2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

On Consensus over Stochastically Switching Directed Topologies

Sundaram Vanka, Vijay Gupta and Martin Haenggi

***Abstract*— We consider average consensus algorithms executed over stochastically varying communication topologies that may be unbalanced. It is known that the state values will reach consensus, under fairly weak conditions. However, the consensus value is a random variable. We provide concentration bounds for the distance of the state vector from the consensus subspace and for the asymptotic distribution of the value to which the various nodes converge as they reach consensus. The results allow the analysis of average consensus over wireless communication networks with more realistic assumptions than before.**

I. Introduction

Consensus algorithms are decentralized algorithms that aim at achieving agreement among all participating agents over the value of a quantity [1]. Such algorithms have seen renewed focus over the last decade [2], [3], [4], [5], [6]. Conditions for convergence of the agents to a common value are now well characterized in both deterministic [5], [6], [7] and stochastic frameworks [8], [9], [10], [11], [12], [13]. Recent work characterizes the rate of convergence by imposing communication-related constraints between nodes. Effects such as quantization [14], delays [15],additive noise [16], packet loss [17], [9] and power constraints [18] have been analyzed.

We will refer to the interconnection topology that determines message passing betweeen nodes as the *communication graph*. Evidently the underlying communication infrastructure of a network determines feasible interconnection topologies. However, in wireless networks, connections formed are often probabilistic and subject to fading, interference, and noise. Given that fact that the decentralized nature of consensus algorithms is well-suited for ad-hoc wireless networks, it is important to study the convergence of properties of these algorithms on wireless networks. If at each iteration the communication graph remains balanced and connected, it is known [6] that the final value is the average of the initial values. From a wireless network standpoint, such a requirement translates to bidirectional link failures or a perfect acknowledgement mechanism. Neither of these assumptions hold in practice. A more realistic approach must account for the inherent randomness in connectivity. Thus it is more natural to view network connectivity - and hence communication graphs - as the function of a possibly unknown network state that is varies randomly with time. In this more general scenario, therefore, the consensus value is not necessarily the average of the initial values. In fact, it is a random variable whose value depends on the set of communication graphs chosen over time. From both an analysis and a design perspective, it is important to characterize this random variable, and the rate of convergence towards the final value. In this paper, we provide concentration bounds on the distribution of the node values at any time, and, in particular, the final value.

The paper is organized as follows. We begin by formulating the problem in the next section. The node value distribution is characterized by the distribution of the values around the average and the distance from the consensus subspace. These characterizations are provided in Sections III and IV respectively. Section V presents a numerical example.

The authors are with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 (e-mail: svanka@nd.edu, vgupta2@nd.edu, mhaenggi@nd.edu). The partial support of NSF (grant CNS 04-47869) and DTRA (grant N00164-07-8510) is gratefully acknowledged. The work for the second author was supported partially by the NSF award 0846631.

II. Problem Formulation

A. Network Model and Geometric Interpretation of Consensus

Consider a collection of N agents (also called nodes) $\mathcal{V} = \{0, 1, \ldots, N-1\}$. For each node i and at any (discrete) time t, we specify an in-neighbor set $\mathcal{N}_i(t) \subseteq \mathcal{V}$ such that for any node $j \in \mathcal{N}_i(t)$, node i has access to the value from node j at time t. Each node i has its initial value $x_i(0)$. If we represent the value held by the i-th node at time t by $x_i(t)$, the (discrete-time) average consensus algorithm proceeds by each node updating its value as

$$x_i(t+1) = x_i(t) - T_s \sum_{j \in \mathcal{N}_i(t)} (x_i(t) - x_j(t)), \qquad (1)$$

where T_s is a small positive constant. If we denote this exchange of information by an edge from node j to node i, we can define a digraph $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$ with $\mathcal{G}(t) \in \mathcal{U}$, where $\mathcal{U} \triangleq \{\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{M-1}\}$ is the feasible set. However, each node has access to its own value at every time step.

Let $\{\mathcal{G}(t)\} = \{\mathcal{G}(0), \mathcal{G}(1), \ldots\}$ be a graph-valued sequence where $\mathcal{G}(t) \in \mathcal{U}$. Defining $L_{\mathcal{G}}$ as the Laplacian of a graph $\mathcal{G}$, the consensus iteration for all nodes together is

$$x(t+1) = W(t)x(t) \qquad (2)$$

where $x(t)$ is an $N \times 1$ state vector formed by stacking the values $x_0(t)$, $x_1(t)$, $\cdots$, $x_{N-1}(t)$ and $W(t) \triangleq I - T_s L_{\mathcal{G}(t)}$ (I is the $N \times N$ identity matrix). $W(t)$ is a stochastic matrix, irrespective of $\mathcal{G}(t)$.

978-1-4244-4524-0/09/$25.00

Thus we can associate with each graph-valued sequence $\{\mathcal{G}(t)\}$ a matrix-valued sequence $\{W(t)\}$ where $W(t) = I - T_s L_{\mathcal{G}(t)}$. Therefore, starting from an initial condition $x(0)$, one can view equation (2) as a linear iteration over $x(t)$ where the switching sequence $W(t)$ is drawn from a set $\mathcal{W} = \{W_0, W_1, \cdots, W_{M-1}\}$, with $W_k = I - T_s L_{\mathcal{G}_k}$. In wireless networks, whether or not a particular edge in $\mathcal{G}(t)$ exists depends on the interference and fading present in the respective wireless channels at time t. For every transmitter, the successful reception of its message at an intended receiver depends on the observed Signal-to-Interference-and-Noise-Ratio (SINR) at the receiver. The receive SINR in turn depends on the state of the channel and the existence of interferers, both of which are random. Consquently, the graph process $\mathcal{G}(t)$ and hence the matrix process $W(t)$, are random. We make the following assumptions in this paper:

• (*Assumption **A1***) The link states, and hence the matrix sequence $W(t)$, are i.i.d. across time.

• (*Assumption **A2***) The diagonal elements $W_{ii} > 0$ for all $W \in \mathcal{W}$. Defining $\bar{W} \triangleq \mathbb{E}[W]$, we also assume that the Laplacian $\bar{L}$ associated with $\bar{W} = I - T_s\bar{L}$ corresponds to a strongly connected weighted digraph.

• (*Assumption **A3***) $\bar{W}$ is doubly stochastic.

Note that, contrary to [12], [13], we do not assume that the edge states (i.e., whether an edge is present or not) are independent at any given time.

Using (2), the state $x(t)$ can be written as

$$x(t) = W(t-1)\cdots W(1)W(0)x(0) = \prod_{k=0}^{t-1} W(k)x(0).$$

Observe that $x(t)$ is a random variable whose value is determined by the stochastic process $W(t)$ and the initial state $x(0)$. Define $\bar{x}(t)$ as the expected value of $x(t)$. Using Assumption **A1**, we obtain (for a given initial condition $x(0)$),

$$\bar{x}(t) = \bar{W}^t x(0). \tag{3}$$

By Assumption **A3**, the average of $\bar{x}(t)$ is preserved, since

$$\mathbf{1}^*\bar{x}(t) = \mathbf{1}^*\bar{W}^t x(0) = \mathbf{1}^* x(0),$$

where $\mathbf{1} \triangleq [\;1 \quad 1 \quad \ldots \quad 1]^*$ and M^* denotes the transpose of the matrix or the vector M. Moreover, $\bar{x}(t)$ converges asymptotically to the average consensus value $x_{av}(0)\mathbf{1}$, where we introduce the notation $x_{av}(t) \triangleq \frac{1}{N}\sum_{n=1}^{N} x_n(t)$. To see this, note that

$$\bar{x}(t) - \frac{\mathbf{1}\mathbf{1}^*}{N}x(0) = (\bar{W} - \frac{\mathbf{1}\mathbf{1}^*}{N})\bar{x}(t-1).$$

Since by Assumption **A2**, $\bar{W}$ corresponds to a strongly connected graph, the spectral radius of $\left(\bar{W} - \frac{\mathbf{1}\mathbf{1}^*}{N}\right)$ is strictly less than one [19], and consequently,

$$\lim_{t\to\infty} \bar{x}(t) = \frac{\mathbf{1}\mathbf{1}^*}{N}x(0) \equiv x_{av}(0)\mathbf{1}.$$

It was also shown recently [11] that $x(t)$ converges almost surely to *some* consensus point if **A2** holds, i.e.,

$$\lim_{t\to\infty} x(t) = \alpha\mathbf{1}, \tag{4}$$

almost surely for some random α. We now build on a geometric interpretation from [11], and introduce notation that will be used later. Let $\|v\|$ denote the Euclidean norm of vector v. We can write

$$\|x(t) - \mathbf{1}x_{av}(0)\|^2 = \|x(t) - \mathbf{1}x_{av}(t) + \mathbf{1}x_{av}(t) - \mathbf{1}x_{av}(0)\|^2.$$

Expanding the right hand side yields

$$\begin{aligned}\|x(t) - \mathbf{1}x_{av}(0)\|^2 &= \|x(k) - \mathbf{1}x_{av}(t)\|^2 \\ &\quad + \|\mathbf{1}x_{av}(t) - \mathbf{1}x_{av}(0)\|^2 \\ &\quad + 2\left\langle x(t) - \mathbf{1}x_{av}(t), \mathbf{1}x_{av}(t) - \mathbf{1}x_{av}(0)\right\rangle,\end{aligned}$$

where $\langle .,. \rangle$ denotes the standard inner product in $\mathbb{R}^N$. This inner product, in turn, vanishes since

$$\begin{aligned}&\left\langle x(t) - \mathbf{1}x_{av}(t), \mathbf{1}x_{av}(t) - \mathbf{1}x_{av}(0)\right\rangle \\ &= \left\langle x(t) - \mathbf{1}x_{av}(t), \mathbf{1}\right\rangle x_{av}(t) - \left\langle x(t) - \mathbf{1}x_{av}(t), \mathbf{1}\right\rangle x_{av}(0) \\ &= Nx_{av}^2(t) - Nx_{av}^2(t) - Nx_{av}(t)x_{av}(0) + Nx_{av}(t)x_{av}(0) \\ &= 0.\end{aligned}$$

Thus, we obtain

$$\begin{aligned}&\|x(t) - \mathbf{1}x_{av}(0)\|^2 \\ &= \|x(t) - \mathbf{1}x_{av}(t)\|^2 + \|\mathbf{1}x_{av}(t) - \mathbf{1}x_{av}(0)\|^2 \\ &= \|r(t)\|^2 + \|e(t)\|^2\end{aligned} \tag{5}$$

where $r(t) \triangleq x(t) - \mathbf{1}x_{av}(t)$ and $e(t) \triangleq \mathbf{1}x_{av}(t) - \mathbf{1}x_{av}(0)$.

A geometric interpretation of this result is shown in Figure 1 for the case of two agents. In the state space of the agent values, the consensus subspace is the straight line $x_1 = x_2 = \cdots = x_N$. The term $\|r(t)\|$ represents the (residual) distance of $x(t)$ from its closest point in the consensus subspace, while $\|e(t)\|$ represents the distance from this closest point and the average consensus point.

B. Specific Problems Considered in this Work

Equating the infinity-norms on both sides in (2), we begin

$$\begin{aligned}\|x(t)\|_\infty &= \|W(t-1)x(t-1)\|_\infty \\ &\le \|W(t-1)\|_\infty \|x(t-1)\|_\infty = \|x(t-1)\|_\infty,\end{aligned}$$

where we exploit the sub-multiplicativity of the infinity-norm and the fact that $W(t)$ is stochastic. A recursive application of this result yields $\|x(t)\|_\infty \le \|x(0)\|_\infty$. Thus, a straightforward application of the Dominated Convergence Theorem [20] allows us to interchange the integral and the limit and write $\mathbb{E}\left[\lim_{t\to\infty} x(t)\right] = \lim_{t\to\infty}\mathbb{E}[x(t)]$. In light of (4) and (3), this simplifies to $\mathbb{E}[\alpha] = x_{av}(0)$. Thus, the expectation of the final state remains the average consensus point. In this paper, we extend this result by answering the following questions:

1) What is the distribution of α?

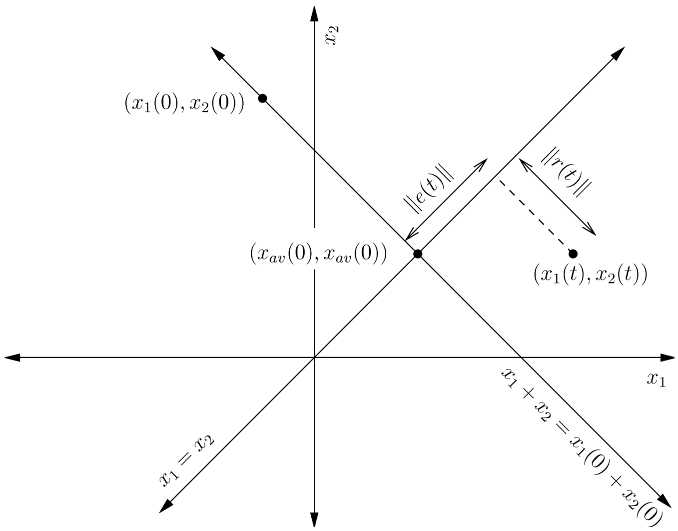

Fig. 1. Geometric interpretation of (5) for $N = 2$. The average consensus point is the intersection of the straight lines $x_1 + x_2 = x_1(0) + x_2(0)$ with the straight line $x_1 = x_2$, shown by the point $(x_{av}(0), x_{av}(0))$. We find concentration bounds on the distributions of r and e for all $t \geq 0$.

2) What is the distance $r(t)$ of the state $x(t)$ from the consensus subspace at any time t given an initial state $x(0)$ and a random switching sequence $\{W(t)\}$?

III. Asymptotic Distribution of α

To find the distribution of α, we need to characterize $e(t) = \mathbf{1}(x_{av}(t) - x_{av}(0))$. We can obtain a slightly easier quantity to consider by rewriting

$$e(t) = \frac{\mathbf{1}}{N}(\mathbf{1}^* x(t) - \mathbf{1}^* x(0)) = \left(\frac{\mathbf{1}^* \prod_{l=0}^{t-1} W(l)x(0) - \mathbf{1}^* x(0)}{N}\right)\mathbf{1} \triangleq \delta(t)\mathbf{1}.$$

Since the sequence $\{W(t)\}$ is uncorrelated and $\bar{W}$ is doubly stochastic,

$$\mathbb{E}\delta(t) = \frac{\mathbf{1}^*(\bar{W}^t - I)x(0)}{N} = 0.$$

Before proving the concentration result, we prove two preliminary results and introduce some notation. The first result provides a partial characterization of the distribution of the weight matrix $W(t)$ at any time t.

Lemma 3.1: Denote the (i, j)-th element of $\bar{W}$ by $\bar{w}_{ij}$. For any $t \geq 0$, if $\mathcal{G}_m \in \mathcal{U}$ is chosen with probability p_m

$$\|W(t) - \bar{W}\|_\infty \leq 2T_s \max_i \sum_{j \neq i}(1 - \bar{w}_{ij}), \sum_{j \neq i} \bar{w}_{ij}.$$

Proof: Let $X_{ij}(t)$ $(i \neq j)$ be a random variable which equals one when there is a directed edge from j to i at time t, and zero otherwise. Thus, X_{ij} is Bernoulli with parameter $\bar{w}_{ij}(t)$ where

$$\bar{w}_{ij} = \sum_{m=0}^{M-1} \mathbf{1}_{(i,j) \in \mathcal{E}_m} p_m.$$

Define

$$X_{ii}(t) \triangleq \sum_{j \neq i} X_{ij}(t)$$
$$\bar{w}_{ii}(t) \triangleq \mathbb{E}X_{ii}(t) = \sum_{j \neq i} \bar{w}_{ij}(t).$$

Therefore $W(t)$ is a random matrix whose entries are $W_{ij}(t) = \Delta_{ij} - T_s X_{ij}(t)$, where Δ_{ij} is the Kronecker delta. We thus obtain

$$\|W(t) - \bar{W}\|_\infty = T_s \max_i \sum_j \left|W_{ij}(t) - \bar{W}_{ij}\right| = T_s \max_i f_i(\{X_{ij}(t)\}), \tag{6}$$

where

$$f_i(\{X_{ij}(t)\}) \triangleq \left|\sum_{j \neq i}(X_{ij}(t) - \bar{w}_{ij})\right| + \sum_{j \neq i}|X_{ij}(t) - \bar{w}_{ij}|$$

is a function on $\{0, 1\}^N$.

Define $\mathcal{I}_i(t)$ to be the neighbor set of node i in the t^{th} iteration, i.e., $\mathcal{I}_i(t) \triangleq \{j : X_{ij}(t) = 1, j \neq i\}$. Evidently, For a fixed value of i, assume that X_{ij}'s in some index set $j \in \mathcal{I} \subseteq \mathcal{N} \setminus \{i\}$ assume the value one and for values of j outside this index set, assume the value zero. Then,

$$f_i(\{X_{ij}(t)\}) = \left||\mathcal{I}_i(t)| - \sum_{j \neq i} \bar{w}_{ij}\right| + |\mathcal{I}_i(t)| - \sum_{j \in \mathcal{I}_i(t)} \bar{w}_{ij} + \sum_{j \in \mathcal{N} \setminus \mathcal{I}_i(t)} \bar{w}_{ij}.$$

Define $Y_i(t) = |\mathcal{I}_i(t)| - \sum_{j \neq i} \bar{w}_{ij}$. If $Y \geq 0$ we obtain

$$f_i(\{X_{ij}(t)\}) = 2\sum_{j \in \mathcal{I}_i(t)}(1 - \bar{w}_{ij}) \leq 2\sum_{j \neq i}(1 - \bar{w}_{ij}). \tag{7}$$

Similarly, if $Y < 0$, then we obtain

$$f_i(\{X_{ij}(t)\}) = 2\sum_{j \in \mathcal{N} \setminus \mathcal{I}_i(t)} \bar{w}_{ij} \leq 2\sum_{j \neq i} \bar{w}_{ij}. \tag{8}$$

The result follows by substituting (7) and (8) in (6). ■

For future reference, define

$$C_1 \triangleq 2T_s \max_i \sum_{j \neq i}(1 - \bar{w}_{ij}), \sum_{j \neq i} \bar{w}_{ij}. \tag{9}$$

Lemma 3.2: Denote the eigenvalues of $\bar{W}$ by $1 = \lambda_0 > \lambda_1 \geq \cdots \geq \lambda_{N-1} > -1$ and let $\bar{W} = S^{-1}\Lambda S$ be its spectral decomposition. Let s_i denote the i^{th} (normalized) eigenvector, and $v_i = [\ v_i^{(0)} \quad v_i^{(1)} \quad \ldots \quad v_i^{(N-1)}\ ]$ the i^{th} row of S^{-1}. If $\bar{W}$ is normal or has no repeated eigenvalues, then for any $t \geq 0$, we have

$$\|\bar{W}^t - \frac{\mathbf{1}\mathbf{1}^*}{N}\|_\infty \leq C_2 \mu^t,$$

where

$$\mu \triangleq \max_{i=1,2,\cdots,N-1} |\lambda_i|$$
$$C_2 \triangleq \sum_{i=1}^{N-1} \sum_{j=0}^{N-1} \left|v_i^{(j)}\right|.$$

Proof: From the spectral decomposition we see that for any positive integer t, $\bar{W} = S^{-1}\Lambda S$ implies $\bar{W}^t = S^{-1}\Lambda^t S$. Since $\bar{W}$ is doubly stochastic, $\mathbf{1}^*$ is a left eigenvector with eigenvalue 1. As a result,

$$\bar{W}^t = N^{-1}\ \mathbf{1}\mathbf{1}^* + \sum_{i=1}^{N-1} \lambda_i^t s_i v_i^*.$$

Define $A = a_{mn} \triangleq \bar{W}^t - N^{-1}\mathbf{1}\mathbf{1}^*$,. Therefore,

$$\begin{aligned}\|A\|_\infty &= \max_m \sum_{n=0}^{N-1} |a_{mn}| \\ &= \max_m \sum_{n=0}^{N-1} \left| \sum_{i=1}^{N-1} \lambda_i^t s_i^{(m)} v_i^{(n)} \right| \\ &\le \sum_{n=1}^{N} \max_{\lambda_i \neq 1} \left|\lambda_i^t\right| \sum_{i=1}^{N-1} \left|v_i^{(n)}\right| \\ &\le \max_{\lambda_i \neq 1} |\lambda_i|^t \sum_{i=1}^{N-1} \sum_{n=0}^{N-1} \left|v_i^{(n)}\right|.\end{aligned}$$

Hence the result follows. ■

This result can be extended to matrices $\bar{W}$ that are not normal and have repeated eigenvalues. Assume that $\bar{W}$ has p (possibly repeated) eigenvalues $\lambda_0 = 1, \lambda_1, \lambda_2, \ldots, \lambda_p$ with multiplicities $1, r_1, \ldots, r_p$ respectively. Denote $r' = \max(\{r_i\}) - 1$. We can use a Jordan decomposition to obtain

$$\bar{W}^t = \frac{\mathbf{1}\mathbf{1}^*}{N} + \sum_{i=1}^{p} \sum_{l=1}^{r_i} \sum_{m=1}^{l} \binom{t}{l-m} \lambda_i^{t+m} s_{i,m} v_{i,l}^*,$$

where $s_{i,j}$ denotes the j^{th} (generalized) eigenvector of the i^{th} eigenvalue and $v_{i,j}$ denotes the corresponding row in S^{-1}. Using similar arguments as above, we can prove that

$$\|\bar{W}^t - N^{-1}\mathbf{1}\mathbf{1}^*\|_\infty \le C_2 \binom{t}{r'} \mu^{t-r'}.$$

For the sake of notational simplicity, unless otherwise stated, we will assume that $\bar{W}$ is normal or has distinct eigenvalues. The results presented below can be extended to more general cases, and we will mention briefly the general statements towards the end.

The main result in this section characterizes the distribution of the projection of the state $x(t)$ on to the consensus space from the average consensus point.

Proposition 3.3: Consider the problem formulation as posed in Section II. Then the distribution of $\delta(t)$ at any time $t \ge 0$ satisfies

$$\mathbb{P}(|\delta(t)| \ge \epsilon) \le \min\left\{1, 2\exp\left(-\epsilon^2 \beta(t)\right)\right\}$$

where

$$\beta(t) \triangleq \frac{(1-\mu^2)}{2C^2 \|x(0)\|_\infty^2 (1-\mu^{2t})} \tag{10}$$

and $C \triangleq C_1 C_2$ with C_1 and C_2 as defined in (9) and Lemma 3.2 respectively.

Proof: The value of $\delta(t)$ depends on the choices $W(0)$, $\cdots$, $W(t-1)$. Denote the conditional expectation of $\delta(t)$ given the past k values $W(t-k)$, $\cdots$, $W(t-1)$ as

$$Z_k(t) \triangleq \mathbb{E}\left[\delta(t) \mid \{W(t-1)\ldots,W(t-k)\}\right].$$

From the definition,

$$Z_k(t) = \frac{\mathbf{1}^*\left(W(t-1)\cdots W(t-k)\bar{W}^{t-k}x(0) - x(0)\right)}{N}.$$

In particular, $Z_0(t) = 0$ and $Z_t(t) = \delta(t)$. By construction, the sequence $\{Z_k(t)\}$ is a martingale in k for a given t [21]. We aim to bound the increase between successive steps of the martingale $\{Z_k(t)\}$. From the definition,

$$\begin{aligned}|Z_k(t) - Z_{k-1}(t)| = \big|N^{-1}\mathbf{1}^* W(t-1)\cdots W(t-k+1) \\ \left(W(t-k) - \bar{W}\right)\bar{W}^{t-k}x(0)\big|. \end{aligned} \tag{11}$$

Using the spectral expansion

$$\bar{W} = \frac{\mathbf{1}\mathbf{1}^*}{N} + \sum_{i=2}^{N} \lambda_i s_i v_i^*$$

and the fact that $W(t)$ is always stochastic, we obtain

$$\left(W(t-k) - \bar{W}\right)\bar{W}^{t-k} = \left(W(t-k) - \bar{W}\right)\sum_{i=2}^{N} \lambda_i^{t-k} s_i v_i^*.$$

Using this result in (11), we obtain

$$\begin{aligned}&|Z_k(t) - Z_{k-1}(t)| \\ &= \big|N^{-1}\mathbf{1}^* W(t-1)\cdots W(t-k+1))\left(W(t-k) - \bar{W}\right) \\ &\quad \cdot \sum_{i=2}^{N} (\lambda_i^{t-k} s_i v_i^*) x(0)\big| \\ &\le \left\|N^{-1}\mathbf{1}^*\right\|_\infty \|W(t-1)\cdots W(t-k+1)\|_\infty \\ &\left\|W(t-k) - \bar{W}\right\|_\infty \left\|\sum_{i=2}^{N} \lambda_i^{t-k} s_i v_i^*\right\|_\infty \|x(0)\|_\infty, \end{aligned} \tag{12}$$

We now upper bound the right hand side. It is easy to verify that $\left\|N^{-1}\mathbf{1}\right\|_\infty = 1$. Moreover, since each $W(k)$ is stochastic and the ∞-norm is sub-multiplicative,

$$\|W(t-1)\cdots W(t-k+1)\|_\infty \le \prod_{l=1}^{k-1} \|W(t-l)\|_\infty \le 1.$$

Upper bounds for the third and the fourth terms have already been obtained in Lemma 3.1 and Lemma 3.2, respectively. Thus for a given $x(0)$,

$$|Z_k(t) - Z_{k-1}(t)| \le C\mu^{t-k} \|x(0)\|_\infty.$$

Applying Azuma's inequality [21] yields

$$\mathbb{P}(|Z_t(t) - Z_0(t)| \ge \epsilon) \le 2\exp\left(-\beta(t)\epsilon^2\right),$$

where

$$\begin{aligned}\beta(t) &= \frac{1}{2C^2 \|x(0)\|_\infty^2 \sum_{k=0}^{t} \mu^{2(t-k)}} \\ &= \frac{1-\mu^2}{2C^2 \|x(0)\|_\infty^2 (1-\mu^{2t})}.\end{aligned}$$

Since the measure is probability, the result follows. ■

Note that this results holds for all t. As a special case, we obtain the asymptotic distribution as $t \to \infty$.

Corollary 3.4: The asymptotic distribution of α obeys the inequality

$$\mathbb{P}(|\alpha| \geq \epsilon) \leq 2\exp\left(-\beta(\infty)\epsilon^2\right) = 2\exp\left(-\frac{\epsilon^2(1-\mu^2)}{2C^2\left\|x(0)\right\|_\infty^2}\right). \quad (13)$$

Proof: Taking the limit $t \to \infty$ in the result from Proposition 3.3. ■

We conclude this section with two remarks. First, (13) can be used to bound any moment of α by using the following identity [20, p. 198] for a random variable X,

$$\mathbb{E}|X|^r = \int |X|^r d\mu \equiv r\int_0^\infty u^{r-1}\lambda(u)du.$$

where $r > 0$ and $\lambda(u) \triangleq \mathbb{P}(|X| \geq u)$. Since Proposition 3.3 provides a function upper bounds $\lambda(u)$ at every point,

$$\mathbb{E}|X|^r \leq r\int_0^\infty u^{r-1}\eta(u)du.$$

For example, the asymptotic variance can be bounded as

$$\lim_{t\to\infty}\mathbb{E}|\delta(t)|^2 \leq 2\lim_{t\to\infty}\left(\int_0^\infty u\exp\left(-u^2\beta(t)\right)du\right) = \frac{1}{\beta(\infty)} = \frac{2C^2\left\|x(0)\right\|_\infty^2}{1-\mu^2}.$$

The second remark concerns the case where $\bar{W}$ may have repeated eigenvalues or is not normal. Using similar arguments as for the case of non-repeated eigenvalues, one obtains a similar bound as in (13). However, instead of a geometric series, we get an expression of the form

$$\mu^{-2r'}\sum_{k=0}^{t}\binom{k}{r'}^2\mu^{2k}.$$

As $t \to \infty$, this converges to a hypergeometric series [22]

$$\lim_{t\to\infty}\sum_{k=0}^{t}\binom{k}{r'}^2\mu^{2k} = {}_3F_2(1,1,1;-r'+1,-r'+1;\mu^2),$$

where

$${}_mF_n(a_1,a_2\ldots,a_m;b_1,b_2\ldots,b_n;z) \triangleq \sum_{l=0}^{\infty}\frac{\alpha_l z^l}{l!},$$

with $\alpha_0 = 1$ and

$$\frac{\alpha_{l+1}}{\alpha_l} = \frac{\prod_{i=1}^m(l+a_i)}{\prod_{j=1}^n(l+b_j)}z.$$

IV. Convergence to Consensus Subspace

We now use similar techniques as above to provide concentration bounds on the distance $r(t)$ of the vector $x(t)$ from the consensus subspace at any time t. We can write

$$\begin{aligned} r(t) &= x(t) - \mathbf{1}x_{av}(t) \\ &= (I - N^{-1}\mathbf{1}\mathbf{1})x(t) \\ &= PW(t-1)W(t-2)\cdots W(0)x(0), \end{aligned}$$

where $P = I - N^{-1}\mathbf{1}\mathbf{1}^*$ is the projection operator. Once again, since $\{W(t)\}$ is uncorrelated, $\mathbb{E}r(t) = P\bar{W}^t x(0)$. The martingale $\{Y_k(t)\}$ in this case is given by

$$\begin{aligned} Y_k(t) &= \mathbb{E}\left[r(t)|\left\{W(t-1),W(t-2),\ldots W(t-k)\right\}\right] \\ &= PW(t-1)W(t-2)\cdots W(t-k)\bar{W}^{t-k}x(0). \end{aligned}$$

In particular, $Y_0(t) = P\bar{W}^t x(0)$ and $Y_t(t) = r(t)$. To bound the increase in successive steps of the martingale, consider

$$\begin{aligned} \|Y_k(t) - Y_{k-1}(t)\| &= \|PW(t-1)\cdots W(t-k+1) \\ &\quad \left(W(t-k)-\bar{W}\right)\bar{W}^{t-k}x(0)\|_\infty \\ &\leq \|P\|_\infty\|W(t-1)\cdots W(t-k+1)\|_\infty \\ &\quad \|\left(W(t-k)-\bar{W}\right)\bar{W}^{t-k}x(0)\|_\infty, \end{aligned}$$

by using the sub-multiplicative property of the ∞-norm. Since P is the projection operator onto the one-dimensional consensus subspace, $\|P\|_\infty = 1$. For the rest of the terms, a similar approach as in (11) yields

$$\|Y_k(t) - Y_{k-1}(t)\|_\infty \leq C\mu^{t-k}\|x(0)\|_\infty.$$

where C and μ are as defined earlier. Applying a vector version of Azuma's inequality we obtain

$$\mathbb{P}\left(\|r(t) - P\bar{W}^t x(0)\|_\infty \geq \epsilon\right) \leq 2\text{exp}(-\epsilon^2\alpha(t)). \quad (14)$$

This result bounds the distribution of $r(t)$ about a deterministic evolution $P\bar{W}^t x(0)$. We summarize this result below.

Proposition 4.1: Consider the average consensus problem with the assumptions as stated in Section II. The state vector $x(t)$ converges almost surely to a point $\alpha\mathbf{1}$ on the consensus subspace. Then the distance $r(t)$ of the state from the consensus subspace has a mean $P\bar{W}^t x(0)$ and satisfies

$$\mathbb{P}\left(\left\|r(t) - P\bar{W}^t x(0)\right\|_\infty \geq \epsilon\right) \leq 2\exp\left(-\epsilon^2\beta(t)\right). \quad (15)$$

where $\beta(t)$ is as defined in (10).

The concentration inequalities in Propositions 3.3 and 4.1 can be used to probabilistically bound $x(t)$ in the state space using equation (5) for any given $x(0)$.

V. Numerical Results

We now present some simulation results for the distribution of α in a simple wireless network. To illustrate our results, we consider the simplest network possible with $N = 2$ nodes that run the average consensus algorithm in the presence of packet losses. In simulations, this presence of drops is modelled as link states being Bernoulli random variables with parameters $p_{01} = 0.2$ and $p_{10} = 0.8$ for links

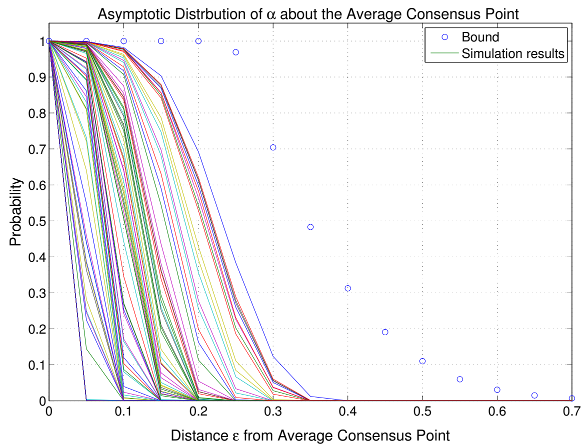

Fig. 2. Simulation results for $N = 2$ nodes running the average consensus algorithm over a fading channel. The plot here shows the distance from the consensus value from the average consensus value for a range of initial states $x(0)$. These results are compared with the theoretically obtained bound in Proposition 3.3.

$0 \to 1$ and $1 \to 0$ respectively. The packet losses are assumed to be independent of each other and across time. The initial values are chosen to be $x_0(0) = 0.4$ and $x_1(0)$ as uniformly random from -0.4 to 0.4. Note that our bound depends only on $\|x(0)\|_\infty = 0.4$ in this example.

To calculate the Monte Carlo estimate of the probability distribution, we run 100 sample runs each starting at a particular initial state chosen as described above, with 10000 simulations for each sample run. The asymptotic limit is approximated by running up to 50 iterations. For each initial state, we obtain an empirical probability distribution of the distance of the asymptotic value from the average consensus value. As expected, we get a family of probability distributions. This family is compared with the theoretical bound in Figure 2. We see that even for this simple network, the observed asymptotic consensus value has a broad support. We verify that the entire family of observed probability distributions is bounded by our result.

VI. Conclusions

In this paper, we have studied the probabilistic evolution of a network of agents implementing the average consensus algorithm. While it is known that these system achieve consensus, our contribution lies in the derivation of new concentration bounds on the state evolution at any time. These results are used to characterize the asymptotic distribution of the consensus point about the average consensus point. Our formulation allows the analysis of consensus algorithms over time-varying fading and interference-limited wireless networks with more realistic assumptions than before.

References

[1] J. N. Tsitsiklis, *Problems in Decentralized Decision Making and Computation*. PhD thesis, Department of EECS, MIT, Nov. 1984.

[2] V. Blondel, J. Hendrickx, A. Olshevsky, and J. Tsitsklis, “Convergence in multiagent co-ordination, consensus and flocking,” *Proc. IEEE Conf. Dec. and Control*, pp. 2996–3000, 2005.

[3] L. Fang and P. Antsaklis, “On communication requirements for multiagent consensus seeking,” *Proc. NESC05: Univ. of Notre Dame, Lecture Notes in Control and Information Sciences (LNCIS), Springer*, vol. 331, pp. 53–68, 2006.

[4] A. Jadbabaie, J. Lin, and A. S. Morse, “Co-ordination of groups of mobile autonomous agents using nearest neighbor rules,” *IEEE Trans. Auto. Control*, vol. 48(6), pp. 998–1001, June 2003.

[5] L. Moreau, “Consensus seeking in multiagent systems using dynamically changing interconnection topologies,” *IEEE Trans. Auto. Control*, vol. 50(2), 2005.

[6] R. Olfati-Saber and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” *IEEE Transactions on Automatic Control*, vol. 43, pp. 1520–33, 2004.

[7] D. B. Kingston and R. W. Beard, “Discrete-time average-consensus under switching network topologies,” *Proc. Am. Control Conf.*, pp. 3551–3556, 2006.

[8] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” *IEEE Trans. Info. Theory*, vol. 52(6), pp. 2508–2530, June 2006.

[9] S. Kar and J. M. F. Moura, “Sensor networks with random links: Topology design for distributed consensus,” *IEEE Trans. Sig. Proc.*, vol. 56 (7), pp. 3315–3326, July 2008.

[10] A. Tahbaz-Salehi and A. Jadbabaie, “A necessary and sufficient condition for consensus over random networks,” *IEEE Trans. Auto. Control.*, vol. 53 (3), pp. 791–795, April 2008.

[11] F. Fagnani and S. Zampieri, “Randomized consensus algorithms over large scale networks,” *IEEE J. Sel. Areas. Comm.*, vol. 26 (4), pp. 634–49, May 2008.

[12] M. Porfiri and D. J. Stilwell, “Consensus seeking over random weighted directed graphs,” *IEEE Trans. Automatic Control*, vol. 52, No. 9, pp. 1767–73, Sept. 2007.

[13] Y. Hatano and M. Mesbahi, “Agreement over random networks,” *IEEE Trans. Automatic Control*, vol. 50, No. 11, pp. 1867–72, Nov. 2005.

[14] R. Carli, F. Fagnani, A. Speranzon, and S. Zampeiri, “Communication constraints in the average consensus problem,” *Automatica*, vol. 44 (3), pp. 671–84, March 2008.

[15] A. Nedich and A. Ozdaglar, “Convergence rate for consensus with delays,” *LIDS Technical Report 2774, MIT Lab. for Inf. and Dec. Systems*.

[16] M. Huang and J. H. Manton, “Stochastic double array analysis and convergence of consensus algorithms with noisy measurements,” *Proc. Am. Control Conf.*, pp. 705–710, July 2007.

[17] P. Hovareshti, V. Gupta, and J. S. Baras, “Average consensus over small world networks: A probabilistic framework,” *IEEE Conf. Dec. and Control*, p. Accepted., 2008.

[18] S. Vanka, V. Gupta, and M. Haenggi, “Power-delay analysis of consensus algorithms on wireless networks with interference,” *Int. J. Sys. Control, and Comm., Spl. Issue on Information Processing and Decision Making in Distributed Control*, 2008 (To Appear).

[19] L. Xiao and S. Boyd, “Fast linear iterations for distributed averaging,” *Systems and Control Letters*, vol. 53, pp. 65–78, 2004.

[20] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*. Wiley-Interscience, 1999.

[21] D. Williams, *Probability with Martingales*. Cambridge University Press, 1991.

[22] M. Petkovsek, H. S. Wilf, and D. Zeilberger, *A=B*. A K Peters Ltd., 1996. Available online at http://www.math.upenn.edu/ wilf/AeqB.html.