

# Stochastic Double Array Analysis and Convergence of Consensus Algorithms with Noisy Measurements

Minyi Huang and Jonathan H. Manton

**Abstract**—This paper considers consensus-seeking of networked agents in an uncertain environment where each agent has noisy measurements of its neighbors' states. We propose stochastic approximation type algorithms with a decreasing step size. We first establish consensus results in a two-agent model via a stochastic double array analysis. Next, we generalize the analysis to a class of well studied symmetric models and obtain consensus results.

## I. INTRODUCTION

The recent years have witnessed an enormous growth of research on the coordination and control of distributed multi-agent systems, and specific topics appear in different forms such as swarming of honeybees, flocking of birds, and formation of autonomous vehicles; see e.g. [28], [9], [11], [17], [24], [20]. A common feature to these systems is that the constituent agents need to maintain a certain coordination so as to cooperatively achieve a group objective.

For coordinating the agents' behavior, it is usually important to propagate shared information within the system by communication rules which may be supported by the interconnection structure between the agents. In this context, of fundamental importance is the so called consensus or agreement problem. In the literature, almost all consensus algorithms are constructed based on averaging rules, and this usually leads to good convergence properties [14], [3].

In its basic formulation, a consensus model consists of a fixed network in which each agent updates its state by forming a convex combination of the states of its neighbors and itself. Starting from this formulation, many generalizations are possible. For instance, the state update may take place asynchronously [21], [2]. In other scenarios, the network topology may change with time [21]. For convergence analysis, stochastic matrix analysis is an important tool [14], and in models with time-dependent communications, set-valued Lyapunov theory is also useful [19].

In this paper, we are interested in consensus-seeking in an uncertain environment. In contrast to most existing work, in our model each agent can only obtain noisy measurements for the states of its neighbors while knowing its own state. Such modelling reflects many practical properties in distributed networks. For instance, the information exchange between different agents may involve the usage of sensors, quantization, and wireless fading channels, which makes it

unlikely to have noise free data delivery. The communication noise issue also arises in the setting of distributed function computation in sensor networks [12]. We note that most previous research on consensus problems has focused on models with perfect measurements, with only a few exceptions (see, e.g., [29], [23], [5]). In particular, the work [29] considered an averaging rule with additive noise and attempted to minimize the long term mean square error by optimizing the coefficients in averaging. In the consensus algorithm of [5], multiplicative noises are introduced to model logarithmic quantization error. In the early work [4], [26], [27] convergence of consensus problems was studied in a stochastic setting, but the inter-agent exchange of random messages was assumed to be error-free; see [13] for more detailed discussions.

In models with noisy measurements, one may easily construct an averaging rule with a constant coefficient matrix. However, this in general leads to no convergence results. For consensus seeking in the stochastic models, the key feature of our algorithm is a decreasing step size, and the algorithm has a gradient descending interpretation. We begin by analyzing a two-agent model. As it turns out, this simple model provides a rich structure for developing convergence analysis and motivates the solution to more general models. In this setup, the key technique is the stochastic double array analysis [25], [7], [22]. Next, we consider the stochastic consensus problem for a class of symmetric networks. In fact, many symmetric models have arisen in practical applications including robot teams, unicycle pursuit models [18], [17], cooperative sensor network deployment for tracking [1] or sampling [16], and consensus problems [5]. Hence, the symmetry assumption not only simplifies the computation, but also has practical importance. For consensus analysis on networks without symmetry, a different approach is developed in [13] via a stochastic Lyapunov analysis.

## II. THE PROBLEM FORMULATION

Consider a set of  $n$  agents distributed according to a directed graph (or digraph)  $G = (\mathcal{N}, \mathcal{E})$  consisting of a set of nodes  $\mathcal{N} = \{1, 2, \dots, n\}$  and a set of edges  $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ . In the digraph, 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 path (from  $i_1$  to  $i_l$ ) consists of a sequence of nodes  $i_1, i_2, \dots, i_l$ ,  $l \geq 2$ , such that  $(i_k, i_{k+1}) \in \mathcal{E}$  for  $k = 1 \dots, l-1$ . We say node  $i$  is connected to node  $j$  ( $\neq i$ ) if there exists a path from  $i$  to  $j$ . The graph  $G$  is said to be strongly connected if for any two distinct nodes  $i$  and  $j$ , there exist a path from  $i$  to  $j$  and also a path from  $j$  to  $i$ .

This work was supported by the Australian Research Council.

M. Huang and J. H. Manton are with the Department of Information Engineering, Research School of Information Sciences and Engineering, The Australian National University, Canberra, ACT 0200, Australia. Email: {minyi.huang, jonathan.manton}@rsise.anu.edu.au.

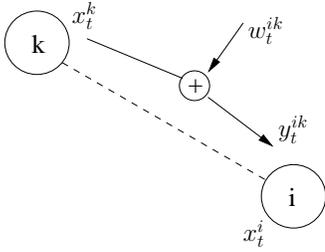


Fig. 1. Measurement with noise  $w_t^{ik}$ .

For convenience of exposition, the two names, agent and node, will be used alternatively. The agent  $A_k$  (resp., node  $k$ ) is a neighbor of  $A_i$  (resp., node  $i$ ) if  $(k, i) \in \mathcal{E}$  where  $k \neq i$ , i.e., there exists an edge from node  $k$  to node  $i$ . Denote the neighbors of node  $i$  by  $\mathcal{N}_i = \{k | (k, i) \in \mathcal{E}\}$ . Throughout the paper we consider digraphs. Note that any undirected graph model<sup>1</sup> can be equivalently represented as a directed graph.

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

$$y_t^{ik} = x_t^k + w_t^{ik}, \quad t \in \mathbb{Z}^+, \quad k \in \mathcal{N}_i, \quad (1)$$

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 [23]. We introduce the assumption:

**(A1)** The noises  $\{w_t^{ik}, t \in \mathbb{Z}^+, i \in \mathcal{N}, k \in \mathcal{N}_i\}$  are independent and identically distributed with respect to the indices  $i, k, t$  and also independent of the initial states  $x_0^i$ ,  $i \in \mathcal{N}$ . In addition,  $\sup_{i \in \mathcal{N}} E|x_0^i|^2 < \infty$ , and each  $w_t^{ik}$  has zero mean and variance  $Q \geq 0$ .  $\square$

Condition **(A1)** means that the noises are i.i.d. with respect to both space (associated with distributed agents) and time.

The state of each agent is updated by:

$$x_{t+1}^i = (1 - a_t)x_t^i + \frac{a_t}{|\mathcal{N}_i|} \sum_{k \in \mathcal{N}_i} y_t^{ik}, \quad (2)$$

where  $i \in \mathcal{N}$ ,  $t \in \mathbb{Z}^+$  and  $a_t \in [0, 1]$ . This gives an averaging rule in that the right hand side is a convex combination of the agent's state and its  $|\mathcal{N}_i|$  observations, where  $|S|$  is used to denote the cardinality of a set  $S$ . The objective of the consensus problem is to select  $\{a_t, t \geq 0\}$  so that the  $n$  individual states  $x_t^i$  converge to a common limit in a certain sense.

To get some insight into the algorithm (2), we rewrite it in the form

$$x_{t+1}^i = x_t^i + a_t(m_t^i - x_t^i), \quad (3)$$

<sup>1</sup>The edge in an undirected graph is denoted as an unordered pair; for definiteness it is assumed there is no self-loop at any node.

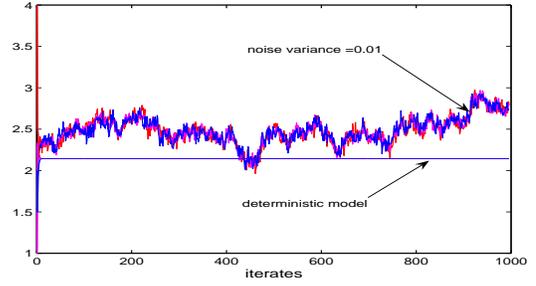


Fig. 2. In the noise free case, the states of the three nodes quickly converge to the same constant level  $\approx 2.143$ . Under Gaussian measurement noises with variance  $\sigma^2 = 0.01$ , the three state trajectories have large fluctuations.

where

$$m_t^i = \frac{1}{|\mathcal{N}_i|} \sum_{k \in \mathcal{N}_i} y_t^{ik}. \quad (4)$$

Note that the structure of (3) is very similar to the recursion in classical stochastic approximation algorithms in that  $m_t^i - x_t^i$  provides a correction term with the step size  $a_t$ . Indeed, by introducing a suitable local potential involving  $A_i$  and its neighbors,  $m_t^i - x_t^i$  may be interpreted as the noisy measurement of a scaled negative gradient along the state of  $A_i$ ; see [13] for details. Since the additive noise is contained in  $\{m_t^i, t \geq 0\}$ , each state  $x_t^i$  will fluctuate randomly. These fluctuations will not die off if the step size  $a_t$  is selected as a constant. We introduce an example for illustration.

*Example 1:* We consider a strongly connected digraph with three nodes  $\{1, 2, 3\}$  arranged on a line where node 2 is at the intermediate position and  $\mathcal{N}_1 = \{2\}$ ,  $\mathcal{N}_2 = \{1, 3\}$ ,  $\mathcal{N}_3 = \{2\}$ . We follow the notation introduced in this section, and the scalar states are updated by:  $x_{t+1}^1 = (x_t^1 + y_t^{12})/2$ ,  $x_{t+1}^2 = (x_t^2 + y_t^{21} + y_t^{23})/3$ , and  $x_{t+1}^3 = (x_t^3 + y_t^{32})/2$ ,  $t \geq 0$ . The i.i.d. Gaussian noises satisfy **(A1)** with variance  $\sigma^2 = 0.01$ .  $\square$

The simulation for Example 1 is shown in Fig. 2 with initial condition  $[x_0^1, x_0^2, x_0^3] = [4, 1, 2]$ . For the noise free case, we change the state updating rule in Example 1 by replacing  $y_t^{ik}$  by  $x_t^k$ , which reduces to a standard algorithm in the literature; see, e.g., [14]. Fig. 2 shows that the measurement noise causes a dramatic loss of convergence for this algorithm where the step size for node  $i$  is in fact  $a^{(i)} = |\mathcal{N}_i| / (|\mathcal{N}_i| + 1)$  to give equal weights  $1 - a^{(i)} = a^{(i)} / |\mathcal{N}_i| = 1 / (|\mathcal{N}_i| + 1)$ .

With the aim of getting a stable behavior for the agents, a decreasing sequence  $\{a_t, t \geq 0\}$  will be used below.

**(A2)** The sequence  $\{a_t, t \geq 0\}$  satisfies i)  $a_t \in [0, 1]$  and ii) there exists  $T_0 \geq 1$  such that

$$\frac{\alpha}{t^\gamma} \leq a_t \leq \frac{\beta}{t^\gamma} \quad (5)$$

for all  $t \geq T_0$ , where  $\gamma \in (0.5, 1]$  and  $0 < \alpha \leq \beta < \infty$ .  $\square$

By starting from a suitable  $T_0$  and requiring  $\frac{\alpha}{t^\gamma} \leq a_t$  only for  $t \geq T_0$ , where  $a_t \in [0, 1]$ , we may allow large values for  $\alpha$ . This gives more flexibility in choosing the step size sequence and otherwise  $\alpha$  greater than one would be excluded. Here the segment  $\{a_t, t < T_0\}$  may be chosen freely as long as

$a_t \in [0, 1]$ . In further analysis, the parameters  $T_0, \alpha, \beta, \gamma$  are treated as fixed constants associated with  $\{a_t, t \geq 0\}$ .

Note that (A2) implies

$$\sum_{t=0}^{\infty} a_t = \infty, \quad \sum_{t=0}^{\infty} a_t^2 < \infty, \quad (6)$$

which is a typical property for the step size sequences used in classical stochastic approximation theory [6], [15].

We introduce the following definitions on the asymptotic behavior of the agents' state evolution.

**Definition 2: (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  $\lim_{t \rightarrow \infty} E|x_t^i - x^*|^2 = 0$  for all  $i \in \mathcal{N}$ .  $\square$

**Definition 3: (strong consensus)** The agents are said to reach strong consensus if there exists a random variable  $x^*$  such that with probability one  $\lim_{t \rightarrow \infty} x_t^i = x^*$  for all  $i \in \mathcal{N}$ .  $\square$

If a sequence converges with probability one (w.p.1), we also say it converges almost surely (a.s.). In both mean square and strong consensus, the limit  $x^*$  may depend on the initial states, the noise sequence and the consensus algorithm.

#### A. The Generalization to Vector States

We give some discussions for the vector case where each individual state  $\mathbf{x}_t^k \in \mathbb{R}^d$  with dimension  $d > 1$ . It is easy to extend (1)-(2) to the vector case by taking a vector noise term. For the vector version of (2), we see that each of the  $d$  components in  $\mathbf{x}_t^k$  is decoupled from the other  $d-1$  components during the iteration. Hence we may decompose the vector equation to  $d$  scalar equations. After adapting assumption (A2) to the vector case, the consensus result in the paper is easily generalized to vector individual states.

### III. MEAN SQUARE CONSENSUS IN A TWO-AGENT MODEL

In this section we analyze a two-agent model. Apart from enabling sharp estimates and shedding light on the selection of the step size, the techniques developed for such a system will provide motivation for analyzing more general models.

For the two agent case, we have  $n = 2$  and  $\mathcal{N} = \{1, 2\}$ . Define  $\xi_t = x_t^1 - x_t^2$ . We notice the relation

$$\xi_{t+1} = (1 - 2a_t)\xi_t + a_t v_t, \quad (7)$$

where  $v_t = w_t^{12} - w_t^{21}$ . By the lower and upper bound condition in (5), we may find an integer  $T_1 > T_0$  such that

$$1 - \frac{2\alpha}{t^\gamma} \geq 1 - 2a_t > 0, \quad \text{for all } t \geq T_1. \quad (8)$$

Denote  $\bar{a}_t = 2a_t$ . In the estimate below, we start with  $T_1$  as the initial time. For  $t \geq T_1$ , it follows from (7) that

$$\begin{aligned} \xi_{t+1} &= (1 - \bar{a}_t)(1 - \bar{a}_{t-1}) \cdots (1 - \bar{a}_{T_1}) \xi_{T_1} \\ &\quad + (1 - \bar{a}_t) \cdots (1 - \bar{a}_{T_1+1}) a_{T_1} v_{T_1} \\ &\quad \vdots \\ &\quad + (1 - \bar{a}_t) a_{t-1} v_{t-1} \\ &\quad + a_t v_t. \end{aligned}$$

Define

$$\Pi_{l,k} = (1 - \bar{a}_l)(1 - \bar{a}_{l-1}) \cdots (1 - \bar{a}_{k+1}) a_k, \quad (9)$$

where  $l > k \geq T_1$ . By convention,  $\Pi_{k,k} = a_k$ .

**Lemma 4:** For  $\Pi_{l,k}$  defined by (9) with  $k \leq l$ , we have the upper bound estimate: (i) If  $\gamma = 1$ , we have

$$\Pi_{l,k} \leq \exp \left\{ -2\alpha \sum_{t=k+1}^l t^{-1} \right\} \frac{\beta}{k} \leq \frac{\beta(k+1)^{2\alpha}}{k(l+1)^{2\alpha}}. \quad (10)$$

(ii) If  $\frac{1}{2} < \gamma < 1$ , we have

$$\Pi_{l,k} \leq \exp \left\{ \frac{-2\alpha}{1-\gamma} [(l+1)^{1-\gamma} - (k+1)^{1-\gamma}] \right\} \frac{\beta}{k^\gamma}. \quad (11)$$

$\square$

Let  $\{c(t), t \geq t_0\}$  and  $\{h(t), t \geq t_0\}$  be two sequences of real numbers indexed by the integer  $t \geq t_0$ , and  $h(t) > 0$  for all  $t \geq t_0$ . We denote  $c(t) = O(h(t))$  (resp.,  $c(t) = o(h(t))$ ) if

$$\overline{\lim}_{t \rightarrow \infty} \frac{|c(t)|}{h(t)} \leq C_d < \infty, \quad \left( \text{resp., } \lim_{t \rightarrow \infty} \frac{|c(t)|}{h(t)} = 0 \right).$$

Here  $C_d$  is called a dominance constant. In practice, it is desirable to pick up a smaller value for  $C_d$  when possible.

**Lemma 5:** We have the upper bound estimate: (i) If  $\gamma = 1$ ,

$$\sum_{k=T_1}^t \Pi_{t,k}^2 = \begin{cases} O(\frac{1}{t^{4\alpha}}) & \text{if } 0 < \alpha < \frac{1}{4} \\ O(\frac{\ln t}{t}) & \text{if } \alpha = \frac{1}{4} \\ O(\frac{1}{t}) & \text{if } \alpha > \frac{1}{4}, \end{cases} \quad (12)$$

and (ii) If  $\frac{1}{2} < \gamma < 1$ ,

$$\sum_{k=T_1}^t \Pi_{t,k}^2 = O(\frac{1}{t^\gamma}). \quad (13)$$

$\square$

**Remark.** We give some discussions on estimating the dominance constant  $C_d$  in Lemma 5. For (12), when  $\alpha \neq \frac{1}{4}$  is close to  $\frac{1}{4}$  from left (resp., right), we need to take a large  $C_d$  associated with  $O(\frac{1}{t^{4\alpha}})$  (resp.,  $\frac{1}{t}$ ). For the case  $\alpha = \frac{1}{4}$  in (12), we may take  $C_d = \beta^2$ . For (13), we take  $C_d = 4\alpha$ , regardless of the particular value of  $\gamma \in (\frac{1}{2}, 1]$ .  $\square$

**Theorem 6:** Assume a system of two agents with the initial condition  $x_0^i, i = 1, 2$ , each with noisy measurements of the state of the other, and assume (A1)-(A2). Then there exists a limit random variable  $x^*$  such that

$$\lim_{t \rightarrow \infty} E|x_t^i - x^*|^2 = 0,$$

where  $i = 1, 2$ . This implies mean square consensus.

*Proof:* Letting  $z_t = \frac{1}{2}(x_t^1 + x_t^2)$  for  $t \geq 0$ , we have

$$z_{t+1} = z_t + a_t \tilde{w}_t, \quad t \geq 0, \quad (14)$$

where  $\tilde{w}_t = \frac{1}{2}(w_t^{12} + w_t^{21})$ . By iterating (14), it follows that  $z_{t+1} = z_0 + \sum_{k=0}^t a_k \tilde{w}_k$ . Since  $\sum_{t=0}^{\infty} a_t^2 < \infty$ , there exists a random variable  $z^*$  such that  $\lim_{t \rightarrow \infty} E|z_t - z^*|^2 = 0$ .

Now we estimate  $\xi_t = x_t^1 - x_t^2$ . We see that

$$E\xi_{t+1}^2 \leq (E\xi_{T_1}^2 \prod_{k=T_1}^t |1 - 2a_k|^2 + \sup_{k \geq T_1} E v_k^2 \times \sum_{k=T_1}^t \Pi_{t,k}^2)$$

and it readily follows from Lemma 5 that  $\lim_{t \rightarrow \infty} E\xi_{t+1}^2 = 0$ . The mean square consensus property follows easily.  $\square$

The i.i.d. noise assumption in Theorem 6 may be relaxed.

#### IV. STRONG CONSENSUS FOR TWO AGENTS

So far we have shown that the states of the agents converge to the same limit in mean square. It is well known that in classical stochastic approximation theory [6], [15], similarly structured algorithms have sample path convergence properties under reasonable conditions. It is tempting to analyze sample path behavior in this consensus context. Compared to the mean square consensus analysis, the proof of strong consensus requires far more analytic labor.

The following lemma is instrumental for our analysis. The proving technique relies on asymptotic analysis of the sum of random variables with weights in a double array.

*Lemma 7:* [25] Let  $\{w, w_t, t \geq 1\}$  be i.i.d. real-valued random variables with zero mean and variance  $Q$ , and  $\{a_{ki}, 1 \leq i \leq l_k \uparrow \infty, k \geq 1\}$  a double array of constants. Assume (i)  $\max_{1 \leq i \leq l_k} |a_{ki}| = O((l_k^{1/p} \log k)^{-1})$ ,  $0 < p \leq 2$ , and  $\log l_k = o(\log^2 k)$ , (ii)  $E|w|^p < \infty$ . Then

$$\lim_{k \rightarrow \infty} \sum_{i=1}^{l_k} a_{ki} w_i = 0, \quad a.s. \quad (15)$$

□

This lemma is an immediate consequence of Theorem 4 and Corollary 3 in [25] (pp. 331 and pp. 340) which dealt with the sum of weighted random variables.

We need to estimate the magnitude of the individual terms  $\Pi_{t,k}$ . Note that for each  $t \geq T_1$ ,  $\Pi_{t,k}$  is defined for  $k$  starting from  $T_1$  up to  $t$ . Hereafter, for notational brevity, we make a convention by setting  $\Pi_{t,k} \equiv 0$ , for  $1 \leq k < T_1$  when  $t \geq T_1$ , and  $\Pi_{t,k} \equiv 0$ , for  $1 \leq k \leq t$  when  $1 \leq t < T_1$ . After this extension, all the entries  $\Pi_{t,k}$  constitute a triangular array.

*Lemma 8:* For case (i) with  $\gamma = 1$ , we have

$$\sup_{1 \leq k \leq t} \Pi_{t,k} = \begin{cases} O(\frac{1}{t^{2\alpha}}) & \text{if } 0 < \alpha < \frac{1}{2}, \\ O(\frac{1}{t}) & \text{if } \alpha \geq \frac{1}{2}, \end{cases} \quad (16)$$

and for case (ii) with  $\frac{1}{2} < \gamma < 1$ , we have  $\sup_{1 \leq k \leq t} \Pi_{t,k} = O(1/t^\gamma)$ . □

*Theorem 9:* Assume all conditions in Theorem 6 hold, and  $\alpha > \frac{1}{4}$  in the case  $\gamma = 1$ . We have (a)  $z_t$  converges a.s.. (b)  $\lim_{t \rightarrow \infty} \xi_t = 0$  a.s.. (c) The two sequences  $\{x_t^1, t \geq 0\}$  and  $\{x_t^2, t \geq 0\}$  converge to the same limit a.s., which implies strong consensus.

*Proof:* Recall that  $z_{t+1} = z_0 + \sum_{k=0}^t a_k \tilde{w}_k$ , where  $\tilde{w}_t = \frac{1}{2}(w_t^{12} + w_t^{21})$ . Since  $\{\tilde{w}_k, k \geq 0\}$  is a sequence of independent random variables with  $E\tilde{w}_k = 0$ ,  $E|\tilde{w}_k|^2 < \infty$ , by Khintchine-Kolmogorov convergence Theorem (see [8], pp. 110), if  $\sum_{k=0}^{\infty} E|a_k \tilde{w}_k|^2 < \infty$ , then  $\sum_{k=0}^t a_k \tilde{w}_k$  converges a.s.. Indeed,  $\sum_{k=0}^{\infty} E|a_k \tilde{w}_k|^2 < \infty$  follows from (A2) and  $\sup_{t \geq 0} E\tilde{w}_t^2 < \infty$ . Hence  $z_t$  converges a.s..

Now we prove (b). Recalling the expression for  $\xi_{t+1}$ , we see that  $\lim_{t \rightarrow \infty} \xi_t = 0$  a.s., if  $\lim_{t \rightarrow \infty} \sum_{k=1}^t \Pi_{t,k} v_k = 0$  a.s.. By Lemma 8 we have  $\sup_{1 \leq k \leq t} \Pi_{t,k} = O((t^{1/2} \log t)^{-1})$ , for both cases: (i)  $\frac{1}{2} < \gamma < 1$ , (ii)  $\gamma = 1$  and  $\alpha > \frac{1}{4}$ .

To apply Lemma 7, we take  $l_k = k$  and  $p = 2$ , which yields  $\lim_{t \rightarrow \infty} \sum_{k=1}^t \Pi_{t,k} v_k = 0$ , a.s. Hence  $\lim_{t \rightarrow \infty} \xi_t = 0$  a.s., and (b) follows. Assertion (c) then follows from (a) and (b). □

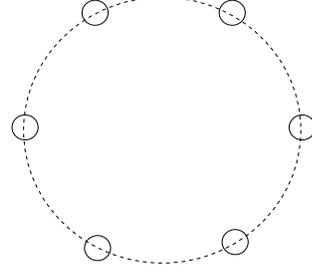


Fig. 3. A ring network where each agent has two neighbors.

The requirement  $\alpha > \frac{1}{4}$  for the case  $\gamma = 1$  is a mild condition, and from an algorithmic point of view, it is not an essential restriction since in applications  $\{a_t, t \geq 0\}$  is a sequence to be designed.

#### V. NETWORKS WITH SYMMETRY STRUCTURES

In this section we consider models where the neighboring relation for the  $n$  agents displays a certain symmetry. A simple example is shown by Fig. 3 with ring-coupled agents.

We specify the associated digraph as follows. First, the  $n$  nodes are listed by the order  $1, 2, \dots, n$ . Here  $i$  and  $i+1$  are unnecessarily neighbors to each other. The  $i$ th node has a set of neighbors  $\mathcal{N}_i$  listed as  $(\alpha_1^i, \alpha_2^i, \dots, \alpha_L^i)$  which is a subset of  $\mathcal{N} = \{1, 2, \dots, n\}$ . The fixed constant  $L \geq 1$  denotes the number of neighbors, which is the same for all nodes. Then the  $(i+1)$ th node's neighbors are given by  $(\alpha_1^i + 1, \alpha_2^i + 1, \dots, \alpha_L^i + 1)$ . In other words, by incrementing each of  $\alpha_k^i$  by one, where  $1 \leq k \leq L$ , we obtain the neighbor set for node  $i+1$ , and after a total of  $n$  steps, we retrieve node  $i$  and its neighbors  $\mathcal{N}_i$ . In fact, the underlying graph may be realized by arranging the  $n$  nodes on a ring and adding the edges appropriately. For this reason, we may term this symmetry as the circulant invariance property for the neighboring structure. Throughout this section, if an index (e.g.,  $\alpha_k^i + 1$ ) for a node or agent exceeds  $n$ , we identify it as an integer between 1 and  $n$  by taking  $\text{mod}(n)$ .

The above symmetry assumption does not ensure strong connectivity of the graph, which may be illustrated by simple examples. For this section, we make the assumption:

(A3) The digraph  $G = (\mathcal{N}, \mathcal{E})$  has both the circulant invariance property and strong connectivity. □

Define the centroid for the state configuration  $(x_t^1, \dots, x_t^n)$  as  $z_t = \frac{1}{n} \sum_{i=1}^n x_t^i$ . Under (A3), we can show that  $z_t$  satisfies

$$z_{t+1} = z_t + \frac{a_t}{nL} \sum_{i \in \mathcal{N}} \sum_{k \in \mathcal{N}_i} w_t^{ik}, \quad t \geq 0.$$

*Lemma 10:* Under (A1)-(A3), the sequence  $\{z_t, t \geq 0\}$  converges both in mean square and almost surely. □

We further denote the difference between  $x_t^{i+1}$  and  $x_t^i$  by

$$\xi_t^i = x_t^{i+1} - x_t^i, \quad (17)$$

for  $1 \leq i \leq n$ , where  $x_t^{n+1}$  is identified as  $x_t^1$  by taking  $\text{mod}(n)$  for the superscript. Thus  $\xi_t^n = x_t^1 - x_t^n$ . Here  $\xi_t^i$ ,  $1 \leq i \leq n$ , are not linearly independent since  $\xi_t^n = -\sum_{i=1}^n \xi_t^i$ . In formation

control the set of  $n-1$  variables  $\xi_t^i$ ,  $1 \leq i \leq n-1$ , is usually called the shape variables. Recall that we have  $|\mathcal{N}_i| = L$ , for all  $i \in \mathcal{N}$ . Specializing the stochastic algorithm (2) to the model of this section, we have

$$x_{t+1}^i = (1-a_t)x_t^i + \frac{a_t}{L} \sum_{k \in \mathcal{N}_i} (x_t^k + w_t^{ik}) \quad (18)$$

for each  $i \in \mathcal{N}$ , and

$$\begin{aligned} x_{t+1}^{i+1} &= (1-a_t)x_t^{i+1} + \frac{a_t}{L} \sum_{k \in \mathcal{N}_{i+1}} (x_t^k + w_t^{i+1,k}) \\ &= (1-a_t)x_t^{i+1} + \frac{a_t}{L} \sum_{k \in \mathcal{N}_i} (x_t^{k+1} + w_t^{i+1,k+1}) \end{aligned} \quad (19)$$

where we get (19) by the circulant invariance property.

Subtracting both sides of (19) by (18) leads to

$$\xi_{t+1}^i = (1-a_t)\xi_t^i + \frac{a_t}{L} \sum_{k \in \mathcal{N}_i} \xi_t^k + \frac{a_t}{L} \tilde{w}_t^i, \quad i \in \mathcal{N}, \quad (20)$$

where

$$\tilde{w}_t^{i,k} = w_t^{i+1,k+1} - w_t^{i,k}, \quad \tilde{w}_t^i = \sum_{k \in \mathcal{N}_i} \tilde{w}_t^{i,k} \quad (21)$$

with  $k \in \mathcal{N}_i$  for  $\tilde{w}_t^{i,k}$ .

*Lemma 11:* Assume **(A3)** and let  $\xi_t^i$  and  $\tilde{w}_t^i$  be defined by (17) and (21), respectively. We have the following zero-sum property:  $\sum_{i \in \mathcal{N}} \xi_t^i = 0$  and  $\sum_{i \in \mathcal{N}} \tilde{w}_t^i = 0$  for all  $t \geq 0$ .  $\square$

We introduce the  $n \times n$  stochastic matrix

$$M(a) = I + aM^c, \quad a \in [0, 1]. \quad (22)$$

The matrix  $M^c$  is given in the form

$$M^c = \begin{bmatrix} -1 & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & -1 & c_1 & \ddots & \vdots \\ c_{n-2} & c_{n-1} & -1 & \ddots & c_2 \\ \vdots & \ddots & \ddots & \ddots & c_1 \\ c_1 & \cdots & c_{n-2} & c_{n-1} & -1 \end{bmatrix}$$

where  $M_{ii}^c = -1$  for  $1 \leq i \leq n$ , and for  $2 \leq k \leq n$ ,

$$M_{1k}^c = c_{k-1} = \begin{cases} \frac{1}{L} & \text{if } k \in \mathcal{N}_1 \\ 0 & \text{otherwise.} \end{cases}$$

Since  $M^c$  is a circulant matrix [10], it is well defined after specifying the first row. Both  $M^c$  and  $M(a)$  are circulant matrices.

*Proposition 12:* Under **(A3)**,  $M(a)$  is doubly stochastic for any  $a \in [0, 1]$ , i.e., both  $M(a)$  and  $[M(a)]^T$  are stochastic matrices. In addition,  $M(a)$  is irreducible for  $a > 0$ .  $\square$

Define two  $\mathbb{R}^n$ -valued vectors  $\zeta_t = (\xi_t^1, \dots, \xi_t^n)^T$  and  $\tilde{w}_t = (\tilde{w}_t^1, \dots, \tilde{w}_t^n)^T$ . We can show that  $\zeta_t$  satisfies the recursion:

$$\zeta_{t+1} = M(a_t)\zeta_t + \frac{a_t}{L}\tilde{w}_t, \quad t \geq 0. \quad (23)$$

*Lemma 13:* Assume **(A2)**-**(A3)** hold, and the real vector  $\theta = [\theta_1, \dots, \theta_n]^T$  has a zero column sum, i.e.,  $\sum_{i=1}^n \theta_i = 0$ . Then for all  $t \geq k \geq 0$ , we have:

(i) The column sum of  $M(a_t) \cdots M(a_k)\theta$  is zero, i.e.,  $\sum_{i=1}^n M_{i,k}^\theta(i) = 0$ , where we denote  $M_{i,k}^\theta = [M_{i,k}^\theta(1), \dots, M_{i,k}^\theta(n)]^T = M(a_t) \cdots M(a_k)\theta$ .

(ii) There exist constants  $\delta^* \in (0, 1)$  and  $T_2 > 0$  which are independent of  $\theta$ , such that for all  $t \geq k \geq T_2$ ,

$$|M(a_t) \cdots M(a_k)\theta| \leq |(1 - \delta^* a_t) \cdots (1 - \delta^* a_k)\theta|,$$

where  $T_2$  is chosen such that  $a_t \leq 1/2$  for all  $t \geq T_2$ .  $\square$

Let  $\omega_n = e^{2\pi i/n}$  where  $i^2 = -1$ , and denote

$$F = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_n & \omega_n^2 & \cdots & \omega_n^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_n^{n-1} & \omega_n^{2(n-1)} & \cdots & \omega_n^{(n-1)(n-1)} \end{pmatrix}$$

which is the so called Fourier matrix of order  $n$  and satisfies  $F_n^* F_n = I$  where  $F_n^*$  is the complex adjoint matrix for  $F_n$ . For  $a \in [0, 1]$ , we introduce the polynomial

$$\varphi(a, z) = (1-a) + a(c_1 z + c_2 z^2 + \cdots + c_{n-1} z^n).$$

Then by well known results for circulant matrices [10], the  $n$  eigenvalues  $\{\lambda_{1,t}, \dots, \lambda_{n,t}\}$  of  $M(a_t)$  are given by

$$\lambda_{k,t} = \varphi(a_t, \omega_n^{k-1})$$

where  $k = 1, \dots, n$ . Obviously,  $\lambda_{1,t} = 1$ . Furthermore,  $M(a_t)$  may be diagonalized in the form

$$M(a_t) = F_n^* \times \text{Diag}(\lambda_{1,t}, \dots, \lambda_{n,t}) \times F_n.$$

Letting  $1_n = [1, \dots, 1]^T$ , it is easy to verify that

$$\begin{aligned} M(a_t) \cdots M(a_k) &= F_n^* \times \prod_{j=k}^t \text{Diag}(\lambda_{1,j}, \dots, \lambda_{n,j}) \times F_n \\ &= F_n^* \times \prod_{j=k}^t \text{Diag}(0, \lambda_{2,j}, \dots, \lambda_{n,j}) \times F_n + \frac{1}{n} 1_n 1_n^T. \end{aligned}$$

*Corollary 14:* Let  $\theta$ ,  $T_2$  and  $\delta^*$  be given as in Lemma 13 and denote  $M(t, k) = M(a_t) \cdots M(a_k)$  for  $t \geq k \geq T_2$ . Then the real matrix  $M^o(t, k) \triangleq M(t, k) - \frac{1}{n} 1_n 1_n^T$  satisfies

$$M(t, k)\theta = M^o(t, k)\theta$$

and  $|M^o(t, k)|_\infty \leq C \prod_{j=k}^t (1 - \delta^* a_j)$  for some  $C > 0$  independent of  $t$  and  $k$ , where  $|\cdot|_\infty$  denotes the largest absolute value of the elements in the matrix.  $\square$

*Theorem 15:* Assume **(A1)**-**(A3)**. Then the algorithm (2) ensures (i) mean square consensus for any  $\alpha > 0$ ; and (ii) strong consensus for (a)  $\gamma \in (0.5, 1)$  associated with any  $\alpha > 0$  in **(A2)**, and (b)  $\gamma = 1$  provided that  $\alpha > 1/(2\delta^*)$ .

*Proof:* We first recall that  $z_t$  converges in mean square and a.s.. We write the recursion (23) for  $\xi_t$ , and show its mean square convergence to zero by Lemma 11, and Lemma 13-(ii). Then mean square consensus follows. For proving almost sure convergence of  $\xi_t$ , we use Lemma 11, Corollary 14 and Lemma 7 to carry out the double array analysis.  $\square$

Without measurement noises, if the coefficient matrix in the consensus algorithm is doubly stochastic, the state average is an invariant, and it is possible to establish average-consensus [29] such that each individual state converges to the initial state average. In our model, the noise causes the limit state to deviate from the initial state average although  $M(a_t)$  is doubly stochastic. We have the deviation estimate.

*Proposition 16:* Under **(A1)**-**(A3)**, the state iterates in the consensus algorithm satisfy  $\lim_{t \rightarrow \infty} E|x_t^i - \frac{1}{n} \sum_{k=1}^n x_0^k|^2 = O(Q)$ , where  $Q$  is the variance of the i.i.d. noises, and  $x_0^i$ ,  $1 \leq i \leq n$ , denotes the initial state at  $t = 0$ .  $\square$

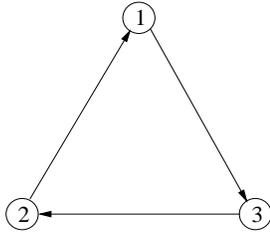


Fig. 4. A digraph with 3 nodes.

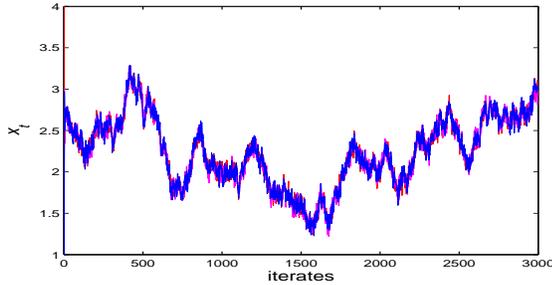


Fig. 5. Equal weights (i.e., 0.5) are used for each agent's state and its observation in the state iteration.

## VI. SIMULATIONS

We consider the model in Fig. 4 where  $\mathcal{N}_1 = \{2\}$ ,  $\mathcal{N}_2 = \{3\}$  and  $\mathcal{N}_3 = \{1\}$ . The initial condition for  $x_t = [x_t^1, x_t^2, x_t^3]$  is  $[4, 3, 1]$  at  $t = 0$ , and the i.i.d. Gaussian noise has variance  $\sigma^2 = 0.01$ . Fig. 5 shows the simulation with equal weights (as in Example 1) in the averaging rule ( $x_{t+1}^1 = (x_t^1 + y_t^{12})/2$ , etc.), without obtaining consensus. Fig. 6 shows the strong consensus result achieved by algorithm (2) with the step size sequence  $\{a_t = (t + 5)^{-0.85}, t \geq 0\}$ .

## REFERENCES

- [1] S. Aranda, S. Martinez, and F. Bullo. On optimal sensor placement and motion coordination for target tracking. *Proc. IEEE Int. Conf. on Robotics and Automat.*, Barcelona, Spain, pp. 4544-4549, April 2005.
- [2] P.-A. Bilman and G. Ferrari-Trecate. Average consensus problems in networks of agents with delayed communications. *Proc. 44th IEEE CDC-ECC'05*, Seville, Spain, pp. 7066-7071, Dec. 2005.
- [3] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. *Proc. IEEE CDC-ECC'05*, Seville, Spain, pp. 2996-3001, Dec. 2005.
- [4] V. Borkar and P. Varaiya. Asymptotic agreement in distributed estimation. *IEEE Trans. Automat. Contr.*, vol. 27, pp. 650-655, June 1982.
- [5] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri. Communication constraints in coordinated consensus problems. *Proc. American Control Conference*, Minneapolis, MN, pp. 4189-4194, June 2006.
- [6] H.-F. Chen. *Stochastic Approximation and Its Applications*. Kluwer, Boston, 2002.
- [7] Y. S. Chow and T. L. Lai. Limiting behavior of weighted sums of independent random variables. *Annals of Probability*, vol. 1, no. 5, pp. 810-824, 1973.
- [8] Y. S. Chow and H. Teicher. *Probability Theory: Independence, Interchangeability, Martingales*. Springer-Verlag, New York, 1978.
- [9] J. Cortes, S. Martinez, and F. Bullo. Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions. *IEEE Trans. Automat. Control*, vol. 51, no. 8, pp. 1289-1298, Aug. 2006.
- [10] P. J. Davis. *Circulant Matrices*. John Wiley, 1979.
- [11] J. P. Desai, V. Kumar, and J. P. Ostrowski. Control of changes in formation for a team of mobile robots. *Proc. IEEE Int. Conf. Robotics and Automation*, Detroit, Michigan, pp. 1556-1561, May 1999.

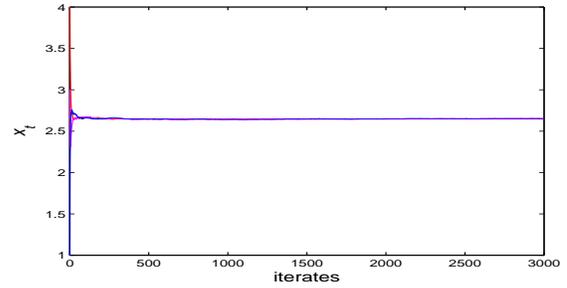


Fig. 6. Achieving strong consensus by stochastic approximation.

- [12] A. Giridhar and P. R. Kumar. Toward a theory of in-network computation in wireless sensor networks. *IEEE Commun. Mag.*, pp. 98-107, April 2006.
- [13] M. Huang and J. H. Manton. Stochastic Lyapunov analysis for consensus algorithms with noisy measurements. *Proc. American Control Conference*, New York, 2007.
- [14] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Automat. Contr.*, vol. 48, no. 6, pp. 988-1000, June 2003.
- [15] H. J. Kushner and G. G. Yin. *Stochastic Approximation Algorithms and Applications*. Springer, New York, 1997.
- [16] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. E. Davis. Collective motion, sensor networks and ocean sampling. *Proc. IEEE: Special Issue on Technology of Networked Control Systems*, vol. 95, no. 1, pp. 48-74, 2007.
- [17] J. A. Marshall and M. E. Broucke. On invariance of cyclic group symmetries in multiagent formations. *Proc. IEEE CDC-ECC'05*, Seville, Spain, pp. 746-751, Dec. 2005.
- [18] J. A. Marshall, M. E. Broucke, and B. A. Francis. Formations of vehicles in cyclic pursuit. *IEEE Trans. Automatic Control*, vol. 49, no. 11, pp. 1963-1974, Nov. 2004.
- [19] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Trans. Automat. Control*, vol. 50, no. 2, pp. 169-182, Feb. 2005.
- [20] R. Olfati-Saber. Flocking for multi-agent dynamic systems: algorithms and theory. *IEEE Trans. Automat. Control*, vol. 51, no. 3, pp. 401-420, Mar. 2006.
- [21] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Automatic Control*, vol. 49, no. 9, pp. 1520-1533, Sep. 2004.
- [22] W. E. Pruitt. Summability of independent random variables. *Journal of Mathematics and Mechanics*, vol. 15, no. 5, pp. 769-776, 1966.
- [23] W. Ren, R. W. Beard, and D. B. Kingston. Multi-agent Kalman consensus with relative uncertainty. *Proc. American Control Conf.*, Portland, OR, pp. 1865-1870, June 2005.
- [24] H. G. Tanner, G. J. Pappas, and V. Kumar. Input-to-state stability on formation graphs. *Proc. IEEE Decision Contr. Conf.*, Las Vegas, NV, pp. 2439-2444, Dec. 2002.
- [25] H. Teicher. Almost certain convergence in double arrays. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 69, pp. 331-345, 1985.
- [26] J. N. Tsitsiklis and M. Athans. Convergence and asymptotic agreement in distributed decision problems. *IEEE Trans. Automat. Control*, vol. 29, no. 1, pp. 42-50, Jan. 1984.
- [27] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Trans. Automat. Contr.*, vol. 31, no. 9, pp. 803-812, 1986.
- [28] P. K. Visscher. How self-organization evolves. *Nature*, vol. 421, pp. 799-800, Feb. 2003.
- [29] L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. *Journal of Parallel and Distributed Computing*, vol. 67, pp. 33-46, 2007.