[6] B. Levy and A. J. Krener, "Dynamics and kinematics of reciprocal diffusions," *J. Math. Phys.*, vol. 34, pp. 1846–1875, 1993.
[7] A. J. Krener, "Reciprocal diffusions in flat space," *Probab. Theory Related Fields*, vol. 107, pp. 243–281, 1997.
[8] F. Carravetta, "Nearest-neighbour modelling of reciprocal chains," *Stochastics: Int. J. Prob. Stoch. Processes*, vol. 80, pp. 525–584, 2008.
[9] P. Levy, "A special problem of Brownian motion and a general theory of Gaussian random functions," in *Proc. 3rd Berkeley Symp. Math. Stat. Prob.*, 1956, vol. 2, pp. 133–175.
[10] B. C. Levy, "Noncausal estimation for discrete Gauss-Markov random fields," in *Proc. Int. Symp. Math. Theory Networks Syst. (MTNS-89)*, Boston, MA, 1990, pp. 13–21.
[11] B. C. Levy, M. B. Adams, and A. S. Willsky, "Solution and linear estimation of 2-D nearest-neighbor models," *Proc. IEEE*, vol. 78, no. 4, pp. 627–641, Apr. 1990.
[12] R. Kindermann and J. L. Snell, *Markov Random Fields and Their Applications, Contemporary Mathematics*, ser. AMS Contemporary Mathematic series. Providence, RI: American Math. Soc., 1980.
[13] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 1, pp. 39–55, Jan. 1987.
[14] X. Guyon, *Random Fields on a Network: Modeling, Statistics and Applications*. New York: Springer Verlag, 1991.
[15] D. A. Castanon, B. C. Levy, and A. S. Willsky, "Algorithms for the incorporation of predictive information in surveillance theory," *Int. J. Syst. Sci.*, vol. 16, pp. 367–382, 1985.
[16] A. J. Krener, "Boundary value linear systems," *Asterisque*, vol. 75–76, pp. 149–165, 1980.
[17] A. J. Krener, "Acausal realization theory, Part 1: Linear deterministic systems," *SIAM J. Control Opt.*, vol. 25, pp. 499–525, 1987.

On the Equivalence and Condition of Different Consensus Over a Random Network Generated by i.i.d. Stochastic Matrices

Qingshuo Song, Guanrong Chen, *Fellow, IEEE*, and Daniel W. C. Ho

***Abstract*—Our objective is to find a necessary and sufficient condition for consensus over a random network generated by i.i.d. stochastic matrices. We show that the consensus problem in all different types of convergence (almost surely, in probability, and in L^p for every $p \geq 1$) are actually equivalent, thereby obtain the same necessary and sufficient condition for all of them. The main technique we used is based on the stability in a projected subspace of the concerned infinite sequences.**

***Index Terms*—Consensus, random network, stability, stochastic matrix.**

I. Introduction

We consider a system described by the stochastic linear difference equation

$$X(t) = A(t)X(t-1), \quad t = 1, 2, \ldots$$

Manuscript received April 15, 2010; revised August 17, 2010 and December 20, 2010; accepted January 11, 2011. Date of publication January 20, 2011; date of current version May 11, 2011. This work was supported in part by the Research Grants Council of Hong Kong under Grant CityU 104007, by the Hong Kong Research Grants Council under Grant CityU1117/10E, and by Grant CityU7002561 and Grant CityU7008114. Recommended by Associate Editor P. Shi.

Q. Song and D. W. C. Ho are with the Department of Mathematics, City University of Hong Kong, Hong Kong, China (e-mail: song.qingshuo@cityu.edu.hk; madaniel@cityu.edu.hk).

G. Chen is with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China (e-mail: gchen@ee.cityu.edu.hk).

Digital Object Identifier 10.1109/TAC.2011.2107130

where $\{X(t)\}$ is a sequence of $\mathbb{R}^N$-valued state vectors and $\{A(t)\}$ is a sequence of i.i.d. (independent and identically distributed) right stochastic matrices (non-negative matrices with each row-sum equal to 1).

The system is said to reach consensus if, for any initial state $X_0 \in \mathbb{R}^N$, $\max_{1\leq i,j\leq N} |X_i(t) - X_j(t)|$ converges to zero as $t \to \infty$ in an appropriate sense. Since $X(t)$ is random, there are different types of consensus available such as almost surely consensus, consensus in probability, and consensus in L^p for some $p \geq 1$, among others.

System or network consensus has been extensively studied in the last decade. The consensus problem for discrete systems has wide applications in random networks described by difference equations (see, e.g.,[2], [4]–[15], [17] and some references cited therein). More specifically, [7] uses the Vicsek model [16] to study consensus in the form of group coordination of mobile autonomous agents, [4] investigates information flow and cooperative control of vehicle formations using directed graph theory, [13] discusses consensus with a switching topology and time-delays, [11] and [12] consider consensus seeking in multi-agent systems with dynamically changing interactions, [17] discusses distributed average consensus, and [2] investigated robust consensus from a graph-theoretic approach.

In particular, closely related to the investigation of the present technical note, [5] studies the consensus in probability, [8] studies almost surely consensus and consensus in L^2 of controllable random networks, and [14] establishes an elegant necessary and sufficient condition for almost surely consensus by investigating the ergodicity of a random matrix sequence

$$|\lambda_2 (\mathbb{E}[A(1)])| < 1 \tag{1}$$

where $\mathbb{E}[\cdot]$ is the expectation operator and $\lambda_2(\cdot)$ is the second largest eigenvalue (in absolute value) of the argument matrix.

At this point, it is noted that since almost surely consensus implies consensus in probability, (1) is clearly a sufficient condition for consensus in probability, but the converse may not be true. Moreover, the related concept of consensus in L^p has not been carefully discussed in the literature. Motivated by this observation, in this technical note we first show that the notion of the aforementioned three different types of consensus are actually equivalent (Theorem III.2). Then, we show that (1) is the necessary and sufficient condition for consensus in L^1 (Theorem III.3). Consequently, we conclude that (1) is more valuable than what it was known before—it actually gives a necessary and sufficient condition for consensus in all three different senses. This is the main contribution of our present work.

In addition, by using a completely different methodology in contrast to [14], we show that our result applies to a more general setting: the restriction imposed in [14] on the space of stochastic matrices with strictly positive diagonal entries can be relaxed, as detailed in Remark II.1 below.

The main technique we used in this technical note is that, based on the observation of a relation between consensus and stability, we first convert the original consensus problem on a sequence to the stability problem on a projected sequence in a subspace. Thereafter, we focus our study on the eigenspace structure of the projection operator. We should note that such a method has been used by others in different settings, for instance [17].

The rest of the technical note is organized as follows: We start with the problem formulation in Section II, where we also present a connection between consensus and stability. We then present the main result, namely a necessary and sufficient condition for consensus over a random network, in Section III. Finally, we conclude the investigation with further discussions in Section IV.

II. Problem Formulation

In the first subsection, we present a useful result (Theorem II.1) on the equivalence of consensus and stability in a subspace, under a general framework. This result actually may be applied to some even more general settings, including nonlinear and random sequences. In the second subsection, we formulate the consensus problem based on a linear stochastic difference equation.

Before proceeding, let us recall some standard notations:

1) In the (column) vector space $\mathbb{R}^N$, x_i represents the ith coordinate of a vector $x \in \mathbb{R}^N$; l^q-norm is $\|x\|_q = (\sum_{i=1}^N |x_i|^q)^{1/q}$, $\forall 1 \le q \le \infty$; x^T denotes the transpose of x. $\mathbf{1}$ denotes a column vector with all ones.
2) In the space of square real matrices, $\mathbb{R}^{N\times N}$, I denotes the identity matrix; for all $A \in \mathbb{R}^{N\times N}$, $\|A\|_p = \max_{\|x\|=1} \|Ax\|_p$ for all $1 \le p \le \infty$; the eigenvalues will be arranged in decreasing order as $|\lambda_1(A)| \ge |\lambda_2(A)| \ge \cdots \ge |\lambda_N(A)|$; the spectral radius refers to $\rho(A) = |\lambda_1(A)|$.
3) $\|\cdot\|$ is used in the formula if it is valid for all l^p-norms.
4) Given a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, denote by $\mathbb{E}[\cdot] = \mathbb{E}^{\mathbb{P}}[\cdot]$ the expectation under $\mathbb{P}$; L^p refers to $L^p(\Omega, \mathcal{F}, \mathbb{P})$: for random vector $Y : \Omega \to \mathbb{R}^N$; the L^p-norm is $\|Y\|_{L^p} = (\int_\Omega \|Y(\omega)\|_2^p \mathbb{P}(d\omega))^{1/p} = (\mathbb{E}[(\|Y\|_2)^p])^{1/p}$.

A. Connection Between Consensus and Stability

Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a filtered probability space, where $\mathbb{F} = \{\mathcal{F}_t : t = 0, 1, 2, \ldots\}$ is a sequence of increasing σ-algebras with $\mathcal{F}_\infty \subset \mathcal{F}$. We consider an $\mathbb{F}$-adapted sequence $\{X(t)\}$ taking values in $\mathbb{R}^N$. In other words, $X(t)$ is a measurable mapping from $(\Omega, \mathcal{F}_t) \to (\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, where $\mathcal{B}(\mathbb{R}^N)$ is the Borel σ-algebra on $\mathbb{R}^N$. Such a sequence includes the general form of

$$X(t) = f_t\left(X(t-1), \ldots, X(1)\right)$$

for some measurable function f_t, which emphasizes its independence of future events.

We start from a precise definition of consensus over a random sequence in three different senses, namely, three different types. As usual, $X(t, \omega)$ will be used instead of $X(t)$ whenever we need to emphasize its dependence on a sample path $\omega \in \Omega$.

Definition II.1 (Consensus of a Sequence): Let $\{X(t)\}$ be an $\mathbb{F}$-adapted $\mathbb{R}^N$-valued random sequence. $\{X(t)\}$ is said to reach consensus

1) in probability, if $\forall \varepsilon > 0$

$$\lim_{t\to\infty} \mathbb{P}\left\{\omega \in \Omega : \max_{1\le i,j\le N} |X_i(t,\omega) - X_j(t,\omega)| > \varepsilon\right\} = 0.$$

2) almost surely (with probability 1), if

$$\mathbb{P}\left\{\omega \in \Omega : \lim_{t\to\infty} \max_{1\le i,j\le N} |X_i(t,\omega) - X_j(t,\omega)| = 0\right\} = 1.$$

3) in L^p for some $p \ge 1$, if

$$\lim_{t\to\infty} \mathbb{E}\left[\max_{1\le i,j\le N} |X_i(t) - X_j(t)|^p\right] = 0.$$

Stability of a sequence can be defined analogously for the above random network: $\{X(t)\}$ is said to be *stable* at zero in probability (respectively almost surely, in L^p), if $\|X(t)\| \to 0$ in probability (respectively almost surely, in L^p) as $t \to \infty$.

Let $\mathbb{R}_0$ be an *agreement space* of $\mathbb{R}^N$, defined by $\mathbb{R}_0 = \{x \in \mathbb{R}^N : x_1 = x_2 = \ldots = x_N\}$, and let Π be a projection operator on $\mathbb{R}_0$, namely, $\Pi x = \langle x, v_0\rangle v_0$, $\forall x \in \mathbb{R}^N$, where $\langle\cdot,\cdot\rangle$ denotes the inner product, and $v_0 \in \mathbb{R}_0$ is an l^2-norm unit vector. By orthogonal decomposition, we have $x = \Pi x + \Pi^\perp x$, where $\Pi^\perp = I - \Pi$. In fact, $\Pi^\perp x$ measures the distance to the agreement space from x.

The following theorem shows that the consensus of a sequence in $\mathbb{R}^N$ is equivalent to the stability of the sequence projected onto the subspace $\mathbb{R}_0^\perp$.

Theorem II.1: Sequence $\{X(t)\}$ reaches consensus almost surely (respectively, in probability, or in L^p) if and only if $\{\Pi^\perp X(t)\}$ is stable almost surely (respectively, in probability, or in L^p).

Proof: We only show the equivalence of consensus and stability almost surely. The equivalence in probability and in L^p can be similarly proved.

($\Longrightarrow$) Suppose $\{X(t)\}$ reaches consensus almost surely. Define $Y(t) \in \mathbb{R}_0$ be a vector with all entries equal to the value of the first coordinate of $X(t)$, namely, $Y(t) = (X_1(t), \ldots, X_1(t))^T$ We have

$$\begin{aligned}\left\|\Pi^\perp X(t)\right\|_\infty &= \min\nolimits_{y\in\mathbb{R}_0} \|X(t) - y\|_\infty \\ &\le \|X(t) - Y(t)\|_\infty \\ &= \max_{1\le i\le N} |X_i(t) - X_1(t)| \\ &\le \max_{1\le i,j\le N} |X_i(t) - X_j(t)| \\ &\to 0 \quad \text{as } t \to \infty \text{ almost surely.}\end{aligned}$$

Therefore, $\{\Pi^\perp X(t)\}$ is stable almost surely.

($\Longleftarrow$) Suppose $\{\Pi^\perp X(t)\}$ is stable almost surely. Let $(\Pi X)_i(t)$ be the ith coordinate of vector $\Pi X(t)$. Note that, since $\Pi X(t) \in \mathbb{R}_0$, all coordinates have the same value, that is, $(\Pi X)_i(t) = (\Pi X)_j(t)$, $\forall i, j$. Consequently, by the triangular inequality, we have

$$\begin{aligned}&\max\nolimits_{i,j} |X_i(t) - X_j(t)| \\ &\quad\le \max\nolimits_{i,j} \left(|X_i(t) - (\Pi X)_i(t)| + |(\Pi X)_j(t) - X_j(t)|\right) \\ &\quad\le \max\nolimits_i |X_i(t) - (\Pi X)_i(t)| + \max\nolimits_j |(\Pi X)_j(t) - X_j(t)| \\ &\quad\le 2\,\|X(t) - \Pi X(t)\|_\infty \\ &\quad= 2\left\|\Pi^\perp X(t)\right\|_\infty \\ &\quad\to 0 \quad \text{as } t \to \infty \text{ almost surely.}\end{aligned}$$

Therefore, $\{X(t)\}$ reaches consensus almost surely. ■

B. A Random Network Generated by i.i.d Stochastic Matrices

We consider a setting similar to [14].

Denote the space of $N \times N$ stochastic matrices by

$$S_N = \left\{A = (a_{ij})_{N\times N} : a_{ij} \ge 0, \sum_{j=1}^N a_{ij} = 1, \forall i, j\right\}. \quad (2)$$

Let $\mathcal{B}(S_N)$ be the Borel σ-algebra on S_N, μ be a given probability distribution on $(S_N, \mathcal{B}(S_N))$, and $\{A(t)\}$ be an S_N-valued i.i.d. sequence with distribution μ. Moreover, let $\Omega = (S_N)^\infty$, $\mathbb{P} = \mu \times \mu \times \cdots$, and $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_t = \sigma(A(1), A(2), \ldots, A(t))$ for $t \ge 1$, with $\mathbb{F} = \cup_{t=0}^\infty \mathcal{F}_t$.

Now, we consider a system given by a random sequence $\{X(t)\}$ via

$$X(t) = A(t)X(t-1), \quad \forall t \in \mathbb{N}. \quad (3)$$

Clearly, $\{X(t)\}$ is an $\mathbb{F}$-adapted sequence in the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. Sometimes, we write $X^x(t)$ to emphasize its dependence on the initial state $X(0) = x$.

Also, note that the distribution of $X(t)$ depends on the distribution μ. We say that $\{X(t)\}$ generated by μ reaches consensus almost surely (respectively, in probability, or in L^p), if $\{X^x(t)\}$ in (3) reaches consensus almost surely (respectively, in probability, or in L^p) for all initial states $x \in \mathbb{R}^N$. Similarly, we can also define the stability for $\{X(t)\}$ generated by μ.

Remark II.1: The space S_N in (2) here is larger than the one formulated in [14], where

$$\hat{S}_N = \{A \in S_N : \text{all diagonal entries are strictly positive}\}$$

Indeed, such a diagonal restriction is crucial in the proof of [14, Theorem 3] for utilizing [1, Perron-Frobenius Theorem] on primitive matrices. In our formulation here, however, the diagonal entries are allowed to be zero. The key idea to get rid of the diagonal restriction is the use of an appropriate projection method where the important mathematical arguments are off-diagonal. Clearly, our results include those of [14] as a special case with $\mu(S_N \setminus \hat{S}_N) = 0$. □

Our next objective is to find a necessary and sufficient condition for consensus of the random sequence generated by μ.

III. Necessary and Sufficient Condition for a Random Sequence

First, we recall some properties of stochastic matrices and consensus in the deterministic case, which will be useful in the rest of the technical note. Then, by studying the fine structure of the random sequence generated by i.i.d. stochastic matrices, we show that almost surely consensus, consensus in probability and consensus in L^p for all $p \geq 1$ classified by Definition II.1 are in fact equivalent to each other. To that end, we will only discuss the case of L^1 consensus using stability on its projected subspace, which in turn implies the same result to other types of consensus by the equivalence.

A. Necessary and Sufficient Condition for a Deterministic Sequence

A deterministic system can be treated as a special case of a random system in the following sense. Let the probability distribution μ on S_N satisfy $\mu(\{A\}) = 1$ for stochastic matrices $A \in S_N$. Then, $A(t) = A$ for all $t = 1, 2, \ldots$, and the sequence $\{X(t)\}$ of (3) becomes deterministic. Consequently

$$X(t) = AX(t-1), \quad \forall t = 0, 1, 2, \ldots$$

In this case, we say that the $\{X(t)\}$ generated by A reaches consensus if $\{X^x(t)\}$ reaches consensus for all initial states $x \in \mathbb{R}^N$.

For this deterministic setting, all the three types of consensus are equivalent since the sample space Ω can be treated as a singleton $\{A\} \times \{A\} \times \cdots$, so this definition is consistent for all the three types with the distribution μ of the form $\mu(\{A\}) = 1$.

We first recall some properties of such matrices. Since each row-sum of such a matrix is equal to 1, its largest eigenvalue is $\rho(A) = \lambda_1(A) = 1$, namely, $Ax = x$ for all $x \in \mathbb{R}_0$. Also, $\|A^t x\|_\infty \leq \|x\|_\infty, \forall x \in \mathbb{R}^N$, $t \in \mathbb{N}$. We next present the following useful properties of stochastic matrices:

Proposition III.1: Let $A \in S_N$ be given. Then

1) Matrix A has a Jordan canonical form:

$$\Lambda = \begin{bmatrix} 1 & 0_{1\times(N-1)} \\ 0_{(N-1)\times 1} & \Lambda_{22} \end{bmatrix} \tag{4}$$

where $0_{m\times n}$ is an $m \times n$ matrix with each entry being zero, and Λ_{22} is a sub-matrix in the Jordan form.

2) The linear operator $\Pi^\perp$ satisfies

$$\Pi^\perp A = \Pi^\perp A \Pi^\perp. \tag{5}$$

3) The matrix $\Pi^\perp A$ has a Jordan form: $\Lambda_0 = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_{22} \end{bmatrix}$ with Λ_{22} defined in (4). In particular, $\rho(\Pi^\perp A) = |\lambda_2(A)|$.

Proof: The first and third properties are standard. Also, due to the fact of $A\Pi x = \Pi x$, for all $x \in \mathbb{R}^N$, we have

$$\begin{aligned} \Pi^\perp Ax &= \Pi^\perp A(\Pi x + \Pi^\perp x) = \Pi^\perp A\Pi x + \Pi^\perp A \Pi^\perp x \\ &= \Pi^\perp \Pi x + \Pi^\perp A\Pi^\perp x = \Pi^\perp A \Pi^\perp x \end{aligned}$$

thus the second property (5) follows. ■

Another well-known fact is a necessary and sufficient condition for stability: $A \in \mathbb{R}^{N\times N}$ is stable if and only if $\rho(A) < 1$. Based on Proposition III.1 and this stability result, we are now ready to derive a necessary and sufficient condition for the consensus of a deterministic system.

Theorem III.1: Let $A \in S_N$ be given. Then, $\{X(t)\}$ generated by A reaches consensus if and only if $|\lambda_2(A)| < 1$.

Proof: By Theorem II.1, $\{X(t)\}$ generated by A reaches consensus if and only if $\{\Pi^\perp X(t)\}$ is stable. Using (5) and $A\Pi x = \Pi x$, for any initial state $x \in \mathbb{R}^N$, we have

$$\begin{aligned} \Pi^\perp X(t) &= \Pi^\perp AX(t-1) = (\Pi^\perp A)\Pi^\perp X(t-1) \\ &= \cdots = (\Pi^\perp A)^t \Pi^\perp x = (\Pi^\perp A)^t x. \end{aligned} \tag{6}$$

Thus, $\{\Pi^\perp X(t)\}$ is a sequence generated by $\Pi^\perp A$. It follows from the stability result that $\{\Pi^\perp X(t)\}$ is stable if and only if $\rho(\Pi^\perp A) < 1$. Observe that Proposition III.1 implies $\rho(\Pi^\perp A) = |\lambda_2(A)|$. This completes the proof. ■

B. Equivalence Among Different Types of Consensus

We first recall some relations about the convergence of random variables in different senses, and we refer to [3] for more detail.

Consider a sequence of random variables $\{a_n, n = 1, 2, \ldots\}$ and a random variable $a \geq 0$. Let $p \geq 1$ is arbitrary number. Both almost surely convergence and convergence in L^p imply convergence in probability, namely, $a_n \to a$ almost surely implies $a_n \to a$ in probability, $a_n \to a$ in L^p implies $a_n \to a$ in probability. However, the reverse directions need further conditions in general. More precisely, $a_n \to a$ in probability together with uniform integrability of $|a_n|^p$ implies $a_n \to a$ in L^p, and $a_n \to a$ in L^p together with monotonicity $0 \leq a_{n+1} \leq a_n$ implies $a_n \to a$ almost surely.

Lemma III.1: Consider the sequence $\{X^x(t)\}$ defined in (3) generated by distribution μ with a given initial state $X(0) = x$. The following statements on the stability of $\{\Pi^\perp X^x(t)\}$ are equivalent:

1) $\{\Pi^\perp X^x(t)\}$ is stable in probability;
2) $\{\Pi^\perp X^x(t)\}$ is stable in L^p for any $p \geq 1$;
3) $\{\Pi^\perp X^x(t)\}$ is stable almost surely.

Proof: Within this proof, we use $X(t)$ to denote $X^x(t)$ for simplicity. Observe that, by (5), $\{\Pi^\perp X(t)\}$ is a sequence generated by the random matrix $\Pi^\perp A(t)$, namely

$$\Pi^\perp X(t) = \Pi^\perp A(t) \Pi^\perp X(t-1). \tag{7}$$

In the following, we prove the equivalence by showing 1) implies 2), 2) implies 3), 3) implies 1), respectively.

1) If the sequence $\{\Pi^\perp X(t)\}$ is stable in probability, then $\|\Pi^\perp X(t)\|_\infty \to 0$ in probability. Together with the uniform boundedness $\|\Pi^\perp X(t)\|_\infty \leq \|x\|_\infty$, the dominated convergence theorem implies that $\|\Pi^\perp X(t)\|_\infty \to 0$ in L^p. Thus, the sequence $\{\Pi^\perp X(t)\}$ is stable in L^p.
2) If the sequence $\{\Pi^\perp X(t)\}$ is stable in L^p, then $\|\Pi^\perp X(t)\|_\infty \to 0$ in L^p. In addition, we can show the monotonicity of

$\|\Pi^\perp X(t)\|_\infty \leq \|\Pi^\perp X(t-1)\|_\infty$, by observing

$$\begin{aligned}\left\|\Pi^\perp X(t)\right\|_\infty &= \left\|\Pi^\perp A(t) X(t-1)\right\|_\infty \\ &= \left\|\Pi^\perp A(t)\Pi^\perp X(t-1)\right\|_\infty \\ &\leq \left\|\Pi^\perp X(t-1)\right\|_\infty \end{aligned} \quad (8)$$

L^p convergence of a non-negative monotone sequence implies that $\|\Pi^\perp X(t)\|_\infty \to 0$ almost surely.

3) It is well known that almost surely convergence implies convergence in probability. ■

We are now in a position to establish our main result about the equivalent consensus in three different senses.

Theorem III.2: Consider the sequence $\{X(t)\}$ defined in (3) generated by distribution μ. The following statements on consensus are equivalent:

1) $\{X(t)\}$ reaches consensus in probability.
2) $\{X(t)\}$ reaches consensus in L^p for any $p \geq 1$.
3) $\{X(t)\}$ reaches consensus almost surely.

Proof: They follow immediately from Theorem II.1 and Lemma III.1. ■

C. Necessary and Sufficient Condition for a Random Sequence

Theorem III.2 shows the equivalence of consensus in different probability senses. As a result, to find a necessary and sufficient condition for consensus, it is enough to study consensus in any one of probability senses. In this subsection, we specifically choose L^1 consensus for the convenience, which will imply any other types of consensus.

Recall that $\{X^x(t)\}$ generated by μ reaches consensus in L^1, if $\mathbb{E}[\|\Pi^\perp X^x(t)\|_q] \to 0$ for arbitrary $q \geq 1$, due to the equivalence of l^q-norms in $\mathbb{R}^N$. Throughout this subsection, we will use different values of q ($= 1$ or ∞) when it is needed.

Also, $\{X(t)\}$ generated by μ reaches consensus if the sequence $\{X^x(t)\}$ reaches consensus for all initial states $x \in \mathbb{R}^N$. The next proposition shows that it is sufficient to check the consensus of $\{X^x(t)\}$ only for all $x \in (\mathbb{R}^+)^N$ to guarantee the consensus of the system.

Proposition III.2: $\{X(t)\}$ generated by distribution μ reaches consensus if and only if $X^x(t)$ defined in (3) reaches consensus for all $x \in (\mathbb{R}^+)^N$.

Proof: Observe that $X^{x+c}(t) = X^x(t) + c$ for all $c \in \mathbb{R}_0$. Hence, $\forall c \in \mathbb{R}_0$

$$\max_{1\leq i,j\leq N} \left|X_i^x(t,\omega) - X_j^x(t,\omega)\right| = \max_{1\leq i,j\leq N}\left|X_i^{x+c}(t,\omega) - X_j^{x+c}(t,\omega)\right|.$$

In other words, to consider the consensus of $\{X^x(t)\}$ for some $x \notin (\mathbb{R}^+)^N$, we can investigate the consensus of $X^{x+c}(t)$ by taking $c = (\|x\|_\infty, \ldots, \|x\|_\infty)^T \in \mathbb{R}_0$. Note that $x + c \in (\mathbb{R}^+)^N$, hence the result follows. ■

Next, we review some useful properties of the expectation operator $\mathbb{E}$. First, $\mathbb{E}$ is commutative with any deterministic matrix. In particular, we have $\Pi^\perp \mathbb{E}[Y] = \mathbb{E}[\Pi^\perp Y]$ for any random vector Y. Furthermore, if A is a random matrix, $\mathbb{E}[AY] = \mathbb{E}[A]\mathbb{E}[Y]$ holds, provided that the random vector $Y : \Omega \to \mathbb{R}^N$ is independent of A. Finally, note that the deterministic sequence $\{\mathbb{E}[X(t)]\}$ is actually a sequence generated by the deterministic matrix $\mathbb{E}[A(1)]$, since we have

$$\mathbb{E}[X(t)] = \mathbb{E}[A(t)]\,\mathbb{E}[X(t-1)] = \mathbb{E}[A(1)] \cdot \mathbb{E}[X(t-1)]. \quad (9)$$

Theorem III.3: $\{X(t)\}$ defined in (3) generated by distribution μ reaches consensus almost surely (also, in probability and in L^p for any $p \geq 1$) if and only if $|\lambda_2(\mathbb{E}^\mu[A(1)])| < 1$.

Proof: As mentioned above, we will not distinguish the three different types of consensus generated by μ.

($\Longrightarrow$) If $\{X(t)\}$ reaches consensus, then Theorem II.1 implies that $\Pi^\perp X(t) \to 0$ in L^1 for all initial states $X(0) = x$, hence $\mathbb{E}[\|\Pi^\perp X(t)\|_\infty] \to 0$. Next, Jensen's inequality, we have

$$\begin{aligned}\left\|\Pi^\perp \mathbb{E}[X(t)]\right\|_\infty &= \left\|\mathbb{E}\left[\Pi^\perp X(t)\right]\right\|_\infty \\ &\leq \mathbb{E}\left[\left\|\Pi^\perp X(t)\right\|_\infty\right] \to 0.\end{aligned}$$

This implies that the deterministic sequence $\{\Pi^\perp \mathbb{E}[X(t)]\}$ is stable. Thus, using (9) and (5), we have

$$\begin{aligned}\Pi^\perp \mathbb{E}[X(t)] &= \Pi^\perp \mathbb{E}[A(1)] \cdot \mathbb{E}[X(t-1)] \\ &= \Pi^\perp \mathbb{E}[A(1)] \cdot \Pi^\perp \mathbb{E}[X(t-1)].\end{aligned}$$

This means that $\{\Pi^\perp \mathbb{E}[X(t)]\}$ is a deterministic sequence generated by matrix $\Pi^\perp \mathbb{E}[A(1)]$. Therefore, by Proposition III.1 together with the stability result, $\rho(\Pi^\perp \mathbb{E}[A(1)]) = |\lambda_2(\mathbb{E}[A(1)])| < 1$.

($\Longleftarrow$) It follows from (9) that the deterministic sequence $\{\mathbb{E}[X(t)]\}$ is generated by matrix $\mathbb{E}[A(1)]$. If $|\lambda_2(\mathbb{E}[A(1)])| < 1$, then by applying Theorem III.1 on (9), we conclude that the sequence $\{\mathbb{E}[X(t)]\}$ reaches consensus. Hence, the deterministic sequence $\{\Pi^\perp \mathbb{E}[X(t)] = \mathbb{E}[\Pi^\perp X(t)]\}$ is stable by Theorem II.1, i.e., $\|\mathbb{E}[\Pi^\perp X(t)]\|_1 \to 0$ By Proposition III.2, we can always assume $x \in (\mathbb{R}^+)^N$. Thus, $\Pi^\perp X(t) \in (\mathbb{R}^+)^N$, and this leads to

$$\mathbb{E}\left[\left\|\Pi^\perp X(t)\right\|_\infty\right] \leq \mathbb{E}\left[\left\|\Pi^\perp X(t)\right\|_1\right] = \left\|\mathbb{E}\left[\Pi^\perp X(t)\right]\right\|_1 \to 0. \quad (10)$$

In other words, $\{\Pi^\perp X(t)\}$ is stable in L^1. This implies the consensus of $\{X(t)\}$ by Theorem II.1. ■

Remark III.1: In (10), we have used the fact that, for all $Y : \Omega \to (\mathbb{R}^+)^N$

$$\mathbb{E}[\|Y\|_1] = \mathbb{E}\left[\sum_{i=1}^N Y_i\right] = \sum_{i=1}^N \mathbb{E}[Y_i] = \|\mathbb{E}Y\|_1.$$

However, one shall not expect the identity $\mathbb{E}[\|Y\|_\infty] = \|\mathbb{E}[Y]\|_\infty$ holds in general. For instance, if $Y(\omega_1) = (0,1)^T$, $Y(\omega_2) = (1,0)^T$, and $\mathbb{P}(\{\omega_1\}) = \mathbb{P}(\{\omega_2\}) = 1/2$, then the inequality is strict: $\mathbb{E}[\|Y\|_\infty] = 1 > (1/2) = \|\mathbb{E}[Y]\|_\infty$. This is the reason why we use the l^1-norm in (10) instead of directly but incorrectly using $\mathbb{E}[\|\Pi^\perp X(t)\|_\infty] = \|\mathbb{E}[\Pi^\perp X(t)]\|_\infty \to 0$

IV. Further Discussions

In this technical note, we have derived a necessary and sufficient condition for consensus over a random network generated by i.i.d. stochastic matrices based on the connection between consensus and stability. To close up, a few important remarks are in order.

Firstly, we note that even for a second-order random network, one can also utilize the results of this work. More precisely, for

$$X(t) = \alpha A(t) X(t-1) + \beta B(t) X(t-2), \quad t = 2, 3, \ldots$$

where $\alpha + \beta = 1$, $\alpha, \beta \geq 0$, $\{A(t)\}$ and $\{B(t)\}$ are i.i.d. stochastic matrix sequences with given distributions μ_A and μ_B on S_N, respectively, the problem is equivalent to

$$Y(t) = C(t)Y(t-1)$$

where $Y(t) = \left[\begin{smallmatrix} X(t) \\ X(t-1) \end{smallmatrix}\right]$ is an $\mathbb{R}^{2N}$-vector and $C(t) = \begin{bmatrix} \alpha A(t) & \beta B(t) \\ I & 0 \end{bmatrix}$ is a stochastic matrix in S_{2N}.

Secondly, we note that [8] also studies almost surely consensus and consensus in L^2 of controllable random networks. The difference between [8] and this paper is as follows: [8] considers average consensus on a space of doubly stochastic matrices, hence that all components of $\{X(t)\}$ converge into an average number of $X(0)$ *a priori* (see [8, Eq. (13)]), that is, $\lim_{t\to\infty} X(t) = X_{avg}\mathbf{1}$, where $X_{avg} = (1/N)(X_1(0) + \cdots + X_N(0))$. In our technical note, however, we did not assume such a constant limit in a general stochastic matrix space. In fact, the following example shows that one cannot assume a constant limit in our consensus problem setting.

Example IV.1: Let $N = 2$ and suppose $A(t)$ follows an i.i.d. distribution:

$$\mathbb{P}\left\{A(t) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\right\} = \mathbb{P}\left\{A(t) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right\} = \frac{1}{2}.$$

Given initially $X(0) = (1,3)^T$, we have $X_{avg} = 2$. Although $X(t) = A(t)X(t-1)$ immediately reaches consensus (in all three senses) by taking either $(1,1)^T$ or $(3,3)^T$ for all $t \geq 1$, it does not have to satisfy $\lim_{t\to\infty} X(t) = (2,2)^T$ in any sense. In fact

$$\mathbb{P}\left\{\lim_{t\to\infty} X(t) = (1,1)^T\right\} = \mathbb{P}\left\{\lim_{t\to\infty} X(t) = (3,3)^T\right\} = \frac{1}{2}.$$

. □

As a result, due to the difference in the problem setting, the entire proof on consensus in [8] relies on the estimation of $\|X(t) - X_{avg}\mathbf{1}\|$ (see [8, Eq. (16) and Eq. (19)]), which is not applicable to our consensus setting, as shown by Example IV.1

Thirdly, it should be noted that although our approach and theory are verified for the discrete-time setting, similar results still hold under the continuous-time framework. Moreover, one can similarly follow the same procedure to obtain consensus conditions based on the stability results of [18] on hybrid switching continuous-time systems.

Last but not least, it is worth mentioning that Kolmogorov's 0-1 law [3], referred by [14] in the context of ergodicity of i.i.d. matrix sequences, are closely related to the present study of consensus and stability problems. Indeed, now we have known that $X^x(t)$ defined in (3) does not reach consensus when $|\lambda_2(\mathbb{E}[A(1)])| = 1$ for a given distribution μ in any of the three senses. In other words

$$\mathbb{P}\left\{\omega \in \Omega : X^x(t,\omega) \text{reaches consensus for all } x \in \mathbb{R}^N\right\} < 1.$$

So, a natural question is: what is the above probability when $|\lambda_2(\mathbb{E}[A(1)])| = 1$? The answer is surprisingly simple: zero. This result can be summarized as follows:

Consider $X^x(t)$ defined in (3) generated by i.i.d. matrices with distribution μ. Then

$$\mathbb{P}\left\{\omega \in \Omega : X^x(t,\omega) \text{ reaches consensus for all } x \in \mathbb{R}^N\right\}$$

is either 1 or 0.

A proof of this result can be accomplished similarly to the proof of [14, Lemma 1], by using the tail σ-field argument on decreasing events of the form

$$B_k = \left\{\omega : \Pi_{t=k}^{\infty} A(t)x \text{ reaches consensus for all } x \in \mathbb{R}^N\right\}.$$

To this end, the investigation of the technical note is completed.

Acknowledgment

The authors would like to thank the referees for insightful comments and suggestions that led to significant improvements in the presentation of the technical note.

References

[1] A. Berman and R. J. Plemmons*, Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 1994, vol. 9.
[2] M. Cao, A. S. Morse, and B. D. O. Anderson, "Reaching a consensus in a dynamically changing environment: A graphical approach," *SIAM J. Control Optim.*, vol. 47, no. 2, pp. 575–600, Mar. 2008.
[3] R.*, Durrett Probability: Theory and Examples*, 3rd ed. Pacific Grove, CA: Wadsworth & Brooks/Cole Advanced Books & Software, 2005.
[4] A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1465–1476, Sep. 2004.
[5] Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1867–1872, Nov. 2005.
[6] H.-C. Hsieh, J.-S. Leu, and W.-K. Shih, "Reaching consensus underlying an autonomous local wireless sensor network," *Int. J. Innovative Comp., Inform. Control*, vol. 6, no. 4, pp. 1905–1914, 2010.
[7] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 985–1001, Jun. 2003.
[8] S. Kar and J. M. F. Moura, "Sensor networks with random links: Topology design for distributed consensus," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pt. 2, pp. 3315–3326, Jul. 2008.
[9] C.-L. Liu and F. Liu, "Asynchronously-coupled consensus of second-order dynamic agents with communication delay," *Int. J. Innovative Comp., Inform. Control*, vol. 6, no. 11, pp. 5035–5046, 2010.
[10] H. Nakano, A. Utani, A. Miyauchi, and H. Yamamoto, "An efficient data gathering scheme using a chaotic PCNN in wireless sensor networks with multiple sink nodes," *ICIC Express Lett.*, vol. 3, no. 3 (B), pp. 805–812, 2009.
[11] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.
[12] W. Ren and R. W. Beard*, Distributed Consensus in Multi-Vehicle Cooperative Control*. London, U.K.: Springer-Verlag, 2008.
[13] R. O. Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.
[14] A. Tahbaz-Salehi and A. Jadbabaie, "A necessary and sufficient condition for consensus over random networks," *IEEE Trans. Autom. Control*, vol. 53, no. 3, pp. 791–795, Apr. 2008.
[15] A. Tahbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 225–230, Jan. 2010.
[16] T. Vicsek, A. Cziok, E. B. Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Phys. Rev. Lett.*, vol. 75, no. 6, pp. 1226–1229, Aug. 1995.
[17] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel Distrib. Comp.*, vol. 67, no. 1, pp. 33–46, Dec. 2007.
[18] C. Zhu, G. Yin, and Q. S. Song, "Stability of random-switching systems of differential equations," *Quart. Appl. Math.*, vol. 67, no. 2, pp. 201–220, Feb. 2009.