International Journal of Advanced Technology in Engineering and Science
Vol. No. 09, Issue No. 02, February 2021
www.ijates.com
ijates
ISSN 2348 - 7550

A Study of Least Mean Square Algorithm

Jaswant[1], Sanjeev Kumar Dhull[2]

[1]*Ph.D. Scholar, ECE Department, GJUS&T, HSR-125001, (India)*

[2]*Professor, ECE Department, GJUS&T, HSR-125001, (India)*

Abstract

This paper deals with analytical approach of LMS algorithm. Stability analysis of LMS for various values of step size has been simulated along with mathematical derivation of stability criteria. Coefficient behavior, MSE, EMSE etc. parameters have been analyzed and simulated. QAM reception of LMS algorithm for various values of step size (μ), SNR along with convergence rate has been simulated.

Keywords: *Algorithm, EMSE, LMS, MSE, QAM, Trained.*

1. INTRODUCTION

In training based data-aided or data supervised equalization techniques [1][8], a pre-known chunk of training data known as training/pilot sequence is introduced to the receiver which helps the receiver adapt to the channel variations and then utilize that pilot sequence to estimate channel and to eliminate or minimize inter-symbol interference (ISI). Merits of data aided based equalization technique include rapid convergence rate, better efficiency, low complexity and has simple application and implementation. This technique is considered suitable for environment where fast fading with high Doppler spread and little coherence time is present. The demerit of this kind of equalizer is requirement of pilot signals constantly. The constant transmission of the pilot data sequence consumes significant bandwidth which leads to slow data transmission rate. A portion of precious and limited bandwidth is occupied by training data. In GSM network, about 18% of the bandwidth is consumed by the pre-known data sequence that is sent to receiver periodically. There are various training or data supervised algorithms that can be used in a training-based adaptive equalizer e.g. LMS [1][3], NLMS[16-18], APA[15] and RLS [2][7]. These algorithms are adaptive in nature to deal with time varying behavior of transmission channel.

2. LEAST MEAN SQURE (LMS)[8]

An N-dimensional input vector, $\alpha(n) = [\ \alpha(n)\ \ \alpha(n-1) \cdots\cdots \alpha(n-N+1)]^T$ Here $[.]^T$ denotes Transpose of a Matrix.

Tap Weight vector, $w = [w_1\ \ w_2 \cdots\cdots\cdots\cdots\cdots\ \ w_N]^T$ and $\alpha(n)$: zero-mean ($\mu_\alpha=0$), Wide Sense Stationary (WSS) input signal. w: N-tap FIR-filter with filter weights: $w_1\ \ w_2 \cdots\cdots\cdots\cdots\cdots\cdots\ w_N$,y(n): output of filter given by, $y(n) = \sum_{i=0}^{N-1} w_i\, \alpha(n-i) = w^T\alpha(n) = \alpha^T(n)w$

d(n): zero-mean ($\mu_d=0$), Wide Sense Stationary (WSS) desired signal and ε(n): error signal

Here, α(n) and d(n) are assumed to be jointly WSS.

$$R_{\alpha\alpha}= E[\alpha(n)\alpha^H(n)] \tag{1}$$

$$R_{\alpha\alpha}=\begin{bmatrix} R_{\alpha\alpha}(0) & \cdots & R_{\alpha\alpha}(N-1) \\ \vdots & \ddots & \vdots \\ R^*_{\alpha\alpha}(N-1) & \cdots & R_{\alpha\alpha}(0) \end{bmatrix}$$

Here, $R_{\alpha\alpha}$ is the correlation matrix of the input vector α(n), E[.] & '*' denote Expectation Operator & Conjugate respectively.

$$P_{\alpha d} = E[\alpha(n)d(n)] \tag{2}$$

$$P_{\alpha d} = [P_{\alpha d}(n,n) \ldots\ldots\ldots\ldots\ldots\ldots.. P_{\alpha d}(n-N+1,n)]^T$$

$P_{\alpha d}$ denotes the cross-correlation matrix between input signal vector α(n) and the desired signal d(n).

2.1 The Wiener-Hopf Equation for WSS

The adaptive filtering situation with non time-varying or stationary filter coefficients is depicted in Fig.1.

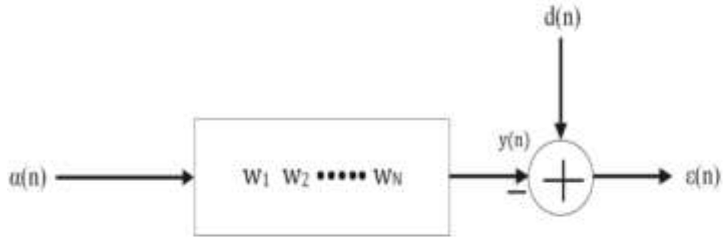

Fig.1.Adaptive Filtering schematic in Wide Sense Stationary (WSS) environment

Then

$$\varepsilon(n) = [d(n) - y(n)] = d(n) - \alpha^T(n)w$$

Now, the Mean Squared Error (MSE) is minimized.

$$E[\varepsilon(n)^2] = E[\{d(n) - y(n)\}^2]$$

$$E[\varepsilon(n)^2] = E[d(n)^2 + y(n)^2 - 2d(n)y(n)]$$

$$E[\varepsilon(n)^2] = \sigma_d^2 + w^H R_{\alpha\alpha} w - 2w^H P_{\alpha d} = \xi^2_{min}(w) \tag{3}$$

Here, $\xi^2_{min}(w)$ denotes cost function that has to be minimized. This notation $\xi^2_{min}(w)$ denotes this is second order equation of weight coefficient. $\blacktriangledown_w[.]$ denotes derivative of function w.r.t. 'w'.

$$\blacktriangledown_w[\xi^2_{min}(w)] = 0$$

$$\blacktriangledown_w[\sigma_d^2] + \blacktriangledown_w[w^H R_{\alpha\alpha} w] - 2\blacktriangledown_w[w^H P_{\alpha d}] = 0$$

Now

$\blacktriangledown_w[\sigma_d^2] = 0$, $\blacktriangledown_w[w^H R_{\alpha\alpha} w] = 2R_{\alpha\alpha} w$ & $\blacktriangledown_w[w^H P_{\alpha d}] = P_{\alpha d}$

$$0 + 2R_{\alpha\alpha} w - 2P_{\alpha d} = 0$$

$$w_{opt} = w = R^{-1}_{\alpha\alpha} P_{\alpha d} \tag{4}$$

w_{opt} corresponds to unique solution to the Wiener-Hopf equation.

For minimum solution, second derivative is calculated.

$$\nabla_w\{\nabla_w[\xi^2_{min}(w)]\}$$

Where: $\nabla_w\{\nabla_w[R_{\alpha\alpha}w]\} = R_{\alpha\alpha}$ & $\nabla_w\{\nabla_w[P_{\alpha d}]\} = 0$

$$\nabla_w\{\nabla_w[\xi^2_{min}(w)]\} = R_{\alpha\alpha}$$

Where, $R_{\alpha\alpha}$ is a positive definite (P.D.) matrix. Hence, it is indeed a global minima.

It can be concluded that:

- $\xi^2_{min}(w)$ is a convex function,
- $R_{\alpha\alpha}$ is not singular but invertible matrix,
- $w_{opt} = w = R^{-1}_{\alpha\alpha}P_{\alpha d}$is the global minimiser,

The solution w_{opt} is often referred to as the least-mean-squares solution or optimal solution.

2.2 The Wiener-Hopf Equation for non-WSS

If α(n) and d(n) are not jointly WSS, Filter coefficients are not fixed but adaptive or time varying in nature.

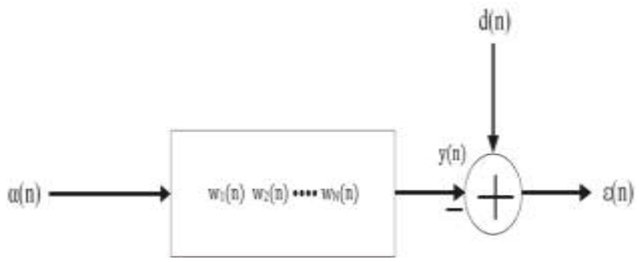

Fig.2. Adaptive Filtering schematic in non-Wide Sense Stationary (WSS) environment

The optimal solution can be calculated:

$$w_{opt}(n) = R^{-1}_{\alpha\alpha}(n)\, P_{\alpha d}(n) \tag{5}$$

However, this method may have following demerits:

- The computational complexity is very high
- The statistics are not known.

The steepest descent (SD) algorithm overcomes with computational complexity problem, but least-mean-square (LMS) adaptive filter overcomes both the problems.

2.3 STEEPEST DESCENT METHOD [8]

The fundamental behind the steepest descent (SD) algorithm is to approach at the unique solution w_{opt} of the Wiener-Hopf equation through a sequence of steps, beginning from some initial point, say, $w_{ini}(0)$. These sequence of steps are followed in negative direction of the gradient $\nabla_w[\xi^2_{min}(w)]$ to reach the minima. This idea is depicted in Fig. 3 and Fig. 4. The filter coefficients are calculated by a recursive update equation given by:

ijates
ISSN 2348 - 7550

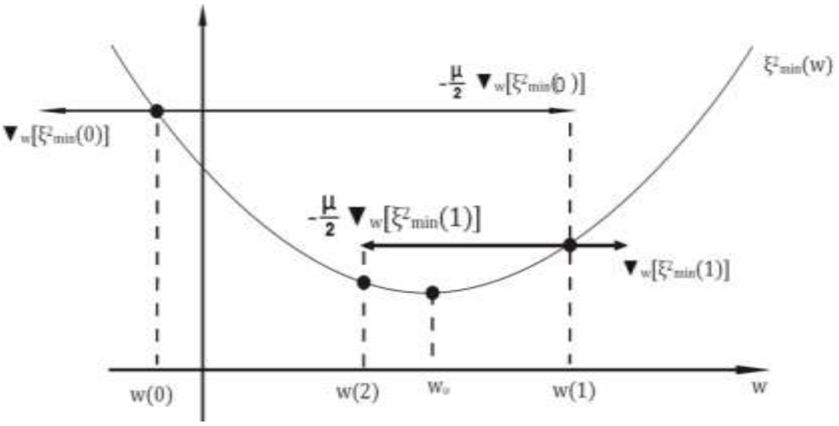

Fig.3.First 3 iterations of the SD Procedure for 1-D filter vector.

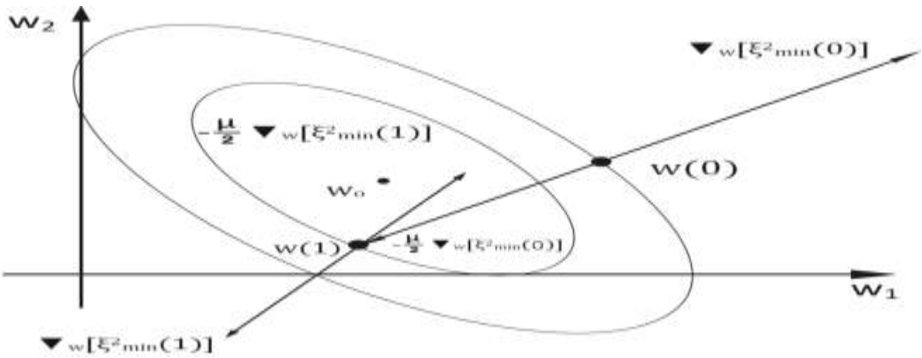

Fig.4.First 3 iterations of SD Procedure for 2-D filter vector.

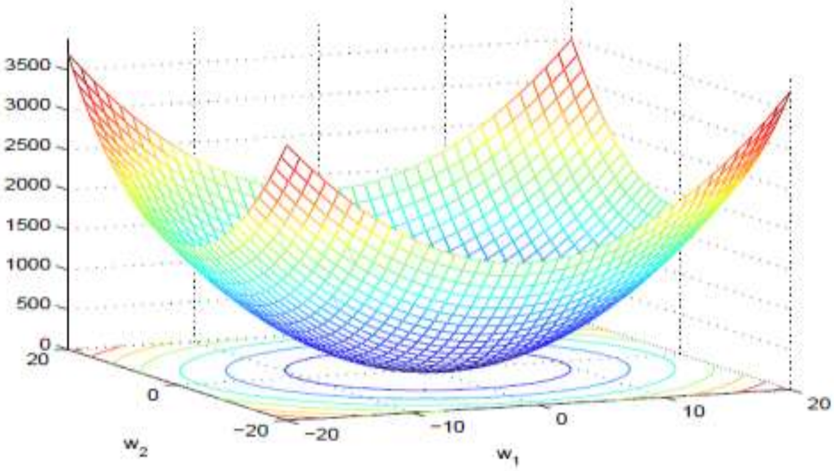

Fig.5. MSE performance curve in 3D

$$w(i+1) = w(i) - \frac{\mu}{2} \nabla_w[\xi^2_{min}(w)] \quad (6)$$

where: $\nabla_w[\xi^2_{min}(w)] = 2R_{\alpha\alpha}w - 2P_{\alpha d}$

$$w(i+1) = w(i) + \mu[P_{\alpha d} - R_{\alpha\alpha}w(i)]$$

Eqn. (6) is an offline procedure. It can further be modified in online LMS procedure or more precisely saying data dependent procedure. For online procedure 'n' is used instead of 'i'.

Steepest Descent Procedure:

1 Initialization of SD algorithm with an initial value $w_{ini}(0)$.

2 Iterating it for i= 0; 1; 2; 3; : : : ; i_{max}

3 Update equation for SDP: $w(i+1) = w(i) + \mu[P_{\alpha d} - R_{\alpha\alpha}w(i)]$

2.4 Least Mean Square Procedure:

Now $R_{\alpha\alpha}$& $P_{\alpha d}$ matrices can be approximated as:

$$R_{\alpha\alpha} = E[\alpha(n)\alpha^H(n)] \approx \alpha(n)\alpha^H(n) \text{ and } P_{\alpha d} = E[\alpha(n)d(n)] \approx \alpha(n)d(n)$$

From eqn. (6)

$$w(n+1) = w(n)+\mu[P_{\alpha d} - R_{\alpha\alpha}w(n)]$$

$$w(n+1) = w(n)+\mu\alpha(n)\varepsilon^*(n) \quad (7)$$

Eqn (7) represents the weight update of LMS algorithm.

2.5 LMS Algorithm:

1 Initialization of SD algorithm with an initial value $w_{ini}(0)$, for example w(0) = 0.

2 Iterating it for n = 0; 1; 2; 3; : : : ; n_{max}

3Filter output: $y(n) = w^T\alpha(n)=\alpha^T(n)w$

4 Output error: $\varepsilon(n) = [d(n) - y(n)] = d(n) - \alpha^T(n)w$

5 Weight update equation: $w(n+1) = w(n)+\mu\alpha(n)\varepsilon^*(n)$

Table 1: Computational Cost of LMS

Terms	×	+ or -
$\alpha^H(n)w$	N	N-1
$\varepsilon(n) = d(n) - \alpha^H(n)w$	-	1
$\mu\varepsilon(n)$	1	-
$\mu\alpha(n)\varepsilon^*(n)$	N	-
$w(n) + \mu\alpha(n)\varepsilon^*(n)$	-	N
Total	**2N+1**	**2N**

Analysis of LMS

Define the weight error:

$$\Delta(n) = w(n) - w_{opt} \tag{8}$$

Step size parameter is chosen such that SD algorithm converges to optimal solution w_{opt}. It follows as:

$$\lim_{n\to\infty} \Delta(n) = 0$$

When α(n) and d(n) are WSS jointly, the SD algorithm is said to be stable. Moreover, step size (μ) is selected such that the Δ(n) approaches to the smallest value as fast as possible.

Convergence Analysis:

$$w(n+1)= w(n)+\mu\alpha(n)\varepsilon^*(n)$$

$$[w(n+1) - w_{opt}] = [w(n) - w_{opt}] +\mu\alpha(n)\varepsilon^*(n)$$

$$\Delta(n+1)= \Delta(n)+\mu\alpha(n)\varepsilon^*(n)$$

$$\Delta(n+1)=\Delta(n)+\mu\alpha(n)d(n) - \mu\alpha(n)\alpha^H(n)w_{opt} - \mu\alpha(n)\alpha^H(n)\Delta(n) \tag{9}$$

We define: $E[\Delta(n)]=\upsilon(n)$

Now taking Expectation operator on both sides:

From eqn.(9), we get $E[\Delta(n+1)]=E[\Delta(n)+\mu\alpha(n)d(n) - \mu\alpha(n)\alpha^H(n)w_{opt}- \mu\alpha(n)\alpha^H(n)\Delta(n)]$

$$\upsilon(n+1)= \upsilon(n)+ \mu[P_{\alpha d} - R_{\alpha\alpha}w_{opt}]-\mu E[\alpha(n)\alpha^H(n)\Delta(n)]$$

From Wiener-Hopf result $w_{opt} = w = R^{-1}_{\alpha\alpha}P_{\alpha d}$, we get: $\upsilon(n+1)= \upsilon(n) - \mu E[\alpha(n)\alpha^H(n)\Delta(n)]$

Using Statistically Independent Assumption:

$$\upsilon(n+1)= \upsilon(n)- \mu E[\alpha(n)\alpha^H(n)]E[\Delta(n)]$$

$$\upsilon(n+1)=[I- \mu R_{\alpha\alpha}]\upsilon(n) \tag{10}$$

$R_{\alpha\alpha}$ is a Harmitian & positive definite matrix. It has an eigen value decomposition $R_{\alpha\alpha}=T\Lambda T^H$. T is orthogonal & unitary matrix & $[I- \mu\Lambda]$ is matrix with diagonal elements. The n'th diagonal of $[I- \mu\Lambda]$ is given by the mode $(1-\mu\lambda_n)$.

$$\upsilon(n+1)=T[I- \mu\Lambda]\, T^H\upsilon(n)$$

Pre-multiplying by T^H. We get: $T^H\upsilon(n+1)=[I- \mu\Lambda]\, T^H\upsilon(n)$

We define: $\acute{\upsilon}(n) = T^H\upsilon(n)$

$$\acute{\upsilon}(n+1)=[I- \mu\Lambda]\acute{\upsilon}(n) \tag{11}$$

$$\|\acute{\upsilon}(n)\|^2 =0, \text{ if n approaches to } \infty.$$

$$\|\acute{\upsilon}(n+1)\|^2 =\sum_{i=0}^{N}(1-\mu\lambda_i)^2\|\upsilon(n)\|^2 \tag{12}$$

If $(1-\mu\lambda_i)^2<1$, then required condition is satisfied. $\|\acute{\upsilon}(n)\|^2 =0$, if n approaches to ∞.

Iff $$0 < \mu <\frac{2}{\lambda_i}$$

For universal bound:

$$0 < \mu < \frac{2}{\lambda_{max}} \tag{13}$$

Step size parameter (μ) plays critical role in the stability of SD algorithm.

Since λ_{max} is not known and difficult to estimate, it can further be restricted by:

$$\lambda_{max} \leq Tr[R_{\alpha\alpha}] = NE[\alpha^2(n)] \leq NS_{max}$$

While the power $E[\alpha^2(n)]$ estimation or the maximum power spectral density S_{max} is moderately easy. It can further be extended:

$$0 < \mu < \frac{2}{Tr[R_{\alpha\alpha}]} \tag{14}$$

$|1 - \mu\lambda_n|$ is close to 0, it corresponds to fast convergence mode.

$|1-\mu\lambda_n|$ is close to 1, it corresponds to slow convergence mode.

Finding optimal step size (μ_o):

The optimal step-size μ_o must satisfy

$$(1- \mu_o\lambda_{min}) = -(1- \mu_o\lambda_{max})$$

$$\mu_o = \frac{2}{\lambda_{max} + \lambda_{min}} \tag{15}$$

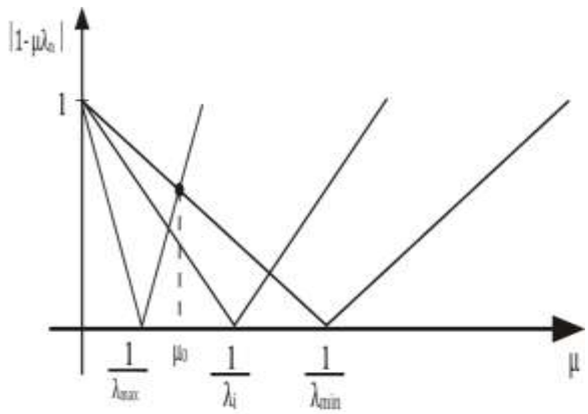

Fig.7. Finding the optimal step-size

2.6 CONDITION NUMBER[$\kappa(R_{\alpha\alpha})$]

The slowest modes $\pm(1- \mu_o \lambda_{max})$ and $\pm(1- \mu_o \lambda_{min})$ for optimal step size is given by:

$$\pm \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} = \pm \frac{\kappa(R_{\alpha\alpha}) - 1}{\kappa(R_{\alpha\alpha}) + 1} \tag{16}$$

where $\kappa(R_{\alpha\alpha}) = \left|\frac{\lambda_{max}}{\lambda_{min}}\right|$ is known as the condition number for the correlation matrix $R_{\alpha\alpha}$. Large condition number leads to the slowest mode approaches to 1and vice versa. When slowest modes are near 0, rate of convergence of SD algorithm is fast.

$$\varepsilon(n) = [d(n) - y(n)] = d(n) - w^T(n)\alpha(n)$$

$$\varepsilon(n) = \varepsilon_{opt}(n) - \Delta(n)^T\alpha(n) \tag{17}$$

$\varepsilon_{opt}(n)$ is the error at $w = w_{opt}$ point.

$$E[\varepsilon^2(n)] = E[\varepsilon^2_{opt}(n)] - 2E[\Delta(n)^T\alpha(n)\varepsilon_{opt}(n)] + E[\{\Delta(n)^T\alpha(n)\}^2] \tag{18}$$

Using Statistically Independent Assumption:

$$E[\varepsilon^2(n)] = E[\varepsilon^2_{opt}(n)] - 2E[\Delta(n)^T]E[\alpha(n)\varepsilon_{opt}(n)] + E[\{\Delta(n)^T\alpha(n)\}^2]$$

Since $\alpha(n)$ &$\varepsilon_{opt}(n)$ are un-correlated and orthogonal to each other. Hence, $E[\alpha(n)\varepsilon_{opt}(n)] = 0$

$$E[\varepsilon^2(n)] = E[\varepsilon^2_{opt}(n)] + E[\{\Delta(n)^T\alpha(n)\}\{\Delta(n)^T\alpha(n)\}^T] \tag{19}$$

$$E[\varepsilon^2(n)] = E[\varepsilon^2_{opt}(n)] + Tr\{E[\Delta(n)\Delta(n)^T]R_{\alpha\alpha}\} \tag{20}$$

Here we define: $E[\Delta(n)\Delta(n)^T] = \mathcal{K}(n)$, $\mathcal{K}(n)$ is Weight Error Co-variance Matrix.

$$E[\varepsilon^2(n)] = E[\varepsilon^2_{opt}(n)] + Tr[\mathcal{K}(n)R_{\alpha\alpha}]$$

Using property of trace of a matrix:

$$E[\varepsilon^2(n)] = E[\varepsilon^2_{opt}(n)] + Tr[T^H\mathcal{K}(n)T\Lambda] \tag{21}$$

We define: $\mathbf{K}(n) = T^H\mathcal{K}(n)T$

$$E[\varepsilon^2(n)] = E[\varepsilon^2_{opt}(n)] + \sum_{i=0}^{N}\mathbf{K}_{ii}(n)\lambda_i$$

Now, we know:

$$\Delta(n+1) = \Delta(n) + \mu\alpha(n)[d(n) - \alpha^H(n)w(n)]$$

$$\Delta(n+1) = \Delta(n) + \mu\alpha(n)[d(n) - \alpha^H\{\Delta(n) + w_{opt}]$$

$$\Delta(n+1) = [I - \mu\alpha(n)\alpha^H(n)]\Delta(n) + \mu\alpha(n)\varepsilon_{opt} \tag{22}$$

Here, we define: $T^H[\Delta(n+1)] = \Delta'(n+1)$, $T^H[\Delta(n)] = \Delta'(n)$, $T^H[\alpha(n)] = \alpha'(n)$,

$$\Delta'(n+1) = [I - \mu\alpha'(n)\alpha'^H(n)]\Delta'(n) + \mu\alpha'(n)\varepsilon_{opt} \tag{23}$$

We want to find out:

$$\mathbf{K}(n) = T^H\mathcal{K}(n)T = E[\Delta'(n+1)\Delta'^H(n+1)] \tag{24}$$

2.7 Excess Mean Square Error (EMSE):

$$\xi^2_{minex}(\infty) = \xi^2_{min}\frac{f(\mu)}{1 - f(\mu)}$$

Assuming input symbols and desired signal are jointly Gaussian. Now, using Matrix Inversion Lemma:

$$\xi^2_{minex}(\infty) \approx \mu\ \xi^2_{min}\left[\frac{Tr[R_{\alpha\alpha}]}{2 - \mu Tr[R_{\alpha\alpha}]}\right] \tag{25}$$

If μ is taken small,

$$\xi^2_{minex}(\infty) \approx \frac{\mu}{2}\,\mathrm{Tr}[R_{\alpha\alpha}]\xi^2_{min} \tag{26}$$

EMSE, the mean-square deviation (MSD), and the misadjustment are calculated when LMS algorithm is in steady state.

2.8 MSE Misadjustment, M(∞)

$$M(\infty) = \frac{\xi^2_{minex}(\infty)}{\xi^2_{min}(w)} \approx \frac{\mu}{2}\mathrm{Tr}[R_{\alpha\alpha}] \tag{27}$$

$\xi^2_{minex}(\infty)$ denotes steady state excess mean error.

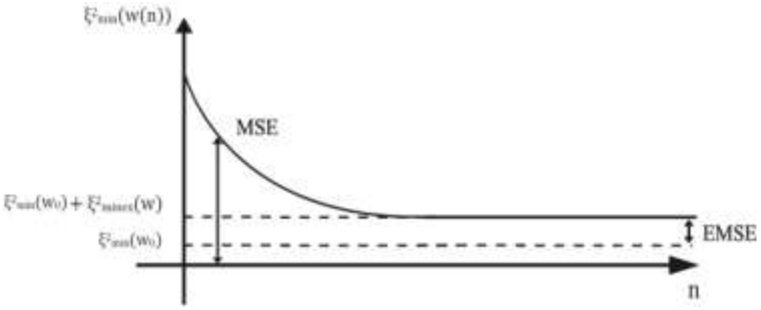

Fig. 8. EMSE & MSE Cost Function

2.9 Mean Square Deviation (MSD)

$$E[||\Delta(n=\infty)||^2] = \frac{\xi^2_{min}}{1-f(\mu)} \sum_{n=1}^{N} \frac{1}{1-\mu\lambda_n} \tag{28}$$

These approximations are valid for considering small value of step size (μ). EMSE is approximately proportional to the input signal power. This is an undesirable problem and termed as gradient noise amplification.

$$E[||\Delta(n=\infty)||^2] \approx \frac{\mu}{2}\xi^2_{min} N \tag{29}$$

3. Simulation Results & Conclusion

QAM reception of LMS algorithm has been simulated for μ = 0.030 at SNR = 25 dB. Convergence rate of LMS for various values of μ has been simulated. It is observed convergence increases up to a certain point and LMS algorithm becomes unstable beyond that point. QAM reception for different values of step sizes and SNR has also been simulated to verify the robustness of LMS. It is observed that LMS is easy to implement than analyze.

International Journal of Advanced Technology in Engineering and Science
Vol. No. 09, Issue No. 02, February 2021
www.ijates.com
ijates
ISSN 2348 - 7550

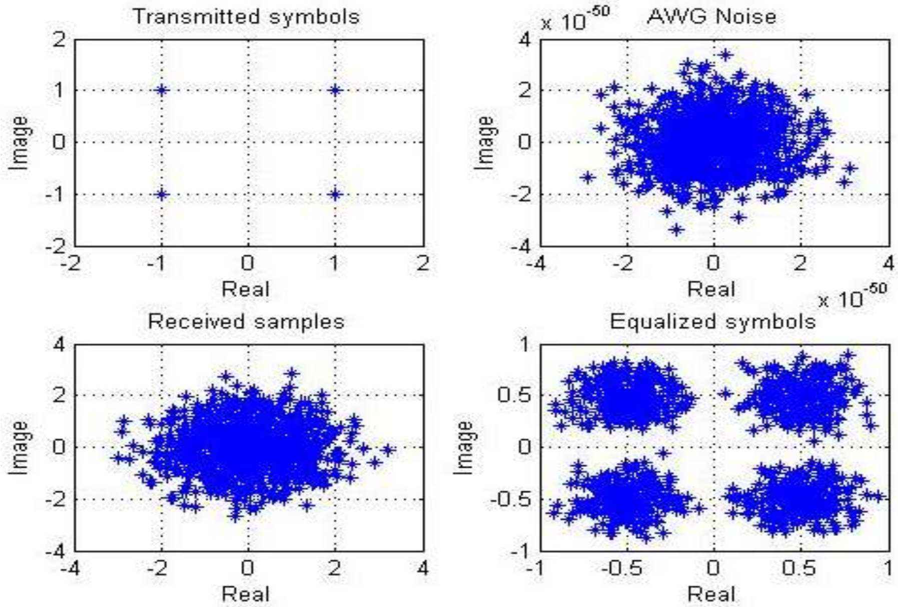

Fig. 9. 4-QAM Reception of LMS Algorithm in noise at SNR=25 dB & μ=0.03

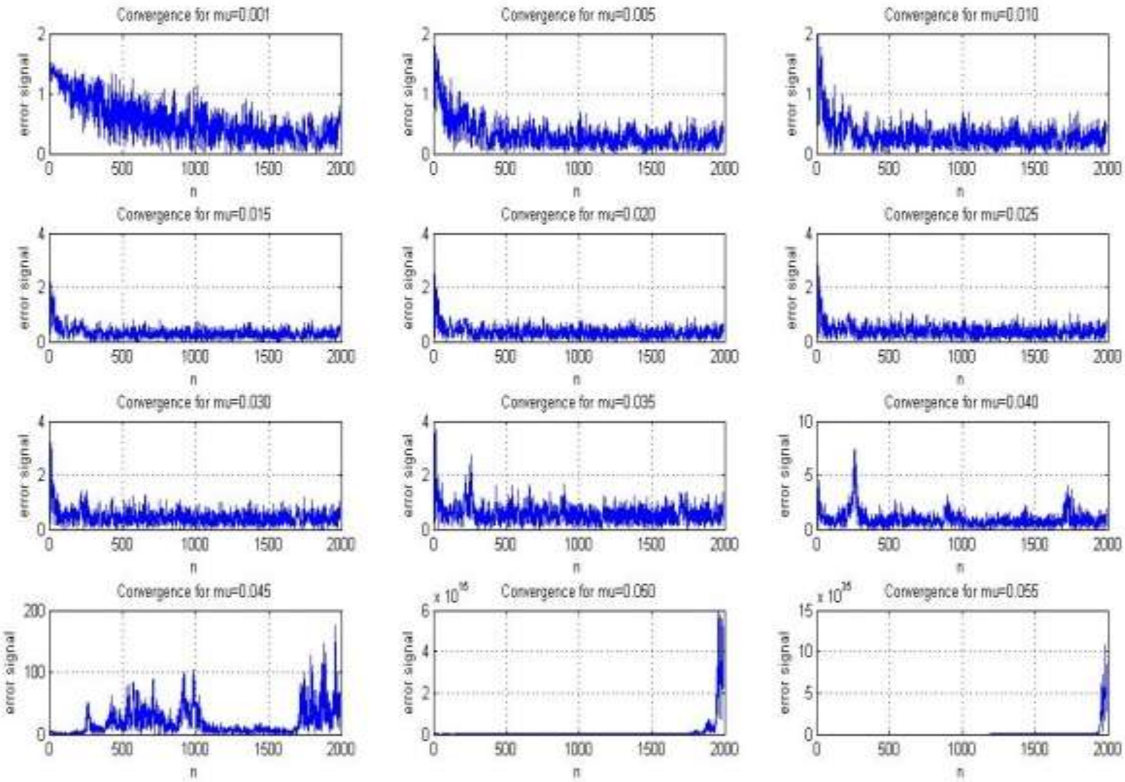

Fig. 10. Convergence behaviour of LMS Algorithm for different values of step sizes (μ)

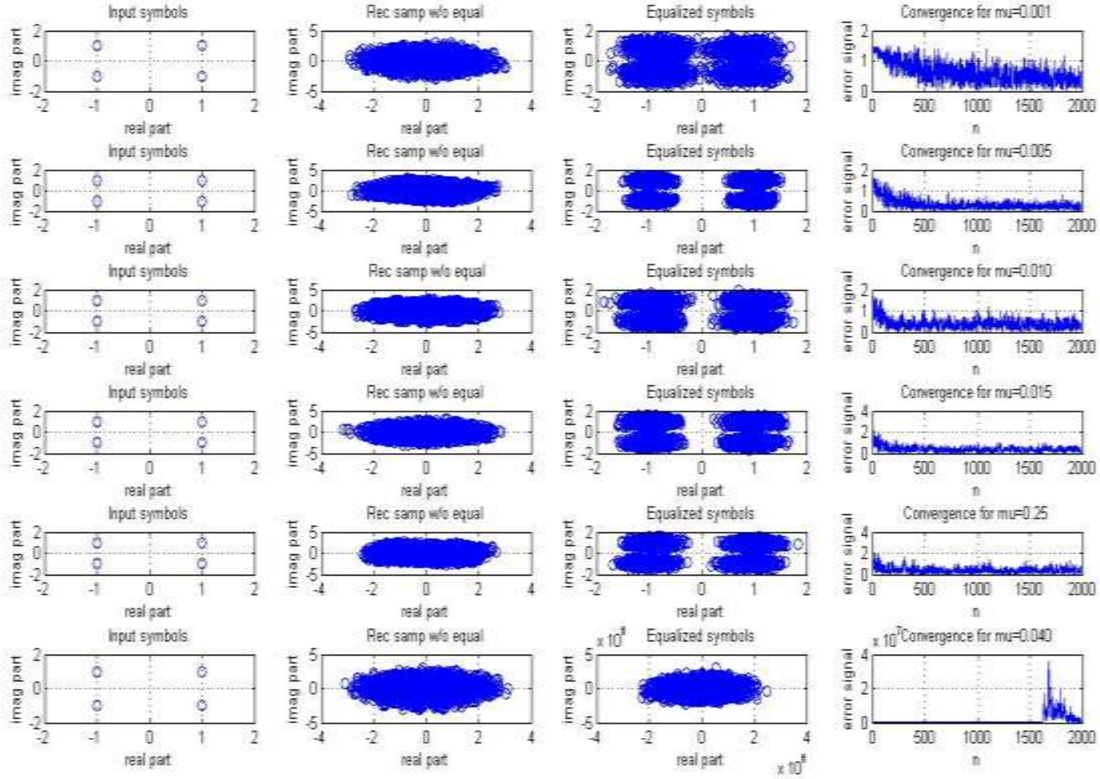

Fig. 11. 4-QAM reception of LMS Algorithm for different values of step sizes (μ)

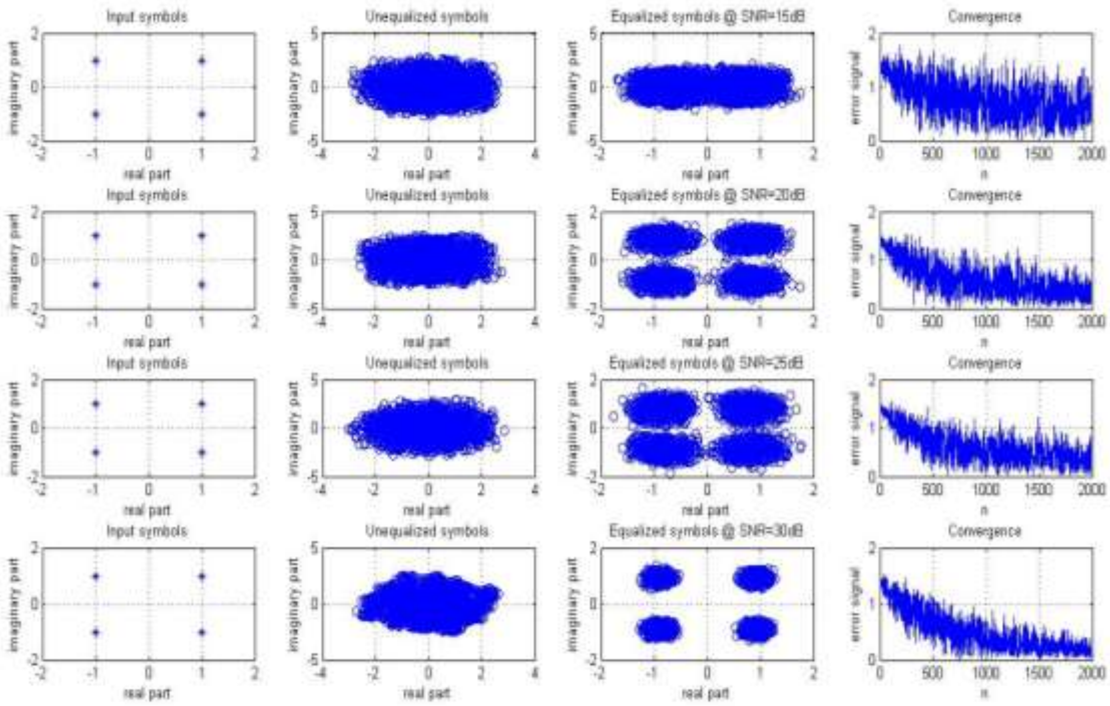

Fig. 12. 4-QAM reception of LMS Algorithm for different values of SNR

International Journal of Advanced Technology in Engineering and Science
Vol. No. 09, Issue No. 02, February 2021
www.ijates.com
ijates
ISSN 2348 - 7550

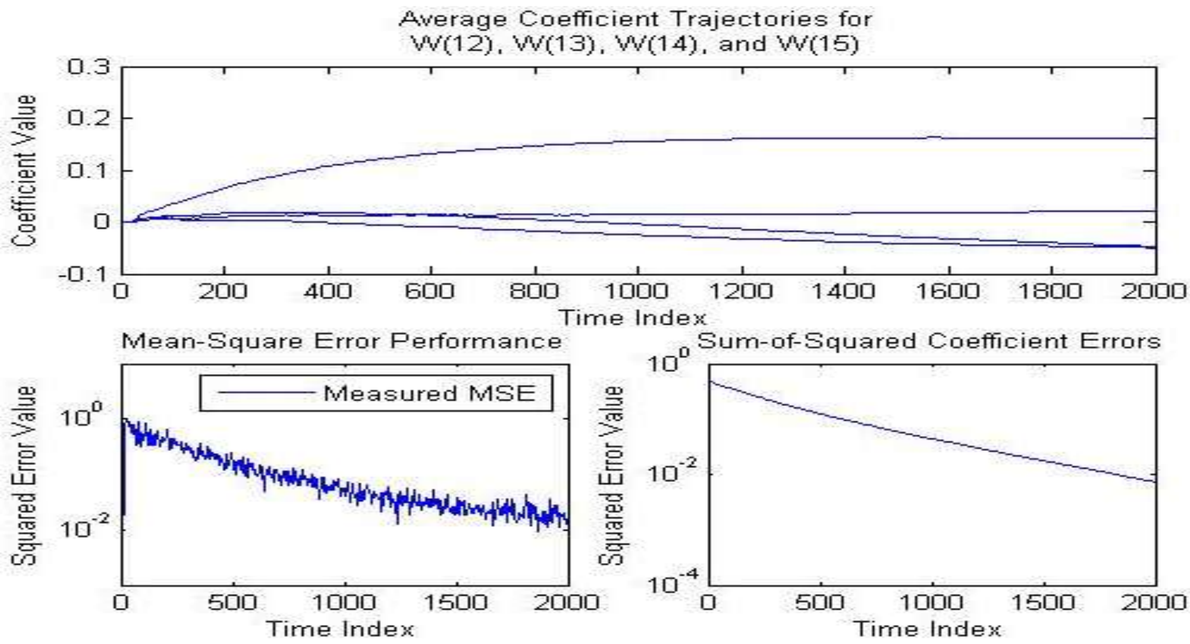

Fig. 13. Average Coefficient Trajectories, MSE and Squared Error of LMS for μ = 0.005

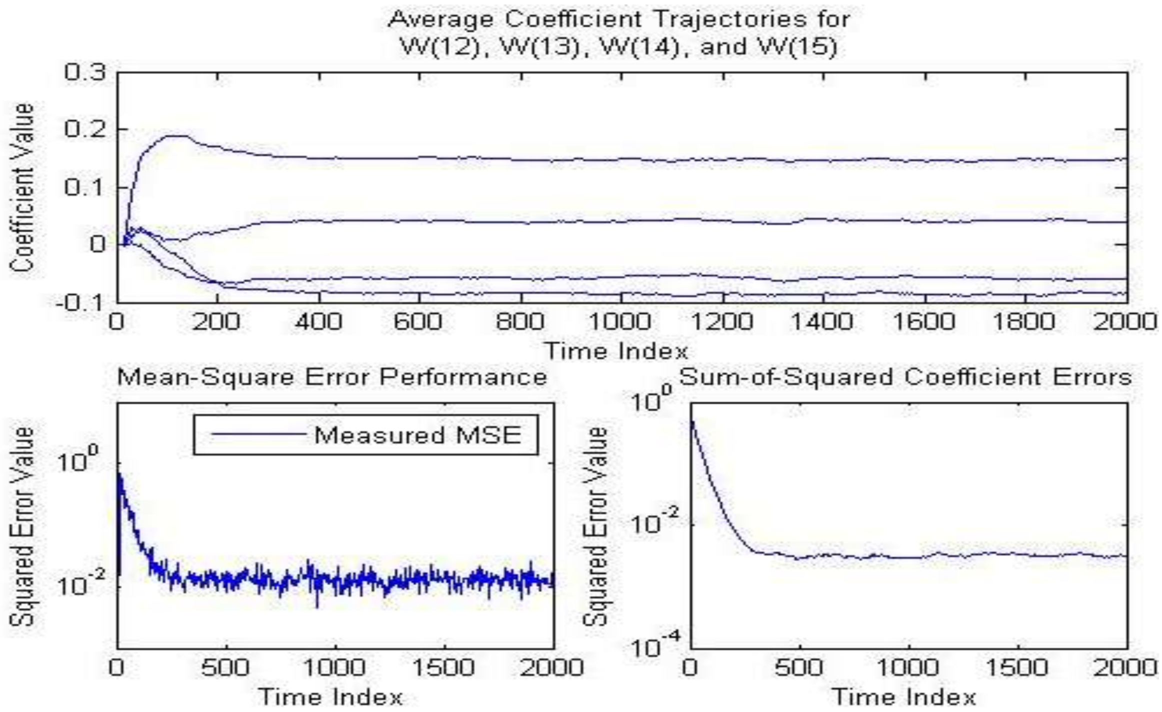

Fig. 14. Average Coefficient Trajectories, MSE and Squared Error of LMS for μ = 0.020

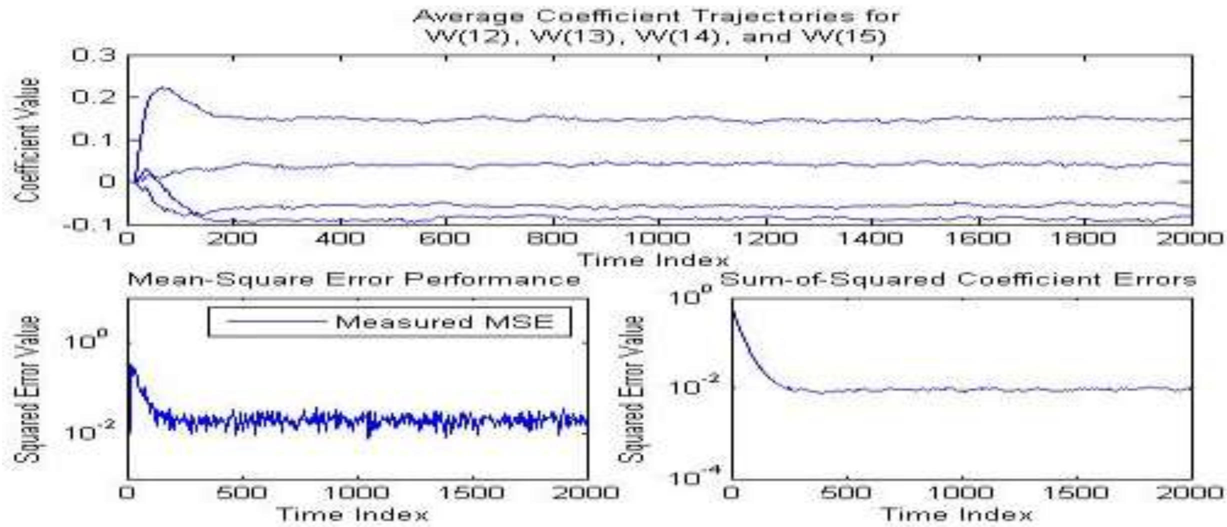

Fig. 15. Average Coefficient Trajectories, MSE and Squared Error of LMS for μ = 0.035

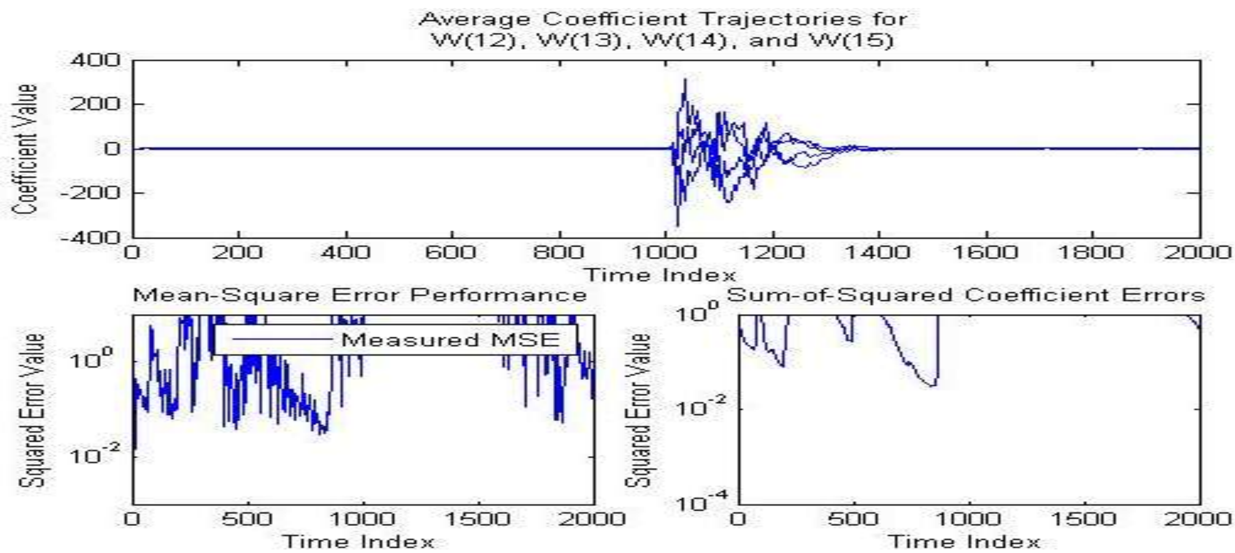

Fig. 16. Average Coefficient Trajectories, MSE and Squared Error of LMS for μ = 0.045

4. REFERENCES

[1] B. Widrow and M.E. Hoff, Jr., "Adaptive switching circuits," in *Proc. IRE WESCON Conf. Rec., part 4, 1960, 96–104.*

[2] Hayes and H. Monson, "Recursive Least Squares," *Statistical Digital Signal Processing & Modeling, Wiley,1996, 541-565.*

[3] Sharma, Prachi, P. Gupta, and P. K. Singh, "Performance Comparison of ZF, LMS and RLS Algorithms for Linear Adaptive Equalizer," *International Journal of Advanced Computer Science and Applications 2(3), 2016.*

[4] V. Ranganathan and K.A.Narayanan kutty, "ADMM Based LMS Algorithm For Constant Modulus Signals In Adaptive Beamforming," *IEEE Trans. on Signal Processing, 2017.*

[5] Zhang, Haijun, Y. Shi, and A. S. Mehr, "Robust equalisation for inter symbol interference communication channels," *Signal Processing, IET 6(2), 2012, 73-78.*

[6] G. Proakis, Digital Communications, 5th Ed., New York, NY, McGraw-Hill," 2005.

[7] Yuxin Chen, Tho Le-Ngoc, Benoit Champagne, and Changjiang Xu, "Recursive Least Squares Constant Modulus Algorithm for Blind Adaptive Array," IEEE Trans. on Signal Processing, 52(5), 2004.

[8] S. Haykin, "Adaptive Filter Theory", Pearson Education India, fourth edition, 2011.

[9] Said E. El-Khamy, M. Shokry, "Enhanced Performance of Blind and Non-Blind Adaptive Arrays using Wavelet Beamforming," *IEEE Trans. Acoust., Speech, Signal Process., vol.4, 2013, 347–351*.

[10] S. Soni, A. Agarwal, "Analysis and Synthesis of Adaptive Equalization Techniques under various Modulation Techniques," *International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE), 2015, 102-106.*

[11] D. C. Lay, Linear algebra and its applications, 3rd ed. Pearson/Addison-Wesley, 2005.

[12] G. Strang, Introduction to Linear Algebra, 4th ed. Wellesley Cambridge Press, 2009.

[16] D. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," *IEEE Trans. Signal Process., 39(1), 1991, 92–114.*

[18] J. C. M. Bermudez and M. H. Costa, "A Statistical Analysis of the "ε-NLMS and NLMS Algorithms for Correlated Gaussian Signals," Journal of the Brazilian Telecommunications Society, 20(2), 2005, 7-13.