IV. Numerical Example

Consider the case when $\alpha = 100$ and

$$\left[\begin{array}{c|c|c} A_1 & B_{01} & B_1 \\ \hline C_1 & D_{01} & D_1 \\ \hline A_2 & B_{02} & B_2 \\ \hline C_2 & D_{02} & D_2 \end{array}\right] = 10^{-3} \times \left[\begin{array}{c|c|c|c} 963.8554 & 96.3855 & 13.5542 & -0.3614 \\ \hline -361.4458 & 963.8554 & 135.5422 & -3.6145 \\ \hline 963.8554 & 96.3855 & 638.5542 & -0.3614 \\ \hline 981.5951 & 98.1595 & 6.9018 & -0.1840 \\ -184.0491 & 981.5951 & 69.0184 & -1.8405 \\ \hline 981.5951 & 98.1595 & 631.9018 & -0.1840 \end{array}\right]$$

$$x_{2k+1}(0) = \begin{bmatrix} -0.5 \\ -0.5 \end{bmatrix},\ x_{2k+2}(\alpha) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \text{ for } k = 0, 1, \ldots$$

$$y_0(p) = 2 \quad \text{for} \quad p = 0, 1, \ldots, 100$$

$$\left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline D_{11} & D_{12} \end{array}\right] = \left[\begin{array}{c|c} 0.10 & 0.20 \\ -0.20 & -0.10 \\ \hline 0.20 & -0.20 \end{array}\right]$$

Fig. 2 shows the disturbance signals. Finally, the reference signal is shown in Fig. 3 (upper plot). The linear matrix inequality of Theorem 2 is feasible in this case and

$$\begin{aligned} K_e &= [1557.2200 \quad 266.6667] \\ [K_{e1}|K_{e2}] &= [206.6831| - 77.4039] \\ K_f &= [3154.5404|533.3333] \\ [K_{f1}|K_{f2}] &= [369.7512| - 152.0100]. \end{aligned}$$

The lower plot in Fig. 3 confirms that the overall design task is achieved and the next stage would be to attempt to tune the design.

V. Conclusions

The major contributions in this short paper are i) the application of lifting techniques to transform the bi-directional dynamics into those of an equivalent uni-directional repetitive process model and hence the availability of a stability theory and control law design to achieve this basic property, and ii) the first results on stability plus performance in the case when there are disturbances present, which are assumed to be periodic over twice the pass length. Also there is clearly much work to do before these results can be evaluated on physical examples. This includes a wide range of algorithms for control law design, robustness analysis, and allowing for more general disturbance terms. Also there could well be alternative lifting type approaches, such as that proposed in [3] for iterative learning control, which may be advantageous for bi-directional repetitive processes. Finally, extending the control design analysis to allow for different reference signals in each direction needs to be addressed.

References

[1] E. Rogers, K. Galkowski, and D. H. Owens, *Control Systems Theory and Applications for Linear Repetitive Processes*, ser. Lecture Notes in Control and Information Sciences. Berlin, Germany: Springer-Verlag, 2007, vol. 349.

[2] M. C. de Oliveira, J. Bernussou, and J. C. Geromel, "A new discrete-time robust stability condition," *Syst. Control Lett.*, vol. 37, no. 4, pp. 261–265, 1999.

[3] J. J. Hatonen, D. H. Owens, and K. L. Moore, "An algebraic approach to iterative learning control," *Int. J. Control*, vol. 77, no. 1, pp. 45–54, 2004.

Stochastic Consensus Seeking With Noisy and Directed Inter-Agent Communication: Fixed and Randomly Varying Topologies

Minyi Huang and Jonathan H. Manton

***Abstract*—We consider consensus seeking of networked agents on directed graphs where each agent has only noisy measurements of its neighbors' states. Stochastic approximation type algorithms are employed so that the individual states converge both in mean square and almost surely to the same limit. We further generalize the algorithm to networks with random link failures and prove convergence results.**

***Index Terms*—Consensus problems, directed graphs, measurement noise, random link failure, stochastic approximation.**

I. Introduction

A fundamental problem in the study of spatially distributed multi-agent systems is the so-called consensus problem, which is crucial for coordinating distributed agents to achieve a group objective. In a generic setting, suppose the kth agent has state x_t^k at time t. Then the primary feature of consensus seeking is for each agent to adjust its own state x_t^k, based on data received from its neighbors, in an endeavor to make all agents' states converge to the same value. Due to their crucial role in such distributed systems, consensus problems and various closely related formulations have been intensively investigated in the context of multi-agent control systems and distributed computing [11], [19], [23]; see [22], [27] for a comprehensive survey on recent research.

In discrete time consensus models [4], [19], [22], typically each node updates its state x_t^k by the rule

$$x_{t+1}^k = \sum_{i \in \mathcal{N}_{kt} \cup \{k\}} a_t^{ki} x_t^i, \qquad t \geq 0$$

where the weights $a_t^{ki} > 0$, $i \in \mathcal{N}_{kt} \cup \{k\}$, add up to one, and $\mathcal{N}_{kt}$ is the set of neighbors of node k at t. The problem formulation may involve asynchronous state updates, dynamic topologies and communication link failures (see [27]). So far, most existing algorithms assume exact state averaging, which in general necessitates perfect state exchange among the agents. There are relatively few works (see e.g. [2], [28], [35]) considering averaging with the presence of noise or disturbance. In [7], the effect of logarithmic quantization error was analyzed. In the early work [5], [32], consensus problems were studied in a stochastic setting, but the exchange of random messages between the agents was assumed to be error-free.

Since information exchange within networks typically involves quantization, wireless channels and/or sensing, perfect state exchange is often impractical. The communication or sensing noise issue also

Manuscript received January 07, 2008; revised August 14, 2008, January 29, 2009, and June 27, 2009. First published December 01, 2009; current version published January 13, 2010. This paper was presented in part at the American Control Conference, 2008. Recommended by Associate Editor Z. Qu.

M. Huang is with the School of Mathematics and Statistics, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: mhuang@math.carleton.ca).

J. H. Manton is with the Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, Victoria 3010, Australia (e-mail: j.manton@ieee.org).

Color versions of one or more of the figures in this technical note are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2009.2036291

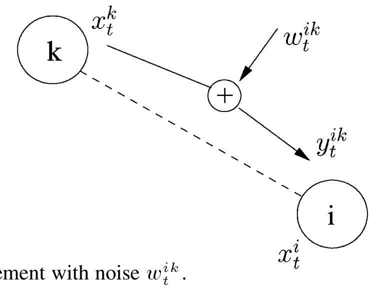

Fig. 1. Measurement with noise w_t^{ik}.

arises in distributed function computation in sensor networks [12] and formation control [1].

When each agent i can only obtain a noisy version of its neighbor's state x_t^k, traditional consensus algorithms involving weights lowered bounded above zero in general cannot ensure convergence. On the other hand, however, in the recursive computation of stochastic systems for the purpose of either function optimization, root-finding, or system identification, it has long been known that a properly decreasing gain for the correction term is essential for obtaining convergence, and there exists a vast literature on these broad areas [3], [6], [8]. Inspired by past advances in stochastic adaptive algorithms, in [16]–[18], a stochastic approximation type algorithm with a decreasing step size was proposed for consensus seeking where the data transmitted from other agents are corrupted by white noise (see Fig. 1). Almost sure (i.e., probability one) convergence results were obtained in [16] via a double array analysis on directed graph (also called digraph) models satisfying a circulant invariance property. Mean square convergence was proved for connected undirected graphs by constructing a stochastic Lyapunov function [17], and this approach was generalized to strongly connected digraphs in [15].

Compared with [18], this note makes the following contributions. First, for both mean square and almost sure convergence, the network topology condition is much weaker by assuming the existence of a spanning tree. Second, for almost sure convergence, our current proving technique can identify the weakest noise conditions. Third, convergence results are proved with both measurement noise and random independent link failures; for related random graph based modeling, see [14], [20], [24], [25], [31], [34]. Concerning the method of analysis, we exploit the fact that the coefficient matrix in the algorithm has exactly one unstable eigenvalue, which is zero, and transform the state recursion into two decoupled parts: a one dimensional real-valued random walk and an $n-1$ dimensional stable linear stochastic approximation model. For treating the model with random link failures, a perturbed stochastic Lyapunov analysis is developed. But on the other hand, we also mention that the approach of double array analysis in [18] has its own advantage when dealing with simultaneous switching network topologies and poor noise conditions (for instance, no finite second moment exists).

The organization of this note is as follows. We introduce the consensus formulation in Section II. Section III gives the equivalent state space model and some preliminary lemmas. In Section IV, we prove mean square and almost sure convergence. Section V considers random link failures. Numerical simulations are presented in Section VI. Section VII presents conclusions.

II. The Stochastic Consensus Problem and Algebraic Preliminaries

Consider n agents distributed according to a digraph $G=(\mathcal{N},\mathcal{E})$ consisting of a set of nodes $\mathcal{N}=\{1,2,\ldots,n\}$ and a set of directed edges $\mathcal{E}\subset\mathcal{N}\times\mathcal{N}$. For brevity, a directed edge will be simply called an edge. An edge from node i to node j is denoted as an ordered pair (i,j) where $i\neq j$ (so there is no edge between a node and itself). A directed path (simply called a path, from i_1 to i_l) consists of a sequence of nodes $i_1,i_2,\ldots,i_l$, $l\geq 2$, such that $(i_k,i_{k+1})\in\mathcal{E}$ for $k=1,\ldots,l-1$. The digraph G is strongly connected if from each node to any other node there exists a path. A directed tree is a digraph where each node, except the root node, has exactly one parent node. Hence, from the root node to any other node there exists a path. The digraph G is said to contain a spanning tree $G_s=(\mathcal{N}_s,\mathcal{E}_s)$ if G_s is a directed tree such that $\mathcal{N}_s=\mathcal{N}$ and $\mathcal{E}_s\subset\mathcal{E}$. A strongly connected digraph always contains a spanning tree. For convenience of exposition, the two names, agent and node, will be used interchangeably. The agent A_k (resp., node k) is a neighbor of A_i (resp., node i) if $(k,i)\in\mathcal{E}$, $k\neq i$. Denote $\mathcal{N}_i=\{k|(k,i)\in\mathcal{E}\}$.

A. Measurement Model

For agent A_i, denote its state at time t by $x_t^i\in\mathbb{R}$, where $t\in\mathbb{Z}^+=\{0,1,2,\ldots\}$. For each $i\in\mathcal{N}$, A_i receives noisy measurements of the states of its neighbors if $\mathcal{N}_i\neq\emptyset$, where $\emptyset$ denotes the empty set. We denote the resulting measurement by A_i of A_k's state by

$$y_t^{ik}=x_t^k+w_t^{ik},\qquad t\in\mathbb{Z}^+,\quad k\in\mathcal{N}_i\neq\emptyset \tag{1}$$

where $w_t^{ik}\in\mathbb{R}$ is the additive noise; see Fig. 1 for illustration. The underlying probability space is denoted by $(\Omega,\mathcal{F},P)$. We call y_t^{ik} the observation of the state of A_k obtained by A_i, and assume each A_i knows its own state x_t^i exactly. There may be various interpretations for the additive noise; a natural one is that x_t^i is corrupted by noise during inter-agent communication [28]. We introduce the assumption:

(**A1**) The digraph $G=(\mathcal{N},\mathcal{E})$ contains a spanning tree. □

For each $t\in\mathbb{Z}^+$, the set of noises $\{w_t^{ik}, i\in\mathcal{N}\text{ and }k\in\mathcal{N}_i\neq\emptyset\}$ is listed into a vector $\mathbf{w}_t$ in which the position of w_t^{ik} depends only on (i,k) and does not change with t. Define the state vector

$$x_t=\left(x_t^1,\ldots,x_t^n\right)^T,\qquad t\geq 0. \tag{2}$$

Denote the σ-algebras $\mathcal{F}_t=\sigma(x_0,\mathbf{w}_k,k=0,\ldots,t)$ (i.e., the set of all events induced by these random variables) for $t\geq 0$. Then obviously $\mathbf{w}_t$ is adapted to (i.e., measurable on) $\mathcal{F}_t$ and $\mathcal{F}_t\subseteq\mathcal{F}_{t+1}$. We introduce the assumption:

(**A2**) The sequence $\{\mathbf{w}_t,t\in\mathbb{Z}^+\}$ satisfies: i) $E[\mathbf{w}_t|\mathcal{F}_{t-1}]=0$ for $t\geq 0$, where $\mathcal{F}_{-1}\triangleq\{\emptyset,\Omega\}$, and ii) $\sup_{t\geq 0}E|\mathbf{w}_t|^2<\infty$. In addition, x_0 satisfies $E|x_0|^2<\infty$. □

By (**A2**)-i) and the fact that $\mathbf{w}_t$ is adapted to $\mathcal{F}_t$ where $\mathcal{F}_t\subseteq\mathcal{F}_{t+1}$, we see that $\{\mathbf{w}_t,t\in\mathbb{Z}^+\}$ is a sequence of martingale differences with respect to (w.r.t.) $\{\mathcal{F}_t,t\in\mathbb{Z}^+\}$. For relevant literature on martingale theory, the reader is referred to [10], [13], [30]. Note that for $t\geq 1$, (**A2**)-i) implies $E[x_0\mathbf{w}_t^T]=E\{E[x_0\mathbf{w}_t^T|\mathcal{F}_{t-1}]\}=E\{x_0E[\mathbf{w}_t^T|\mathcal{F}_{t-1}]\}=0$. The following assumption with independent noises holds as a special case of (**A2**).

($\mathbf{A2}^o$) The noises $\{w_t^{ik},t\in\mathbb{Z}^+,i\in\mathcal{N}\text{ and }k\in\mathcal{N}_i\neq\emptyset\}$ are independent w.r.t. the indices i,k,t and also independent of x_0, and $Ew_t^{ik}=0$, $\sup_{i,k,t}E|w_t^{ik}|^2<\infty$. In addition, $E|x_0|^2<\infty$. □

B. Stochastic Approximation Algorithm

The state of each agent is updated by the rule

$$x_{t+1}^i=(1-a_tb_{ii})x_t^i+a_t\sum_{k\in\mathcal{N}_i}b_{ik}y_t^{ik},\qquad t\geq 0 \tag{3}$$

where $i\in\mathcal{N}$, the step size $a_t>0$, and the parameters b_{ij} will be specified subsequently. We only consider scalar individual states and the generalization to vector individual states is obvious. Throughout our analysis, we adopt the convention: $\sum_{k\in\emptyset} = 0$. For specifying b_{ij} in (3), we consider two cases for $\mathcal{N}_i$.

Case 1) If $\mathcal{N}_i \neq \emptyset$, we take

$$\begin{cases} b_{ik} > 0, & \text{if } k \in \mathcal{N}_i \\ b_{ik} = 0, & \text{if } k \notin \mathcal{N}_i \cup \{i\} \\ b_{ii} = \sum_{k\in\mathcal{N}_i} b_{ik}. \end{cases}$$

The right hand side of (3) is a linear combination of A_i's state and its $|\mathcal{N}_i|$ observations. Here $|S|$ denotes the cardinality of a set S.

Case 2) If $\mathcal{N}_i = \emptyset$, the state of agent i is fixed as its initial value

$$x_t^i \equiv x_0^i. \tag{4}$$

For instance, (4) arises in a leader following situation where the leader's state is fixed as a constant at all times. In the case $\mathcal{N}_i = \emptyset$, we define $b_{ik} \equiv 0$ for all $k \in \mathcal{N}$. By our earlier convention $\sum_{k\in\emptyset} = 0$, (4) may be interpreted as a special case of (3) after x_0^i is given.

Define the matrix

$$B = \begin{pmatrix} -b_{11} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{n1} & \cdots & -b_{nn} \end{pmatrix}. \tag{5}$$

Let $w_t^i = \sum_{k\in\mathcal{N}_i} b_{ik} w_t^{ik}$ for $i \in \mathcal{N}$, and define $w_t = (w_t^1, \ldots, w_t^n)^T$. If $\mathcal{N}_i = \emptyset$, we accordingly have $w_t^i = \sum_{k\in\emptyset} = 0$. We write (3) in the vector form

$$x_{t+1} = x_t + a_t B x_t + a_t w_t, \qquad t \geq 0. \tag{6}$$

We introduce the following assumption on the step size sequence.

(**A3**) The sequence $\{a_t, t \geq 0\}$ satisfies: i) $a_t > 0$, and ii) $\sum_{t=0}^{\infty} a_t = \infty$, $\sum_{t=0}^{\infty} a_t^2 < \infty$. □

Remark: Under (**A1**), $\max_{i\in\mathcal{N}} b_{ii} > 0$ holds. If one further restricts $a_t \leq (\max_{i\in\mathcal{N}} b_{ii})^{-1}$ for all $t \geq 0$, the coefficients for x_t^i and y_t^{ik} in (3) will be maintained nonnegative, leading to a weighted averaging consistent with typical consensus algorithms (see, e.g., [4]).

Definition 1: (Mean Square Consensus): The agents are said to reach mean square consensus if $E|x_t^i|^2 < \infty, t \geq 0, i \in \mathcal{N}$, and there exists a random variable x^* such that $\lim_{t\to\infty} E|x_t^i - x^*|^2 = 0$ for all $i \in \mathcal{N}$. □

Definition 2: (Strong Consensus): The agents are said to reach strong consensus if there exists a random variable x^* such that $\lim_{t\to\infty} x_t^i = x^*$ almost surely (a.s.), for all $i \in \mathcal{N}$. □

III. The Equivalent State Space Model and Auxiliary Lemmas

First of all, we see that B may be interpreted as the generator of an associated continuous time Markov chain $X_t, t \geq 0$, with state space $S = \{1, \ldots, n\}$. Since $b_{ij} > 0$ if and only if $(j, i) \in \mathcal{E}$, the existence of a spanning tree in G is equivalent to the property that X_t has one nonempty communicating class $\mathbf{C} = \{l_1, \ldots, l_{d_1}\}$ and all other states are transient. Now we may summarize the following properties based on standard results on Markov chains [29].

Proposition 3: If G contains a spanning tree, then we have i) B has a unique zero eigenvalue and all other $n-1$ eigenvalues have negative real parts, and ii) there exists a unique probability measure π such that $\pi B = 0$; in addition, π takes a positive value at a state i if and only if $i \in \mathbf{C}$. □

Remark: Part i) of Proposition 3 is proved in [26] by determinant calculations and mathematical induction.

Although (6) is a linear model, most existing convergence results in stochastic approximation cannot be directly applied since B is not Hurwitz. We introduce the following class of $n \times (n-1)$ matrices

$$\mathcal{C}(B) = \left\{\phi \in \mathbb{R}^{n\times(n-1)} | \operatorname{span}\{\phi\} = \operatorname{span}\{B\}\right\}. \tag{7}$$

Under (**A1**), $\operatorname{rank}(B) = n-1$ and each $\phi \in \mathcal{C}(B)$ has rank $n-1$. Let $1_n \in \mathbb{R}^n$ be the vector with all n entries equal to one.

Lemma 4: Assuming (**A1**), for algorithm (6) we have:

(i) For any given $\phi_{n\times(n-1)} \in \mathcal{C}(B)$, the matrix $\Phi = (1_n, \phi_{n\times(n-1)})$ is nonsingular and

$$\Phi^{-1} B \Phi = \begin{pmatrix} 0 & \\ & \tilde{B}_{n-1} \end{pmatrix} \tag{8}$$

where $\tilde{B}_{n-1} \in \mathbb{R}^{(n-1)\times(n-1)}$ is Hurwitz.

(ii) Letting $z_t = (z_t^1, \ldots, z_t^n)^T = \Phi^{-1} x_t$ and $v_t = (v_t^1, \ldots, v_t^n)^T = \Phi^{-1} w_t$, we have the relation

$$z_{t+1}^1 = z_t^1 + a_t v_t^1 \tag{9}$$

$$z_{t+1}^{(n-1)} = (I + a_t \tilde{B}_{n-1}) z_t^{(n-1)} + a_t v_t^{(n-1)}, \quad t \geq 0 \tag{10}$$

where $z_t^{(n-1)} = (z_t^2, \ldots, z_t^n)^T$ and $v_t^{(n-1)} = (v_t^2, \ldots, v_t^n)^T$.

Proof:

(i) Denote $\mathbb{S} \triangleq \operatorname{span}\{\phi_{n\times(n-1)}\} = \operatorname{span}\{B\}$. We show that $1_n \notin \mathbb{S}$; otherwise, there exists $\xi \in \mathbb{R}^n$ such that $1_n = B\xi$, which gives the contradiction $0 < \pi 1_n = \pi B \xi = 0$ where π is determined in Proposition 3. Hence Φ is nonsingular. Let $\Phi^{-1} = \left[\begin{smallmatrix}\Psi_1 \\ \Psi_2\end{smallmatrix}\right]$, where Ψ_1 is the first row, and then $\Psi_1 1_n = 1$ and $\Psi_1 \phi_{n\times(n-1)} = 0$. Since $\phi_{n\times(n-1)} \in \mathcal{C}(B)$, there exists an $(n-1)\times n$ matrix Γ such that $B = \phi_{n\times(n-1)}\Gamma$, which implies $\Psi_1 B = 0$. Recalling Proposition 3, we necessarily have $\Psi_1 = \pi$ to get the representation

$$\Phi^{-1} = \begin{bmatrix} \pi \\ \Psi_2 \end{bmatrix}. \tag{11}$$

Now (8) is easily verified. By the eigenvalue distribution of B in Proposition 3, $\tilde{B}_{n-1}$ is Hurwitz.

(ii) This part follows from (6) and (8).

□

Equation (9) can be viewed as a random walk with increment controlled by the parameter a_t. Equation (10) is a linear stochastic approximation model, and since $\tilde{B}_{n-1}$ is Hurwitz, the convergence of $z_t^{(n-1)}$ can be handled by existing methods (see, e.g., [3], [9]). We have the following equivalence relation:

Lemma 5: Assume (**A1**) and let z_t be defined in Lemma 4. The n agents reach mean square (resp., strong) consensus if and only if z_t^1 converges in mean square (resp., a.s.) to a random variable z_∞^1 and $z_t^{(n-1)}$ converges in mean square (resp., a.s.) to 0, as $t \to \infty$.

Proof: Sufficiency is obvious since we have $x_t = \Phi z_t = z_t^1 1_n + \phi_{n\times(n-1)} z_t^{(n-1)}$. We now prove necessity. Assume $x_t \to x_\infty^1 1_n$ in mean square (resp., a.s.), as $t \to \infty$. By (11), we have $z_t^1 = \pi x_t$ and $z_t^{(n-1)} = \Psi_2 x_t = \Psi_2(x_t - x_\infty^1 1_n) + \Psi_2(x_\infty^1 1_n) = \Psi_2(x_t - x_\infty^1 1_n)$ since $\Psi_2 1_n = 0$. The necessity part follows. □

The convergence of the state trajectory may be completely determined by the noise sample path.

Lemma 6: Following the notation in Lemma 4, algorithm (6) ensures strong consensus if and only if the two conditions hold: (i) $\sum_{t=0}^{k} a_t v_t^1$ converges a.s., as $k \to \infty$, (ii) $\lim_{k\to\infty} \sup_{k\leq t\leq\nu(k,T)} |\sum_{j=k}^{t} a_j v_j^{(n-1)}| = 0$ a.s., where $\nu(k,T) = \max\{j | a_k + \cdots + a_j \leq T\}$ for some constant $0 < T < \infty$.

Proof: We write (10) in the equivalent form $z_{t+1}^{(n-1)} = z_t^{(n-1)} + a_t[\tilde{B}_{n-1} z_t^{(n-1)} + v_t^{(n-1)}]$. Since $\tilde{B}_{n-1}$ is Hurwitz, for any fixed sample $\omega_0 \in \Omega$, $z_t^{(n-1)}(\omega_0)$ converges to zero if and only if

$$\lim_{k\to\infty} \sup_{k\le t\le \nu(k,T)} \left| \sum_{j=k}^{t} a_j v_j^{(n-1)}(\omega_0) \right| = 0. \quad (12)$$

For a proof of this fact, see [33, Theorem 4] and [9, Theorem 1]. Equality (12) is usually called the Kushner-Clark condition [21], [33] along that noise sample path.

Hence, if (i) and (ii) hold, we have the a.s. convergence of z_t^1 to z_∞^1, and $z_t^{(n-1)}$ to 0, as $t\to\infty$. By Lemma 5, strong consensus follows. Conversely, if algorithm (6) ensures strong consensus, Lemma 5 implies condition (i) and also $z_t^{(n-1)} \to 0$ a.s., which further implies condition (ii). □

IV. Mean Square and Almost Sure Convergence

Theorem 7: Under (**A1**)–(**A3**), algorithm (6) achieves mean square consensus.

Proof: By $\sum_{i=0}^{\infty} a_i^2 < \infty$, the mean square convergence of z_t^1 readily follows. We now analyze $z_t^{(n-1)}$. For any $K>0$, the Lyapunov equation $Q\tilde{B}_{n-1} + \tilde{B}_{n-1}^T Q = -K$ has a unique solution $Q>0$ since $\tilde{B}_{n-1}$ is Hurwitz. Let $V_t = E(z_t^{(n-1)})^T Q z_t^{(n-1)}$. Similar to the stochastic Lyapunov analysis in [18], we can show that there exist constants $c_1>0, c_2>0$ such that $V_{t+1} \le (1-c_1 a_t)V_t + c_2 a_t^2$ and $1-c_1 a_t \ge 0$, for $t\ge T_0$ where T_0 is a large constant. We show $\lim_{t\to\infty} V_t = 0$ by contradiction. Assume

$$\limsup_{t\to\infty} V_t = \eta > 0. \quad (13)$$

By (**A3**), we may take a large $T_1 > T_0$ such that $c_2\sum_{t=T_1}^{\infty} a_t^2 < \eta/2$. It is straightforward to show that $V_{t+1} \le \prod_{i=T_1}^{t}(1-c_1 a_i)V_{T_1} + c_2\sum_{i=T_1}^{t} a_t^2 \le \prod_{i=T_1}^{t}(1-c_1 a_i)V_{T_1} + (\eta/2)$; since $\sum_{t=0}^{\infty} a_t = \infty$, we have $\lim_{t\to\infty}\prod_{i=T_1}^{t}(1-c_1 a_i) = 0$, which implies $\limsup_{t\to\infty} V_t < \eta$ and contradicts (13). Hence, we conclude that $\lim_{t\to\infty} V_t = 0$ implying $\lim_{t\to\infty} E|z_t^{(n-1)}|^2 = 0$. Mean square consensus follows from Lemma 5. □

Theorem 8: Under (**A1**)–(**A3**), algorithm (6) ensures strong consensus.

Proof: Since v_t defined in Lemma 4 is linear in $\mathbf{w}_t$ and $\sum_{t=0}^{\infty} a_t^2 E|\mathbf{w}_t|^2 < \infty$ by (**A2**), it follows that $\sum_{t=0}^{\infty} a_t^2 E|v_t^1|^2 < \infty$ and $\sum_{t=0}^{\infty} a_t^2 E|v_t^{(n-1)}|^2 < \infty$. By the martingale convergence theorem [13, pp. 18–19], it follows that both $\sum_{t=0}^{k} a_t v_t^1$ and $\sum_{t=0}^{k} a_t v_t^{(n-1)}$ converge a.s., as $k\to\infty$. The convergence of $\sum_{t=0}^{k} a_t v_t^{(n-1)}$ clearly implies condition (ii) in Lemma 6. Hence, strong consensus follows from Lemma 6. □

Remark: For the special case of leader following, let agent k_0 be the leader with $x_t^{k_0} \equiv x_0^{k_0}$. So the k_0th row of B is zero. In (11), π is accordingly given as $e_{k_0} = (0,\ldots,0,1,0,\ldots,0)$ satisfying $\pi B = 0$. By $z_t = \Phi^{-1}x_t$, it follows that $z_t^1 = x_t^{k_0} \equiv x_0^{k_0}$. In this case, we can further verify that $x_\infty = (1_n, \phi_{n\times(n-1)})(z_\infty^1, 0)^T = x_0^{k_0} 1_n$.

We state the following corollary by using a pth conditional moment condition of the noise while strengthening the condition for the step size sequence.

Corollary 9: Letting $p\in(1,2]$, we assume (**A1**) holds, (**A2**) holds after replacing $\sup_{t\ge 0} E|\mathbf{w}_t|^2 < \infty$ by $\sup_{t\ge 0} E[|\mathbf{w}_t|^p|\mathcal{F}_{t-1}] < \infty$, and (**A3**) holds after replacing $\sum_{i=0}^{\infty} a_i^2 < \infty$ by $\sum_{i=0}^{\infty} a_i^p < \infty$. Then algorithm (6) ensures strong consensus.

Proof: We can first show the a.s. convergence of $\sum_{t=0}^{k} a_t v_t^1$ and $\sum_{t=0}^{k} a_t v_t^{(n-1)}$ (see, e.g., [30, pp. 67]). Then strong consensus follows as in Theorem 8. □

V. Randomly Time-Varying Communication Links

Let us use the fixed digraph $G=(\mathcal{N},\mathcal{E})$ in Section II to describe the maximal set of communication links when there is no link failure. At time t the inter-agent communication is described by a subgraph of G denoted by $G_t = (\mathcal{N},\mathcal{E}_t)$ where $\mathcal{E}_t \subset \mathcal{E}$; the edge $(i,j)\in\mathcal{E}_t$ if and only if there exists a communication link from i to j at time t where $(i,j)\in\mathcal{E}$. The digraph G_t is generated as the outcome of random link failures and hence depends on the probability sample. Denote $\mathcal{N}_{it} = \{k|(k,i)\in\mathcal{E}_t\}$ at time t.

At $t\ge 0$, the adjacency matrix of G_t is defined as $A_t^{\mathcal{N}} = (a_t^{ij})_{1\le i,j\le n}$, where $a_t^{ij}=1$ if $(i,j)\in\mathcal{E}_t$, and $a_t^{ij}=0$ otherwise. The digraph G_t is completely characterized by the random matrix $A_t^{\mathcal{N}}$. Now, the measurement relation is given as

$$y_t^{ik} = x_t^k + w_t^{ik} \quad \text{if} \quad a_t^{ki}=1 \ (\text{i.e., } k\in\mathcal{N}_{it})$$

where w_t^{ik} is the noise. The state x_t^i is updated by the rule

$$x_{t+1}^i = (1-a_t|\mathcal{N}_{it}|)\, x_t^i + a_t \sum_{k\in\mathcal{N}_{it}} y_t^{ik}, \qquad t\ge 0. \quad (14)$$

When $\mathcal{N}_{it}=\emptyset$, (14) is interpreted as $x_{t+1}^i = x_t^i$. Here for simplicity we assign the same weight to the $|\mathcal{N}_{it}|$ observations y_t^{ik}.

In order to specify the statistical properties of the noise, we introduce the array of measurement noises as a square matrix: $W_t = (w_t^{ik})_{1\le i,k\le n}$, where $w_t^{ik}\equiv 0$ if $(k,i)\notin\mathcal{E}$. It is sufficient to further specify w_t^{ik} with $(k,i)\in\mathcal{E}$. The combined link and noise assumption is stated below.

(**A4**) (i) For $(i,j)\in\mathcal{E}$, $P\{a_t^{ij}=1\} = P\{(i,j)\in\mathcal{E}_t\} = p_{ij} > 0$, and for each t, $\{a_t^{ij}|(i,j)\in\mathcal{E}\}$ are independent binary random variables. (ii) The pair $(A_t^{\mathcal{N}}, W_t)$ is independent of $(x_0, A_k^{\mathcal{N}}, W_k, k\le t-1)$, where $t\ge 0$. (iii) Conditioned on $A_t^{\mathcal{N}} = (a_t^{ij})_{1\le i,j\le n}$, the noises $\{w_t^{ik}|(k,i)\in\mathcal{E}\}$ are independent and satisfy

$$P\left(w_t^{ik}=0|a_t^{ki}=0\right)=1, \quad E\left(w_t^{ik}|a_t^{ki}=1\right)=0$$

$$\sup_{i,k,t} E\left(\left|w_t^{ik}\right|^2 |a_t^{ki}=1\right) \le C_w$$

where $C_w<\infty$ is a constant. The term $(x_0, A_k^{\mathcal{N}}, W_k, k\le t-1)$ is interpreted as x_0 when $t=0$. □

If we further define the distribution of w_t^{ik} conditioned on $\{a_t^{ki}=1\}$, then any finite dimensional distribution of $(x_0, A_k^{\mathcal{N}}, W_k, k\le t)$ is well defined. We still denote $x_t = (x_t^1,\ldots,x_t^n)^T$, and define the noise vector $w(A_t^{\mathcal{N}}, W_t) = (w_t^1,\ldots,w_t^n)^T$, where $w_t^i = \sum_{k\in\mathcal{N}_{it}} a_t^{ki} w_t^{ik}$. Denote $D(A_t^{\mathcal{N}}) = \mathrm{Diag}(\sum_{k\in\mathcal{N}} a_t^{k1},\ldots,\sum_{k\in\mathcal{N}} a_t^{kn})$. We write (14) in the vector form

$$x_{t+1} = x_t + a_t\left[\left(A_t^{\mathcal{N}}\right)^T - D\left(A_t^{\mathcal{N}}\right)\right] x_t + a_t w\left(A_t^{\mathcal{N}}, W_t\right). \quad (15)$$

Let $B_t = [(A_t^{\mathcal{N}})^T - D(A_t^{\mathcal{N}})]$, $\overline{B} = EB_t$ and $\Delta B_t = B_t - \overline{B}$, and it can be shown that all row sums of both B_t and $\overline{B}$ are zero. Now for $t\ge 0$, (15) may be written in the form

$$x_{t+1} = x_t + a_t\overline{B}x_t + a_t\Delta B_t x_t + a_t w\left(A_t^{\mathcal{N}}, W_t\right). \quad (16)$$

Lemma 10: Assuming (**A4**) holds, we have:

(i) For $t\ge 0$, the pair $(\Delta B_t, w(A_t^{\mathcal{N}}, W_t))$ is independent of x_t, and $Ew(A_t^{\mathcal{N}}, W_t)) = 0$.

(ii) If, in addition, G contains a spanning tree, then $\overline{B}$ satisfies: a) it has a zero eigenvalue of algebraic multiplicity equal to one, and $n-1$ eigenvalues with negative real parts; b) there exists a unique probability measure such that $\bar{\pi}\overline{B}=0$; c) $\mathrm{Null}(\overline{B}) = \mathrm{span}\{1_n\}$.

Proof:

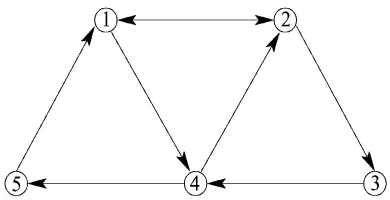

Fig. 2. Digraph with five nodes.

(i) By (15), we see that x_t depends on $(x_0, A_k^{\mathcal{N}}, W_k, k \le t-1)$, $t \ge 1$, and the independence part follows from part (ii) of (**A4**). By part (iii) of (**A4**), we can show $Ew(A_t^{\mathcal{N}}, W_t) = 0$.

(ii) Let $\overline{B} = (\bar{b}_{ij})_{1\le i,j\le n}$. Since each row sum of $\overline{B}$ is zero, and for $j \ne i$, $\bar{b}_{ij} > 0$ if and only if $P\{a_t^{ji} = 1\} > 0$ if any only if $(j,i) \in \mathcal{E}$, Proposition 3 can be applied in an obvious manner. □

By Lemma 10, (16) can be viewed as a perturbed version of (6) where the additional term $a_t \Delta B_t x_t$ is unbiased in the sense $E[\Delta B_t x_t] = 0$ and is controlled by the decreasing step size a_t.

Theorem 11: Under (**A1**), (**A3**) and (**A4**), algorithm (15) ensures mean square and strong consensus.

Proof: See Appendix. □

VI. Simulations

We consider a digraph with 5 nodes as shown in Fig. 2. The variance of the i.i.d. zero mean Gaussian measurement noises is $\sigma^2 = 0.01$, and the constant initial condition is $x_t|_{t=0} \equiv [4,3,1,6,1]^T$. So assumptions (**A2**) and ($\mathbf{A2}^o$) are satisfied. Fig. 3(a) shows the simulation of the standard averaging rule with equal weights to an agent's neighbors and itself (for instance, $x_{t+1}^1 = (x_t^1 + y_t^{12} + y_t^{15})/3, t \ge 0$), and no convergence is achieved. Fig. 3(b) shows the 5 trajectories all converge to the same constant when algorithm (6) is applied with $b_{ij} = |\mathcal{N}_i|^{-1}$, $j \in \mathcal{N}_i$, and $\{a_t = (t+5)^{-0.85}, t \ge 0\}$.

In the next simulation, the network topology is based on Fig. 2 where each link fails independently with a failure probability p_f, and algorithm (14) is implemented using $\{a_t = (t+5)^{-0.85}, t \ge 0\}$. The independent Gaussian noise has zero mean and variance $\sigma^2 = 0.01$. In Fig. 4(a) and 4(b), the initial conditions are still given by $x_t|_{t=0} = [4,3,1,6,1]^T$, but the failure probabilities are, respectively, given by 0.3 and 0.75. Due to poorer connectivity conditions, the convergence rate in Fig. 4(b) is much slower than in Fig. 4(a).

VII. Conclusion

We consider consensus problems on digraphs with noisy measurements. We apply stochastic approximation algorithms and establish their mean square and almost sure convergence. Furthermore, the modeling and algorithms are generalized to networks with independent random link failures.

Appendix
Proof of Theorem 11

In this appendix, we use C_0 to denote a generic constant independent of t which may vary from place to place.

Lemma 12: Let $\{V_t, t \ge 0\}$ and $\{s_t, t \ge 0\}$ be two sequences of nonnegative numbers, and $c_0 > 0$ be a constant such that $\inf_{t\ge 0}(1 - c_0 a_t) \ge 0$, where a_t is given in (**A3**). Suppose

$$\begin{cases} V_{t+1} \le (1 - c_0 a_t)V_t + C a_t^2(1 + V_t + s_t) \\ s_{t+1} \le s_t + C_1 a_t^2 (1 + s_t + V_t) \end{cases} \quad \text{(A.1)}$$

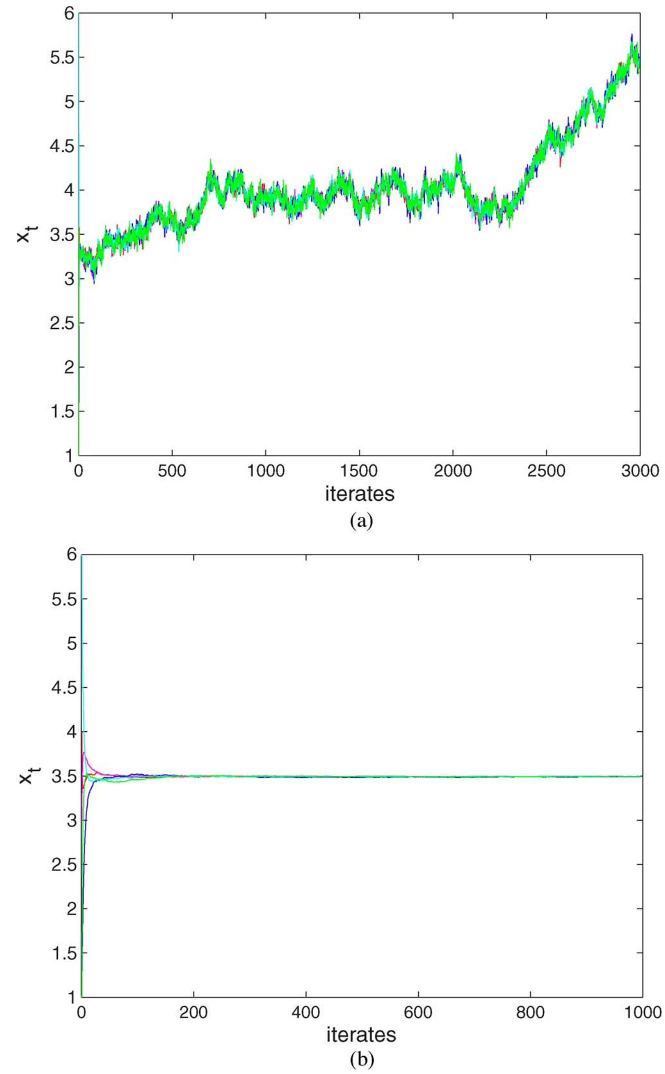

Fig. 3. (a) The five trajectories fail to converge when fixed weights are used. (b) The five trajectories converge to the same constant level with a decreasing step size.

holds for $t \ge 0$, where $C > 0$ and $C_1 > 0$ are constants. Then $\lim_{t\to\infty} V_t = 0$ and $\sup_{t\ge 0} s_t < \infty$.

Proof: By the second part of (A.1), we have

$$s_{t+1} \le s_0 \prod_{k=0}^{t}\left(1 + C_1 a_k^2\right) + C_1 \sum_{i=0}^{t}\left[\prod_{k=i+1}^{t}\left(1 + C_1 a_k^2\right)\right] a_i^2 (1 + V_i).$$

Since s_0 is fixed and $\sum_{i=0}^{\infty} a_i^2 < \infty$ implies $\prod_{k=0}^{\infty}(1 + C_1 a_k^2) < \infty$, it follows that

$$s_{t+1} \le C_0 \left(1 + \sum_{i=0}^{t} a_i^2 V_i\right) \le C_0 (1 + \max_{0\le i\le t} V_i), \qquad t \ge 0 \quad \text{(A.2)}$$

where C_0 is a generic constant by our earlier convention. Combining the first part of (A.1) with (A.2), we have

$$\begin{aligned} V_{t+1} &\le (1 - c_0 a_t)V_t + C_0 a_t^2 (1 + \max_{0\le i\le t} V_i) \\ &\le \left(1 - c_0 a_t + C_0 a_t^2\right) \max_{0\le i\le t} V_i + C_0 a_t^2. \end{aligned} \quad \text{(A.3)}$$

Since $\lim_{t\to\infty} a_t = 0$, we can choose a large $T > 0$ (depending on c_0, C_0 in (A.3)) such that for all $t \ge T$, $V_{t+1} \le \max_{0\le i\le t} V_i + C_0 a_t^2$, which implies that

$$\max_{0\le i\le t+1} V_i \le \max_{0\le i\le t} V_i + C_0 a_t^2, \quad t \ge T. \quad \text{(A.4)}$$

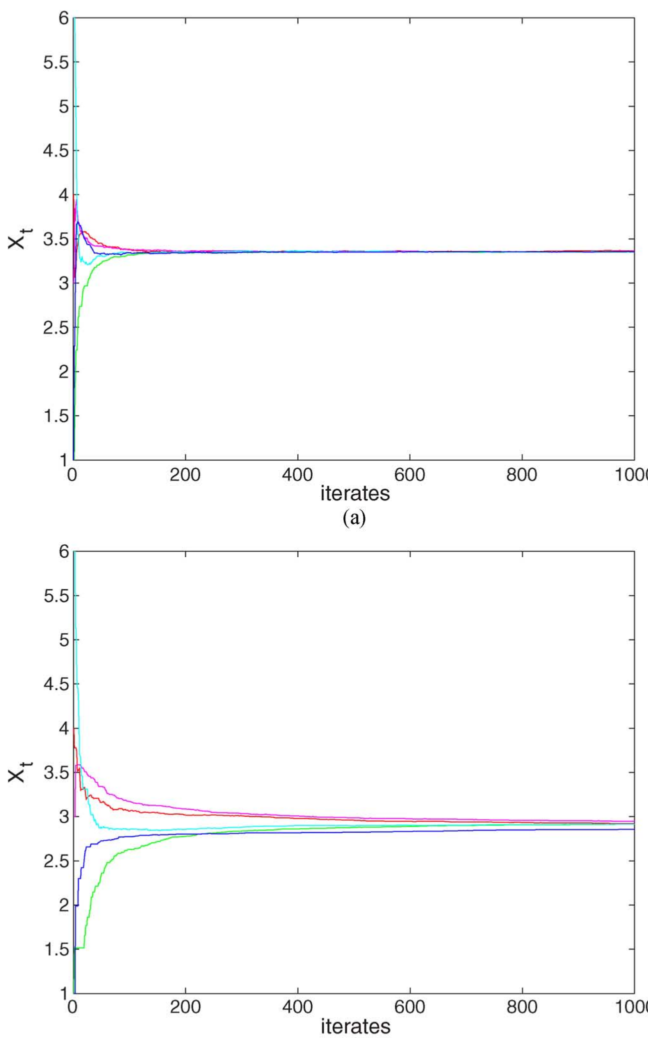

Fig. 4. (a) The five trajectories converge with a decreasing step size and failure probability $p_f = 0.3$. (b) The five trajectories converge with a deceasing step size and $p_f = 0.75$.

Hence it follows from (A.4) and $\sum_{t=0}^{\infty} a_t^2 < \infty$ that

$$\sup_{t\geq 0} V_t < \infty. \tag{A.5}$$

Consequently, (A.2) and (A.5) yield $\sup_{t\geq 0} s_t < \infty$.

By (A.5) and the boundedness of s_t, it follows from (A.1) that $V_{t+1} \leq (1 - c_0 a_t)V_t + C_0 a_t^2$, which further implies $\lim_{t\to\infty} V_t = 0$ since $\sum_{i=0}^{\infty} a_i = \infty$ and $\sum_{i=0}^{\infty} a_i^2 < \infty$. □

Proof of Theorem 11: Step 1—(Change of coordinates). In parallel to (7), we introduce the class $\mathcal{C}(\overline{B})$ and take $\bar{\phi}_{n\times(n-1)} \in \mathcal{C}(\overline{B})$. Similar to the case of Φ in Lemma 4, it can be checked that $\bar{\Phi} = (1_n, \bar{\phi}_{n\times(n-1)})$ is nonsingular, and the first row in $\bar{\Phi}^{-1}$ is $\bar{\pi}$ given in Lemma 10. We may write $\bar{\Phi}^{-1} = \left[\begin{smallmatrix}\bar{\pi}\\ \bar{\Psi}_2\end{smallmatrix}\right]$. Moreover, there exists an $(n-1)\times(n-1)$ Hurwitz matrix $\overline{B}_{n-1}$ such that $\bar{\Phi}^{-1}\overline{B}\bar{\Phi} = \mathrm{Diag}(0, \overline{B}_{n-1}) \triangleq \widetilde{B}$.

Let $z_t = (z_t^1, \ldots, z_t^n)^T = \bar{\Phi}^{-1} x_t$ and $z_t^{(n-1)} = (z_t^2, \ldots, z_t^n)^T$. It follows that for $t \geq 0$

$$z_{t+1} = z_t + a_t \widetilde{B} z_t + a_t \bar{\Phi}^{-1}\Delta B_t \bar{\Phi} z_t + a_t \bar{\Phi}^{-1} w\left(A_t^{\mathcal{N}}, W_t\right). \tag{A.6}$$

Denote $\bar{\Phi}^{-1}\Delta B_t\bar{\Phi} = \left[\begin{smallmatrix}J_1(t)\\ J_2(t)\end{smallmatrix}\right]$, where $J_1(t)$ is the first row. By Lemma 10, $EJ_1(t) = 0$, $EJ_2(t) = 0$ and the pair $(J_1(t), J_2(t))$ is independent of z_t. Now (A.6) can be written in the form

$$z_{t+1}^1 = z_t^1 + a_t J_1(t) z_t + a_t \bar{\pi} w\left(A_t^{\mathcal{N}}, W_t\right) \tag{A.7}$$

$$z_{t+1}^{(n-1)} = z_t^{(n-1)} + a_t \overline{B}_{n-1} z_t^{(n-1)} + a_t J_2(t) z_t + a_t \bar{\Psi}_2 w\left(A_t^{\mathcal{N}}, W_t\right). \tag{A.8}$$

Step 2—(Estimates for the Lyapunov function). Since $\overline{B}_{n-1}$ is Hurwitz, there exists a unique solution $Q > 0$ to the algebraic Lyapunov equation $Q\overline{B}_{n-1} + \overline{B}_{n-1}^T Q = -I$. Denote $s_t = E|z_t^1|^2$ and $V_t = E(z_t^{(n-1)})^T Q z_t^{(n-1)}$ for $t \geq 0$.

Since $(J_1(t), \bar{\pi} w(A_t^{\mathcal{N}}, W_t))$ is independent of z_t, we have $E[z_t^1 J_1(t) z_t] = E[z_t^1 \bar{\pi} w(A_t^{\mathcal{N}}, W_t)] = 0$. By the expectation of w_t^{ij} conditioned on $A_t^{\mathcal{N}}$, it can be shown that $E[J_1(t) z_t \bar{\pi} w(A_t^{\mathcal{N}}, W_t)] = 0$. Now for $t \geq 0$, (A.7) leads to

$$s_{t+1} = s_t + a_t^2 E\left|J_1(t) z_t\right|^2 + a_t^2 E\left|\bar{\pi} w\left(A_t^{\mathcal{N}}, W_t\right)\right|^2. \tag{A.9}$$

Hence, $\{s_t, t \geq 0\}$ is a monotonically increasing sequence, and by the boundedness of $J_1(t)$, we can find a fixed $C_1 > 0$ such that

$$s_{t+1} \leq s_t + C_1 a_t^2 (1 + s_t + V_t), \quad t \geq 0 \tag{A.10}$$

since $|z_t^1|^2 + (z_t^{(n-1)})^T Q z_t^{(n-1)} \geq \eta |z_t|^2$ for some $\eta > 0$.

We proceed to estimate V_{t+1}. Similar to the derivation of (A.9), we eliminate crossing terms to get

$$\begin{aligned}
V_{t+1} = {} & E\left[\left(z_t^{(n-1)}\right)^T \left(I + a_t \overline{B}_{n-1}^T\right) Q (I + a_t \overline{B}_{n-1}) z_t^{(n-1)}\right] \\
& + a_t^2 E\left[z_t^T J_2^T(t) Q J_2(t) z_t\right] \\
& + a_t^2 E\left[w^T\left(A_t^{\mathcal{N}}, W_t\right) \bar{\Psi}_2^T Q \bar{\Psi}_2 w\left(A_t^{\mathcal{N}}, W_t\right)\right] \\
\leq {} & V_t - a_t E\left|z_t^{(n-1)}\right|^2 \\
& + a_t^2 E\left[\left(z_t^{(n-1)}\right)^T \overline{B}_{n-1}^T Q \overline{B}_{n-1} z_t^{(n-1)}\right] \\
& + C_0 a_t^2 \left(1 + E|z_t|^2\right) \\
\leq {} & V_t - a_t E\left|z_t^{(n-1)}\right|^2 + C_0 a_t^2 \left(1 + E|z_t|^2\right).
\end{aligned} \tag{A.11}$$

By (A.11), we can pick a fixed constant $c_0 > 0$ satisfying $\inf_{t\geq 0}(1 - c_0 a_t) \geq 0$ such that $V_{t+1} \leq (1 - c_0 a_t)V_t + C_0 a_t^2 (1 + E|z_t|^2)$ for $t \geq 0$, and therefore, for a fixed constant $C_2 > 0$, we have

$$V_{t+1} \leq (1 - c_0 a_t) V_t + C_2 a_t^2 (1 + s_t + V_t). \tag{A.12}$$

Step 3—(Mean square convergence of x_t). By (A.10), (A.12), Lemma 12 and (A.7), we see that z_t converges in mean square to $z_\infty = (z_\infty^1, 0)^T$. Hence, x_t converges in mean square to $\bar{\Phi} z_\infty = (1_n, \bar{\phi}_{n\times(n-1)}) z_\infty = z_\infty^1 1_n$.

Step 4—(Strong consensus). For reasons of space, we briefly sketch the key steps. First, letting $\mathcal{F}_t$ be the σ-algebra generated by $(x_0, A_k^{\mathcal{N}}, W_k, k = 0, \ldots, t-1)$, we see that z_t is adapted to $\mathcal{F}_t$. Second, letting Q be determined in Step 2, by elementary calculations, we may check that

$$E\left[\left(z_{t+1}^{(n-1)}\right)^T Q z_{t+1}^{(n-1)} | \mathcal{F}_t\right] \leq \left(z_t^{(n-1)}\right)^T Q z_t^{(n-1)} + C a_t^2 \left(1 + \left|z_t^1\right|^2 + \left|z_t^{(n-1)}\right|^2\right) \tag{A.13}$$

for some $C > 0$ and all t. Third, it follows from Step 3 that $\sum_{t=0}^{\infty} a_t^2 E[|z_t^1|^2 + |z_t^{(n-1)}|^2] < \infty$. Now, by Lemma A.1 in [32], it follows from (A.13) that $(z_t^{(n-1)})^T Q z_t^{(n-1)}$ and hence $z_t^{(n-1)}$ converge a.s. The a.s. limit of $z_t^{(n-1)}$ is necessarily 0 by Step 3.

Furthermore, Step 1 implies that $\{J_1(t)z_t, t \geq 0\}$ and $\{\bar{\pi}w(A_t^{\mathcal{N}}, W_t), t \geq 0\}$ in (A.7) form two martingale difference sequences (w.r.t. $\mathcal{F}_t' = \sigma(x_0, A_k^{\mathcal{N}}, W_k, k = 0, \ldots, t)$), each with bounded second moments. Hence, by martingale convergence theorem [30], z_t^1 converges to a limit z_∞^1 a.s. Finally, strong consensus follows from the coordinate change between x_t and z_t. □

REFERENCES

[1] P. Barooah and J. P. Hespanha, "Graph effective resistence and distributed control: Special properties and applications," in *Proc. 45th IEEE CDC Conf.*, San Diego, CA, Dec. 2006, pp. 3479–3485.

[2] D. Bauso, L. Giarré, and R. Pesenti, "Consensus for switched networks with unknown but bounded disturbances," arXiv:math/0612834v1, Dec. 2006.

[3] A. Benveniste, M. Métivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*. New York: Springer-Verlag, 1990.

[4] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *Proc. 44th IEEE CDC-ECC'05*, Seville, Spain, Dec. 2005, pp. 2996–3001.

[5] V. Borkar and P. Varaiya, "Asymptotic agreement in distributed estimation," *IEEE Trans. Autom. Control*, vol. AC-27, pp. 650–655, Jun. 1982.

[6] P. E. Caines, *Linear Stochastic Systems*. New York: Wiley, 1988.

[7] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri, "Communication constraints in the average consensus problem," *Automatica*, vol. 44, pp. 671–684, 2008.

[8] H.-F. Chen, *Stochastic Approximation and Its Applications*. Boston, MA: Kluwer, 2002.

[9] E. K. P. Chong, I.-J. Wang, and S. R. Kulkarni, "Noise conditions for prescribed convergence rates of stochastic approximation algorithms," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 810–814, Mar. 1999.

[10] Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*. New York: Springer-Verlag, 1978.

[11] J. Cortes, S. Martinez, and F. Bullo, "Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions," *IEEE Trans. Autom. Control*, vol. 51, no. 8, pp. 1289–1298, Aug. 2006.

[12] A. Giridhar and P. R. Kumar, "Toward a theory of in-network computation in wireless sensor networks," *IEEE Commun. Mag.*, pp. 98–107, Apr. 2006.

[13] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application*. New York: Academic Press, 1980.

[14] Y. Hatano and M. Mesbahi, "Agreement in random networks," *IEEE Trans. Autom. Control*, vol. 51, no. 11, pp. 1867–1872, Nov. 2005.

[15] M. Huang and J. H. Manton, "Stochastic approximation for consensus-seeking: Mean square and almost sure convergence," in *Proc. 46th IEEE CDC Conf.*, New Orleans, LA, Dec. 2007, pp. 121–126.

[16] M. Huang and J. H. Manton, "Stochastic double array analysis and convergence of consensus algorithms with noisy measurements," in *Proc. Amer. Control Conf.*, New York, NY, Jul. 2007, pp. 705–710.

[17] M. Huang and J. H. Manton, "Stochastic Lyapunov analysis for consensus algorithms with noisy measurements," in *Proc. Amer. Control Conf.*, New York, NY, Jul. 2007, pp. 1419–1424.

[18] M. Huang and J. H. Manton, "Coordination and consensus of networked agents with noisy measurements: Stochastic algorithms and asymptotic behavior," *SIAM J. Control Optim.: Special Issue Control Optim. Cooperative Networks*, vol. 48, no. 1, pp. 134–161, 2009.

[19] 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. 988–1000, Jun. 2003.

[20] S. Kar and J. M. Moura, "Distributed average consensus in sensor networks with random link failures," in *Proc. IEEE Int. Conf. Acoust., Speech, Sig. Process. (ICASSP)*, Honolulu, HI, Apr. 2007, vol. II, pp. 1013–1016.

[21] H. J. Kushner and D. S. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. New York: Springer-Verlag, 1978.

[22] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[23] R. Olfati-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.

[24] S. Patterson, B. Bamieh, and A. E. Abbadi, "Distributed average consensus with stochastic communication failures," in *Proc. 46th IEEE CDC*, New Orleans, LA, Dec. 2007, pp. 4215–4220.

[25] M. Porfiri and D. J. Stilwell, "Consensus seeking over random weighted directed graphs," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1767–1773, Sep. 2007.

[26] 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.

[27] W. Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proc. Amer. Control Conf.*, Portland, OR, Jun. 2005, pp. 1859–1864.

[28] W. Ren, R. W. Beard, and D. B. Kingston, "Multi-agent Kalman consensus with relative uncertainty," in *Proc. Amer. Control Conf.*, Portland, OR, Jun. 2005, pp. 1865–1870.

[29] E. Seneta, *Non-Negative Matrices and Markov Chains*, 2nd ed. New York: Springer, 2006.

[30] W. F. Stout, *Almost Sure Convergence*. New York: Academic Press, 1974.

[31] 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.

[32] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. AC-31, no. 9, pp. 803–812, Sep. 1986.

[33] I.-J. Wang, E. K. P. Chong, and S. R. Kulkarni, "Weighted averaging and stochastic approximation," *Math. Control Signals Syst.*, vol. 10, pp. 41–60, 1997.

[34] C. W. Wu, "Synchronization and convergence of linear dynamics in random directed networks," *IEEE Trans. Autom. Control*, vol. 51, no. 7, pp. 1207–1210, Jul. 2006.

[35] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *J. Parallel Distrib. Comput.*, vol. 67, pp. 33–46, 2007.