Block Diffusion Adaptation over Distributed Adaptive Networks under Imperfect Data Transmission

Sananda Kumar, Ajit Kumar Sahoo, Debiprasad Priyabrata Acharya
Electronics and Communication Department
National Institute of Technology
Rourkela, Odisha
India, 769008
Email: sananda4u@gmail.com, {ajitsahoo, dpacharya@nitrkl.ac.in}

***Abstract*—A distributed Block LMS estimate strategy is developed by appealing to collaboration techniques that exploits both space and time structures of data. In diffusion strategies, information are exchanged among the nodes, usually containing noisy links. The weight combination of the neighboring nodes play a crucial role in adaptation and tracking ability of the network. The paper investigates for general adaptive diffusion algorithm, in presence of various sources of imperfect information exchange, then moves on to investigate for BLMS diffusion method, which plays a crucial role in reducing the burden of communication by a factor of block length.**

I. Introduction

Distributed adaptive algorithms are incorporated on large geographic area, in order to reduce the energy consumption among the nodes, decrease load of computation as compared to centralized processing network. The sensors used in distributed adaptation have the capability of data acquisition, self computation and sharing of information in between the nodes of network. Various mode of cooperation among the nodes have been proposed like incremental, diffusion, probabilistic diffusion mode of cooperation. Being immune node and link failure [1] diffusion mode of cooperation is selected compared to incremental mode.

In diffusion mode of cooperation each node tries to estimate the global unknown parameter and in turn shares information of regressor, desired, estimated weight amongst its neighboring connected nodes [1]–[4]. In practice these sharing of data among the nodes are not idle in nature. For presence of link noises and quantization error among the links the shared data are perturbed which in turn hampers for proper estimation of parameter. In [5], the paper tries to study the effect of these link noises using LMS algorithm of diffusion mode of cooperation. In [6], it has implemented BLMS method for incremental mode of cooperation.

This paper tries to implement Block Least Mean Squares (BLMS) algorithm in diffusion mode of cooperation, for amount of data communication is reduced by a factor of block length and thus reduces the energy consumption among the nodes. One more advantage of BLMS is it averages out the result for block size of data thus shows a bit better result compared to LMS algorithm.

We use boldface letters to denote random variables and normal font for deterministic (non-random) quantities. Capital letters are used for matrices and small letters for vectors and scalars. All vectors are column vectors except for regression vectors, which are denoted by throughout. The superscript $(.)^T$ represents the transpose of a matrix or a vector. The notation $col\{...\}$ stands for a vector obtained by stacking the specified vectors. Similarly, we use $col\{...\}$ to denote the (block) diagonal matrix consisting of the specified vectors or matrices. The trace of a matrix is denoted by $Tr(.)$, expectation is denoted by $E(.)$ and $\otimes$ denotes Kronecker product.

II. Problem Formulation

The objective of the WSN is to estimate the parameter of the environment using the measured data. A connected network with N nodes is considered. Each node k collects a scalar measurement $\mathbf{y}_k(i)$ and $1 \times M$ regression data vector $\mathbf{x}_{k,i}$. Converting them to block format of both input regressor data and desired data we have

$$\mathbf{X}_{k,i} = col\left\{\mathbf{x}_{k,(i-1)L+1}, \mathbf{x}_{k,(i-1)L+2}, \ldots\ldots, \mathbf{x}_{k,iL}\right\} \quad (1)$$

$$\mathbf{y}_{k,i} = col\left\{\mathbf{y}_k\,(i{-}1)\,L{+}1, \mathbf{y}_k\,(i{-}2)\,L{+}2, \ldots\ldots, \mathbf{y}_k(iL)\right\} \quad (2)$$

with dimensions $L \times M$ and $L \times 1$ respectively over successive time instant $i \geq 0$ and for L denotes the block size. With the linear regression model the unknown parameter w^0 of $M \times 1$ across all the nodes in the network is given by relation

$$\mathbf{y}_{k,i} = \mathbf{X}_{k,i} w^0 + \mathbf{v}_{k,i} \quad (3)$$

where $\mathbf{v}_{k,i}$ is the associated model noise with zero mean and variance $\sigma^2_{v,k}$. The nodes in the network will estimate w^0 by solving the minimization problem as given below

$$\min_w \sum_{k=1}^{N} E\left|\mathbf{y}_{k,i} - \mathbf{X}_{k,i} w\right|^2 \quad (4)$$

Solving above problem with adaptation and learning property with immune to node link failures, diffusion mode of cooperation is selected.

III. Diffusion under perfect data exchange

The combined CTA and ATC strategies for diffusion network together is defined as

$$\mathbf{\Phi}_{k,i-1} = \sum_{l \in \aleph_k} a_{1,lk} \mathbf{w}_{l,i-1} \tag{5}$$

$$\mathbf{\Psi}_{k,i} = \mathbf{\Phi}_{k,i-1} + \mu'_k \sum_{l \in \aleph_k} c_{lk} . \mathbf{X}^*_{l,i} \left(\mathbf{y}_{l,i} - \mathbf{X}_{l,i} . \mathbf{\Phi}_{k,i-1} \right) \tag{6}$$

$$\mathbf{w}_{k,i} = \sum_{l \in \aleph_k} a_{2,lk} \mathbf{\Psi}_{l,i} \tag{7}$$

for μ'_k is small positive block step-size and $a_{1,lk}, c_{lk}, a_{2,lk}$ are non-negative entries of the $N \times N$ matrices with (A_1, C, A_2) respectively. The coefficients $a_{1,lk}, c_{lk}, a_{2,lk}$ are zero, wherever l is not connected the node k is $l \notin \aleph_k$, where $\aleph_k$ denotes the neighborhood of node k. For ATC implementation $A_1 = I$ taken and for CTA implementation $A_2 = I$ is taken and for noncooperation implementation all the coefficients are taken to be I.

The matrices (A_1, C, A_2) are either left or right stochastic i.e.

$$\mathrm{A}_1^T 1_N = 1_N, \mathrm{A}_2^T 1_N = 1_N, C 1_N = 1_N \tag{8}$$

Defining the error vectors

$$\tilde{\mathbf{\Phi}}_{k,i-1} = w^0 - \mathbf{\Phi}_{k,i-1} \tag{9}$$

$$\tilde{\mathbf{\Psi}}_{k,i} = w^0 - \mathbf{\Psi}_{k,i} \tag{10}$$

$$\tilde{\mathbf{w}}_{k,i} = w^0 - \mathbf{w}_{k,i} \tag{11}$$

Substituting the linear model (3) into the adaptation step (6), we get

$$\tilde{\mathbf{\Psi}}_{k,i} = \left(I_M - \mu'_k \mathbf{R}_{k,i} \right) \tilde{\mathbf{\Phi}}_{k,i-1} - \mu'_k \mathbf{s}_{k,i} \tag{12}$$

where $\mathbf{R}_{k,i}$, $\mathbf{s}_{k,i}$ and μ'_k are defined as

$$\mathbf{R}_{k,i} = \sum_{l \in \aleph_k} c_{lk} . \mathbf{X}^*_{l,i} . \mathbf{X}_{l,i} \tag{13}$$

$$\mathbf{s}_{k,i} = \sum_{l \in \aleph_k} c_{lk} . \mathbf{X}^*_{l,i} . \mathbf{v}_{l,i} \tag{14}$$

$$\mu'_k = \frac{1}{L} \mu_k \tag{15}$$

Globalising these above vectors and matrices we have

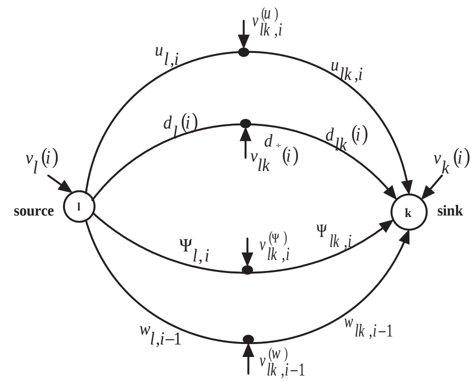


Fig. 1: Noisy Link in between l and k

$$\begin{aligned} &\tilde{\mathbf{\Psi}}_i \triangleq col \left\{ \tilde{\mathbf{\Psi}}_1, \tilde{\mathbf{\Psi}}_2, ..., \tilde{\mathbf{\Psi}}_N \right\} \\ &\tilde{\mathbf{w}}_i \triangleq col \left\{ \tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, ..., \tilde{\mathbf{w}}_N \right\} \\ &\tilde{\mathbf{\Phi}}_i \triangleq col \left\{ \tilde{\mathbf{\Phi}}_1, \tilde{\mathbf{\Phi}}_2, ..., \tilde{\mathbf{\Phi}}_N \right\} \\ &\mu' \triangleq diag \left\{ \mu'_1 I_M, \mu'_2 I_M, ..., \mu'_N I_M \right\} \\ &\mathbf{R}_i \triangleq diag \left\{ \mathbf{R}_{1,i}, \mathbf{R}_{2,i}, ..., \mathbf{R}_{N,i} \right\} \\ &\mathbf{s}_i \triangleq diag \left\{ \mathbf{s}_{1,i}, \mathbf{s}_{2,i}, ..., \mathbf{s}_{N,i} \right\} \end{aligned} \tag{16}$$

Finally the recursive error vector $\tilde{\mathbf{w}}_i$ is given by

$$\tilde{\mathbf{w}}_i = \mathcal{A}_2^T \left(I_{NM} - \mu' \mathbf{R}_i \right) \mathcal{A}_1^T \tilde{\mathbf{w}}_{i-1} - \mathcal{A}_2^T \mu' \mathcal{C}^T \mathbf{s}_i \tag{17}$$

where

$$\mathcal{A}_1 = \mathrm{A}_1 \otimes I_M, \mathcal{A}_2 = \mathrm{A}_2 \otimes I_M, \mathcal{C} = C \otimes I_M \tag{18}$$

IV. Diffusion under noisy information exchange

In diffusion mode of cooperation for estimation of the global parameter w^0 data of node k is shared with its neighboring nodes l for $l \in \aleph_k$ and vice-versa. In practice, these transmissions are subjected to additive noise and quantization error. The objective of this work is to analyze the aggregate effect of these perturbations on general BLMS diffusion strategies.

Considering exchange of information over links with additive noise as in Fig. 1, we have

$$\mathbf{w}_{lk,i-1} = \mathbf{w}_{l,i-1} + \mathbf{v}^{(w)}_{lk,i-1} \tag{19}$$

$$\mathbf{\Psi}_{lk,i} = \mathbf{\Psi}_{l,i} + \mathbf{v}^{(\Psi)}_{lk,i} \tag{20}$$

$$\mathbf{y}_{lk,i} = \mathbf{y}_{l,i} + \mathbf{v}^{(y)}_{lk,i} \tag{21}$$

$$\mathbf{X}_{lk,i} = \mathbf{X}_{l,i} + \mathbf{v}^{(X)}_{lk,i} \tag{22}$$

with dimensions of $\mathbf{v}^{(w)}_{lk,i-1}, \mathbf{v}^{(\psi)}_{lk,i}, \mathbf{v}^{(y)}_{lk,i}, \mathbf{v}^{(X)}_{lk,i}$ as $M \times 1, M \times N, L \times 1, L \times M$ respectively.

with the inclusion of noisy links the overall BLMS diffusion algorithm becomes

$$\mathbf{\Phi}_{k,i-1} = \sum_{l \in \aleph_k} a_{1,lk} \mathbf{w}_{lk,i-1} \tag{23}$$

$$\begin{aligned} &\mathbf{\Psi}_{k,i} = \mathbf{\Phi}_{k,i-1} + \\ &\mu_k' \sum_{r=0}^{L-1} \sum_{l \in \aleph_k} c_{lk}.\mathbf{X}^*_{lk,(i-1)L+r} \begin{pmatrix} \mathbf{y}_{lk,(i-1)L+r} \\ -\mathbf{X}_{lk,(i-1)L+r}.\mathbf{\Phi}_{k,i-1} \end{pmatrix} \end{aligned} \tag{24}$$

$$\mathbf{w}_{k,i} = \sum_{l \in \aleph_k} a_{2,lk} \mathbf{\Psi}_{lk,i} \tag{25}$$

Further with introduction of scalar zero-mean noise signal we have

$$\mathbf{v}_{lk,i} = \mathbf{v}_{l,i} + \mathbf{v}^{(y)}_{lk,i} - \mathbf{v}^{(X)}_{lk,i}.w^0 \tag{26}$$

for $l \in \aleph_k \backslash \{k\}$ it's variance is defined as

$$\sigma^2_{lk} \triangleq \sigma^2_{v,l} + \sigma^2_{v,lk} + w^{0*} \mathbf{R}^{(X)}_{X,lk} w^0 \tag{27}$$

The noisy linear model is now redefined with link noises as

$$\mathbf{y}_{lk,i} = \mathbf{X}_{lk,i} w^0 + \mathbf{v}_{lk,i} \tag{28}$$

with this the error weight recursion will be defined as

$$\tilde{\mathbf{\Psi}}_{k,i} = \left(I_M - \mu_k' \mathbf{R}'_{k,i}\right) \tilde{\mathbf{\Phi}}_{k,i-1} - \mu_k' \mathbf{z}_{k,i} \tag{29}$$

where $\mathbf{R}'_{k,i}$ is a $M \times M$ matrix with $\mathbf{z}_{k,i}$ as $M \times 1$ vector. given as

$$\mathbf{R}_{k,i} = \sum_{l \in \aleph_k} c_{lk}.\mathbf{X}^*_{lk,i}.\mathbf{X}_{lk,i} \tag{30}$$

$$\mathbf{z}_{k,i} = \sum_{l \in \aleph_k} c_{lk}.\mathbf{X}^*_{lk,i}.\mathbf{v}_{lk,i} \tag{31}$$

Globalizing the above error weight recursion with globalizing the regressor data and desired data and other parameters we have

$$\begin{aligned} &\mathbf{R}'_i = diag\left\{\mathbf{R}'_{1,i}, ..., \mathbf{R}'_{N,i}\right\} \\ &\mathbf{z}_i = col\left\{\mathbf{z}_{1,i}, ..., \mathbf{z}_{N,i}\right\} \\ &\mathbf{v}^{(w)}_i = col\left\{\mathbf{v}^{(w)}_{1,i}, ..., \mathbf{v}^{(w)}_{N,i}\right\} \\ &\mathbf{v}^{(\psi)}_i = col\left\{\mathbf{v}^{(\psi)}_{1,i}, ..., \mathbf{v}^{(\psi)}_{N,i}\right\} \end{aligned} \tag{32}$$

So, the final global error weight recursion for diffusion algorithm with BLMS method is defined as

$$\begin{aligned} \tilde{\mathbf{w}}_i = &A_2^T \left(I_{NM} - \mu' \mathbf{R}'_i\right) A_1^T \tilde{\mathbf{w}}_{i-1} - \\ &A_2^T \left(I_{NM} - \mu' \mathbf{R}'_i\right) \mathbf{v}^{(w)}_{i-1} - A_2^T \mu' z_i - \mathbf{v}^{(\psi)}_i \end{aligned} \tag{33}$$

V. Mean convergence

Finding the expectation of weight error recursion of (33) we have

$$\begin{aligned} E\left(\tilde{\mathbf{w}}_i\right) = &G.E\left(\tilde{\mathbf{w}}_{i-1}\right) - A_2^T \left(I_{NM} - \mu' R'\right).E\left(\mathbf{v}^{(w)}_{i-1}\right) \\ &- A_2^T \mu' E\left(\mathbf{z}_i\right) - E\left(\mathbf{v}^{(\psi)}_i\right) \end{aligned} \tag{34}$$

where

$$G = A_2^T \left(I_{NM} - \mu' R'\right) A_1^T \tag{35}$$

$$R' = E\left(\mathbf{R}'_i\right) = diag\left(R'_1, ..., R'_N\right) \tag{36}$$

$$R'_k \triangleq E\left(\mathbf{R}'_{k,i}\right) = \sum_{l \in \aleph_k} c_{lk} \left(R_{X,l} + R^{(X)}_{v,lk}\right) \tag{37}$$

for Gaussian noises with zero mean

$$E\left(\mathbf{v}^{(w)}_{i-1}\right) = E\left(\mathbf{v}^{(\psi)}_i\right) = 0 \tag{38}$$

So, the expression of (34) becomes

$$E\left(\tilde{\mathbf{w}}_i\right) = G.E\left(\tilde{\mathbf{w}}_{i-1}\right) - A_2^T \mu' E\left(\mathbf{z}_i\right) \tag{39}$$

with the presence of the driving force z, which will disappear from (39) if there were no noise present during the exchange of data among the nodes. For convergence to occur in mean the coefficient of the matrix G must be stable, i.e. $\rho(G) < 1$. The matrix A_1^T and A_2^T are right stochastic matrices, so the matrix G is stable whenever $(I_{NM} - \mu' R')$ is stable. With this the upper bound of step-sizes μ_k to guarantee the convergence of $E\left(\tilde{\mathbf{w}}_i\right)$ to steady state value must satisfy the condition

$$\mu_k' < \frac{2}{\lambda_{\max}\left(R'_k\right)} \tag{40}$$

$$\begin{aligned} \lambda_{\max}\left(R'_k\right) &= \lambda_{\max}\left(\sum_{l \in \aleph_k} c_{lk} \left(R_{X,l} + R^{(X)}_{v,lk}\right)\right) \\ &\leq \left(\sum_{l \in \aleph_k} c_{lk} \lambda_{\max}\left(R_{X,l} + R^{(X)}_{v,lk}\right)\right) \\ &\leq \max_{l \in \aleph_k} \lambda_{\max}\left(R_{X,l} + R^{(X)}_{v,lk}\right) \end{aligned} \tag{41}$$

Convergence condition for BLMS diffusion algorithm is to satisfy the condition of

$$\mu_k < \frac{2}{L.\max_{l \in \aleph_k}\left[\lambda_{\max}\left(R_{X,l} + R^{(X)}_{v,lk}\right)\right]} \tag{42}$$

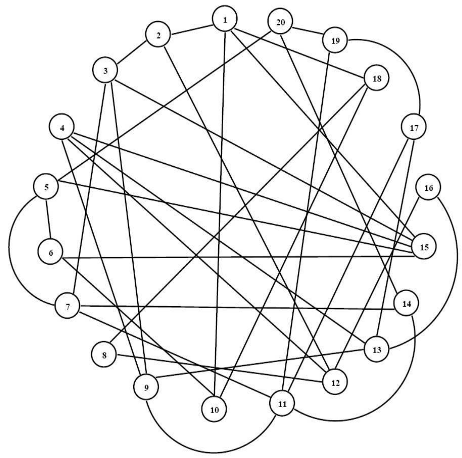
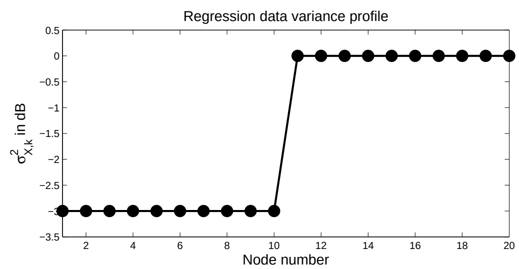
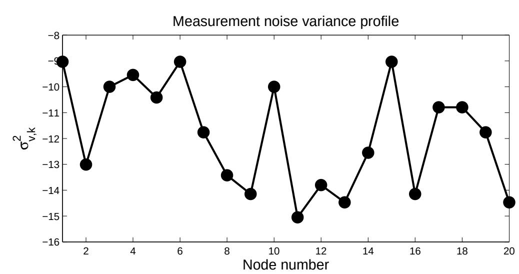


Fig. 2: Network structure of 20 nodes

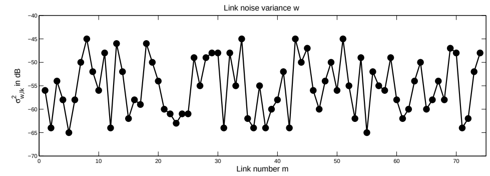


(a) regression variance

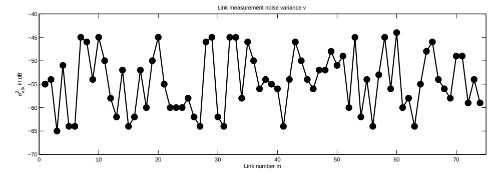


(b) measurement noise variance

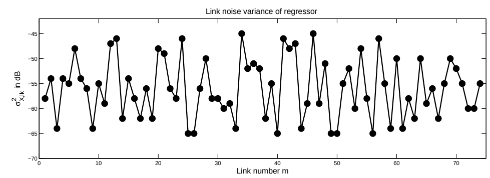
Fig. 3: variance of input regressor and measurement noise



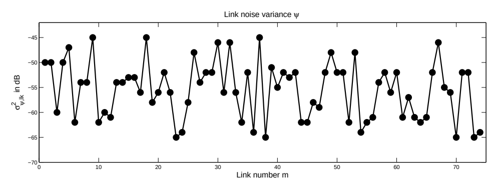
(a) $\sigma^2_{w,lk}$



(b) $\sigma^2_{v,lk}$



(c) $\sigma^2_{x,lk}$



(d) $\sigma^2_{y,lk}$

Fig. 4: Node link noises in dB

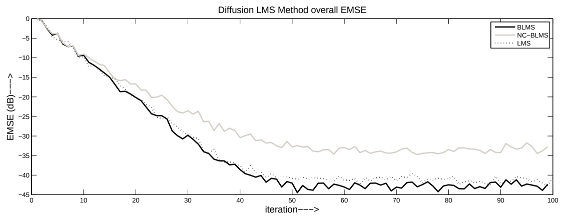
VI. Simulation and Results

A network consisting of 20 nodes as shown in Fig. 2 is taken for simulation study of BLMS diffusion algorithm. The size of unknown global vector w^0 is chosen to be 4 and the weights are considered to be normalized as equals to $[1/2, 1/2, 1/2, 1/2]$. The regressor $\mathbf{X}_{k,i}$ is chosen to be Gaussian and their corresponding variance $\sigma^2_{X,k}$ of nodes are as shown in Fig. 3a and the measurement noise variances $\sigma^2_{v,k}$ are as shown in Fig. 3b The block size is chosen to be $L = 16$, number of samples taken is 1600 and the step size is selected to be $\mu = 0.01$ for LMS case. Variances of various link noise present in between the node links are shown in Fig. 4

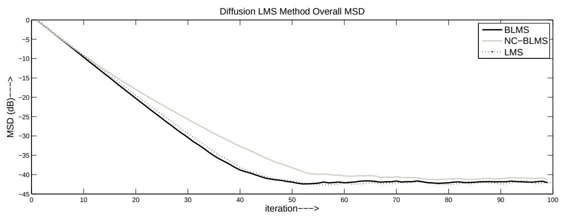
The BLMS diffusion algorithm is performed for 100 independent experiments and their Excess Mean Square Error (EMSE) and Mean square Deviation (MSD) is computed by taking average over all the experiments and all the nodes. Implementing diffusion algorithm all the coefficients A_1, C, A_2 are chosen in general as $A_1 = I$, C =metropolis and A_2 =relative degree for LMS as well as for BLMS. Similarly for noncooperation mode of diffusion adaptation all the coefficients are chosen as identity matrix (I). The simulation results of EMSE and MSD obtained are given in Fig. 5.

VII. Conclusion

From the simulation results of EMSE and MSD obtained in Fig. 5 it is clear that both LMS and BLMS have relatively same steady state performance for diffusion mode of cooperation. The BLMS noncooperation mode of diffusion algorithm is having lower steady state performance as opposed to BLMS diffusion mode despite of no communication in between the nodes i.e. remaining immune to node link noises and quantization error. The main benefit of using BLMS algorithm is decrease in cost of data communication by a factor of block length thus extending battery life of nodes as compared to LMS diffusion strategy.



(a) EMSE



(b) MSD

Fig. 5: Overall EMSE and MSD plots done taking mean over all nodes

REFERENCES

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

[2] F. S. Cattivelli and A. H. Sayed, "Diffusion lms strategies for distributed estimation," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1035–1048, 2010.

[3] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks." in *ICASSP (3)*, 2007, pp. 917–920.

[4] C. G. Lopes, "Diffusion adaptive networks with changing topologies." Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, 2008.

[5] X. Zhao, S.-Y. Tu, and A. H. Sayed, "Diffusion adaptation over networks under imperfect information exchange and non-stationary data," *Signal Processing, IEEE Transactions on*, vol. 60, no. 7, pp. 3460–3475, 2012.

[6] A. Khalili, M. A. Tinati, and A. Rastegarnia, "An incremental block lms algorithm for distributed adaptive estimation," in *Communication Systems (ICCS), 2010 IEEE International Conference on*. IEEE, 2010, pp. 493–496.

[7] A. H. Sayed, *Fundamentals of adaptive filtering*. John Wiley & Sons, 2003.

[8] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2008.