2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

On Existence of a Quadratic Comparison Function for Random Weighted Averaging Dynamics and Its Implications

Behrouz Touri and Angelia Nedić

***Abstract*— In this paper we study the stability and limiting behavior of discrete-time deterministic and random weighted averaging dynamics. We show that any such dynamics admits infinitely many comparison functions including a quadratic one. Using a quadratic comparison function, we establish the stability and characterize the set of equilibrium points of a broad class of random and deterministic averaging dynamics. This class includes a set of balanced chains, which itself contains many of the previously studied chains. Finally, we provide some implications of the developed results for products of independent random stochastic matrices.**

I. Introduction

A mathematical framework for studying weighted averaging dynamics models has been provided in [1]. Later, such models have found applications in distributed networked problems such as decentralized computation [2], [3], distributed optimization [2], [4], as well as in modeling of opinion dynamics [5].

A question on existence of a quadratic Lyapunov function for averaging dynamics has been raised in [6]. This question has been addressed in [7], where it was shown that a quadratic Lyapunov function does not exist for such dynamics in general. It is well-known that a quadratic Lyapunov function does exist for a special class of averaging dynamics driven by doubly stochastic chains. Convergence time of such averaging dynamics has been established in [8].

In this paper, we deal with discrete-time weighted averaging dynamics driven by deterministic chains or independent random chains. We are interested in studying the asymptotic behavior of the dynamics including stability and ergodicity. In our study, we make use of comparison functions for the dynamic models under consideration. In particular, we show that any averaging dynamics admits infinitely many comparison functions, among which there exists a quadratic comparison function (Section III). We then focus on the quadratic function and characterize its decrease over time. Using this characterization and some results from [9], [10], we prove the stability of a class of "balanced" random dynamics and characterize the equilibrium points for the models in this class. We show that the class includes many of the previously studied weighted averaging dynamics (Section IV). Finally, we provide two implications of the developed results (Section V). In particular, we apply our results to study ergodicity of random averaging dynamics.

Department of Industrial and Enterprise Systems Engineering, University of Illinois, Urbana, IL 61801, email: {touri1,angelia}@illinois.edu. The work is funded by the National Science Foundation grant CMMI 07-42538.

II. Problem Formulation and Basic Terminology

We introduce our notation, formulate the problem of our interest and provide some basic terminology that will be used in the subsequent development.

A. Notation and Basic Terminology

We use subscripts for indexing elements of vectors and matrices. We write $x \geq 0$ or $x > 0$ if, respectively, $x_i \geq 0$ or $x_i > 0$ for all i. We use e to denote the vector with all entries equal to one. A vector a is stochastic if $a \geq 0$ and $\sum_i a_i = 1$. We use I for the identity matrix. A matrix W is stochastic if all of its rows are stochastic vectors, and it is doubly stochastic if its rows and columns are stochastic vectors. We use $[m]$ to denote the set $\{1, \ldots, m\}$. We denote a proper subset of $[m]$ by $S \subset [m]$, its cardinality by $|S|$, and its complement by $\bar{S}$. A set $S \subset [m]$ such that $S \neq \emptyset$ is a *nontrivial* subset of $[m]$. For an $m \times m$ matrix W, we let $W_S = \sum_{i \in S, j \in \bar{S}} (W_{ij} + W_{ji})$. For a matrix sequence $\{W(k)\}$, we define $W(t:k) = W(t-1) \cdots W(k)$ for $t > k$ and $W(k:k) = I$ for $k \geq 0$. For a vector v, we use $\mathrm{diag}(v)$ to denote the diagonal matrix with the ith diagonal entry equal to v_i. We write $\bar{X} = \mathsf{E}[X]$ for the expected value of a random variable (matrix) X. We often use *a.s.* to denote *almost surely*.

B. Dynamics, Stability and Ergodicity

Throughout the paper we deal with stochastic matrices $W(k)$. Let $\{W(k)\}$ be an independent random chain of $m \times m$ stochastic matrices, i.e., $W_{ij}(k)$ is a random variable on some probability space for all $i, j \in [m]$ and all $k \geq 0$, $\{W(k)\}$ is an independent chain, and $W(k)$ is a stochastic matrix almost surely for all $k \geq 0$. For a starting time $t_0 \geq 0$ and a starting point $x(t_0) \in \mathbb{R}^m$, consider the following dynamics

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

We refer to $\{x(k)\}$ as a (random) dynamics driven by the chain $\{W(k)\}$.

We are interested in the limiting behavior of dynamics (1) and, especially, in the *stability* and *ergodicity* of the dynamics. These concepts are defined as follows.

Stability: $\{W(k)\}$ is a stable chain if $\lim_{k\to\infty} x(k)$ exists a.s. for every $t_0 \geq 0$ and $x(t_0) \in \mathbb{R}^m$.

Ergodicity: $\{W(k)\}$ is ergodic if $\lim_{k\to\infty} (x_i(k) - x_j(k)) = 0$ a.s. for all $t_0 \geq 0$, all $x(t_0) \in \mathbb{R}^m$, and all $i, j \in [m]$.

To ensure stability, we often assume that the given chain $\{W(k)\}$ has a form of feedback property as defined below.

Definition 1: A random chain $\{W(k)\}$ has *feedback property* if there is a scalar $\gamma > 0$ such that $\mathsf{E}[W_{ii}(k)W_{ij}(k)] \geq \gamma\mathsf{E}[W_{ij}(k)]$ for all $i, j \in [m]$ with $i \neq j$ and all $k \geq 0$.

We refer to γ as a feedback coefficient. For a deterministic chain $\{A(k)\}$, the feedback property is equivalent to $A_{ii}(k) \geq \gamma$ for all $i \in [m]$ and $k \geq 0$.

III. Comparison Functions for Random Chains

In this section, we introduce absolute probability sequences associated with a random chain of stochastic matrices. Then, using these sequences, we show that infinitely many comparison functions can be constructed for dynamics driven by stochastic chains. Subsequently, we focus on quadratic comparison functions and study the decrease of these functions along trajectories of the weighted-average dynamics of our interest.

A. Absolute Probability Sequences

In his elegant work [11], Kolmogorov has introduced and studied an *absolute probability sequence*, which is a special sequence of stochastic vectors associated with an inhomogeneous Markov chain. We adopt Kolmogorov's definition of absolute probability sequence with a slight adjustment to fit in the framework of random chains.

Definition 2: A sequence $\{\pi(k)\}$ of stochastic vectors is an *absolute probability* sequence for a random chain $\{W(k)\}$ if the following relation holds:

$$\pi^T(k+1)\mathsf{E}[W(k)] = \pi^T(k) \qquad \text{for all } k \geq 0.$$

Generally, a chain may have infinitely many such sequences. As an example, consider A to be a non-singular stochastic matrix. Starting from any stochastic vector $\pi(0)$, we can define $\pi^T(k) = \pi^T(0)A^{-k}$. Then, $\{\pi(k)\}$ is an absolute probability sequence for the static chain $\{A\}$. Hence, this chain has infinitely many absolute probability sequences.

In [11], Kolmogorov proved that every (deterministic) chain $\{A(k)\}$ of stochastic matrices has an absolute probability sequence. His elegant proof uses the fact that the set of stochastic matrices is a compact subset of $\mathbb{R}^{m\times m}$. In view of this, starting from any time $k \geq 0$, the (backward) product

$$A(t:k) = A(t-1)\cdots A(k+1)A(k)$$

has a convergent subsequence as t goes to the infinity. Now, by some diagonalization argument (see [12] for details), we can find a subsequence $\{t_k\}$ such that

$$R(k) = \lim_{r\to\infty} A(t_r : k) \tag{2}$$

exists for any time $k \geq 0$. Having the sequence $\{R(k)\}$ and choosing any stochastic vector $\hat{\pi}$, we define $\pi(k) = R^T(k)\hat{\pi}$ for all k. Then, the sequence $\{\pi(k)\}$ is an absolute probability sequence for $\{A(k)\}$. This holds since

$$\begin{aligned}\pi^T(k+1)A(k) &= \hat{\pi}^T R(k+1)A(k)\\ &= \hat{\pi}^T \lim_{r\to\infty} A(t_r : k+1)A(k) = \hat{\pi}^T \lim_{r\to\infty} A(t_r : k),\end{aligned}$$

and $\hat{\pi}^T \lim_{r\to\infty} A(t_r : k) = \hat{\pi}^T R(k) = \pi^T(k)$. Note that the definition of the absolute probability sequence involves only the expected chain. Thus, by Kolmogorov's argument, it follows that *any random chain $\{W(k)\}$ has an absolute probability sequence*.

B. Constructing Comparison Functions

Using an absolute probability sequence $\{\pi(k)\}$ of $\{W(k)\}$, we can come up with a rich family of comparison functions for the random dynamics in (1). For this, let $g : \mathbb{R} \to \mathbb{R}$ be an *arbitrary convex function* and let us define

$$V_{g,\pi}(x,k) = \sum_{i=1}^m \pi_i(k)g(x_i) - g(\pi^T(k)x). \tag{3}$$

We will show that $V_{g,\pi}(x(k),k)$ is a supermartingale along any trajectory $\{x(k)\}$ of the dynamics in (1). Before proving this, let us take a closer look at the comparison function in Eq. (3). Suppose that g is a differentiable convex function. Then, we have $\sum_{i=1}^m \pi_i(k)g'(\pi^T(k)x)(x_i - \pi^T(k)x) = g'(\pi^T(k)x)\sum_{i=1}^m \pi_i(k)(x_i - \pi^T(k)x) = 0$. Therefore, by the stochasticity of $\pi(k)$, we obtain

$$\begin{aligned}V_{g,\pi}(x,k) &= \sum_{i=1}^m \pi_i(k)\left(g(x_i) - g(\pi^T(k)x)\right)\\ &= \sum_{i=1}^m \pi_i(k)\left(g(x_i) - g(\pi^T(k)x) - g'(\pi^T(k)x)(x_i - \pi^T(k)x)\right)\\ &= \sum_{i=1}^m \pi_i(k)B_g(x_i \parallel \pi^T(k)x),\end{aligned} \tag{4}$$

where $B_g(\alpha\|\beta) = g(\alpha)-g(\beta)-g'(\beta)(\alpha-\beta)$ is the Bregman divergence (distance) of α and β with respect to the convex function $g(\cdot)$ as defined in [13]. Equation (4) shows that our comparison function is in fact a weighted average of the Bregman divergence of the mass points $x_1,\ldots,x_m$ with respect their weighted center of the mass.

We now shows that $V_{g,\pi}$ is a comparison function for the random dynamics in (1).

Lemma 1: For a dynamic $\{x(k)\}$ driven by an independent random chain $\{W(k)\}$ with an absolute probability sequence $\{\pi(k)\}$, we have $\mathsf{E}[V_{g,\pi}(x(k+1),k+1) \mid x(k)] \leq V_{g,\pi}(x(k),k)$.

Proof: By the definition of $V_{g,\pi}$ in Eq. (3), we have

$$\begin{aligned}&V_{g,\pi}(x(k+1),k+1)\\ &= \sum_{i=1}^m \pi_i(k+1)g(x_i(k+1)) - g(\pi^T(k+1)x(k+1))\\ &= \sum_{i=1}^m \pi_i(k+1)g([W(k)x(k)]_i) - g(\pi^T(k+1)x(k+1))\\ &\leq \sum_{i=1}^m \pi_i(k+1)\sum_{j=1}^m W_{ij}(k)g(x_j(k)) - g(\pi^T(k+1)x(k+1)),\end{aligned}$$

where the inequality follows by convexity of $g(\cdot)$ and stochasticity of matrices $W(k)$ for any sample point. Now, since $\{W(k)\}$ is independent and $\pi^T(k+1)\mathsf{E}[W(k)] = \pi^T(k)$, by taking the conditional expectation on $x(k)$ and using Jensen's inequality ([14] page 225) for the convex function g, we have

$$\begin{aligned}&\mathsf{E}[V_{g,\pi}(x(k+1),k+1)\mid x(k)]\\&\le\sum_{j=1}^{m}\pi_j(k)g(x_j(k))-\mathsf{E}\big[g(\pi^T(k+1)x(k+1))\mid x(k)\big]\\&\le\sum_{j=1}^{m}\pi_j(k)g(x_j(k))-g(\pi^T(k+1)\mathsf{E}[x(k+1)\mid x(k)]).\end{aligned}$$

By using the dynamics equation for $x(k+1)$ and the defining property of the absolute probability sequence, we obtain

$$\begin{aligned}&\mathsf{E}[V_{g,\pi}(x(k+1),k+1)\mid x(k)]\\&\le\sum_{j=1}^{m}\pi_j(k)g(x_j(k))-g(\pi^T(k)x(k))=V_{g,\pi}(x(k),k).\end{aligned}$$

■

Lemma 1 provides us with infinitely many choices for constructing comparison functions for random and deterministic averaging dynamics through the use of an absolute probability sequence and a convex function g. Here, we mention two functions which might be of particular interest:

Quadratic function: Let $g(s)=s^2$ and let us consider the formulation provided in Eq. (4). Then, it can be seen that

$$V_{g,\pi}(x,k)=\sum_{i=1}^{m}\pi_i(k)(x_i-\pi^T(k)x)^2. \tag{5}$$

For dynamics $\{x(k)\}$ in (1), the function $V_{g,\pi}(x(k),k)=\sum_{i=1}^{m}\pi_i(k)(x_i(k)-\pi^T(k)x(k))^2$ is a bounded supermartingale in $\mathbb{R}$ and, hence, convergent almost surely.

Kullback-Leibler divergence: Let $x(0)\in[0,1]^m$. One can view $x(0)$ as a vector of positions of m particles in $[0,1]$. Intuitively, by the successive weighted averaging of the m particles, the entropy of such system should not increase. Mathematically, this corresponds to choosing $g(s)=s\ln(s)$ in Eq. (3) (with $g(0)=0$). Then, it can be seen that

$$V_{g,\pi}(x,k)=\sum_{i=1}^{m}\pi_i(k)D_{KL}(x_i\parallel\pi^T(k)x),$$

where $D_{KL}(\alpha\parallel\beta)=\alpha\ln(\frac{\alpha}{\beta})$ is the Kullback-Leibler divergence of α and β.

C. Quadratic Comparison Functions

Lyapunov function $d(x)=\max_i x_i-\min_i x_i$ has often been used to prove convergence and establish convergence rate for weighted averaging dynamics (1). However, convergence rate results obtained by using $d(x)$ are often not the best and tend to have an unfavorable growth with the dimension m of the underlying space. In [8], using a quadratic Lyapunov function, tight convergence rate results were shown for the averaging dynamics driven by doubly stochastic matrices. Unfortunately, these results do not extend to stochastic chains in general since not all stochastic chains (even those with rich structures) have a quadratic Lyapunov function, as shown in [7]. However, using Lemma 1 with the quadratic function $g(s)=s^2$, we see that *any deterministic and random stochastic chain has a quadratic comparison function*. We study the properties of this comparison function, which in a sense extends the role of Lyapunov function in [8].

In our further development, we consider a *quadratic comparison function* as in (5). To simplify the notation, we omit the subscript g in $V_{g,\pi}$ and define

$$V_\pi(x,k)=\sum_{i=1}^{m}\pi_i(k)(x_i-\pi^T(k)x)^2.$$

For a chain $\{W(k)\}$ with an absolute probability sequence $\{\pi(k)\}$, we refer to $V_\pi(x,k)$ as *the quadratic comparison function associated with* $\{\pi(k)\}$.

By the existence of a probability sequence for a stochastic chain and by Lemma 1, the following corollary is immediate.

Corollary 1: Every independent random chain $\{W(k)\}$ admits a quadratic comparison function.

Lemma 1 implies that $\{V_\pi(x(k),k)\}$ is a supermartingale along any trajectory of dynamics $\{x(k)\}$. However, in order to show stability and to develop rate of convergence results, we need to quantify the amount of decrease at each time step. The following result gives a lower bound for the amount of decrease, which is *exact* for deterministic chains.

Theorem 1: Let $\{W(k)\}$ be an independent random chain and let $\{\pi(k)\}$ be an absolute probability sequence for $\{W(k)\}$. Then, for any trajectory $\{x(k)\}$ under $\{W(k)\}$, we have

$$\begin{aligned}&\mathsf{E}[V_\pi(x(k+1),k+1)\mid x(k)]\\&\le V_\pi(x(k),k)-\sum_{i<j}H_{ij}(k)(x_i(k)-x_j(k))^2,\end{aligned}$$

where $H(k)=\mathsf{E}\big[W^T(k)\mathrm{diag}(\pi(k+1))W(k)\big]$. Furthermore, if $\pi^T(k+1)W(k)=\pi^T(k)$ almost surely, then the above inequality holds as equality.

Proof: Let $D(k)=\mathrm{diag}(\pi(k))-\mathsf{E}\big[W^T(k)\mathrm{diag}(\pi(k+1))W(k)\big]$. Then, we have:

$$\begin{aligned}D(k)e&=\big[\mathrm{diag}(\pi(k))-\mathsf{E}\big[W^T(k)\mathrm{diag}(\pi(k+1))W(k)\big]\big]\,e\\&=\pi(k)-\mathsf{E}\big[W^T(k)\mathrm{diag}(\pi(k+1))e\big]\,,\end{aligned}$$

where we used the fact that $W(k)$ is stochastic, almost surely. Therefore, since $\{\pi(k)\}$ is an absolute probability sequence for $\{W(k)\}$, we have $D(k)e=\pi(k)-\mathsf{E}\big[W^T(k)\pi(k+1)\big]=0$. Also, note that $D(k)$ is a symmetric matrix. Thus, by Proposition 3.1.3 in [15], and since $D_{ij}(k)=-H_{ij}(k)$ for $i\neq j$, we have:

$$x^T(k)D(k)x^T(k)=\sum_{i<j}H_{ij}(k)(x_i(k)-x_j(k))^2. \tag{6}$$

In the other hand, $V_\pi(x,k)=x^T\mathrm{diag}(\pi(k))x-(\pi^T(k)x)^2$. Thus,

$$\begin{aligned}x^T(k)D(k)x^T(k)&=V_\pi(x(k),k)-\mathsf{E}[V_\pi(x(k+1),k+1)\mid x(k)]\\&-\left\{(\pi^T(k)x(k))^2-\mathsf{E}\big[(\pi^T(k+1)x(k+1))^2\mid x(k)\big]\right\}.\end{aligned}$$

By the convexity of $s\to s^2$ and Jensen's inequality, we have

$$\mathsf{E}\big[(\pi^T(k+1)x(k+1))^2\mid x(k)\big]\ge\mathsf{E}\big[\pi^T(k+1)x(k+1)\mid x(k)\big]^2,$$

and since, $\{\pi(k)\}$ is an absolute probability sequence, we have $\mathsf{E}\big[\pi^T(k+1)W(k)x(k) \mid x(k)\big]^2 = (\pi^T(k)x(k))^2$. Note that the equality holds if $\pi^T(k+1)W(k) = \pi(k)$ almost surely. Thus, overall, we have

$$x^T(k)D(k)x^T(k) \geq V_\pi(x(k),k) - \mathsf{E}[V_\pi(x(k+1),k+1) \mid x(k)]\,.$$

Combining this relation and (6) concludes the assertion. ■

Theorem 1 is a generalization of Lemma 4 in [8], which has played a central role in the convergence rate result for deterministic doubly stochastic chains. In case of these chains, by letting $\pi(k) = \frac{1}{m}e$ for all k in Theorem 1, we obtain Lemma 4 in [8].

From Theorem 1 we have the following important result.

Corollary 2: Let $\{\pi(k)\}$ be an absolute probability sequence for an independent random chain $\{W(k)\}$. Then, for any starting time $t_0 \geq 0$ and starting point $x(t_0) \in \mathbb{R}^m$, we have $\sum_{k=t_0}^{\infty}\sum_{i<j} H_{ij}(k)\mathsf{E}\Big[(x_i(k)-x_j(k))^2\Big] \leq \mathsf{E}[V_\pi(x(t_0),t_0)] < \infty$.

Now, let $\{W(k)\}$ be an independent random chain with an absolute probability sequence $\{\pi(k)\}$. We say that $\{\pi(k)\}$ is a *uniformly bounded* absolute probability sequence if there is $p^* > 0$ such that $\pi_i(k) \geq p^* > 0$ for all $i \in [m]$ and $k \geq 0$. We denote by $\mathcal{P}^*$ *the set of independent random chains that have uniformly bounded absolute probability sequence*. These chains are central in our further development. In Section IV, we will derive two conditions that certify when a chain is in $\mathcal{P}^*$.

D. Infinite-Flow Stability

In [9], [10], we have shown that for certain chains, the dynamics in (1) is stable almost surely. Also, we have characterized the equilibrium points of such a dynamics using *the infinite flow graph* of the chain that drives the dynamics. Here, under an additional condition, we extend the characterization to an arbitrary chain in $\mathcal{P}^*$. To do so, we introduce infinite flow graph, as appeared in [10].

Definition 3: For an independent random chain $\{W(k)\}$, the infinite flow graph $G^\infty = ([m],\mathcal{E}^\infty)$ is the graph with the vertex set $[m]$ and the edge set $\mathcal{E}^\infty = \{\{i,j\} \mid \sum_{k=0}^{\infty}(W_{ij}(k)+W_{ji}(k)) = \infty \quad a.s.,\ i \neq j \in [m]\}$.

Basically, the infinite flow graph of an independent random chain is the graph consisting of the edges that carry infinite accumulated weights almost surely. As discussed in [10], an independent random chain $\{W(k)\}$ and its expected chain have the same infinite flow graph.

For a given random chain$\{W(k)\}$, we define infinite-flow stability as follows.

Definition 4: An independent random chain $\{W(k)\}$ is infinite-flow stable if the dynamics $\{x(k)\}$ is stable and $\lim_{k\to\infty}(x_i(k)-x_j(k)) = 0$ a.s. for all $t_0 \geq 0$, all $x(t_0) \in \mathbb{R}^m$ and every $\{i,j\} \in \mathcal{E}^\infty$, where $\mathcal{E}^\infty$ is the edge set of the infinite flow graph of $\{W(k)\}$.

We are now ready to state one of our main results.

Theorem 2: Let $\{W(k)\} \in \mathcal{P}^*$ and let $\{W(k)\}$ have feedback property. Then, $\{W(k)\}$ is infinite-flow stable.

Proof: Let $t_0 \geq 0$ and $x(t_0) \in \mathbb{R}^m$ be arbitrary starting time and starting point for the dynamics $\{x(k)\}$ defined by Eq. (1). Also, let $\{\pi(k)\}$ be a uniformly bounded absolute probability sequence for $\{\bar{W}(k)\}$. By Corollary 2, we have $\sum_{k=t_0}^{\infty}\sum_{i<j} H_{ij}(k)\mathsf{E}\big[(x_i(k)-x_j(k))^2\big] < \infty$. We let $L(k) = \mathsf{E}\big[W(k)^T W(k)\big]$ and note that

$$H(k) = \mathsf{E}\big[W(k)^T \mathrm{diag}(\pi(k+1))W(k)\big] \geq p^* L(k).$$

Therefore, $\sum_{k=t_0}^{\infty}\sum_{i<j} L_{ij}(k)\mathsf{E}\big[(x_i(k)-x_j(k))^2\big] < \infty$ and, hence, $\{W(k)\}$ belongs to the $\mathscr{M}_2$-class of chains, as defined in [10]. Furthermore, $\{W(k)\}$ has weak feedback property (see [10]), which is implied by its feedback property. Then, using a similar approach as in Theorem 5 of [10], we can see that $\{W(k)\}$ is an infinite-flow stable chain. ■

Theorem 2 not only shows that dynamics (1) is stable almost surely for chains in $\mathcal{P}^*$ that have feedback property but also characterizes the equilibrium points of this dynamics.

IV. Balanced Chains

In this section, we characterize a sub-class of chains in $\mathcal{P}^*$, namely the set of *balanced chains* (in expectation) with feedback property. This class includes many of the chains that have been studied in the existing literature. Our goal is to show that any balanced chain with feedback property has a uniformly bounded absolute probability sequence, or in other words, it belongs to the class $\mathcal{P}^*$.

We start our development by considering deterministic chains at first and we discuss random chains later. For this, let $\{A(k)\}$ be a deterministic chain of stochastic matrices and $A_{S\bar{S}}(k) = \sum_{i\in S, j\in\bar{S}} A_{ij}(k)$ for $S \subset [m]$.

Definition 5: A chain $\{A(k)\}$ is *balanced*[1] if there exists a scalar $\alpha > 0$ such that

$$A_{S\bar{S}}(k) \geq \alpha\, A_{\bar{S}S}(k) \qquad \text{for any } S \subset [m] \text{ and } k \geq 0.$$

We refer to α as a balancedness coefficient, and we denote the set of balanced chains by $\mathcal{B}$.

Note that in Definition 5, the scalar α is time-independent. Furthermore, due to the inter-changeability of any subset S with its complement $\bar{S}$, for a balanced chain $\{A(k)\}$ we have

$$A_{S\bar{S}}(k) \geq \alpha A_{\bar{S}S}(k) \geq \alpha^2 A_{S\bar{S}}(k),$$

implying $\alpha \leq 1$.

A. Absolute Probability Sequence of Balanced Chains

The main result of this section is that every balanced chain with feedback property has a uniformly bounded absolute probability sequence, i.e., belongs to the class of $\mathcal{P}^*$ chains.

We start by establishing an auxiliary result in forthcoming Lemma 2, which proof builds on the approach used in [17]. However, this approach needed some significant extensions to fit in our more general assumption of balancedness. Also, the lemma shows that the bound for the nonnegative entries given in Proposition 4 of [17] can be reduced.

[1]Through a personal discussion of the first author of this paper with the first author of [16] at CDC 2010, we noticed that work in [16] discusses the same property for chains in continuous time dynamics.

To this end, let $S_j(k) = \{\ell \in [m] \mid A_{\ell j}(k:0) > 0\}$ be the set of indices of the positive entries in the jth column of $A(k:0)$ for $j \in [m]$ and $k \geq 0$. Also, let $\mu_j(k)$ be the minimum of these positive entries, i.e., $\mu_j(k) = \min_{\ell \in S(k)} A_{\ell j}(k:0) > 0$. We have the following result.

Lemma 2: Let $\{A(k)\}$ be a balanced chain with feedback property and with uniformly bounded positive coefficients, i.e., there exists a scalar $\gamma > 0$ such that $A_{ij}(k) \geq \gamma$ when $A_{ij}(k) > 0$. Then, $S_j(k) \subseteq S_j(k+1)$ and $\mu_j(k) \geq \gamma^{|S_j(k)|-1}$ for all $j \in [m]$ and $k \geq 0$.

Proof: Let $j \in [m]$ be arbitrary but fixed. By induction on k, we prove that $S_j(k) \subseteq S_j(k+1)$ for all $k \geq 0$ as well as the desired relation. For $k = 0$, we have $A(0:0) = I$ (see notation), so $S_j(0) = \{j\}$. Then, $A(1:0) = A(1)$ and by the feedback property of the chain $\{A(k)\}$ we have $A_{jj}(1) \geq \gamma$, implying $\{j\} = S_j(0) \subseteq S_j(1)$. Furthermore, we have $|S_j(0)| - 1 = 0$ and $\mu_j(0) = 1 \geq \gamma^0$. Hence, the claim is true for $k = 0$.

Now suppose that the claim is true for some $k \geq 0$. By the feedback property and the boundedness of the positive entries of $A(k)$, we have for $i \in S_j(k)$: $A_{ij}(k+1:0) = \sum_{\ell=1}^m A_{i\ell}(k) A_{\ell j}(k:0) \geq A_{ii}(k) A_{ij}(k:0) \geq \gamma \mu_j(k) > 0$. Thus, $i \in S_j(k+1)$, implying $S_j(k) \subset S_j(k+1)$. To show the relation for $\mu_j(k+1)$, we consider two cases:
Case 1: $A_{S_j(k)\bar{S}_j(k)}(k) = 0$. In this case for any $i \in S_j(k)$, we have:

$$A_{ij}(k+1:0) = \sum_{\ell \in S_j(k)} A_{i\ell}(k) A_{\ell j}(k:0)$$
$$\geq \mu_j(k) \sum_{\ell \in S_j(k)}^{m} A_{i\ell}(k) = \mu_j(k), \qquad (7)$$

where the inequality follows by the fact that $i \in S_j(k)$ and $A_{S_j(k)\bar{S}_j(k)}(k) = 0$, and the definition of $\mu_j(k)$. Furthermore, by the balancedness of $A(k)$ and $A_{S_j(k)\bar{S}_j(k)}(k) = 0$, it follows that $0 = A_{S_j(k)\bar{S}_j(k)}(k) \geq \alpha A_{\bar{S}_j(k)S_j(k)} \geq 0$. Hence, $A_{\bar{S}_j(k)S_j(k)}(k) = 0$. Thus, for any $i \in \bar{S}_j(k)$, we have $A_{ij}(k+1:0) = \sum_{\ell=1}^m A_{i\ell}(k) A_{\ell j}(k:0) = \sum_{\ell \in \bar{S}_j(k)} A_{i\ell}(k) A_{\ell j}(k:0) = 0$, where the second equality follows from $A_{\ell j}(k:0) = 0$ for all $j \in \bar{S}_j(k)$. Therefore, in this case we have $S_j(k+1) = S_j(k)$, which by (7) implies $\mu_j(k+1) \geq \mu_j(k)$. In view of $S_j(k+1) = S_j(k)$ and the inductive hypothesis, we further have $\mu_j(k) \geq \gamma^{|S_j(k)|-1} = \gamma^{|S_j(k+1)|-1}$, implying $\mu_j(k+1) \geq \gamma^{|S_j(k+1)|-1}$.
Case 2: $A_{S_j(k)\bar{S}_j(k)}(k) > 0$. Since the chain is balanced, we have $A_{\bar{S}_j(k)S_j(k)}(k) \geq \alpha A_{S_j(k)\bar{S}_j(k)}(k) > 0$ implying that $A_{\bar{S}_j(k)S_j(k)}(k) > 0$. Therefore, by the uniform boundedness of the positive entries of $A(k)$, there exists $\hat{\xi} \in \bar{S}_j(k)$ and $\hat{\ell} \in S_j(k)$ such that $A_{\hat{\xi}\hat{\ell}}(k) \geq \gamma$. Hence, we have

$$A_{\hat{\xi}j}(k+1:0) \geq A_{\hat{\xi}\hat{\ell}}(k) A_{\hat{\ell}j}(k:0) \geq \gamma \mu_j(k) = \gamma^{|S_j(k)|},$$

where the equality follows by the induction hypothesis. Thus, $\hat{\xi} \in S(k+1)$ while $\hat{\xi} \notin S_j(k)$, implying $|S_j(k+1)| \geq |S_j(k)| + 1$. This, together with $A_{\hat{\xi}j}(k+1:0) \geq \gamma^{|S_j(k)|}$, yields $\mu_j(k+1) \geq \gamma^{|S_j(k)|} \geq \gamma^{|S_j(k+1)|-1}$. ∎

It can be seen that Lemma 2 holds for products $A(k:t_0)$ starting with any $t_0 \geq 0$ and $k \geq t_0$ (with appropriately defined $S_j(k)$ and $\mu_j(k)$). An immediate corollary of Lemma 2 is the following result.

Corollary 3: Under the assumptions of Lemma 2, we have $\frac{1}{m} e^T A(k:t_0) \geq \min(\frac{1}{m}, \gamma^{m-1}) e^T$ for $k \geq t_0 \geq 0$.

Proof: Without loss of generality, let $t_0 = 0$. By Lemma 2, for any $j \in [m]$, we have $\frac{1}{m} e^T A^j(k:0) \geq \frac{1}{m} |S_j(k)| \gamma^{|S_j(k)|-1}$, where A^j denotes the jth column of A. For $\gamma \in [0,1]$, the function $t \mapsto t\gamma^{t-1}$ defined on $[0,m]$ attains its minimum at either $t = 0$ or $t = m$. Therefore, $\frac{1}{m} e^T A^j(k:0) A(k:0) \geq \min(\frac{1}{m}, \gamma^{m-1}) e^T$. ∎

Now, we relax the assumption on the bounded entries in Corollary 3.

Theorem 3: Let $\{A(k)\}$ be a balanced chain with feedback property. Let $\alpha, \beta > 0$ be balancedness and feedback coefficients for $\{A(k)\}$, respectively. Then, there is a scalar $\gamma = \gamma(\alpha, \beta) \in (0,1]$ such that $\frac{1}{m} e^T A(k:0) \geq \min(\frac{1}{m}, \gamma^{m-1}) e^T$ for any $k \geq 0$.

Proof: Let $\mathcal{B}_{\alpha,\beta}$ be the set of balanced matrices with the balancedness coefficient α and feedback property with coefficient $\beta > 0$, i.e.,

$$\mathcal{B}_{\alpha,\beta} := \{Q \in \mathbb{R}^{m \times m} \mid Q \geq 0, Qe = e, \qquad (8)$$
$$Q_{S\bar{S}} \geq \alpha Q_{\bar{S}S} \text{ for all } S \subset [m], Q_{ii} \geq \beta \text{ for all } i \in [m]\}.$$

The description in relation (8) shows that $\mathcal{B}_{\alpha,\beta}$ is a bounded polyhedral set in $\mathbb{R}^{m \times m}$. Let $\{Q^\xi \in \mathcal{B}_{\alpha,\beta} \mid \xi \in [n_{\alpha,\beta}]\}$ be the set of vertices of this polyhedral set indexed by the positive integers between 1 and $n_{\alpha,\beta}$, which is the number of extreme points of $\mathcal{B}_{\alpha,\beta}$.

Since $A(k) \in \mathcal{B}_{\alpha,\beta}$ for all $k \geq 0$, we can write $A(k)$ as a convex combination of the extreme points in $\mathcal{B}_{\alpha,\beta}$, i.e., there exist coefficients $\lambda_\xi(k) \in [0,1]$ such that

$$A(k) = \sum_{\xi=1}^{n_{\alpha,\beta}} \lambda_\xi(k) Q^\xi \qquad \text{with} \quad \sum_{\xi=1}^{n_{\alpha,\beta}} \lambda_\xi(k) = 1. \qquad (9)$$

Now, consider the following independent random matrix process defined by:

$$W(k) = Q^\xi \qquad \text{with probability } \lambda_\xi(k).$$

In view of this definition any sample path of $\{W(k)\}$ consists of extreme points of $\mathcal{B}_{\alpha,\beta}$. Thus, every sample path of $\{W(k)\}$ has bounded coefficient bounded by the minimum positive entry of the matrices in $\{Q^\xi \in \mathcal{B}_{\alpha,\beta} \mid \xi \in [n_{\alpha,\beta}]\}$, denoted by $\gamma = \gamma(\alpha, \beta) > 0$, where $\gamma > 0$ since $n_{\alpha,\beta}$ is finite. Therefore, by Corollary 3, we have $\frac{1}{m} e^T W(k:t_0) \geq \min(\frac{1}{m}, \gamma^{m-1}) e^T$ for all $k \geq t_0 \geq 0$. Furthermore, by Eq. (9) we have $\mathsf{E}[W(k)] = A(k)$ for all $k \geq 0$, implying

$$\frac{1}{m} e^T A(k:t_0) = \frac{1}{m} e^T \mathsf{E}[W(k:t_0)] \geq \min\left(\frac{1}{m}, \gamma^{m-1}\right) e^T,$$

where we also use independence of $\{W(k)\}$. ∎

We are now ready to prove the main result of this section.

Theorem 4: For any balanced chain $\{A(k)\}$ with feedback property, we have $\{A(k)\} \in \mathcal{P}^*$.

Proof: Consider a balanced chain $\{A(k)\}$ with a balancedness coefficient α and a feedback coefficient β. Let $\{t_r\}$ be an increasing sequence of positive integers such that Eq. (2) holds, i.e., $R(k) = \lim_{r\to\infty} A(t_r : k)$. Then, as discussed in Section III, any sequence $\{\hat{\pi}^T R(k)\}$ is an absolute probability sequence for $\{A(k)\}$, where $\hat{\pi}$ is a stochastic vector. Let $\hat{\pi} = \frac{1}{m}e$. Then, by Theorem 3,

$$\frac{1}{m}e^T R(k) = \frac{1}{m}\lim_{r\to\infty} e^T A(t_r : k) \geq p^* e^T,$$

with $p^* = \min(\frac{1}{m}, \gamma^{m-1}) > 0$. Thus, $\{\frac{1}{m}e^T R(k)\}$ is a uniformly bounded absolute probability sequence for $\{A(k)\}$. ■

We conclude this section by considering an implication of Theorem 4 for random chains.

Theorem 5: Let $\{W(k)\}$ be an independent random model with feedback property. Also, let the expected chain $\{\bar{W}(k)\}$ be balanced. Then, $\{W(k)\} \in \mathcal{P}^*$.

Proof: By Lemma 7 in [9], the expected chain has feedback property. Therefore, by Theorem 4, the chain $\{\bar{W}(k)\}$ is balanced and, hence, the result follows. ■

Using Theorem 4, it can be shown that many of the models studied in consensus literature, such as those in [2], [3], [6], [4], [9], as well as Hegselmann-Krause model for opinion dynamics in [5] and its generalization in [17], are instances of chains in $\mathcal{P}^*$. Therefore, the theory and results developed in this section can be applied to them.

V. Implications

In this section, we provide two implications of Theorem 2 for the stability and convergence of deterministic and random weighted averaging dynamics.

Let $\{W(k)\}$ be an independent random chain. If $\{W(k)\}$ is an ergodic chain almost surely, then the ergodicity of its expected chain $\{\bar{W}(k)\}$ follows (Lemma 6 in [9]). However, the converse statement is not necessarily true in general. For example, let $\{W(k)\}$ be an independent identically distributed random chain with each $W(k)$ uniformly distributed over the set of permutation matrices. In this case, $W(k:0)$ does not converge as $k\to\infty$, while $\{\bar{W}(k)\}$ is the ergodic static chain $\{\frac{1}{m}ee^T\}$. Nevertheless, in the following theorem, we show that the converse assertion is true for a broad subclass of independent random chains with feedback property.

Theorem 6: Let $\{W(k)\}$ be an independent random chain with feedback property. Let $\{\bar{W}(k)\}$ be an ergodic chain and let $\lim_{t\to\infty} \bar{W}(t:k) = e\pi^T(k)$ for a stochastic vector $\pi(k)$ and $k \geq 0$. Then, if the sequence $\{\pi(k)\}$ is uniformly bounded by some p^*, then the chain $\{W(k)\}$ is ergodic almost surely.

Proof: As shown in [11], when $\lim_{t\to\infty} \bar{W}(t:k) = e\pi^T(k)$, then $\{\pi(k)\}$ is an absolute probability sequence for $\{\bar{W}(k)\}$. Since the chain $\{W(k)\}$ has feedback property, Theorem 2 implies that the ergodicity of the chain $\{W(k)\}$ and its expected chain are equivalent. ■

Some consensus and ergodicity results for deterministic weighted averaging dynamics rely on uniform boundedness of the positive entries of the averaging matrices and the existence of a periodical connectivity of the graphs associated with the matrices (see [2], [3] and [6]). Using Theorem 6, we can extend these results to independent random chains.

Theorem 7: Let $\{W(k)\}$ be an independent random chain with feedback property. Assume that for the expected chain $\{\bar{W}(k)\}$, there exist a scalar $\alpha > 0$ and an integer $B > 0$ such that the graph $G(k) = ([m], E_B(k))$ is strongly connected for all $k \geq 0$, where $E_B(k) = \{(i,j) \mid \bar{W}_{ij}(t) \geq \alpha, \text{for some } t \in [kB, (k-1)B)\}$. Then, $\{W(k)\}$ is ergodic almost surely.

Proof: When a random chain $\{W(k)\}$ has feedback property with coefficient $\beta > 0$, its expected chain $\{\bar{W}(k)\}$ has the following (strong feedback) property $\bar{W}_{ii}(k) \geq \beta/m$ for all i and k (cf. Lemma 7 in [9]). Then, by Lemma 4 of [4], it follows that the expected chain $\{\bar{W}(k)\}$ is ergodic. It can be shown that $\lim_{t\to\infty} \bar{W}(t:k) = e\pi^T(k) \geq \gamma^m$ for any $k \geq 0$, where $\gamma = \min(\alpha, \beta/m)$ ([3], Lemma 2.1). By Theorem 6, it follows that $\{W(k)\}$ is ergodic a.s. ■

References

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

[2] J. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Dept. of Electrical Engineering and Computer Science, MIT, 1984.

[3] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, pp. 803–812, 1986.

[4] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

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

[6] A. Jadbabaie, J. Lin, and S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[7] A. Olshevsky and J. Tsitsiklis, "On the nonexistence of quadratic lyapunov functions for consensus algorithms," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2642–2645, dec. 2008.

[8] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2506–2517, nov. 2009.

[9] B. Touri and A. Nedić, "On ergodicity, infinite flow and consensus in random models," 2010, to appear in IEEE Transactions on Automatic Control.

[10] B. Touri and A. Nedić, "On approximations and ergodicity classes in random chains," 2010, under review.

[11] A. Kolmogoroff, "Zur theorie der markoffschen ketten," *Mathematische Annalen*, vol. 112, no. 1, pp. 155–160, 1936.

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

[13] L. Bregman, "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 3, pp. 200 –217, 1967.

[14] R. Durrett, *Probability: Theory and Examples*, 3rd ed. Curt Hinrichs, 2005.

[15] D. Spielman, "Spectral graph theory, the laplacian," University Lecture Notes, 2009, available at: http://www.cs.yale.edu/homes/spielman/561/lect02-09.pdf.

[16] J. Hendrickx and J. Tsitsiklis, "Convergence of type-symmetric and cut-balanced consensus seeking systems," 2011, http://arxiv.org/abs/1102.2361.

[17] J. Lorenz, "A stabilization theorem for continuous opinion dynamics," *Physica A: Statistical Mechanics and its Applications*, vol. 355, p. 217223, 2005.