A Robust Distributed Estimation Algorithm under Alpha-Stable Noise Condition

Azam Khalili[1], Amir Rastegarnia[2], and Md Kafiul Islam[3]
[1] *Department of Electrical Engineering, Malayer University, a.khalili@ieee.org*
[2] *Department of Electrical Engineering, Malayer University, a_rastegar@ieee.org*
[3] *Department of Electrical and Computer Engineering, National University of Singapore, kafiul_islam@nus.edu.sg*
Corresponding author: Azam Khalili

***Abstract*— Robust adaptive estimation of unknown parameter has been an important issue in recent years for reliable operation in the distributed networks. The conventional adaptive estimation algorithms that rely on mean square error (MSE) criterion exhibit good performance in the presence of Gaussian noise, but their performance drastically decreases under impulsive noise. In this paper, we propose a robust adaptive estimation algorithm for networks with cyclic cooperation. We model the impulsive noise as the realization of alpha-stable distribution. Here, we move beyond MSE criterion and define the estimation problem in terms of a modified cost function which exploits higher order moments of the error. To derive a distributed and adaptive solution, we first recast the problem as an equivalent form amenable to distributed implementation. Then, we resort to the steepest-descent and statistical approximation to obtain the proposed algorithm. We present some simulations results which reveal the superior performance of the proposed algorithm than the incremental least mean square (ILMS) algorithm in impulsive noise environments.**

***Index Terms*—** adaptive networks; distributed estimation; impulsive noise.

I. INTRODUCTION

In recent years, great attention has been devoted to distributed estimation problem where the objective is to estimate an unknown parameter using data collected by the nodes [1]. In general, distributed estimation problem can be typically solved by either a centralized approach or a decentralized approach (see [2] and references therein). In many applications, however, sensors need to perform estimation task in a environment without any statistical information about the underlying processes of interest. This issue motivated the development of distributed adaptive estimation algorithms which are also known as adaptive networks [3,4]. So far, different distributed adaptive

Manuscript received 14-May-2015 and revised 28-Aug.-2015 *P- ISSN: 2322-4088*
Accepted on 6- Oct.-2015 *E- ISSN: 2322-3936*

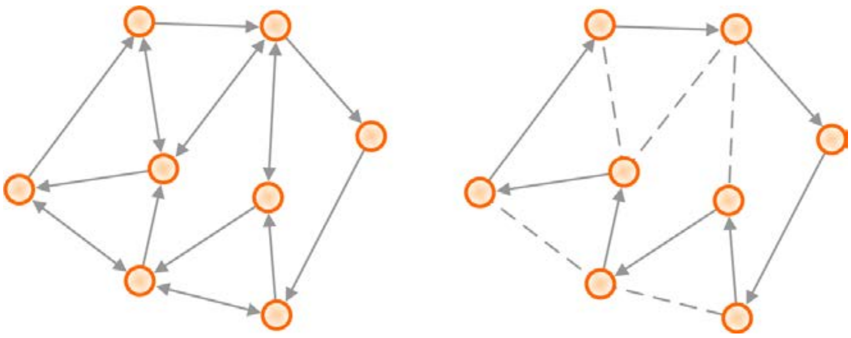

Fig. 1. Different cooperation modes for adaptive networks: Incremental (left) and diffusion (right).

estimation algorithms have been introduced in the literature. These algorithms can be categorized, in general, based on the mode of cooperation between nodes as incremental networks [4-10] and diffusion networks [11-15]. The incremental LMS (ILMS) algorithm [4,26], distributed recursive least square (DRLS) algorithm [5] and distributed affine projection algorithm [6] are examples of distributed adaptive estimation algorithms that use incremental cooperation between nodes.

These schemes inherently require a Hamiltonian cycle through which signal estimates are sequentially circulated from sensor to sensor (See Fig. 1). On the other hand, in diffusion based schemes, each node updates its estimate using all available estimates from its neighbors, as well as data and its own past estimate [11-15].

The conventional gradient-based distributed adaptive estimation algorithms exhibit good performance in the presence of Gaussian noise but their performance drastically decreases in impulsive noise environments [16]. Robust LMS algorithm have been reported in [17,18] which rely on Wilcoxon norm. This class of algorithms are difficult to analyze, and therefore it is common to resort to different methods and assumptions [16]. In [19], it has been shown that the error saturation nonlinearity-based LMS algorithm provides good performance in the presence of impulsive noise. However, the robust adaptive algorithms discussed so far are not inherently distributed in nature.

In this paper we consider the problem of distributed estimation with incremental LMS adaptive network in the presence of impulsive noise. Our aim is to develop an incremental algorithm which is robust to measurements that are corrupted by impulsive noise. We model the impulsive noise as the realizations of alpha-stable distribution [20,21]. Unlike the ILMS algorithm which relies on the MSE cost function, in the proposed algorithm we use a modified cost function which exploits higher order moments of the error. To derive a distributed and adaptive solution, we recast the problem as an equivalent form amenable to distributed implementation. Then, we resort to the steepest-descent and statistical approximations to finally obtain a fully distributed adaptive estimation algorithm. We compare the performance of the proposed algorithm with some available algorithms. Numerical examples shows that the proposed algorithms outperform existing ILMS algorithm.

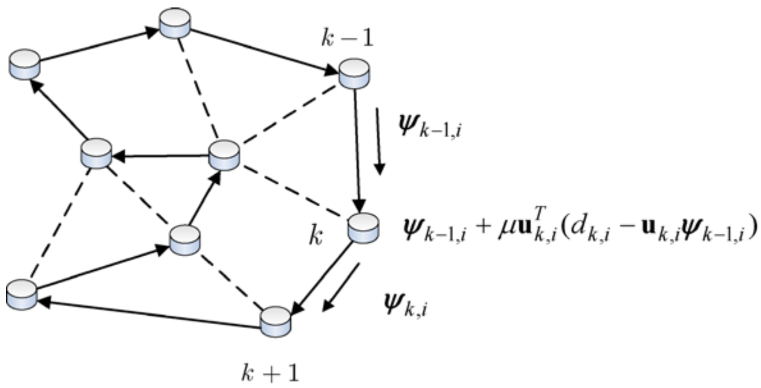

Fig. 2. The structure of incremental LMS algorithm.

Notation Throughout the paper, We adopt small boldface letters for vectors and bold capital letters for matices. The symbol * denotes conjugation for scalars and Hermitian transpose for matrices. We also use $(\cdot)^T$ to denote the transpose of a vector or a matrix and $E\{\cdot\}$ for the statistical expectation.

II. DISTRIBUTED ESTIMATION

Consider a network with N nodes that collaborate to estimate an unknown vector $\mathbf{w}^o \in R^{M\times 1}$, from streaming data. At time instant i each node k has access to measurement data as $\{d_{k,i}, \mathbf{u}_{k,i}\}$ where each $d_{k,i}$ is a scalar measurement and each $\mathbf{u}_{k,i}$ is a $1\times M$ row regression vector that satisfies the following linear model:

$$d_{k,i} = \mathbf{u}_{k,i}\mathbf{w}^o + v_{k,i} \tag{1}$$

where $v_{k,i}$ denotes the observation (measurement) noise. The linear regression model in (1) appears in many practical applications such as spectrum sensing, target tracking, and source localization [3]. Now we can rewrite the estimation on unknown parameter as the following optimization problem:

$$\mathbf{w}^o = \arg\min_{\mathbf{w}} J(\mathbf{w}), \quad \text{where} \quad J(\mathbf{w}) = \sum_{k=1}^{N} E\{\| \mathbf{d}_{k,i} - \mathbf{u}_{k,i}\mathbf{w} \|^2\} \tag{2}$$

The optimal solution $\mathbf{w}^o$ of the unconstrained optimization problem (2) satisfies the following normal equation [22]:

$$\mathbf{r}_{du} = \mathbf{R}_u \mathbf{w}^o \tag{3}$$

where

$$\mathbf{R}_u = \sum_{k=1}^{N} E\{\mathbf{u}_{k,i}^T \mathbf{u}_{k,i}\}, \qquad \mathbf{r}_{du} = \sum_{k=1}^{N} E\{d_{k,i}\mathbf{u}_{k,i}^T\} \tag{4}$$

Starting with gradient-descent implementation and applying instantaneous approximations $\mathbf{R}_{u,k} \approx \mathbf{u}_{k,i}^T \mathbf{u}_{k,i}$ and $\mathbf{r}_{du,k} \approx d_{k,i}\mathbf{u}_{k,i}^T$, the ILMS algorithm with a cyclic estimation structure can be derived as [4]

$$\boldsymbol{\psi}_{k,i} = \boldsymbol{\psi}_{k-1,i} + \mu_k \mathbf{u}_{k,i}^T [d_{k,i} - \mathbf{u}_{k,i}\boldsymbol{\psi}_{k-1,i}] \tag{5}$$

where $\boldsymbol{\psi}_{k,i}$ represents the local estimate at the node k and time i. A schematic for the ILMS algorithm is shown in Fig. 2.

The ILMS algorithm works as follows: At time i, node k utilizes the local data $\{d_{k,i}, \mathbf{u}_{k,i}\}$ and $\boldsymbol{\psi}_{k-1,i}$ received from the node $k-1$ to calculate its local estimate, i.e. $\boldsymbol{\psi}_{k,i}$. Note that at the beginning of every iteration i, due to cyclic cooperation, node k uses $\boldsymbol{\psi}_{N,i-1}$ to update its local estimate $\boldsymbol{\psi}_{1,i}$. Although adaptive networks that rely on second order statistics (like MSE) exhibit good performance in the presence of Gaussian noise, however, their performance drastically decreases for non-Gaussian data such as impulsive noise environments [23, 24]. To model the impulsive noise environments, in this paper we assume that the measurement noise term in (1) follows the alpha-stable distribution. The characteristic function of alpha-stable process [20] is described as:

$$\phi(t) = \exp\{j\delta t - \gamma |t|^{\alpha} \left[1 + j\beta \mathrm{sgn}(t) f(t,\alpha)\right]\} \tag{6}$$

where sgn denotes the sign function, and

$$f(t,\alpha) = \begin{cases} \tan\left(\dfrac{\alpha\pi}{2}\right) & \alpha \neq 1 \\ \dfrac{2}{\pi}\log|t| & \alpha = 1 \end{cases} \tag{7}$$

Moreover, in (7) $\alpha \in (0,2]$ is the characteristic exponent and describes the tail of the distribution. In addition, $-\infty < \delta < \infty$ is the location parameter of the distribution, and $\beta \in [-1,1]$ is the symmetry parameter. $\gamma > 0$ is the dispersion, which plays a role similar to the variance of the Gaussian distribution. The distribution is symmetric around its location parameter δ when $\beta = 0$. Throughout this paper, we assume that the alpha stable noise is symmetric $\beta = 0$ and the location parameter $\delta = 0$. Fig. 3 show the symmetric and skewed alpha-stable densities for different values of parameters. In the next section we present our proposed algorithm which is robust to impulsive noise.

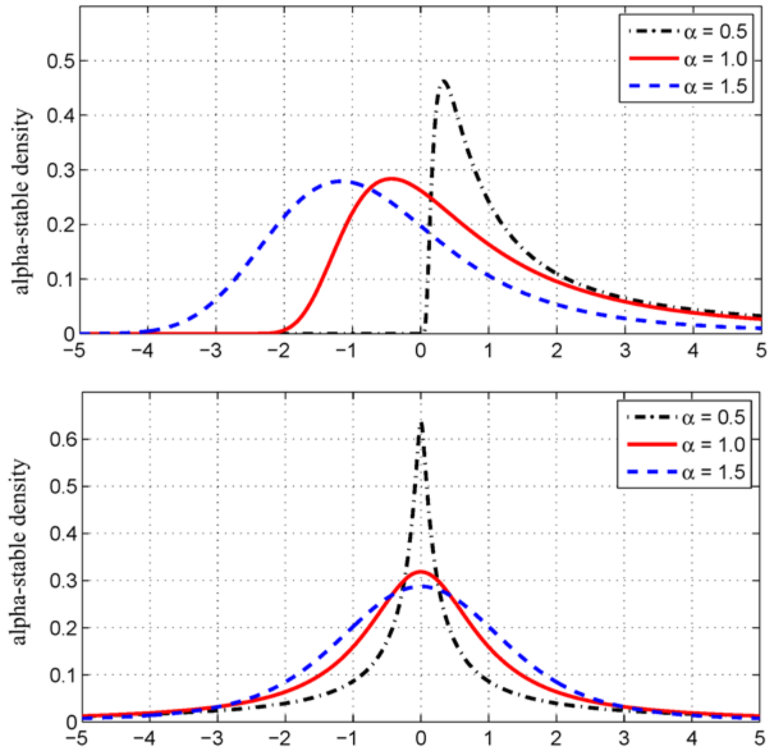

Fig. 3. Symmetric alpha-stable densities, for $\beta = 0$, $\gamma = 1$, $\delta = 0$ (top), and skewed alpha-stable densities, for $\beta = 0.5$, $\gamma = 1$, $\delta = 0$ (bottom).

III. Proposed Algorithm

Let us consider a network with N nodes where at any time i, node k measures data $\{d_{k,i}, \mathbf{u}_{k,i}\}$ that satisfy linear model of the (1). To move beyond mean squared error and exploit higher order moments of the error, we define the following modified cost function as

$$J(\mathbf{w}) = \sum_{k=1}^{N} E\{|d_k - \mathbf{u}_k \mathbf{w}|^p\} \tag{8}$$

where $p > 0$. To derive the proposed algorithm, which is a distributed, adaptive solution for the optimization problem in (8) we firstly use iterative steepest-descent method to find $\mathbf{w}^o$ as

$$\mathbf{w}_i = \mathbf{w}_{i-1} - \mu[\nabla_w J(\mathbf{w}_{i-1})]^T \tag{9}$$

where $\mu > 0$ is a step-size parameter and $\mathbf{w}_i$ is an estimate for $\mathbf{w}^o$ at iteration i, and $\nabla_{\mathbf{w}} J$ denotes the gradient of $J(\mathbf{w})$ with respect to $\mathbf{w}$ which can be obtained as

$$\nabla_{\mathbf{w}} J(\mathbf{w}) = -\sum_{k=1}^{N} E\{|\zeta_{k,i}|^{p-2}\zeta_{k,i}\mathbf{u}_{k,i}\} \tag{10}$$

where the error signal $\zeta_{k,i}$ is defined as

$$\zeta_{k,i} = d_{k,i} - \mathbf{u}_{k,i}\mathbf{w}_{i-1} \tag{11}$$

Substituting (10) into (9) leads to

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu\sum_{k=1}^{N} E\{|\zeta_{k,i}|^{p-2}\zeta_{k,i}\mathbf{u}_{k,i}^T\} \tag{12}$$

Note that each iteration step in (12) involves sum of N terms. Equivalently, we can obtain the same result by splitting the update into N separate steps whereby each step adds one term at $E\{|\zeta_{k,i}|^{p-2}\zeta_{k,i}\mathbf{u}_{k,i}^T\}$ to get an intermediate value $\boldsymbol{\psi}_{k,i}$. Thus, we can rewrite (12) as the following form

$$\begin{cases} \boldsymbol{\psi}_{0,i} \leftarrow \mathbf{w}_{i-1} \\ \boldsymbol{\psi}_{k,i} = \boldsymbol{\psi}_{k-1,i} + \mu E\{|\zeta_{k,i}|^{p-2}\zeta_{k,i}\mathbf{u}_{k,i}^T\} \\ \mathbf{w}_i \leftarrow \boldsymbol{\psi}_{N,i} \end{cases} \tag{13}$$

This implementation is not a distributed solution as it requires every node to access to the global information $\mathbf{w}_{i-1}$. A distribute solution can be obtained by replacing the $\mathbf{w}_{i-1}$ at each node by a local estimate $\boldsymbol{\psi}_{k,i}$ where the global error signal is replaced in it with local error signal $e_{k,i}$ where

$$e_{k,i} = d_{k,i} - \boldsymbol{\psi}_{k-1,i}\mathbf{u}_{k,i}^T \tag{14}$$

Using local error signal and local estimates a distributed (incremental) solution for (10) can be obtained as follows

$$\begin{cases} \boldsymbol{\psi}_{0,i} \leftarrow \mathbf{w}_{i-1} \\ \boldsymbol{\psi}_{k,i} = \boldsymbol{\psi}_{k-1,i} + \mu E\{|e_{k,i}|^{p-2}e_{k,i}\mathbf{u}_{k,i}^T\} \\ \mathbf{w}_i \leftarrow \boldsymbol{\psi}_{N,i} \end{cases} \tag{15}$$

Now note that the incremental algorithm (15) requires knowledge of the statistical moment $E\{|e_{k,i}|^{p-2}e_{k,i}\mathbf{u}_{k,i}^T\}$. An adaptive implementation of (15) can be obtained by replacing the required moment by its instantaneous approximations

$$E\{|e_{k,i}|^{p-2}e_{k,i}\mathbf{u}_{k,i}^T\} \approx |e_{k,i}|^{p-2}e_{k,i}\mathbf{u}_{k,i} \tag{16}$$

Using the above approximation leads to the following update equation for the proposed algorithm

Table I. Pseudo code for the proposed algorithm

```
Initialization: ψ_{1,1} = 0
For i = 1, 2, ⋯
    For k = 1, 2, ⋯, N
        Receive from ψ_{k-1,i} previous node.
        Update ψ_{k,i} as ψ_{k,i} = ψ_{k-1,i} + μ{|e_{k,i}|^{p-2} e_{k,i} u^T_{k,i}}
        Send ψ_{k,i} to the next node
    end
end
```

$$\begin{cases} \boldsymbol{\psi}_{0,i} \leftarrow \mathbf{w}_{i-1} \\ \boldsymbol{\psi}_{k,i} = \boldsymbol{\psi}_{k-1,i} + \mu\{\left|e_{k,i}\right|^{p-2} e_{k,i}\mathbf{u}_{k,i}^T\} \\ \mathbf{w}_i \leftarrow \boldsymbol{\psi}_{N,i} \end{cases} \tag{17}$$

Pseudo code for the proposed algorithm is shown in Table I.

Remark 1. Note that as it is discussed in [25] the LMP algorithm has similar degrees of complexity as LMS algorithm. Thus, in terms of complexity per iteration per node, it needs 2M+2 multiplications and $2M$ summations.

In the next section we attempt to show the performance of our algorithm.

IV. Simulation Results

In this section we present the simulation results to evaluate the performance of the proposed algorithm. To this end, we consider a distributed network with $N = 20$ nodes and assume $M = 6$.

We assume that $\mathbf{u}_{k,i}$'s are Gaussian regressors with $\mathrm{Tr}[R_{u,k}] = 1$. We also select $\mu = 0.002$ for both ILMS and proposed algorithm. The observation noise are be drawn from alpha-stable distribution with $\alpha = 1.2$, $\beta = 0$, $\gamma = 0$ and $\delta = 1$. Each curve is obtained by averaging over 100 independent experiments. We examine the network performance by the global average mean-square deviation (MSD) and excess mean square error (MESE) which are defined respectively as

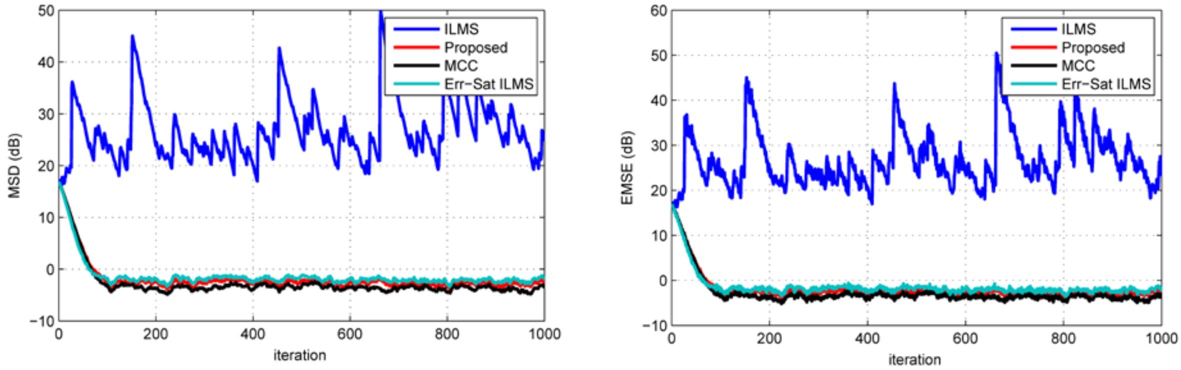

Fig. 4. The global average MSD and global average EMSE for different algorithms in the presence of impulsive noise.

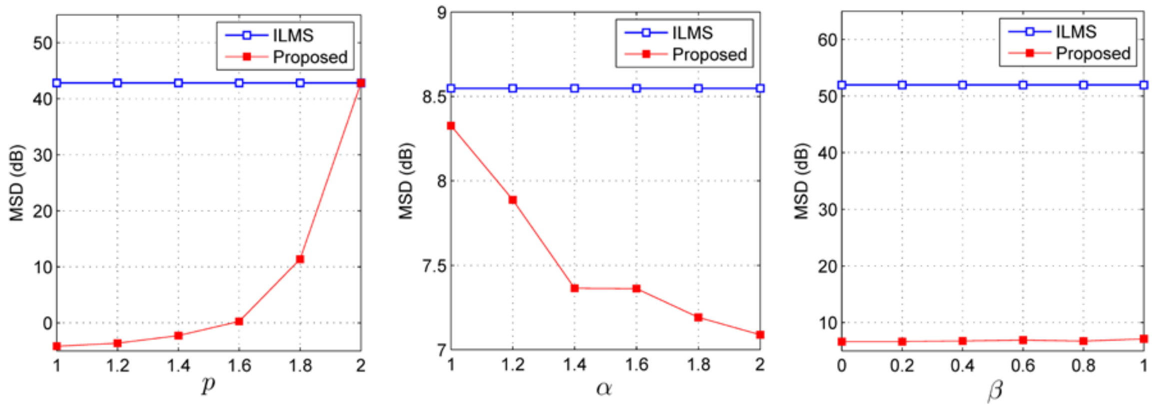

Fig. 5. Performance of the proposed algorithm for different values of noise parameters in comparison with the ILMS algorithm.

$$\mathrm{MSD} = \frac{1}{N}\sum_{k=1}^{N} E\left\{\left\| w^{o} - \psi_{k-1,i} \right\|^{2}\right\} \tag{18}$$

$$\mathrm{EMSE} = \frac{1}{N}\sum_{k=1}^{N} E\left\{\left| \mathbf{u}_{k,i}(\mathbf{w}^{o} - \boldsymbol{\psi}_{k-1,i}) \right|^{2}\right\} \tag{19}$$

Fig. 4 shows the global average MSD and global average EMSE for different algorithms including the ILMS algorithm, error saturation nonlinearity ILMS [23], incremental algorithm based on the Maximum Correntropy Criterion (MCC). As it is clear from Fig. 4, the performance of ILMS algorithm is strongly affected by the impulsive noise. Nevertheless, the proposed algorithm is able to overcome this problem. Specially, the steady-state error is clearly improved by means of our proposed method. Note that the error saturation nonlinearity ILMS and incremental MCC algorithm can also provide similar performance. The main drawback of error saturation nonlinearity ILMS is that it requires more operation than the proposed algorithm, while the MCC-based algorithm needs careful tuning of kernel parameter (See [24]). Fig. 5 shows the performance of the proposed algorithm for different values of noise parameters in comparison with the ILMS algorithm. We observe that, the proposed algorithm provides better performance than ILMS algorithm for different values of noise parameters. Note that as p increases, the performance of the proposed algorithm decrease and the proposed algorithm changes to the conventional ILMS algorithm for $p = 2$.

V. CONCLUSIONS

In this paper we proposed an incremental-based adaptive network for distributed estimation in alpha-stable noise environments. To alleviate the effect of impulsive noise, in the proposed algorithm we used modified cost function which considers higher order moments of the error. Numerical examples showed that the proposed algorithm outperforms existing online estimation scheme such as ILMS algorithm.

REFERENCES

[1] C. G. Lopes and A. H. Sayed, "Distributed processing over adaptive networks," *Proc. Adaptive Sensor Array Processing Workshop*, MIT Lincoln Lab., Lexington, MA, Jun. 2006.

[2] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, pp. 27-41, 2006.

[3] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Foundations and Trends in Machine Learning*, vol. 7, issue 4-5, pp. 311-801, NOW Publishers, Boston-Delft, 2014.

[4] A. H. Sayed, "Adaptive networks," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 460-497, April 2014.

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

[6] A. H. Sayed and C. G. Lopes, "Distributed recursive least-squares strategies over adaptive networks," *Proc. Asilomar Conf. Signals, Systems, Computers*, Monterey, CA, pp. 233-237, Oct. 2006.

[7] 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 Trans. on Signal Processing*, vol. 58, no. 1, pp. 151-164, Jan. 2010.

[8] 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.

[9] M. A. Tinati, a. Rastegarnia, and a. Khalili, "An incremental least-mean square algorithm with adaptive combiner," IET 3rd International Conference on Wireless, Mobile and Multimedia Networks (ICWMMN 2010), pp. 266–269, 2010.

[10] A. Khalili, A. Rastegarnia, and W. Bazzi, "A collaborative adaptive algorithm for the filtering of noncircular complex signals," in Telecommunications (IST), 7th International Symposium on, pp. 96–99, Sept. 2014.

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

[12] F. S. Cattivelli and A. H. Sayed, “Multilevel diffusion adaptive networks,” *Proc. ICASSP*, Taipei, Taiwan, April 2009.

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

[14] M. A. Tinati, A. Khalili, and A. Rastegarnia, “A diffusion block least-mean square algorithm for distributed adaptive estimation,” Wireless, Mobile and Multimedia Networks (ICWMNN 2010), IET 3rd International Conference on, pp.270-273, Sept. 2010.

[15] A. Khalili, R. G. Rahmati, and A. Rastegarnia, “An adaptive diffusion algorithm based on augmented QLMS for distributed filtering of hypercomplex processes,” *American Journal of Signal Processing*, vol. 5, no. 2A, pp. 1-8, 2015.

[16] T. Panigrahi, G. Panda and B. Mulgrew, “Error saturation nonlinearities for robust incremental LMS over wireless sensor networks,” *ACM Transaction on Sensor Network*, vol. 11, no. 2, pp. 1-20, 2014.

[17] B. Majhi, G. Panda, and B. Mulgrew, “Robust identification using new Wilcoxon least mean square algorithm”, *IEEE Electronics Letters*, vol. 45, no.6, pp. 334-335, 2009.

[18] S. R. Kim and A. Efron, “Adaptive robust impulsive noise filtering,” *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp.1855-1866, 1995.

[19] N. J. Bershad. “On error saturation nonlinearities for LMS adaptation in impulsive noise,” *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4526-4530, 2008.

[20] F Wen, “Diffusion least mean P-power algorithms for distributed estimation in alpha-stable noise environments,” *Electronics Letters*, vol. 49, no. 21, pp.1355-1356, 2013.

[21] M. Shao, and C. L. Nikias, “Signal processing with fractional lower order moments: stable processes and their applications,” *Proc. IEEE*, vol. 30, no. 3, pp. 986-1010, 1993.

[22] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York: Wiley, 2003.

[23] T. Panigrahi, G. Panda, B Mulgrew, and Babita Majhi, “Robust incremental LMS over wireless sensor network in impulsive noise,” *International Conference on Computational Intelligence and Communication Networks*, pp. 205-209, 2010.

[24] W. M. Bazzi, A. Rastegarnia, A. Khalili, “A robust diffusion adaptive network based on the maximum correntropy criterion,” *24th International conference on computer communication and networks (ICCCN)*, pp. 2015.

[25] Yegui Xiao , Y. Tadokoro , K. Shida, “Adaptive algorithm based on least mean p-power error criterion for Fourier analysis in additive noise,” *IEEE Transactions on Signal Processing*, vol. 47, no. 4, pp. 1172-1181, 1999.

[26] E. Mostafapour, C. Ghobadi, J. Nourinia and M. Chehel Amirani, “Tracking performance of incremental LMS algorithm over adaptive distributed sensor networks,” Journal of Communication Engineering, vol. 4, no. 1, pp. 55-66, Jan.-June 2015.