

# On Averaging Dynamics in General State Spaces

Behrouz Touri, Tamer Başar, and Angelia Nedić

**Abstract**—In this paper, we present a framework for studying distributed averaging dynamics over general state spaces. We define several modes of ergodicity and consensus for such dynamics and show that, unlike for a finite dimensional space, these modes are not equivalent. Motivated by the role of the infinite flow property in ergodicity in finite dimensional spaces, we define the infinite flow property for averaging dynamics in general state spaces. We show that this property is a necessary condition for the weakest form of ergodicity. Also, we characterize a class of quadratic Lyapunov comparison functions for certain averaging dynamics and provide a relation capturing the decrease of these functions along the trajectories of the dynamics.

## I. INTRODUCTION

Ergodicity of a sequence of stochastic matrices is one of the fundamental concepts in the study of various phenomena in distributed averaging dynamics including consensus and stability of the dynamics. As discussed in [1], in finite-dimensional spaces, one can define weak and strong ergodicity for the left products of stochastic matrices which, however, are the same.

In this paper, we study ergodicity and consensus for averaging dynamics in general state spaces, where the averaging dynamics is driven by a sequence of stochastic kernels. We introduce several modes of ergodicity and consensus, and we show that these modes are not equivalent to each other. Inspired by the infinite flow property [6], we define an analogous property for dynamics over general state spaces and prove that this property is necessary for the weakest form of ergodicity. Finally, we show that the averaging dynamics admits a quadratic Lyapunov function under some conditions on the sequence of stochastic kernels. The development in this paper is restricted to the setting where the sequences of stochastic kernels are deterministic. The work is motivated by the theory of Markov chains over general state spaces [4].

The structure of this paper is as follows: In Section II, we introduce several modes of ergodicity and consensus for averaging dynamics over general state spaces. In Section III, we discuss the generalization of the infinite flow property for a general state space. In Section IV, we characterize quadratic Lyapunov functions for a class of averaging dynamics and provide a fundamental relation for the decrease of such functions. We conclude the paper with Section V.

## II. FRAMEWORK

In this section, we introduce averaging dynamics together with several notions of ergodicity and consensus for general

state spaces. Let us start by revisiting such dynamics in a finite-dimensional space. Averaging dynamics in  $\mathbb{R}^m$  is given by

$$x(k+1) = A(k)x(k) \quad \text{for } k \geq t_0, \quad (1)$$

where  $t_0 \geq 0$  is an initial time and  $x(t_0) \in \mathbb{R}^m$  is an initial state of the dynamics. The sequence  $\{A(k)\}$  consists of row-stochastic matrices i.e., each  $A(k)$  has non-negative entries and the sum of the entries in each row of  $A(k)$  is equal to 1.

Studying the limiting behavior of the dynamics (1) is closely related to studying the convergence properties of the matrix products

$$A(k : t_0) = A(k-1) \cdots A(t_0+1)A(t_0) \quad \text{for } k > t_0.$$

This leads to our operator theoretic viewpoint to averaging dynamics over general state spaces. Note that to show that  $\lim_{k \rightarrow \infty} A(k : t_0)$  exists, it suffices to verify the convergence of  $\{x(k)\}$  only for all  $x(t_0) = e_\ell$  with  $\ell \in [m]$ , where  $e_\ell$  is a vector with all zero entries except for the  $\ell$ th entry which is equal to 1. (The preceding condition simply means that each column of  $A(k : t_0)$  converges.) Our development builds on a counterpart of such a result in general state spaces.

Before formulating averaging dynamics over general state spaces, let us discuss the notation that we use throughout this paper. Instead of sequences of stochastic matrices, we will deal with sequences of stochastic kernels in a measure space, also referred to as chains. We use the subscripts for indexing time variables; for example,  $\{\mathcal{K}_k\}$  denotes a sequence of kernels indexed by time variable  $k$ . For a kernel  $\mathcal{K}$ , we use  $\mathcal{K}(\xi, \eta)$  to denote the value of the kernel  $\mathcal{K}$  at the point  $(\xi, \eta)$ . For averaging dynamics in  $\mathbb{R}^m$ , however, we use  $\{A(k)\}$  to denote the sequence of matrices. We use  $\mathbb{R}_+$  to denote the set of non-negative scalars and  $Z_+$  to denote the set of non-negative integers.

To describe an averaging dynamics in a general state space, let  $X$  be a set and  $\mathcal{M}$  be a  $\sigma$ -algebra of subsets of  $X$ . Throughout the paper, the measurable space  $(X, \mathcal{M})$  serves as the underlying state space for the averaging dynamics.

*Definition 1:* [4] We say that a function  $\mathcal{K} : X \times \mathcal{M} \rightarrow \mathbb{R}_+$  is a *stochastic kernel* if

- (a) the function  $f_S : X \rightarrow \mathbb{R}^+$  defined by  $f_S(\xi) = \mathcal{K}(\xi, S)$  is measurable, for every  $S \in \mathcal{M}$ ;
- (b) the function  $\mathcal{K}(\xi, \cdot)$  (defined on sets  $S \in \mathcal{M}$ ) is a measure on  $X$ , for every  $\xi \in X$ ;
- (c)  $\mathcal{K}(\xi, X) = 1$ , for every  $\xi \in X$ , i.e., the measure  $\mathcal{K}(\xi, \cdot)$  is a probability measure for every  $\xi \in X$ .

Furthermore, if we have

$$\mathcal{K}(\xi, S) = \int_S \tilde{\mathcal{K}}(\xi, \eta) d\mu(\eta)$$

Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801, Email: {touri1,basar1,angelia}@illinois.edu. This work has been partially supported by grants NSF CMMI 07-42538 and NSF CCF 11-11342.

for a measurable function  $\tilde{\mathcal{K}} : X \times X \rightarrow \mathbb{R}_+$  and a measure  $\mu$  on  $(X, \mathcal{M})$ , then we refer to  $\mathcal{K}$  as a *stochastic integral kernel* with density  $\tilde{\mathcal{K}}$  and basis  $\mu$ .

Let  $L_\infty$  be the space of all measurable functions  $x$  from  $(X, \mathcal{M})$  to  $\mathbb{R}$  such that  $\sup_{\xi \in X} |x(\xi)| < \infty$ , and let

$$\|x\|_\infty = \sup_{\xi \in X} |x(\xi)|.$$

To define an averaging dynamics, let  $\{\mathcal{K}_k\}$  be a chain of stochastic kernels. Let  $t_0 \geq 0$  be a given starting time and  $x_{t_0} \in L_\infty$  be a starting point. Consider the averaging dynamics given as follows:

$$x_{k+1} = \mathcal{K}_k x_k \quad \text{for all } \xi \in X \text{ and } k \geq t_0, \quad (2)$$

where

$$x_{k+1}(\xi) = \int_X \mathcal{K}_k(\xi, d\eta) x_k(\eta)$$

(the integral is for the function  $x_k$  with respect to the measure  $\mathcal{K}(\xi, \cdot)$ ). When  $\mathcal{K}_k$  in (2) is an integral kernel with density  $\tilde{\mathcal{K}}_k$  and a basis  $\mu$ , the dynamics in (2) can be written as

$$x_{k+1}(\xi) = \int_X \tilde{\mathcal{K}}_k(\xi, \eta) x_k(\eta) d\mu(\eta). \quad (3)$$

We now show that the dynamics (2) is well-defined in the sense that  $\{x_k\} \subset L_\infty$  whenever the dynamics is started at a point  $x_{t_0} \in L_\infty$ .

*Lemma 1:* Let  $\{x_k\}$  be generated by dynamics (2) started at a time  $t_0 \geq 0$  and a point  $x_{t_0} \in L_\infty$ . Then,  $\|x_{k+1}\|_\infty \leq \|x_k\|_\infty$  for all  $k \geq 0$  (so  $\{x_k\} \subset L_\infty$ ).

*Proof:* Since  $x_{t_0} \in L_\infty$ , by using the induction on  $k \geq t_0$ , we can see that for any  $\xi \in X$ ,

$$|x_{k+1}(\xi)| \leq \int_X |\mathcal{K}_k(\xi, d\eta) x_k| \leq \|x_k\|_\infty,$$

where the last inequality holds since  $\mathcal{K}_k(\xi, \cdot)$  is a probability measure. Therefore,  $|x_{k+1}(\xi)| \leq \|x_{t_0}\|_\infty$  for all  $\xi \in X$ , implying  $x_{k+1} \in L_\infty$ ; hence  $x_k \in L_\infty$  for all  $k \geq t_0$ . ■

*Example 1:* Let  $X = [m] = \{1, \dots, m\}$ ,  $\mathcal{M} = \mathcal{P}([m])$  (the set of all subsets of  $[m]$ ) and let  $\mu$  be the counting measure on  $(X, \mathcal{M})$ , i.e.,  $\mu(S) = |S|$  for all  $S \subseteq [m]$ . For a chain of stochastic matrices  $\{A(k)\}$ , if we define  $\tilde{\mathcal{K}}_k(i, j) = A_{ij}(k)$  for all  $i, j \in [m]$  and  $k \geq 0$ , then  $\{\tilde{\mathcal{K}}_k\}$  is a chain of density functions with basis  $\mu$ . Any vector  $x_{t_0} \in \mathbb{R}^m$  is in  $L_\infty$  and, hence, the finite-dimensional averaging dynamics (1) is a special case of the averaging dynamics (3).

For a probability measure  $\pi$  on  $(X, \mathcal{M})$ , let us denote the stochastic kernel  $\mathcal{K}(\xi, S) = \pi(S)$  by  $\mathcal{K} = \mathbf{1}_X \pi^T$  where  $\mathbf{1}_S$  is the characteristic function of a set  $S \in \mathcal{M}$ . Also, let  $\mathcal{K}_{k:t_0}$  be the transfer function of the dynamics (2) from time  $t_0$  to time  $k > t_0$ , given by

$$\mathcal{K}_{k:t_0}(\xi, \eta) = \int \mathcal{K}_{k-1}(\xi, d\psi) \mathcal{K}_{k-1:t_0}(\psi, \eta) \quad \text{for all } k > t_0,$$

where  $\mathcal{K}_{t_0+1:t_0} = \mathcal{K}_{t_0}$ .

*Definition 2:* Let  $\{\mathcal{K}_k\}$  be a chain of stochastic kernels on  $(X, \mathcal{M})$ . We say that  $\{\mathcal{K}_k\}$  is:

*Uniformly Ergodic* if  $\lim_{k \rightarrow \infty} \|\mathcal{K}_{k:t_0} - \mathbf{1}_X \pi^T\|_\infty = 0$  for

some probability measure  $\pi_{t_0}$  on  $(X, \mathcal{M})$ , and for all  $t_0 \geq 0$  and  $x_{t_0} \in L_\infty$  (the equality is understood as the induced operator norm);

*Strongly Ergodic* if  $\lim_{k \rightarrow \infty} x_k(\xi) = c(x_{t_0}, t_0)$  for some  $c(x_{t_0}, t_0) \in \mathbb{R}$ , and for all  $\xi \in X$ ,  $t_0 \geq 0$ , and  $x_{t_0} \in L_\infty$ ;

*Weakly Ergodic* if  $\lim_{k \rightarrow \infty} (x_k(\xi) - x_k(\eta)) = 0$  for all  $\xi, \eta \in X$ ,  $t_0 \geq 0$ , and  $x_{t_0} \in L_\infty$ .

Letting  $t_0 = 0$  in Definition 2, we obtain several modes of consensus.

*Definition 3:* We say that  $\{\mathcal{K}_k\}$  admits consensus with the following qualifiers:

*Uniformly* if  $\lim_{k \rightarrow \infty} \|\mathcal{K}_{k:0} - \mathbf{1}_X \pi^T\|_\infty = 0$  for some probability measure  $\pi$  on  $(X, \mathcal{M})$ ;

*Strongly* if  $\lim_{k \rightarrow \infty} x_k(\xi) = c(x_0)$  for some  $c(x_0) \in \mathbb{R}$ , and for all  $\xi \in X$  and  $x_0 \in L_\infty$ ;

*Weakly* if  $\lim_{k \rightarrow \infty} (x_k(\xi) - x_k(\eta)) = 0$  for all  $\xi, \eta \in X$  and  $x_0 \in L_\infty$ .

For a finite-dimensional state space, all the notions of ergodicity and consensus of Definition 2 and Definition 3, respectively, are equivalent. However, in general state spaces they lead to different properties. Specifically, from Definition 2 it can be seen that uniform ergodicity implies strong ergodicity, which in turn implies weak ergodicity. Similar relations hold for the different modes of consensus. However, the reverse implications do not necessarily hold. The following example shows that strong ergodicity (consensus) does not imply uniform ergodicity (consensus).

*Example 2:* Consider the set  $\mathbb{Z}_+$  of non-negative integers. Let  $\mathcal{M} = \mathcal{P}(\mathbb{Z}_+)$  and  $\mu$  be the counting measure on  $\mathbb{Z}_+$ . Let  $\{\mathcal{K}_k\}$  be the chain of stochastic kernels with density functions given by:

$$\tilde{\mathcal{K}}_k(i, j) = \begin{cases} \delta_{0j} & \text{if } i \leq k, \\ \delta_{ij} & \text{if } i > k, \end{cases} \quad \text{for } k > 1,$$

with  $\delta_{ij} = 1$  if  $i = j$  and  $\delta_{ij} = 0$  otherwise. For  $k = 0$ , define  $\tilde{\mathcal{K}}_0(i, j) = \delta_{ij}$ . Then,  $\tilde{\mathcal{K}}_k$  can be viewed as an infinite stochastic matrix of the form:

$$\tilde{\mathcal{K}}_k = \begin{bmatrix} 1 & \cdots & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots \\ 1 & \cdots & 0 & 0 & 0 & 0 & \cdots \\ 1 & \cdots & 0 & 0 & 0 & 0 & \cdots \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots \\ 0 & \cdots & 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (4)$$

Now, let  $\{x_k\}$  be the dynamics generated by  $\{\mathcal{K}_k\}$  started at an arbitrary initial pair  $(t_0, x_{t_0}) \in \mathbb{Z}_+ \times L_\infty$ . Then, for any  $\xi \in \mathbb{Z}_+$ , by letting  $k > \max(t_0, \xi)$ , we have  $x_k(\xi) = x_{t_0}(\xi)$ , implying  $\lim_{k \rightarrow \infty} x_k(\xi) = x_{t_0}(\xi)$ . Therefore,  $\{\mathcal{K}_k\}$  is strongly ergodic and admits consensus strongly.

However, in the light of Eq. (4), the only candidate for the limiting kernel  $\lim_{k \rightarrow \infty} \mathcal{K}_{k:t_0}$  is the integral kernel  $\mathbf{1}_X \pi^T$ , where  $\pi$  is the probability measure concentrated at  $\{0\}$ , i.e.,  $\pi(S) = 1$  if  $0 \in S$  and  $\pi(S) = 0$  otherwise. However, letting  $x_{t_0}(i) = \delta_{0i}$ , we find that  $\|(\mathcal{K}_k - \mathbf{1}_X \pi^T) x_{t_0}\|_\infty = 1$ . Hence,

$\{\mathcal{K}_k\}$  is not uniformly ergodic and does not admit consensus uniformly.

The following example shows that weak ergodicity (consensus) does not imply strong ergodicity (consensus).

*Example 3:* Consider the measure space  $(\mathbb{Z}_+, \mathcal{M})$  of Example 2 and the chain of stochastic integral kernels  $\{\mathcal{K}_k\}$  with density kernels given by:

$$\tilde{\mathcal{K}}_k(i, j) = \begin{cases} \delta_{jk} & \text{if } i \leq k, \\ \delta_{ij} & \text{if } i > k, \end{cases} \quad \text{for } k \geq 0.$$

The density kernel  $\tilde{\mathcal{K}}_k$  has the following form

$$\tilde{\mathcal{K}}_k = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & 0 & \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots \\ 0 & \cdots & 0 & 1 & 0 & 0 & \cdots \\ 0 & \cdots & 0 & 1 & 0 & 0 & \cdots \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots \\ 0 & \cdots & 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

For any two  $\xi, \eta \in \mathbb{Z}_+$  and any initial pair  $(t_0, x_{t_0}) \in \mathbb{Z}_+ \times L_\infty$ , we have  $x_k(\xi) = x_k(\eta) = x_k(k)$  for  $k \geq K = \max(\xi, \eta)$ , implying  $\lim_{k \rightarrow \infty} (x_k(\xi) - x_k(\eta)) = 0$ . Nevertheless, if we let  $\{\alpha_k\}$  be a sequence of scalars in  $[0, 1]$  that is not convergent, and we start the dynamics at time  $t = 0$  with the starting point  $x_0 = (\alpha_0, \alpha_1, \alpha_2, \dots)$  (which belongs to  $L_\infty$ ), then we can see that  $x_k = (\alpha_k, \alpha_k, \dots, \alpha_k, \alpha_{k+1}, \alpha_{k+2}, \dots)$ . Therefore, for any  $\xi \in \mathbb{Z}_+$ , we have  $x_k(\xi) = \alpha_k$  for  $k > \xi$ . Since  $\{\alpha_k\}$  is not convergent, we conclude that  $\lim_{k \rightarrow \infty} x_k(\xi)$  does not exist. Thus, the chain  $\{\mathcal{K}_k\}$  is not strongly ergodic and does not admit consensus uniformly.

### III. INFINITE FLOW PROPERTY

The infinite flow property has been defined for a chain of stochastic matrices in [6], where the necessity of this property for ergodicity in finite-dimensional averaging dynamics has been shown (see Theorem 1 in [6]). In this section, we show an analogous result for averaging dynamics in a general state space.

For a stochastic chain  $\{A(k)\}$ , the infinite flow property requires that  $\sum_{k=0}^{\infty} A_S(k) = \infty$  for every non-trivial set  $S \subset [m]$ , where

$$\begin{aligned} A_S(k) &= A_{S\bar{S}}(k) + A_{\bar{S}S}(k) \\ &= \sum_{i \in S, j \in \bar{S}} A_{ij}(k) + \sum_{i \in \bar{S}, j \in S} A_{ij}(k), \end{aligned} \quad (5)$$

with  $\bar{S}$  being the complement of  $S$  with respect to  $[m]$  ( $\bar{S} = [m] \setminus S$ ). In a general state space  $(X, \mathcal{M})$ , the analogous quantity for a stochastic kernel  $\mathcal{K}$ , denoted by  $\mathcal{K}(S, \bar{S})$ , is not well-defined when the space is not equipped with a measure. However, for an integral kernel  $\mathcal{K}$ , with density  $\tilde{\mathcal{K}}$  and a basis  $\mu$ , we can define

$$\mathcal{K}(S, \bar{S}) = \int_S \mathcal{K}(\xi, \bar{S}) d\mu(\xi) = \int_S \int_{\bar{S}} \tilde{\mathcal{K}}(\xi, \eta) d\mu(\eta) d\mu(\xi).$$

Thus, for a chain of stochastic integral kernels  $\{\mathcal{K}_k\}$  with density kernels  $\{\tilde{\mathcal{K}}_k\}$  and basis  $\mu$ , it is tempting to define the flow over a non-trivial set  $S \in \mathcal{M}$  (i.e.,  $S \neq \emptyset$  and  $S \neq X$  almost everywhere) to be

$$\begin{aligned} \bar{\mathcal{K}}^f(S) &= \mathcal{K}(S, \bar{S}) + \mathcal{K}(\bar{S}, S) \\ &= \int_S \int_{\bar{S}} \tilde{\mathcal{K}}(\xi, \eta) d\mu(\xi) d\mu(\eta) + \int_{\bar{S}} \int_S \tilde{\mathcal{K}}(\xi, \eta) d\mu(\xi) d\mu(\eta), \end{aligned} \quad (6)$$

and conclude the necessity of the infinite flow for (at least) uniform ergodicity of any chain of stochastic integral kernels. However, with the definition of flow in Eq. (6), such a result is not true, as seen from the following example.

*Example 4:* Let  $X = [0, 1]$  with  $\mathcal{M}$  consisting of the Borel sets in  $[0, 1]$  and  $\mu$  be the Borel-measure on  $[0, 1]$ . Let the densities  $\tilde{\mathcal{K}}_k$  be defined as follows: for  $k \geq 0$ ,

$$\tilde{\mathcal{K}}_k(\xi, \eta) = \begin{cases} 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\eta) & \xi \in [0, 2^{-k}], \\ 2^k \cdot \mathbf{1}_{[0, 2^{-k}]}(\eta) & \xi \in (2^{-k}, \frac{1}{2}], \\ 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\eta) & \xi \in (\frac{1}{2}, 1]. \end{cases} \quad (7)$$

First, we prove that  $\{\mathcal{K}_k\}$  is uniformly ergodic by showing that  $\tilde{\mathcal{K}}_{k+1} \tilde{\mathcal{K}}_k(\xi, \eta) = 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\eta)$  for any  $k \geq 0$  and  $\xi \in [0, 1]$ . We show this by considering the following cases: (i)  $\xi \in [0, 2^{-(k+1)}]$ : In this case, by Eq. (7), we have  $\tilde{\mathcal{K}}_{k+1}(\xi, \psi) = 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\psi)$ , implying

$$\begin{aligned} \tilde{\mathcal{K}}_{k+1} \tilde{\mathcal{K}}_k(\xi, \eta) &= \int_{[0, 1]} \tilde{\mathcal{K}}_{k+1}(\xi, \psi) \tilde{\mathcal{K}}_k(\psi, \eta) d\mu(\psi) \\ &= 2 \int_{(\frac{1}{2}, 1]} \tilde{\mathcal{K}}_k(\psi, \eta) d\mu(\psi) \\ &= 2 \int_{(\frac{1}{2}, 1]} 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\eta) d\mu(\psi) \\ &= 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\eta). \end{aligned}$$

(ii)  $\xi \in (2^{-(k+1)}, \frac{1}{2}]$ : We have  $\tilde{\mathcal{K}}_{k+1}(\xi, \psi) = 2^{k+1} \cdot \mathbf{1}_{[0, 2^{-(k+1)}]}(\psi)$ , so that

$$\begin{aligned} \tilde{\mathcal{K}}_{k+1} \tilde{\mathcal{K}}_k(\xi, \eta) &= \int_{[0, 1]} \tilde{\mathcal{K}}_{k+1}(\xi, \psi) \tilde{\mathcal{K}}_k(\psi, \eta) d\mu(\psi) \\ &= 2^{k+1} \int_{[0, 2^{-(k+1)}]} \tilde{\mathcal{K}}_k(\psi, \eta) d\mu(\psi). \end{aligned}$$

Since  $[0, 2^{-(k+1)}] \subset [0, 2^{-k}]$  it follows that

$$\begin{aligned} \tilde{\mathcal{K}}_{k+1} \tilde{\mathcal{K}}_k(\xi, \eta) &= 2^{k+1} \int_{[0, 2^{-(k+1)}]} 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\eta) d\mu(\psi) \\ &= 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\eta). \end{aligned}$$

(iii)  $\xi \in (\frac{1}{2}, 1]$ : We have  $\tilde{\mathcal{K}}_{k+1}(\xi, \psi) = 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\psi)$  and a similar result as in case (i) holds.

Thus,  $\tilde{\mathcal{K}}_{k+1} \tilde{\mathcal{K}}_k(\xi, \eta) = 2 \cdot \mathbf{1}_{(\frac{1}{2}, 1]}(\eta)$  implying that  $\{\mathcal{K}_k\}$  is uniformly ergodic. Nevertheless, if we let  $S = [0, \frac{1}{2}]$ , then  $\bar{\mathcal{K}}_k^f(S) = 2^{-k}$  and  $\sum_{k=0}^{\infty} \bar{\mathcal{K}}_k^f(S) = 2 < \infty$ .

What makes the chain in Example 4 uniformly ergodic is the fact that there may exist sequences of sets with measures approaching zero and yet, such small measure sets can significantly contribute and lead to ergodicity. Thus, this straightforward generalization of the infinite flow property

need not be necessary for the weakest form of ergodicity in a general state space.

Nevertheless, with a proper definition of the infinite flow property, the necessity of the infinite flow is still true for the weak ergodicity. To adequately define the infinite flow property, for a stochastic kernel  $\mathcal{K}$  and any non-trivial set  $S \in \mathcal{M}$ , we define the flow from  $S$  to  $\bar{S}$  as follows:

$$\mathcal{K}^f(S, \bar{S}) = \sup_{\xi \in S} \mathcal{K}(\xi, \bar{S}). \quad (8)$$

Note that  $\mathcal{K}^f(S, \bar{S}) \leq 1$  since  $\mathcal{K}$  is a stochastic kernel. Using Eq. (8), we define the flow between  $S$  and  $\bar{S}$  to be  $\mathcal{K}^f(S) = \mathcal{K}^f(S, \bar{S}) + \mathcal{K}^f(\bar{S}, S)$ . Now, we define the infinite flow property.

*Definition 4:* We say that a chain of stochastic kernels  $\{\mathcal{K}_k\}$  has infinite flow property if  $\sum_{k=0}^{\infty} \mathcal{K}_k^f(S) = \infty$  for all non-trivial set  $S \in \mathcal{M}$ .

For example, the chain  $\{\mathcal{K}_k\}$  discussed in Example 4 has infinite flow property over the set  $S = [0, \frac{1}{2}]$ . In this case  $\mathcal{K}_k^f(S) = 1$  for all  $k \geq 0$  and  $\sum_{k=0}^{\infty} \mathcal{K}_k^f([0, \frac{1}{2}]) = \infty$ .

Definition 4 is a generalization of the infinite flow property in the finite-dimensional averaging dynamics. In the case of a stochastic chain, we have  $\{\mathcal{K}_k\} = \{A(k)\}$  and a non-trivial set  $S \in \mathcal{M}$  would be a set with  $S \neq \emptyset$  and  $S \neq [m]$ . For any such  $S$ , we have

$$\frac{1}{|S|} (A_{S\bar{S}}(k) + A_{\bar{S}S}(k)) \leq \mathcal{K}_k^f(S) \leq A_{S\bar{S}}(k) + A_{\bar{S}S}(k).$$

Thus,  $\{A(k)\}$  has the infinite flow property (in terms of Definition 4) if and only if  $\sum_{k=0}^{\infty} A_{S\bar{S}}(k) + A_{\bar{S}S}(k) = \infty$  for every non-trivial  $S \subset [m]$ , which coincides with the definition of the infinite flow property given in (5).

Next, we prove the necessity of the infinite flow property for the weak ergodicity.

*Theorem 1:* The infinite flow property is necessary for weak ergodicity.

*Proof:* Let  $\{\mathcal{K}_k\}$  be a chain that does not have infinite flow property. Then, there exists a non-trivial  $S \in \mathcal{M}$  with  $\sum_{k=0}^{\infty} \mathcal{K}_k^f(S) < \infty$ . Let  $t_0 \geq 0$  be such that  $\sum_{t=t_0}^{\infty} \mathcal{K}_t^f(S) \leq \frac{1}{4}$ . Let  $x_{t_0} = \mathbf{1}_S - \mathbf{1}_{\bar{S}}$  which is in  $L_{\infty}$ . Then, for any  $k \geq t_0$  and any  $\xi \in S$ , we have

$$\begin{aligned} x_{k+1}(\xi) &= \int_X \mathcal{K}_k(\xi, d\eta) x_k(\eta) \\ &= \int_S \mathcal{K}_k(\xi, d\eta) x_k(\eta) + \int_{\bar{S}} \mathcal{K}_k(\xi, d\eta) x_k(\eta) \\ &\geq (\inf_{\eta \in S} x_k(\eta)) \mathcal{K}_k(\xi, S) - \mathcal{K}_k(\xi, \bar{S}) \\ &= (\inf_{\eta \in S} x_k(\eta)) (1 - \mathcal{K}_k(\xi, \bar{S})) - \mathcal{K}_k(\xi, \bar{S}), \end{aligned} \quad (9)$$

where the inequality follows from  $\|x_k\|_{\infty} \leq \|x_{t_0}\|_{\infty} = 1$  (Lemma 1), and the last equality follows from  $\mathcal{K}_k(\xi, \cdot)$  being a probability measure. Thus, by using the induction on  $k \geq t_0$  and the preceding relation, assuming  $\inf_{\eta \in S} x_k(\eta) \geq 0$

(which holds for  $k = t_0$ ), we have

$$\begin{aligned} x_{k+1}(\xi) &\geq (\inf_{\eta \in S} x_k(\eta)) (1 - \mathcal{K}_k(\xi, \bar{S})) - \mathcal{K}_k(\xi, \bar{S}) \\ &\geq (\inf_{\eta \in S} x_k(\eta)) (1 - \mathcal{K}_k^f(S, \bar{S})) - \mathcal{K}_k^f(S, \bar{S}) \\ &\geq \inf_{\eta \in S} x_k(\eta) - 2\mathcal{K}_k^f(S, \bar{S}), \end{aligned}$$

where the last inequality follows by  $\|x(k)\|_{\infty} \leq 1$  and the fact that  $\mathcal{K}_k(\xi, \bar{S}) \leq \mathcal{K}_k^f(S, \bar{S})$ . Thus, by taking the infimum over  $\xi \in S$  in the preceding relation, we can see that for  $k \geq t_0$ ,

$$\begin{aligned} \inf_{\xi \in S} x_{k+1}(\xi) &\geq \inf_{\eta \in S} x_k(\eta) - 2\mathcal{K}_k^f(S) \\ &\geq \inf_{\eta \in S} x_{t_0}(\eta) - 2 \sum_{t=t_0}^k \mathcal{K}_t^f(S) \geq \frac{1}{2}, \end{aligned}$$

and hence,  $\inf_{\xi \in S} x_k(\xi) \geq \frac{1}{2} > 0$  for any  $k \geq t_0$ . Using the same line of argument, we have

$$\sup_{\eta \in \bar{S}} x_k(\eta) \leq \sup_{\eta \in \bar{S}} x_{t_0}(\eta) + 2 \sum_{t=t_0}^k \mathcal{K}_t^f(S) \leq -\frac{1}{2}.$$

Therefore, for any  $\xi \in S$  and any  $\eta \in \bar{S}$ , we have

$$\liminf_{k \rightarrow \infty} (x_k(\xi) - x_k(\eta)) \geq 1,$$

which shows that  $\{\mathcal{K}_k\}$  is not weakly ergodic.  $\blacksquare$

#### IV. QUADRATIC LYAPUNOV FUNCTION

Quadratic Lyapunov functions are often used to perform stability analysis and to establish bounds for the convergence of averaging dynamics. Moreover, as discussed in [5], the decrease rate for a class of time-varying Lyapunov functions can be established precisely. In this section, we show that an analogous result can be provided for quadratic Lyapunov functions in an arbitrary state space.

Let  $\pi$  be a probability measure on  $(X, \mathcal{M})$ . For a stochastic kernel  $\mathcal{K}$ , let us define a function  $\lambda$  on  $\mathcal{M}$  by

$$\lambda(S) = \int_X \pi(d\xi) \mathcal{K}(\xi, S) \quad \text{for all } S \in \mathcal{M}.$$

Then,  $\lambda$  induces a measure on  $(X, \mathcal{M})$  and it is a probability measure since

$$\lambda(X) = \int_X \pi(d\xi) \mathcal{K}(\xi, X) = \int_X \pi(d\xi) = 1.$$

Next, we have the following definition.

*Definition 5:* We say that a sequence  $\{\pi_k\}$  of probability measures on  $(X, \mathcal{M})$  is an absolute probability sequence for  $\{\mathcal{K}_k\}$  if

$$\pi_k(S) = \int_X \pi_{k+1}(d\xi) \mathcal{K}_k(\xi, S) \quad \text{for all } S \in \mathcal{M} \text{ and } k \geq 0.$$

As in the case of finite-dimensional averaging dynamics [5], using an absolute probability sequence and a convex function  $g : \mathbb{R} \rightarrow \mathbb{R}$ , we can construct a Lyapunov

comparison function for the dynamics driven by  $\{\mathcal{K}_k\}$ . For this, let us define  $V_{g,\pi} : L_\infty \times \mathbb{Z}_+ \rightarrow \mathbb{R}_+$  by

$$\begin{aligned} V_{g,\pi}(x, k) &= \int_X \pi_k(d\xi)g(x(\xi)) - g(\pi_k x) \\ &= \int_X \pi_k(d\xi) (g(x(\xi)) - g(\pi_k x)), \end{aligned} \quad (10)$$

where  $\{\pi_k\}$  is an absolute probability sequence for  $\{\mathcal{K}_k\}$  and

$$\pi x = \int_X \pi(d\xi)x(\xi).$$

The given comparison function  $V_{g,\pi}(x, k)$  can be interpreted as the weighted distance of a measurable function  $x \in L_\infty$  from the average point  $\pi_k x$ .

The following result shows that  $V_{g,\pi}(x, k)$  is a comparison function for the dynamics in (2).

*Theorem 2:* Let  $\{\pi_k\}$  be an absolute probability sequence for  $\{\mathcal{K}_k\}$ . Then, the function  $V_{g,\pi}(x, k)$  of Eq. (10) is a Lyapunov comparison function for the averaging dynamics driven by  $\{\mathcal{K}_k\}$ .

*Proof:* Let  $\{x_k\}$  be the dynamics driven by  $\{\mathcal{K}_k\}$ . Then, for any  $k \geq t_0$ , we have

$$\begin{aligned} &\int_X \pi_{k+1}(d\xi)g(x_{k+1}(\xi)) \\ &= \int_X \pi_{k+1}(d\xi)g\left(\int_X \mathcal{K}_k(\xi, d\eta)x_k(\eta)\right) \\ &\leq \int_X \pi_{k+1}(d\xi) \int_X \mathcal{K}_k(\xi, d\eta)g(x_k(\eta)) \\ &= \int_X \pi_k(d\eta)g(x_k(\eta)), \end{aligned} \quad (11)$$

where the inequality follows by the application of the Jensen's inequality ([2], page 461), and the last equality follows by  $\{\pi_k\}$  being absolute probability sequence for  $\{\mathcal{K}_k\}$ . On the other hand, we have  $\pi_{k+1}x_{k+1} = \pi_k x_k$  and hence,  $g(\pi_{k+1}x_{k+1}) = g(\pi_k x_k)$ . Using this observation and relation (11), we conclude that

$$\begin{aligned} V_{g,\pi}(x_{k+1}, k+1) &= \int_X \pi_{k+1}(d\xi)g(x_{k+1}(\xi)) - g(\pi_{k+1}x_{k+1}) \\ &\leq \int_X \pi_k(d\xi)g(x_k(\xi)) - g(\pi_k x_k) \\ &= V_{g,\pi}(x_k, k). \end{aligned}$$

Thus,  $V_{g,\pi}$  is non-increasing along any trajectory of the dynamics.  $\blacksquare$

As a special case, we will consider the quadratic Lyapunov function, given by

$$\begin{aligned} V_\pi(x, k) &= \int_X \pi_k(d\xi)(x(\xi) - \pi_k x)^2 \\ &= \int_X \pi_k(d\xi)x^2(\xi) - (\pi_k x)^2. \end{aligned} \quad (12)$$

Assume we are given a probability measure  $\pi$  on  $(X, \mathcal{M})$  and a stochastic kernel  $\mathcal{K}$ . Our next step is to define a probability measure  $H$  on the product space  $(X \times X, \mathcal{M} \otimes \mathcal{M})$

using  $\pi$  and  $\mathcal{K}$ . Let us define  $H$  on the product of two measurable sets  $S, T \in \mathcal{M}$  by

$$H(S \times T) = \int_X \pi(d\xi)\mathcal{K}(\xi, S)\mathcal{K}(\xi, T).$$

Moreover, for a collection of disjoint sets  $S_1 \times T_1, \dots, S_n \times T_n \subset X \times X$ , where  $S_i, T_i \in \mathcal{M}$  for all  $i \in [n]$ , let

$$\begin{aligned} H\left(\bigcup_{i=1}^n S_i \times T_i\right) &= \sum_{i=1}^n H(S_i \times T_i) \\ &= \sum_{i=1}^n \int_X \pi(d\xi)\mathcal{K}(\xi, S_i)\mathcal{K}(\xi, T_i). \end{aligned} \quad (13)$$

Equation (13) provides a pre-measure on the algebra of rectangular sets on  $X \times X$ . By Theorem 1.14 in [3],  $H$  can be extended to an outer measure on  $X \times X$  such that its restriction on  $\mathcal{M} \otimes \mathcal{M}$  is a measure. Note that  $H(X \times X) = \int_X \pi(d\xi)\mathcal{K}(\xi, X)\mathcal{K}(\xi, X) = 1$  and hence,  $H$  is a probability measure on the product space  $(X \times X, \mathcal{M} \otimes \mathcal{M})$ . We refer to the measure  $H$  constructed this way as the measure induced by  $\mathcal{K}$  and  $\pi$  on  $(X \times X, \mathcal{M} \otimes \mathcal{M})$ .

Now, we can quantify the decrease rate of the quadratic comparison function along a trajectory of the dynamics (2).

*Theorem 3:* Suppose that  $\pi$  and  $\lambda$  are probability measures on  $(X, \mathcal{M})$  such that  $\lambda(S) = \int_X \pi(d\xi)\mathcal{K}(\xi, S)$  for all  $S \in \mathcal{M}$ . Then, for all  $x, y \in L_\infty$  with  $y = \mathcal{K}x$ , we have

$$\begin{aligned} \int_X \pi(d\xi)y^2(\xi) - (\pi y)^2 &= \int_X \lambda(d\xi)x^2(\xi) - (\lambda x)^2 \\ &\quad - \frac{1}{2} \int_{X \times X} H(d\eta_1 \times d\eta_2)(x(\eta_1) - x(\eta_2))^2, \end{aligned} \quad (14)$$

where  $H$  is the probability measure induced by  $\mathcal{K}$  and  $\pi$  on  $(X \times X, \mathcal{M} \otimes \mathcal{M})$ .

*Proof:* We first prove relation (14) for an arbitrary simple function. Let  $x = \sum_{i=1}^m \alpha_i \mathbf{1}_{S_i}$  where  $\{S_i \mid i \in [m]\} \subseteq \mathcal{M}$  is a partition for  $X$  with  $m \geq 1$  disjoint sets (i.e.,  $S_i \cap S_j = \emptyset$  almost everywhere for all  $i, j \in [m]$  with  $i \neq j$  and  $\bigcup_{i \in [m]} S_i = X$ ). Also, let  $\alpha = (\alpha_1, \dots, \alpha_m)^T \in \mathbb{R}^m$ . Then,

$$y(\xi) = \mathcal{K}(\xi, \cdot)x = \int_X \mathcal{K}(\xi, d\eta)x(\eta) = \sum_{i=1}^m \alpha_i \mathcal{K}(\xi, S_i).$$

Therefore,

$$\begin{aligned} \pi y^2 &= \int_X \pi(d\xi)y^2(\xi) = \int_X \pi(d\xi) \left( \sum_{i=1}^m \alpha_i \mathcal{K}(\xi, S_i) \right)^2 \\ &= \int_X \pi(d\xi) \left( \sum_{i=1}^m \alpha_i^2 \mathcal{K}_i^2(\xi, S_i) \right) \\ &\quad + \int_X \pi(d\xi) \sum_{\substack{i,j \in [m] \\ i \neq j}} \alpha_i \alpha_j \mathcal{K}(\xi, S_i)\mathcal{K}(\xi, S_j). \end{aligned} \quad (15)$$

By the linearity of integral we have

$$\begin{aligned}
& \int_X \pi(d\xi) \sum_{\substack{i,j \in [m] \\ i \neq j}} \alpha_i \alpha_j \mathcal{K}(\xi, S_i) \mathcal{K}(\xi, S_j) \\
&= \sum_{\substack{i,j \in [m] \\ i \neq j}} \alpha_i \alpha_j \int_X \pi(d\xi) \mathcal{K}(\xi, S_i) \mathcal{K}(\xi, S_j) \\
&= \sum_{\substack{i,j \in [m] \\ i \neq j}} \alpha_i \alpha_j H(S_i \times S_j), \tag{16}
\end{aligned}$$

which follows by the definition of the induced measure  $H$ .

On the other hand, for any  $\xi \in X$  we have  $\mathcal{K}(\xi, S_i) = 1 - \sum_{j \neq i} \mathcal{K}(\xi, S_j)$  which follows from  $\mathcal{K}(\xi, \cdot)$  being a probability measure. Therefore,

$$\begin{aligned}
& \int_X \pi(d\xi) \sum_{i=1}^m \alpha_i^2 \mathcal{K}^2(\xi, S_i) \\
&= \int_X \pi(d\xi) \sum_{i=1}^m \alpha_i^2 \left( \mathcal{K}(\xi, S_i) - \sum_{j \neq i} \mathcal{K}(\xi, S_j) \mathcal{K}(\xi, S_i) \right).
\end{aligned}$$

We also have

$$\begin{aligned}
\int_X \pi(d\xi) \sum_{i=1}^m \alpha_i^2 \mathcal{K}(\xi, S_i) &= \sum_{i=1}^m \alpha_i^2 \int_X \pi(d\xi) \mathcal{K}(\xi, S_i) \\
&= \sum_{i=1}^m \alpha_i^2 \lambda(S_i) = \lambda x^2, \tag{17}
\end{aligned}$$

which follows from the definition of the probability measure  $\lambda$  and  $x = \sum_{i=1}^m \alpha_i \mathbf{1}_{S_i}$ . Also, by a similar argument to the one we have used to derive Eq. (16), we can show that

$$\begin{aligned}
& \int_X \pi(d\xi) \sum_{i=1}^m \alpha_i^2 \sum_{\substack{j \in [m] \\ j \neq i}} \mathcal{K}(\xi, S_i) \mathcal{K}(\xi, S_j) \\
&= \sum_{\substack{i,j \in [m] \\ i \neq j}} H(S_i \times S_j) \alpha_i^2. \tag{18}
\end{aligned}$$

Substituting relations (16)–(18) in Eq. (15) yields

$$\begin{aligned}
\pi y^2 &= \lambda x^2 - \sum_{\substack{i,j \in [m] \\ i \neq j}} H(S_i \times S_j) (\alpha_i^2 - \alpha_i \alpha_j) \\
&= \lambda x^2 - \frac{1}{2} \sum_{i,j \in [m]} H(S_i \times S_j) (\alpha_i - \alpha_j)^2,
\end{aligned}$$

where the last equality follows from  $H(S_i \times S_j) = H(S_j \times S_i)$  for any  $i, j \in [m]$ . Since the function  $x_i(\eta_1) - x_j(\eta_2)$  is equal to the constant  $\alpha_i - \alpha_j$  over  $S_i \times S_j$ , it follows that

$$\begin{aligned}
& \sum_{i,j \in [m]} H(S_i \times S_j) (\alpha_i - \alpha_j)^2 \\
&= \int_{X \times X} H(d\eta_1 \times d\eta_2) (x(\eta_1) - x(\eta_2))^2. \tag{19}
\end{aligned}$$

Note that  $(\pi y)^2 = (\pi(\mathcal{K}x))^2 = (\lambda x)^2$ . By subtracting this relation from both sides of Eq. (19), we arrive at the desired relation (14) for simple functions.

Now, we prove that the assertion holds for an arbitrary  $x \in L_\infty$ . The function  $T(z) = \pi z = \int_X \pi(d\xi) z(\xi)$  is a continuous function on  $(L_\infty, \|\cdot\|_\infty)$  for any probability measure  $\pi$ . Also,  $\tilde{T}(x) = x^2$  is a continuous function from  $L_\infty$  to  $L_\infty$ . Thus, the functions  $\pi y^2 - (\pi y)^2$  and  $\lambda x^2 - (\lambda x)^2$  are continuous functionals over  $L_\infty$ . Similarly, the functional  $\bar{T}(x) = \int_{X \times X} H(d\eta_1 \times d\eta_2) (x(\eta_1) - x(\eta_2))^2$  is a continuous functional from  $(L_\infty, \|\cdot\|_\infty)$  to  $\mathbb{R}$ , which follows from  $H$  being a probability measure on  $(X \times X, \mathcal{M} \otimes \mathcal{M})$ . Therefore, all the functionals involved in relation (14) are continuous over  $L_\infty$ . Since the relation holds for a dense subset of  $L_\infty$  (i.e., simple functions), we conclude that the relation holds for any  $x \in L_\infty$ . ■

The following corollary quantifies the total decrease of the quadratic Lyapunov function along the trajectories of an averaging dynamics. It is a consequence of Theorem 3.

*Corollary 1:* Let  $\{\mathcal{K}_k\}$  be a chain of stochastic integral kernels and let  $\{\pi_k\}$  be an absolute probability sequence for  $\{\mathcal{K}_k\}$ . Then, for the dynamics  $\{x_k\}$  driven by  $\{\mathcal{K}_k\}$  started at a time  $t_0$  and a point  $x_{t_0} \in L_\infty$ , we have for all  $k \geq t_0$ ,

$$\begin{aligned}
V_\pi(x_{k+1}, k+1) &= V_\pi(x_k) \\
&\quad - \frac{1}{2} \int_{X \times X} H_k(d\eta_1 \times d\eta_2) (x_k(\eta_1) - x_k(\eta_2))^2,
\end{aligned}$$

where  $H_k$  is the induced measure on  $(X \times X, \mathcal{M} \otimes \mathcal{M})$  by  $\mathcal{K}_k$  and  $\pi_{k+1}$ . Furthermore, we have

$$\begin{aligned}
& \sum_{k=t_0}^{\infty} \int_{X \times X} H_k(d\eta_1 \times d\eta_2) (x_k(\eta_1) - x_k(\eta_2))^2 \\
&\leq 2V_\pi(x_{t_0}, t_0).
\end{aligned}$$

## V. CONCLUSION

In this paper, we provided a framework for averaging dynamics over general state spaces. We defined several modes of ergodicity and consensus and showed that, in contrast to finite state spaces, these modes are not necessarily the same. We introduced the infinite flow property for averaging dynamics and showed that this property is necessary for the weakest form of ergodicity. Finally, we characterized quadratic Lyapunov functions for a class of averaging dynamics and precisely quantified the total decrease such a function along the trajectories of the dynamics.

## 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*, 3 ed., Curt Hinrichs, 2005.
- [3] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2 ed., Wiley, 1999.
- [4] E. Nummelin, *General Irreducible Markov Chains and Non-Negative Operators*, Cambridge University Press, 1984.
- [5] B. Touri, *Product of Random Stochastic Matrices and Distributed Averaging*, PhD Thesis, University of Illinois at Urbana-Champaign, to appear in Springer Theses series, 2012.
- [6] B. Touri and A. Nedić, *On ergodicity, infinite flow and consensus in random models*, IEEE Transactions on Automatic Control **56** (2011), no. 7, 1593–1605.