FAST MEAN SQUARE CONVERGENCE OF CONSENSUS ALGORITHMS IN WSNs WITH RANDOM TOPOLOGIES

Silvana Silva Pereira, Alba Pagès-Zamora

Signal Processing and Communications group.
Technical University of Catalonia (UPC), Barcelona (SPAIN).
E-mail: silvana@gps.tsc.upc.edu, alba.pages@upc.edu

ABSTRACT

The average consensus in wireless sensor networks is achieved under assumptions of symmetric or balanced topology at every time instant. However, communication and/or node failures, as well as node mobility or changes in the environment make the topology vary in time, and instantaneous symmetry of the links is not guaranteed unless an acknowledgment protocol or an equivalent approach is implemented. In this paper, we evaluate the convergence in the mean square sense of a well-known consensus algorithm assuming a random topology and asymmetric communication links. A closed form expression for the mean square error of the state is derived as well as the optimum choice of parameters to guarantee fastest convergence of the mean square error.

Index Terms— Wireless sensor networks, random topologies, asymmetric links, mean average consensus, mean square convergence.

1. INTRODUCTION

Consensus algorithms are iterative algorithms where neighboring nodes interact with each other to reach an agreement regarding a certain value of interest. These algorithms are well suited for distributed estimation of parameters in wireless sensor networks (WSNs), as the nodes can make a decision without the necessity of conveying the information to a fusion center. We focus on the time-varying topology model of the average consensus algorithm by Olfati-Saber and Murray in [1]. Important contributions based on this model can be found on literature (see [2] and references therein). When the topology of the network is random however, the convergence of the algorithm should be studied in probabilistic terms. For instance, Hatano and Mesbahi use stochastic stability notions in [3] to study the convergence in probability of the consensus algorithm over random graphs. Kar and Moura in [4] relate mean square convergence of the consensus algorithm to the second smallest eigenvalue of the average Laplacian matrix. In both contributions, the topology is assumed symmetric at every time instant. For networks with non-symmetric random topologies, Tahbaz-Salehi and Jadbabaie in [5] use ergodicity properties to show a necessary and sufficient condition for almost sure convergence to a common value. Rabbat et al. in [6] show that the average consensus can be achieved with both symmetric and asymmetric links under certain parameter conditions, but at the cost of increasing the convergence time and thus, the overall energy consumption of the network.

In this paper we study the mean square convergence of the algorithm in [1] for WSNs with random time-varying topologies. The constraint on instantaneous link symmetry in [3, 4] and the constraints in [6] are relaxed, leading to a faster convergence of the algorithm. Since the physical parameter to be estimated is modelled as a random variable (r.v.), the convergence of the state value is evaluated with respect to its statistical mean. The novelty in the analysis is that, assuming equal probability of connection for all the links, a closed-form expression for the mean square error (MSE) of the state is derived, which allows us to determine the convergence conditions of the algorithm and the value of the design parameters that minimize the convergence time.

The paper is organized as follows. In Section 2 we introduce some basic definitions of graph theory and in Section 3 we present the problem statement. In Section 4 we study the MSE of the state vector and present our main result. In Section 5 we analyze the convergence conditions and the asymptotic MSE. Simulation results and conclusions are included in Section 6 and 7 respectively.

2. GRAPH THEORY CONCEPTS

The information flow among the nodes of a network can be described by a directed graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ where $\mathscr{V} = \{1, \cdots, N\}$ is the set of vertices (nodes) and $\mathscr{E}$ is the set of edges (links) e_{ij}, $\forall\ i, j = \{1, \cdots, N\}$, such that the information flows from node j to node i [7]. We assume $\mathscr{G}$ has no loops or multiple edges. The set of neighbors of node i is denoted $\mathscr{N}_i = \{j \in \mathscr{V} : e_{ij} \in \mathscr{E}\}$, and represents the set of

This work has been partially funded by the European Commission (033914-WINSOC), the Spanish Ministry of Education and Science and FEDER funds (TEC2006-06481 and CONSOLIDER CSD2008-00010 COMONSENS) and the Catalan Government (2005SGR-00639).

nodes sending information to node i. The adjacency matrix of $\mathscr{G}$, denoted $\mathbf{A} \in \mathbb{R}^{N\times N}$, has entries equal to

$$[\mathbf{A}]_{ij} = \begin{cases} 1 & \text{if } e_{ij} \in \mathscr{E} \quad \forall i,j = \{1,\cdots,N\} \\ 0 & \text{otherwise.} \end{cases}$$

The degree matrix $\mathbf{D} \in \mathbb{R}^{N\times N}$ is a diagonal matrix whose entries are the sum of the incoming edges for each node, i.e. $[\mathbf{D}]_{ii} = \sum_{j=1}^{N} [\mathbf{A}]_{ij}$. The Laplacian matrix of the graph is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, with corresponding eigenvalues denoted λ_i, $i = \{1,\cdots,N\}$, and smallest eigenvalue $\lambda_1 = 0$. If the graph is connected, λ_1 has algebraic multiplicity one and $\mathbf{L}$ is an irreducible matrix [8]. When the topology of the network varies randomly with time, the communication among the nodes can be described by a dynamic graph $\mathscr{G}(k) = \{\mathscr{V}, \mathscr{E}(k)\}$, where $\mathscr{E}(k)$ is the set of edges at time k and $\mathscr{V}$ is the constant set of nodes. In this paper, we assume the Erdős-Rényi random graph model [9], where the existence of a link between any pair of nodes of the network is probabilistic, i.e. $e_{ij} \in \mathscr{E}(k)$ with probability $0 < p \leq 1$. The resulting adjacency matrix at time k is therefore a random matrix with entries

$$[\mathbf{A}(k)]_{ij} = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

and mean[1] $\bar{\mathbf{A}} = \mathbf{P}$, where $\mathbf{P}$ is the probability matrix with entries $[\mathbf{P}]_{ij} = p \ \forall i \neq j$ and $[\mathbf{P}]_{ii} = 0$. The instantaneous Laplacian is also random and given by $\mathbf{L}(k) = \mathbf{D}(k) - \mathbf{A}(k)$ with mean $\bar{\mathbf{L}} = \bar{\mathbf{D}} - \mathbf{P}$, $\mathbf{D}(k)$ denoting the degree matrix at time k.

3. CONSENSUS IN RANDOM TOPOLOGIES

Consider a WSN composed of N nodes indexed with $i = \{1,\cdots,N\}$ and a scalar value $x_i(k)$ defined as the state of node i at time k. The state is initialized at each node at time $k = 0$ with the value of a single measurement and evolves in time according to the difference equation in [1]. Let $\mathbf{x}(k) \in \mathbb{R}^{N\times 1}$ denote the vector containing all the states of the network at time $k > 0$. Assuming a random time-varying topology, the evolution of $\mathbf{x}(k)$ can be written in matrix form as follows

$$\mathbf{x}(k) = \mathbf{W}(k-1)\mathbf{x}(k-1) \tag{1}$$

where the weight matrix is given by

$$\mathbf{W}(k) = \mathbf{I} - \epsilon\mathbf{L}(k), \tag{2}$$

$\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix, $\mathbf{L}(k)$ is the instantaneous random Laplacian defined in Section 2 and ϵ is a positive constant equal for all the iterations (the range of values of ϵ that guarantee the convergence of (1) will be determined in Section 5). The set of matrices $\{\mathbf{W}(k), \forall k\}$ in (2) are by construction independent of each other and row stochastic but not necessarily symmetric (since $\mathbf{L}(k)$ may have asymmetric links). Due to the Perron-Frobenius theorem [8], they have largest eigenvalue $|\lambda_1(\mathbf{W}(k))| = 1$ and associated right eigenvector an all-ones vector $\mathbf{1} \in \mathbb{R}^{N\times 1}$. However, the expected matrix of $\mathbf{W}(k)$ is symmetric and row-stochastic since $\bar{\mathbf{W}} = \mathbf{I} - \epsilon\bar{\mathbf{L}}$. Let $\mathbf{x}(0) = [x_1(0) \ x_2(0) \ \cdots \ x_N(0)]^T$ be the vector of measurements taken by the sensors, modeled as independent identically distributed (i.i.d.) r.v. with mean x_m and variance σ_0^2. The iterative algorithm in (1) can be rewritten as

$$\mathbf{x}(k) = \mathbf{M}_{\mathbf{w}}(k)\mathbf{x}(0), \tag{3}$$

where $\mathbf{M}_{\mathbf{w}}(k) = \prod_{l=1}^{k} \mathbf{W}(k-l)$. Due to the random nature of both $\mathbf{x}(0)$ and the matrix $\mathbf{M}_{\mathbf{w}}(k)$, we study the convergence of (3) analyzing the MSE of $\mathbf{x}(k)$ with respect to the mean average consensus given by $\mathbf{x}_m = \frac{1}{N}\mathbf{1}^T\mathbb{E}[\mathbf{x}(0)]\mathbf{1} = x_m\mathbf{1}$, where $\mathbb{E}[.]$ denotes expected value.

4. MEAN SQUARE ANALYSIS

The MSE of the state vector $\mathbf{x}(k)$ with respect to the mean average consensus averaged over N nodes is defined as

$$\mathsf{MSE}(x(k)) = \frac{1}{N}\mathbb{E}\left[\left\|\mathbf{x}(k) - \mathbf{x}_m\right\|_2^2\right]. \tag{4}$$

Replacing equation (3) in (4) and expanding the expression yields

$$\begin{aligned}\mathsf{MSE}(x(k)) = \frac{1}{N}\mathbb{E}\big[&\mathbf{x}^T(0)\mathbf{M}_{\mathbf{w}}^T(k)\mathbf{M}_{\mathbf{w}}(k)\mathbf{x}(0) \\ &-\mathbf{x}^T(0)\mathbf{M}_{\mathbf{w}}^T(k)\mathbf{x}_m - \mathbf{x}_m^T\mathbf{M}_{\mathbf{w}}(k)\mathbf{x}(0) - \mathbf{x}_m^T\mathbf{x}_m\big]\end{aligned}$$

The matrix $\mathbf{W}(k)$ and therefore $\mathbf{M}_{\mathbf{w}}(k)$, are assumed independent of $\mathbf{x}(0)$ $\forall k$. Considering that $\mathbb{E}[\mathbf{M}_{\mathbf{w}}(k)] = \bar{\mathbf{W}}^k$ is a symmetric row-stochastic matrix, the expression above can be rewritten as follows

$$\begin{aligned}\mathsf{MSE}(x(k)) = \ &\mathrm{tr}\left(\left(\sigma_0^2\mathbf{I} + \mathbf{x}_0\mathbf{x}_0^T\right)\mathbf{R}_{\mathbf{w}}(k)\right) - \mathbf{x}_0^T\bar{\mathbf{W}}^k\mathbf{x}_m \\ &-\mathbf{x}_m^T\bar{\mathbf{W}}^k\mathbf{x}_0 + \mathbf{x}_m^T\mathbf{x}_m\end{aligned} \tag{5}$$

where $\mathbf{x}_0 = \mathbb{E}[\mathbf{x}(0)]$ and

$$\mathbf{R}_{\mathbf{w}}(k) = \mathbb{E}\left[\mathbf{M}_{\mathbf{w}}^T(k)\mathbf{M}_{\mathbf{w}}(k)\right] \tag{6}$$

is a symmetric, nonnegative and double stochastic matrix $\forall k$. Since $\mathbb{E}[\mathbf{x}(0)] = x_m\mathbf{1}$ and considering further that $\mathbf{1}^T\mathbf{R}_{\mathbf{w}}(k) = \mathbf{1}^T$, equation (5) reduces to

$$\mathsf{MSE}(x(k)) = \frac{\sigma_0^2}{N}\mathrm{tr}\left(\mathbf{R}_{\mathbf{w}}(k)\right). \tag{7}$$

We present now our main result in the following theorem:

Theorem 1. *Consider the iterative algorithm in (1) with N nodes, probability of connection $0 < p \leq 1$ equal for all the links and i.i.d. initial values $\mathbf{x}(0)$ with mean x_m and variance*

[1] The bar denotes expected matrix.

σ_0^2. *The* MSE *of the state vector averaged over N nodes in (4) is equal to*

$$\mathsf{MSE}(x(k)) = \sigma_0^2 \left(\frac{b}{1-a+b} - \frac{(a-1)}{1-a+b}(a-b)^k \right) \quad (8)$$

with

$$\begin{aligned} a &= 1 - 2(N-1)p\epsilon + 2(N-1)p\epsilon^2 + (N-1)(N-2)p^2\epsilon^2 \\ b &= 2p\epsilon - Np^2\epsilon^2. \end{aligned} \quad (9)$$

Proof. Expanding the expression of $\mathbf{R_w}(k)$ in (6) and applying the linearity properties of the trace and the expected value operators[2], equation (7) can be rewritten as follows

$$\frac{\sigma_0^2}{N}\mathrm{tr}(\mathbf{R_w}(k)) = \frac{\sigma_0^2}{N}\mathrm{tr}(\mathbf{R_w}(k-1)\cdot\mathbf{C_w}) \quad (10)$$

where $\mathbf{C_w} = \mathbb{E}\left[\mathbf{W}(k)\mathbf{W}^T(k)\right]$, equal $\forall k$. After some matrix manipulations, $\mathbf{C_w}$ can be analytically expressed as

$$\mathbf{C_w} = \mathbf{I} - 2\epsilon\bar{\mathbf{L}} + \epsilon^2\mathbb{E}\left[\mathbf{L}(k)\mathbf{L}^T(k)\right]. \quad (11)$$

Assuming equal probability of connection $0 < p \leq 1$ for all the links we have that

$$[\bar{\mathbf{L}}]_{ij} = \begin{cases} (N-1)p & \text{for } i=j \\ -p & \text{for } i \neq j \end{cases} \quad (12)$$

and

$$\mathbb{E}\left[\mathbf{L}(k)\mathbf{L}^T(k)\right]_{ij} = \begin{cases} 2(N-1)p + (N-1)(N-2)p^2 & i=j \\ -Np^2 & i\neq j \end{cases} \quad (13)$$

Replacing (12) and (13) in (11), we obtain a matrix $\mathbf{C_w}$ of the form

$$\mathbf{C_w} = b\cdot\mathbf{1}\mathbf{1}^T + \mathbf{I}(a-b) \quad (14)$$

with a and b as in (9). Replacing equation (14) in (10) yields

$$\begin{aligned} \frac{\sigma_0^2}{N}\mathrm{tr}(\mathbf{R_w}(k)) &= \frac{\sigma_0^2}{N}\Big(\mathrm{tr}(\mathbf{R_w}(k-1)\mathbf{1}\mathbf{1}^T b + \mathbf{R_w}(k-1)(a-b))\Big) \\ &= \frac{\sigma_0^2}{N}(Nb + \mathrm{tr}(\mathbf{R_w}(k-1)(a-b))) \end{aligned}$$

where we have used the row-stochastic property of $\mathbf{R_w}(k)$. Substituting the trace above recursively we obtain

$$\begin{aligned} \frac{\sigma_0^2}{N}\mathrm{tr}(\mathbf{R_w}(k)) &= \frac{\sigma_0^2}{N}\cdot N\left(\sum_{l=0}^{k-2} b(a-b)^l + a(a-b)^{k-1}\right) \\ &= \sigma_0^2\left(\frac{b}{1-a+b} - \frac{(a-1)}{1-a+b}(a-b)^k\right) \end{aligned}$$

and the proof is completed. □

The MSE expression in (8) allows us to compute the mean square error of the state at every time instant offline, as it requires knowledge of general parameters only. Since a and b in (9) are rather cumbersome, in the following section we evaluate (8) analytically, to provide a better understanding of the MSE in terms of both convergence and asymptotic behaviour.

[2]The proof of (10) and the properties of $\mathbf{R_w}(k)$ are not included here because of lack of space.

5. FASTEST CONVERGENCE AND ASYMPTOTIC ANALYSIS OF THE MSE

In this section, we determine the dynamical range of ϵ that guarantees the convergence of (4) and the value of ϵ that gives fastest convergence of the MSE in (8). Then, we study the impact of N and p on the asymptotic MSE.

Recently, the authors in [4] relate the convergence in the mean square sense to the second smallest eigenvalue of the average Laplacian. On the other hand, the authors in [5] relate the almost sure convergence of the consensus algorithm to the second largest eigenvalue of the average weight matrix. In our case however, the convergence time of the MSE, and therefore the choice of ϵ, are related to the term $(a-b)$, since from equation (8) we observe that the MSE converges whenever $(a-b)^k \to 0$ as $k \to \infty$. For given N and p, consider the function

$$f(\epsilon) = a - b \quad (15)$$

with a and b as defined in (9). We observe that $(f(\epsilon))^k \to 0$ as $k\to\infty$ whenever $|f(\epsilon)| < 1$. It is not difficult to check that $f(\epsilon)$ is a quadratic function, nonnegative $\forall\epsilon$. The dynamical range of ϵ that guarantees $|f(\epsilon)| < 1$ corresponds to the interval $(0, 2\epsilon^*)$, where ϵ^* is the value that minimizes the function $f(\epsilon)$. We can now state the following corollary:

Corollary 1. *For a given number of nodes N and a given probability $0 < p \leq 1$, the value of ϵ that minimizes the function $f(\epsilon)$ in (15), with a and b defined in (9) is given by*

$$\epsilon^* = \frac{N}{2(N-1) + (N-1)^2 p + p}. \quad (16)$$

Summing up, if we choose ϵ inside the interval $(0, 2\epsilon^*)$ with ϵ^* as defined in (16), we can guarantee that as $k \to \infty$ the averaged MSE of the state vector in (4) will converge. Under this assumption, the limit of the MSE expression in (8) is

$$\lim_{k\to\infty} \mathsf{MSE}(x(k)) = \sigma_0^2\left(\frac{b}{1-a+b}\right).$$

Now, substituting for the values of a and b above yields

$$\begin{aligned} \lim_{k\to\infty} \mathsf{MSE}(x(k)) &= \frac{\sigma_0^2}{N}\left(\frac{2N - N^2 p\epsilon}{2N - (2(N-1) + (N-1)^2 p + p)\,\epsilon}\right) \\ &= \frac{\sigma_0^2}{N}\cdot g(\epsilon). \end{aligned} \quad (17)$$

Clearly, the function $g(\epsilon)$ approaches 1 as ϵ approaches 0, so the MSE at each node tends to σ_0^2/N as the value of ϵ approaches 0. Our contribution with respect to [6] is that (17) provides the deviation of the MSE with respect to the optimum σ_0^2/N when ϵ does not tend to 0. That is, equation (17) shows that whenever ϵ is larger than 0, the asymptotic MSE will be higher than σ_0^2/N by a factor equal to $g(\epsilon)$. Actually, it can be seen that $g(\epsilon)$ increases monotonically for $\epsilon \in (0, 2\epsilon^*)$ and tends to ∞ as ϵ approaches the upper limit

$2\epsilon^*$. In order to gain intuitive insight on the impact of N and p on the asymptotic MSE, we assume that ϵ is sufficiently small to approximate $g(\epsilon)$ using a first-order Taylor series expansion, such that $g(0)=1$ and $g'(0)=\frac{N-1}{N}(1-p)$. Note that in the proximity of $\epsilon=0$, the limit in (17) behaves as

$$\lim_{k\to\infty} \mathsf{MSE}(x(k)) \approx \frac{\sigma_0^2}{N}\left(1+\frac{N-1}{N}(1-p)\epsilon\right).$$

This result shows that the impact of the number of nodes to the asymptotic MSE becomes negligible for a relatively high number of nodes. On the contrary, the higher the probability of connection of the links, the closer the asymptotic MSE will be to the optimum σ_0^2/N.

6. SIMULATIONS

The analytical results obtained in the previous sections are supported here with computer simulations. We simulate a WSN composed of $N = 20$ nodes randomly deployed in a squared area, where each entry of the vector $\mathbf{x}(0)$ is modeled as an independent Gaussian r.v. with mean $x_m=20$ and variance $\sigma_0^2 = 5$. The probability of connection is set equal to $p=0.4$. A total of 100 thousand independent realizations were run to obtain the empirical MSE, where the position of the nodes and the connection probabilities were kept fixed for all the realizations while a new Laplacian matrix was generated in every iteration.

Figure 1 shows the empirical MSE computed with (4) (dotted lines) along with the theoretical MSE derived with (8) (patterns), plotted in dB as a function of the iteration index for different values of ϵ. The benchmark value of σ_0^2/N is included in solid line. As expected, the empirical values match the theoretical values found using equation (8). The optimum $\epsilon^* = 0.1094$ is computed with (16). In the first curve we let $\epsilon=0.01$ and in the second one $\epsilon=0.2$. The curve for $\epsilon=\epsilon^*$ is depicted with '*'. Clearly, choosing $\epsilon=\epsilon^*$ we obtain fastest convergence of the MSE, as less than ten iterations are required. We observe that the smallest ϵ gives the slowest convergence but the reached value is closest to the benchmark. In the cases of higher ϵ (0.1094 and 0.2), the gap corresponding to the term $g(\epsilon)$ of equation (17) can be clearly observed.

7. CONCLUDING REMARKS

We have shown that a closed form expression for the MSE of the state with respect to the mean average consensus in (4) can be found whenever the links have the same probability of connection. The convergence in the mean square sense of (1) is assured for appropriate values of ϵ, whose dynamical range and optimum value have been established. The deviation of the asymptotic MSE with respect to the optimum σ_0^2/N can be computed offline since it depends on the number of nodes, the ϵ parameter and the probability of link connection only.

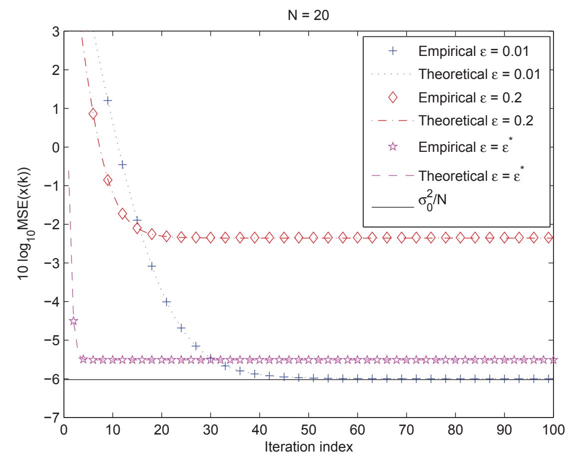

Fig. 1. Empirical and theoretical MSE of the state in dB as a function of the iterations for $p=0.4$ and different values of ϵ.

8. REFERENCES

[1] R. Olfati-Saber and R.M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1520–1533, Sept. 2004.

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

[3] Y. Hatano and M. Mesbahi, "Agreement over random networks," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, 14-17 Dec. 2004, vol. 2, pp. 2010–2015Vol.2.

[4] S. Kar and J.M.F. Moura, "Distributed average consensus in sensor networks with random link failures," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 15-20 April 2007, vol. 2, pp. II–1013–II–1016.

[5] A. Tahbaz-Salehi and A. Jadbabaie, "A necessary and sufficient condition for consensus over random networks," vol. 53, no. 3, pp. 791–795, 2008.

[6] M.G. Rabbat, R.D. Nowak, and J.A. Bucklew, "Generalized consensus computation in networked systems with erasure links," in *Signal Processing Advances in Wireless Communications, 2005 IEEE 6th Workshop on*, 5-8 June 2005, pp. 1088–1092.

[7] C. Godsil and G. Royle, *Algebraic graph theory*, vol. 207, Graduate Texts in Mathematics. Berlin, Germany: Springer-Verlag, 2001.

[8] R.A. Horn and C.R. Johnson, *Matrix analysis*, Cambridge University Press, 2006.

[9] P. Erdös and A. Rényi, "On random graphs I," *Publ. Math. Debrecen*, vol. 6, pp. 290–297, 1959.