Research Article

Performance Analysis of Distributed Incremental LMS Algorithm with Noisy Links

Azam Khalili, Mohammad Ali Tinati, and Amir Rastegarnia

Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51664, Iran

Correspondence should be addressed to Amir Rastegarnia, a_rastegar@ieee.org

Recently proposed distributed adaptive estimation algorithms for wireless sensor networks (WSNs) do not consider errors due to noisy links, which occur during the transmission of local estimates between sensors. In this paper, we study the effect of noisy links on the performance of distributed incremental least-mean-square (DILMS) algorithm for the case of Gaussian regressors. More specifically, we derive theoretical relations which explain how steady-state performance of DILMS algorithm (in terms of mean-square deviation (MSD), excess mean-square error (EMSE), and mean-square error (MSE)) is affected by noisy links. In our analysis, we use a spatial-temporal energy conservation argument to evaluate the steady-state performance of the individual nodes across the entire network. Our simulation results show that there is a good match between simulations and derived theoretical expressions. However, the important result is that unlike the ideal links case, the steady-state MSD, EMSE and MSE are not monotonically increasing functions of step size parameter when links are noisy. In addition, the optimal step size is found in a closed form for a special case which minimizes the steady-state values of MSD, EMSE, and MSE in each node.

1. Introduction

In many wireless sensor network applications, multiple displaced sensors are used to estimate and track an unknown parameter, for example, average temperature, level of water contaminants, or a target position [1, 2]. In general, parameter estimation in WSN can be solved by either a centralized approach or a decentralized approach [3]. In a centralized approach, the spatially distributed sensor send their locally processed data to a fusion center to form the final estimate [4–6]. As the number of nodes increases, centralized processing becomes computationally prohibitive, since it would require communications over longer range which leads to reduced battery life. On the other hand, in decentralized estimation, spatially displaced sensors provide local estimates by collaborating with other nodes in the network [7–9].

In some WSN applications, the parameter estimation task must be done whereas the statistical model for the underlying processes of interest is not available, or it changes over time. This issue motivated the development of special class of decentralized approaches known as distributed adaptive estimation schemes [10–14]. In these schemes, every node is equipped with local computing ability to derive a local estimate and share it with its predefined neighbors. Using cooperative processing in conjunction with adaptive filtering per node enables the entire network and also each individual node to track not only the variations of the environment but also the topology of the network.

Depending on the manner by which the nodes communicate with each other, they may be referred to as incremental algorithms or diffusion algorithms. In the incremental mode, a cyclic path through the network is required, and nodes communicate with neighbors within this path [10–12]. The given algorithms in [10–12] use different adaptive filter in their structure, such as LMS, recursive least-squares (RLS), and affine projection. When more communication and energy resources are available, a diffusion cooperative scheme can be applied where nodes communicate with all of their neighbors, and no cyclic path is required. Both LMS-based and RLS-based diffusion algorithms have been considered in the literature [13, 14]. In [15], we have considered the quantization effects on the steady-state performance of DILMS algorithm.

1.1. Problem Description. Analysis of distributed estimation algorithms in the presence of noisy links is an important

physical layer issue which has been considered for different algorithms in the literature [16–21]. Nevertheless, this issue has not been considered in distributed adaptive estimation algorithms. Hence, in this paper, we study the performance of DILMS algorithm in a WNS with noisy links between sensors. The importance of such a study arises from the following facts:

(i) The WSN with noisy links is more realistic assumption than WSN with ideal links.

(ii) The performance of distributed adaptive estimation algorithm (e.g., DILMS) can vary significantly when they are used in noisy links WSN, which makes it vital to analyze their performance.

Our aim in this paper is to derive some theoretical expressions that describe the steady-state performance of DILMS algorithm with noisy links. It must be noted that analyzing the DILMS algorithm in a WSN with noisy links is challenging task, since an adaptive network comprises a system of systems that processes data cooperatively in both time and space in presence of noisy links. To be more specific, the contributions of this paper are listed below.

1.2. Contributions.

(1) We show that the performance of distributed adaptive estimation algorithms drastically decreases when links are noisy.

(2) We show that unlike the ideal link case [10], the steady-state MSD, EMSE, and MSE curves are not monotonically increasing functions of step size when links are noisy. This result is very important in real-world DILMS implementation and highlights the importance of our work.

(3) The optimal step size is found in a closed form for a special case which minimizes the steady-state values of MSD, EMSE, and MSE in each sensor.

1.3. Paper Organization. The remainder of this paper is organized as follows: in Section 2, we introduce the DILMS algorithm. In Section 3, we present our analysis and explain some results of the derived theoretical results in Section 4. Simulation results are given in Section 5, and Finally, the conclusions are drawn in Section 6.

1.4. Notation. Throughout the paper, we adopt boldface letters for random quantities and normal font for nonrandom (deterministic) quantities. The "∗" symbol is used for both complex conjugation for scalars and Hermitian transpose for matrices. The main symbols used in this paper are listed in Table 1.

2. The DILMS Algorithm

Suppose that a WSN is deployed to estimate an unknown vector w^o from measurements collected at N nodes in a network. Each node k has access to time realizations

TABLE 1: Symbols and their descriptions.

symbol	description
$(\cdot)^T$	transposition
$\| \cdot \|^2$	absolute squared operation
$\mathrm{Re}\{a\}$	real value of a
$E(a)$	statistical expectation of a
$\mathrm{col}\{a, b\}$	column vector with entries a and b
$\|x\|_{\Sigma}^2$	weighted norm for a column vector x, $x^*\Sigma x$
$\mathrm{Tr}[A]$	trace of matrix A
μ	step size parameter in DILMS algorithm
I_M	$M \times M$ identity matrix

$\{d_k(i), u_{k,i}\}$ of zero-mean spatial data $\{\mathbf{d}_k, \mathbf{u}_k\}$, where each $\mathbf{d}_k$ is a scalar measurement and each $\mathbf{u}_k$ is a $1 \times M$ row regression vector. We collect the regression and measurement data into global matrices as

$$\mathbf{U} \triangleq \mathrm{col}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N\} \ (N \times M),$$
$$\mathbf{d} \triangleq \mathrm{col}\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_N\} \ (N \times 1), \tag{1}$$

then the estimation problem is formulated as

$$\min_w J(w), \quad \text{where } J(w) = E\left(\|\mathbf{d} - \mathbf{U}w\|^2\right). \tag{2}$$

In the appendix, we introduce a motivating application where an estimation problem (2) arises. The optimal solution w^o is given by [10]

$$R_{du} = R_u w^o, \tag{3}$$

where $R_{du} = E(\mathbf{U}^*\mathbf{d})$ and $R_u = E(\mathbf{U}^*\mathbf{U})$. Note that the cost function (2) can be decomposed as [10]

$$J(w) = \sum_{k=1}^{N} J_k(w), \quad \text{where } J_k(w) = E\left(|\mathbf{d}_k - \mathbf{u}_k w|^2\right). \tag{4}$$

Using this property, in [10], a distributed incremental LMS strategy with a cyclic estimation structure is proposed, as follows:

$$\boldsymbol{\psi}_0^{(i)} \longleftarrow \mathbf{w}_{i-1},$$
$$\boldsymbol{\psi}_k^{(i)} = \underset{k=1,2,\dots,N}{\boldsymbol{\psi}_{k-1}^{(i)}} + \mu_k \mathbf{u}_{k,i}^* \left[\mathbf{d}_k(i) - \mathbf{u}_{k,i}\boldsymbol{\psi}_{k-1}^{(i)}\right], \tag{5}$$
$$\mathbf{w}_i \longleftarrow \boldsymbol{\psi}_N^{(i)},$$

where $\boldsymbol{\psi}_k^{(i)}$ indicates the local estimate at the node k and time i and $\mathbf{w}_i$ indicates the overall estimate at iteration i. For each time i, each node utilizes the local data $\mathbf{d}_k(i), \mathbf{u}_{k,i}$ and $\boldsymbol{\psi}_{k-1}^{(i)}$ received from the node $k-1$ to obtain $\boldsymbol{\psi}_k^{(i)}$. At the end of this cycle, $\boldsymbol{\psi}_N^{(i)}$ is employed as both the global estimate $\mathbf{w}_i$ and the initial condition for the next time instant. Note that to implement the DILMS, the time realizations $\{d_k(i), u_{k,i}\}$
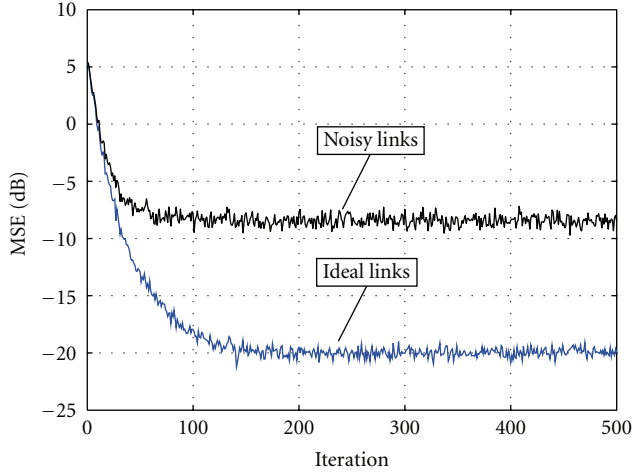
FIGURE 1: The effect of noisy links on the performance of DILMS algorithm.

are used. The update equation for DILMS in noisy link condition changes to

$$\boldsymbol{\psi}_k^{(i)} = \boldsymbol{\psi}_{k-1}^{(i)} + \mathbf{q}_{k,i} + \mu_k \mathbf{u}_{k,i}^* \left[\mathbf{d}_k(i) - \mathbf{u}_{k,i} \left(\boldsymbol{\psi}_{k-1}^{(i)} + \mathbf{q}_{k,i} \right) \right], \quad (6)$$

where the $M \times 1$ vector $\mathbf{q}_{k,i}$ is the channel noise term between sensor k and $k-1$ which is assumed to be additive with zero mean and covariance matrix Q_k. No distributional assumptions are required on the noise sequence. To show the effect of noisy links on the performance of DILMS algorithms, we consider a network with $N = 20$ nodes. The observation noise has a variance of $\sigma_{v,k}^2 = 10^{-2}$, and also we assume $Q_k = 10^{-3} I_M$ for channel noise. The curves are obtained by averaging over 300 experiments with $\mu = 0.001$ (see Figure 1). As it is clear from Figure 1, the performance of DILMS algorithms drastically decreases when links are noisy.

3. Performance Analysis

3.1. Data Model and Assumptions. In order to pursue the performance analysis, we will rely on the energy conservation approach of [10, 22]. Also, to carry out the performance analysis, we first need to assume a model for the data as is commonly done in the literature of adaptive algorithms. In the subsequent analysis, the following assumptions will be considered.

(A.1) The desired unknown vector w^o relates to the $\{\mathbf{d}_k(i), \mathbf{u}_{k,i}\}$ via

$$\mathbf{d}_k(i) = \mathbf{u}_{k,i} w^o + \mathbf{v}_k(i), \quad (7)$$

where $\mathbf{v}_k(i)$ is white noise term with variance $\sigma_{v,k}^2$ and is independent of $\{\mathbf{d}_l(j), \mathbf{u}_{l,j}\}$ for all l, j.

(A.2) $\mathbf{u}_{k,i}$ is independent of $\mathbf{u}_{l,i}$ for $k \neq l$.

(A.3) $\mathbf{u}_{k,i}$ is independent of $\mathbf{u}_{k,j}$ for $i \neq j$.

(A.4) $\mathbf{q}_{k,i}$ is independent of $\{\mathbf{u}_{l,j}, \mathbf{v}_l(j), \mathbf{q}_{l,j}\}$ for all l, j.

3.2. Weighted Energy Conservation Relation. In steady-state analysis, we are interested in evaluating the MSD, EMSE, and MSE for every node k which are defined as

$$\eta_k \triangleq E\left(\left\| \widetilde{\boldsymbol{\psi}}_{k-1}^{(\infty)} \right\|^2 \right) = E\left(\left\| \widetilde{\boldsymbol{\psi}}_{k-1}^{(\infty)} \right\|_I^2 \right) \quad \text{(MSD)},$$

$$\zeta_k \triangleq E\left(\left| \mathbf{e}_{a,k}(\infty) \right|^2 \right) = E\left(\left\| \widetilde{\boldsymbol{\psi}}_{k-1}^{(\infty)} \right\|_{R_{u,k}}^2 \right) \quad \text{(EMSE)}, \quad (8)$$

$$\xi_k \triangleq E\left(\left| \mathbf{e}_k(\infty) \right|^2 \right) = \zeta_k + \sigma_{v,k}^2 \quad \text{(MSE)},$$

where $\mathbf{e}_k(i) \triangleq \mathbf{d}_k(i) - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k-1}^{(i)}$, $\mathbf{e}_{a,k}(i) \triangleq \mathbf{u}_{k,i} \widetilde{\boldsymbol{\psi}}_{k-1}^{(i)}$, and $\widetilde{\boldsymbol{\psi}}_k^{(i)} \triangleq w^o - \boldsymbol{\psi}_k^{(i)}$. Note that due to incremental cooperation (cyclic path), for $k = 1$ we use $k - 1 = N$ in (8). We further define the weighted *a priori* and *a posteriori* local errors for each node k as follows:

$$\mathbf{e}_{a,k}^\Sigma(i) = \mathbf{u}_{k,i} \Sigma \widetilde{\boldsymbol{\psi}}_{k-1}^{(i)}, \qquad \mathbf{e}_{p,k}^\Sigma(i) = \mathbf{u}_{k,i} \Sigma \widetilde{\boldsymbol{\psi}}_k^{(i)}. \quad (9)$$

Using (6) and the definition of $\mathbf{e}_k(i)$, we have

$$\boldsymbol{\psi}_k^{(i)} = \boldsymbol{\psi}_{k-1}^{(i)} + \mu_k \mathbf{u}_{k,i}^* \mathbf{e}_k(i) + \mathbf{q}_{k,i} - \mu_k \mathbf{u}_{k,i}^* \mathbf{u}_{k,i} \mathbf{q}_{k,i}. \quad (10)$$

By subtracting w^o from both sides of (10), we get

$$\widetilde{\boldsymbol{\psi}}_k^{(i)} = \widetilde{\boldsymbol{\psi}}_{k-1}^{(i)} - \mu_k \mathbf{u}_{k,i}^* \mathbf{e}_k(i) - \mathbf{q}_{k,i} + \mu_k \mathbf{u}_{k,i}^* \mathbf{u}_{k,i} \mathbf{q}_{k,i}. \quad (11)$$

Multiplying the previous equation from left by $\mathbf{u}_{k,i} \Sigma$ and using the definitions in (9), we have

$$\mathbf{e}_{p,k}^\Sigma(i) = \mathbf{e}_{a,k}^\Sigma(i) - \mu_k \|\mathbf{u}_{k,i}\|_\Sigma^2 \mathbf{e}_k(i) - \mathbf{u}_{k,i} \Sigma \mathbf{q}_{k,i}$$
$$+ \mu_k \|\mathbf{u}_{k,i}\|_\Sigma^2 \mathbf{u}_{k,i} \mathbf{q}_{k,i}. \quad (12)$$

By replacing the $\mathbf{e}_k(i)$ from (11) and equating the weighted norm of both sides of the resultant equation, we arrive to the following relation:

$$\left\| \widetilde{\boldsymbol{\psi}}_k^{(i)} \right\|_\Sigma^2 + \frac{\left| \mathbf{e}_{a,k}^\Sigma(i) \right|^2}{\|\mathbf{u}_{k,i}\|_\Sigma^2}$$

$$= \left\| \widetilde{\boldsymbol{\psi}}_{k-1}^{(i)} \right\|_\Sigma^2 + \frac{\left| \mathbf{e}_{p,k}^\Sigma(i) \right|^2}{\|\mathbf{u}_{k,i}\|_\Sigma^2} + \|\mathbf{q}_{k,i}\|_\Sigma^2 - 2 \operatorname{Re}\left\{ \widetilde{\boldsymbol{\psi}}_{k-1}^{(i)*} \Sigma \mathbf{q}_{k,i} \right\}$$

$$+ 2 \operatorname{Re}\left\{ \frac{\widetilde{\boldsymbol{\psi}}_{k-1}^{(i)*} \Sigma \mathbf{u}_{k,i}^* \mathbf{u}_{k,i} \Sigma \mathbf{q}_{k,i}}{\|\mathbf{u}_{k,i}\|_\Sigma^2} \right\} - \frac{\mathbf{q}_{k,i}^* \Sigma \mathbf{u}_{k,i}^* \mathbf{u}_{k,i} \Sigma \mathbf{q}_{k,i}}{\|\mathbf{u}_{k,i}\|_\Sigma^2}. \quad (13)$$

We find from (13) that the cross terms are canceled out. Equation (13) is a space-time version of the weighted energy conservation relation used in [10] in the context of regular adaptive implementations.

3.3. Weighted Variance Relation. In this section, we use the energy conservation relation to evaluate the steady-state performance of the DILMS algorithm in every node when the links between nodes are noisy. To this aim, we need to have

a recursive equation for $\|\widetilde{\boldsymbol{\psi}}_k^{(i)}\|_\Sigma^2$. To obtain such a recursion we replace $\mathbf{e}_{p,k}^\Sigma(i)$ from (12) into (13) to get

$$
\begin{aligned}
\left\|\widetilde{\boldsymbol{\psi}}_k\right\|_\Sigma^2 = {} & \left\|\widetilde{\boldsymbol{\psi}}_{k-1}\right\|_\Sigma^2 - 2\,\mathrm{Re}\left\{\widetilde{\boldsymbol{\psi}}_{k-1}^* \Sigma \mathbf{q}_k\right\} + \|\mathbf{q}_k\|_\Sigma^2 \\
& + 2\mu_k\,\mathrm{Re}\left\{\mathbf{e}_{a,k}^{\Sigma *}\mathbf{u}_k\mathbf{q}_k\right\} - \mu_k\mathbf{q}_k^*\Sigma\mathbf{u}_k^*\mathbf{u}_k\mathbf{q}_k \\
& - \mu_k\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{u}_k\Sigma\mathbf{q}_k + \mu_k^2\|\mathbf{u}_k\|_\Sigma^2\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{u}_k\mathbf{q}_k \\
& - \mu_k\mathbf{e}_{a,k}^{\Sigma *}\mathbf{e}_k - \mu_k\mathbf{e}_k^*\mathbf{e}_{a,k}^\Sigma + \mu_k^2\|\mathbf{u}_k\|_\Sigma^2|\mathbf{e}_k|^2 \\
& + \mu_k\mathbf{e}_k^*\mathbf{u}_k\Sigma\mathbf{q}_k - \mu_k^2\|\mathbf{u}_k\|_\Sigma^2\mathbf{e}_k^*\mathbf{u}_k\mathbf{q}_k \\
& + \mu_k\mathbf{q}_k^*\Sigma\mathbf{u}_k^*\mathbf{e}_k - \mu_k^2\|\mathbf{u}_k\|_\Sigma^2\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{e}_k.
\end{aligned}
\tag{14}
$$

We have dropped the time index i for compactness of notation. Now, we can relate the $\mathbf{e}_k$ to $\mathbf{e}_{a,k}$ via

$$
\mathbf{e}_k = \mathbf{e}_{a,k} + \mathbf{v}_k.
\tag{15}
$$

If we replace (15) into (14), take expectation of both sides, and use assumptions (A.1)–(A.4), we obtain

$$
\begin{aligned}
E\left(\left\|\widetilde{\boldsymbol{\psi}}_k\right\|_\Sigma^2\right) = {} & E\left(\|\widetilde{\boldsymbol{\psi}}_{k-1}\|_\Sigma^2\right) + E\left(\|\mathbf{q}_k\|_\Sigma^2\right) \\
& - \mu_k E\left(\mathbf{e}_{a,k}^{\Sigma *}\mathbf{e}_{a,k}\right) - \mu_k E\left(\mathbf{e}_{a,k}^*\mathbf{e}_{a,k}^\Sigma\right) \\
& + \mu_k^2 E\left(\|\mathbf{u}_k\|_\Sigma^2\,|\mathbf{e}_{a,k}|^2\right) + \mu_k^2\sigma_{v,k}^2 E\left(\|\mathbf{u}_k\|_\Sigma^2\right) \\
& + \mu_k^2 E\left(\|\mathbf{u}_k\|_\Sigma^2\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{u}_k\mathbf{q}_k\right) \\
& - \mu_k E\left(\mathbf{q}_k^*\Sigma\mathbf{u}_k^*\mathbf{u}_k\mathbf{q}_k\right) - \mu_k E\left(\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{u}_k\Sigma\mathbf{q}_k\right).
\end{aligned}
\tag{16}
$$

By considering the error definitions $\mathbf{e}_{a,k}^\Sigma$ and $\mathbf{e}_{a,k}$, we can obtain the following relations:

$$
\begin{aligned}
E\left(\mathbf{e}_{a,k}^*\mathbf{e}_{a,k}^\Sigma\right) &= E\left(\left\|\widetilde{\boldsymbol{\psi}}_{k-1}\right\|_{\mathbf{u}_k^*\mathbf{u}_k\Sigma}^2\right), \\
E\left(\mathbf{e}_{a,k}^{\Sigma *}\mathbf{e}_{a,k}\right) &= E\left(\left\|\widetilde{\boldsymbol{\psi}}_{k-1}\right\|_{\Sigma\mathbf{u}_k^*\mathbf{u}_k}^2\right), \\
E\left(\|\mathbf{u}_k\|_\Sigma^2\,|\mathbf{e}_{a,k}|^2\right) &= E\left(\|\mathbf{u}_k\|_\Sigma^2\left\|\widetilde{\boldsymbol{\psi}}_{k-1}\right\|_{\mathbf{u}_k^*\mathbf{u}_k}^2\right).
\end{aligned}
\tag{17}
$$

Using the property $\|x\|_A^2 + \|x\|_B^2 = \|x\|_{A+B}^2$ and (17), we can expand (16) as

$$
\begin{aligned}
E\left(\left\|\widetilde{\boldsymbol{\psi}}_k\right\|_\Sigma^2\right) = {} & E\left(\left\|\widetilde{\boldsymbol{\psi}}_{k-1}\right\|_{\Sigma'}^2\right) + E\left(\|\mathbf{q}_k\|_\Sigma^2\right) + \mu_k^2\sigma_{v,k}^2 E\left(\|\mathbf{u}_k\|_\Sigma^2\right) \\
& + \mu_k^2 E\left(\|\mathbf{u}_k\|_\Sigma^2\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{u}_k\mathbf{q}_k\right) \\
& - \mu_k E\left(\mathbf{q}_k^*\Sigma\mathbf{u}_k^*\mathbf{u}_k\mathbf{q}_k\right) - \mu_k E\left(\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{u}_k\Sigma\mathbf{q}_k\right),
\end{aligned}
\tag{18}
$$

where in (18), we have

$$
\boldsymbol{\Sigma}' \triangleq \Sigma - \mu_k\Sigma\mathbf{u}_k^*\mathbf{u}_k - \mu_k\mathbf{u}_k^*\mathbf{u}_k\Sigma + \mu_k^2\|\mathbf{u}_k\|_\Sigma^2\mathbf{u}_k^*\mathbf{u}_k.
\tag{19}
$$

Invoking the independence of the regression data $\{\mathbf{u}_k\}$ allows us to write

$$
E\left(\left\|\widetilde{\boldsymbol{\psi}}_{k-1}\right\|_{\Sigma'}^2\right) = E\left(\left\|\widetilde{\boldsymbol{\psi}}_{k-1}\right\|_{E(\Sigma')}^2\right),
\tag{20}
$$

so that (18) becomes

$$
\begin{aligned}
E\left(\left\|\widetilde{\boldsymbol{\psi}}_k\right\|_\Sigma^2\right) = {} & E\left(\left\|\widetilde{\boldsymbol{\psi}}_{k-1}\right\|_{\Sigma'}^2\right) + E\left(\|\mathbf{q}_k\|_\Sigma^2\right) + \mu_k^2\sigma_{v,k}^2 E\left(\|\mathbf{u}_k\|_\Sigma^2\right) \\
& + \mu_k^2 E\left(\|\mathbf{u}_k\|_\Sigma^2\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{u}_k\mathbf{q}_k\right) \\
& - \mu_k E\left(\mathbf{q}_k^*\Sigma\mathbf{u}_k^*\mathbf{u}_k\mathbf{q}_k\right) - \mu_k E\left(\mathbf{q}_k^*\mathbf{u}_k^*\mathbf{u}_k\Sigma\mathbf{q}_k\right).
\end{aligned}
\tag{21}
$$

Note that in (21), Σ' is a deterministic matrix as

$$
\Sigma' \triangleq \Sigma - \mu_k E\left(\Sigma\mathbf{u}_k^*\mathbf{u}_k + \mathbf{u}_k^*\mathbf{u}_k\Sigma\right) + \mu_k^2 E\left(\|\mathbf{u}_k\|_\Sigma^2\mathbf{u}_k^*\mathbf{u}_k\right).
\tag{22}
$$

Recursion (21) is a variance relation that can be used to infer the steady-state performance at every node k. Note that Σ' is solely regressors dependent and, therefore, decoupled from the weight error vector. For simplicity, in this work, we consider the following assumption.

(A.5) The regressors $\{\mathbf{u}_k\}$ arise from a source with circular Gaussian distribution with covariance matrix $R_{u,k}$.

We introduce the eigndecomposition $R_{u,k} = U_k\Lambda_k U_k^*$, where Λ_k is a diagonal matrix with the eigenvalues of $R_{u,k}$ and U_k is unitary, that is, $U_k U_k^* = U_k^* U_k = I$. Define the transformed quantities

$$
\begin{aligned}
\overline{\boldsymbol{\psi}}_k &\triangleq U_k^*\widetilde{\boldsymbol{\psi}}_k, & \overline{\boldsymbol{\psi}}_{k-1} &\triangleq U_k^*\widetilde{\boldsymbol{\psi}}_{k-1}, & \overline{\mathbf{u}}_k &\triangleq \mathbf{u}_k U_k, \\
\overline{\mathbf{q}}_k &\triangleq U_k^*\mathbf{q}_k, & \overline{\Sigma} &\triangleq U_k^*\Sigma U_k, & \overline{\Sigma}' &\triangleq U_k^*\Sigma' U_k.
\end{aligned}
\tag{23}
$$

Using the above definitions, (21) and (22) can be rewritten in the equivalent forms

$$
\begin{aligned}
E\left(\left\|\overline{\boldsymbol{\psi}}_k\right\|_{\overline{\Sigma}}^2\right) &= E\left(\left\|\overline{\boldsymbol{\psi}}_{k-1}\right\|_{\overline{\Sigma}'}^2\right) + E\left(\|\overline{\mathbf{q}}_k\|_{\overline{\Sigma}}^2\right) \\
&\quad + \mu_k^2\sigma_{v,k}^2 E\left(\|\overline{\mathbf{u}}_k\|_{\overline{\Sigma}}^2\right) + \mu_k^2 E\left(\|\overline{\mathbf{u}}_k\|_{\overline{\Sigma}}^2\overline{\mathbf{q}}_k^*\overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k\overline{\mathbf{q}}_k\right) \\
&\quad - \mu_k E\left(\overline{\mathbf{q}}_k^*\overline{\Sigma}\overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k\overline{\mathbf{q}}_k\right) - \mu_k E\left(\overline{\mathbf{q}}_k^*\overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k\overline{\Sigma}\overline{\mathbf{q}}_k\right), \\
\overline{\Sigma}' &= \overline{\Sigma} - \mu_k E\left(\overline{\Sigma}\overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k + \overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k\overline{\Sigma}\right) + \mu_k^2 E\left(\|\overline{\mathbf{u}}_k\|_{\overline{\Sigma}}^2\overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k\right).
\end{aligned}
\tag{24}
$$

To proceed, we need to evaluate the moments in (24) as follows:

$$
\begin{aligned}
E\left(\|\overline{\mathbf{q}}_k\|_{\overline{\Sigma}}^2\right) &= \mathrm{Tr}\left[B_k\overline{\Sigma}\right], \\
E\left(\|\overline{\mathbf{u}}_k\|_{\overline{\Sigma}}^2\right) &= \mathrm{Tr}\left[\Lambda_k\overline{\Sigma}\right], \\
E\left(\|\overline{\mathbf{u}}_k\|_{\overline{\Sigma}}^2\overline{\mathbf{q}}_k^*\overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k\overline{\mathbf{q}}_k\right) &= \mathrm{Tr}\left[B_k\left(\Lambda_k\,\mathrm{Tr}\left[\overline{\Sigma}\Lambda_k\right] + \gamma\Lambda_k\overline{\Sigma}\Lambda_k\right)\right], \\
E\left(\overline{\mathbf{q}}_k^*\overline{\Sigma}\overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k\overline{\mathbf{q}}_k\right) &= \mathrm{Tr}\left[\Lambda_k B_k\overline{\Sigma}\right], \\
E\left(\|\overline{\mathbf{u}}_k\|_{\overline{\Sigma}}^2\overline{\mathbf{u}}_k^*\overline{\mathbf{u}}_k\right) &= \left(\Lambda_k\,\mathrm{Tr}\left[\overline{\Sigma}\Lambda_k\right] + \gamma\Lambda_k\overline{\Sigma}\Lambda_k\right),
\end{aligned}
\tag{25}
$$

where $B_k = U_k^* Q_k U_k$ and $\gamma = 1$ for circular complex data and $\gamma = 2$ for real data. Replacing these moments, (24) can be written as

$$E\left(\left\|\overline{\psi}_k\right\|_{\overline{\Sigma}}^2\right) = E\left(\left\|\overline{\psi}_{k-1}\right\|_{\overline{\Sigma}}^2\right) + \mathrm{Tr}\left[B_k\overline{\Sigma}\right] + \mu_k^2 \sigma_{v,k}^2 \mathrm{Tr}\left[\Lambda_k\overline{\Sigma}\right]$$
$$+ \mu_k^2 \mathrm{Tr}\left[B_k\left(\Lambda_k \mathrm{Tr}\left[\overline{\Sigma}\Lambda_k\right] + \gamma\Lambda_k\overline{\Sigma}\Lambda_k\right)\right]$$
$$- 2\mu_k \mathrm{Tr}\left[\Lambda_k B_k\overline{\Sigma}\right], \tag{26}$$

$$\overline{\Sigma}' = \overline{\Sigma} - \mu_k\left(\overline{\Sigma}\Lambda_k + \Lambda_k\overline{\Sigma}\right)$$
$$+ \mu_k^2\left(\Lambda_k \mathrm{Tr}\left[\overline{\Sigma}\Lambda_k\right] + \gamma\Lambda_k\overline{\Sigma}\Lambda_k\right). \tag{27}$$

Note from (27) that choosing $\overline{\Sigma}$ to be diagonal, $\overline{\Sigma}'$ will be diagonal as well, suggesting a more compact notation. Thus, we introduce the $M \times 1$ column vectors

$$\overline{\sigma} \triangleq \mathrm{diag}\{\overline{\Sigma}\}, \qquad \overline{\sigma}' \triangleq \mathrm{diag}\{\overline{\Sigma}'\}, \qquad \lambda_k \triangleq \mathrm{diag}\{\Lambda_k\}, \tag{28}$$

where the diag$\{\}$ notation will be used in two ways: $\lambda = \mathrm{diag}\{\Lambda\}$ is a column vector containing the main diagonal of Λ, and $\Lambda = \mathrm{diag}\{\lambda\}$ is a diagonal matrix whose entries are those of the vector λ. Therefore, using the diagonal notation, we obtain the following linear relation between the corresponding vectors $\{\overline{\sigma}, \overline{\sigma}'\}$:

$$\overline{\sigma}' = \overline{F}_k\overline{\sigma}, \tag{29}$$

where $\overline{F}_k$ is a $M \times M$ matrix that includes statistics of local data and given by

$$\overline{F}_k \triangleq I - 2\mu_k\Lambda_k + \gamma\mu_k^2\Lambda_k^2 + \mu_k^2\lambda_k\lambda_k^T. \tag{30}$$

As a result, (26) becomes

$$E\left(\left\|\overline{\psi}_k\right\|_{\mathrm{diag}\{\overline{\sigma}\}}^2\right) = E\left(\left\|\overline{\psi}_{k-1}\right\|_{\mathrm{diag}\{\overline{F}_k\overline{\sigma}\}}^2\right) + g_k\overline{\sigma}, \tag{31}$$

where g_k is a row vector as

$$g_k \triangleq \mu_k^2 \sigma_{v,k}^2 \lambda_k^T + \left(\mathrm{diag}\{B_k\}\right)^T\left[I + \mu_k^2\lambda_k\lambda_k^T + \gamma\mu_k^2\Lambda_k^2 - 2\mu_k\Lambda_k\right]. \tag{32}$$

Using (30), g_k can be rewritten in a more compact form as

$$g_k = \mu_k^2 \sigma_{v,k}^2 \lambda_k^T + \left(\mathrm{diag}\{B_k\}\right)^T\overline{F}_k. \tag{33}$$

For the sake of clarity, we reintroduce the time index i but drop the diag$\{\}$ notation from the subscripts in (31) for compactness. Expression (31) becomes

$$E\left(\left\|\overline{\psi}_k^{(i)}\right\|_{\overline{\sigma}_k}^2\right) = E\left(\left\|\overline{\psi}_{k-1}^{(i)}\right\|_{\overline{F}_k\overline{\sigma}_k}^2\right) + g_k\overline{\sigma}_k. \tag{34}$$

We replaced $\{\overline{\sigma}, \overline{\sigma}'\}$ by $\{\overline{\sigma}_k, \overline{\sigma}'_k\}$ in order to indicate that the weighting matrix can be node dependent.

3.4. Steady-State Behavior.
By comparing (34) with the similar equation for DILMS algorithm with ideal links (i.e., equation (55) in [10]), we can conclude that the desired steady-state MSD, EMSE, and MSE for DILMS algorithm with noisy links at node k can be expressed as

$$\eta_k = a_k(I - \Pi_{k,1})^{-1}r, \quad \text{(MSD)}, \tag{35}$$

$$\zeta_k = a_k(I - \Pi_{k,1})^{-1}\lambda_k, \quad \text{(EMSE)}, \tag{36}$$

$$\xi_k = \zeta_k + \sigma_{v,k}^2, \quad \text{(MSE)}, \tag{37}$$

where

$$\Pi_{k,l} \triangleq \overline{F}_{k+l-1}\overline{F}_{k+l}\cdots\overline{F}_N\overline{F}_1\cdots\overline{F}_{k-1}, \quad l = 1,\ldots,N, \tag{38}$$

$$a_k \triangleq g_k\Pi_{k,2} + g_{k+1}\Pi_{k,3} + \cdots + g_{k-2}\Pi_{k,N} + g_{k-1}, \tag{39}$$

$$r \triangleq \mathrm{diag}\{I\}, \qquad \lambda_k \triangleq \mathrm{diag}\{\Lambda_k\}. \tag{40}$$

Note that in (38), the subscripts are all mod$N$. It is evident that the effect of channel noise is addition of term $(\mathrm{diag}\{B_k\})^T\overline{F}_k$ to g_k in the case of ideal link.

4. Discussion on Derived Theoretical Results

An important result is that unlike the ideal links case, in the presence of noisy links, the MSD, EMSE, and MSE curves are not monotonically increasing functions of step size. To show this result more clearly and to make (35)–(37) analytically more tractable, we assume that

(A.6) $\mu_k = \mu$, $R_{u,k} = \lambda I$, $Q_k = \sigma_{c,k}^2 I$.

We further assume that μ is small enough so that $\overline{F}_k$ can be approximated as

$$\overline{F}_k \approx I - 2\mu\lambda I. \tag{41}$$

So, $\overline{F}_k$ is now a diagonal matrix and as a result, matrix $\Pi_{k,\ell} = \Pi = \overline{F}_1\overline{F}_2\cdots\overline{F}_N$ will be diagonal as well. For small μ, we have $(1 - 2\mu\lambda)^N \approx (1 - 2\mu N\lambda)$, then Π can be approximated as

$$\Pi = (I - 2\mu\lambda I)(I - 2\mu\lambda I)\cdots(I - 2\mu\lambda I)$$
$$\approx (1 - 2\mu N\lambda)I, \tag{42}$$

so that

$$I - \Pi \approx \sum_{k=1}^{N} 2\mu\lambda I = 2N\mu\lambda I. \tag{43}$$

Similarly, using the assumptions in (A.6), we have

$$g_k = \mu^2 \sigma_{v,k}^2 \lambda r^T + \sigma_{c,k}^2(1 - 2\mu\lambda)r^T, \tag{44}$$

so that a_k becomes

$$a_k \approx \sum_{k=1}^{N} g_k = \left(\sum_{k=1}^{N}\left(\mu^2 \sigma_{v,k}^2 \lambda + \sigma_{c,k}^2(1 - 2\mu\lambda)\right)\right)r^T. \tag{45}$$

Now, replacing (43) and (45) into (35) and using $r^T r = M$, we obtain

$$\eta_k = \frac{M}{2\mu\lambda N} \sum_{k=1}^{N} \left(\mu^2 \sigma_{v,k}^2 \lambda + \sigma_{c,k}^2 (1 - 2\mu\lambda) \right), \quad (46)$$

similarly, we can find the following approximations for EMSE and MSE as

$$\zeta_k = \frac{M}{2\mu N} \sum_{k=1}^{N} \left(\mu^2 \sigma_{v,k}^2 \lambda + \sigma_{c,k}^2 (1 - 2\mu\lambda) \right),$$
$$\xi_k = \zeta_k + \sigma_{v,k}^2. \quad (47)$$

We can easily conclude from (46)-(47) that the MSD, EMSE, and MSE curves are not monotonically increasing function of step size parameter. In Figure 2, we have shown the MSD as a function of μ when $N = 20$, $M = 5$, $\lambda = 1$, $\sigma_v^2 = 0.02$ and for different values of σ_c^2. As it is clear from Figure 2, for $\sigma_c^2 = 0$ (i.e., noiseless links), the MSD curve is a monotonically increasing function of μ.

Remark 1. To explain this behavior, we consider again the update (6). For small μ, the channel noise term say $q_{k,i}$ is dominant term in update equation, so as $\mu \to 0$ the steady-state performance deteriorates. As μ increases, the effect of channel noise term decreases, and finally as μ becomes larger, the steady-state performance deteriorates again like any adaptive algorithm.

Remark 2. The optimal step size for MSD is given by

$$\mu_o = \sqrt{\frac{\sum_{k=1}^{N} \sigma_{c,k}^2}{\sum_{k=1}^{N} \lambda \sigma_{v,k}^2}}. \quad (48)$$

It must be noted that (48) is also the optimal step size for EMSE and MSE curves.

Remark 3. Note that according to the given results in [10], as step size becomes smaller ($\mu \to 0$), steady-state values of MSD, EMSE, and MSE in each node should be more smaller too, but this is not the case in the presence of noisy links. In fact, μ must be chosen more carefully in real world.

5. Simulation

In this section, we provide computer simulations to compare the theoretical expressions with simulation results. To conduct our simulation results, we consider the following steps:

(1) consider a network with noisy link and generate the measurement and regression data,

(2) select a parameter w^o (which is known for us but unknown for DILMS algorithm),

(3) let the DILMS algorithm estimate w^o in WSN with noisy links and data (generated in step 1),
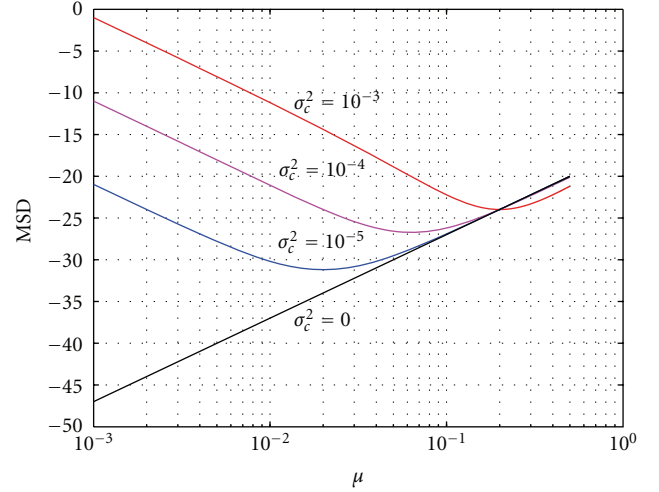
(4) obtain the MSD, EMSE, and MSE simulation results,

FIGURE 2: The steady-state MSD (in dB) curve as a function of μ and for different values of σ_c^2.

(5) we apply the data (generated in step 1) to our derived theoretical results,

(6) finally, we compare the resultant simulation results with derived theoretical results.

To this aim, we consider a distributed network with $N = 20$ nodes, and choose $M = 4$ and $w^o = [1 \quad 1 \quad \cdots \quad 1]^T / \sqrt{M}$.

5.1. Regressors with Shift Structure. Although the analysis relied on the independence assumptions, simulations presented in this subsection were carried out using regressors with shift structure to cope with realistic scenarios. The regressors are generated at each node k according to the following recursion:

$$u_k(i) = \alpha_k u_k(i-1) + \beta_k \tau_k(i). \quad (49)$$

The expression above describes a first-order autoregressive (AR) process with a pole at α_k, τ_k is a white, zero-mean, Gaussian random sequence with unity variance or a uniform random sequence with unity variance, $\alpha_k \in (0, 0.5]$ and $\beta_k = \sqrt{\sigma_{u,k}^2 (1 - \alpha_k^2)}$. In this way, the covariance matrix $R_{u,k}$ of the regressors $u_{k,i}$ is $M \times M$ Toeplitz matrix with entries $r_k(m) = \sigma_{u,k}^2 \alpha_k^{|m|}$, $m = 0, 1, 2, \ldots, M-1$ with $\sigma_{u,k}^2 \in [0, 0.5]$. The MSD, EMSE, and MSE are obtained by averaging the last 200 samples. Each curve is obtained by averaging over 100 independent experiments. The steady-state curves are generated by running the network learning process for 2000 iterations. We consider real data ($\gamma = 2$).

We assume that for node k, covariance matrix Q_k is a diagonal matrix and has different values at the diagonal so

$$Q_k = \begin{bmatrix} \delta_{k,1} & 0 & \cdots & 0 \\ 0 & \delta_{k,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_{k,M} \end{bmatrix}_{M \times M}. \quad (50)$$
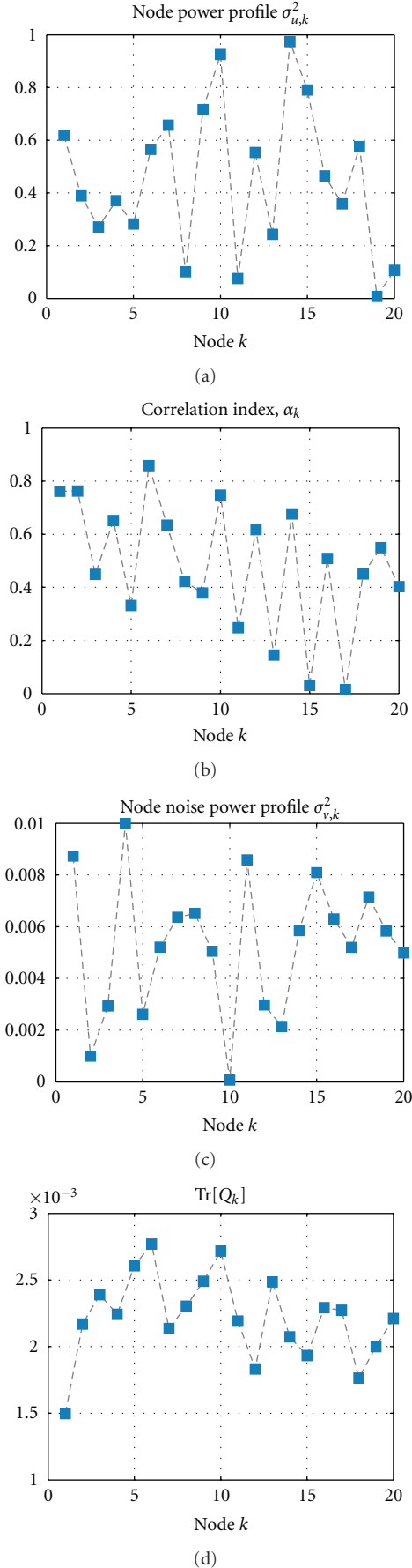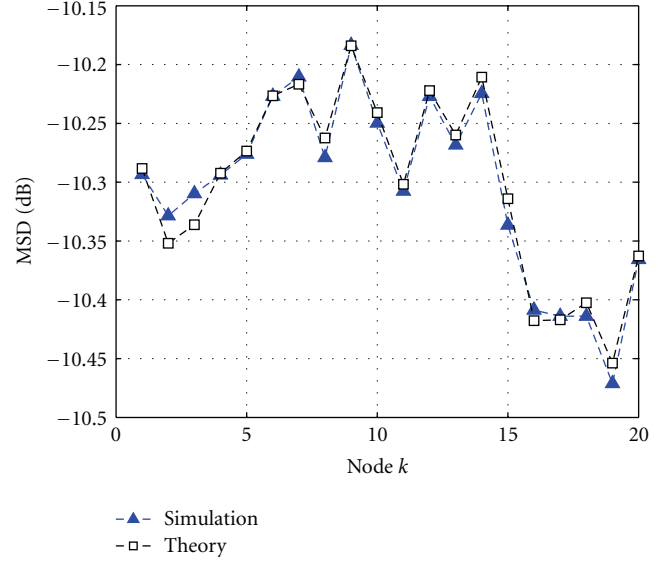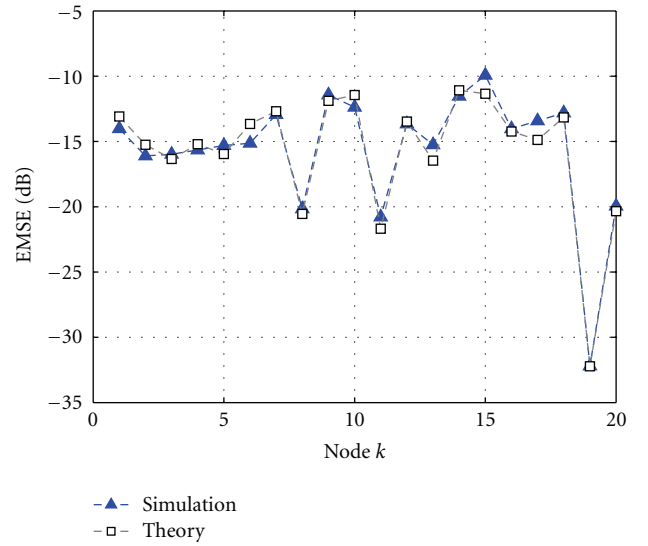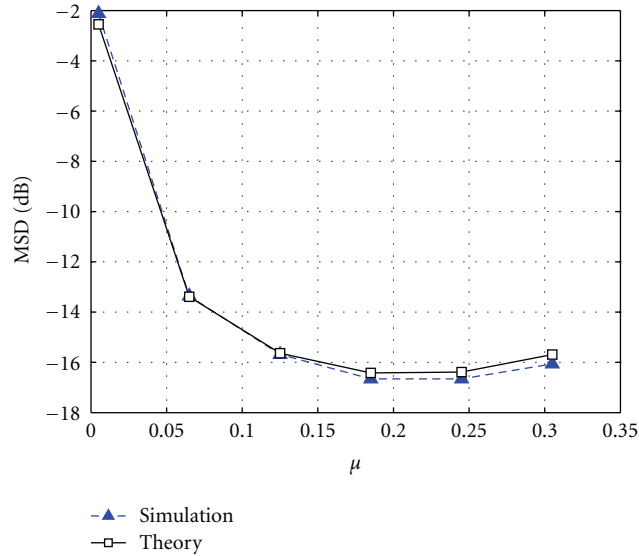
(a)

(b)

(c)

(d)

Figure 3: Node profile and channel noise information.

Figure 4: Steady-state MSD versus node, $\mu = 0.03$.

Figure 5: Steady-state EMSE versus node, $\mu = 0.03$.

In fact, at each node k, the noise term $q_{k,i}$ is generated to result the required (assumed) covariance matrix Q_k. The statistical profiles for the mentioned parameters are illustrated in Figure 3.

In Figures 4–6, the steady-state of MSD, EMSE, and MSE for $\mu = 0.03$ are plotted, respectively. It is clear from Figures 4–6 that there is a good match between simulation and theory. Note also that despite the diverse statistical profile, the MSD in Figure 4 is roughly even over the network. On the other hand, the EMSE and the MSE are more sensitive to local statistics, as depicted in Figures 5 and 6.
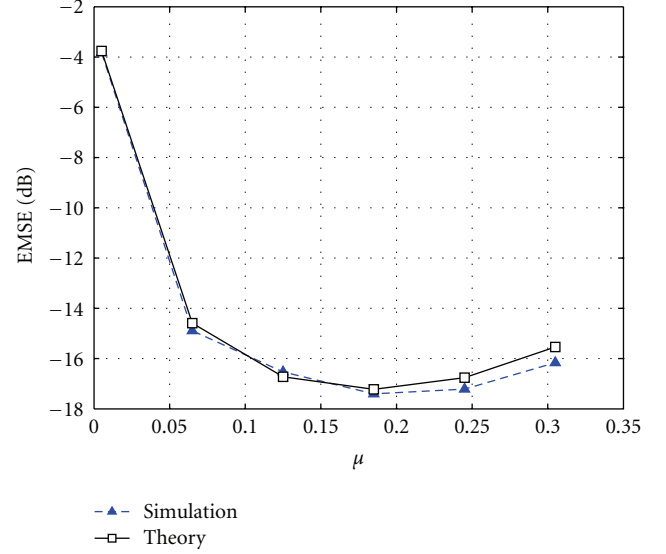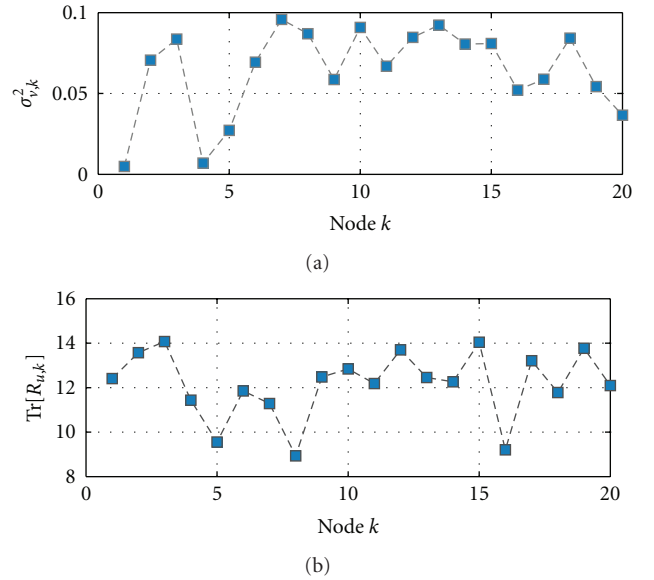
In Figures 7 and 8, the MSD and EMSE for different values of μ for node $k = 10$ are plotted. We note that unlike the ideal link case [10], here, the steady-state MSD and EMSE (and also MSE) are not a monotonically increasing function of μ.
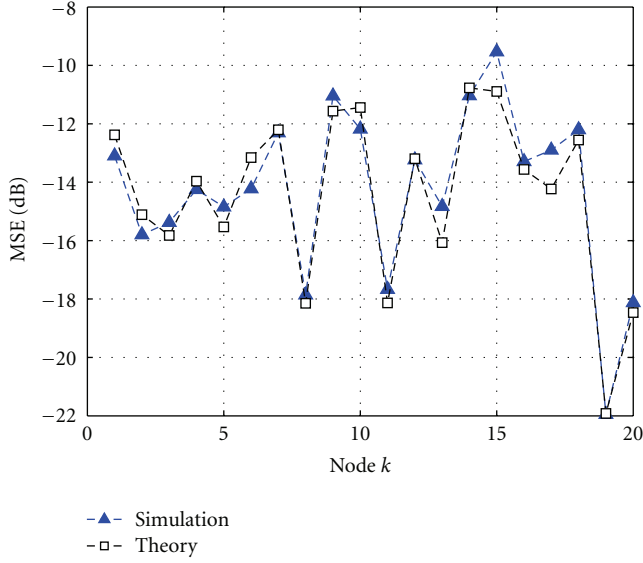
FIGURE 6: Steady-state MSE versus node, $\mu = 0.03$.

FIGURE 7: Steady-state MSD versus μ, node $k = 10$.

FIGURE 8: Steady-state EMSE versus μ, node $k = 10$.

(a)

(b)

FIGURE 9: The Observation noise power profile, (a) $\sigma_{v,k}^2$ and (b) $\text{Tr}[R_{u,k}]$.

5.2. Independent Regressors. In this case, we assume that the regressors data arise from independent Gaussian, where their eigenvalue spread is $\rho = 5$. We assume $Q_k = 10^{-3} I_M$ for channel noise. The observation noise variance $\sigma_{v,k}^2$ and $\text{Tr}[R_{u,k}]$ are shown in Figure 9. In Figures 10 and 11, the steady-state of MSD and EMSE for $\mu = 0.001$ are plotted. It is clear from Figures 10 and 11 that there is a good match between simulation and theory.

6. Conclusions and Future Work

In this paper, we derived theoretical relations to predict the performance of incremental distributed least-mean square algorithm (DILMS) when the links between nodes were noisy. Starting with the weighted energy conservation relation, we derived a variance relation for our problem. However, The important result is that unlike the ideal link case, the steady-state MSD, EMSE, and MSE are not a monotonically increasing function of step size when links are noisy. The simulation results revealed that there is good match between the derived closed-form expressions for MSD, EMSE, and MSE for every node in the network and the simulation results. Note that although in this paper we focused on the LMS adaptive filter, we can extend it for other adaptive filters. In our future work, we will consider more sophisticated cooperation modes (rather than incremental mode), such as diffusion mode.
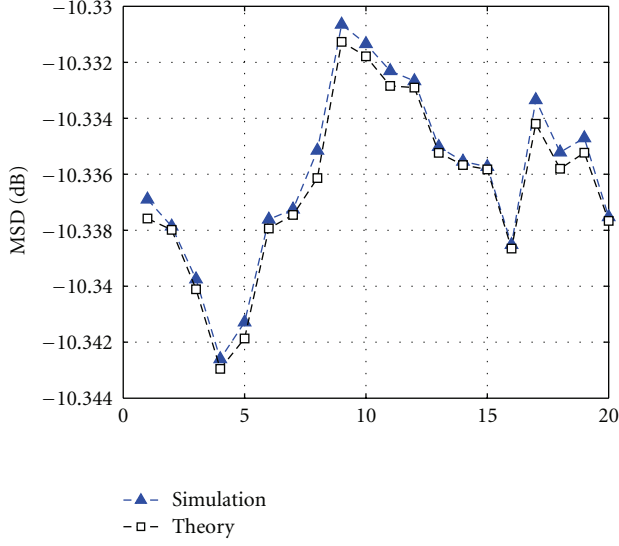
FIGURE 10: Steady-state MSD versus node, $\mu = 0.001$.
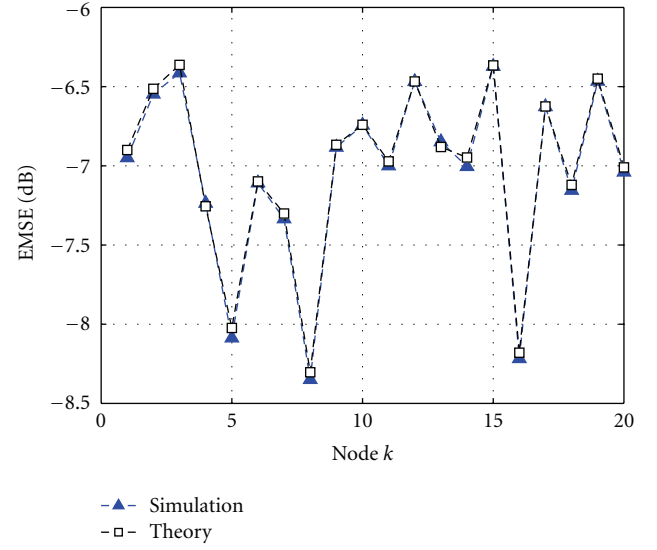
FIGURE 11: Steady-state EMSE versus node, $\mu = 0.001$.

Appendix

Consider that a network with N sensors is deployed to observe a physical phenomenon such as temperature, humidity, or any other events in a specified environment. At time i, the kth node collects a measurement $d_k(i)$ (a time-realization data). By assuming an autoregressive (AR) model to represent these measurements, we have

$$\mathbf{d}_k(i) = \sum_{m=1}^{M} \beta_m \mathbf{d}_k(i - m) + \mathbf{v}_k(i), \qquad \text{(A.1)}$$

where $\mathbf{v}_k(i)$ is additive zero-mean noise and the coefficients $\{\beta_m\}$ are parameters of the underlying model. If we define the $M \times 1$ parameter vector

$$w^o = \mathrm{col}\{\beta_1, \beta_2, \ldots, \beta_M\}, \qquad \text{(A.2)}$$

and the $1 \times M$ regression vector

$$\mathbf{u}_{k,i} = \begin{bmatrix} \mathbf{d}_k(i-1) & \mathbf{d}_k(i-2) & \cdots & \mathbf{d}_k(i-M) \end{bmatrix}, \qquad \text{(A.3)}$$

then (A.1) at each node k can be rewritten as an equivalent linear measurement model

$$\mathbf{d}_k(i) = \mathbf{u}_{k,i} w^o + \mathbf{v}_k(i). \qquad \text{(A.4)}$$

The objective becomes to estimate the model parameter vector w^o from the measurements $\mathbf{d}_k(i)$ and $\mathbf{u}_{k,i}$ over the network and thereby has the form of a system identification problem.

Acknowledgment

References

[1] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Proceedings of IEEE Interntional Conference on Acoustics, Speech, and Signal Processing*, pp. 2033–2036, Salt Lake City, Utah, USA, May 2001.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[3] J. J. Xiao, A. Ribeiro, Z. Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 27–41, 2006.

[4] J. Li and G. AlRegib, "Distributed estimation in energy-constrained wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3746–3758, 2009.

[5] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks—part I: Gaussian case," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1131–1143, 2006.

[6] A. Ribeiro and G. B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks—part II: Unknown probability density function," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2784–2796, 2006.

[7] Z. Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2210–2219, 2005.

[8] J. J. Xiao and Z. Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3564–3575, 2005.

[9] L. Xiao, S. Boyd, and S. Lall, "A spacetime diffusion scheme for peer-topeer least-squares estimation," in *Proceedings of the International Conference on Information Processing in Sensor Networks*, pp. 168–176, Nashville, Tenn, USA, 2006.

[10] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4064–4077, 2007.

[11] A. H. Sayed and C. G. Lopes, "Distributed recursive least-squares strategies over adaptive networks," in *Proceedings*

of the 40th Asilomar Conference on Signals, Systems, and Computers (ACSSC '06), pp. 233–237, Pacific Grove, Calif, USA, October-November 2006.

[12] L. Li, J. A. Chambers, C. G. Lopes, and A. H. Sayed, "Distributed estimation over an adaptive incremental network based on the affine projection algorithm," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, Article ID 5071198, pp. 151–164, 2010.

[13] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: formulation and performance analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3122–3136, 2008.

[14] F. S. Cattivelli, C. G. Lopes, and A. H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1865–1877, 2008.

[15] A. Rastegarnia, M. A. Tinati, and A. Khalili, "Performance analysis of quantized incremental LMS algorithm for distributed adaptive estimation," *Signal Processing*, vol. 90, no. 8, pp. 2621–2627, 2010.

[16] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in ad hoc WSNs with noisy links—part I: distributed estimation of deterministic signals," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 350–364, 2008.

[17] I. D. Schizas, G. B. Giannakis, S. I. Roumeliotis, and A. Ribeiro, "Consensus in Ad hoc WSNs with noisy links—part II: distributed estimation and smoothing of random signals," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1650–1666, 2008.

[18] T. C. Aysal and K. E. Barner, "Constrained decentralized estimation over noisy channels for sensor networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1398–1410, 2008.

[19] T. C. Aysal and K. E. Barner, "Blind decentralized estimation for bandwidth constrained wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, Article ID 4524301, pp. 1466–1471, 2008.

[20] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: link failures and channel noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 355–369, 2009.

[21] J. -F. Chamberland and V. V. Veeravalli, "The impact of fading on decentralized detection in power constrained wireless sensor networks," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 3, pp. 837–840, Montreal, Canada, May 2004.

[22] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, Hoboken, NJ, USA, 2003.