A Quality-aware Incremental LMS Algorithm for Distributed Adaptive Estimation

Wael M. Bazzi[1]   Amir Rastegarnia[2]   Azam Khalili[3]

[1]Electrical and Computer Engineering Department, American University in Dubai, Dubai, United Arab Emirates
[2]Department of Electrical Engineering, Malayer University, Malayer 65719-95863, Iran
[3]Department of Electrical Engineering, Malayer University, Malayer 65719-95863, Iran

Abstract: In this paper, we consider the problem of unknown parameter estimation using a set of nodes that are deployed over an area. The recently proposed distributed adaptive estimation algorithms (also known as adaptive networks) are appealing solutions to the mentioned problem when the statistical information of the underlying process is not available or it varies over time. In this paper, our goal is to develop a new incremental least-mean square (LMS) adaptive network that considers the quality of measurements collected by the nodes. Thus, we use an adaptive combination strategy which assigns each node a step size according to its quality of measurement. The adaptive combination strategy improves the robustness of the proposed algorithm to the spatial variations of signal-to-noise ratio (SNR). The performance of our algorithm is more remarkable in inhomogeneous environments when there are some nodes with low SNRs in the network. The simulation results indicate the efficiency of the proposed algorithm.

Keywords: Adaptive networks, distributed estimation, least mean-square (LMS), incremental cooperation, quality aware.

1   Introduction

Consider a set of nodes that are deployed to estimate an unknown parameter using the data collected at nodes. Such an estimation problem (known as distributed estimation problem) arises in several applications, e.g., environment monitoring, target localization and sensor network application[1−4]. Different distributed estimation algorithms have been proposed in [5–10]. In general, the available algorithms can be categorized based on the presence or absence of a fusion center (FC) for data processing (see for example [5] and references therein). In many applications, we need to perform estimation when the statistical information of the underlying processes of interest is not available or it varies over time. This issue motivated the development of novel distributed estimation algorithms that are known as adaptive networks.

We adopt the term adaptive networks from [6, 7] to refer to a collection of nodes that interact with each other and function as a single adaptive entity that is able to track statistical variations of data in real-time. In fact, in an adaptive network, each node collects local observations and at the same time, interacts with its immediate neighbors. At every instant, the local observation is combined with information from the neighboring nodes in order to improve the estimate at the local node. By repeating this process of simultaneous observation and consultation, the nodes are constantly exhibiting updated estimates that respond to observations in real time[7].

Based on the mode of cooperation between nodes, distributed adaptive estimation algorithms may be referred to as incremental algorithms or diffusion algorithms (see

Fig. 1). The distributed incremental least-mean square (DILMS)[8], distributed recursive least square (DRLS) algorithm[9], affine projection algorithm[10] and parallel projections[11] are examples of distributed adaptive estimation algorithms that use incremental cooperation. In these algorithms, there is a Hamiltonian cycle through which the estimates are sequentially circulated from node to node.
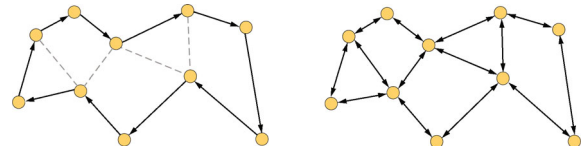

Fig. 1   Different types of cooperation mode in adaptive networks: incremental (left), diffusion (right)

On the other hand, in the diffusion cooperative mode each node updates its estimate using all available estimates from its neighbors as well as data and its own past estimate. Both least-mean square (LMS)-based and recursive least square (RLS)-based diffusion algorithms have been considered in [12, 13]. In [13], a multi-level diffusion algorithm was proposed, where a network running a diffusion algorithm is enhanced by adding special nodes that can perform different processing. In [14], an efficient adaptive combination strategy for diffusion algorithms over adaptive networks was proposed in order to improve the robustness against the spatial variation of signal-to-noise ratio (SNR) over the network.

In the previous incremental adaptive estimation algorithms[7−11], the node profile has not been considered in the algorithm structure. Therefore, the performance of such algorithms will deteriorate if the SNR at some nodes is lower than others, because the noisy estimates of such nodes pervade the entire network by incremental coopera-

tion among the nodes. To address this issue, Panigrahi and Panda have proposed an efficient step size assignment for every node in the incremental least mean squares (ILMS) adaptive learning algorithm[14]. The aim in [14] is to improve the robustness of the ILMS algorithm against the spatial variation of observation noise statistics over the network. The algorithm provides improved steady-state performance in comparison with ILMS algorithm with static step-sizes. However, it needs the noise variance information of all the nodes. In [15], the authors have proposed an efficient step size assignment for every node in the ILMS algorithm so that the convergence of the proposed algorithm is the same as the ILMS algorithm with static step-sizes.

In this paper, we propose a new incremental LMS algorithm with an adaptive combination strategy, which assigns each node a suitable weight according to its reliability of measurement. We formulate the weight assignment as a constrained optimization problem and then recast it into a distributed form and give finally an adaptive solution to the problem. Simulation results are also provided to show the performance of the proposed algorithm.

Notations. We adopt boldface letters for random quantities and normal font for nonrandom (deterministic) quantities. The symbol * denotes conjugation for scalars and Hermitian transpose for matrices. The notation $\|x\|_\Sigma^2 = x^*\Sigma x$ stands for the weighted square norm of x. The exact meaning of this notation will be clear from the context.

2 Estimation problem and the adaptive distributed solution

2.1 Incremental LMS solution

Consider a network composed of N distributed nodes as shown in Fig. 2. The purpose is to estimate an unknown parameter $M \times 1$ vector w^o from multiple spatially independent but possibly time-correlated measurements collected at N nodes in a network. Each node k has access to time-realizations $\{d_k(i), u_{k,i}\}$ of zero-mean spatial data $\{\boldsymbol{d}_k, \boldsymbol{u}_k\}$ where each $\boldsymbol{d}_k$ is a scalar measurement and each $\boldsymbol{u}_k$ is a $1 \times M$ row regression vector that are related via

$$d_k(i) = u_{k,i}w^o + v_k(i) \qquad (1)$$

where w^o is the $M \times 1$ unknown vector parameter and $v_k(i)$ is the observation noise term with zero-mean and variance of $\sigma_{v,k}^2$. In practice, w^o may represent different physical quantities, e.g., location of a target, parameter of an autoregressive (AR) model taps of a communication channel, or the location of food sources[16]. Collecting regression and measurement data into global matrices results in

$$\boldsymbol{U} \triangleq \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_N \end{bmatrix}_{N \times M}, \quad \boldsymbol{d} \triangleq \begin{bmatrix} \boldsymbol{d}_1 \\ \boldsymbol{d}_2 \\ \vdots \\ \boldsymbol{d}_N \end{bmatrix}_{N \times 1}. \qquad (2)$$
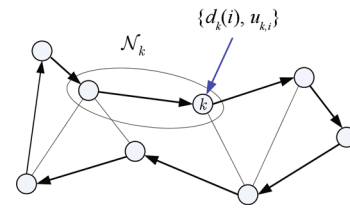

Fig. 2   A distributed network with N nodes. Note to our definition for the neighborhood of a node $\mathcal{N}_k$ in the incremental cooperation

The objective is to estimate the $M \times 1$ vector w that solves

$$\min_w J(w) \quad \text{where } J(w) = \mathrm{E}\{\|\boldsymbol{d} - \boldsymbol{U}w\|^2\}. \qquad (3)$$

The optimal solution w^o satisfies the normal equations[6]

$$R_{du} = R_u w^o \qquad (4)$$

where

$$R_{du} = \mathrm{E}\{\boldsymbol{U}^*\boldsymbol{d}\} = \sum_{k=1}^N R_{du,k} \qquad (5)$$

$$R_u = \mathrm{E}\{\boldsymbol{U}^*\boldsymbol{U}\} = \sum_{k=1}^N R_{u,k}. \qquad (6)$$

Note that in order to use (4) to compute w^o, each node must have access to the global statistical information $\{R_u, R_{du}\}$ which, in turn, needs a lot of communication and computational resources. Moreover, such an approach, does not enable the network to response to the changes in statistical properties of data.

2.2 Incremental LMS solution

In [6, 8], a fully distributed adaptive estimation algorithm, known as DILMS algorithm has been proposed. To develop our proposed algorithm, we introduce the DILMS algorithm first. To this end, we first consider the gradient-descent method. The standard gradient-descent implementation to solve the normal equation (4) is

$$w_i = w_{i-1} + \mu_k \left[\nabla J(w_{i-1})\right]^* \qquad (7)$$

where μ_k is a suitably chosen step-size parameter, w_i is an estimate of w^o at iteration i and $\nabla J(\cdot)$ denotes the gradient vector of $J(w)$ evaluated at w_{i-1}. In order to obtain a distributed solution, first the cost function $J(w)$ is decomposed as follows:

$$J(w) = \sum_{k=1}^N J_k(w) \qquad (8)$$

where

$$J_k \triangleq \mathrm{E}\{|d_k - u_k w|^2\}. \qquad (9)$$

Using (8) and (9) the standard gradient-descent implementation of (7) can be rewritten as (8)

$$w_i = w_{i-1} - \mu_k \left[\sum_{k=1}^N \nabla J_k(w_{i-1})\right]^*. \qquad (10)$$

By defining $\psi_k^{(i)}$ as the local estimate of w^o at node k and time i, w_i can be evaluated as[6, 8]

$$\psi_k^{(i)} = \psi_{k-1}^{(i)} - \mu_k \left[\nabla J_k (w_{i-1})\right]^*, \; k = 1, 2, \cdots, N. \quad (11)$$

This scheme still requires all nodes to share the global information w_{i-1}. The fully distributed solution can be achieved by using the local estimate $\psi_{k-1}^{(i)}$ at each node k instead of w_{i-1}[6]:

$$\psi_k^{(i)} = \psi_{k-1}^{(i)} - \mu_k \left[\nabla J_k \left(\psi_{k-1}^{(i)}\right)\right]^*, \; k = 1, 2, \cdots, N. \quad (12)$$

Now, we need to determine the gradient of J and replace it in (10). To do this, the following approximations are used

$$R_{du,k} \approx d_k(i) u_{k,i}^*$$
$$R_{u,k} \approx u_{k,i}^* u_{k,i}. \quad (13)$$

The resulting DILMS algorithm is as[6, 8]

$$\begin{cases} \psi_0^{(i)} \leftarrow w_{i-1} \\ \psi_k^{(i)} = \psi_{k-1}^{(i)} + \mu_k u_{k,i}^* \left[d_k(i) - u_{k,i}\psi_{k-1}^{(i)}\right] \\ \quad\quad k = 1, 2, \cdots, N \\ w_i \leftarrow \psi_N^{(i)}. \end{cases} \quad (14)$$

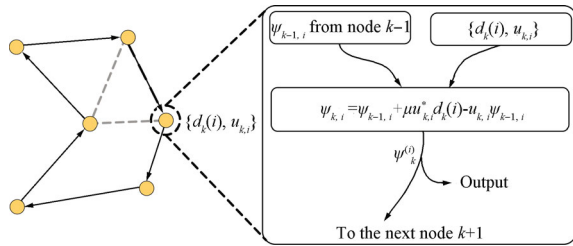The block diagram of the distributed incremental least-mean square (DILMS) algorithm is shown in Fig. 3.

Fig. 3   The block diagram of the DILMS algorithm

3   Effect of noisy nodes

Suppose that we have a network so that the SNR at some nodes is lower than others (noisy nodes). The performance of DILMS algorithm in such a network deteriorates because the noisy estimate of such a node pervades the entire network by incremental cooperation among the nodes. To see this, let us consider a network with $N = 20$ nodes and assume $M = 5$, $\mu = 0.01$, and $w^o = \mathbf{1}_M/\sqrt{M}$. The regression vectors $u_{k,i}$ are considered to be independent Gaussian where their eigenvalue spread is $\rho = 3$[1]. We assume that there are 5 noisy nodes with $\sigma_{v,k}^2 \in (0, 2)$ in the network while for other nodes we consider $\sigma_{v,k}^2 \in (0, 0.1)$. Fig. 4 shows the node profile for our set up. Fig. 5 shows the global average excess mean-square error (ESME) in different conditions when noisy nodes are included and excluded from the network. The global average EMSE is defined as

$$\zeta_g(i) = \frac{1}{N} \sum_{k=1}^{N} \mathrm{E} \left\{ |u_{k,i}(w^o - \psi_{k-1,i})|^2 \right\}$$

---

[1]Note that each $R_{u,k}$ is a diagonal matrix where its minimum eigenvalue is set to 1 and its maximum eigenvalue is set to 3. The mean of regression vectors is zero. For more information about regression vectors, please see [17].
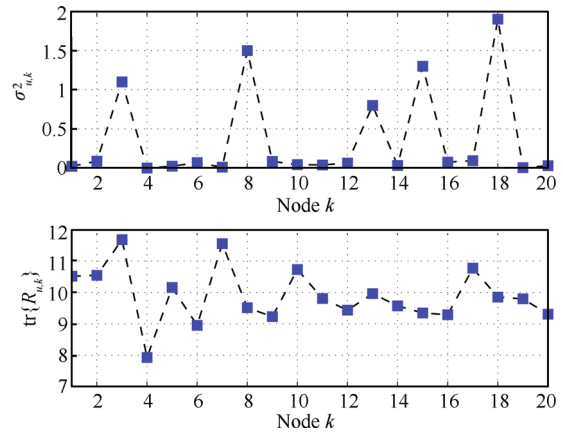

Fig. 4   The node profile, $\sigma_{v,k}^2$ and $\mathrm{tr}\{R_{u,k}\}$

We can see from Fig. 5 that the noisy nodes strongly affect the performance of DILMS algorithm. More precisely, the performance of DILMS algorithm (in terms of steady-state EMSE) is reduced by about 9 dB because of noisy nodes.
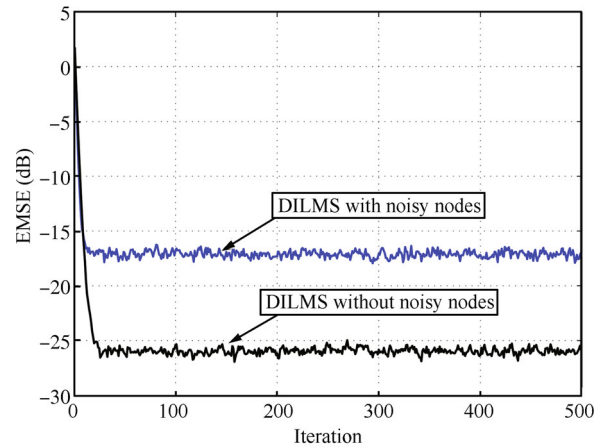

Fig. 5   The global average EMSE for DILMS algorithm in different conditions

4   Proposed algorithm

The block diagram of the proposed algorithm is shown in Fig. 6. As we can see, at the first step, we need to modify the DILMS algorithm (14) as

$$\begin{cases} \phi_{k,i} = c_{k,k-1}\psi_{k-1,i} + c_{k,k}\psi_{k,i-1} \\ \psi_{k,i} = \phi_{k,i} + \mu_k u_{k,i}^*(d_k(i) - u_{k,i}\phi_{k,i}) \end{cases} \quad (15)$$

where $\{c_{k,k-1}, c_{k,k}\} \in \mathbf{R}$ are combination coefficients used at node k. Thus, in the modified DILMS, each node updates its local estimate by combing the local estimate of previous node (i.e., $\psi_{k-1,i}$) and its previous time local estimate (i.e., $\psi_{k,i-1}$). It must be noted that at node $k = 1$, due to incremental cooperation in the DILMS algorithm, we have $c_{1,0} = c_{1,N}$.
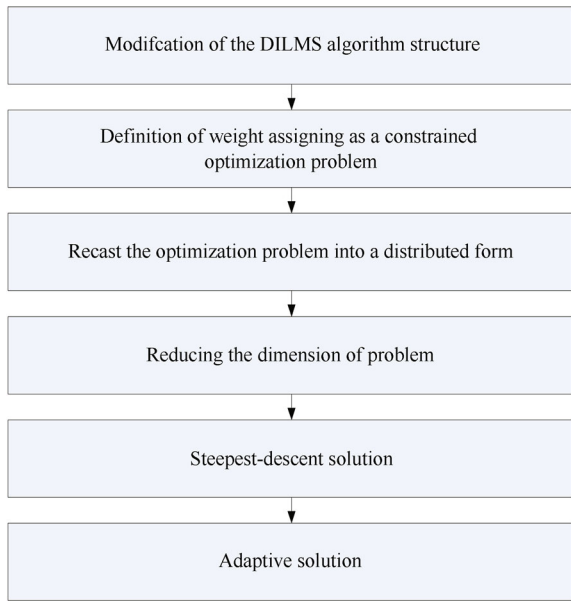
Fig. 6   The block diagram of the proposed algorithm

To formulate the problem of finding combination coefficients, we define the $N \times 1$ vector $c_k = [c_{k,1}, c_{k,2}, \cdots, c_{k,N}]^{\mathrm{T}} \in \mathbf{R}^N$ for every node k, where we have $c_{k,\ell} = 0$, $\ell \neq \{k-1, k\}$. Now, suppose that for each node k, the local estimates $\{\psi_{k,i}, \ i = 0, 1, \cdots\}$ are realizations of some random vector ψ_k. Then, we would like to find a vector of coefficients $c_k \in \mathbf{R}^N$ that solves the following problem

$$\begin{cases} \min\limits_{\{c_1, \cdots, c_N\} \in \mathbf{R}^N} \quad \sum\limits_{k=1}^{N} \mathrm{E}\{\|\Psi c_k - w^o\|^2\} \\ \text{subject to} \quad c_{k,\ell} = 0 \quad \text{for } \ell \notin \mathcal{N}_k, \ \mathcal{N}_k = \{k-1, k\} \end{cases} \tag{16}$$

where $\Psi = [\psi_1, \psi_2, \cdots, \psi_N]$ is an $M \times N$ random matrix. To have a distributed solution we use the fact that the optimization problem (16) can be decomposed into N subproblems for each $k = \{1, \cdots, N\}$ as

$$\begin{cases} \min\limits_{c_k \in \mathbf{R}^N} \quad J(c_k) = \mathrm{E}\{\|\Psi c_k - w^o\|^2\} \\ \text{subject to} \quad c_{k,\ell} = 0 \quad \text{for } \ell \notin \mathcal{N}_k, \ \mathcal{N}_k = \{k-1, k\}. \end{cases} \tag{17}$$

There is a difficulty to use the optimization problem (17): due to the presence of the unknown quantity w^o the optimization problem (17) cannot be solved directly. To address this issue, we can assume that every local estimate ψ_k is unbiased, i.e., $\mathrm{E}\{\psi_k\} = w^o$ for all $k = 1, \cdots, N$. Now we can use this assumption to say $\mathrm{E}\{\Psi\} = w^o \mathbf{1}_N^{\mathrm{T}}$. Then, we apply the bias-variance decomposition[18, 19]

$$J(c_k) = c_k^{\mathrm{T}} Q_{\Psi} c_k + \left\| (\mathbf{1}_N^{\mathrm{T}} c_k - 1) w^o \right\|^2 \tag{18}$$

where Q_{Ψ} is an $N \times N$ matrix defined as

$$Q_{\Psi} \triangleq \mathrm{E}\{(\Psi - \mathrm{E}\{\Psi\})^*(\Psi - \mathrm{E}\{\Psi\})\}. \tag{19}$$

Therefore, by imposing $\mathbf{1}_N^{\mathrm{T}} c_k = 1$, the second term involving the unknown quantity w^o is eliminated and we arrive at the following problem

$$\begin{cases} \min\limits_{c_k \in \mathbf{R}^N} \quad c_k^{\mathrm{T}} Q_{\Psi} c_k \\ \text{subject to} \quad \mathbf{1}_N^{\mathrm{T}} c_k = 1 \quad \text{and } c_{k\ell} = 0 \text{ for } \ell \notin \mathcal{N}_k. \end{cases} \tag{20}$$

As it is shown in [18], the dimension of the problem like (20) can be reduced from N unknowns to the cardinality of $\mathcal{N}_k$, say 2, by introducing the following $N \times 2$ auxiliary variable

$$P_k = [\tau_{k-1}, \tau_k]_{N \times 2} \tag{21}$$

where τ_k is the $N \times 1$ vector whose components are all zeros except k, for example $\tau_2 = [0 \ 1 \ 0 \ \cdots \ 0]^{\mathrm{T}}$. Note that due to incremental cooperation we define P_1 as $P_1 = [\tau_1 \tau_N]$.

Then, any vector c_k that satisfies $c_{k\ell} = 0$ for $\ell \in \mathcal{N}_k$ can be represented as[18, 19]

$$c_k = P_k a_k \text{ with some } a_k \ \in \mathbf{R}^2. \tag{22}$$

Therefore, substituting (22) into (20) we get

$$\begin{cases} \min\limits_{c_k \in \mathbf{R}^N} \quad f_k(a_k) \triangleq a_k^{\mathrm{T}} Q_{\Psi,k} a_k \\ \text{subject to} \quad a_k \in V_N \triangleq \{a \in \mathbf{R}^2 | \mathbf{1}_2^{\mathrm{T}} a = 1\} \end{cases} \tag{23}$$

where $Q_{\Psi,k}$ is the 2×2 matrix which is defined as

$$Q_{\Psi,k} \triangleq P_k^{\mathrm{T}} Q_{\Psi} P_k \tag{24}$$

and $\mathbf{1}_2 \triangleq P_k^{\mathrm{T}} \mathbf{1}_N$ is the vector whose components are all unity. The solution to (23) is well-known to be[18, 19]

$$a_k^o \triangleq \frac{Q_{\Psi,k}^{-1} \mathbf{1}_2}{\mathbf{1}_2^{\mathrm{T}} Q_{\Psi,k}^{-1} \mathbf{1}_2}. \tag{25}$$

To have an adaptive distributed solution, we need to eliminate the constraint V_k from the constrained optimization problem in (23). By applying a similar technique introduced in [18, 19], we finally arrive at the following iterative solution

$$\begin{cases} b_{k,i} = b_{k,i-1} - \lambda_k(i) \Lambda Q_{\Psi,k} b_{k,i-1} \\ c_k = P_k b_{k,i} \end{cases} \tag{26}$$

where $\lambda_k(i)$ is the step-size and Λ is the following matrix

$$\Lambda = \begin{bmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{bmatrix}. \tag{27}$$

To derive an adaptive solution, we replace $Q_{\Psi,k}$ by its instantaneous approximation as

$$Q_{\Psi,k} \approx (\Delta \phi)^* (\Delta \phi) \tag{28}$$

where $\Delta \phi$ is an $M \times 2$ matrix as

$$\Delta \phi = [\psi_{k-1,i} - \psi_{k-1,i-1} \quad \psi_{k,i-1} - \psi_{k,i-2}]. \tag{29}$$

Now by substituting (28) into (26) and using the modified DILMS algorithm (10), our proposed algorithm with adaptive combination algorithm yields

$$
\begin{cases}
g_{k,i} = \Lambda(\Delta\phi)^*(\Delta\phi)b_{k,i-1} \\
b_{k,i} = b_{k,i-1} - \lambda_k(i)g_{k,i} \\
c_k = P_k b_{k,i} \\
\phi_{k,i} = c_{k,k-1}\psi_{k-1,i} + c_{k,k}(i)\psi_{k-1,i} \\
\psi_{k,i} = \phi_{k,i} + \mu_k u_{k,i}^*(d_k(i) - u_{k,i}\phi_{k,i}).
\end{cases}
\tag{30}
$$

A possible choice for $\lambda_k(i)$ is the normalized step-size[18, 19]

$$
\lambda_k(i) = \gamma \frac{\min\{b_{k,i-1}(1), b_{k,i-1}(2)\}}{\| g_{k,i} \|_\infty + \varepsilon}
\tag{31}
$$

where $\gamma \in (0,1)$ and ε are constants, $\| \cdot \|_\infty$ denotes the maximum norm, and $b_{k,i-1}(m)$ is the m-th component of $b_{k,i-1}$. The pseudo code of the proposed algorithm is shown in sequel.

Algorithm 1. The pseudo code of the proposed algorithm

    Initialize $\psi_{k,0} = 0$
    for $k = 1, 2, \cdots, N$
       for $i = 1, 2, \cdots$
          compute $\Delta\phi$ using (29)
          $g_{k,i} = \Lambda(\Delta\phi)^*(\Delta\phi)b_{k,i-1}$
          $b_{k,i} = b_{k,i-1} - \lambda_k(i)g_{k,i}$
          $c_k = P_k b_{k,i}$
          $\phi_{k,i} = c_{k,k-1}\psi_{k-1,i} + c_{k,k}(i)\psi_{k-1,i}$
          $\psi_{k,i} = \phi_{k,i} + \mu_k u_{k,i}^*(d_k(i) - u_{k,i}\phi_{k,i})$
       end for
    end for

5 Simulation results

In this section, we present the simulation results to evaluate the performance of our proposed algorithm. To this aim, we consider again the set up in Section 3. We further choose $\gamma = 0.01$ and $\varepsilon = 10^{-5}$. We examine the network performance by the global average mean-square deviation (MSD) and global average EMSE. Note that global average MSD is defined as

$$
\eta_g(i) = \frac{1}{N} \sum_{k=1}^{N} \mathrm{E}\left\{ \|w^o - \psi_{k-1,i}\|^2 \right\}.
\tag{32}
$$

The curves are generated by averaging over 100 independent experiments. In Fig. 7, we have shown the global average MSD and global average EMSE for the DILMS with noisy nodes, the proposed algorithm, and DILMS without noisy nodes. As it is clear from Fig. 7, the proposed algorithm outperforms the DILMS algorithm in a sense of steady-state performance. In fact, in this special set up, the performance of the proposed algorithm is increased by about 7 dB in comparison with the DILMS algorithm. The main disadvantage of the proposed algorithm is its convergence rate, as it is obvious from Fig. 8. In Fig. 8, we have plotted the EMSE learning curves for the DILMS with noisy

nodes and the proposed algorithm by manipulating the step size so that both the algorithms have the same convergence
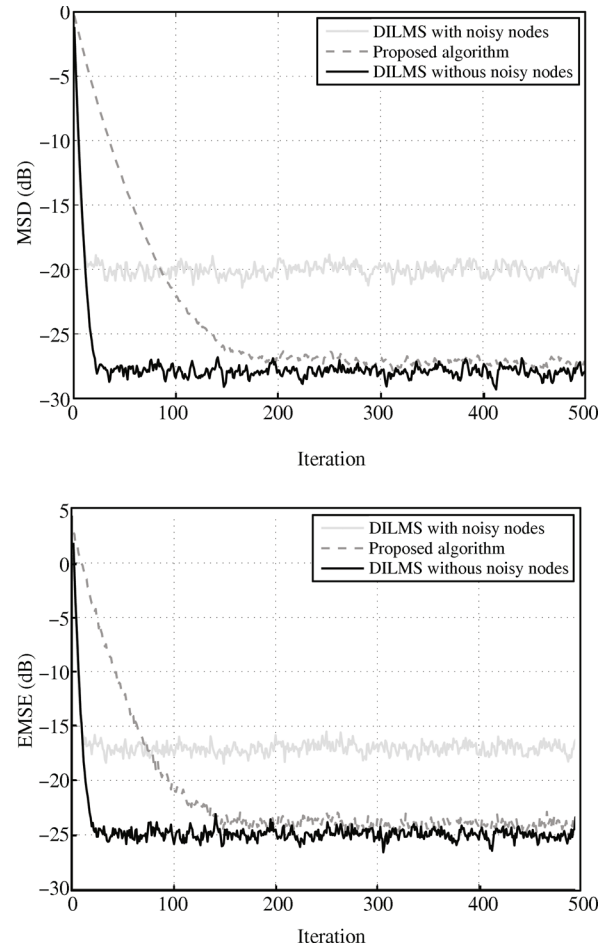

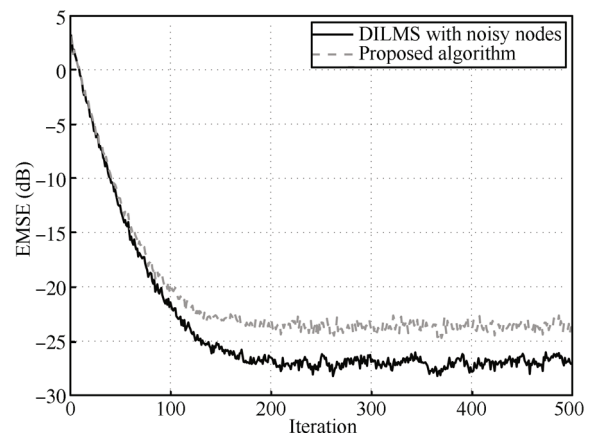Fig. 7   The global average MSD and ESME for DILMS and our proposed algorithm

Fig. 8   The EMSE learning curves for the DILMS with noisy nodes and the proposed algorithm. Note that we have manipulated the step size so that both the algorithm have same convergence speed

speed. The approximation of $Q_{\Psi,k}$ in (28) is the main reason for low convergence speed of the proposed algorithm, which needs time to know the exact variance of a node. The smaller convergence rate in adaptive networks means that the algorithm needs more iterations to converge, which in turn increases the required power, and reduces the network life. The convergence rate of the proposed algorithm can be improved by using a variable step-size LMS algorithm. It is possible to use the block adaptive algorithm concept to minimize the communication overhead involved in the algorithm[20, 21].

6 Conclusions and future work

In this paper, we proposed a new incremental LMS algorithm with adaptive combination strategy. The adaptive combination strategy improves the robustness of the proposed algorithm to the spatial variation of SNR. We formulated the weight assignment as a constrained optimization problem and then changed it into a distributed form and finally derived an adaptive solution to the problem. The simulation results showed the effectiveness of our proposed algorithm. In this paper, we assumed ideal links between nodes in the networks. As we have shown in [22, 23], the performance of incremental adaptive networks deteriorates in the presence of noisy links. Moreover, the performance of adaptive networks can vary significantly when they are implemented in finite-precision arithmetic, which makes it vital to analyze their performance in a quantized environment. In the future work, we will consider these issues to modify the structure of our proposed algorithm.

References

[1] D. Estrin, G. Pottie, M. Srivastava. Instrumenting the world with wireless sensor networks. In *Proceedings of International Conference on Acoustics, Speech, Signal Processing*, IEEE, Salt Lake City, UT, USA, pp. 2033–2036, 2001.

[2] H. Leung, S. Chandana, S. Wei. Distributed sensing based on intelligent sensor networks. *IEEE Circuits and Systems Magazine*, vol. 8, no. 2, pp. 38–52, 2008.

[3] R. P. Yadav, P. V. Varde, P. S. V. Nataraj, P. Fatnani, C. P. Navathe. Model-based tracking for agent-based control systems in the case of sensor failures. *International Journal of Automation and Computing*, vol. 9, no. 6, pp. 561–569, 2012.

[4] C. L. Wang, T. M. Wang, J. H. Liang, Y. C. Zhang, Y. Zhou. Bearing-only visual SLAM for small unmanned aerial vehicles in GPS-denied environments. *International Journal of Automation and Computing*, vol. 10, no. 5, pp. 387–396, 2013.

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

[6] A. H. Sayed, C. G. Lopes. Distributed recursive least-squares strategies over adaptive networks. In *Proceedings of the 40th Asilomar Conference on Signals, Systems and Computers*, IEEE, Pacific Grove, CA, USA, pp. 233–237, 2006.

[7] A. H. Sayed, C. G. Lopes. Adaptive processing over distributed networks. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E90-A, no. 8, pp. 1504–1510, 2007.

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

[9] C. G. Lopes, A. H. Sayed. Diffusion least-mean-squares over adaptive networks. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, Honolulu, Hawaii, USA, vol. 3, pp. 917–920, 2007.

[10] L. L. Li, J. A. Chambers, C. G. Lopes, 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, pp. 151–164, 2010.

[11] N. Takahashi, I. Yamada. Incremental adaptive filtering over distributed networks using parallel projection onto hyperslabs. *IEICE Technical Report*, vol. 108, pp. 17–22, 2008.

[12] C. G. Lopes, 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.

[13] F. S. Cattivelli, C. G. Lopes, 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.

[14] T. Panigrahi, G. Panda. Robust distributed learning in wireless sensor network using efficient step size. *Procedia Technology*, vol. 6, pp. 864-871, 2012.

[15] A. Khalili, A. Rastegarnia, J. A. Chambers, W. M. Bazzi. An optimum step-size assignment for incremental LMS adaptive networks based on average convergence rate constraint. *AEU–International Journal of Electronics and Communications*, vol. 67, no. 3, pp. 263–268, 2013.

[16] A. H. Sayed. Diffusion adaptation over networks. *E-Reference Signal Processing*, R. Chellapa, S. Theodoridis, Eds., New York: Elsevier, 2013.

         International Journal of Automation and Computing 11(6), December 2014

[17] A. H. Sayed. *Adaptive Filters*, New Jersey: Wiley, 2008.

[18] I. Yamada, N. Ogura. Adaptive projected subgradient method for asymptotic minimization of sequence of nonnegative convex functions. *Numerical Functional Analysis and Optimization*, vol. 25, no. 7–8, pp. 593–617, 2004.

[19] N. Takahashi, I. Yamada, A. H. Sayed. Diffusion least-mean squares with adaptive combiners. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, IEEE, Taipei, Taiwan, China, pp. 2845–2848, 2009.

[20] A. Khalili, M. A. Tinati, A. Rastegarnia. An incremental block LMS algorithm for distributed adaptive estimation. In *Proceedings of IEEE International Conference on Communication Systems*, IEEE, Singapore, pp. 493–496, 2010.

[21] M. A. Tinati, A. Khalili, A. Rastegarnia. A diffusion block least-mean square algorithm for distributed adaptive estimation. In *Proceedings of the 3rd IET International Conference on Wireless, Mobile and Multimedia Networks*, IEEE, Beijing, China, pp. 270–273, 2010.

[22] A. Khalili, M. A. Tinati, A. Rastegarnia. Performance analysis of distributed incremental LMS algorithm with noisy links. *International Journal of Distributed Sensor Networks*, vol. 2011, Article number 756067, 2011.

[23] A. Khalili, M. A. Tinati, A. Rastegarnia. Steady-state analysis of incremental LMS adaptive networks with noisy links. *IEEE Transactions Signal Processing*, vol. 59, no. 5, pp. 2416–2421, 2011.

Wael M. Bazzi graduated from the American University of Beirut (AUB), Lebanon in 1996. He received the M. Eng. degree from AUB in 1999 and the Ph. D. degree from the University of Waterloo, Canada, in 2001. He is currently an associate professor at the American University in Dubai.

His research interests include wireless communication and networks, especially the optimization and modeling aspects of communication networks and systems

E-mail: wbazzi@and.edu

Amir Rastegarnia completed his Ph. D. degree in the electrical engineering at the University of Tabriz, Tabriz, Iran in 2011. In 2011, he joined the Department of Electrical Engineering, Malayer University, as assistant professor. He is a member of IEEE.

His research interests include theory and methods for adaptive and statistical signal processing, distributed adaptive estimation, as well as signal processing for communications.

E-mail: a_rastegar@ieee.org (Corresponding author)

Azam Khalili received the Ph. D. degree in electrical engineering from the University of Tabriz, Tabriz, Iran in 2011. In 2011, she joined the Department of Electrical Engineering, Malayer University, as assistant professor. She is a member of IEEE.

Her research interests include theory and methods for adaptive filtering, distributed adaptive estimation, as well as signal processing for communications.

E-mail: a.khalili@ieee.org