THE EFFECT OF ADDITIVE NOISE ON CONSENSUS ACHIEVEMENT IN WIRELESS SENSOR NETWORKS

Antonio Fasano, Gesualdo Scutari

INFOCOM Dpt., University of Rome Sapienza, Via Eudossiana 18, 00184 Rome, Italy
E-mail: `<fasano,scutari>@infocom.uniroma1.it`

ABSTRACT

Achieving consensus on common global parameters through totally decentralized algorithms is a topic that has attracted considerable attention in the last few years, in view of its potential application in sensor networks. Several algorithms, along with their convergence properties, have been studied in the literature, among which the most popular are the (weighted) average consensus based schemes. One of the most critical aspects of these algorithms is that they suffer from catastrophic noise propagation. We show that the noise affecting the system state variables has a variance that grows linearly with the time index. In addition, we prove that encoding the information on the first forward difference of the state variables rather than on the state itself improves noise resilience, since it guarantees that the asymptotic value of the consensus is affected by noise with bounded variance. The results of our in-depth analysis of the effect of additive noise on consensus algorithms are valid *regardless* of the noise statistics and for *arbitrary* network topologies, i.e., arbitrary Laplacian matrices, and contain as special cases previously known results.

Index Terms— Sensor networks, Distributed algorithms, Distributed detection, Distributed estimation, Multisensor systems.

1. INTRODUCTION AND MOTIVATION

Endowing a sensor network with self-organizing capabilities is undoubtedly a useful goal to increase the resilience of the network against node failure (or simply the switch to sleep mode) and avoid potentially dangerous congestion conditions around the sink nodes. Decentralizing decisions decrease also the vulnerability of the network against damages to the sink or control nodes. Distributed computation over a network and its application to statistical consensus theory has a long history, starting with the pioneering work of Tsitsiklis, Bertsekas and Athans on asynchronous agreement problem for discrete-time distributed decision-making systems [4] and parallel computing [5]. A simple, yet significant, form of in-network distributed computing is achieving a consensus about one common observed phenomenon, without the presence of a fusion center. Distributed consensus algorithms have received great attention in the recent years in view of their potential application in sensor networks [6]. Excellent tutorials on distributed consensus techniques and their applications are given in [7, 8].

The main drawbacks of classical consensus protocols [6]−[8] are the lack of robustness against propagation delays and the high sensitivity to additive noise affecting the state variables of the nodes. More specifically, in the presence of propagation delays, these algorithms either may not converge [6] or converge to a final consensus value that depends on delays, network topology and initial conditions of each node [9]. In the presence of additive noise [10] authors show that, when considering average consensus problems on undirected connected graphs, the average of the state variables undergoes a random walk with linearly growing variance, thus preventing the node values to converge to the average of the initial values in any useful sense. These limitations make classical consensus algorithms [7, 8] not suitable to be applied in a real sensor network scenario, where in general both propagation delays and noise effects are not negligible. In [11, 12], authors propose a novel distributed consensus algorithm whose distinctive characteristic is the fact that consensus is achieved on the asymptotic value of the first derivative of the state variables, rather than on the state variables themselves, as classical consensus algorithms do [7, 8]. Moreover, this approach was shown to be robust against propagation delays [12].

In this work, we carry out an in-depth analysis of the effect of additive noise on distributed consensus algorithms, considering both classical ones [7, 8], and the one proposed in [11, 12]. In particular we focus on discrete-time systems and, as a consequence, we consider the first forward difference instead of the derivative, for the system in [11, 12]. Differently from the current works in the literature [10, 13, 14], we do not impose any constraint both on the noise statistics and the network topology, as long as this last is described in terms of a graph Laplacian. We will show that consensus algorithms on the state exhibit catastrophic noise propagation, whereas the corresponding algorithms on the first difference of the state exhibit noise resilience. The reader is assumed to be familiar with the graphs terminology, in particular with reference to consensus problems. Anyway, the relevant definitions can be found in [8] or [11].

2. SYSTEM MODEL

Consider a wireless sensor network composed of N nodes (sensors), where the interaction mechanisms among different sensors is described in terms of a directed graph (digraph) with Laplacian matrix $\boldsymbol{L}$ [11]. Now let use define the matrix

$$\boldsymbol{P} = \boldsymbol{I} - \epsilon \boldsymbol{L}\,, \tag{1}$$

with $\epsilon > 0$ such that $\boldsymbol{P}$ is stochastic with positive diagonal entries, i.e., $0 < \epsilon < [\max_k\{l_{kk}\}]^{-1}$, where l_{kk} is the k-th diagonal element of $\boldsymbol{L}$. Denoting by $\boldsymbol{x}$ the N-size vector gathering the state variables of all sensors, we consider systems where the dynamics of the overall network is governed by the following discrete-time dynamical model

$$\begin{cases} \boldsymbol{x}(n+1) = \boldsymbol{P}\boldsymbol{x}(n) + \boldsymbol{b} + \boldsymbol{w}(n)\,, \quad n \geq 0\,, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R} \end{cases} \tag{2}$$

where $\boldsymbol{b}$ denotes a possible constant input vector, $\boldsymbol{x}_0$ is the initial state and $\boldsymbol{w}(n)$ is the observation noise. The noise is a stochastic process with the following characteristics

$$\mathbb{E}\{\boldsymbol{w}(n)\} = \boldsymbol{0}\,, \qquad \mathbb{E}\left\{\boldsymbol{w}(n)\boldsymbol{w}(k)^T\right\} = \boldsymbol{K}_{\boldsymbol{w}}\delta_{n,k}\,, \tag{3}$$

where $\boldsymbol{K}_{\boldsymbol{w}}$ is assumed to be positive definite, and $\delta_{n,k}$ is the Kronecker delta. Note that we do not impose any constraint on the statistics of the noise. As a consequence $\boldsymbol{w}(n)$ in (2) can model, for example, the additive Gaussian noise of noisy links, or the quantization error in the possibly finite-precision representation of the state variables, or any other kind of random noise. Moreover the noise affecting different sensors can be correlated. It should be noted that (2) is a quite general model and subsumes, as special cases and when we omit the noise, several different systems proposed in the literature. In particular it can model both systems where consensus is achieved on the state $\boldsymbol{x}(n)$ (in this case $\boldsymbol{b} = \mathbf{0}$) [7, 8], and systems where consensus is achieved on the first difference of the state $\Delta\boldsymbol{x}(n)=\boldsymbol{x}(n+1) - \boldsymbol{x}(n)$ (in this case, in general $\boldsymbol{b} \neq \mathbf{0}$), which constitute the discrete version of the system proposed in [11, 12]. In both cases the evolution of (2) is governed by the behavior of $\boldsymbol{P}^n$ as n goes to infinity, and $\boldsymbol{P}$, see (1), depends on the network topology through $\boldsymbol{L}$. So, it will be useful to exploit the following general result, proven in [15].

Theorem 1. *Assume that $\boldsymbol{L}$ is the Laplacian matrix of an arbitrary directed*[1] *graph $\mathcal{G}$ and let $K \geq 1$ be the multiplicity of the zero eigenvalue of $\boldsymbol{L}$. Then, there always exists an ordering of the vertices of $\mathcal{G}$ such that $\boldsymbol{L}$ can be written as*

$$\boldsymbol{L} = \left[\begin{array}{c|c} \operatorname{diag}\{\boldsymbol{L}_1, \cdots, \boldsymbol{L}_K\} & \mathbf{0} \\ \hline \boldsymbol{R} & \boldsymbol{A} \end{array}\right], \tag{4}$$

where $\boldsymbol{L}_1, \cdots, \boldsymbol{L}_K$ are irreducible Laplacian matrices corresponding to strongly connected components of $\mathcal{G}$ and $\boldsymbol{A}$, if present, is nonsingular. Then, for the matrix $\boldsymbol{P}$ defined in (1) *the following asymptotic expression holds true*

$$\boldsymbol{P}^\infty = \lim_{n\to\infty} \boldsymbol{P}^n = \left[\begin{array}{c|c} \boldsymbol{\Gamma} & \mathbf{0} \\ \hline -\boldsymbol{A}^{-1}\boldsymbol{R}\boldsymbol{\Gamma} & \mathbf{0} \end{array}\right], \tag{5}$$

where $\boldsymbol{\Gamma} = \operatorname{diag}\left\{\mathbf{1}_1\boldsymbol{\gamma}_1^T, \cdots, \mathbf{1}_K\boldsymbol{\gamma}_K^T\right\}$, and $\boldsymbol{\gamma}_i$ is the left eigenvector of $\boldsymbol{L}_i$ corresponding to the zero eigenvalue, satisfying[2] $\boldsymbol{\gamma}_i^T\mathbf{1}_i = 1$.

It is worth noting that Theorem 1 provide us with a very general result. Indeed, it is able to fully predict the structure of $\boldsymbol{P}^\infty$, *whichever* is the network topology, and includes, as special cases, *all* the typical topologies encountered in the literature on consensus algorithms, like undirected graphs, quasi-strongly connected and strongly connected digraphs, and even not connected graphs. It is possible to prove [15], and we will see in Section 4, that the structure (4) of the Laplacian causes the node belonging to the strongly connected components corresponding to $\boldsymbol{L}_1, \ldots, \boldsymbol{L}_K$, to achieve consensus independently of each other, whereas the remaining nodes converge to values that depend on the matrices $\boldsymbol{A}$ and $\boldsymbol{R}$. When $K = 1$, i.e., $\mathcal{G}$ is quasi-strongly connected, $\boldsymbol{A}$ and $\boldsymbol{R}$ are such that all nodes achieve consensus [15]. In the sequel, without loss of generality, we assume that nodes are ordered to comply with (4). Moreover, we will refer to consensus in the broader sense of convergence as described above.

3. THE EFFECT OF ADDITIVE NOISE

In this section we analyze the effect of additive noise on consensus algorithms. As we will see, the evolution of (2) and in particular the impact of noise, strongly depend on the fact that we consider consensus on the state $\boldsymbol{x}(n)$ or on its first forward difference $\Delta\boldsymbol{x}(n)$. For the sake of precision, both $\boldsymbol{x}(n)$ and $\Delta\boldsymbol{x}(n)$ are stochastic processes so, we are legitimate to speak about consensus, assumed it can be achieved, only on the average. In the sequel we will consider separately the two cases, computing for each the expected value and the covariance matrix.

[1] We consider weighted loopless digraphs.

[2] $\mathbf{1}_i$ is the vector, of the proper size, with all entries equal to 1.

3.0.1. Consensus on $\boldsymbol{x}(n)$ ($\boldsymbol{b} = \mathbf{0}$)

The system is

$$\begin{cases} \boldsymbol{x}(n+1) = \boldsymbol{P}\boldsymbol{x}(n) + \boldsymbol{w}(n), & n \geq 0, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{cases} \tag{6}$$

with the following evolution

$$\boldsymbol{x}(n) = \boldsymbol{P}^n\boldsymbol{x}_0 + \sum_{k=0}^{n-1} \boldsymbol{P}^{n-1-k}\boldsymbol{w}(k). \tag{7}$$

Denoting by $\boldsymbol{m}_x(n) = \mathbb{E}\{\boldsymbol{x}(n)\}$ and taking the expectation of (7) we get

$$\boldsymbol{m}_{\boldsymbol{x}}(n) = \boldsymbol{P}^n\boldsymbol{x}_0, \tag{8}$$

so that the expected value evolves as the state of the original noiseless system, with corresponding asymptotic (vector) value

$$\lim_{n\to\infty} \boldsymbol{m}_{\boldsymbol{x}}(n) = \boldsymbol{P}^\infty\boldsymbol{x}_0, \tag{9}$$

where $\boldsymbol{P}^\infty$ is as in (5).

Now let us consider the covariance matrix of $\boldsymbol{x}(n)$. From (7) and (8) we get

$$\begin{aligned} \boldsymbol{K}_{\boldsymbol{x}}(n) &= \mathbb{E}\left\{[\boldsymbol{x}(n) - \boldsymbol{m}_{\boldsymbol{x}}(n)]\,[\boldsymbol{x}(n) - \boldsymbol{m}_{\boldsymbol{x}}(n)]^T\right\} = \\ &= \sum_{k=0}^{n-1} \boldsymbol{P}^k\boldsymbol{K}_{\boldsymbol{w}}(\boldsymbol{P}^k)^T. \end{aligned} \tag{10}$$

Using (10) we can compute the covariance matrix of $\boldsymbol{x}(n)$ for any value of the index n. However, we need to determine how variances evolve as n goes to infinity. The following theorem [15] gives such a characterization.

Theorem 2. *Denoting by $\sigma_1^2(n), \cdots, \sigma_N^2(n)$ the diagonal entries of the covariance matrix $\boldsymbol{K}_{\boldsymbol{x}}(n)$, the following inequalities hold true for $i = 1, \ldots, N$,*

$$\frac{\lambda_{\min}(\boldsymbol{K}_{\boldsymbol{w}})}{N} \cdot n \leq \sigma_i^2(n) \leq \lambda_{\max}(\boldsymbol{K}_{\boldsymbol{w}}) \cdot n \tag{11}$$

where $\lambda_{\min}(\boldsymbol{K}_{\boldsymbol{w}})$ and $\lambda_{\max}(\boldsymbol{K}_{\boldsymbol{w}})$ are the minimum[3] *and the maximum eigenvalues of $\boldsymbol{K}_{\boldsymbol{w}}$, respectively.*

The result in Theorem 2 is consistent with the observations made in [10] and greatly generalize them. In [10] authors consider the particular case of average consensus with an underlying graph undirected and connected, and they show that the average of the state variables, namely $\mathbf{1}^T\boldsymbol{x}(n)/N$, undergoes a random walk with a linearly growing variance. More generally, Theorem 2 states that the noise affecting *every* sensor has the variance that *grows linearly* with the time index n, and moreover this result holds *regardless* of the noise statistics and the network topology, as long as this last is described in terms of a graph Laplacian.

From Theorem 2 we conclude that system in (6) when we look at consensus on state variables is not robust with respect to the additive noise, since conditions (11) prevent the system to converge to (9) in any useful sense.

[3] Note that $\lambda_{\min}(\boldsymbol{K}_{\boldsymbol{w}}) > 0$ since $\boldsymbol{K}_{\boldsymbol{w}}$ is assumed positive definite.

3.0.2. Consensus on $\Delta\boldsymbol{x}(n)$ ($\boldsymbol{b} \neq \boldsymbol{0}$ in general)

The system is

$$\begin{cases} \boldsymbol{x}(n+1) = \boldsymbol{P}\boldsymbol{x}(n) + \boldsymbol{b} + \boldsymbol{w}(n)\,, \quad n \geq 0\,, \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{cases} \tag{12}$$

with the following evolution

$$\boldsymbol{x}(n) = \boldsymbol{P}^n\boldsymbol{x}_0 + \sum_{k=0}^{n-1}\boldsymbol{P}^k\boldsymbol{b} + \sum_{k=0}^{n-1}\boldsymbol{P}^{n-1-k}\boldsymbol{w}(k)\,. \tag{13}$$

In this case the presence of $\boldsymbol{b}$ makes the asymptotic state diverging, in general. Nonetheless, the first forward difference of the state vector $\Delta\boldsymbol{x}(n)$ turns out to be a stochastic process with bounded variance. In fact, from (13) we get

$$\begin{aligned} \Delta\boldsymbol{x}(n) = \boldsymbol{x}(n+1) - \boldsymbol{x}(n) = \boldsymbol{P}^n(\boldsymbol{P}-\boldsymbol{I})\boldsymbol{x}_0 + \boldsymbol{P}^n\boldsymbol{b} + \\ + \boldsymbol{w}(n) + (\boldsymbol{P}-\boldsymbol{I})\sum_{k=0}^{n-1}\boldsymbol{P}^{n-1-k}\boldsymbol{w}(k)\,. \end{aligned} \tag{14}$$

Now, let us evaluate the expected value and the covariance matrix of $\Delta\boldsymbol{x}(n)$. From (14) we have

$$\boldsymbol{m}_{\Delta\boldsymbol{x}}(n) = \mathbb{E}\left\{\Delta\boldsymbol{x}(n)\right\} = \boldsymbol{P}^n(\boldsymbol{P}-\boldsymbol{I})\boldsymbol{x}_0 + \boldsymbol{P}^n\boldsymbol{b}\,. \tag{15}$$

Thus, the corresponding asymptotic (vector) value is given by

$$\lim_{n\to\infty}\boldsymbol{m}_{\Delta\boldsymbol{x}}(n) = \boldsymbol{P}^\infty(\boldsymbol{P}-\boldsymbol{I})\boldsymbol{x}_0 + \boldsymbol{P}^\infty\boldsymbol{b} = \boldsymbol{P}^\infty\boldsymbol{b}\,, \tag{16}$$

since $\boldsymbol{P}^\infty\boldsymbol{P} = \boldsymbol{P}^\infty$. Note that (16) is the counterpart of (9) and expresses the consensus value as a function of the sole vector $\boldsymbol{b}$, regardless of the initial state $\boldsymbol{x}_0$. Now, let us consider the covariance matrix of $\Delta\boldsymbol{x}(n)$. From (14) and (15) we get

$$\boldsymbol{K}_{\Delta\boldsymbol{x}}(n) = \boldsymbol{K}_{\boldsymbol{w}} + \sum_{k=0}^{n-1}(\boldsymbol{P}-\boldsymbol{I})\boldsymbol{P}^k\boldsymbol{K}_{\boldsymbol{w}}(\boldsymbol{P}^T)^k(\boldsymbol{P}-\boldsymbol{I})^T\,, \tag{17}$$

which allows us to compute the covariance matrix of $\Delta\boldsymbol{x}(n)$ for any value of the index n. Actually, we are interested in evaluating the asymptotic behavior of $\boldsymbol{K}_{\Delta\boldsymbol{x}}(n)$ as the index n goes to infinity. However, in this case (17) becomes a series and convergence issues arise. The interesting result is that, as $n \to \infty$, (17) converges to a limit matrix. To demonstrate this result, we have to resort to the properties of the solution of the discrete-time Lyapunov equation [2] which are summarized in the following theorem

Theorem 3. *Consider the discrete-time Lyapunov equation in the indeterminate $\boldsymbol{X}$*

$$\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}^T + \boldsymbol{Q} = \boldsymbol{X} \tag{18}$$

where $\boldsymbol{A}$, $\boldsymbol{Q}$, $\boldsymbol{X} \in \mathbb{R}^{n\times n}$, and $\boldsymbol{Q}$ is symmetric positive semidefinite, and denote by $\lambda_i(\boldsymbol{A})$ the i-th eigenvalue of $\boldsymbol{A}$. If $|\lambda_i(\boldsymbol{A})| < 1$, $i = 1, \ldots, n$, then (18) has a unique solution $\boldsymbol{X}$, which is symmetric and positive semidefinite. Moreover $\boldsymbol{X}$ can be expressed in terms of the following convergent series

$$\boldsymbol{X} = \sum_{k=0}^{\infty}\boldsymbol{A}^k\boldsymbol{Q}(\boldsymbol{A}^T)^k\,. \tag{19}$$

Proof. See [2]. □

Exploiting the results of Theorem 3 we are ready to fully characterize the behavior of $\boldsymbol{K}_{\Delta\boldsymbol{x}}(n)$ when n goes to infinity, as the following theorem shows.

Theorem 4. *From the decomposition $\boldsymbol{P} = \boldsymbol{C}\boldsymbol{J}\boldsymbol{C}^{-1}$, where $\boldsymbol{J}$ is the Jordan canonical form of $\boldsymbol{P}$ and $\boldsymbol{C}$ is a similarity transformation, let us denote by $\tilde{\boldsymbol{P}} = \boldsymbol{C}\tilde{\boldsymbol{J}}\boldsymbol{C}^{-1}$, where $\tilde{\boldsymbol{J}}$ is the matrix obtained from $\boldsymbol{J}$ by zeroing the eigenvalues equal to one. Then, the covariance matrix $\boldsymbol{K}_{\Delta\boldsymbol{x}}(n)$ in (17) satisfies the following two equivalent asymptotic expressions*

$$\begin{aligned} \lim_{n\to\infty}\boldsymbol{K}_{\Delta\boldsymbol{x}}(n) &= \boldsymbol{K}_{\boldsymbol{w}} + \sum_{k=0}^{\infty}\Big[(\boldsymbol{P}-\boldsymbol{I})\boldsymbol{P}^k\boldsymbol{K}_{\boldsymbol{w}}(\boldsymbol{P}^T)^k(\boldsymbol{P}^T-\boldsymbol{I})\Big] \\ &= \boldsymbol{K}_{\boldsymbol{w}} + (\boldsymbol{P}-\boldsymbol{I})\boldsymbol{X}(\boldsymbol{P}^T-\boldsymbol{I})\,, \end{aligned} \tag{20}$$

where[4] *$\boldsymbol{X} = \sum_{k=0}^{\infty}\tilde{\boldsymbol{P}}^k\boldsymbol{K}_{\boldsymbol{w}}(\tilde{\boldsymbol{P}}^T)^k$ is the unique solution of the following discrete-time Lyapunov equation*

$$\tilde{\boldsymbol{P}}\boldsymbol{X}\tilde{\boldsymbol{P}}^T + \boldsymbol{K}_{\boldsymbol{w}} = \boldsymbol{X}\,. \tag{21}$$

Proof. First of all let us consider how to compute the matrix $\tilde{\boldsymbol{P}}$. Since $\boldsymbol{P}$ is a stochastic matrix, its spectrum is contained in the unit disc of the complex plane [1]. Moreover, an application of the Geršgorin disc theorem to the definition (1), leads to the conclusion that the only eigenvalue of $\boldsymbol{P}$ with unit modulus is 1. As a consequence, without loss of generality, it is always possible to write the Jordan canonical form of $\boldsymbol{P}$ as

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_{\boldsymbol{J}} \end{bmatrix} \tag{22}$$

where the identity matrix $\boldsymbol{I}_r$ collects the eigenvalues equal to one, and $\boldsymbol{D}_{\boldsymbol{J}}$ is a block diagonal matrix composed of Jordan blocks associated with the eigenvalues with modulus strictly less than 1. Denoting by $\boldsymbol{C}$ the similarity transformation leading to (22), matrix $\boldsymbol{P}$ can be represented as

$$\boldsymbol{P} = \boldsymbol{C}\begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_{\boldsymbol{J}} \end{bmatrix}\boldsymbol{C}^{-1}\,. \tag{23}$$

As a consequence $\tilde{\boldsymbol{P}}$ is obtained from (23) substituting $\boldsymbol{I}_r$ with $\boldsymbol{0}_r$

$$\tilde{\boldsymbol{P}} = \boldsymbol{C}\begin{bmatrix} \boldsymbol{0}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}_{\boldsymbol{J}} \end{bmatrix}\boldsymbol{C}^{-1}\,. \tag{24}$$

It is worth noting that due to (24) the eigenvalues of $\tilde{\boldsymbol{P}}$ have modulus strictly less than 1.

Now, consider (17) and in particular the term $(\boldsymbol{P}-\boldsymbol{I})\boldsymbol{P}^k$ within the summation. Exploiting (23) we get

$$\begin{aligned} (\boldsymbol{P}-\boldsymbol{I})\boldsymbol{P}^k &= \boldsymbol{C}\begin{bmatrix} \boldsymbol{0}_r & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{D}_{\boldsymbol{J}}-\boldsymbol{I}) \end{bmatrix}\begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{D}_{\boldsymbol{J}})^k \end{bmatrix}\boldsymbol{C}^{-1} \\ &= \boldsymbol{C}\begin{bmatrix} \boldsymbol{0}_r & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{D}_{\boldsymbol{J}}-\boldsymbol{I}) \end{bmatrix}\begin{bmatrix} \boldsymbol{0}_r & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{D}_{\boldsymbol{J}})^k \end{bmatrix}\boldsymbol{C}^{-1} \\ &= (\boldsymbol{P}-\boldsymbol{I})\tilde{\boldsymbol{P}}^k\,. \end{aligned} \tag{25}$$

Substituting (25) in (17) we have

$$\boldsymbol{K}_{\Delta\boldsymbol{x}}(n) = \boldsymbol{K}_{\boldsymbol{w}} + (\boldsymbol{P}-\boldsymbol{I})\left[\sum_{k=0}^{n-1}\tilde{\boldsymbol{P}}^k\boldsymbol{K}_{\boldsymbol{w}}(\tilde{\boldsymbol{P}}^T)^k\right](\boldsymbol{P}-\boldsymbol{I})^T\,. \tag{26}$$

[4] Note that $\tilde{\boldsymbol{P}}$ has eigenvalues with modulus strictly less than 1.

Now, taking the limit of (26) for $n \to \infty$ and considering that the sum within square brackets becomes a convergent series, since $\tilde{\boldsymbol{P}}$ satisfies the hypothesis of Theorem 4, we get

$$\lim_{n\to\infty} \boldsymbol{K}_{\Delta \boldsymbol{x}}(n) = \boldsymbol{K}_{\boldsymbol{w}} + (\boldsymbol{P}-\boldsymbol{I})\boldsymbol{X}(\boldsymbol{P}-\boldsymbol{I})^T\,, \tag{27}$$

with

$$\boldsymbol{X} = \sum_{k=0}^{\infty} \tilde{\boldsymbol{P}}^k \boldsymbol{K}_{\boldsymbol{w}} (\tilde{\boldsymbol{P}}^T)^k\,. \tag{28}$$

Moreover, due to (28) matrix $\tilde{\boldsymbol{P}}$ can be thought as the unique solution of the following discrete-time Lyapunov equation

$$\tilde{\boldsymbol{P}}\boldsymbol{X}\tilde{\boldsymbol{P}}^T + \boldsymbol{K}_{\boldsymbol{w}} = \boldsymbol{X}\,, \tag{29}$$

and the theorem is proved. □

Theorem 4 is a strong result, since it proves that systems like (2) when considering the evolution of the first forward difference are robust with respect to additive noise, and this property holds true *regardless* of the noise statistics and the network topology, as long as this last is described in terms of a graph Laplacian. The resilience against coupling noise was already observed in [11, 12] and, in this regard, the analysis carried out in [13, 14] when referencing [12] holds true in the special case of undirected connected graphs.

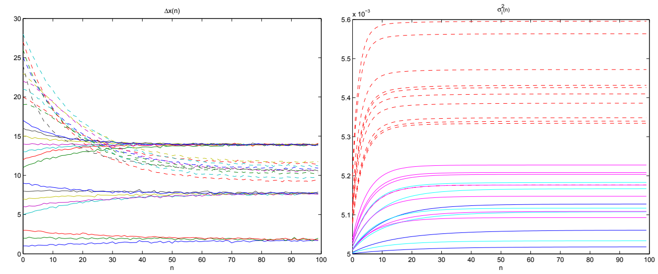
Furthermore, Theorem 4 allows the computation of the asymptotic covariance matrix by means of the solution of the discrete-time Lyapunov equation (21), and this last can be solved by efficient numerical algorithms [3]. From the computational point of view it is useful to remark that we do not need to compute the Jordan canonical form of $\boldsymbol{P}$ in order to determine $\tilde{\boldsymbol{P}}$. In fact, the following result can be proved [15]. Denote by $\boldsymbol{Z}_R$ the matrix whose columns constitute a basis for the eigenspace associated with the right eigenvectors of $\boldsymbol{P}$ corresponding to the eigenvalue 1, and denote by $\boldsymbol{Z}_L$ the corresponding matrix associated with the left eigevectors. Then, matrix $\tilde{\boldsymbol{P}}$ can be computed through the following formula

$$\tilde{\boldsymbol{P}} = \boldsymbol{P} - \boldsymbol{Z}_R(\boldsymbol{Z}_L^T\boldsymbol{Z}_R)^{-1}\boldsymbol{Z}_L^T\,. \tag{30}$$

Moreover, $\boldsymbol{Z}_R$ and $\boldsymbol{Z}_L$ can be easily determined since their columns constitute a basis[5] for $\mathcal{N}(\boldsymbol{P}-\boldsymbol{I})$ and $\mathcal{N}(\boldsymbol{P}^T-\boldsymbol{I})$ respectively.

4. SIMULATION RESULTS AND CONCLUSION

As an illustrative example, we have considered a network composed of 25 nodes characterized by a directed graph $\mathcal{G}$ (randomly generated) whose Laplacian has the zero eigenvalue with multiplicity 3. As a result, in the decomposition (4) there are three irreducible Laplacian matrices, namely $\boldsymbol{L}_1, \boldsymbol{L}_2, \boldsymbol{L}_3$, which correspond to three strongly connected components of $\mathcal{G}$. In this case, these last are composed of 3, 5, and 7 nodes respectively, as it is evident from Figure 1(a). We have considered the model (2) with $\boldsymbol{b} = \boldsymbol{x}_0 = [1\cdots 3\,, 5\cdots 9\,, 11\cdots 17\,, 19\cdots 28]^T$, affected by Gaussian noise having covariance matrix $\sigma_w^2\boldsymbol{I}$, with $\sigma_w^2 = 5\times 10^{-3}$. From Figure 1(a) we see that nodes belonging to the three strongly connected components (solid lines) achieve consensus independently, and each consensus value is a convex combination of the corresponding entries in $\boldsymbol{b}$, whereas the remaining nodes (dashed lines) converge to values that depend on matrices $\boldsymbol{A}$ and $\boldsymbol{R}$ in (4). From Figure 1(b), we have a confirmation of the results of our theoretical analysis. The variances of the noise affecting $\Delta\boldsymbol{x}(n)$ remain bounded as n increases,



(a) $\Delta\boldsymbol{x}(n)$ (b) Noise variances

Fig. 1. Sensor network composed of 25 nodes with a digraph whose Laplacian has the zero eigenvalue with multiplicity 3.

and the asymptotic values are the ones predicted by (20). The spread of the variance values, in this case, is of the order of 12% of σ_w^2 and interestingly the lower values pertain to the nodes belonging to the three strongly connected components (solid lines).

In this work, we carried out an in-depth analysis of the effect of additive noise on consensus algorithms. We proved that the noise affecting the system state variables has a variance that grows linearly with the time index, thus making consensus algorithms on the state variables catastrophically sensitive even to very low noise. Conversely, the corresponding algorithms on the first difference of the state exhibit a favorable resilience to noise, which makes them suitable for practical applications. Finally, it is worth noting that all our results hold true *regardless* of the noise statistics and for *arbitrary* network topology, as long as it is described in terms of a graph Laplacian.

5. REFERENCES

[1] A. Berman, R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1994.

[2] P. Lancaster, M. Tismenetsky, *The Theory of Matrices*, Academic Press, 2nd Ed., 1985.

[3] D. Kressner, "Block variants of Hammarling's method for solving Lyapunov equations", to appear in *ACM Trans. Math. Software*, 34(1), 2008.

[4] J. N. Tsitsiklis, D. P. Bertsekas, M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. on Automatic Control*, pp. 803-812, Sep. 1986.

[5] D. P Bertsekas and J.N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Athena Scientific, 1989.

[6] R. Olfati-Saber, and R. M. Murray, "Consensus Problems in Networks of Agents with Switching Topology and Time-Delays," *IEEE Trans. on Automatic Control*, vol. 49, pp. 1520-1533, Sep., 2004.

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

[8] W. Ren, R. W. Beard, and E. M. Atkins, "Information Consensus in Multivehicle Cooperative Control: Collective Group Behavior Through Local Interaction," *IEEE Control Systems Mag.*, vol. 27, no. 2, pp. 71-82, April 2007.

[9] V. Chellaboina, W. M. Haddad, Q. Hui, and J. Ramakrishnan, "On System State Equipartitioning and Semistability in Network Dynamical Systems with Arbitrary Time-Delays,"*Proc. of CDC 2006*, Dec. 13-15, 2006.

[10] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed Average Consensus with Least-Mean-Square Deviation," *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33-46, Jan. 2007.

[11] S. Barbarossa and G. Scutari, "Bio-inspired Sensor Network Design: Distributed Decision Through Self-synchronization," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 26-35, May 2007.

[12] G. Scutari, S. Barbarossa, and L. Pescosolido, "Distributed Decision Through Self-Synchronizing Sensor Networks in the Presence of Propagation Delays and Nonreciprocal Channels,"to appear in *IEEE Trans. on Signal Processing*. Available at `http://arxiv.org/abs/0709.2410.`

[13] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus Based Distributed Parameter Estimation in Ad Hoc Wireless Sensor Networks with Noisy Links,"*Proc. of ICASSP 2007*, April 15-20, 2007.

[14] I. D. Schizas, A. Ribeiro and G. B. Giannakis, "Consensus in Ad Hoc WSNs with Noisy Links - Part I: Distributed Estimation of Deterministic Signals," to appear in *IEEE Trans. on Signal Processing*, 2007.

[15] A. Fasano and G. Scutari, "The Effect of Additive Noise on Consensus Achievement in Wireless Sensor Networks," in preparation.

[5] $\mathcal{N}(\cdot)$ denotes the nullspace.