On Approximations and Ergodicity Classes in Random Chains

Behrouz Touri and Angelia Nedić*

September 7, 2010

Abstract

We study the limiting behavior of a random dynamic system driven by a stochastic chain. Our main interest is in the chains that are not necessarily ergodic but rather decomposable into ergodic classes. To investigate the conditions under which the ergodic classes of a model can be identified, we introduce and study an ℓ_1-approximation and infinite flow graph of the model. We show that the ℓ_1-approximations of random chains preserve certain limiting behavior. Using the ℓ_1-approximations, we show how the connectivity of the infinite flow graph is related to the structure of the ergodic groups of the model. Our main result of this paper provides conditions under which the ergodicity groups of the model can be identified by considering the connected components in the infinite flow graph. We provide two applications of our main result to random networks, namely broadcast over time-varying networks and networks with random link failure.

1 Introduction

The dynamic systems driven by stochastic matrices have found its use in many problems in decentralized communication [5, 8, 24, 3], decentralized control [15, 22, 26], distributed optimization [33, 34, 25, 20, 16], and information diffusion in social networks [14, 1]. In many of these applications, the ergodicity plays a central role in ensuring that the local "agent" information diffuses eventually over the entire network of agents. The conditions under which the ergodicity happens has been subject of some recent studies [29, 30]. However, there are no studies that investigate the limiting behavior of the dynamics driven by time-varying chains when the ergodicity does not happen. Such studies are important for understanding the group formation in both deterministic and random time-varying networks, such as multiple-leaders/multiple-followers networked systems.

The main objective of this paper is to investigate the limiting behavior of the linear dynamics driven by random independent chains of stochastic matrices in the absence

*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 grant CMMI 07-42538.

of ergodicity. Our goal is to study the conditions under which the ergodic groups are formed and to characterize these groups. To do so, we introduce an ℓ_1-approximation and the infinite flow graph of a random model, and we study the properties of these objects. Using the established properties, we extend the main result of the previous work in [31] to a broader class of independent random models. We then proceed to show that for certain random models, although the ergodicity might not happen, the dynamics of the model still converges almost surely and partial ergodicity happens almost surely. In other words, under certain conditions, *ergodic groups* are formed and we characterize these groups through the connected components of the infinite flow graph. We then apply the results to a broadcast-gossip algorithm over time-changing networks and to networks with link failures.

The work in this paper is related to the literature on ergodicity of random models. A discussion on the ergodicity of deterministic (forward and backward) chains can be found in [28]. The earliest occurrence of the study of random models dates back to the work of Rosenblatt [27], where the algebraic and topological structure of the set of stochastic matrices is employed to investigate the limiting behavior of the product of independent identically distributed (i.i.d.) random matrices. Later, in [17, 18, 10], such a product is studied extensively under a more general assumption of stationarity, and a necessary and sufficient condition for the ergodicity is developed. In [7, 29] the class of i.i.d. random models with almost sure positive diagonal entries were studied. In particular, in [29] is has been showed that such a random model is ergodic if and only if its expected chain is ergodic. Later, this result has been extended to the stationary ergodic models in [30].

Unlike the work on i.i.d. models or stationary processes studied in [27, 17, 18, 12, 29, 30], the focus of the work in this paper is on independent random models. The work is a continuation of our work in [31], where we showed that for a class of independent random models the ergodicity is is equivalent to the connectivity of the *infinite flow graph*, a graph which can be derived from the random model or its expected model. Furthermore, unlike the studies that provide conditions for ergodicity of deterministic or random chains, such as [9, 33, 15, 6, 12, 29, 30, 31], the work presented in this paper considers the limiting behavior of deterministic and random models that are not necessarily ergodic.

The main contribution of this work lies in the following:

(1) The establishment of *conditions on random models under which the ergodicity classes are fully characterized*. This result not only implies that certain random models almost always converge, but also provides the structure of the limiting matrix. The structure is revealed through the connectivity topology of the infinite flow graph of the model. Although the model is not ergodic, the ergodicity happens *locally for groups of indices*, which are characterized as the vertices in the same connected component of the infinite flow graph (Theorem 6).

(2) The introduction and study of a class of perturbations (ℓ_1-approximations) of chains that preserve ergodicity classes, as seen is Section 3 (Lemma 2).

(3) The introduction of a class $\mathscr{M}_2$ random models whose ergodicity can be characterized through the infinite flow property (Theorem 4). This class encircles many of the known ergodic deterministic and random models, as discussed in Section 4.

The structure of this paper is as follows: in Section 2, we discuss the problem of our interest, introduce some terminology and motivate our work by considering the limiting behavior of the gossip protocol extended to time-varying graphs. Motivated by this example, we establish some notions for the limiting behavior of the model. In Section 3, we define and investigate ℓ_1-approximations of random models which play an important role in the later development. In Section 4, we extend the infinite flow theorem [31] to a larger class of random models. In Section 5, we introduce and study an infinite flow graph of a model. Using this graph, together with the results of Sections 3 and 4, we establish our main result that characterizes the ergodic classes of certain random models. In Section 6, we demonstrate the use of our results on two different random network models.

Notation and Basic Terminology. We view all vectors as columns. For a vector x, we write x_i to denote its ith entry, and we write $x \geq 0$ ($x > 0$) to denote that all its entries are nonnegative (positive). We use x^T to denote the transpose of a vector x. For a vector $x \in \mathbb{R}^m$, we use $\|x\|_p = (\sum_{i=1}^m |x_i|^p)^{1/p}$ for $p \geq 1$ and $\|x\|$ when $p = 2$. For a matrix A, we write $\|A\|_p$ to denote the matrix norm induced by $\|\cdot\|_p$ vector norm. We use e_i to denote the vector with the ith entry equal to 1 and all other entries equal to 0. We write e to denote the vector with all entries equal to 1. We write $\{x(k)\}$ to denote a sequence $x(0), x(1), \ldots$ of some elements, and we write $\{x(k)\}_{k\geq t}$ to denote the truncated sequence $x(t), x(t+1), \ldots$ for $t > 0$. We say that a sequence of scalars, vectors, or matrices is an ℓ_1-sequence if the sequence generated by each entry of the corresponding object is absolutely summable.

For a given set C and a subset S of C, we write $S \subset C$ when S is a proper subset of C. A set $S \subset C$ with $S \neq \emptyset$ is a *nontrivial* subset of C. We use $[m]$ for the integer set $\{1, \ldots, m\}$. We let $\bar{S}$ denote the complement of a given set $S \subseteq [m]$ with respect to $[m]$.

We denote the identity matrix by I. For a finite collection $A_1, \ldots, A_\tau$ of square matrices, we write $A = \text{diag}(A_1, \ldots, A_\tau)$ to denote the block diagonal matrix with rth diagonal block being A_r for $1 \leq r \leq \tau$. For a matrix W, we write W_{ij} to denote its (i,j)th entry, W^i to denote its ith column vector, and W^T to denote its transpose. For an $m \times m$ matrix W, we let $\sum_{i<j} W_{ij} = \sum_{i=1}^{m-1} \sum_{j=i+1}^m W_{ij}$. For such a matrix and a nontrivial subset $S \subset [m]$, we define $W_S \triangleq \sum_{i\in S, j\in\bar{S}} (W_{ij} + W_{ji})$. A vector $v \in \mathbb{R}^m$ is stochastic if $v \geq 0$ and $\sum_{i=1}^m v_i = 1$. A matrix W is *stochastic* when all its rows are stochastic, and it is *doubly stochastic* when both W and W^T are stochastic. We let $\mathbb{S}^m$ denote the set of $m \times m$ stochastic matrices. We refer to a sequence $\{W(k)\}$ of matrices as *model* or *chain* interchangeably.

We write $\mathsf{E}[X]$ to denote the expected value of a random variable X. For an event $\mathscr{A}$, we use $\mathsf{Pr}(\mathscr{A})$ to denote its probability. If $\mathsf{Pr}(\mathscr{A}) = 1$, we say that the event $\mathscr{A}$ happens almost surely. We often abbreviate "almost surely" by *a.s.*

2 Problem Formulation and Motivation

In this section, we formulate the problem of our interest, introduce some notions and provide an example that motivates the further development.

2.1 Problem Description

We consider the dynamics of a linear system driven by a random stochastic chain, i.e.,

$$x(k+1) = W(k)x(k) \qquad \text{for } k \geq t_0. \tag{1}$$

where $W(k) \in \mathbb{S}^m$ for all k, the time t_0 is an initial time and $x(t_0) \in \mathbb{R}^m$ is an initial state of the system. It is well known that when the matrix sequence $\{W(k)\}$ is ergodic, the dynamics in (1) is convergent almost surely for any initial time t_0 and any initial state $x(t_0)$. Furthermore, the limiting value of each coordinate $x_i(k)$ is the same, which is often referred to as consensus, agreement, or synchronization. In this case, the sequence $\{W(k)\}$ has a single ergodic class, the class $[m]$ consisting of all coordinate indices $\{1, \ldots, m\}$.

A natural question arises: what happens if $\{W(k)\}$ is not ergodic? In particular, what can we say about the limiting dynamics of the coordinates $x_i(t)$? Can we determine the ergodicity classes based on the properties of the matrices $W(k)$? Our motivation in this paper is to answer these questions. To do this, we formalize the probabilistic model for the matrices $\{W(k)\}$ and introduce some terminology.

Let $(\Omega, \mathscr{F}, \mathsf{Pr}(\cdot))$ be a probability space and let $W : \Omega \to \Pi_{i=0}^{\infty}\mathbb{S}^m$ be a random matrix process such that $W_{ij}(k)$ is a Borel-measurable function for all $i, j \in [m]$ and $k \geq 0$. We call such a process a *random chain* or a *random model*. We often denote such a process by its coordinate process representation $\{W(k)\}$. If matrices $W(k)$ are independent, we say that the model is independent and, in addition, if $W(k)$s are identically distributed we say that $\{W(k)\}$ an *independent identically distributed (i.i.d.)* random model.

Given a random model $\{W(k)\}$, 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 random dynamic system (1) and, in particular, to characterize the ergodicity classes of the model $\{W(k)\}$. In the development, we consider some generalizations of the notions of ergodicity.

2.2 Terminology

A deterministic chain $\{A(k)\}$ can be viewed as a special independent random chain by setting $\Omega = \{\omega\}$, $\mathscr{F} = \{\{\omega\}, \emptyset\}$, $\mathsf{Pr}(\{\omega\}) = 1$ and $W(k)(\omega) = A(k)$. Then, the dynamic system in Eq. (1) is deterministic and we have the following definition.

Definition 1. *A chain $\{A(k)\}$ is ergodic chain or ergodic model if* $\lim_{k\to\infty}(x_i(k) - x_j(k)) = 0$ *for any $i, j \in [m]$, any starting time $t_0 \geq 0$ and any starting point $x(t_0) \in \mathbb{R}^m$. We say that the chain admits consensus if the above assertion is true for $t_0 = 0$. A random model $\{W(k)\}$ is ergodic (admits consensus) if the event $\mathcal{E}$ ($\mathcal{C}$) happens almost surely.*

To justify Definition 1 for a random model, note that we can speak about subsets $\mathcal{E}$ and $\mathcal{C}$ of Ω on which the ergodicity and consensus happen, respectively. These sets are given by

$$\mathcal{E} = \cap_{t_0=0}^{\infty} \left(\cap_{\ell=1}^{m} \{\omega \mid \lim_{k\to\infty}(x_i(k) - x_j(k)) = 0 \;\text{ for all } i, j \in [m],\, x(t_0) = e_\ell\} \right),$$

$$\mathcal{C} = \cap_{\ell=1}^{m}\{\omega \mid \lim_{k\to\infty}(x_i(k) - x_j(k)) = 0 \ \text{ for all } i,j \in [m], x(0) = e_\ell\}.$$

The scalars $x_i(k)-x_j(k)$ are random variables since $W(k)$ are Borel-measurable. Thus, $\mathcal{E}$ and $\mathcal{C}$ are events (for the discussion on why it suffice to consider only $x(0) = e_\ell$ see [31]).

The ergodicity of a random chain $\{W(k)\}$ is closely related to the limiting behavior of the product of the matrices $W(0), W(1), \ldots$. To see this, we let $\Phi(t,s) = W(t-1)\cdots W(s)$ for $t > s$. Now, if in Definition 1, we set $x(t_0) = e_\ell$ for $\ell \in [m]$, then we conclude that for an ergodic chain $\{W(k)\}$, the ith column of $\Phi(t,t_0)$, i.e. $\Phi^i(t,t_0)$, converges almost surely to a vector that is co-linear with the vector e. Therefore, the ergodicity of a chain implies the almost sure convergence of all the columns of $\Phi(t,t_0)$ to a vector co-linear with e for any starting time t_0, or in other words $\lim_{t\to\infty}\Phi(t,t_0) = e\phi(t_0)^T$ almost surely for a random vector $\phi(t_0) \in \mathbb{R}^m$. Note that in this case $\phi(t_0)$ is a stochastic vector almost surely by the stochasticity of the matrices $W(k)$. Due to the finite dimensionality of the space, the converse of the above statement is also true. Therefore, a chain is ergodic if and only if $\lim_{t\to\infty}\Phi(t,t_0) = e\phi(t_0)^T$ almost surely for a random stochastic vector $\phi(t_0)$ for all $t_0 \geq 0$.

In [31, 32], we have shown that the ergodicity of certain models is closely related to the *infinite flow property*, as defined below.

Definition 2. *(Infinite Flow Property) A deterministic chain $\{A(k)\}$ has the infinite flow property if $\sum_{k=0}^{\infty} A_S(k) = \infty$ for any nonempty $S \subset [m]$. A random model $\{W(k)\}$ has the infinite flow property if the model has infinite flow property almost surely.*

As in the case of consensus and ergodicity events, the subset of Ω over which the infinite flow happens is an event since $W_{ij}(k)$s are Borel-measurable. We denote this event by $\mathscr{F}$.

In our further development, we also use some additional properties of random models such as a weak feedback property and a common steady state in expectation, as introduced in [31]. For convenience, we provide them in the following definition.

Definition 3. *Let $\{W(k)\}$ be a random model. We say that the model has:*

(a) Weak feedback property if it there exists $\gamma > 0$ such that
$$\mathsf{E}\left[W^i(k)^T W^j(k)\right] \geq \gamma(\mathsf{E}[W_{ij}(k)] + \mathsf{E}[W_{ji}(k)]) \qquad \text{for all } k \geq 0 \text{ and } i \neq j,\ i,j \in [m].$$

(b) A common steady state π in expectation if $\pi^T\mathsf{E}[W(k)] = \pi^T$ for all $k \geq 0$.

Apparently, any random model with $W_{ii}(k) \geq \gamma > 0$ almost surely for all $k \geq 0$ and $i \in [m]$ has weak feedback property. Also i.i.d. models with almost sure positive diagonal entries have weak feedback property, as shown in [31]. As another example, consider the homogeneous deterministic chain $\{A(k)\}$ defined by $A(k) = \frac{1}{m-1}(ee^T - I)$ for all $k \geq 0$. It can be seen that this model has weak feedback property with $\gamma = \frac{m-2}{2(m-1)}$.

As an example of a model with a common steady state π in expectation, one can consider any i.i.d. random model $\{W(k)\}$. Another example are the models where $W(k)$ is doubly stochastic almost surely for any k, for which we have $\pi = \frac{1}{m}e$.

With a random model, we associate an *infinite flow graph*. We define the infinite flow graph to be an undirected simple graph, i.e., a graph with no self-loops and multiple edges, and with links that have sufficient information flow.

Definition 4. *(Infinite Flow Graph) For a random model $\{W(k)\}$, the infinite flow graph of the model is the simple 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.*

The infinite flow graph has been (silently) used in [33] mainly to establish the ergodicity of a certain deterministic chains. Here, however, we make use of this graph to establish ergodicity classes for a class of independent random chains. In particular, as we will see in the later sections, the infinite flow graph and its connected components play important role in identifying the ergodicity classes of the model. In the sequel, *a connected component of a graph will always be maximal* with respect to the set inclusion, i.e., it will be the largest connected component that is not properly contained in any other connected component.

2.3 Infinite Flow Graph and Gossip Algorithm on Time-Varying Network

To illustrate the use of the infinite flow graph in determining the ergodicity classes of a random model, we consider a gossip algorithm over a time-varying network, as discussed in [31, 32] to extend the original gossip algorithm [4, 5]. We revisit this algorithm briefly and provide some background results. We then investigate the algorithm for the case when the convergence does not occur to a common vector, but rather a group of agents' values converge to a common vector that may be different for different groups of agents.

We next provide the infinite flow theorem from [31] (see also [32]), which we use in the further development.

Theorem 1. *(Infinite Flow Theorem) Let the random model $\{W(k)\}$ be independent, and have a common steady state $\pi > 0$ in expectation and weak feedback property. Then, the following conditions 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.

Note that Theorem 1 applies to a deterministic model $\{A(k)\}$ that has a common steady state π (i.e., $\pi^T A(k) = \pi^T$) and has the weak feedback property. In view of this observation and the infinite flow theorem we have the following corollary.

Corollary 1. *Let $\{W(k)\}$ be a random model. Suppose that $\pi^T W(k) = \pi^T$ almost surely for a vector $\pi > 0$ and all $k \geq 0$. Also, suppose that $W^i(k)^T W^j(k) \geq \gamma(W_{ij}(k) + W_{ji}(k))$ almost surely for some $\gamma > 0$ and for all $k \geq 0$ and $i, j \in [m]$ with $i \neq j$. Then, the infinite flow and the ergodicity events coincide almost surely, i.e., $\mathcal{E} = \mathscr{F}$ almost surely.*

Now, we focus on the random gossip algorithm over time-varying network. Suppose that we have the set $[m]$ of m agents, and each agent $i \in [m]$ has a scalar value $x_i(0)$ at time $t = 0$. We also have a sequence $\{P(k)\}$ of matrices such that $\sum_{i<j} P_{ij}(k) = 1$ for all k. At any time instance $k > 0$, the value $P_{ij}(k)$ is the probability that link $\{i, j\}$ is activated (independent of the link realizations in other time instances). When the link $\{i, j\}$ is activated at time k, agents i and j exchange their values and, then, update as follows:

$$x_\ell(k) = \frac{1}{2}\,(x_i(k-1) + x_j(k-1)) \qquad \text{for } \ell = i, j,$$

while the other agents do not update, i.e., $x_\ell(k) = x_\ell(k-1)$ for $\ell \neq i, j$.

In our dynamic system form, we have an independent random model $\{W(k)\}$, where

$$W(k) = I - \frac{1}{2}(e_i - e_j)(e_i - e_j)^T \qquad \text{with probability } P_{ij}(k). \tag{2}$$

By looking at the connected components of the infinite flow graph of the gossip model, we can characterize the ergodicity classes of the dynamics $x(t)$. Specifically, let G^∞ be the infinite flow graph of the gossip model. Let $S_1, \ldots, S_\tau \subset [m]$ be the connected components of G^∞, where $\tau \geq 1$ is the number of the components. We have the following result.

Lemma 1. *Consider the time-varying gossip algorithm given by* (2). *Then, for any initial vector $x(0) \in \mathbb{R}^m$, the dynamics $\{x(t)\}$ converges almost surely. Furthermore, we have $\lim_{t\to\infty}(x_i(t) - x_j(t)) = 0$ almost surely for $i, j \in S_r$ and $r \in [\tau]$.*

Proof. Let $m_r = |S_r|$ be the number of vertices in the connected component S_r of G^∞. Without loss of generality assume that the vertices are ordered such that $S_1 = \{1, \ldots, a_1\}$, $S_2 = \{a_1 + 1, \ldots, a_2\}, \ldots, S_\tau = \{a_{\tau-1} + 1, \ldots, a_\tau\}$ where $m_r = a_r - a_{r-1}$ and $a_0 = 0 < a_1 < a_2 < \cdots < a_{\tau-1} < a_\tau = m$. By the Borel-Cantelli lemma [11], page 46, and finiteness of the dimension of the space, for almost all $\omega \in \Omega$ there exists $N(\omega) < \infty$ such that no communication link $\{i, j\}$ will appear between two different components S_α and S_β for $\alpha \neq \beta$ and $\alpha, \beta \in [\tau]$ for any time $k \geq N(\omega)$. Therefore, for almost all $\omega \in \Omega$, the chain $\{W(k)\}(\omega)$ can be written in the form $W(k)(\omega) = \mathrm{diag}(W^{(1)}(k)(\omega), \ldots, W^{(\tau)}(k)(\omega))$ for $k \geq N(\omega)$ where $W^{(r)}(k)(\omega)$ are $m_r \times m_r$ stochastic matrices. From the dynamics system perspective, this means that after a while (after time $N(\omega)$), the dynamics $\{x(t)\}$ driven by $\{W(k)\}$ can be decoupled into τ disjoint dynamics systems $\{x^{(r)}\}$, where $\{x^{(r)}\}$ is a dynamic evolving in $\mathbb{R}^{m_r}$ and it is governed by $\{W^{(r)}(k)(\omega)\}$. Also, based on the Borel-Cantelli lemma [11], the deterministic models $\{W^{(r)}(k)(\omega)\}_{k\geq N(\omega)}$ have infinite flow property for $r \in [\tau]$. Now, note that for each $r \in [\tau]$ the deterministic model $\{W^{(r)}(k)(\omega)\}_{k\geq N(\omega)}$ is doubly stochastic and $W_{ii}^{(r)}(k) \geq \frac{1}{2}$ for all $i \in [m]$ and $k \geq 0$. Therefore by Corollary 1, those individual chains are ergodic almost surely. Hence, starting from any initial point $x(0) \in \mathbb{R}^m$, we have $x(t_0) = \Phi(t_0, 0)x(0)$. From this point, the m_r coordinates of $x(k)$ belonging to S_r, will evolve by the chain $\{W^{(r)}(k)\}$. Since each chain $\{W^{(r)}(k)\}$ is ergodic, it follows that $\lim_{k\to\infty}(x_i(k) - x_j(k)) = 0$ for $i, j \in S_r$ and $r \in [\tau]$. Note that although the time $N(\omega)$ is a random (stopping) time, the sets S_r are deterministic. **Q.E.D.**

Basically, Lemma 1 states that for any starting point $x(0) \in \mathbb{R}^m$, any two agents in the same connected component of G^∞ will eventually consent. From the algebraic point of view, the preceding assertion means that the matrix $\Phi(t, 0) = W(t) \cdots W(0)$ converges almost surely to a random matrix $\Psi(0)$ with the property that the ith and jth row of $\Psi(0)$ are the same vectors if $i, j \in S_r$ for some $r \in [\tau]$. Observe that the same argument holds for any starting time $t_0 \geq 0$. The choice of starting time zero was not crucial.

Our goal in the upcoming sections is to provide results similar to Lemma 1 but for a class of random models larger than the gossip model in (2).

2.4 Mutual Ergodicity and Weak Ergodicity

In order to define the ergodicity classes of a model, we introduce the following notions.

Definition 5. *Let $\{A(k)\}$ be a deterministic chain, and let $\{x(k)\}$ in (1) be driven by the chain $A(k)$. We say that:*

(a) Two indices $i, j \in [m]$ are mutually weakly ergodic indices if $\lim_{k\to\infty}(x_i(k) - x_j(k)) = 0$ for any initial time $t_0 \geq 0$ and initial point $x(t_0) \in \mathbb{R}^m$. We write $i \leftrightarrow_A j$ when the indices i and j are mutually weakly ergodic indices for the chain $\{A(k)\}$.

(b) The index $i \in [m]$ is an ergodic index for the chain $\{A(k)\}$ if the limit $\lim_{k\to\infty} x_i(k)$ 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 index, we say the chain is partially-ergodic.

(c) Two indices $i, j \in [m]$ are mutually ergodic indices if i and j are ergodic indices and $i \Leftrightarrow_A j$ for any initial time $t_0 \geq 0$ and initial point $x(t_0) \in \mathbb{R}^m$. We write $i \Leftrightarrow_A j$ when the indices i and j are mutually ergodic for the chain $\{A(k)\}$.

The relation $\leftrightarrow_A$ is induced by a chain $\{A(k)\}$ and it defines an equivalence relation on $[m]$. Hence, one can consider the equivalency classes of such a relation. We refer to the classes of this equivalence relation as *ergodicity classes* and to the partitioning of $[m]$ induced by such an equivalence relation as the *ergodicity pattern* of the model.

Definition 5 extends naturally to a random model. Specifically, if any of the properties in Definition 5 holds almost surely for a random model $\{W(k)\}$, we say that the model $\{W(k)\}$ has the corresponding property. In the further development, when unambiguous, we will omit the explicit dependency of the relation $\leftrightarrow$ and $\Leftrightarrow$ on the underlying chain.

For a random model $\{W(k)\}$, the set of realizations over which $i \leftrightarrow j$ (or $i \Leftrightarrow j$) is a measurable set; hence, an event. Since such events are tail events for an independent random model $\{W(k)\}$, each of these events happens with either probability zero or one. Hence, for a random independent model, we say $i \leftrightarrow j$ (or $i \Leftrightarrow j$) if $\mathsf{Pr}(i \leftrightarrow j) = 1$ (or $\mathsf{Pr}(i \Leftrightarrow j) = 1$). Conversely, we say $i \not\leftrightarrow j$ (or $i \not\Leftrightarrow j$), if $\mathsf{Pr}(i \leftrightarrow j) = 0$ (or $\mathsf{Pr}(i \Leftrightarrow j) = 0$). Thus, the ergodicity pattern of an independent random model is well-defined.

Now, let us discuss two examples to illustrate Definition 5.

Example 1. *By Definition 5 and the ergodicity definition (Definition 1), a chain is ergodic if and only if its ergodicity class is a singleton or, equivalently, its ergodicity pattern is* $\{[m]\}$.

Example 2. *For the time-varying gossip algorithm in Eq.* (2)*, we proved that if* i *and* j *belong to the same connected component of the infinite flow graph* G^{∞} *of the gossip model* $\{W(k)\}$*, then* $i \Leftrightarrow_W j$*. We prove in Section 5 that this statement holds for any random model satisfying the conditions of the infinite flow theorem (Theorem 1).*

3 Approximation of chains

Two random models may have the same ergodic properties while the dynamics of two chains might be completely different. For example, all the ergodic chains have the same ergodic properties, and yet, their dynamics can be quite different. Next, we specify what we mean by "the same ergodic properties" for two chains.

Definition 6. *Two chains* $\{W(k)\}$ *and* $\{U(k)\}$ *have the same ergodic properties if there exists a bijection* $\theta : [m] \to [m]$ *between the indices of* $\{W(k)\}$ *and* $\{U(k)\}$ *such that:*

(a) $i \leftrightarrow_W j$ *if and only if* $\theta(i) \leftrightarrow_U \theta(j)$*, and*

(b) i *is an ergodic index for* $\{W(k)\}$ *if and only if* $\theta(i)$ *is an ergodic index for* $\{U(k)\}$.

Note that if θ is a bijection of Definition 6 for two chains, then the indices of one of the chains can be permuted according to the bijection θ, so that the bijection θ can always be taken as identity. We assume that this is the case for the rest of the paper.

3.1 ℓ_1-Approximation

Our aim in this section is to prove that if two chains do not "differ much", then they have the same ergodic properties. In the above statement, by "not differ much", we mean that the two chains are ℓ_1-approximation of each other as defined below.

Definition 7. *A deterministic chain* $\{B(k)\}$ *is an* ℓ_1*-approximation of a chain* $\{A(k)\}$ *if* $\{A(k) - B(k)\}$ *is an* ℓ_1*-sequence, i.e.,* $\sum_{k=0}^{\infty} |A_{ij}(k) - B_{ij}(k)| < \infty$ *for all* $i, j \in [m]$.

Definition 7 extends to random chains by requiring that ℓ_1-approximation is almost sure, i.e., a random chain $\{U(k)\}$ is an ℓ_1-approximation of a random chain $\{W(k)\}$ if $\sum_{k=0}^{\infty} |W_{ij}(k) - U_{ij}(k)| < \infty$ almost surely for all $i, j \in [m]$.

We have several side remarks about Definition 7. First, we note that the ℓ_1-approximation is an equivalence relation. This is because the set of all absolutely summable sequences in $\mathbb{R}$ is a vector space over $\mathbb{R}$. The second remark is that, for two independent random models $\{W(k)\}$ and $\{U(k)\}$ that are adapted to the same sigma-field, we have $\sum_{k=0}^{\infty} |W_{ij}(k) - U_{ij}(k)| < \infty$ for all $i, j \in [m]$ with either probability zero or one, due to Kolmogorov's 0-1 law ([11], page 61). The third remark is about an alternative formulation of ℓ_1-approximation. Since the matrices have finite dimension, if $\sum_{k=0}^{\infty} |A_{ij}(k) - B_{ij}(k)| < \infty$ for all $i, j \in [m]$, then an equivalent characterization of

$\{B(k)\}$ being ℓ_1-approximation of $\{A(k)\}$ is that $\sum_{k=0}^{\infty} \|A(k) - B(k)\|_p < \infty$ for any $p \geq 1$.

We now discuss some examples of ℓ_1-approximation.

Example 3. *Consider two random models $\{W(k)\}$ and $\{U(k)\}$ that differ only in finitely many coordinate maps, i.e., $W(k) = U(k)$ for all $k \geq t$ and some $t \geq 0$. Since any entry of each random matrix $W(k)$ and $U(k)$ is bounded in the interval $[0,1]$, it follows that $\sum_{k=0}^{\infty} |W_{ij}(k) - U_{ij}(k)| \leq t$. Hence, the two models are ℓ_1-approximation of each other. We use such an approximation in the proof of Theorem 6.*

For another example of ℓ_1-approximation, consider the time-varying gossip model, as discussed in Section 2.

Example 4. *Let $\{W(k)\}$ be the chain of the gossip model in Eq. (2) with the infinite flow graph G^{∞}. Assume that the connected components of G^{∞} are $S_1, \ldots, S_{\tau}$. Now, define the approximate gossip model $\{U(k)\}$ to be*

$$U(k) = \begin{cases} W(k) & \text{if link } \{i,j\} \text{ is activated at time } k \text{ with } i,j \in S_r \text{ for some } r, \\ I & \text{otherwise.} \end{cases} \tag{3}$$

In this case, by the definition of the infinite flow graph we have $\sum_{k=0}^{\infty} |W_{ij}(k) - U_{ij}(k)| < \infty$ almost surely for all i and j that do not belong to the same connected component of G^{∞}. Basically, in the approximate gossip model, we cut the links between the agents that belong to different connected components of G^{∞}. In this way, we have an approximate dynamic consisting of τ decoupled dynamics (one per connected component of G^{∞}). At the same time, the original and the approximate dynamics have the same ergodic properties. This will be seen for a more general case in forthcoming Lemma 2.

Now, we present the main result of this section. The result shows that if two chains are ℓ_1-approximations of each other, then their ergodic properties are identical. Such a result might be helpful in determining the ergodic properties of a chain that may not be easily dealt with directly, but it may admit a suitable l_1-approximation.

Lemma 2. *(Approximation lemma) Let 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.*

Proof. Since $\{B(k)\}$ is an ℓ_1-approximation of $\{A(k)\}$, for any given $\epsilon > 0$, there exists integer $N_{\epsilon} \geq 0$ such that $\sum_{k=N_{\epsilon}}^{\infty} \|A(k) - B(k)\|_{\infty} < \epsilon$.

Suppose that $i \leftrightarrow_B j$. For the given $x(0) \in [0,1]^m$, let $\{x(k)\}$ be the sequence generated by matrices $\{A(k)\}$ and the dynamics defined in Eq. (1). Also, let $\{z(k)\}_{k \geq N_{\epsilon}}$ be the dynamics driven by $\{B(k)\}$ started at time N_{ϵ} by vector $z(N_{\epsilon}) = x(N_{\epsilon})$ as defined in Eq. (1). For any k, we have:

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

But $|x_i(k)| \leq 1$ for any $k \geq N_{\epsilon}$. Thus, we have

$$\|x(k) - B(k-1)x(k-1)\|_{\infty} \leq \|A(k-1) - B(k-1)\|_{\infty}. \tag{4}$$

Therefore, $\|x(N_\epsilon+1) - B(N_\epsilon)x(N_\epsilon)\|_\infty \le \|A(N_\epsilon) - B(N_\epsilon)\|_\infty$. Since $z(N_\epsilon) = x(N_\epsilon)$, it follows $\|x(N_\epsilon+1) - z(N_\epsilon+1)\|_\infty \le \|A(N_\epsilon) - B(N_\epsilon)\|_\infty$. We now proceed by induction and we assume that $\|x(k) - z(k)\|_\infty \le \sum_{t=N_\epsilon}^{k-1} \|A(k) - B(k)\|_\infty$ for $k > N_\epsilon$. Using Eq. (4) and triangle inequality, we have

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

Now, by the induction assumption and the relation $\|B(k)\|_\infty \le 1$, which holds since $B(k)$ is a stochastic, it follows that

$$\|x(k+1) - z(k+1)\|_\infty \le \sum_{t=N_\epsilon}^{k} \|A(k) - B(k)\|_\infty \le \sum_{t=N_\epsilon}^{\infty} \|A(k) - B(k)\|_\infty \le \epsilon. \tag{5}$$

Therefore, $|x_i(k) - z_i(k)| \le \epsilon$ and $|z_j(k) - x_j(k)| \le \epsilon$ and hence, by the triangle inequality $|(x_i(k) - x_j(k)) + (z_i(k) - z_j(k))| \le 2\epsilon$ for any $k \ge N_\epsilon$. But $i \leftrightarrow_B j$ and, hence, $\lim_{k\to\infty}(z_i(k) - z_j(k)) = 0$ and we have

$$\limsup_{k\to\infty} |x_i(k) - x_j(k)| \le 2\epsilon.$$

The preceding inequality holds for any $\epsilon > 0$ and any starting time t_0 for $\{B(k)\}$, implying that $i \leftrightarrow_A j$.

Using the same argument, and based on inequality (5), one can deduce that if i is ergodic index for $\{B(k)\}$, then it is also ergodic index for $\{A(k)\}$. Since ℓ_1-approximation is symmetric with respect to the chains, the result follows. **Q.E.D.**

An immediate consequence of ℓ_1-approximation lemma is the following result.

Corollary 2. *The class of ergodic deterministic chains is closed under ℓ_1-approximations and hence, the class of ergodic random models is closed under ℓ_1-approximations.*

3.2 Infinite Flow Graph and Mutual Ergodicity

We now explore the relation between mutual weak ergodicity and the infinite flow graph. More specifically, we consider a random model $\{W(k)\}$ and its infinite flow graph $G^\infty = ([m], \mathcal{E}^\infty)$, as given in Definition 4. We show that if indices i and j are mutually weakly ergodic, then i and j belong to the same connected component of G^∞. This result is general as it applies to any random model. We establish it for arbitrary deterministic chain.

Lemma 3. *Let $\{A(k)\}$ be a deterministic chain. Then, $i \leftrightarrow_A j$ implies that i and j belong to the same connected component of the infinite flow graph of $\{A(k)\}$.*

Proof. To arrive at a contradiction, suppose that i and j belong to two different connected components $S, T \subset [m]$ of G^∞. Therefore, $T \subset \bar{S}$ implying that $\bar{S}$ is not empty. Also, since S is a (maximal) connected component of G^∞, it follows that $\sum_{k=0}^{\infty} A_S(k) < \infty$. Without loss of generality, we can assume that $S = \{1, \ldots, i^*\}$ for some $i^* < m$. Then, consider the chain $\{B(k)\}$ as defined below:

$$B_{ij}(k) = \begin{cases} A_{ij}(k) & \text{if } i \neq j \text{ and } i, j \in S \text{ or } i, j \in \bar{S}, \\ 0 & \text{if } i \neq j \text{ and } i \in S, j \in \bar{S} \text{ or } i \in \bar{S}, j \in S, \\ A_{ii}(k) + \sum_{\ell \in \bar{S}} A_{i\ell}(k) & \text{if } i = j \in S, \\ A_{ii}(k) + \sum_{\ell \in S} A_{i\ell}(k) & \text{if } i = j \in \bar{S}. \end{cases} \tag{6}$$

The above approximation simply sets the cross terms between S and $\bar{S}$ to zero and adds the corresponding values to the diagonal entries to maintain the stochasticity of the matrix. Therefore, we have new stochastic chain $B(k)$ given by

$$B(k) = \begin{bmatrix} B_1(k) & 0 \\ 0 & B_2(k) \end{bmatrix},$$

where $B_1(k)$ and $B_2(k)$ are respectively $i^* \times i^*$ and $(m - i^*) \times (m - i^*)$ matrices for all k. Note that by the assumption $\sum_{k=0}^{\infty} A_S(k) < \infty$, the chain $\{B(k)\}$ is an ℓ_1-approximation of $\{A(k)\}$. Now, let u_{i^*} be the vector which has the first i^* coordinates equal to one and the rest equal to zero, i.e., $u_{i^*} = \sum_{\ell=1}^{i^*} e_{i^*}$. Then $B(k)u_{i^*} = u_{i^*}$ for any $k \geq 0$ and hence, $i \not\leftrightarrow_B j$. By Approximation Lemma (Lemma 2), we have $i \not\leftrightarrow_A j$ and the result follows. **Q.E.D.**

Lemma 3 shows that the ergodic components of a random model $\{W(k)\}$ are subsets of the connected components of the infinite flow graph of the model. Therefore, if the model is ergodic, then the infinite flow graph of the model is connected which proves the necessity of the infinite flow property for the ergodicity of a model. Thus, as a special consequence of Lemma 2, we obtain Theorem 1 in [31].

The converse result of Lemma 3 is not true. For example, let

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

In this case, the infinite flow graph is connected while the model is not ergodic. In the resulting dynamics, the agents 1 and 2 keep swapping their initial values $x_1(0)$ and $x_2(0)$.

4 Ergodicity in class $\mathscr{M}_2$

In this section, we show that the infinite flow property is sufficient for ergodicity of a class of random models. Assume that $\{W(k)\}$ is an independent random chain with common steady state distribution π in expectation. Let $x(0) \in \mathbb{R}^m$ and let $\{x(k)\}$ be the resulting dynamic governed by $\{W(k)\}$. In [31] (Theorem 5), it was shown that for any independent random model $\{W(k)\}$ and any starting point $x(0) \in \mathbb{R}^m$, we almost surely have

$$\sum_{k=0}^{\infty} \sum_{i<j} \bar{H}_{ij}(k)(x_i(k) - x_j(k))^2 < \infty,$$

where $\bar{H}(k) = \mathsf{E}\big[W^T(k)\mathrm{diag}(\pi)W(k)\big]$. This result have played a crucial role when proving that, for certain independent models with common steady state $\pi > 0$ in expectation, the infinite flow is both necessary and sufficient for ergodicity. We will show in Theorem 3 that the infinite flow property is in fact necessary and sufficient for the ergodicity of the models $\{W(k)\}$ such that $\sum_{k=0}^{\infty}\sum_{i<j}\mathsf{E}\big[W^T(k)W(k)\big]_{ij}\mathsf{E}[(x_i(k)-x_j(k))^2] < \infty$. We refer to such models as class $\mathscr{M}_2$-models and we formally introduce this class, as follows.

Definition 8. *The class $\mathscr{M}_2$ of random models is the set of all independent random models $\{W(k)\}$ such that for the dynamic system* (1) *we have for any $t_0 \geq 0$ and $x(t_0) \in \mathbb{R}^m$,*

$$\sum_{k=t_0}^{\infty}\sum_{i<j} H_{ij}(k)\mathsf{E}\big[(x_i(k)-x_j(k))^2\big] < \infty, \tag{7}$$

where $H(k) = \mathsf{E}\big[W^T(k)W(k)\big]$.

If a model has a common steady state $\pi > 0$ in expectation, then $\pi_{\min}\mathsf{E}\big[W^T(k)W(k)\big] \leq \mathsf{E}\big[W^T(k)\mathrm{diag}(\pi)W(k)\big]$, where $\pi_{\min} = \min_{i\in[m]}\pi_i$. Therefore, any model with a common steady state $\pi > 0$ belongs to class $\mathscr{M}_2$. The following result shows that any ℓ_1-approximation of such models also belongs to class $\mathscr{M}_2$.

Lemma 4. *Let $\{W(k)\}$ be an independent random model with a common steady state $\pi > 0$ in expectation. Let an independent random model $\{U(k)\}$ be an ℓ_1-approximation of $\{W(k)\}$. Then $\{U(k)\} \in \mathscr{M}_2$.*

The proof of Lemma 4 can be found in the appendix.

Next, we show that the infinite flow property is sufficient for any model in class $\mathscr{M}_2$ model with weak feedback property. We establish this by using the following intermediate result whose proof is provided in the appendix.

Theorem 2. *Let $\{A(k)\}$ be a deterministic chain with infinite flow property, and let $z(k) = A(k-1)z(k-1)$ for $k \geq 1$ and $z(0) = e_\ell$ for any $\ell \in [m]$. Assume that $\lim_{k\to\infty}(z_{\max}(k) - z_{\min}(k)) > 0$. Then,*

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

By using Theorem 2, we establish the following result showing that the infinite flow property is necessary and sufficient for the ergodicity of class $\mathscr{M}_2$-models.

Theorem 3. *Let $\{W(k)\}$ be a class $\mathscr{M}_2$ model with weak feedback property. Then, the infinite flow property is both necessary and sufficient for the ergodicity of the model.*

Proof. The necessity of the infinite flow property follows by Lemma 3. For the converse, assume that the model has the infinite flow property. Note that by the definition of an $\mathscr{M}_2$-model, for any $x(0) \in \mathbb{R}^m$,

$$\sum_{k=0}^{\infty}\sum_{i<j} H_{ij}(k)\mathsf{E}\big[(x_i(k)-x_j(k))^2\big] < \infty,$$

where $H(k) = \mathsf{E}\big[W^T(k)W(k)\big]$. Due to the feedback property, we have $\mathsf{E}\big[W^i(k)^T W^j(k)\big] \geq \gamma \mathsf{E}[W_{ij}(k) + W_{ji}(k)]$ for some $\gamma > 0$. Therefore, using the independency of the model and the relation $H_{ij}(k) = \mathsf{E}\big[W^i(k)^T W^j(k)\big]$, for any $x(0) \in \mathbb{R}^m$ we obtain

$$\begin{aligned}
&\sum_{k=0}^{\infty} \sum_{i<j} \mathsf{E}\big[(W_{ij}(k) + W_{ji}(k))(x_i(k) - x_j(k))^2\big] \\
&= \mathsf{E}\left[\sum_{k=0}^{\infty} \sum_{i<j} (W_{ij}(k) + W_{ji}(k))(x_i(k) - x_j(k))^2\right] < \infty,
\end{aligned}$$

where the equality holds by $(W_{ij}(k) + W_{ji}(k))(k)(x_i(k) - x_j(k))^2 \geq 0$ and the monotone convergence theorem [13], page 50. As a result, for any $x(0)$,

$$\sum_{k=0}^{\infty} \sum_{i<j} (W_{ij}(k) + W_{ji}(k))(k)(x_i(k) - x_j(k))^2 < \infty \qquad \text{almost surely.}$$

Since the model is assumed to have the infinite flow property, by Theorem 2 we have $\lim_{k\to\infty}(x_{\max}(k) - x_{\min}(k)) = 0$ almost surely for $x(0) = e_\ell$ with $\ell \in [m]$, implying that the system reaches consensus almost surely. Since $\sum_{k=t}^{\infty} \sum_{i<j} H_{ij}(k) \mathsf{E}[(x_i(k) - x_j(k))^2] < \infty$ for any starting time $t \geq 0$, by the same argument it follows that the model is ergodic. **Q.E.D.**

Using Theorem 2, we can extend the Infinite Flow Theorem 1 as follows.

Theorem 4. *The Infinite Flow Theorem 1 holds for any random model $\{W(k)\}$ in class $\mathscr{M}_2$ with weak-feedback property.*

Proof. By Theorem 3, (d) implies (a). The implications (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) is true for any independent random model as proven in [31]. **Q.E.D.**

As an example of a model in the class $\mathscr{M}_2$ that does not have a common steady state, consider the class of deterministic chains $\{A(k)\}$ that satisfy a bounded-connectivity condition and have a uniform lower-bound on positive entries, as discussed in [33, 15, 19, 20, 21]. In these models, the sequence $d(x(k)) = \max_{i\in[m]} x_i(k) - \min_{j\in[m]} x_j(k)$ is (sub)geometric and, thus, it is absolutely summable. Furthermore, $H_{ij}(k) = [A^T(k)A(k)]_{ij} \leq m$. This and $|x_i(k) - x_j(k)| \leq d(x(k))$ for all $i, j \in [m]$ imply $\sum_{k=0}^{\infty} \sum_{i<j} [A^T(k)A(k)]_{ij}(x_i(k) - x_j(k))^2 < \infty$, i.e., the defining property of $\mathscr{M}_2$-class in Eq. (7) holds.

5 Infinite Flow Graph

So far, we have been focused on the infinite flow property of independent chains and showed that for a class of random models, infinite flow property and ergodicity are equivalent. We next study models with a common steady state $\pi > 0$ in expectation and weak-feedback property that *do not* have infinite flow property. Our goal is to investigate the limiting behavior of such models and, in particular, to characterize their ergodicity classes.

Let $\{W(k)\}$ be a random model with common steady state $\pi > 0$ in expectation. Let G^∞ be the infinite flow graph of the model as in Definition 4. In Section 2.3, we showed that on extended gossip algorithm, the connected components of this graph are closely related to the limiting behavior of the model. We aim to show that the same results hold for any independent random model with a common steady state $\pi > 0$ in expectation and weak feedback property.

Suppose that G^∞ has $\tau \geq 1$ connected components. Let $S_1, \ldots, S_\tau \subset [m]$ be the sets of vertices of each connected component of G^∞. Let $S_1 = \{1, \ldots, a_1\}$, $S_2 = \{a_1 + 1, \ldots, a_2\}, \ldots, S_\tau = \{a_{\tau-1} + 1, \ldots, a_\tau = m\}$ for $1 \leq a_1 \leq \ldots \leq a_\tau = m$, and let $m_r = |S_r| = a_r - a_{r-1}$ be the number of vertices in the rth component, where $a_0 = 0$.

Using the connected components of the infinite flow graph of $\{W(k)\}$, we define the *diagonal approximation* $\{\tilde{W}(k)\}$ of $\{W(k)\}$, as follows.

Definition 9. *(Diagonal approximation) For $1 \leq r \leq \tau$, define the random model $\{W^{(r)}(k)\}$ in $\mathbb{R}^{m_r}$ as follows: for $i, j \in [m_r]$,*

$$W_{ij}^{(r)}(k) = \begin{cases} W_{(i+a_{r-1})(i+a_{r-1})}(k) + \sum_{\ell \in \bar{S}_r} W_{(i+a_{r-1})\ell}(k) & \text{if } j = i, \\ W_{(i+a_{r-1})(j+a_{r-1})}(k) & \text{if } j \neq i. \end{cases} \tag{8}$$

The diagonal approximation of the model $\{W(k)\}$ is the model $\{\tilde{W}(k)\}$ defined by

$$\tilde{W}(k) = diag(W^{(1)}(k), \ldots, W^{(\tau)}(k)) = \begin{pmatrix} W^{(1)}(k) & 0 & \cdots & 0 \\ 0 & W^{(2)}(k) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & W^{(\tau)}(k) \end{pmatrix}. \tag{9}$$

Basically, in the diagonal approximation, we are decoupling the components of our dynamics that have weak interactions with each other.

In fact, the diagonal approximation of a given model is an ℓ_1-approximation of the model.

Lemma 5. *Let $\{W(k)\}$ be an independent model and $\{\tilde{W}(k)\}$ be its diagonal approximation. Then, $\{\tilde{W}(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$. Furthermore, the models $\{W^{(r)}(k)\}$, $r = 1, \ldots, \tau$ defined in Eq. (8) have infinite flow property.*

The proof of Lemma 5 is presented in the Appendix.

In Lemma 3, we showed that if i and j are mutually weakly ergodic then i and j belong to the same connected component of G^∞. In Lemma 1, we showed that for the generalized gossip algorithm, i and j are mutually ergodic if i and j belong to the same connected component of the infinite flow graph of the model. Hence, in this particular case, i and j are mutually ergodic if and only if i and j belong to the same connected component of G^∞. The following theorem shows that the same result holds for any random model satisfying the conditions of the Infinite Flow Theorem 1.

Theorem 5. *Let $\{W(k)\}$ be an independent model with a common steady state $\pi > 0$ in expectation and with weak feedback property. Then $i \Leftrightarrow j$ if and only if i and j are in the same connected component of the infinite flow graph G^∞ of the model.*

Proof. The "if" part follows from Lemma 3 and the fact that mutual ergodicity implies mutual weak ergodicity. To show the "only if" part, we use two ℓ_1-approximations successively. We proceed through the following steps to prove this result. First, in order to have the weak-feedback property, we construct an ℓ_1-approximation of the diagonal approximation for the original model $\{W(k)\}$ and we prove that the resulting chain has weak-feedback property. Then, using the developed tools in the preceding sections, we prove the statement.

Let G^∞ be the infinite flow graph of $\{W(k)\}$ 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 $\tilde{W}^{(r)}(k)$ defined as in Eq. (8) for $r \in [\tau]$. Let $M(k) = \mathsf{E}\Big[\max_{i,j\in[m]} |\tilde{W}_{ij}(k) - W_{ij}(k)|\Big]$. Since $\{\tilde{W}(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$, we have $\sum_{k=0}^\infty M(k) < \infty$. 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 a 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 the $[0, 1]$ interval, 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 $U(k) = \mathrm{diag}(U^{(1)}(k), \ldots, U^{(\tau)}(k))$, $k \geq 0$ is an ℓ_1-approximation of $\{\tilde{W}(k)\}$. By Lemma 5, $\{\tilde{W}(k)\}$ 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:

$$\begin{aligned} {Q^i}^T Q^j &= \Big((1-d(k))\tilde{W}^{(r)i_r}(k) + d(k)\pi^{(r)}\Big)^T \Big((1-d(k))\tilde{W}^{(r)j_r}(k) + d(k)\pi^{(r)}\Big) \\ &\geq (1-d(k))^2(\tilde{W}^{(r)i_r}(k))^T \tilde{W}^{(r)j_r}(k) + (1-d(k))d(k)\pi^{(r)T}(\tilde{W}^{(r)i_r}(k) + \tilde{W}^{(r)j_r}(k)). \end{aligned} \tag{10}$$

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,

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

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. (10) and using the preceding inequalities, we obtain

$$\begin{aligned}\mathsf{E}\left[Q^{iT}Q^j\right] &\geq (1-d(k))^2\mathsf{E}\left[(\tilde{W}^{i_r}(k))^T\tilde{W}^{j_r}(k)\right] + (1-d(k))d(k)(\pi_i+\pi_j-2M(k)) \\ &\geq (1-d(k))^2\mathsf{E}\left[(\tilde{W}^{i_r}(k))^T\tilde{W}^{j_r}(k)\right] + (1-d(k))d(k)\frac{\pi_i+\pi_j}{2}, \end{aligned} \tag{11}$$

which holds by $M(k) \leq \frac{\pi_{\min}}{8m} \leq \frac{\pi_i}{4}$ for any $i \in [m]$ 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. (11), we have:

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

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

$$2mM(k) = d(k)\frac{\pi_{\min}}{2} \leq d(k)\,\frac{\pi_i+\pi_j}{4}.$$

By combining the preceding two relations and using $(1-d(k))^2 \leq (1-d(k))$, we have

$$\begin{aligned}\mathsf{E}\left[Q^{iT}Q^j\right] &\geq (1-d(k))^2\mathsf{E}\left[(W^{i_r}(k))^T W^{j_r}(k)\right] + (1-d(k))d(k)\,\frac{\pi_i+\pi_j}{4} \\ &\geq (1-d(k))^2\gamma\mathsf{E}[W_{i_rj_r}(k) + W_{j_ri_r}(k)] + (1-d(k))d(k)\,\frac{\pi_i+\pi_j}{4}, \end{aligned} \tag{12}$$

where the last inequality follows by weak feedback property of $\{W(k)\}$.

Since $i_r, j_r \in S_r$ and $i_r \neq j_r$, by the construction of $\tilde{W}(k)$, we have $\tilde{W}_{i_rj_r}(k) = W_{i_rj_r}(k)$. Hence, $\mathsf{E}[Q_{ij}+Q_{ji}] = (1-d(k))(\mathsf{E}[W_{i_rj_r}(k)] + \mathsf{E}[W_{j_ri_r}(k)]) + d(k)(\pi_i+\pi_j)$. By combining this with Eq. (12), we have

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

Without loss of generality, we assume $\gamma \leq \frac{1}{4}$ (otherwise, we can replace γ by $\frac{1}{4}$). Thus,

$$\mathsf{E}\left[Q^{iT}Q^j\right] \geq \frac{\gamma}{2}\,\mathsf{E}[Q_{ij}+Q_{ji}]\,,$$

which follows from $d(k) \leq \frac{1}{2}$. Note that we defined $Q = U^{(r)}(k)$ 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 4, it follows that $\{U(k)\} \in \mathscr{M}_2$. Hence, $\sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)\mathsf{E}[(x_i(k) - x_j(k))^2] < \infty$ almost surely, where $L(k) = \mathsf{E}\big[U^T(k)U(k)\big]$. Hence, for any $r \in [\tau]$,

$$\sum_{k=0}^{\infty} \sum_{\substack{i<j \\ i,j \in S_r}} L_{ij}(k)\mathsf{E}\big[(x_i(k) - x_j(k))^2\big] < \infty.$$

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 $r \in [\tau]$,

$$\begin{aligned} L_{i_r j_r}(k) &= \mathsf{E}\left[\sum_{\ell \in [m]} \pi_\ell U_{\ell i_r}(k) U_{\ell j_r}(k)\right] = \mathsf{E}\left[\sum_{\ell \in S_r} \pi_\ell U^{(r)}_{\ell i_r}(k) U^{(r)}_{\ell j_r}(k)\right] \\ &= \mathsf{E}\left[\sum_{\bar{\ell} \in [m_r]} \pi^{(r)}_{\bar{\ell}} U^{(r)}_{\bar{\ell} i}(k) U^{(r)}_{\bar{\ell} j}(k)\right]. \end{aligned} \tag{13}$$

Therefore, by the above observations, the random dynamics in $\mathbb{R}^m$ induced by $\{U(k)\} \in \mathscr{M}_2$, decomposes into τ random dynamics in $\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_\tau}$ induced by $\{U^{(1)}(k)\}, \ldots, \{U^{(\tau)}(k)\}$ all of which belong to class $\mathscr{M}_2$. 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 2, $\{U^{(r)}(k)\}$ is ergodic chain for any $r \in [\tau]$ which implies $i \Leftrightarrow j$ for any $i, j \in S_r$. Since, $U(k) = \operatorname{diag}(U^{(1)}(k), \ldots, U^{(\tau)}(k))$, hence, $i \Leftrightarrow j$ in $\{U(k)\}$. Therefore, by Approximation Lemma 2, $i \Leftrightarrow j$ 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 $r \in [\tau]$. **Q.E.D.**

By the above result, we have the following characterization of the ergodicity classes.

Theorem 6. *(Extended Infinite Flow Theorem) Let $\{W(k)\}$ be any ℓ_1-approximation of an independent random model with common steady state $\pi > 0$ in expectation and weak feedback property. Let G^∞ be the infinite flow graph of $\{W(k)\}$ and $\bar{G}^\infty$ be the infinite flow graph of the expected model $\{\bar{W}(k)\}$, where $\bar{W}(k) = \mathsf{E}[W(k)]$. Then, the following statements are equivalent:*

(a) $i \Leftrightarrow_W j$.

(b) $i \Leftrightarrow_{\bar{W}} j$.

(c) i and j belong to the same connected component of $\bar{G}^\infty$.

(d) i and j belong to the same connected component of G^∞.

By Theorem 5 and Theorem 6 it follows that, any dynamics driven by a random model satisfying the assumptions of the Extended Infinite Flow theorem (Theorem 6), converges almost surely. In general, this consequence does not happen if any assumption of the theorem is violated, as seen in the following examples.

Example 5. *Let the matrices* $W(k)$ *be given by*

$$W(k) = \begin{bmatrix} 1 & 0 & 0 \\ u_1(k) & u_2(k) & u_3(k) \\ 0 & 0 & 1 \end{bmatrix},$$

where $u(k) = (u_1(k), u_2(k), u_3(k))^T$ *are i.i.d. random vectors distributed uniformly in the probability simplex of* $\mathbb{R}^3$*. Then, starting from a point* $x(0) = (0, \frac{1}{2}, 1)^T$*, the dynamics never converges. Note that this model has infinite flow property and satisfies all assumptions of Theorem 6 except for the assumption* $\pi > 0$*.*

Example 6. *Consider the random permutation model. Specifically, let* $W(k)$ *be the i.i.d. model with* $W(k)$ *randomly and uniformly chosen from the set of permutation matrices in* $\mathbb{R}^m$*. Starting from any initial point, this model just permutes the coordinates of the initial point. Therefore, the dynamic is not converging for any* $x(0)$ *that lies outside the subspace spanned by the vector* e*. The model has the infinite flow property and has the common steady state* $\pi = \frac{1}{m}e$ *in expectation. However, the model does not have weak feedback property, since* $\mathsf{E}\left[W^i(k)^T W^j(k)\right] = 0$ *for* $i \neq j$ *while* $\mathsf{E}[W_{ij}(k)] + \mathsf{E}[W_{ji}(k)] > 0$*.*

6 Applications

Here, we consider some applications of the Infinite Flow Theorem 1 and its extended variant to ergodicity classes in Theorem 6. First, we visit the broadcast-gossip model over time-changing networks and we derive a necessary and sufficient condition for the ergodicity of the model. Then, we consider the effect of a link failure process on random networks and provide an infinite flow type condition for the ergodicity.

6.1 Broadcast Gossip Algorithm on Time-Changing Networks

Broadcast gossip algorithm has been presented and analyzed in [2, 3]. As in the case of gossip algorithm, the broadcast gossip algorithm has been proposed for consensus over a static network. Here, we propose broadcast gossip algorithm over time-changing networks and provide a necessary and sufficient condition for ergodicity.

Suppose that we have a sequence of simple undirected graphs (networks) $\{G(k)\}$ on $[m]$, where $G(k) = ([m], \mathcal{E}(k))$, so that $G(k)$ represents the topology of the network at time k. The sequence $\{G(k)\}$ is assumed to be deterministic. Suppose that at time k, agent $i \in [m]$ wakes up with probability $\frac{1}{m}$ (independently of earlier choices and other agents) and broadcasts its value to its neighboring agents $N_i(k) = \{j \in [m] \mid \{i,j\} \in \mathcal{E}(k)\}$. At this time, each agent $j \in N_i(k)$ updates its estimate as follows:

$$x_j(k+1) = \gamma(k)x_i(k) + (1 - \gamma(k))x_j(k),$$

where $\gamma(k) \in (0, \gamma]$ is a mixing parameter of the system at time k and $\gamma \in (0, 1)$. The other agents keep their values unchanged, i.e., $x_j(k+1) = x_j(k)$ for $j \notin N_i(k)$. Therefore, in this case the vector $x(k)$ of agents' estimates $x_i(k)$ evolves in time according to (1) where

$$W(k) = I - \gamma(k) \sum_{j \in N_i(k)} e_j (e_j - e_i)^T \qquad \text{with probability } \tfrac{1}{m}. \tag{14}$$

Let G_b^∞ be the infinite flow graph of the broadcast gossip model, and suppose that this graph has τ connected components, namely $S_1, \ldots, S_\tau$. Using Theorem 6, we have the following result.

Lemma 6. *In time-varying broadcast gossip model of* (14)*, any two agents are in the same ergodicity class if and only if they belong to the same connected component of G_b^∞. In particular, the model is ergodic if and only if G_r^∞ is connected,*

Proof. In view of Theorem 6, it suffices to show that the broadcast gossip model in (14) has a common steady state $\pi > 0$ in expectation and weak feedback property. Since each node (agent) is chosen uniformly at each time instance and the graph $G(k)$ is undirected, the (random) entries $W_{ij}(k)$ and $W_{ji}(k)$ have the same distribution. Therefore, the expected matrix $\mathsf{E}[W(k)]$ is a doubly stochastic matrix for any time $k \geq 0$. Also, due to the condition $\gamma(k) \leq \gamma < 1$, we have $W_{ii}(k) \geq 1 - \alpha(k) \geq \gamma$ for all $i \in [m]$ and all $k \geq 0$. If a model is such that $W_{ii}(k) \geq \gamma > 0$ for all i and k, then the model has weak feedback property with $\frac{\gamma}{m}$, as implied by Lemma 7 in [31]. **Q.E.D.**

As a matter of fact, we can provide a characterization of the connected components S_r for the infinite flow graph G_b^∞. By Theorem 6, the model and its expected model have the same connected components. Therefore, it suffices to determine the infinite flow graph $\bar{G}_b^\infty$ of the expected model. A link $\{i, j\}$ is in the edge-set of the graph $\bar{G}_b^\infty$ if and only if $\sum_{k=0}^\infty (\mathsf{E}[W_{ij}(k)] + \mathsf{E}[W_{ji}(k)]) = \infty$. By (14), we have $\mathsf{E}[W_{ij}(k)] = \frac{1}{m}\gamma(k)$ if $j \in N(k)$ and otherwise $\mathsf{E}[W_{ij}(k)] = 0$. Thus, $\{i, j\} \in \bar{G}_b^\infty$ if and only if $\sum_{k:\{i,j\} \in \mathcal{E}(k)} \gamma(k) = \infty$.

Two instances of the time-varying broadcast gossip algorithm that might be of practical interest are: (1) The case where the underlying sequence of graphs is not time-changing, i.e., $G(k) = G$ for all $k \geq 0$. Then, the random model is ergodic if and only if G is connected and $\sum_{k=0}^\infty \gamma(k) = \infty$. (2) The case where the sequence $\gamma(k)$ is also bounded below i.e., $\gamma(k) \in [\gamma_b, \gamma]$ with $0 < \gamma_b \leq \gamma < 1$. Then, the model is ergodic if and only if in the sequence $\{G(k)\}$ infinitely many edges exist between S and $\bar{S}$ for any nonempty $S \subset [m]$.

6.2 Link Failure Models

Suppose that we are given a random model $\{W(k)\} \subset \mathbb{S}^m$ and a matrix sequence $\{F(k)\} \subset \mathbb{S}^m$ such that $F_{ij} = 0$ or $F_{ij} = 1$ for all i, j. Thus, $\{F(k)\}$ is a binary matrix sequence, which we refer to as a *failure process*. We define the *link-failure model* as the random model $\{U(k)\}$ given by

$$U(k) = W(k) \cdot (ee^T - F(k)) + \operatorname{diag}([W(k) \cdot F(k)]e), \tag{15}$$

where "$\cdot$" denotes the element-wise product of two matrices. To give a motivation for this definition, suppose that we have a random model $\{W(k)\}$ and suppose that each entry $W_{ij}(k)$ is set to zero, or in other words *fails*, whenever $F_{ij}(k) = 1$. In this way, $F(k)$ induces a failure pattern on $W(k)$. The term $W(k) \cdot (ee^T - F(k))$ in Eq. (15) reflects this effect. Therefore, $W(k) \cdot (ee^T - F(k))$ misses some of the entries of $W(k)$. This lack is compensated by introducing the feedback term which is equal to the sum of the failed links, i.e., the term $\text{diag}([W(k) \cdot F(k)]e)$. This can be viewed as adding $\sum_{j \neq i}[W(k) \cdot F(k)]_{ij}$ to the self-feedback term of the agent i at time k (which is $W_{ii}(k)$) in order to maintain the stochasticity of the resulting matrices.

By the above discussion, the stochasticity of the link-failure model $\{U(k)\}$ follows. Here, we should highlight the fact that the failure process is the binary matrix process $\{F(k)\}$ while the link-failure model $\{U(k)\}$ is a stochastic matrix process derived from a random model $\{W(k)\}$ and the underlying link-failure process $\{F(k)\}$.

Now, let us define models with *feedback property*. The random model $\{W(k)\}$ has feedback property if there is $\gamma > 0$ such that $\mathsf{E}[W_{ii}(k)W_{ij}(k) + W_{jj}(k)W_{ji}(k)] \geq \gamma \mathsf{E}[W_{ij}(k) + W_{ji}(k)]$ for any $k \geq 0$ and $i, j \in [m]$ with $i \neq j$. In general, this property is stronger than weak feedback property, as proved in [31]. However, as seen in [31], the class of models with feedback property contains i.i.d. models with almost sure positive diagonal entries.

Our next goal is to study the effect of the *uniform link failure process* $\{F(k)\}$ on independent random models with common steady state $\pi > 0$ in expectation and with feedback property.

Definition 10. *A uniform link-failure process is the link-failure process $\{F(k)\}$ that satisfies the following conditions:*

(a) The random variables $\{F_{ij}(k) \mid i, j \in [m],\ i \neq j\}$ are binary i.i.d. for any fixed $k \geq 0$.

(b) The failure process $\{F(k)\}$ is an independent process in time.

Note that the i.i.d. condition in Definition 10 is assumed for a fixed time. Therefore, the uniform link-failure model can have a time-dependent distribution but for any given time the distribution of the link-failure should be identical across the different edges.

Lemma 7. *Let $\{W(k)\}$ be an independent model with common steady state $\pi > 0$ in expectation and feedback property. Let $\{F(k)\}$ be a uniform-link failure process that is independent of $\{W(k)\}$. Then, the failure model $\{U(k)\}$ is ergodic if and only if $\sum_{k=0}^{\infty}(1 - p_k)W_S(k) = \infty$ for all nonempty $S \subset [m]$, where $p_k = \Pr(F_{ij}(k) = 1)$.*

Proof. By the definition of the failure model $\{U(k)\}$ in (15), since both random processes $\{W(k)\}$ and $\{F(k)\}$ are independent, the failure model $\{U(k)\}$ is also independent. Then, for $i \neq j$ and for any $k \geq 0$, we have

$$\mathsf{E}[U_{ij}(k)] = \mathsf{E}[W_{ij}(k)(1 - F_{ij}(k))] = (1 - p_k)\mathsf{E}[W_{ij}(k)]\,,$$

where the last equality holds since $W_{ij}(k)$ and $F_{ij}(k)$ are independent, and $\mathsf{E}[F_{ij}(k)] = p_k$. The matrix $U(k)$ is stochastic, so $1 - \mathsf{E}\big[U_{ii(k)}\big] = \sum_{j \neq i} \mathsf{E}[U_{ij}(k)]$. By the preceding relation, we have $\sum_{j \neq i} \mathsf{E}[U_{ij}(k)] = (1 - p_k) \sum_{j \neq i} \mathsf{E}[W_{ij}(k)]$, which by stochasticity of

$W(k)$ implies $\sum_{j\neq i} \mathsf{E}[U_{ij}(k)] = (1-p_k)(1-\mathsf{E}[W_{ii}(k)])$. Therefore, $\mathsf{E}[U_{ii}(k)] = p_k + (1-p_k)\mathsf{E}[W_{ii}(k)]$, or in matrix notation:

$$\mathsf{E}[U(k)] = p_k I + (1-p_k)\mathsf{E}[W(k)]\,. \tag{16}$$

But since π is a common steady state of $\{\mathsf{E}[W(k)]\}$, it follows that $\pi^T\mathsf{E}[U(k)] = p_k\pi^T + (1-p_k)\pi^T = \pi^T$. Hence, the model $\{U(k)\}$ has the same common steady state $\pi > 0$ in expectation as the original model $\{W(k)\}$.

We next show that the failure model has feedback property. By the definition of $U(k)$, we have $U_{ii}(k) \geq W_{ii}(k)$ for all $i \in [m]$ and $k \geq 0$. Hence, $\mathsf{E}[U_{ii}(k)U_{ij}(k)] \geq \mathsf{E}[W_{ii}(k)U_{ij}(k)]$. Since $\{F(k)\}$ and $\{W(k)\}$ are independent, we have

$$\begin{aligned}\mathsf{E}[W_{ii}(k)U_{ij}(k)] &= \mathsf{E}[\mathsf{E}[W_{ii}(k)U_{ij}(k) \mid F_{ij}(k)=0]] = \mathsf{E}[\mathsf{E}[W_{ii}(k)W_{ij}(k) \mid F_{ij}(k)=0]]\\ &= (1-p_k)\mathsf{E}[W_{ii}(k)W_{ij}(k)]\,.\end{aligned}$$

A similar relation holds for $\mathsf{E}[U_{jj}(k)U_{ji}(k)]$. By the feedback property of $\{W(k)\}$, we have

$$\mathsf{E}[U_{ii}(k)U_{ij}(k) + U_{jj}(k)U_{ji}(k)] \geq (1-p_k)\gamma\mathsf{E}[W_{ij}(k) + W_{ji}(k)] = \gamma\mathsf{E}[U_{ij}(k) + U_{ji}(k)]\,,$$

where the last equality follows from Eq. (16) and $\gamma > 0$ is the feedback constant for $\{W(k)\}$. Thus, the failure model $\{U(k)\}$ has feedback property with constant γ. Hence, the model satisfies the assumptions of Infinite Flow Theorem 1. By this theorem, the model $\{U(k)\}$ is ergodic if and only if $\sum_{k=0}^{\infty} \mathsf{E}[U_S(k)] = \infty$ for any nontrivial $S \subset [m]$. By Eq. (16) we have $\mathsf{E}[U_S(k)] = (1-p_k)\mathsf{E}[W_S(k)]$. Hence, the failure model $\{U(k)\}$ is ergodic if and only if $\sum_{k=0}^{\infty}(1-p_k)\mathsf{E}[W_S(k)] = \infty$ for any nontrivial $S \subset [m]$. **Q.E.D.**

When the failure probabilities p_k are uniformly bounded away from one, i.e., $p_k \leq \bar{p}$ for all k and some $\bar{p} < 1$, we have that $\sum_{k=0}^{\infty}(1-p_k)\mathsf{E}[W_S(k)] = \infty$ if and only if $\sum_{k=0}^{\infty} \mathsf{E}[W_S(k)] = \infty$. Thus, when the failure probabilities p_k are uniformly bounded away from one, by Lemma 7 we have: the failure model $\{U(k)\}$ is ergodic if and only if the original model $\{W(k)\}$ is ergodic.

.1 Proof of Lemma 4

Proof. Define the function $V(x) = \sum_{i=1}^m \pi_i(x_i - \pi^T x)^2$. Also, let $D = \text{diag}(\pi)$, $H(k) = \mathsf{E}\big[U^T(k)U(k)\big]$ and $L(k) = \mathsf{E}\big[W^T(k)DW(k)\big]$ for $k \geq 0$. Let $x(0) \in [0,1]^m$ and $\{x(k)\}$ be the dynamics resulted from chain $\{U(k)\}$. Then we have:

$$\begin{aligned}\mathsf{E}[V(x(k+1))|x(k)] &= \mathsf{E}\big[x^T(k+1)(D-\pi\pi^T)x(k+1)|x(k)\big] \\ &= \mathsf{E}\big[(W(k)x(k)+y(k))^T(D-\pi\pi^T)(W(k)x(k)+y(k))|x(k)\big]\\ &\leq \mathsf{E}[V(W(k)x(k))|x(k)] + 2\mathsf{E}\big[x(k+1)^T(D-\pi\pi^T)y(k)|x(k)\big]\,,\end{aligned} \tag{17}$$

where $y(k) = (U(k)-W(k))x(k)$ and the last inequality follows by positive semidefiniteness of $D - \pi\pi^T$. Based on Theorem 4 in [31], we have $\mathsf{E}[V(W(k)x(k))|x(k)] \leq$

$V(x(k)) - \sum_{i<j} L_{ij}(k)(x_i(k) - x_j(k))^2$. Thus:

$$\mathsf{E}[V(x(k+1))|x(k)] \leq V(x(k)) - \sum_{i<j} L_{ij}(k)(x_i(k) - x_j(k))^2 \\ + 2\mathsf{E}\left[x(k+1)^T(D - \pi\pi^T)y(k)|x(k)\right], \qquad (18)$$

On the other hand, since $U(k)$s are stochastic, $x(k) \in [0,1]^m$. Hence, $\|y(k)\|_\infty \leq \|W(k) - U(k)\|_\infty$ and thus we have $-\|W(k) - U(k)\|_\infty \leq y_i(k) \leq \|W(k) - U(k)\|_\infty$. Therefore, since the summation of the absolute value of the ith row of $D - \pi\pi^T$ is $\pi_i(2 - \pi_i) \leq 1$, we have $\|(D - \pi\pi^T)y(k)\|_\infty \leq \|W(k) - U(k)\|_\infty$ and hence, because $x(k+1) \in [0,1]^m$, $x(k+1)^T(D - \pi\pi^T)y(k) \leq m\|W(k) - U(k)\|_\infty$. But since, $U(k)$ and $W(k)$ are independent of $x(k)$, from the above discussion and Eq. (17) and Eq. (18), it follows that

$$\mathsf{E}[V(x(k+1))|x(k)] \leq V(x(k)) - \sum_{i<j} L_{ij}(k)(x_i(k) - x_j(k))^2 + 2m\mathsf{E}[\|W(k) - U(k)\|_\infty]. \qquad (19)$$

But $\{U(k)\}$ is an ℓ_1-approximation of $\{W(k)\}$, and hence $\sum_{k=0}^{\infty} \mathsf{E}[\|W(k) - U(k)\|_\infty] < \infty$. Therefore, by the Robbins-Siegmund theorem ([23] page 164), $\sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)(x_i(k) - x_j(k))^2 < \infty$ almost surely.

The last step is to show that the difference between the two sums $\sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)(x_i(k) - x_j(k))^2$ and $\sum_{k=0}^{\infty} \sum_{i<j} H_{ij}(k)(x_i(k) - x_j(k))^2$ is finite. Since $|W_{\ell i}(k) - U_{\ell i}(k)| \leq \|W(k) - U(k)\|_\infty$ for any $i, \ell \in [m]$ and based on the definition of $H(k)$, we have

$$\begin{aligned}
\pi_{\min} H_{ij}(k) &\leq \sum_{\ell=1}^{m} \mathsf{E}[\pi_\ell U_{\ell i}(k) U_{\ell j}(k)] \\
&= \sum_{\ell=1}^{m} \mathsf{E}[\pi_\ell(W_{\ell i}(k) + [W(k) - U(k)]_{\ell i})(W_{\ell j}(k) + [W(k) - U(k)]_{\ell j})] \\
&= L_{ij}(k) + \sum_{\ell=1}^{m} \mathsf{E}[\pi_\ell U_{\ell i}(k)[W(k) - U(k)]_{\ell j}] + \sum_{\ell=1}^{m} \mathsf{E}[\pi_\ell [W(k) - U(k)]_{\ell i} U_{\ell j}(k)] \\
&\leq L_{ij}(k) + \mathsf{E}\left[\|U(k) - W(k)\|_\infty \sum_{\ell=1}^{m} \pi_\ell(U_{\ell i}(k) + U_{\ell j}(k))\right] \\
&\leq L_{ij}(k) + \mathsf{E}[\|U(k) - W(k)\|_\infty],
\end{aligned}$$

where $\pi_{\min} = \min_{i \in [m]} \pi_i > 0$ and the last inequality follows from the fact that $U(k)$ is a stochastic matrix and π is a stochastic vector. Therefore,

$$\begin{aligned}
&\pi_{\min} \sum_{k=0}^{\infty} \sum_{i<j} H_{ij}(k)(x_i(k) - x_j(k))^2 \\
&\qquad \leq \sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)(x_i(k) - x_j(k))^2 + \sum_{k=0}^{\infty} \sum_{i<j} \mathsf{E}[\|U(k) - W(k)\|_\infty] (x_i(k) - x_j(k))^2 \\
&\qquad \leq \sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)(x_i(k) - x_j(k))^2 + m^2 \sum_{k=0}^{\infty} \mathsf{E}[\|U(k) - W(k)\|_\infty],
\end{aligned}$$

where the last inequality holds since $(x_i(k)-x_j(k))^2 \le 1$. But again, since $\sum_{k=0}^{\infty} \mathsf{E}[\|U(k) - W(k)\|_\infty] < \infty$, and we have shown that $\sum_{k=0}^{\infty} \sum_{i<j} L_{ij}(k)(x_i(k)-x_j(k))^2 < \infty$ almost surely, hence, $\sum_{k=0}^{\infty} \sum_{i<j} H_{ij}(k)(x_i(k) - x_j(k))^2 < \infty$ almost surely. **Q.E.D.**

.2 Proof of Theorem 2

Proof. Let $d_k = z_{\max}(k) - z_{\min}(k)$ and $d_\infty = \lim_{k\to\infty}(z_{\max}(k) - z_{\min}(k)) > 0$. For $t \ge 0$, let σ (which is a function of t) be a permutation on $[m]$ such that $z_{\sigma_1}(t) \le \cdots \le z_{\sigma_m}(t)$. Since the matrices $\{A(k)\}$ are stochastic, the sequence $\{d(k)\}$ is non-increasing and hence,

$$d_t = z_{\sigma_m}(t) - z_{\sigma_1}(t) = \sum_{i=1}^{m-1}(z_{\sigma_{i+1}}(t) - z_{\sigma_i}(t)) \ge d_\infty.$$

Therefore, there exists an index i^* such that $z_{\sigma_{i^*+1}}(t) - z_{\sigma_{i^*}}(t) \ge \frac{1}{m-1} d_\infty$. Let $S = \{\sigma_i \mid i \le i^*\}$, so that $d_S(t) \ge \frac{1}{m-1} d_\infty$. Now, let $t' = \arg\min_{s>t} \sum_{k=t}^{s-1} A_S(k) \ge \frac{1}{3} d_S(t)$. Then, by Lemma 1 in [31], for $k \in [t, t']$, we have $d_S(k) \ge d_S(t) - 2d_t \sum_{k=t}^{t'} A_S(k) \ge \frac{1}{3}\, d_S(t)$. Since $d_S(t) \ge \frac{1}{m-1}\, d_\infty$, we have $d_S(s) \ge \frac{1}{3(m-1)}\, d_\infty$. Therefore,

$$\begin{aligned}\sum_{k=t}^{t'-1}\sum_{i<j}(A_{ij}(k) + A_{ji}(k))(z_i(k) - z_j(k))^2 &\ge \sum_{k=t}^{t'-1}\sum_{i\in S, j\in\bar{S}}(A_{ij}(k) + A_{ji}(k))(z_i(k) - z_j(k))^2 \\ \ge \sum_{k=t}^{t'-1}\sum_{i\in S, j\in\bar{S}}(A_{ij}(k) + A_{ji}(k)) d_S^2(k) &= \sum_{k=t}^{t'-1} d_S^2(k) \sum_{i\in S, j\in\bar{S}}(A_{ij}(k) + A_{ji}(k)),\end{aligned} \tag{20}$$

where the last inequality holds since $(z_j(k) - z_i(k)) \ge d_S(k)$ for $i \in S$ and $j \in \bar{S}$. Also note that by the definition $A_S(k) = \sum_{i\in S, j\in\bar{S}}(A_{ij}(k) + A_{ji}(k))$. Therefore, we have

$$\begin{aligned}\sum_{k=t}^{t'-1}\sum_{i<j}(A_{ij}(k) + A_{ji}(k))(z_i(k) - z_j(k))^2 &\ge \sum_{k=t}^{t'-1} A_S(k)\, d_S^2(k) \\ \ge \left(\frac{1}{3(m-1)}\, d_\infty\right)^2 \sum_{k=t}^{t'-1} A_S(k) &\ge \left(\frac{1}{3(m-1)}\, d_\infty\right)^3 = \frac{1}{27(m-1)^3}\, d_\infty^3.\end{aligned} \tag{21}$$

The preceding argument holds for arbitrary $t \ge 0$. Hence, by letting $t_0 = 0$ and recursively defining $t_s = t'_{s-1}$ for $s > 0$, we have

$$\begin{aligned}&\sum_{k=0}^{\infty}\sum_{i<j}(A_{ij}(k) + A_{ji}(k))(z_i(k) - z_j(k))^2 \\ &\quad= \sum_{s=0}^{\infty}\sum_{k=t_s}^{t_{s+1}-1}\sum_{i<j}(A_{ij}(k) + A_{ji}(k))(z_i(k) - z_j(k))^2 \\ &\quad\ge \sum_{s=0}^{\infty}\frac{1}{27(m-1)^3}\, d_\infty^3 = \infty,\end{aligned}$$

where the last inequality follows from Eqs. (20) and (21). **Q.E.D.**

.3 Proof of Lemma 5

Proof. First we show that $\tilde{W}(k)$ is a stochastic matrix for any $k \geq 0$. To do this, in view of the diagonal structure of the matrix $\tilde{W}(k)$ (Eq. (8)), it suffices to show that $W^{(r)}(k)$ is stochastic for $1 \leq r \leq \tau$. Note that based on the definition of $W^{(r)}(k)$, we have $W^{(r)}(k) \geq 0$. Also for any $i \in [m_r]$, we have:

$$\begin{aligned}\sum_{j=1}^{m_r} W_{ij}^{(r)}(k) &= W_{ii}^{(r)}(k) + \sum_{j \neq i, j \in [m_r]} W_{ij}^{(r)}(k) \\ &= W_{(i+a_{r-1})(i+a_{r-1})}(k) + \sum_{\ell \in \bar{S}_v} W_{(i+a_{r-1})\ell}(k) + \sum_{\ell \neq i+a_{r-1}, \ell \in S_r} W_{(i+a_{r-1})\ell}(k) \\ &= \sum_{\ell=1}^{m} W_{(i+a_r)\ell}(k) = 1. \end{aligned} \tag{22}$$

Now, let $i \in S_r$ for $1 \leq r \leq \tau$. Then, for any $j \neq i$, we have two cases:

(i) If $j \in S_r$, by the definition of $\tilde{W}(k)$, we have $W_{ij}(k) = \tilde{W}_{ij}(k)$ and hence, $|W_{ij}(k) - \tilde{W}_{ij}(k)| = 0$.

(ii) If $j \notin S_r$, we have $\tilde{W}_{ij}(k) = 0$ and hence, $|W_{ij}(k) - \tilde{W}_{ij}(k)| = W_{ij}(k)$.

Finally, for $j = i$, we have $\tilde{W}_{ii}(k) = W_{ii}(k) + \sum_{j \notin S_r} W_{ij}(k)$. Hence, $|W_{ij}(k) - \tilde{W}_{ij}(k)| = \sum_{j \notin S_r} W_{ij}(k)$. Therefore, we have:

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

By summing up over all $i \in S_r$, we have:

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

Again, by summing up the two sides of the preceding inequality for $r = 1, \ldots, \tau$, we have:

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

But $\sum_{r=1}^{\tau} \sum_{i \in S_r} \sum_{j=1}^{m} |W_{ij}(k) - \tilde{W}_{ij}(k)| = \sum_{i,j \in [m]} |W_{ij}(k) - \tilde{W}_{ij}(k)|$. Also note that $S_1, \ldots, S_\tau$ are the sets of vertices of each connected components of G^∞ and hence, we obtain $\sum_{k=0}^{\infty} \sum_{r=1}^{\tau} W_{S_r}(k) < \infty$ almost surely. Therefore, by combining the above facts:

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

This proves that $\{\tilde{W}(k)\}$ is in fact an ℓ_1-approximation of $\{W(k)\}$. To prove that $\{W^{(r)}(k)\}$ has infinite flow property, let $V \subset S_r$. Then since S_r is the set of vertices of one of the connected components of G^∞, there exists an edge $\{i,j\} \in \mathcal{E}^\infty$ where $i \in V$ and $j \in \bar{V}$. But based on the definition of $W^{(r)}(k)$, for $i_r = i - a_{r-1}$ and $j_r = j - a_{r-1}$, we have $W^{(r)}_{i_r j_r}(k) + W^{(r)}_{j_r i_r}(k) = W_{ij}(k) + W_{ji}(k)$. Since $\{i,j\} \in \mathcal{E}^\infty$, it follows

$$\sum_{k=0}^{\infty}(W^{(r)}_{i_r j_r}(k) + W^{(r)}_{j_r i_r}(k)) = \sum_{k=0}^{\infty}(W_{ij} + W_{ji}) = \infty.$$

Therefore, the infinite graph of $\{W^{(r)}(k)\}$ is connected and, hence, $\{W^{(r)}(k)\}$ has infinite flow property. **Q.E.D.**

References

[1] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, *Bayesian learning in social networks*, MIT LIDS Working Paper 2780, 2010.

[2] T.C. Aysal, M.E. Yildiz, A.D. Sarwate, and A. Scaglione, *Broadcast gossip algorithms: Design and analysis for consensus*, Proceedings of the 47th IEEE Conference on Decision and Control, 2008, pp. 4843–4848.

[3] T.C. Aysal, M.E. Yildriz, A.D. Sarwate, and A. Scaglione, *Broadcast gossip algorithms for consensus*, IEEE Transactions on Signal processing **57** (2009), 2748–2761.

[4] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, *Gossip algorithms: Design, analysis, and applications*, Proceedings IEEE Infocom, 2005, pp. 1653–1664.

[5] ———, *Randomized gossip algorithms*, IEEE Transactions on Information Theory **52** (2006), no. 6, 2508–2530.

[6] M. Cao, D.A. Spielman, and A.S. Morse, *A lower bound on convergence of a distributed network consensus algorithm*, Proceedings of IEEE CDC, 2005.

[7] R. Carli, F. Fagnani, P. Frasca, T. Taylor, and S. Zampieri, *Average consensus on networks with transmission noise or quantization*, Proceedings of IEEE American Control Conference, 2007, pp. 4189–4194.

[8] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri, *Communication constraints in the average consensus problem*, Automatica **44** (2008), no. 3, 671–684.

[9] S. Chatterjee and E. Seneta, *Towards consensus: Some convergence theorems on repeated averaging*, Journal of Applied Probability **14** (1977), no. 1, 89–97.

[10] R. Cogburn, *On products of random stochastic matrices*, In Random matrices and their applications (1986), 199–213.

[11] R. Durrett, *Probability: Theory and examples*, third ed., Curt Hinrichs, 2005.

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

[13] G. B. Folland, *Real analysis : modern techniques and their applications*, 2 ed., Wiley, 1999.

[14] B. Golub and M.O. Jackson, *Naive learning in social networks: Convergence, influence and the wisdom of crowds*, The American Economic Journal: Microecnomics **2** (2010), no. 1), 112–149.

[15] 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.

[16] B. Johansson, *On distributed optimization in networked systems*, Ph.D. thesis, Royal Institute of Technology, Stockholm, Sweden, 2008.

[17] K. Nawrotzki, *Discrete open systems on markov chains in a random environment. I*, Elektronische Informationsverarbeitung und Kybernetik **17** (1981), 569–599.

[18] ———, *Discrete open systems on markov chains in a random environment. II*, Elektronische Informationsverarbeitung und Kybernetik **18** (1982), 83–98.

[19] 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.

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

[21] ———, *Asynchronous consensus with delays*, Journal of Global Optimization **47** (2010), no. 3, 437–456.

[22] R. Olfati-Saber and R.M. Murray, *Consensus problems in networks of agents with switching topology and time-delays*, IEEE Transactions on Automatic Control **49** (2004), no. 9, 1520–1533.

[23] Alexander S. Poznyak, *Advanced mathematical tools for automatic control engineers: Stochastic techniques*, Elsevier, 2009.

[24] P. Frasca R. Carli, F. Fagnani and S. Zampieri, *Gossip consensus algorithms via quantized communication*, Automatica **46** (2010), no. 1, 70–80.

[25] S. Sundhar Ram, A. Nedić, and V.V. Veeravalli, *Distributed stochastic subgradient algorithm for convex optimization*, To appear in JOTA, Available at http://arxiv.org/abs/0811.2595, 2008.

[26] W. Ren and R.W. Beard, *Consensus seeking in multi-agent systems under dynamically changing interaction topologies*, IEEE Transactions on Automatic Control **50** (2005), no. 5, 655–661.

[27] M. Rosenblatt, *Products of independent identically distributed stochastic matrices*, Journal of Journal of Mathematical Analysis and Applications **11** (1965), no. 1, 1–10.

[28] E. Seneta, *Non-negative Matrices and Markov Chains*, 3 ed., Springer Series in Statistics, 2006.

[29] 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.

[30] ———, *Consensus over ergodic stationary graph processes*, IEEE Transactions on Automatic Control **55** (2010), no. 1, 225–230.

[31] B. Touri and A. Nedić, *On ergodicity, infinite flow and consensus in random models*, `http://arxiv.org/PS_cache/arxiv/pdf/1001/1001.1890v3.pdf`, 2010.

[32] ———, *When infinite flow is sufficient for ergodicity*, `https://netfiles.uiuc.edu/touri1/www/Papers/ergodicityCDC.pdf`, 2010.

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

[34] 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.