STATISTICAL ANALYSIS OF THE LMS ALGORITHM WITH A ZERO-MEMORY NONLINEARITY AFTER THE ADAPTIVE FILTER

Márcio H. Costa[1], *José C. M. Bermudez*[2] *and Neil J. Bershad*[3]

[1]Biomedical Instrumentation Group, Universidade Católica de Pelotas, Pelotas-RS, Brazil*
[2]Department of Electrical Engineering, Universidade Federal de Santa Catarina, Florianópolis-SC, Brazil
[3]Department of Electrical and Computer Engineering, University of California, Irvine, CA, USA
E-mails: costa@atlas.ucpel.tche.br, bermudez@fastlane.com.br, bershad@ece.uci.edu

ABSTRACT

This paper presents a statistical analysis of the Least Mean Square (LMS) algorithm when a zero-memory nonlinearity appears at the adaptive filter output. The nonlinearity is modelled by a scaled error function. Deterministic nonlinear recursions are derived for the mean weight and mean square error (MSE) behavior for white gaussian inputs and slow adaptation. Monte Carlo simulations show excellent agreement with the behavior predicted by the theoretical models. The analytical results show that a small nonlinear effect has a significant impact on the converged MSE.

1. INTRODUCTION

The LMS algorithm is the most popular algorithm for real-time adaptive system implementations. It has been employed in many areas, such as modelling, control, beamforming and equalization. In recent years, this algorithm and its variants have often been applied to active noise and vibration control (ANC) [1]. The behavior of the LMS algorithm in ANC sytems can be affected by distortions caused by the power amplifiers (including saturation) and the nonlinear behavior of the loudspeakers and transducers.

Several researchers have studied the statistical behavior of the LMS algorithm with nonlinearities applied to the error or data signals [2], [3], [4]. These results, however, do not explain the behavior of the algorithm with a nonlinearity at the adaptive filter output.

This paper investigates the statistical behavior of the system in Fig.1. $g(y)$ is a zero-memory saturation-type nonlinearity. This block diagram could model the nonlinear effects of wide-band acoustical transducers in ANC systems, for example.

Deterministic nonlinear recursions are derived for the mean weight and MSE behaviors for white gaussian inputs, slow adaptation and the degree of nonlinearity (DN) for the saturation nonlinearity. Here DN is defined as the ratio of the power at the input to the nonlinearity divided by σ^2, the saturation parameter of $g(y)$.

As DN approaches zero, the well-known linear system identification equations are obtained from the new recursions. The results show that very small DN can substantially affect the algorithm behavior and the achievable level of cancellation. The theoretical predictions show excellent agreement with Monte Carlo simulations.

2. ANALYSIS

2.1. The Analysis Model

$\mathbf{W^o} = \left[w_0^o \quad w_1^o \quad \dots \quad w_{N-1}^o \right]^T$ is the impulse response (N-dimensional) to be identified in Fig.1; $\mathbf{W}(n) = \left[w_0(n) \quad w_1(n) \quad \dots \quad w_{N-1}(n) \right]^T$ is the adaptive weight vector; $d(n)$ is the primary signal; $x(n)$ is stationary, white, zero mean and gaussian with variance σ_x^2; $\mathbf{X}(n) = \left[x(n) \quad x(n-1) \quad \dots \quad x(n-N+1) \right]^T$ is the observed data vector; $z(n)$ is the measurement noise, which is stationary, white, gaussian, zero mean with variance σ_z^2 and uncorrelated to any other signal; and $e(n)$ is the error signal. The nonlinearity is modelled by the function

$$g(y) = \int_0^y e^{-\frac{z^2}{2\,\sigma^2}} dz$$

Note that $\lim_{\sigma \to \infty} g(y) = y$ and $\lim_{\sigma \to 0} g(y) = \sigma \sqrt{\frac{\pi}{2}} \operatorname{sgn}(y)$. Hence, the behavior of $g(y)$ can be varied between that of a linear device and that of a hard limiter by changing σ.

The LMS weight update equation is:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu\, e(n)\, \mathbf{X}(n) \tag{1}$$

where:

$$\begin{aligned} e(n) &= d(n) - y_q(n) \\ &= \mathbf{W^o}^T \mathbf{X}(n) + z(n) - g\left[\mathbf{W}^T(n) \mathbf{X}(n) \right] \end{aligned} \tag{2}$$

* This work was supported in part by CAPES (Brazilian Ministry of Education) under Grant No. PICDT 0129/97-9, and by CNPq (Brazilian Ministry of Science and Technology) under Grant No. 352084/92-8.

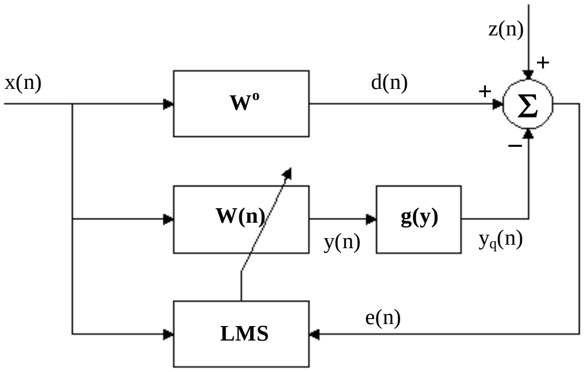

Figure 1. Block diagram of the system analyzed.

resulting in:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu \left\{ \mathbf{W}^{\mathbf{o}T}\mathbf{X}(n) + z(n) - \mathrm{g}\left[\mathbf{W}^T(n)\mathbf{X}(n)\right] \right\} \mathbf{X}(n) \tag{3}$$

$x(n)$ is assumed white and gaussian, $E\left\{\mathbf{X}(n)\mathbf{X}^T(n)\right\} = \sigma_x^2\,\mathbf{I}$.

2.2. Mean Weight Behavior

Taking the expectation of (3) conditioned on $\mathbf{W}(n)$ yields:

$$E\left\{\mathbf{W}(n+1) \,\middle|\, \mathbf{W}(n)\right\} = \mathbf{W}(n) + \mu\,\sigma_{\mathrm{x}}^2\,\mathbf{W}^{\mathbf{o}} - \mu\,E\left\{g\left[\mathbf{W}^T(n)\mathbf{X}(n)\right]\mathbf{X}(n) \,\middle|\, \mathbf{W}(n)\right\} \tag{4}$$

The expectation of the nonlinear term can be evaluated using Bussgang's theorem [5], yielding:

$$E\left\{g\left[\mathbf{W}^T(n)\mathbf{X}(n)\right]\mathbf{X}(n) \,\middle|\, \mathbf{W}(n)\right\} = \frac{\sigma_x^2\,\sigma}{\sqrt{\sigma_x^2\,\mathbf{W}^T(n)\,\mathbf{W}(n) + \sigma^2}}\mathbf{W}(n) \tag{5}$$

Substituting (5) in (4) we have:

$$E\left\{\mathbf{W}(n+1) \,\middle|\, \mathbf{W}(n)\right\} = \mu\,\sigma_{\mathrm{x}}^2\,\mathbf{W}^{\mathbf{o}} + \left[1 - \mu\frac{\sigma_x^2\,\sigma}{\sqrt{\sigma_x^2\,\mathbf{W}^T(n)\,\mathbf{W}(n) + \sigma^2}}\right]\mathbf{W}(n) \tag{6}$$

Assuming μ sufficiently small so that the weights change slowly, the fluctuations of $\mathbf{W}(n)$ about $E\{\mathbf{W}(n)\}$ have a negligible effect on the average behavior of the weights over time [6]. Thus, (6) can be approximated by:

$$E\{\mathbf{W}(n+1)\} = \mu\,\sigma_{\mathrm{x}}^2\,\mathbf{W}^{\mathbf{o}} + \left[1 - \mu\frac{\sigma_x^2\,\sigma}{\sqrt{\sigma_{\mathrm{x}}^2\,\mathrm{E}\left\{\mathbf{W}^T(n)\right\}\mathrm{E}\{\mathbf{W}(n)\} + \sigma^2}}\right]E\{\mathbf{W}(n)\} \tag{7}$$

Now DN = $\eta^2 = \dfrac{\sigma_x^2\,\mathbf{W}^{\mathbf{o}T}\mathbf{W}^{\mathbf{o}}}{\sigma^2}$ (7a)

Recursion (7) reduces to the well-known LMS mean weight recursion [7 - Ch. 6], [8] as the DN approaches zero. Using the orthogonality principle ($\lim_{n\to\infty} E\{e(n)\mathbf{X}(n)\} = 0$), (2) and (5), it can be shown that

$$\lim_{n\to\infty} E\{\mathbf{W}(n)\} = \frac{1}{\sqrt{1-\eta^2}}\mathbf{W}^{\mathbf{o}} = k\,\mathbf{W}^{\mathbf{o}} \tag{8}$$

Hence, the mean weights converge to a scaled version of the unknown system response. As $\eta^2 \to 1$, the mean converged weights will grow without bound. Eq. (7) is unstable and has no stationary points for $\eta^2 > 1$. As $\sigma \to \infty$, $\eta^2 \to 0$ and $k \to 1$ as expected for the linear case.

If $\mathbf{W}(0) = \mathbf{0}$, the solution of (7) is given by

$$E\{\mathbf{W}(n)\} = k(n)\,\mathbf{W}^{\mathbf{o}} \quad \text{for all } n \tag{9}$$

Thus, using (9) in (7), a family of recursions in the scale factor $k(n)$ can be defined as

$$k(n+1) = \mu\,\sigma_x^2 + \left[1 - \mu\frac{\sigma_x^2}{\sqrt{\eta^2 k^2(n) + 1}}\right]k(n) \tag{10}$$

Eq. (10) has stationary points at $k = \dfrac{1}{\sqrt{1-\eta^2}}$. If $\eta^2 \geq 1$, (10) is unstable and $k(n)$ does not converge. The behavior of (10) is parametrized on η and is essentially independent of $\mathbf{W}^{\mathbf{o}}$.

2.3. Mean Square Error

Squaring (2) and taking its expectation conditioned on $\mathbf{W}(n)$ yields:

$$\begin{aligned} E\left\{e^2(n) \,\middle|\, \mathbf{W}(n)\right\} &= \sigma_x^2\,\mathbf{W}^{\mathbf{o}T}\mathbf{W}^{\mathbf{o}} + \sigma_z^2 \\ &\quad - 2\mathbf{W}^{\mathbf{o}T} E\left\{g\left[\mathbf{W}^T(n)\mathbf{X}(n)\right]\mathbf{X}(n) \,\middle|\, \mathbf{W}(n)\right\} \\ &\quad + E\left\{g^2\left[\mathbf{W}^T(n)\mathbf{X}(n)\right] \,\middle|\, \mathbf{W}(n)\right\} \end{aligned} \tag{11}$$

The first expectation is given by (5). The second one can be evaluated following the steps in [9-Appendix I] as:

$$E\left\{g^2\left[\mathbf{W}^T(n)\mathbf{X}(n)\right] \,\middle|\, \mathbf{W}(n)\right\} = \sigma^2 \arcsin\left(\frac{\sigma_x^2\,\mathbf{W}^T(n)\,\mathbf{W}(n)}{\sigma_x^2\,\mathbf{W}^T(n)\,\mathbf{W}(n)+\sigma^2}\right) \tag{12}$$

Using (5) and (12), (11) becomes:

$$E\left\{e^2(n) \,\middle|\, \mathbf{W}(n)\right\} = \sigma_x^2\,\mathbf{W}^{\mathbf{o}T}\mathbf{W}^{\mathbf{o}} + \sigma_z^2 - 2\frac{\sigma_x^2\,\sigma}{\sqrt{\sigma_x^2\,\mathbf{W}^T(n)\,\mathbf{W}(n)+\sigma^2}}\mathbf{W}^{\mathbf{o}T}\mathbf{W}(n) + \sigma^2 \arcsin\left(\frac{\sigma_x^2\,\mathbf{W}^T(n)\,\mathbf{W}(n)}{\sigma_x^2\,\mathbf{W}^T(n)\,\mathbf{W}(n)+\sigma^2}\right) \tag{13}$$

For μ sufficiently small (slow adaptation), (13) reduces to:

$$\zeta(n) = E\left\{e^2(n)\right\} = \sigma_x^2\,\mathbf{W}^{\mathbf{o}T}\mathbf{W}^{\mathbf{o}} + \sigma_z^2 - 2\frac{\sigma_x^2\,\mathbf{W}^{\mathbf{o}T}}{\sqrt{\sigma_x^2\,\mathrm{E}\left\{\mathbf{W}^T(n)\right\}\mathrm{E}\left\{\mathbf{W}(n)\right\}+\sigma^2}}E\left\{\mathbf{W}(n)\right\} + \sigma^2 \arcsin\left(\frac{\sigma_x^2\,\mathrm{E}\left\{\mathbf{W}^T(n)\right\}\mathrm{E}\left\{\mathbf{W}(n)\right\}}{\sigma_x^2\,\mathrm{E}\left\{\mathbf{W}^T(n)\right\}\mathrm{E}\left\{\mathbf{W}(n)\right\}+\sigma^2}\right) \tag{14}$$

where $E\left\{\mathbf{W}(n)\right\}$ can be evaluated using (7). Here again, it can be easily verified that (11) reduces to the conventional LMS MSE recursion for slow adaptation as $\sigma \to \infty$. Also, for $\mathbf{W}(0) = \mathbf{0}$, (9) can be used in (14) leading to

$$\zeta(n) = E\left\{e^2(n)\right\} = \sigma_x^2\,\mathbf{W}^{\mathbf{o}T}\mathbf{W}^{\mathbf{o}}\left\{1 - \frac{2}{\sqrt{\eta^2 k^2(n)+1}}k(n) + \frac{1}{\eta^2}\arcsin\left(\frac{\eta^2 k^2(n)}{\eta^2 k^2(n)+1}\right)\right\} + \sigma_z^2 \tag{15}$$

3. SIMULATIONS

This section presents simulations which verify the accuracy of the analytical models (7) and (14), as well as (9), (10) and (15). Consider the system in Fig. 1 with the following parameters:

$\mathbf{W}^{\mathbf{o}} = [0.4130 \quad 0.4627 \quad 0.4803 \quad 0.4627 \quad 0.4130]^T$,

$\mathbf{W}^{\mathbf{o}T}\mathbf{W}^{\mathbf{o}} = 1$, $\mathbf{W}(0) = \mathbf{0}$, $\mu = 0.01$, $\sigma_x^2 = 1$ and $\sigma_z^2 = 10^{-6}$.

Fig. 2 shows the mean behavior of the third weight for $\sigma^2 = 2$, 4 and 1000 ($\eta^2 = 0.5$, 0.25 and 0.001). These cases correspond to a large DN, a small DN and nearly linear system. The theoretical curves using (7) (continuous curves) and Monte Carlo simulations (10 runs) (ragged curves) show excellent agreement. The behavior of the other weights are similar. The asymptotes (derived from (8)) are 0.679, 0.554 and 0.481, in agreement with the curves in Fig. 2.

Fig. 3 shows the MSE behavior for σ^2 = 2, 15 and 1000 ($\eta^2 = 0.5$, 0.067 and 0.001). Again, there is excellent agreement between theory (Eq. (14)) and simulation results (averaged over 100 runs). Note the dramatic reduction in converged cancellation performance for the nonlinear cases (a) and (b) as compared to the nearly linear case (c)

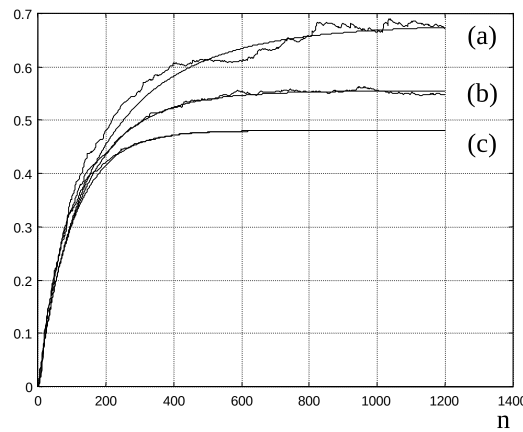

Fig. 2. Mean behavior of the third coefficient. (a) η^2 = 0.5; (b) η^2 = 0.25 and (c) η^2 = 0.001.

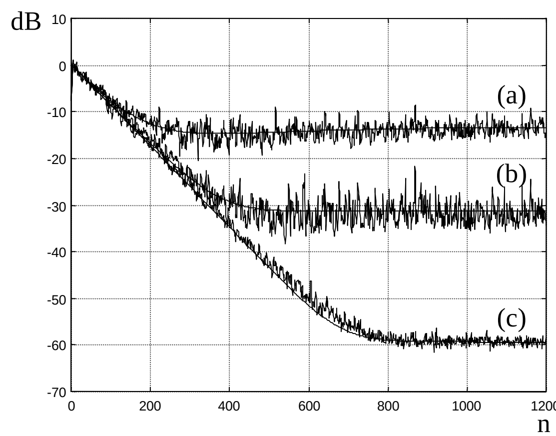

Figure 3. Mean square error. (a) η^2 = 0.5; (b) η^2 = 0.067 and (c) η^2 = 0.001.

Fig. 4 shows the function $g(y)$ and the amplitude histograms of $y(n)$ obtained from the simulations for all 1200 iterations and averaged over 10 runs. Fig. 4 clearly shows that the system was driven into nonlinear operation. Small DN can have a significant effect on the achievable cancellation level as shown by Figures 2-4 (> 20 dB from curves (c) to (b) in Fig. 3). Hence, the present analysis is quite important for design purposes.

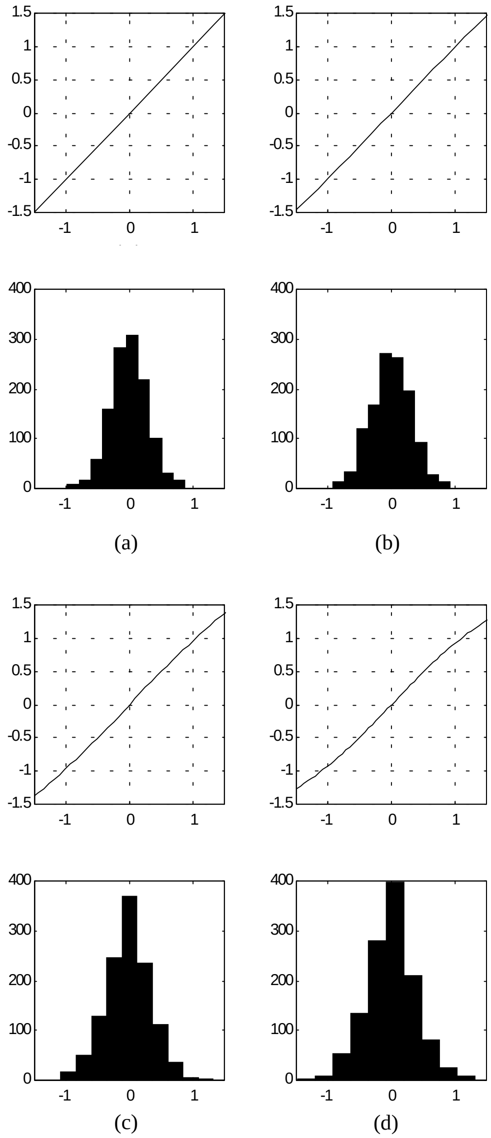

Figure 4. Functions $g(y)$ and respective histograms of $y(n)$ obtained from simulation. (a) η^2 = 0.001, (b) η^2 = 0.067, (c) η^2 = 0.25 and (d) η^2 = 0.5.

4. CONCLUSIONS

This paper presented a statistical analysis of the LMS algorithm with a zero-memory nonlinearity at the adaptive filter output. Recursive equations have been derived for the mean weight and mean square error behaviors for white gaussian inputs and slow adaptation. These expression were shown to be a simple function of the system's DN.

The theoretical expressions predict the statistical behavior of the algorithm during all phases of the adaptation process for all values of DN. The analytical results show that small DN has a significant impact on the achievable cancellation level.

5. REFERENCES

[1] Kuo S. M. and Morgan D.R., Active Noise Control Systems: Algorithms and DSP Implementations, New York: John Wiley, 1996.

[2] Duttweiler D.L. "Adaptive Filter Performance with Nonlinearities in the Correlation Multiplier", IEEE Transactions on Acoustic, Speech and Signal Processing, 30(4):578-586, 1982.

[3] Bershad N.J. "On Error Saturation Nonlinearities in LMS Adaptation", IEEE Transactions on Acoustics, Speech and Signal Processing, 36(4):440-452, 1988.

[4] Douglas S.C. and Meng T.H.Y "Normalized Data Nonlinearities for LMS Adaptation", IEEE Transactions on Signal Processing, 42(6):1352-1365, 1994.

[5] Bussgang J.J., "Cross-correlation functions of amplitude-distorted gaussian signals," Tech Rep. 216, Research Laboratory of Electronics, MIT, Cambridge, MA, March 1952.

[6] Bershad N.J., Shynk J.J., Feintuch P.L. "Statistical Analysis of the Single Layer Backpropagation Algorithm: Part I – Mean Weight Behavior", IEEE Transactions on Signal Processing, 41(2):573-582, 1993.

[7] Widrow B. and Stearns S.D., Adaptive Signal Processing, Prentice-Hall, N.J., 1985

[8] Haykin S. Adaptive Filter Theory, Prentice Hall, second edition, 1991.

[9] Bershad N.J., Shynk J.J., Feintuch P.L. "Statistical Analysis of the Single Layer Backpropagation Algorithm: Part II – MSE and Classification Performance", IEEE Transactions on Signal Processing, 41(2):583-591, 1993.