Distributed Average Consensus in Sensor Networks with Random Link Failures and Communication Channel Noise

Soummya Kar and José M. F. Moura
Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA 15213 USA
Email: soummyak@andrew.cmu.edu, moura@ece.cmu.edu

***Abstract*— In this paper we study distributed average consensus type algorithms in sensor networks with random network link failures and communication channel. Specifically, the network links fail randomly across iterations, and communication through an active link incurs additive stochastic noise. We consider the $\mathcal{A}-\mathcal{ND}$ algorithm for distributed average consensus under such imperfect communication scenario. Using results from the theory of controlled Markov processes and stochastic approximation, we show that the $\mathcal{A}-\mathcal{ND}$ algorithm leads to consensus of the sensor states. In particular, all the sensor states converge a.s. to a finite random variable θ, the latter being an unbiased estimate of the desired average. We explicitly characterize the resulting the mean-squared error (m.s.e.) and show that the m.s.e. can be made arbitrarily small by tuning certain parameters of the algorithm. But, reducing the m.s.e. in this way, decrease the convergence rate of the algorithm, and we obtain an interesting trade-off between the m.s.e. and the convergence rate of the algorithm. Our results show that the sensor network topology plays a significant role in determining the convergence rate of these algorithms.**

***Index Terms*— Distributed Consensus, Random Link Failures, Communication Noise, Topology, Laplacian.**

I. INTRODUCTION

Distributed computation in sensor networks in the context of signal processing and control is a well-studied field with an extensive body of literature (see, for example, [1] for early work.) A problem that has received renewed interest recently is average consensus. It computes iteratively the global average of distributed data in a sensor network by using only local communications. In [2], a continuous time state update model was adopted for consensus and the results were extended to situations involving switching sensor network topology and delayed communication. In [3], the problem of designing the optimal link weights were addressed for a fixed sensor network topology, the optimality criterion being the convergence rate of the consensus algorithm.

The $\mathcal{A}-\mathcal{ND}$ algorithm addresses the case, where, simultaneously, the network links fail randomly (random topology change) and the transmitted data is corrupted by additive noise. This happens, for example, in an erasure network, where the transmissions are occasionally lost, and, in the case of a successful transmission, the data is distorted due to channel imperfections. We show that under the modelling assumptions considered in the paper, the $\mathcal{A}-\mathcal{ND}$ algorithm leads to almost sure (a.s.) convergence of the sensor state vector sequence to the *consensus subspace* if the expected network is connected, i.e., $\lambda_2(\mathbb{E}L) > 0$, where L is the random network Laplacian matrix. In other words, the sensor states reach consensus asymptotically, and converge a.s. to the same finite random variable. The $\mathcal{A}-\mathcal{ND}$ algorithm consists of distributed linear iterations, where each sensor updates its current state by a weighted fusion of its current neighbors' states (which are distorted when they reach it) and these fusion weights decrease to zero in an appropriate way, as time progresses. We show that the $\mathcal{A}-\mathcal{ND}$ algorithm falls under the purview of controlled Markov processes and the convergence analysis uses stochastic approximation techniques. We explicitly characterize the mean-squared error (m.s.e.) between the desired average and the resulting consensus value reached by the sensors and show that, by properly tuning the fusion weight sequence, the m.s.e. can be made arbitrarily small. However, reducing the m.s.e. in this way, decreases the convergence rate of the algorithm and we find an interesting trade-off between the m.s.e. and the convergence rate. In this context, we note that [4] also uses a decreasing sequence of weights for consensus in presence of additive noise, but considers a fixed network topology, while we allow the topology to vary randomly simultaneously. Also, our approach is much more general and applies to a wider range of situations, for example, data dependent noise etc., as will be detailed in the paper.

We comment briefly on the organization of the paper. In Sections II we summarize relevant spectral graph theoretic results, required for the development of the paper. The average consensus problem with additive noise and random link failures is treated in Section III, where we present the $\mathcal{A}-\mathcal{ND}$ algorithm and the underlying assumptions. We prove the a.s. convergence of the $\mathcal{A}-\mathcal{ND}$ algorithm in Section IV. The m.s.e. is explicitly characterized in Section V, while Section VI studies the trade-off between m.s.e. and the con-

This work was partially supported by the DARPA DSO Advanced Computing and Mathematics Program Integrated Sensing and Processing (ISP) Initiative under ARO grant # DAAD 19-02-1-0180, by NSF under grants # ECS-0225449 and # CNS-0428404, by the Office of Naval Research under MURI N000140710747, and by an IBM Faculty Award.

vergence rate. We suggest generalizations of our approach in Section VII. Finally, Section VIII concludes the paper.

II. Algebraic Graph Theory

In this section we summarize several spectral graph theoretic properties to be used in the paper. We model the sensor network as a graph, $G = (V, E)$, where V is the set of sensor nodes, with $|V| = N$. The edge set is the set of inter-sensor links (which may be random because of link failures) and is denoted by E. The connectivity structure imposed by the network can be represented by a symmetric $N \times N$ adjacency matrix, A, given by

$$A_{nl} = \begin{cases} 1 & \text{if } (n,l) \in E \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The neighborhood of sensor n, Ω_n, is defined as

$$\Omega_n = \{l \in V \mid (n,l) \in E\}, \quad n \in [1 \cdots N] \quad (2)$$

and its degree as $d_n = |\Omega_n|$. We define the graph Laplacian matrix, L, as

$$L = D - A \quad (3)$$

where, $D = \text{diag}\,(d_1 \cdots d_N)$ is the diagonal matrix of node degrees. By construction, the Laplacian L is symmetric positive semidefinite and we arrange its eigenvalues as

$$0 = \lambda_1(L) \le \lambda_2(L) \le \cdots \le \lambda_N(L) \quad (4)$$

The multiplicity of the zero eigenvalue is equal to the number of connected components of the network, and, in particular, for a connected graph, $\lambda_2(L) > 0$. The second eigenvalue, $\lambda_2(L)$ is referred to as the algebraic connectivity or the Fiedler value of the network. References [5], [6], [7] provide detailed treatment of graphs and their spectral theory.

III. Problem Formulation - Algorithm A-ND

In distributed average consensus, the sensor nodes start with some initial set of values, $\{x_n(0)\}_{1 \le n \le N}$ (which may be measurements of some real process), and iteratively compute their average. The iterations are distributed, so that, at each iteration, a sensor can access only the states of its neighbors for updating its current state. We define the vector of initial sensor states as

$$\mathbf{x}(0) = [x_1(0) \cdots x_N(0)]^T \in \mathbb{R}^{N \times 1} \quad (5)$$

and the corresponding average as

$$r = \frac{1}{N}\mathbf{1}^T\mathbf{x}(0) \quad (6)$$

where $\mathbf{1}$ is the vector of ones. In the case of perfect communication (static network, no channel noise), the sensors may update their states according to the following linear iterations:

$$x_n(i+1) = (1 - d_n\alpha)x_n(i) + \alpha \sum_{l \in \Omega_n} x_l(i) \quad (7)$$

If the network is connected and the weight α is chosen appropriately, the above sequence of iterations converges to the desired average r (see, for example, [3].) However, in the imperfect communication case, where the network links fail randomly and communication is corrupted by additive noise, the nodes have access only to a random subset of neighboring states and in the event of an active communication, the transmitted data is corrupted by additive noise. In a situation like this, we denote the status of a potential network link between sensors n and l at an iteration i, by an erasure random variable, $e_{nl}(i)$ (we assume $e_{nl}(i) = e_{ln}(i)$), where

$$e_{nl}(i) = \begin{cases} 1 & \text{if sensors } n \text{ and } l \text{ communicate at iteration } i \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Then, the data available to sensor n from sensor l at iteration i is given by

$$y_{nl}(i) = e_{nl}(i)(x_l(i) + v_{nl}(i)) \quad (9)$$

where, $v_{nl}(i)$ denotes the additive stochastic channel noise. The distributed average consensus algorithm needs to be modified to accommodate such imperfect communication scenarios and to this end, we propose the $\mathcal{A} - \mathcal{ND}$ algorithm for average consensus in the presence of random inter-sensor link failures and communication channel noise.

A-ND Algorithm: Recall from eqn. (9), the data available to sensor n from sensor l at iteration i is given by $y_{nl}(i)$. The $\mathcal{A} - \mathcal{NC}$ is given by the following set of iterations:

$$\begin{aligned} x_n(i+1) &= [1 - \alpha(i)d_n(i)]\,x_n(i) + \alpha(i)\sum_{l \neq n} e_{nl}(i)\,(x_l(i) \\ &\quad + v_{nl}(i))\,, \;\; 1 \le n \le N \end{aligned} \quad (10)$$

where, $\{\alpha(i)\}_{i \ge 0}$ is an appropriately chosen sequence (to be explained later) weights. The above set of iterations for the $\mathcal{A} - \mathcal{ND}$ algorithm can be written in a compact form as

$$\mathbf{x}(i+1) = \mathbf{x}(i) - \alpha(i)\,[L(i)\mathbf{x}(i) + \mathbf{n}(i)] \quad (11)$$

with initial state as $\mathbf{x}(0)$ and

$$n_l(i) = -\sum_{k \neq l} e_{lk}(i)v_{lk}(i), \quad 1 \le l \le N, \; i \ge 0 \quad (12)$$

The $\{L(i)\}_{i \ge 0}$ is a random sequence of Laplacian matrices, given by,

$$L_{nl}(i) = \begin{cases} \sum_{k \neq n} e_{nk}(i) & \text{if } n = l \\ -e_{nl}(i) & \text{otherwise} \end{cases} \quad (13)$$

Clearly, the convergence properties of the $\mathcal{A} - \mathcal{ND}$ algorithm is determined by the statistical properties of the random Laplacian sequence, $\{L(i)\}_{i \ge 0}$ (or, equivalently the erasure process, $\{e_{nl}(i)\}_{n \neq l,\, i \ge 0}$, and the additive noise process, $\{v_{nl}(i)\}_{n \neq l,\, i \ge 0}$. We show that, under fairly broad assumptions on these stochastic processes, the sensor states reach consensus asymptotically and converge a.s. to a finite random variable, arbitrarily close to the desired average r in a mean-squared sense. We state the assumptions as follows:

1) Random Network Failure: The graph Laplacians are

$$L(i) = \overline{L} + \widetilde{L}(i), \ \forall i \geq 0 \tag{14}$$

where $\{L(i)\}_{i \geq 0}$ is a sequence of independent identically distributed (i.i.d.) Laplacian matrices with mean $\overline{L} = \mathbb{E}[L(i)]$, such that $\lambda_2 \left(\overline{L}\right) > 0$. Note that, during the same iteration, we are not restricting the link failures to be independent, i.e., they may be correlated. In other words, for a fixed i, the erasure random variables, $\{e_{nl}(i)\}_{n \neq l}$, may be correlated to each other, but independent across iterations.

2) Independent Noise Sequence: The additive noise $\{v_{nl}(i)\}_{1 \leq n,l \leq N, \ i \geq 0}$ is an independent sequence

$$\mathbb{E}[v_{nl}(i)] = 0, \ \forall 1 \leq n, l \leq N, \ i \geq 0 \tag{15}$$

$$\sup_{n,l,i} \mathbb{E}\left[v_{nl}^2(i)\right] = \mu < \infty \tag{16}$$

From eqn. (12), it then follows that

$$\mathbb{E}[\mathbf{n}(i)] = \mathbf{0}, \ \forall i, \ \sup_i \mathbb{E}\left[\|\mathbf{n}(i)\|^2\right] = \eta \leq N(N-1)\mu < \infty \tag{17}$$

3) Persistence Condition:

$$\alpha(i) > 0, \ \sum_{i \geq 0} \alpha(i) = \infty, \ \sum_{i \geq 0} \alpha^2(i) < \infty \tag{18}$$

This condition, commonly assumed in the adaptive control and adaptive signal processing literature, assumes that the weights decay to zero, but not too fast. Examples of such sequences include

$$\alpha(i) = \frac{1}{i^\beta}, \ 0.5 < \beta \leq 1 \tag{19}$$

We define the consensus subspace, $\mathcal{C}$, as

$$\mathcal{C} = \left\{x \in \mathbb{R}^{N \times 1} \mid x = a\mathbf{1}, \ a \in \mathbb{R}\right\} \tag{20}$$

In Section IV, we show that, under the $\mathcal{A}-\mathcal{ND}$ algorithm, the state vector sequence, $\{\mathbf{x}(i)\}_{i \geq 0}$, converges a.s. to the consensus subspace $\mathcal{C}$. In other words, there exists a finite random variable, θ, such that

$$\mathbb{P}\left[\lim_{i \to \infty} \mathbf{x}(i) = \theta\mathbf{1}\right] = 1 \tag{21}$$

The m.s.e, ζ, is then given by

$$\zeta = \mathbb{E}[\theta - r]^2 \tag{22}$$

and this is explicitly characterized in Section V.

IV. Consensus - A.S. Convergence of A-ND Algorithm

We state a theorem on the convergence of Markov process sample paths, which will be used to prove the a.s. convergence of the $\mathcal{A}-\mathcal{ND}$ algorithm.

Theorem 1 Consider a Markov process, $\{\mathbf{x}(i)\}_{i \geq 0}$ on $\mathbb{R}^{N \times 1}$. Define the operator $\mathcal{L}$, which acts on non-negative functions $V(i, \mathbf{x}), \ i \geq 0, \ x \in \mathbb{R}^{N \times 1}$ by

$$\mathcal{L}V(i, \mathbf{x}) = \mathbb{E}[V(i+1, \mathbf{x}(i+1)) | \mathbf{x}(i) = \mathbf{x}] - V(i, \mathbf{x}) \ \ a.s. \tag{23}$$

Now suppose there exists a non-negative function $V(i, \mathbf{x}), \ i \geq 0, \ x \in \mathbb{R}^{N \times 1}$ and a set $B \subset \mathbb{R}^{N \times 1}$ with the following properties:

1)

$$\inf_{i \geq 0, x \in V_\epsilon(B)} V(i, \mathbf{x}) > 0, \ \forall \epsilon > 0 \tag{24}$$

$$V(i, \mathbf{x}) \equiv 0, \ \mathbf{x} \in B, \quad \lim_{\mathbf{x} \to B} \sup_{i \geq 0} V(i, \mathbf{x}) = 0 \tag{25}$$

where $V_\epsilon(B) = \{x \in \mathbb{R}^{N \times 1} \mid \inf_{y \in B} \rho(x, y) \geq \epsilon\}$.

2)

$$\mathcal{L}V(i, \mathbf{x}) \leq g(i)(1 + V(i, \mathbf{x})) - \alpha(i)\varphi(i, \mathbf{x}) \tag{26}$$

where $\varphi(i, \mathbf{x}), i \geq 0, \ \mathbf{x} \in \mathbb{R}^{N \times 1}$ is a non-negative function such that

$$\inf_{i, \mathbf{x} \in V_\epsilon(B)} \varphi(i, \mathbf{x}) > 0, \ \forall \epsilon > 0 \tag{27}$$

3)

$$\alpha(i), g(i) > 0, \quad \sum_{i \geq 0} \alpha(i) = \infty, \quad \sum_{i \geq 0} g(i) < \infty \tag{28}$$

Then, the Markov process $\{\mathbf{x}_i\}_{i \geq 0}$ with arbitrary initial distribution converges a.s. to B as $i \to \infty$.

Proof: The proof is detailed in [8] and builds on [9]. ■

Recall the consensus subspace, $\mathcal{C}$, given in eqn. (20). We now show that the under the $\mathcal{A}-\mathcal{ND}$ algorithm the sensor states reach consensus a.s., or, in other words, the sensor states approach the consensus subspace with probability one. We formalize this in the following theorem.

Theorem 2 Consider the $\mathcal{A}-\mathcal{ND}$ distributed average consensus algorithm given in Section III with arbitrary initial state $\mathbf{x}(0) \in \mathbb{R}^{N \times 1}$. Then,

$$\mathbb{P}\left[\lim_{i \to \infty} \rho(\mathbf{x}(i), \mathcal{C}) = 0\right] = 1 \tag{29}$$

where, $\rho(\cdot)$ is the standard Euclidean metric.

Proof: Clearly, the state vector sequence, $\{\mathbf{x}(i)\}_{i \geq 0}$, generated by the $\mathcal{A}-\mathcal{ND}$ algorithm is a Markov process, under the assumptions stated in Section III. We now use Theorem 1 to prove the result. To this end, we define the stochastic potential function,

$$V(i, \mathbf{x}) = \mathbf{x}^T \overline{L} \mathbf{x} \tag{30}$$

where, $\overline{L}$, is the mean Laplacian matrix, defined in eqn. (14). Then, by taking $B = \mathcal{C}$ in Theorem 1, it can be shown that all the assumptions in Theorem 1 are satisfied (see, [8]) and the theorem follows. ■

Theorem 2 shows that with probability one, the sensor states reach consensus asymptotically, i.e., they eventually merge to the consensus subspace, $\mathcal{C}$. In the following theorem, we strengthen this notion and show that the sensor states, not only approach $\mathcal{C}$, but in fact, converge a.s. to a finite random variable θ.

Theorem 3 Consider the $\mathcal{A}-\mathcal{ND}$ distributed average consensus algorithm given in Section III with arbitrary initial state $\mathbf{x}(0) \in \mathbb{R}^{N\times 1}$. Then, there exists an almost sure finite real random variable θ such that

$$\mathbb{P}\left[\lim_{i\to\infty} \mathbf{x}(i) = \theta\mathbf{1}\right] = 1 \quad (31)$$

Proof: We first note that the a.s. consensus implied by Theorem 2 is equivalent to the statement,

$$\mathbb{P}\left[\lim_{i\to\infty} \mathbf{x}(i) = x_{\text{avg}}(i)\mathbf{1}\right] = 1 \quad (32)$$

where, $\{x_{\text{avg}}(i)\}_{i\geq 0}$ is the sequence of instantaneous averages, given by

$$x_{\text{avg}}(i) = \frac{1}{N}\mathbf{1}^T\mathbf{x}(i) \quad (33)$$

We now show that, there exists a finite random variable θ, such that,

$$\mathbb{P}\left[\lim_{i\to\infty} x_{\text{avg}}(i) = \theta\right] = 1 \quad (34)$$

The average update is then given by the following recursion:

$$x_{\text{avg}}(i+1) = x_{\text{avg}}(i) - \alpha(i)\overline{\mathbf{n}}(i), \quad x_{\text{avg}}(0) = r \quad (35)$$

where

$$\overline{\mathbf{n}}(i) = \frac{1}{N}\mathbf{1}^T\mathbf{n}(i), \ \forall i \quad (36)$$

and

$$\mathbb{E}\left[\overline{\mathbf{n}}(i)\right] = 0, \quad \mathbb{E}\left[\overline{\mathbf{n}}^2(i)\right] \leq \frac{\eta}{N^2} \quad (37)$$

(This follows by multiplying both sides of eqn. (11) by $\frac{1}{N}\mathbf{1}^T$ and noting that, $\mathbf{1}^T L(i) = 0, \ \forall i$, from the properties of Laplacian matrices.) It can be shown (see, [8]) that the sequence, $\{x_{\text{avg}}(i)\}_{i\geq 0}$, is a martingale with respect to the filtration [1]

$$\mathcal{F}_i = \sigma\{\mathbf{x}(0), \{L(j)\}_{0\leq j<i}, \{\mathbf{n}(j)\}_{0\leq j<i}\} \quad (39)$$

We now have

$$\begin{aligned}\mathbb{E}\left[x^2_{\text{avg}}(i+1)\right] &= \mathbb{E}\left[x_{\text{avg}}(i) - \alpha(i)\overline{\mathbf{n}}(i)\right]^2 \\ &= \mathbb{E}\left[x^2_{\text{avg}}(i)\right] + \alpha^2(i)\mathbb{E}\left[\overline{\mathbf{n}}^2(i)\right] \\ &\leq \mathbb{E}\left[x^2_{\text{avg}}(i)\right] + \frac{\alpha^2(i)\eta}{N^2} \quad (40)\end{aligned}$$

where we have used the independence assumptions and eqn. (37). Continuing the recursion and using the fact that $\sum_{i\geq 0}\alpha^2(i) < \infty$, we have

$$\mathbb{E}\left[x^2_{\text{avg}}(i)\right] \leq r^2 + \frac{\eta}{N^2}\sum_{j\geq 0}\alpha^2(j) \quad (41)$$

Thus, the sequence $\{x_{\text{avg}}(i)\}_{i\geq 0}$ is an $\mathcal{L}_2$-bounded martingale, and converges a.s. to a finite random variable θ (see, [10].) This, together with eqn. (32) implies that

$$\mathbb{P}\left[\lim_{i\to\infty} \mathbf{x}(i) = \theta\mathbf{1}\right] = 1 \quad (42)$$

and proves the theorem. ∎

V. Mean-Squared Error

In Section IV we have shown that under the $\mathcal{A}-\mathcal{ND}$ algorithm, the sensor states reach consensus a.s. and converge to a finite random variable θ. Viewing θ as an estimate of the desired average r, we now investigate its statistical properties. In other words, we would expect θ to possess desirable properties, including unbiasedness and small mean-squared error (m.s.e.). To this end, we note that from eqn. (35) it follows

$$\mathbb{E}\left[x_{\text{avg}}(i)\right] = r, \ \forall i \geq 0 \quad (43)$$

Since, the sequence $\{x_{\text{avg}}(i)\}_{i\geq 0}$ converges to θ in $\mathcal{L}_2$, it converges also in $\mathcal{L}_1$, and we have

$$\begin{aligned}\mathbb{E}\left[\theta\right] &= \lim_{i\to\infty}\mathbb{E}\left[x_{\text{avg}}(i)\right] \quad (44)\\ &= r\end{aligned}$$

Thus, θ is an unbiased estimate of the desired average r. To compute the m.s.e. ζ (see eqn. (22)), we note that the sequence of non-negative functions $(x_{\text{avg}}(i) - r)^2$ converges a.s. to $(\theta - r)^2$. Hence, by Fatou's lemma,

$$\mathbb{E}\left[\theta - r\right]^2 \leq \liminf_{i\to\infty}\mathbb{E}\left[x_{\text{avg}}(i) - r\right]^2 \quad (45)$$

Using exactly similar manipulations, as used in the derivation of eqn. (41), it can be shown that

$$\mathbb{E}\left[x_{\text{avg}}(i) - r\right]^2 \leq \frac{\eta}{N^2}\sum_{j\geq 0}\alpha^2(j), \ \forall i \quad (46)$$

Combining eqns. (45,46) it follows that

$$\zeta \leq \frac{\eta}{N^2}\sum_{j\geq 0}\alpha^2(j) \quad (47)$$

which gives an explicit upper bound on the m.s.e. From eqn. (47), we note that, for a given η and N, the bound on the noise variance, ζ can be made arbitrarily small by properly scaling the weight sequence, $\{\alpha(j)\}_{j\geq 0}$. As an example, consider the weight sequence,

$$\alpha(j) = \frac{1}{j+1}, \ \forall j$$

Clearly, this choice of $\alpha(i)$ satisfies the persistence conditions of eqn. (**??**) and, in fact,

$$\sum_{j\geq 0}\alpha^2(j) = \sum_{j\geq 1}\frac{1}{j^2} = \frac{\pi^2}{6}$$

[1] A filtration, $\mathcal{F}$, is a non-decreasing sequence of sigma algebras. A stochastic process, $\{\mathbf{z}(i)\}_{i\geq 0}$, is $\mathcal{F}$ adapted, if $\mathbf{z}(i)$ is $\mathcal{F}_i$ measurable for each i. An integrable process, $\{\mathbf{z}(i)\}_{i\geq 0}$, which is adapted to a filtration $\mathcal{F}$, is a martingale if

$$\mathbb{E}\left[\mathbf{z}(i+1) \,|\, \mathcal{F}_i\right] = \mathbf{z}(i) \ \ a.s. \quad (38)$$

Then, for any $\epsilon > 0$, the scaled weight sequence, $\{\widetilde{\alpha}(j)\}_{j\geq 0}$,

$$\widetilde{\alpha}(j) = \frac{\sqrt{6\epsilon}N}{\sqrt{\eta}\pi(j+1)}$$

will guarantee that $\zeta \leq \epsilon$. However, reducing the m.s.e. by scaling the weights in this way will reduce the convergence rate of the algorithm; this trade-off is considered in Section VI.

VI. Convergence Rate vs M.S.E. Trade-off

In this section, we present an informal study of the rate at which the sensor states reach consensus, or the convergence rate of the state vector sequence, $\{\mathbf{x}(i)\}_{i\geq 0}$, to the random vector $\theta\mathbf{1}$. A detailed convergence analysis can be done by invoking the ODE method (see [11]), which we skip here. For preciseness and clarity, we present a simpler convergence rate analysis, involving the mean state vector sequence only. From the asymptotic unbiasedness of θ, it follows that

$$\lim_{i\to\infty} \mathbb{E}\left[\mathbf{x}(i)\right] = r\mathbf{1} \tag{48}$$

We now study the rate at which

$$\|\mathbb{E}\mathbf{x}(i) - r\mathbf{1}\| \longrightarrow 0 \tag{49}$$

Then, assuming that, $\alpha(i) \leq \frac{2}{\lambda_2(\overline{L})+\lambda_N(\overline{L})}, \ \forall i$ (this is eventually the case, as the $\alpha(i)$'s decay to zero), it can be shown that (see, [8])

$$\|\mathbb{E}\left[\mathbf{x}(i)\right] - r\mathbf{1}\| \leq \left(e^{-\lambda_2(\overline{L})\left(\sum_{0\leq j\leq i-1}\alpha(j)\right)}\right)\|\mathbb{E}\left[\mathbf{x}(0)\right] - r\mathbf{1}\| \tag{50}$$

Eqn. (50) shows that the rate of convergence depends on the topology through the algebraic connectivity $\lambda_2\left(\overline{L}\right)$ of the graph and through the weights $\alpha(i)$. Eqns. (50) and (47) show a tradeoff between the m.s.e. and the rate of convergence at which the sequence $\{\mathbb{E}\left[\mathbf{x}(i)\right]\}_{i\geq 0}$ converges to $r\mathbf{1}$. Eqn. (50) shows that this rate of convergence is closely related to the rate at which the weight sequence, $\alpha(i)$, sums to infinity. For a faster rate, we want the weights to sum up fast to infinity, i.e., the weights to be large. In contradistinction, eqn. (47) shows that, to achieve a small ζ, the weights should be small.

VII. Generalizations

We now comment on extensions of the $\mathcal{A}-\mathcal{ND}$ algorithm to accommodate more general additive noise processes and link failures. The algorithm can be immediately extended to the case of data-dependent noise (in other words, the distribution of $\mathbf{n}(i)$ depends on $\mathbf{x}(i)$), provided that

$$\mathbb{E}\left[\mathbf{n}(i) \mid \mathbf{x}(i)\right] = 0 \text{ a.s. } \forall i \tag{51}$$

Clearly, in this case the process $\{\mathbf{x}(i)\}_{i\geq 0}$ remains Markov, and the analysis goes through.

Regarding the link failure model, we assumed that the failures are independent across iterations, but may be correlated across different links at a particular iteration. This can be extended similarly to data dependent link failures, provided that

$$\mathbb{E}\left[\widetilde{L}(i) \mid \mathbf{x}(i)\right] = 0 \text{ a.s. } \forall i \tag{52}$$

More general cases of correlated noise or link failures (across iterations) may be handled by this approach, by possibly augmenting the *state*, so that the resulting process is a Markov process w.r.t. the new state. Also, in this case, the potential function $V(\cdot)$ needs to be modified accordingly. The approach developed in this paper applies to other cases of imperfect communication in sensor networks, see, for example, [12], where we develop a randomized algorithm for average consensus with quantized inter-sensor communication.

VIII. Conclusion

In this paper, we consider the distributed average consensus problem, when simultaneously inter-sensor communication links fail randomly and communication through an active link incurs additive noise. We show, that, if the mean network is connected, the $\mathcal{A}-\mathcal{ND}$ algorithm leads to a.s. consensus of the sensor states. We explicitly characterize the resulting m.s.e. and find an interesting trade-off between the m.s.e. and the convergence rate of the algorithm. In other words, the m.s.e can be made arbitrarily small, though at a cost of lower convergence rate. Finally, we note, that the approach used in this paper, may be applied to other problems of distributed computation in sensor networks with imperfect inter-sensor communication, for example, distributed load balancing in parallel processing, distributed network flow etc. These may provide avenues of further research and we would like to pursue these in the future.

References

[1] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D., Massachusetts Institute of Technology, Cambridge, MA, 1984.

[2] R. O. Saber and R. M. Murray, "Consensus protocols for networks of dynamic agents," in *42nd IEEE Conference on Decision and Control*, 2003.

[3] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Contr. Lett.*, vol. 53, pp. 65–78, 2004.

[4] Y. Hatano, A. K. Das, and M. Mesbahi, "Agreement in presence of noise: pseudogradients on random geometric networks," in *44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05*, Seville, Spain, December 2005.

[5] F. R. K. Chung, *Spectral Graph Theory*. Providence, RI : American Mathematical Society, 1997.

[6] B. Mohar, "The Laplacian spectrum of graphs," in *Graph theory, Combinatorics, and Applications*, Y. Alavi, G. Chartrand, O. R. Oellermann, and A. J. Schwenk, Eds. New York: J. Wiley & Sons, 1991, vol. 2, pp. 871–898.

[7] B. Bollobás, *Modern Graph Theory*. New York, NY: Springer Verlag, 1998.

[8] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks: Link failures and channel noise," *Submitted to the IEEE Transactions*.

[9] M. Nevel'son and R. Has'minskii, *Stochastic Approximation and Recursive Estimation*. Providence, Rhode Island: American Mathematical Society, 1973.

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

[11] H. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.

[12] S. Kar and J. M. F. Moura, "Distributed average consensus in sensor networks with quantized inter-sensor communication," Oct. 2007, submitted for publication.