}$ is partially ergodic chain and by Lemma 2, it follows that the expected chain $\{\mathsf{E}[W(k)]\}$ is partially ergodic. By Lemma 2, (a) implies (d). Next, (d) implies (c) by Lemma 6. Since, $W_{ji}(k) + W_{ij}(k) \in [0, 2]$, by use of Lemma 3, the infinite flow graph of $\{W(k)\}$ and $\{\mathsf{E}[W(k)]\}$ are the same and hence, (c) and (b) are equivalent. Finally, since $\{W(k)\}$ satisfies the conditions of Theorem 4, (b) implies (a). ∎

Under the conditions of Theorem 4, as a special case, the dynamics driven by deterministic chains will always have a limit point. As an example we have the following corollary.

Corollary 3: Let $\{A(k)\}$ be a deterministic chain of doubly stochastic matrices. Let $\gamma > 0$ be such that $A_{ii}(k) \geq \gamma > 0$ for any $k \geq 0$ and $i \in [m]$. Then, $\Phi(\infty, t_0) = \lim_{k\to\infty} \Phi(k, t_0)$ exists for any t_0. Furthermore, if two indices i and j are in the same connected component of the infinite flow graph of $\{A(k)\}$, then the ith row and jth row of $\Phi(\infty, t_0)$ are the same.

As another example, consider an application of Theorem 4 to an i.i.d. random model $\{W(k)\}$ with positive diagonal entries (almost surely). In [6], it was shown that such a chain is ergodic if and only if its expected chain is ergodic. In [7], it was shown that such a chain has weak-feedback property. For this model the expected matrix $\mathsf{E}[W(k)]$ is time invariant (and stochastic). Hence, the model has a common steady state π in expectation, but π may have zero entries.

Corollary 4: Let $\{W(k)\}$ be an i.i.d. model with $W_{ii}(0) > 0$ almost surely for all $i \in [m]$. Let $\pi > 0$ be a vector such that $\pi^T \mathsf{E}[W(k)] = \pi^T$. Then, the limit $\lim_{k\to\infty} \Phi(k, 0)$ exists almost surely.

We next provide an example showing that the condition $\pi > 0$ is indeed needed in Corollary 4. Consider

$$
W(k) = \begin{pmatrix} 1 & 0 & 0 \\ \alpha(k) & \frac{1}{2} & \beta(k) \\ 0 & 0 & 1 \end{pmatrix},
$$

where the vector $(\alpha(k), \beta(k))$ has two random realizations, $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$, each occurring with probability $\frac{1}{2}$ independently and identically for $k \geq 0$. If the dynamics is started at $x(0) = (0, \frac{1}{2}, 1)$, then for any $k \geq 1$, $x_1(k) = x_1(0) = 0$, $x_3(k) = x_3(0) = 1$ and $x_2(k)$ would equally likely be either $\frac{x_2(k-1)+1}{2}$ or $\frac{x_2(k-1)}{2}$, which shows that $x(k)$ will never converge.

VI. CONCLUSION

In this paper, we investigated the limiting behavior of the random linear dynamics driven by independent random stochastic matrices. We proved that for a class of independent random models, such dynamics always converge. To do so, we introduced notion of partial ergodicity and mutual ergodicity as well as ℓ_1-approximation of a given chain. We proved that ℓ_1-approximation of a given chain preserves ergodic behaviors of the chains. Finally, we proved that such dynamics converge and, also, we gave a characterization of the limiting points through the use of the infinite flow graph that we associated with the model.

REFERENCES

[1] S. Chatterjee and E. Seneta, *Towards consensus: Some convergence theorems on repeated averaging*, Journal of Applied Probability **14** (1977), no. 1, 89–97.
[2] R. Durrett, *Probability: Theory and examples*, third ed., Curt Hinrichs, 2005.
[3] A. Jadbabaie, J. Lin, and S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Transactions on Automatic Control **48** (2003), no. 6, 988–1001.
[4] B.T. Polyak, *Introduction to optimisation*, Optimization Software, Inc., New York, 1987.
[5] E. Seneta, *Non-negative Matrices and Markov Chains*, 3 ed., Springer Series in Statistics, 2006.
[6] A. Tahbaz-Salehi and A. Jadbabaie, *A necessary and sufficient condition for consensus over random networks*, IEEE Transactions on Automatic Control **53** (2008), no. 3, 791–795.
[7] B. Touri and A. Nedić, *On ergodicity, infinite flow and consensus in random models*, available at: http://arxiv.org/abs/1001.1890.
[8] _____, *When infinite flow is sufficient for ergodicity*, Submitted to the 49th IEEE Conference on Decision and Control, available at: https://netfiles.uiuc.edu/touri1/www/Papers/ergodicityCDC.pdf.
[9] J.N. Tsitsiklis, *Problems in decentralized decision making and computation*, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, MIT, 1984.