A New Leaky-LMS Algorithm with Analysis

Tajuddeen Gwadabe and Mohammad Salman
Department of Electrical and Electronics Engineering, Mevlana University, Turkey

Abstract: *Though the Leaky Least-Mean-Square (LLMS) algorithm mitigates the drifting problem of the LMS algorithm, its performance is similar to that of the LMS algorithm in terms of convergence rate. In this paper, we propose a new LLMS algorithm that has a better performance than the LLMS algorithm in terms of the convergence rate and at the same time solves the drifting problem in the LMS algorithm. This better performance is achieved by expressing the cost function in terms of a sum of exponentials at a negligible increase in the computational complexity. The convergence analysis of the proposed algorithm is presented. Also, a normalized version of the proposed algorithm is presented. The performance of the proposed algorithm is compared to those of the conventional LLMS algorithm and a Modified version of the Leaky Least-Mean-Square (MLLMS) algorithm in channel estimation and channel equalization settings in additive white Gaussian and white and correlated impulsive noise environments.*

Keywords: *LLMS algorithm, channel estimation, channel equalization, impulsive noise.*

Received September 24, 2014; accepted June 2, 2015

1. Introduction

The Least-Mean-Square (LMS) algorithm [9] is one of the well-known adaptive algorithms due to its simplicity and ease of analysis. One of the main drawbacks of the LMS algorithm is the drifting problem [9, 17]. The drifting problem is a situation where by the weight update of the LMS algorithm diverges as a result of inadequate information in the input sequence [14]. For example, inadequacy of excitation in the input sequence can result in unbounded parameter estimates [17]. This behaviour can cause numerical problems due to overflow as well as degraded performance as a consequence of possibly unbounded prediction error.

Introducing a leakage term in the LMS algorithm stabilizes the system. It can be also seen as a situation where the LMS algorithm generates unbounded parameter estimates for a bounded input sequence [4]. The Leaky Least-Mean-Square (LLMS) algorithm is an improved version of the LMS algorithm [9, 13]. It was proposed to mitigate the drifting problem in LMS algorithm. LLMS-type algorithms have been applied to several application areas [2, 15, 16, 18] and have shown significant performances. Another method of mitigating the drifting problem was proposed in [10].

This paper is organized as follows: In section 2, problem statement and motivation is given. In section 3, the proposed algorithm is derived. In section 4, the convergence analysis of the proposed algorithm is presented. In section 5, the computational complexities of the proposed algorithm, LLMS and Modified Leaky Least-Mean-Square (MLLMS) algorithms are calculated. In section 6, a normalized version of the proposed algorithm is presented which improves the step-size selection criteria. Experimental results are presented and discussed in section 7. Finally, the conclusions are drawn.

2. Problem Statement and Motivation

Despite the fact that the LLMS algorithm mitigates the drifting problem in the LMS algorithm, its convergence rate is similar to that of the LMS algorithm. In order to improve this convergence rate, we propose a new LLMS based algorithm. The proposed algorithm employs a sum of exponentials [3] in the cost function of the conventional LLMS; which in turn provides a higher performance in terms of convergence rate.

3. The Proposed Algorithm

For a system identification setting, the output of a linear system with input signal $x(k)$ is given by:

$$d(k) = h^T x(k) + v(k) \tag{1}$$

Where h is the impulse response of the unknown system, $x(k)$ is the tap-input vector, $v(k)$ is an additive noise and $[.]^T$ is the transposition operator. The cost function of the proposed algorithm is given by:

$$J(k) = \left(exp(e(k)) + exp(-e(k))\right)^2 + \gamma w^T(k) w(k) \tag{2}$$

Where $w(k)$ is the filter-tap weight vector, γ is the leakage factor ($0< \gamma< 1$) and $e(k)$ is the instantaneous error and defined by:

$$e(k) = d(k) - w^T(k) x(k) \tag{3}$$

Where $d(k)$ is the desired response of the adaptive filter. Deriving Equation 2 with respect to $w(k)$ gives:

$$\frac{\partial J(k)}{\partial w(k)} = -4x(k)\sinh(e(k)) + 2\gamma w(k) \quad (4)$$

The tap-update is given by [9]:

$$w(k+1) = w(k) - \frac{\mu}{2}\frac{\partial J(k)}{\partial w(k)} \quad (5)$$

Substituting Equation 4 in Equation 5 and rearranging; the update equation of the proposed algorithm becomes:

$$w(k+1) = (1-\gamma\mu)w(k) + 2\mu x(k)sinh(e(k)) \quad (6)$$

In Equation 6, replacing the error $e(k)$ by its sinh guarantees faster convergence especially at the beginning where $e(k)$ is relatively high.

Table 1 below shows a summary of the proposed algorithm.

Table 1. A summary of the proposed algorithm.

initialize the parameters μ, γ and initialize w=0 for k=1, 2, ... $w(k-1)=(1-\gamma\mu)w(k)+2\mu x(k)sinh(e(k))$. Where $e(k)=d(k)-w^T(k)x(k)$.

4. Convergence Analysis

In this section the convergence analysis is presented and the stability criterion of the proposed algorithm is derived. Starting from Equation 6 and using the Taylor series expansion of $sinh(e(k))$ we get:

$$w(k+1) = w(k) - \gamma\mu w(k) + 2\mu x(k)\sum_{l=0}^{\infty}\frac{e(k)^{2l+1}}{(2l+1)!} \quad (7)$$

We define the translated weight vector $\delta w(k)=w(k)-w_0$, where $w_0=R^{-1}p$ is the optimal weight vector, $R=E[x(k)x^T(k)]$ is the autocorrelation matrix of the tap-input and p is the cross-correlation vector defined as $p=E[d(k)\ x(k)]$. Subtracting w_0 from both sides of Equation 7:

$$\delta w(k+1) = (1-\gamma\mu)\delta w(k) - \gamma\mu w_0 + 2\mu x(k)\sum_{l=0}^{\infty}\frac{e(k)^{2l+1}}{(2l+1)!} \quad (8)$$

The optimal error $e_0(k)$ can be defined by:

$$e_0(k) = d(k) - x^T(k)w_0 \quad (9)$$

And from Equation 3:

$$e(k) = e_0(k) - x^T(k)\delta w(k) \quad (10)$$

Substituting Equation 10 in Equation 8 and simplifying by neglecting the high powers of the Taylor series due to small step-size assumption:

$$\begin{aligned}\delta w(k+1) &= (1-\gamma\mu)\delta w(k) - \gamma\mu w_0 + \\ &2\mu x(k)\left[e_0(k) - x^T(k)v(k)\right]\end{aligned} \quad (11)$$

Defining the rotated vectors, $Q^T\delta w(k)=\delta w'(k)$, $Q^Tx(k)=x'(k)$ and $Q^Tw_0=w'_0$, where Q is the eigenvectors matrix of R, we get:

$$\begin{aligned}\delta w'(k+1) &= (1-\gamma\mu)\delta w'(k) - \gamma\mu w'_0 + \\ &2\mu x'(k)e_0(k) - 2\mu x'(k)x'^T(k)\delta w'(k))\end{aligned} \quad (12)$$

Taking the expectation of Equation 12 yields:

$$\begin{aligned}E[\delta w'(k+1)] &= (1-\gamma\mu)E[\delta w'(k)] - \gamma\mu E[w'_0] \\ &-2\mu E[x'(k)x'^T(k)\delta w'(k))] + 2\mu E[x'(k)e_0(k)]\end{aligned} \quad (13)$$

By the independence assumption [9] of $x'(k)$ and $\delta w'(k)$ and also assuming $E[x'(k)x'^T(k)]=\Lambda$ yields:

$$E[\delta w'(k+1)] = [I - \mu(\gamma I + 2\Lambda)]E[\delta w'(k))] - \gamma\mu w'_0 \quad (14)$$

In order to find the convergence criterion, our concern will be in finding the bound when Equation 14 is bounded for all modes. Solving Equation 14 gives:

$$E[\delta w'(k+1)] = [I - \mu(\gamma I + 2\Lambda)]^k E[\delta w'(0)] \quad (15)$$

From Equation 15 it is noted that, $E[\delta w'(k+1)] \rightarrow$ constant value if $|1-\mu\gamma-2\mu\gamma_i|<1$, which shows that the coefficients converge to their optimum solution in the mean sense and hence:

$$0 < \mu < \frac{2}{\gamma + 2\lambda_{max}} \quad (16)$$

Where λ_{max} is the maximum eigenvalue of R. A more practical way of calculating λ_{max} could be [9]:

$$\lambda_{max} \le Tr(R) = \sum_{i=1}^{N}\lambda_i = N\sigma_x^2 \quad (17)$$

Where $tr(.)$ denotes trace of a matrix and σ_x^2 is the power of the input signal.

$$0 < \mu < \frac{2}{N\left(\gamma + 2\sigma_x^2\right)} \quad (18)$$

Following the derivation of the convergence in the mean sense, the convergence in the mean-square sense is derived in this section. The mean-square-error (mse) is given by:

$$\varepsilon = \varepsilon_{min} + E\left[\delta w'^T(k)\Lambda\delta w'(k)\right] \quad (19)$$

Where ε_{min} is the minimum mse. It is necessary that the diagonal elements of $E[\delta w'^T(k)\Lambda\delta w'(k)]$ to converge. Therefore, the convergence condition of $E[\delta w'^T(k)\Lambda\delta w'(k)]$ is obtained by post multiplying both sides of Equation 12 by $\delta w'^T(k+1)$ and taking the expectation provides:

$$\begin{aligned}E[\delta w'(n+1)\delta w'^T(n+1)] &= E[\delta w'(n)\delta w'^T(n)]\\ &\left[I-2\gamma\mu I-2\mu\Lambda+\gamma^2\mu^2 I+2\gamma\mu^2\Lambda\right]\\ &+\left[2\gamma\mu^2\Lambda-2\mu\Lambda\right]E[\delta w'(n)\delta w'^T(n)]\\ &+8\mu^2\Lambda E[\delta w'(n)\delta w'^T(n)]\Lambda\\ &+4\mu^2\Lambda\, tr\left(\Lambda E[\delta w'(n)\delta w'^T(n)]\right)\\ &+E[\delta w'(n)]\left[\gamma^2\mu^2 p'^T\Lambda-\gamma\mu p'^T\Lambda\right]\\ &+\left[\gamma^2\mu^2\Lambda^{-1}p'-\gamma\mu\Lambda^{-1}p'\right]E[\delta w'^T(n)]\\ &+2\gamma\mu^2\Lambda E[\delta w'(n)]p'^T\Lambda\\ &+2\gamma\mu^2\Lambda^{-1}p'E[\delta w'^T(n)]\Lambda\\ &+\gamma^2\mu^2\Lambda^{-1}p'p'^T\Lambda+4\mu^2 E[e_0^2(n)]\Lambda\end{aligned} \tag{20}$$

Defining w'_0 as $w'_0=\Lambda^{-1}p'$ where $p'=Q^Tp$ and employing the Gaussian moment factoring theorem [9] to simplify the expression $E\{x'(k)x'^T(k)\delta w'(k)\delta w'^T(k)x'(k)x'^T(k)\}$, $\Psi_1(k)$ can be defined as an $N\times 1$second moment vector whose components are the diagonal elements of $E[\delta w'^T(k)\ \delta w'(k)]$, and $\Psi_2(k)=E[\delta w'(k)]$, a state vector matrix is constructed as:

$$\Psi(n+1)=\begin{bmatrix}\Psi_1(n+1)\\ \Psi_2(n+1)\end{bmatrix} \tag{21}$$

Using Equations 14, 20, and 21 can be expressed as:

$$\Psi(n+1)=K\Psi(n)+C \tag{22}$$

Where:

$$K=\begin{bmatrix}A & B\\ 0 & D\end{bmatrix} \tag{23}$$

And:

$$A=[1-2\gamma\mu+\gamma^2\mu^2]I-4\mu\Lambda[I-\gamma\mu I]+8\mu^2\Lambda^2+4\mu^2\Gamma\Gamma^T \tag{24}$$

$$B=[(2\gamma^2\mu^2-2\gamma\mu)\Lambda^{-1}+4\gamma\mu^2 I]Y \tag{25}$$

$$D=I-\mu[\gamma I-2\Lambda] \tag{26}$$

Where, $Y=diag\{p'_1, p'_2, \ldots, p'_N\}$ and $\Gamma=[\lambda_1, \lambda_2, \ldots, \lambda_N]^T$. Since C is bounded, we may neglect it in the rest of the analysis. Furthermore, a_j is defined as:

$$a_j=1-2\mu(\gamma+2\lambda_j)+\mu^2((\gamma+2\lambda_j)^2+4\lambda_j^2) \tag{27}$$

Where a_j is the jj^{th} element of A.

In order to study the convergence of $E[\delta w'^T(k)\delta w'(k)]$, it should be noted that Equation 12 is exponentially stable if the roots ρ_j of $det[K\text{-}\rho I]$ lies inside the unit circle [8]. Therefore:

$$det[K-\rho I]=det[D-\rho I]det[A-\rho I] \tag{28}$$

Where from Equation 25:

$$det[D-\rho I]=\prod_{j=1}^{N}(1-\mu(\gamma+2\lambda_j)-\rho). \tag{29}$$

Equations 28 and 29 show the first convergence condition for convergence of $\Psi(k)$ is the same as the condition found for the convergence in the mean. Also knowing that $det[\Gamma\Gamma^T]=0$, then by [7]:

$$\begin{aligned}det[A-\rho I]&=\left(det[diag\{a_1-\rho,a_2-\rho,\ldots,a_N-\rho\}]\right)\\ &\bullet\left(1+4\mu^2\Gamma^T diag\{a_1-\rho,a_2-\rho,\ldots,a_N-\rho\}^{-1}\Gamma\right)\\ &=\left[\prod_{j=1}^{N}(a_j-\rho)\right]\left[1+4\mu^2\sum_{j=1}^{N}\frac{\lambda_j^2}{a_j-\rho}\right]\end{aligned} \tag{30}$$

Following the approach in [6] it can be shown that the necessary and sufficient conditions for the roots of $det[A\text{-}\rho I]$ to be inside the unit circle are:

$$-1<a_j<1 \quad j=1, 2, \ldots, N \tag{31}$$

And:

$$1+4\mu^2\sum_{j=1}^{N}\frac{\lambda_j^2}{a_j-\rho}>0 \tag{32}$$

From the definition of a_j in (27) as a function of μ, it can be seen that it is a convex function and is greater than zero $\forall\mu$ and γ since it has a minimum non-negative value of:

$$\frac{4\lambda_j^2}{(\gamma+2\lambda_j)^2+4\lambda_j^2} \tag{33}$$

At:

$$\mu=\frac{(\gamma+2\lambda_j)}{(\gamma+2\lambda_j)^2+4\lambda_j^2} \tag{34}$$

This changes the condition to:

$$a_j<1, \quad j=1, 2, \ldots, N \tag{35}$$

And:

$$1+4\mu^2\sum_{j=1}^{N}\frac{\lambda_j^2}{a_j-1}>0 \tag{36}$$

The condition in Equation 35 results in:

$$\mu(\mu((\gamma+2\lambda_j)^2+4\lambda_j^2)-2(\gamma+2\lambda_j))<0 \tag{37}$$

Using Equation 16, $\mu>0$; then from Equation 37 we get:

$$0<\mu<\frac{2(\gamma+2\lambda_j)}{(\gamma+2\lambda_j)^2+4\lambda_j^2} \quad j=1, 2, \ldots, N \tag{38}$$

Equation 36 leads to the second condition on μ for convergence in the mean square:

$$\sum_{j=1}^{N}\frac{4\mu\lambda_j^2}{2(\gamma+2\lambda_j)-\mu((\gamma+2\lambda_j)^2+4\lambda_j^2)}-1<0 \tag{39}$$

In order to convert this condition into a direct bound on μ we denote the left hand side of Equation 39 by $\chi(\mu)$ knowing that $\chi(\mu)$ is a monotonically non-decreasing function of μ since:

$$\frac{\delta\chi(\mu)}{\delta\mu}=\sum_{j=1}^{N}\frac{8\lambda_j^2(\gamma+2\lambda_j)}{(2(\gamma+2\lambda_j)-\mu((\gamma+2\lambda_j)^2+4\lambda_j^2))^2} \tag{40}$$

Also, $\chi(\mu)$ has poles at:

$$\rho_j=\frac{2(\gamma+2\lambda_j)}{(\gamma+2\lambda_j)^2+4\lambda_j^2},\quad j=1,2,...,N \tag{41}$$

And is equal to -1 at μ=0 and $\lim_{\mu\to\infty}\chi(\mu)<0$. $\chi(\mu)$=0 can be written as:

$$\chi(\mu)=\prod_{i=1}^{N}(\mu-\mu_i)=0 \tag{42}$$

Where μ_i, i=1, 2, …, N are the roots of $\chi(\mu)$ and $0<\lambda\le\lambda_{N-1}\le\lambda_{N-2}\le\ldots\le\lambda_1$. It can be seen that:

$$0<\mu_1<\rho_1<\mu_2<\rho_2<\ldots<\mu_N<\rho_N \tag{43}$$

In order for the conditions in Equations 35 and 36 to hold, μ should be bounded by:

$$0<\mu<\mu_1 \tag{44}$$

The condition in Equation 39 provides a more stringent bound on μ as given in Equation 44 than in Equation 38 since $\mu_1<\rho_1$. A closed-form expression for μ_1 cannot be found but using the theorem in [11] to obtain a tight lower bound on μ_1. For Equation 42 where $0<\mu_1<\mu_2<\ldots<\mu_N$, the smallest root μ_1 is lower bounded by:

$$\mu_1\ge\frac{N}{s_{-1}+\sqrt{(N-1)(Ns_{-2}-s_{-1}^2)}} \tag{45}$$

Where,

$$s_{-1}=\sum_{i=1}^{N}\frac{1}{\mu_i} \tag{46}$$

And

$$s_{-2}=\sum_{i=1}^{N}\left(\frac{1}{\mu_i}\right)^2=\left(\sum_{i=1}^{N}\frac{1}{\mu_i}\right)^2-\sum_{i\ne}^{N}\sum_{j}^{N}\frac{1}{\mu_i}\cdot\frac{1}{\mu_j} \tag{47}$$

Rewriting $\chi(\mu)$=0 as:

$$\prod_{i=1}^{N}\left(\frac{1}{\mu}-\frac{1}{\mu_i}\right)=0 \tag{48}$$

And:

$$\left(\frac{1}{\mu}\right)^N-b_1\left(\frac{1}{\mu}\right)^{N-1}+b_2\left(\frac{1}{\mu}\right)^{N-2}\cdots+b_N(-1)^N=0 \tag{49}$$

Then expanding the first part of Equation 48 and comparing it with the second part of the same equation:

$$b_1=\sum_{i=1}^{N}\frac{1}{\mu_i} \tag{50}$$

And:

$$b_2=\frac{1}{2}\sum_{i\ne}^{N}\sum_{j}^{N}\frac{1}{\mu_i}\cdot\frac{1}{\mu_j} \tag{51}$$

Therefore:

$$\begin{aligned}s_{-1}&=b_1\\ s_{-2}&=b_1^2-2b_2\end{aligned} \tag{52}$$

Comparing the LHS of the second part of Equation 48 with the LHS of Equation 39, it can be seen after some manipulations that:

$$b_1=\sum_{i=1}^{N}\frac{8\lambda_i^2+(\gamma+2\lambda_i)^2}{2(\gamma+2\lambda_i)} \tag{53}$$

And:

$$b_2=\sum_{i\ne}^{N}\sum_{j}^{N}\frac{\left(12\lambda_i^2+(\gamma+2\lambda_i)^2\right)\left(4\lambda_j^2+(\gamma+2\lambda_j)^2\right)}{8(\gamma+2\lambda_i)(\gamma+2\lambda_j)} \tag{54}$$

Substituting Equation 52 in Equation 45 results in:

$$\mu_1\ge\frac{N}{b_1+\sqrt{b_1^2(N-1)^2-2b_2N(N-1)}}=\mu^* \tag{55}$$

Thus, to ensure convergence in the mean square, μ should be bounded by:

$$0<\mu<\mu^* \tag{56}$$

To make Equation 56 more practical:

$$\mu^*\ge\frac{1}{b_1} \tag{57}$$

Then:

$$0<\mu<\frac{1}{\sum_{j=1}^{N}\frac{8\lambda_j^2+(\gamma+2\lambda_j)^2}{2(\gamma+2\lambda_j)}} \tag{58}$$

A tighter and more practical bound can be expressed in terms of the input signal power σ_x^2 as:

$$0<\mu<\frac{2}{N\left(6\sigma_x^2+\gamma\right)} \tag{59}$$

5. Computational Complexity

In this section, computational complexities of the LLMS [9], MLLMS in [12] and the proposed algorithm are discussed.

For the proposed algorithm, neglecting the higher order terms of the Taylor series expansion of exp($e(k)$), it can be approximated by:

$$exp(e(k))\approx1+e(k)+\frac{e^2(k)}{2}+\frac{e^3(k)}{6} \tag{60}$$

Table 2 shows that the computational complexity of the proposed algorithm is very comparable to that of the LLMS algorithm and lower than that of the MLLMS algorithm if the filter length is relatively high.

Table 2. Computational complexities of the proposed, LLMS and MLLMS algorithms.

	add./sub.	**mult./div.**
Proposed	2N+6	3N+6
LLMS	2N+1	3N+2
MLLMS	3N	3N+2

6. Normalized Version of The Algorithm

For the proposed algorithm to converge the step-size parameter needs to be very small, this small step-size is an unfavorable condition that limits the application of the proposed algorithm in many situations. In order to improve the step-size selection range, we proposed a normalized version of the algorithm that though does not improve convergence nevertheless it improves the stability of the proposed algorithm. The normalization is given by:

$$w(k+1) = (1-\gamma\mu)w(k) + 2\frac{\mu}{\delta + x^T(k)x(k)}x(k)sinh(e(k)) \quad (61)$$

Where δ is a small number to avoid division by zero.

7. Simulation Results

In this section, we compare the performance of the proposed algorithm to those of the LLMS [9] and MLLMS [11] algorithms in channel estimation and channel equalization settings under white Gaussian and white and correlated impulsive noise environments.

7.1. Channel Estimation

In the channel estimation setting shown in Figure 1, the aim is estimate the impulse response of the unknown system h. The input signal is created using a first order autoregressive model (AR(1)) given by $x(k) = 0.8x(k-1) + \eta_0(k)$, where $\eta_0(k)$ is a white Gaussian process with zero mean and variance $\sigma^2_{\eta_0} = 0.36$.

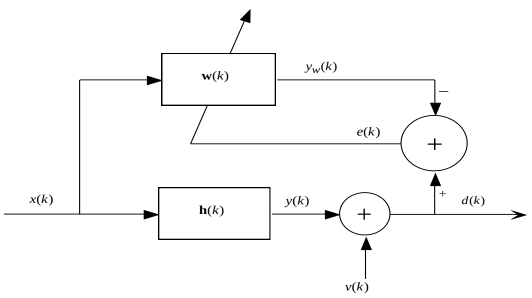

Figure 1. Block diagram of a channel estimation setting.

The impulse response of the system is modeled by a low pass filter of 16 taps (N=16) with the transfer function shown in Figure 2. The convergence rate and the mse are considered as the performance measures. The simulations were done for stationary signals corrupted with white and correlated impulsive noises.

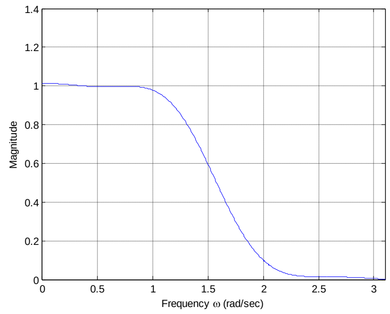

Figure 2. Transfer function of the channel's impulse response h.

7.1.1. Additive White Gaussian Noise

In this experiment, the input signal is assumed to be corrupted by an Additive White Gaussian Noise (AWGN) process with zero mean and variance $\sigma^2_v = 2.25\times10^{-4}$. The simulations were done with: μ=0.007 for all algorithms and γ=0.001 for the proposed and LLMS algorithms. Figure 3 shows that all the algorithms converge to the same mse=-38.5dB. However, the proposed algorithm converges much faster than the other algorithms (800 iterations faster).

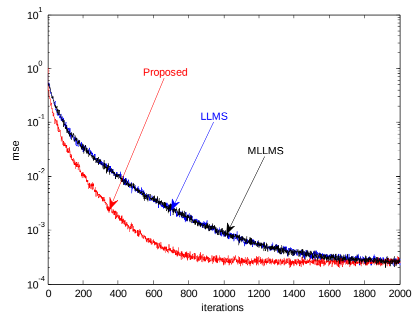

Figure 3. The ensemble mse of the proposed, LLMS and MLLMS algorithms in AWGN, N=16, μ=0.007 for all algorithms and γ=0.001 for the proposed and LLMS algorithms.

7.1.2. Additive White Impulsive Noise

Noises such as atmospheric noise, under water acoustic noise, man-made noise, etc. are not usually modeled as Gaussian noises; they are better modeled as impulsive noise [1]. An impulsive noise process can be generated using the probability density function [5, 19]: $f = (1-\varepsilon)G(0,\sigma^2_v) + \varepsilon G(0,\kappa\sigma^2_v)$, with variance $\sigma^2_f = (1-\varepsilon)\sigma^2_v + \varepsilon\kappa\sigma^2_v$. The impulsive noise comprises of nominal background Gaussian noise represented by $G(0,\sigma^2_v)$ with zero mean and variance σ^2_v and an impulsive part represented by $G(0,\kappa\sigma^2_v)$ where $\kappa \geq 1$ and ε are the strength and the probability of the impulsive components, respectively. In this experiment, an Additive White Impulsive Noise (AWIN) process with zero mean and variance $\sigma^2_v = 2.25\times10^{-4}$ is used with κ=100 and ε=0.2. The simulations were done with: μ=0.004 for all algorithms and γ=0.001 for the proposed and LLMS algorithms. Figure 4 shows that all the algorithms converge to the same mse=-25dB. However, the proposed algorithm converges much faster than the other algorithms (600 iterations faster).

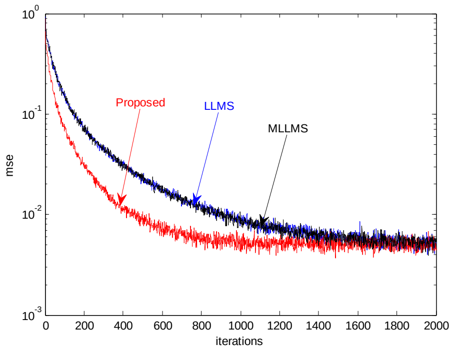

Figure 4. The ensemble mse of the proposed, LLMS and MLLMS algorithms in AWIN, N=16, μ=0.004 for all algorithms and γ=0.001 for the proposed and LLMS algorithms.

7.1.3. Additive Correlated Impulsive Noise

In this experiment, the signal created in section 6.1 is assumed to be corrupted by an Additive Correlated Impulsive Noise (ACIN) process. The ACIN is generated by AR(1) process, $v(k+1)=\rho v(k)+v_0(k)$, where $v_0(k)$ is an AWIN process with zero mean and variance $\sigma_{v_0}^2 = 2.25\times10^{-4}$, and ρ is the correlation coefficient (ρ=0.7). The simulations were done with: μ=0.007 for all algorithms and γ=0.001 for the proposed and LLMS algorithms. Figure 5 shows that all of the algorithms converge to the same mse=-20dB but with faster convergence rate of the proposed algorithm (800 iterations faster).

From the previous experiments, for the channel estimation setting in Figure 1, the proposed algorithm, always, converges faster than the LLMS and the MLLMS algorithms. This shows the achieved improvement in the performance by modifying the cost function of the proposed algorithm. Also, in ACIN case, the convergence rate difference between the proposed algorithm and the other algorithms is higher. This shows the ability of the proposed algorithm in suppressing the noise is prominent when the noise is correlated.

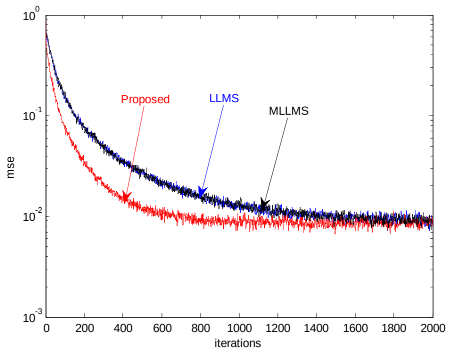

Figure 5. The ensemble mse of the proposed, LLMS and MLLMS algorithms in ACIN, N=16, μ=0.007 for all algorithms and γ=0.001 for the proposed and LLMS algorithms.

7.2. Channel Equalization

For the channel equalization setting shown in Figure 6, the input is generated using a Bernoulli sequence of zero mean and unity variance where $x(k)$= ±1 and the impulse response of the channel $h(k)$ is given by:

$$h(k) = \begin{cases} \frac{1}{2}\left[1+\cos\left(\frac{2\pi}{W}(k-2)\right)\right], & k=1,2,3 \\ 0, & \text{otherwise.} \end{cases} \tag{62}$$

Where W controls the eigenvalue spread $\chi(R)$ of the autocorrelation of the tap-input sequence. A filter of length N=11 is used. W=3.5 is used for the simulations in this section.

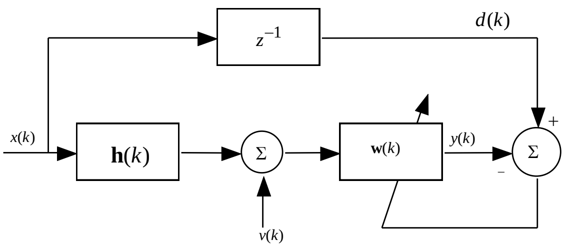

Figure 6. Block diagram of a channel equalization setting.

7.2.1. Additive White Gaussian Noise

In this experiment, an AWGN process with zero mean and variance $\sigma_v^2 = 1\times10^{-4}$ is used. The simulations were done with: μ=0.01 for all algorithms and γ=0.001 for the proposed and LLMS algorithms. Figure 7 shows that all the algorithms converge to the same mse=-30dB with much faster convergence rate of the proposed algorithm (3000 iterations faster than the other algorithms).

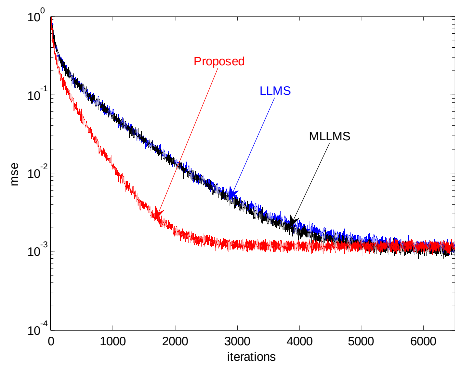

Figure 7. The ensemble mse of the proposed, LLMS and MLLMS algorithms in AWGN, N=11, μ=0.01 for all algorithms and γ=0.001 for the proposed and LLMS algorithms.

7.2.2. Additive White Impulsive Noise

In this section, the performance of the proposed algorithm is compared to those of LLMS and MLLMS algorithms in the channel equalization setting shown in Figure 6 where $v(k)$ is an AWIN. The noise process is generated as in section 7.1.2 with zero mean and variance $\sigma_v^2 = 1\times10^{-4}$. The simulations were done with:

μ=0.01 for all algorithms and γ=0.001 for the proposed and LLMS algorithms. Figure 8 shows that all algorithms converge to the same mse=-20dB with a faster convergence rate for the proposed algorithm than the LLMS and MLLMS algorithms (1500 iterations faster).

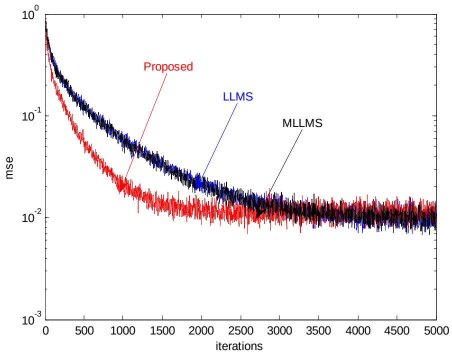

Figure 8. The ensemble mse of the proposed, LLMS and MLLMS algorithms in AWIN, N=11, μ=0.01 for all algorithms and γ=0.001 for the proposed and LLMS algorithms.

7.2.3. Additive Correlated Impulsive Noise

In this experiment, the ACIN generated in section 6.1.3 is used with zero mean and variance $\sigma_{v_0}^2 = 1\times10^{-4}$. The simulations were done with: μ=0.01 for all algorithms and γ=0.001 for the proposed and LLMS algorithms. From Figure 9, all algorithms converge to the same mse=-23dB with the proposed algorithm having a faster convergence rate than the LLMS and MLLMS algorithms (1500 iterations faster).

Sections 7.2.1 to 7.2.3 show simulation results, for channel equalization setting, with white AWGN and white and correlated impulsive noise processes. The convergence rate of the proposed algorithm is faster than those of the LLMS and MLLMS algorithms (3000, 1500 and 1500 iterations, respectively, faster). The significant performance in all cases shows the robustness of the proposed algorithm in AWGN and impulsive noise environments. Furthermore, the computational complexity of the proposed algorithm is very comparable to those of the LLMS and MLLMS algorithms as shown in Table 2.

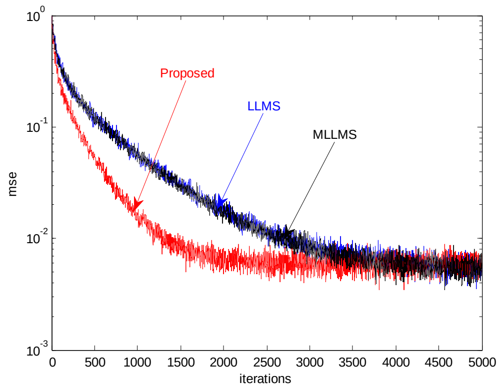

Figure 9. The ensemble mse of the proposed, LLMS and MLLMS algorithms in ACIN, N=11, μ=0.01 for all algorithms and γ=0.001 for the proposed and LLMS algorithms.

8. Conclusions

In this paper, a new leaky LMS adaptive filtering algorithm is proposed. The proposed algorithm employs a sum of exponentials in its cost function which leads to an improved performance. The convergence analysis of the proposed algorithm is presented and the convergence criteria are derived. Also, a normalized version of the proposed algorithm is shown which improves the step-size selection range. Simulation results show that the proposed algorithm outperforms the LLMS and MLLMS algorithms in different experimental settings with very comparable computational complexity.

References

[1] Ahmad M., Kukrer O., and Hocanin A., "An Efficient Recursive Inverse Adaptive Filtering Algorithm for Channel Equalization," *in Proceeding of European Wireless*, Lucca, pp. 88-92, 2010.

[2] Arya S. and Singh B., "Performance of DSTATCOM using Leaky-LMS Control Algorithm," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 1, no. 2, pp. 104-113, 2013.

[3] Boukis C., Mandic D., and Constantinides A., "A Generalised Mixed Norm Stochastic Gradient Algorithm," *in Proceeding of 15th International Conference on Digital Signal Processing*, Wales, pp. 27-30. 2007.

[4] Cartes D., Ray L., and Collier R., "Lyapunov Tuning of the Leaky LMS Algorithm for Single-Source, Single-Point Noise Cancellation," *in Proceeding of the American Control Conference*, Alington, pp. 3600-3605, 2001.

[5] Delic H. and Hocanin A., "Robust Detection in DS/CDMA," *Journal of IEEE Transactions on Vehicular Technology*, vol. 51, no. 1, pp. 155-170, 2002.

[6] Feuer A. and Weinstein E., "Convergence Analysis of LMS Filters with Uncorrelated Gaussian Data," *Journal of IEEE Transactions on Acoustic, Speech and Signal Processing*, vol. 33, no. 1, pp. 222-230, 1985.

[7] Foley J. and Boland F., "A Note on Convergence Analysis of LMS Adaptive Filters with Gaussian Data," *Journal of IEEE Transactions on Acoustic, Speech and Signal Processing*, vol. 36, no. 7, pp. 1087-1089, 1988.

[8] Goodwin G., Hill D., and Xianya X., "Stochastic Adaptive Control for Exponentially Convergent Time-Varying Systems," *SIAM Journal on Control and Optimization*, vol. 24, no. 4, pp. 589-603, 1986.

[9] Haykin S., *Adaptive Filter Theory*, Prentice Hall, 2002.

[10] Hovd M. and Bitmead R., "Directional Leakage and Parameter Drift," *International Journal of Adaptive Control and Signal Processing*, vol. 20, no. 1, pp. 27-39, 2006.

[11] Jagerman D., "Nonstationary Blocking in Telephone Traffic," *Journal of Bell System Technology*, vol. 54, no. 3, pp. 625-661, 1975.

[12] Jo S., Choi J., and Lee Y., "Modified Leaky LMS Algorithm for Channel Equalization in DS-CDMA Systems," *Journal of IEEE Communication Letters*, vol. 6, no. 5, pp. 202-204, 2002.

[13] Mayyas K. and Aboulnasr T., "Leaky-LMS: A Detailed Analysis," *in Proceeding of IEEE International Symposium on Circuits and Systems*, Seatlle, pp. 1255-1258, 1995.

[14] Nascimento V. and Sayed A., "Unbiased and Stable Leaky-based Adaptive Filters," *Journal of IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3261-3276, 2002.

[15] Rigling B. and Schniter P., "Subspace Leaky LMS," *Journal of IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 136-139, 2004.

[16] Rojas H. and Cortes C., "Adaptive Filtering for Lightning Electric Field Signals in Fractional Fourier Domain," *in Proceeding of IEEE Statistical Signal Processing Workshop*, Ann Arbor, pp. 213-216, 2012.

[17] Sethares W., Lawrence D., Johnson C., and Bitmead R., "Parameter Drift in LMS Adaptive Filters," *Journal of IEEE Transactions on Acoustic, Speech and Signal Processing*, vol. 34, no. 4, pp. 868-879, 2003.

[18] Varade S. and Kulat K., "Improvement in Interference Rejection Capability using Space-Time Turbo Code and Adaptive Beamforming for MIMO-OFDM System," *in Proceeding of IEEE Conference on Signal Processing, Computing and Control*, Waknaghat, pp. 1-8, 2012.

[19] Zhu Z., Zhang X., Wan X., and Wang Q., "A Random-Valued Impulse Noise Removal Algorithm with Local Deviation Index and Edge-Preserving Regularization," *Journal of Signal, Image and Video Processing*, vol. 9, no. 1, pp. 221-228, 2015.

Tajuddeen Gwadabe was born in 1987 in Nigeria. He received his B.E. in electrical engineering from Bayero University Kano (BUK), Nigeria, in 2011 and master's degree in electrical and electronics at Mevlana University, Turkey. His research interests include signal processing, adaptive filtering techniques and image processing.

Mohammad Salman received the B.Sc., M.Sc., and Ph.D. Degrees in Electrical and Electronics Engineering from Eastern Mediterranean University (EMU), in 2006, 2007, and 2011, respectively. From 2006 to 2010, he was a teaching assistant of Electrical and Electronics Engineering department at EMU. In 2010, he joined the Department of Electrical and Electronic Engineering at European University of Lefke (EUL) as a senior lecturer for the Department. Since 2011, he worked as an Assist. Prof. Dr. in the Department of Electrical and Electronics Engineering, Mevlana (Rumi) University, Turkey. He served as a general chair, program chair, and a TPC member for many international conferences. He has supervised 6 M.Sc. Theses and currently supervising 4 Ph.D. Theses. His research interests include signal processing, adaptive filters, image processing, and communications systems.