Cooperation Gain in Incremental LMS Adaptive Networks with Noisy Links

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

[1] Department of Electrical Engineering, Malayer University
Malayer, 65719-95863, Iran
{a_rastegar,a.khalili}@ieee.org
[2] Electrical and Computer Engineering Department, American University in Dubai
P.O. Box 28282, Dubai, United Arab Emirates
wbazzi@aud.edu

Abstract. In this paper, we study the influence of noisy links on the effectiveness of cooperation in incremental LMS adaptive network (ILMS). The analysis reveals the important fact that under noisy communication, cooperation among nodes may not necessarily result in better performance. More precisely, we first define the concept of cooperation gain and compute it for the ILMS algorithm with ideal and noisy links. We show that the ILMS algorithm with ideal links outperforms the non-cooperative scheme for all values of step-size (cooperation gain is always bigger than 1). On the other hand, in the presence of noisy links, cooperation gain is not always bigger than 1 and based on the channel and data statistics, for some values of step-size, non-cooperative scheme outperforms the ILMS algorithm. We presented simulation results to clarify the discussions.

Keywords: adaptive networks, cooperation gain, incremental, step-size.

1 Introduction

An adaptive network is a collection of spatially distributed nodes that interact with each other, and function as a single adaptive entity that is able to respond to data in real-time and also track variations in their statistical properties [1-3]. Based on the mode of cooperation between nodes, adaptive networks can be roughly classified into incremental [1-6], diffusion [6-11], and hierarchical [12], [13] algorithms. In incremental based adaptive networks, a Hamiltonian cycle is established through the nodes and each node cooperates only with one adjacent node to exploit the spatial dimension, whilst performing local computations in the time dimension [3]. This approach reduces communications among nodes and improves the network autonomy as compared to a centralized solution [1-3]. In the diffusion based adaptive networks, on the other hand, nodes communicate with all of their neighbors, and no cyclic path is required. The incremental adaptive networks in [1-6] assume ideal links between nodes. However, as we have shown in [14-18], the performance of incremental adaptive network changes considerably in the presence of noisy links. In fact, we show that

V.V. Das and P.M. El-Kafrawy (Eds.): SPIT 2012, LNICST 117, pp. 107–114, 2014.

- noisy links lead to a larger residual MSD, as expected.
- reducing the adaptation step size may actually increase the residual MSD.

In this work, we present other interesting results about the performance of incremental adaptive networks with noisy links. To this aim, we first define the concept of *cooperation gain* for incremental adaptive networks. Then, we calculate the cooperation gain for incremental adaptive networks with ideal links and noisy links. We observe that, when links are ideal, incremental adaptive networks always have a better steady-state performance than non-cooperative scheme, while in the presence of noisy links, depending on data and channel statistics, non-cooperative scheme may have better performance. We also present simulation results to support the derived expressions.

Notation: Bold uppercase letters denote matrices, whereas bold lowercase letters stand for vectors. Symbol * is used for both complex conjugation for scalars and Hermitian transpose for matrices. $\|\mathbf{x}\|_{\Sigma}^{2}=\mathbf{x}^{*}\Sigma\mathbf{x}$ denotes weighted norm for a column vector $\mathbf{x}$. $\mathbf{I}_M$ is $M\times M$ identity matrix and 1_N is $N\times 1$ vector with unit entries.

2 Incremental LMS Adaptive Network

Let's denote by $\mathcal{N}=\{1,...,N\}$ a set of nodes that communicate according to a given network topology. At time i, each node k has access to scalar measurement $d_k(i)$ and $1\times M$ regression vector u_k that are related via

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

where $M\times 1$ vector $\mathbf{w}^{o}\in R^{M}$ is an unknown parameter and $v_k(i)$ is the observation noise term with variance $\sigma_{v,k}^{2}$. The objective of the network is to estimate $\mathbf{w}^{o}$ from measurements collected at N nodes. The collected data at all nodes are

$$\mathbf{U}=\begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_N\end{bmatrix}^{T}(N\times M),\quad \mathbf{d}=\begin{bmatrix}d_1 & d_2 & \cdots & d_N\end{bmatrix}^{T}(N\times 1) \tag{2}$$

It must be noted that $\mathbf{w}^{o}$ is the solution of the following optimization problem

$$\arg\min_{\mathbf{w}} J(\mathbf{w}) \text{ where } J(\mathbf{w})=E\{\|\mathbf{d}-\mathbf{U}\mathbf{w}\|^{2}\} \tag{3}$$

The optimal solution of (3), is given by normal equations [1]

$$\mathbf{w}^{o}=\mathbf{R}_{u}^{-1}\mathbf{R}_{du} \tag{4}$$

where

$$\mathbf{R}_{du}=E\left\{\mathbf{U}^{*}\mathbf{d}\right\}, \text{and } \mathbf{R}_{u}=E\left\{\mathbf{U}^{*}\mathbf{U}\right\} \tag{5}$$

In order to use (4), each node must have access to the global statistical information $\left\{\mathbf{R}_u,\mathbf{R}_{du}\right\}$ which in many applications is not available. To address this issue, the incremental LMS adaptive network is proposed in [3]. The update equation in the ILMS algorithm is given by

$$\hat{\mathbf{w}}_k^{(i)} = \hat{\mathbf{w}}_{r,k}^{(i)} + \mu_k \mathbf{u}_{k,i}^* (d_k(i) - \mathbf{u}_{k,i} \hat{\mathbf{w}}_{r,k}^{(i)}) \tag{6}$$

where $\hat{\mathbf{w}}_k^{(i)}$ denotes the local estimate of $\mathbf{w}^o$ at node k at time i, μ_k is the step size parameter and $\hat{\mathbf{w}}_{r,k}^{(i)}$ is the received local estimate which is given by

$$\hat{\mathbf{w}}_{r,k}^{(i)} = \begin{cases} \hat{\mathbf{w}}_{k-1}^{(i)} & \text{ideal links} \\ \hat{\mathbf{w}}_{k-1}^{(i)} + \mathbf{q}_{k,i} & \text{noisy links} \end{cases} \tag{7}$$

where $\mathbf{q}_{k,i} \in R^{M\times 1}$, is the (time-realization) of channel the noise term between sensor k and $k-1$ which is assumed to have zero mean and covariance matrix $\mathbf{Q}_k = E\{\mathbf{q}_k \mathbf{q}_k^*\}$. Replacing (7) in (6), the update equation of ILMS algorithm with the noisy links changes to

$$\hat{\mathbf{w}}_k^{(i)} = \hat{\mathbf{w}}_{k-1}^{(i)} + \mathbf{q}_{k,i} + \mu_k \mathbf{u}_{k,i}^* (d_k(i) - \mathbf{u}_{k,i} (\hat{\mathbf{w}}_{k-1}^{(i)} + \mathbf{q}_{k,i})) \tag{8}$$

As we have shown in [14, 15], noisy links lead to a larger residual MSE, and also, reducing the adaptation step size may actually increase the residual MSE.

2.1 Steady-State Performance

A good measure of the adaptive network performance is the MSD which for each node k is defined as follows

$$\eta_k = E\{\|\tilde{\mathbf{w}}_{k-1}^{(\infty)}\|_{\mathbf{I}}^2\} \tag{9}$$

where

$$\tilde{\mathbf{w}}_{k-1}^{(i)} = \mathbf{w}^o - \hat{\mathbf{w}}_{k-1}^{(i)} \tag{10}$$

In [14, 15], the mean-square performance of ILMS adaptive network with noisy links has been investigated using the space-time energy conservation argument that was initially proposed in [2]. The analysis relies on the following assumptions data

(A.1) The regression data $\mathbf{u}_{k,i}$ are temporally and spatially independent and identically distributed (i.i.d.) circular white Gaussian random variables with zero mean and diagonal covariance matrix $\lambda \mathbf{I}_M$.

(A.2) $\mathbf{u}_{k,i}$ and $v_k(j)$ are independent of each other for all i and j.

In [14, 15], a complex closed-form expression for MSD has been derived. However, if we consider the following assumption

$$\mu_k = \mu, \;\; \mathbf{R}_{u,k} = \lambda \mathbf{I}, \;\; \mathbf{Q}_k = \sigma_{c,k}^2 \mathbf{I}$$

and also assuming small μ, we can approximate η_k as

$$\eta_k^{\text{inc,noisy}} = \frac{M}{2\mu\lambda N} \sum_{k=1}^{N} \left(\mu^2 \sigma_{v,k}^2 \lambda + \sigma_{c,k}^2 (1 - 2\mu\lambda) \right) \tag{11}$$

Obviously, the steady-state MSD for an ILMS adaptive network with ideal links can be extracted from (11) for $\sigma_{c,k}^2 = 0$ as

$$\eta_k^{\text{inc,ideal}} = \frac{M\mu}{2N} \sum_{k=1}^{N} \sigma_{v,k}^2 \tag{12}$$

Note that (12) reveals an equalization effect on the MSD throughout the network, i.e. for $k, \ell \in \mathcal{N}$, we have $\eta_k = \eta_\ell$; thus, the average MSD is given as

$$\bar{\eta}^{\text{inc,noisy}} = \frac{1}{N}\sum_{k=1}^{N}\eta_k^{\text{inc,noisy}} = \frac{M}{2\mu\lambda N}\sum_{k=1}^{N}\left(\mu^2\sigma_{v,k}^2\lambda + \sigma_{c,k}^2(1-2\mu\lambda)\right) \tag{13}$$

Similarly, the average MSD over all nodes for the ILMS with ideal links becomes

$$\bar{\eta}^{\text{inc,ideal}} = \frac{1}{N}\sum_{k=1}^{N}\eta_k^{\text{inc,ideal}} = \frac{M\mu}{2N}\sum_{k=1}^{N}\sigma_{v,k}^2 \tag{14}$$

2.2 Non-cooperation Scheme

It is noticeable that each node in the network can individually estimate $\mathbf{w}^o$ using its own data $\{d_k, \mathbf{u}_k\}$ and its previous time local estimate $\hat{\mathbf{w}}_{\text{nc},k}^{(i-1)}$ via

$$\hat{\mathbf{w}}_{\text{nc},k}^{(i)} = \hat{\mathbf{w}}_{\text{nc},k}^{(i-1)} + \mu_k \mathbf{u}_{k,i}^*(d_k(i) - \mathbf{u}_{k,i}\hat{\mathbf{w}}_{\text{nc},k}^{(i-1)}) \tag{15}$$

The steady-state MSD for non-cooperative scheme is given by [19]

$$\eta_k^{\text{nc}} = \frac{M\mu\sigma_{v,k}^2}{2} \tag{16}$$

where in this case (non-cooperative scheme), the steady-state MSD is given by

$$\eta_k^{\text{nc}} = \lim_{i\to\infty} E\{\|\tilde{\mathbf{w}}_k^{(i)}\|_I^2\} \tag{17}$$

The average MSD over all nodes of network is given as

$$\bar{\eta}^{\text{nc}} = \frac{M\mu\sum_{k=1}^{N}\sigma_{v,k}^2}{2N} \tag{18}$$

3 Cooperation Gain

In this section we compare the steady-state MSD performance of the ILMS algorithm (6) with a non-cooperative scheme (15). It must be noted that to compare the MSD of non-cooperative scheme with incremental LMS algorithm, we need to replace μ with μN in (18) . This is because the incremental algorithm uses N iterations for every measurement time. So we have

$$\bar{\eta}^{\text{nc}} = \frac{M\mu\left(\sum_{k=1}^{N}\sigma_{v,k}^2\right)}{2} \tag{19}$$

Now, to define the cooperation gain for incremental LMS algorithm, consider a network composed of $N \geq 2$ nodes with a space–time data $\{\mathbf{d}, \mathbf{U}\}$ satisfying the model (1) and the assumptions (A1)-(A3). Let's denote by $\bar{\eta}^{\text{nc}}$, $\bar{\eta}^{\text{inc,ideal}}$ and $\bar{\eta}^{\text{inc,noisy}}$, the average steady-state MSD provided by a non-cooperative scheme, the ILMS algorithm with ideal links and the ILMS algorithm with noisy links respectively. Thus, we can define the cooperation gain for ILMS algorithm with ideal links as

$$\mathcal{G}^{\text{inc,ideal}} = \frac{\bar{\eta}^{\text{nc}}}{\bar{\eta}^{\text{inc,ideal}}} \tag{20}$$

Replacing (19) and (14) in (20) we obtain

$$\mathcal{G}^{\text{inc,ideal}} = N \tag{21}$$

We can conclude from (21) that for all values of μ, the ILMS adaptive network with ideal links has better MSD performance than a non-cooperative scheme, or in formal terms

$$\mathcal{G}^{\text{inc,ideal}} > 1 \tag{22}$$

In addition, the cooperation gain is proportional to the number of nodes N and increasing the number of nodes increases the cooperation gain $\mathcal{G}^{\text{inc,ideal}}$. Similarly, the cooperation gain for the ILMS algorithm with noisy links can be defined as

$$\mathcal{G}^{\text{inc,noisy}} = \frac{\bar{\eta}^{\text{nc}}}{\bar{\eta}^{\text{inc,noisy}}} \tag{23}$$

Replacing (19) and (13) in (23) we obtain

$$\mathcal{G}^{\text{inc,noisy}} = \frac{\mu^2 \lambda N \sum_{k=1}^{N} \sigma_{v,k}^2}{\mu^2 \lambda \sum_{k=1}^{N} \sigma_{v,k}^2 + (1-2\mu\lambda) \sum_{k=1}^{N} \sigma_{c,k}^2} \tag{24}$$

We can conclude from (24) that in the presence of noisy links, the cooperation gain is not always bigger than 1. We have

$$0 < \mathcal{G}^{\text{inc,noisy}} < N \tag{25}$$

The above equation indicates that for some values of data and channel statistics we may have $\mathcal{G}^{\text{inc,noisy}} < 1$. In fact, the required condition for the ILMS algorithm to outperform the non-cooperative scheme is

$$\mathcal{G}^{\text{inc,noisy}} > 1 \tag{26}$$

or equivalently

$$\mathcal{G}^{\text{inc,noisy}} = \frac{\mu^2 \lambda N \sum_{k=1}^{N} \sigma_{v,k}^2}{\mu^2 \lambda \sum_{k=1}^{N} \sigma_{v,k}^2 + (1-2\mu\lambda) \sum_{k=1}^{N} \sigma_{c,k}^2} > 1 \tag{27}$$

The above equation is a quadratic equation in μ which can be rewritten as

$$a\mu^2 + b\mu + c > 0 \tag{28}$$

where

$$a = (N-1)\lambda \left(\sum_{k=1}^{N} \sigma_{v,k}^2 \right), b = 2\lambda \left(\sum_{k=1}^{N} \sigma_{v,k}^2 \right), c = -\left(\sum_{k=1}^{N} \sigma_{c,k}^2 \right) \tag{29}$$

Now, two different cases are possible:

Case I: $\Delta = b^2 - 4ac < 0$: Since $a > 0$, in this case, for $\forall \mu \in \mathcal{D}$ we have $\mathcal{G}^{\text{inc,noisy}} > 1$ so cooperation yields better steady-state performance.

Case II: $\Delta = b^2 - 4ac > 0$: Let x_1 and x_2 be the roots of equation $a\mu^2 + b\mu + c = 0$. Since $x_1 x_2 = \frac{c}{a} < 0$, roots are of opposite sign. If we assume $x_1 < 0$ and $x_2 > 0$, the inequality (26) holds when

$$x_2 < \mu < \sup\{\mathcal{D}\} \tag{30}$$

Therefore, the above discussion reveals that under noisy communication, cooperation among nodes may not necessarily result in better performance.

4 Simulation Results

We consider a distributed network with $N = 20$ nodes, and choose $M = 4$, $\mathbf{w}^o = 1_M / \sqrt{M}$, $\sigma_{v,k}^2 = 10^{-1}$, and $\sigma_{c,k}^2 = 10^{-4}$. Moreover, we assume that the regressors data arise from independent Gaussian, where $\mathbf{R}_{u,k} = \mathbf{I}$. Fig. 2 shows $\overline{\eta}^{\mathrm{nc}}$, $\overline{\eta}^{\mathrm{inc,ideal}}$ and $\overline{\eta}^{\mathrm{inc,noisy}}$ as a function of step size parameter μ. As we can see, both $\overline{\eta}^{\mathrm{nc}}$ and $\overline{\eta}^{\mathrm{inc,ideal}}$ are monotonically increasing function of μ and $\overline{\eta}^{\mathrm{inc,ideal}} > \overline{\eta}^{\mathrm{nc}}$ for all μ. Moreover, for all μ, the difference between $\overline{\eta}^{\mathrm{nc}}$ and $\overline{\eta}^{\mathrm{inc,ideal}}$ is constant, so that the cooperation gain is constant $\mathcal{G}^{\mathrm{inc,ideal}} = N = 20$ (see Fig. 1).

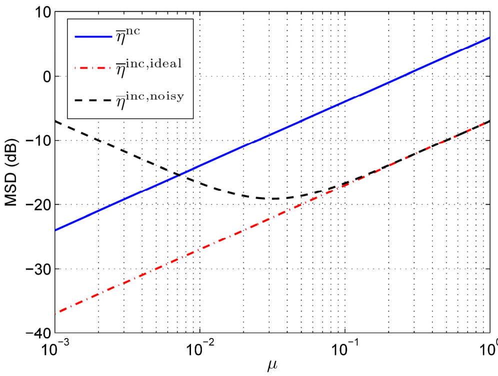

Fig. 1. $\overline{\eta}^{\mathrm{nc}}$, $\overline{\eta}^{\mathrm{inc,ideal}}$ and $\overline{\eta}^{\mathrm{inc,noisy}}$ as a function of step size parameter μ

On the other hand, in noisy links case, the steady-state MSD ($\overline{\eta}^{\mathrm{inc,noisy}}$) is not a monotonically increasing function of step size (see from Fig. 2). Specifically, for some values of μ, the non-cooperative scheme provides better performance $\mathcal{G}^{\mathrm{inc,noisy}} < 1$; while for some values of μ the ILMS algorithm has better performance ($\mathcal{G}^{\mathrm{inc,noisy}} > 1$). Fig. 2 also shows $\overline{\eta}^{\mathrm{inc,noisy}}$ in terms of μ.

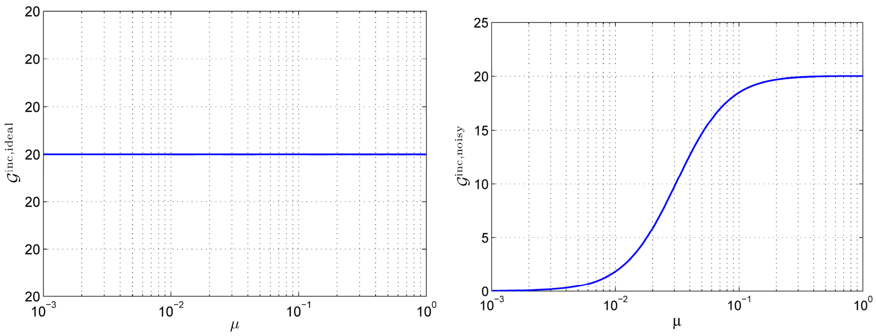

Fig. 2. $\overline{\eta}^{\text{inc,ideal}}$ versus μ (left) and $\overline{\eta}^{\text{inc,noisy}}$ versus μ (right)

5 Conclusion

In this paper, we considered the performance of incremental LMS adaptive networks in the presence of noisy links. We first defined the concept of cooperation gain for incremental adaptive networks. Then we showed that when the communication links are ideal, the ILMS algorithm has better performance than the non-cooperative scheme for every step size value, or equivalently cooperation gain is always bigger than 1. On the other hand, in the presence of noisy links, cooperation gain is not a constant function of μ and depending on data and channel statistics, non-cooperative scheme may have better performance. Finally we presented simulation results to support the derived expressions.

References

1. Lopes, C.G., Sayed, A.H.: Incremental adaptive strategies over distributed networks. IEEE Transactions on Signal Processing 55(8), 4064–4077 (2007)
2. Sayed, A.H., Lopes, C.G.: Distributed recursive least-squares strategies over adaptive networks. In: Proc. Asilomar Conference on Signals, Systems and Computers, pp. 233–237 (2006)
3. Li, L., Chambers, J.A., Lopes, C.G., Sayed, A.H.: Distributed estimation over an adaptive incremental network based on the affine projection algorithm. IEEE Transactions on Signal Processing 58(1), 151–164 (2010)
4. Ram, S.S., Nedic, A., Veeravalli, V.V.: Stochastic incremental gradient descent for estimation in sensor networks. In: Proc. Asilomar Conference on Signals, Systems and Computers, pp. 582–586 (2007)
5. Lopes, C.G., Sayed, A.H.: Randomized incremental protocols over adaptive networks. In: Proc. IEEE ICASSP, Dallas, TX, pp. 3514–3517 (2010)
6. Cattivelli, F., Sayed, A.H.: Analysis of spatial and incremental LMS processing for distributed estimation. IEEE Transactions on Signal Processing 59(4), 1465–1480 (2011)
7. Lopes, C.G., Sayed, A.H.: Diffusion least-mean squares over adaptive networks: Formulation and performance analysis. IEEE Trans. on Signal Processing 56(7), 3122–3136 (2008)

8. Cattivelli, F.S., Lopes, C.G., Sayed, A.H.: Diffusion recursive least-squares for distributed estimation over adaptive networks. IEEE Trans. on Signal Processing 56(5), 1865–1877 (2008)
9. Cattivelli, F.S., Sayed, A.H.: Diffusion LMS strategies for distributed estimation. IEEE Transactions on Signal Processing 58(3), 1035–1048 (2010)
10. Stankovic, S.S., Stankovic, M.S., Stipanovic, D.M.: Decentralized parameter estimation by consensus based stochastic approximation. In: Proc. IEEE Conference on Decision and Control, New Orleans, pp. 1535–1540 (2007)
11. Shin, Y.-J., Sayed, A.H., Shen, X.: Adaptive models for gene networks. PLoS ONE 7(2), e31657 (2012), doi:10.1371/journal.pone.0031657
12. Cattivelli, F.S., Sayed, A.H.: Multilevel diffusion adaptive networks. In: Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP), Taipei, Taiwan (2009)
13. Schizas, I.D., Mateos, G., Giannakis, G.B.: Distributed LMS for consensus-based in-network adaptive processing. IEEE Transactions on Signal Processing 57(6), 2365–2382 (2009)
14. Khalili, A., Tinati, M.A., Rastegarnia, A.: Performance analysis of distributed incremental LMS algorithm with noisy links. Inter. Journal of distributed sensor networks 2011, 1–10 (2011)
15. Khalili, A., Tinati, M.A., Rastegarnia, A.: Steady-state analysis of incremental LMS adaptive networks with noisy links. IEEE Trans. Signal Processing 59(5), 2416–2421 (2012)
16. Khalili, A., Tinati, M.A., Rastegarnia, A.: Analysis of incremental RLS adaptive networks with noisy links. IEICE Electron. Express 8(9), 623–628 (2011)
17. Khalili, A., Tinati, M.A., Rastegarnia, A., Chambers, J.A.: Steady-state analysis of diffusion LMS adaptive networks with noisy links. IEEE Trans. Signal Processing 60(2), 974–979 (2012)
18. Khalili, A., Tinati, M.A., Rastegarnia, A., Chambers, J.A.: Transient analysis of diffusion least-mean squares adaptive networks with noisy channels. Wiley Int. Journal of Adaptive Control and Signal Processing 26(2), 171–180 (2012)
19. Sayed, A.H.: Fundamentals of Adaptive Filtering. John Wiley and Sons, New York (2003)