VARIABLE STEP-SIZE LEAST MEAN SQUARE ALGORITHMS OVER ADAPTIVE NETWORKS

Muhammad Omer Bin Saeed, Azzedine Zerguine, and Salam A. Zummo

Department of Electrical Engineering
King Fahd University of Petroleum & Minerals
Dhahran, 31261, Saudi Arabia
Email: {mobs, azzedine, zummo}@kfupm.edu.sa

ABSTRACT

This paper presents two new variations for algorithms over adaptive networks using a variable step-size strategy in order to enhance the overall performance. Variable step-size least mean square (VSSLMS) algorithms for incremental as well as diffusion strategies are studied and the results are compared with existing results. A comparison is done with the recently proposed Diffusion LMS algorithm with adaptive combiners and it is shown that VSSLMS provides a simplified solution than that of the Diffusion LMS algorithm with adaptive combiners. Also, a great improvement in performance is obtained when compared with the fixed step-size Incremental and Diffusion LMS techniques.

Index Terms – Adaptive filters, Variable step-size least mean square, diffusion algorithm, incremental algorithm, adaptive networks

1. INTRODUCTION

Here, we study the problem of distributed estimation over adaptive networks [1–4], which uses cooperation between neighboring nodes of a network in order to estimate a particular parameter. Thus, the spatial and temporal diversities of the network are exploited to enhance the performance of the system. The performance of such adaptive networks is based strongly on the adaptive strategy being used, namely, incremental [1], diffusion [2], or probability diffusion [3]. This paper focuses on the incremental and diffusion modes. The incremental LMS algorithm is prone to node failure but provides better performance [1, 2, 4]. The diffusion strategy, however, is more robust to node and link failure [3].

In [4], a new systematic approach was provided to vary the combination weights of the network in order to improve performance of the adaptive network, particularly in the case where a node suffers from low Signal-to-Noise Ratio (SNR). However, the solution was complex and the improvement was not very significant.

In this paper, we use the VSSLMS algorithm [5] to show that varying the step-size can provide a much simpler solution for improving the performance of the system.

The paper begins by defining the system model in Section 2. Section 3 introduces the variable step-size variations for both schemes. Section 4 presents mathematical analysis of the two schemes in comparison with previous algorithms. Simulation results are shown in Section 5 and the paper concludes with some remarks in Section 6.

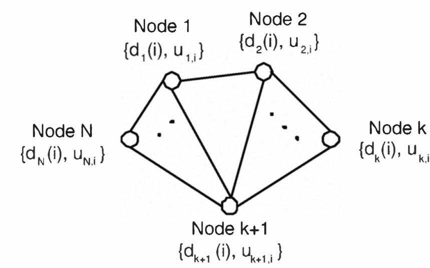

Figure 1: An adaptive network with *N* nodes.

2. VSSLMS ALGORITHMS OVER ADAPTIVE NETWORKS

Let us introduce the notation that is being used in the paper. Boldface letters are used for vectors/matrices and normal font for scalar quantities. Matrices are defined by capital letters and small letters are used for vectors. The notation $(.)^T$ stands for transposition for vectors and matrices and expectation operation is denoted by $E[.]$. Let us suppose we have N nodes in a network with a predefined topology. For each node k, the number of neighbors is given by N_k, including the node k itself, as shown in Fig. 1. At each iteration, i, the output of the system at each node is given by

$$d_k(i) = \mathbf{u}_{k,i}\mathbf{w}^o + v_k(i), \tag{1}$$

where $\mathbf{u}_{k,i}$ is a known regression row vector of length M, $\mathbf{w}^o$ is an unknown column vector of length M and $v_k(i)$ is

978-1-4244-7167-6/10/$26.00 ©2010 IEEE

noise. The output and regression data are used to produce an estimate of the unknown vector, given by $\psi_{k,i}$.

The adaptation can be performed using two different strategies. The first one is called Incremental Least Mean Squares (ILMS) [1] where each node updates its own estimate at every iteration and then passes on its estimate to the next node. The estimate of the last node is taken as the final estimate of that iteration. The second strategy is called Diffusion LMS (DLMS) [2] where each node combines its own estimate with the estimates of its neighbors using some combination technique and then the combined estimate is used for updating the node estimate. This technique is referred to as Combine-then-Adapt (CTA) diffusion [2]. It is also possible to first update the estimate using the estimate from the previous iteration and then combine the updates from all neighbors to form the final estimate for the iteration. This technique is known as Adapt-then-Combine (ATC) diffusion [4]. Simulation results suggest that ATC scheme outperforms the CTA scheme. Therefore, this paper uses ATC diffusion technique only. Using LMS, the ATC diffusion algorithm is given by

$$\left\{ \begin{array}{lcl} \boldsymbol{\phi}_{k,i} & = & \boldsymbol{\psi}_{k,i-1} + \mu_k \mathbf{u}_{k,i}^T \left(d_k(i) - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k,i-1} \right) \\ \boldsymbol{\psi}_{k,i} & = & \Sigma_{l \epsilon N_k} c_{lk}(i) \, \boldsymbol{\phi}_{l,i} \end{array} \right. , \tag{2}$$

where $\{c_{lk}\}_{l \epsilon N_k}$ is a combination weight for node k, that may be fixed [2] or adapting every iteration [4], while $\{\boldsymbol{\phi}_{l,i}\}_{l \epsilon N_k}$ is the local estimate for each node neighboring node k, and μ_k is the node step-size.

The strategy in [4] uses adaptive combiners in order to improve the performance, given the fact that nodes with low SNR perform poorly and so should be given lesser weightage. The final algorithm is complex and does not provide a lot of improvement. In comparison, varying the step-size for each node provides a simpler solution and gives better results as well. This will be demonstrated via simulation results.

3. VARIABLE STEP-SIZE LMS

In [5], a variable step-size lms (VSSLMS) algorithm was introduced in which the step-size was adapted at each iteration using the instantaneous power of the error. Since then various VSS strategies have been proposed. However, the authors of [6] proved that an optimally designed VSSLMS algorithm of [5] outperformed the other VSS variations. Therefore, this paper uses the VSS algorithm of [5] for both strategies. These changes are incorporated in both ILMS algorithm and DLMS algorithm as shown in the ensuing analysis.

3.1. Incremental VSS LMS (VSSILMS)

The ILMS algorithm is defined by

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{\psi}_{k-1,i} + \mu_k \mathbf{u}_{k,i}^T \left(d_k(i) - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k-1,i} \right). \tag{3}$$

Every node updates its estimate based on the updated estimate coming from the previous node. Here, the error is given by:

$$e_k(i) = d_k(i) - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k-1,i} \tag{4}$$

Using the algorithm in [5], the step-size is updated according to

$$\mu_k(i) = \alpha \mu_k(i-1) + \gamma e^2(i), \tag{5}$$

where α and γ are controlling parameters. Using this update in the ILMS equation results in the VSSILMS algorithm defined by the following recursion:

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{\psi}_{k-1,i} + \mu_k(i) \mathbf{u}_{k,i}^T \left(d_k(i) - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k-1,i} \right). \tag{6}$$

3.2. Diffusion VSS LMS (VSSDLMS)

In the DLMS algorithm, defined by (2), the error is given by

$$e_k(i) = d_k(i) - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k,i-1}. \tag{7}$$

Using this error in (5) gives the VSSDLMS algorithm which is given by the following update equations:

$$\left\{ \begin{array}{lcl} \boldsymbol{\phi}_{k,i} & = & \boldsymbol{\psi}_{k,i-1} + \mu_k(i) \mathbf{u}_{k,i}^T \left(d_k(i) - \mathbf{u}_{k,i} \boldsymbol{\psi}_{k,i-1} \right) \\ \boldsymbol{\psi}_{k,i} & = & \Sigma_{l \epsilon N_k} c_{lk}(i) \, \boldsymbol{\phi}_{l,i} \end{array} \right. . \tag{8}$$

The update equation is the same as (5).

4. PERFORMANCE ANALYSIS

4.1. VSSILMS

Following the procedure given in [1], and including the VSS update equation (5) in the analysis, the variance relation in [1] is given by

$$E\left[\|\overline{\boldsymbol{\psi}}_k\|_{\overline{\boldsymbol{\Sigma}}}^2 \right] = E\left[\|\overline{\boldsymbol{\psi}}_{k-1}\|_{\overline{\boldsymbol{\Sigma}}'}^2 \right] + \mu_k^2 \sigma_{v,k}^2 Tr\left(\boldsymbol{\Lambda}_k \overline{\boldsymbol{\Sigma}} \right) \tag{9}$$

$$\overline{\boldsymbol{\Sigma}}' = \overline{\boldsymbol{\Sigma}} - \mu_k \left(\boldsymbol{\Lambda}_k \overline{\boldsymbol{\Sigma}} + \overline{\boldsymbol{\Sigma}} \boldsymbol{\Lambda}_k \right) + \mu_k^2 \left(\boldsymbol{\Lambda}_k Tr\left(\overline{\boldsymbol{\Sigma}} \boldsymbol{\Lambda}_k \right) + 2 \boldsymbol{\Lambda}_k \overline{\boldsymbol{\Sigma}} \boldsymbol{\Lambda}_k \right) \tag{10}$$

where $\overline{\boldsymbol{\Sigma}}$ is a Gaussian normalized weighting matrix, $\boldsymbol{\Lambda}_k$ is a diagonal matrix containing the eigenvalues for the regressor vector at node k, and $\sigma_{v,k}^2$ is the noise variance. These equations show how the variance of the Gaussian normalized error vector $\overline{\boldsymbol{\psi}}_k$ iterates from node to node. When (5) is incorporated into the results, the independence assumption [5] is invoked

$$E\left[\mu_k(i) \mathbf{u}_{k,i}^T e_k(i) \right] = E\left[\mu_k(i) \right] E\left[\mathbf{u}_{k,i}^T e_k(i) \right], \tag{11}$$

and therefore, (9) and (10) respectively look like the new following recursions:

$$E\left[\left\|\overline{\psi}_k^i\right\|_{\overline{\Sigma}}^2\right] = E\left[\left\|\overline{\psi}_{k-1}^i\right\|_{\overline{\Sigma}'}^2\right] + E\left[\mu_{k,i-1}^2\right]\sigma_{v,k}^2 Tr\left(\mathbf{\Lambda}_k \overline{\mathbf{\Sigma}}\right) \tag{12}$$

$$\begin{aligned} \overline{\mathbf{\Sigma}}' &= \overline{\mathbf{\Sigma}} - E\left[\mu_{k,i-1}\right]\left(\mathbf{\Lambda}_k\overline{\mathbf{\Sigma}} + \overline{\mathbf{\Sigma}}\mathbf{\Lambda}_k\right) \\ &\quad + E\left[\mu_{k,i-1}^2\right]\left(\mathbf{\Lambda}_k Tr\left(\overline{\mathbf{\Sigma}}\mathbf{\Lambda}_k\right) + 2\mathbf{\Lambda}\overline{\mathbf{\Sigma}}\mathbf{\Lambda}\right). \end{aligned} \tag{13}$$

At steady-state, the step-size expectation values in (12) and (13) $\left(\overline{\mu} = E\left[\mu_{k,\infty}\right] \text{ and } \overline{\mu^2} = E\left[\mu_{k,\infty}^2\right]\right)$ are given by [5]

$$\overline{\mu_k} = \frac{\gamma\left(\mathcal{M}_k + \sigma_{v,k}^2\right)}{1-\alpha} \tag{14}$$

$$\overline{\mu_k^2} = \frac{2\alpha\gamma\overline{\mu}\left(\sigma_{v,k}^2 + \mathcal{M}_k\right) + 3\gamma^2\left(\sigma_{v,k}^2 + \mathcal{M}_k\right)^2}{1-\alpha^2} \tag{15}$$

where $\mathcal{M}$ is the steady-state misadjustment for the step-size and is given by

$$\mathcal{M}_k = \sqrt{\frac{1 - \left[1 - 2\frac{(3-\alpha)\gamma\sigma_{v,k}^2}{1-\alpha^2}Tr\left(\mathbf{\Lambda}\right)\right]}{1 + \left[1 - 2\frac{(3-\alpha)\gamma\sigma_{v,k}^2}{1-\alpha^2}Tr\left(\mathbf{\Lambda}\right)\right]}}. \tag{16}$$

These equations can be directly incorporated into steady-state equations for Mean Square Deviation (MSD) and Excess Mean Square Error (EMSE) in [1] to get the values of MSD and EMSE for the newly proposed VSS algorithm.

4.2. VSSDLMS

As was done for the VSSILMS algorithm, the variance relation for the the DLMS algorithm is given by [2]

$$E\left[\left\|\overline{\psi}^i\right\|_{\overline{\sigma}}^2\right] = E\left[\left\|\overline{\psi}^{i-1}\right\|_{\boldsymbol{F}\overline{\sigma}}^2\right] + \boldsymbol{b}^T\overline{\boldsymbol{\sigma}} \tag{17}$$

$$\begin{aligned} \overline{\boldsymbol{F}} &= \left(\overline{\boldsymbol{G}}^T \odot \overline{\boldsymbol{G}}^*\right)\left[\boldsymbol{I}_{N^2M^2} - \left(\boldsymbol{I}_{NM} \odot \mathbf{\Lambda}\boldsymbol{D}\right)\right. \\ &\quad \left. - \left(\mathbf{\Lambda}\boldsymbol{D} \odot I_{NM}\right) + \left(\boldsymbol{D} \odot \boldsymbol{D}\right)\boldsymbol{A}\right] \end{aligned} \tag{18}$$

$$\overline{\boldsymbol{\sigma}} = bvec\left\{\overline{\mathbf{\Sigma}}\right\} \tag{19}$$

$$\boldsymbol{b} = bvec\left\{\boldsymbol{R}_v\boldsymbol{D}^2\mathbf{\Lambda}\right\} \tag{20}$$

where $\boldsymbol{R}_v$ is the noise auto-correlation matrix, $bvec\{.\}$ is the block vectorization operator [2], $\boldsymbol{D}$ is the block diagonal step-size matrix for the whole network, $\overline{\boldsymbol{G}}$ is the block combiner matrix, $\boldsymbol{A}$ is the block vectorized form of the fourth order weighted moment of the regressor vectors, and $\odot$ is the block Kronecker product [2]. Again using (11), (18) and (20) become

$$\begin{aligned} \overline{\boldsymbol{F}} &= \left(\overline{\boldsymbol{G}}^T \odot \overline{\boldsymbol{G}}^*\right)\left[\boldsymbol{I}_{N^2M^2} - \left(\boldsymbol{I}_{NM} \odot \mathbf{\Lambda}E\left[\boldsymbol{D}\right]\right)\right. \\ &\quad \left. - \left(\mathbf{\Lambda}E\left[\boldsymbol{D}\right] \odot \boldsymbol{I}_{NM}\right) + E\left[\left(\boldsymbol{D} \odot \boldsymbol{D}\right)\right]\boldsymbol{A}\right] \end{aligned} \tag{21}$$

$$\boldsymbol{b} = bvec\left\{\boldsymbol{R}_v E\left[\boldsymbol{D}^2\right]\mathbf{\Lambda}\right\} \tag{22}$$

Since the step-size matrix is block-diagonal, this operation becomes straight-forward. Steady-state matrices can be formed for the step-sizes using (14) and (15). Invoking these matrices directly into the steady-state equations of [2] gives us MSD and EMSE values for the new algorithm.

5. SIMULATIONS

Simulation examples are reported here to illustrate the performance of the VSSLMS algorithms over adaptive networks. Results are compared with the fixed step-size algorithms in [1, 2]. A network topology with $N = 15$ nodes is considered (Fig. 2). For VSSILMS algorithm, alpha is set to 0.997 and gamma to 2×10^{-4} while for VSSDLMS these are set to 0.998 and 2×10^{-5}, respectively. These values ensure that the convergence rate is the same for all algorithms. Two scenarios are presented in the results. In the first one, the SNR and noise power profile is as given in Fig. 3. In the second one, the noise power for Node 5 is increased from 6×10^{-4} to 1×10^{-2}, which reduces the SNR to about 18 dB. As a result, there is a deterioration in performance.

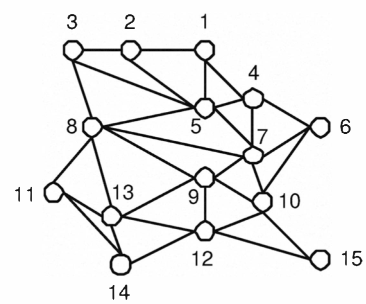

Figure 2: Network topology for simulation example.

Fig. 4 depicts the results obtained for the case of scenario 1. As can be seen from this figure, great improvement in performance is obtained through the use of the VSS strategy in both incremental and diffusion techniques. About 25 dB difference in favor of the VSSDLMS as compared to its fixed step-size counterpart is obtained. In the case of scenario 2, the performance of the VSSDLMS deteriorates only by about 3 dB whereas that of the VSSILMS deteriorates by nearly 9 dB, as can be seen from Fig. 5. Fig. 6 details the steady-state MSD values at each node for all the algorithms, which shows the superiority of the VSSDLMS among all.

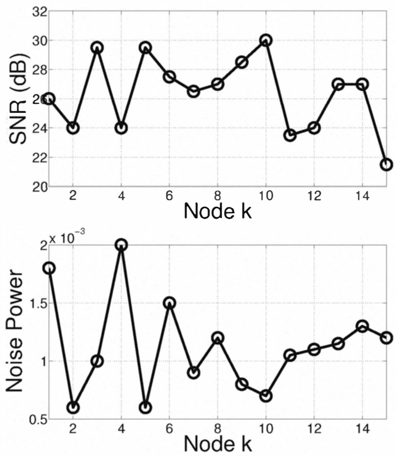

Figure 3: SNR (top) and Noise Power (bottom) for N = 15 nodes.

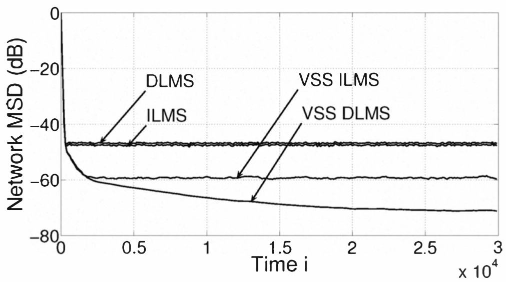

Figure 4: Learning behavior of network MSDs for Scenario 1.

6. CONCLUDING REMARKS

Here in this work, we have introduced the variable step-size LMS algorithm as an improvement over the fixed step-size algorithms previously being used for estimation over adaptive networks. Moreover, the proposed VSS strategy avoids the calculation of the error-correlation-matrix in diffusion algorithm [4], and therefore, a reduction in the computational complexity is achieved through the use of this technique.

Acknowledgment: This research work is funded by King Fahd University of Petroleum & Minerals (KFUPM) under Research Grant (FT100012).

7. REFERENCES

[1] C.G. Lopes and A.H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4064-4077, Aug. 2007.

[2] C.G. Lopes and A.H. Sayed, "Diffusion least-mean squares over adaptive networks: formulation and performance analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3122-3136, July 2008.

[3] C.G. Lopes and A.H. Sayed, "Diffusion adaptive networks with changing topologies," in *Proc. IEEE ICASSP*, Las Vegas, NV, Apr. 2008, pp. 3285-3288.

[4] N. Takahashi, I. Yamada and A.H. Sayed, "Diffusion least-mean squares with adaptive combiners," in *Proc. ICASSP*, Taipei, Taiwan, April 2009, pp. 2845-2848.

[5] R.H. Kwong and E.W. Johnston, "A variable step size LMS algorithm," in *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1633-1642, 1992.

[6] C.C.G. Lopes and J.C.M. Bermudez, "Evaluation and design of variable step size adaptive algorithms," *Proc. IEEE ICASSP*, Salt Lake City, UT, May 2001, pp. 3845-3848.

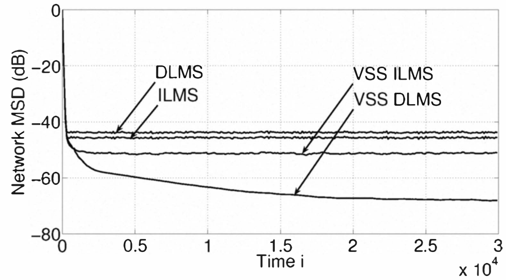

Figure 5: Learning behavior of network MSDs for Scenario 2.

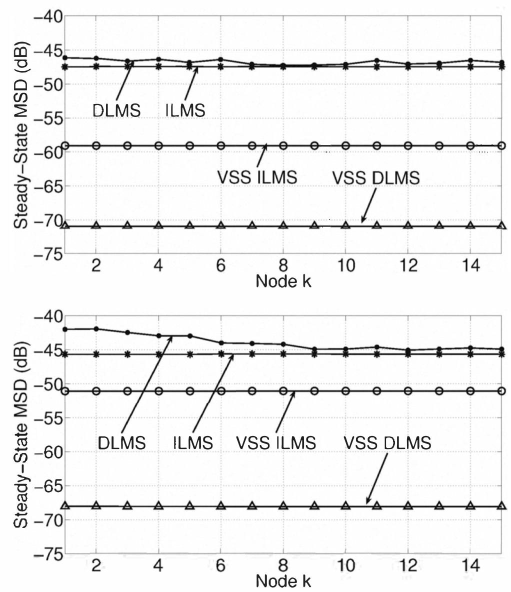

Figure 6: Comparison of steady-state MSD for scenario 1 (top) and scenario 2 (bottom).