# A TRUE STOCHASTIC GRADIENT ADAPTIVE ALGORITHM FOR APPLICATIONS USING NONLINEAR ACTUATORS

*Márcio H. Costa[1] and José C. M. Bermudez[2]*

[1]Núcleo de Eng. Biomédica, Escola de Engenharia e Arquitetura, Universidade Católica de Pelotas, Pelotas-RS, Brazil*
[2]Department of Electrical Engineering, Universidade Federal de Santa Catarina, Florianópolis-SC, Brazil
E-mails: costa@atlas.ucpel.tche.br, bermudez@fastlane.com.br

## ABSTRACT

This work considers the practical situation where adaptive systems are subject to a saturation nonlinearity at the output of the adaptive filter. Such is the case in active control of noise and vibration. A new adaptive algorithm is proposed which implements the true stochastic gradient approach to the nonlinear problem. Deterministic nonlinear recursions are derived which model the mean weight and mean square error behaviors. The steady-state behavior is also studied. The practical aspects of nonlinearity estimation and hardware implementation are addressed. It is shown that the new algorithm outperforms the LMS algorithm even for considerable errors in estimating the nonlinearity parameters.

## 1. INTRODUCTION

Most adaptive system analyses and designs assume that nonlinear effects can be neglected. Linearity simplifies the mathematical problem and often permits detailed system analyses and accurate designs in practical circumstances. However, when nonlinear effects are significant in determining system behavior, more elaborate models and algorithms must be used. Such is the case when the electrical signal at the adaptive filter output is converted to a signal of different nature. Then, amplifiers and transducers are required, which can be driven into a nonlinear region of operation. Examples of such situation are active noise control (ANC) and active vibration control (AVC) systems.

Active control of noise and vibration using adaptive filters has attracted increasing interest from researchers and application engineers. Active control consists of attenuating pressure (sound) or vibrational waves in a space region through an artificial acoustic field or vibration of the same intensity and in anti-phase. The Least Mean Square (LMS) adaptive algorithm is widely employed in such systems [1] for its simplicity [2].

ANC and AVC systems present nonlinearities introduced by power amplifiers and transducers such as loudspeakers and piezoelectric actuators [3,4]. Thus, their nonlinear nature must be considered unless they are overdesigned to avoid large signal amplitudes. Many works have addressed the influence of nonlinearities on the behavior of adaptive systems [5-8]. Ref. [6] studied the behavior of an adaptive system with a nonlinearity at the adaptive filter output. The mean-square error (MSE) surface properties derived were used to evaluate the performance of the LMS algorithm under these circumstances (common in ANC systems). It was shown that the LMS algorithm produces a biased estimate of the optimum controller. The steady-state weight vector is a scaled version of the optimum weight vector. The scaling factor depends on the system's degree of nonlinearity. This multiplicative bias occurs because nonlinearity at the adaptive filter output is not used for estimating the gradient of the performance surface. The LMS weight update equation is the same derived for the linear case.

This work proposes a new stochastic gradient algorithm that incorporates the nonlinearity in estimating the gradient of the cost function. Implementation issues and the estimation of the nonlinearity parameters are addressed. Deterministic recursive equations for the mean weight and mean square error (MSE) behavior are developed for gaussian inputs and slow adaptation. Steady-state behavior is derived from these results. Monte Carlo simulations show excellent agreement with the analytical model. It is shown that the new algorithm outperforms the LMS algorithm for realistic error levels in the estimate of the nonlinearity parameter.

## 2. ALGORITHM

Consider the system in Fig. 1. $\mathbf{W^o} = \left[ w_0^o \quad w_1^o \quad \ldots \quad w_{N-1}^o \right]^T$ is the unknown impulse response; $d(n)$ is the primary signal; $z(n)$ is a stationary, white, gaussian and zero-mean measurement noise with variance $\sigma_z^2$ and uncorrelated with any other signal. $\mathbf{W}(n) = \left[ w_0(n) \quad w_1(n) \quad \ldots \quad w_{N-1}(n) \right]^T$ is the adaptive weight vector. $x(n)$ is stationary, zero-mean and Gaussian. $\mathbf{X}(n) = \left[ x(n) \quad x(n-1) \quad \ldots \quad x(n-N+1) \right]^T$ is the observed data vector and $\varepsilon(n)$ is the error signal. The nonlinearity $g(\bullet)$ is modeled by the error function [7]:

$$g(y) = \int_0^y e^{-\frac{z^2}{2\sigma^2}} dz \tag{1}$$

---

* Thanks to CAPES (Brazilian Ministry of Education) under Grant No. PICDT 0129/97-9 and to CNPq (Brazilian Ministry of Science and Technology) under Grant No. 352084/92-8 for funding.

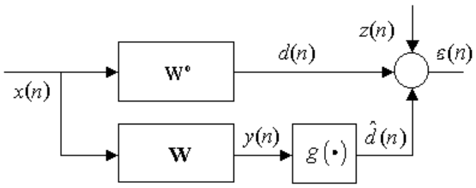


**Figure 1**. Block diagram of the system analyzed.

The behavior of $g(y)$ can be varied between that of a linear device and that of a hard limiter by changing σ and by using a suitable multiplicative constant (considered unit here for simplicity). The error signal is given by:

$$\varepsilon(n) = \mathbf{W}^{\mathbf{o}T}\mathbf{X}(n) + z(n) - g\left(\mathbf{W}^T(n)\mathbf{X}(n)\right) \tag{2}$$

The weight adaptive equation is obtained through:

$$\mathbf{W}(n+1) = \mathbf{W}(n) - \frac{\mu}{2}\nabla_{\mathbf{W}} E\left\{\varepsilon^2(n)\right\} \tag{3}$$

where the gradient of the performance surface $\nabla_{\mathbf{W}} E\left\{\varepsilon^2(n)\right\}$ is estimated by the stochastic gradient $\partial\varepsilon^2(n)/\partial\mathbf{W}$, resulting in:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu\,\varepsilon(n)\mathbf{X}(n)\cdot e^{-\frac{\left(\mathbf{W}^T(n)\mathbf{X}(n)\right)^2}{2\hat{\sigma}^2}} \tag{4}$$

where $\hat{\sigma}$ is an estimate of the parameter $\sigma$.

### 2.1 Practical Implementation Issues

Eq. (4) is the update equation for the proposed adaptive algorithm. Since it incorporates a better estimate of the gradient of the MSE surface, it is expected to perform better than the LMS algorithm in the system of Fig. 1. However, its practical applicability depends on two aspects: (i) How can the parameter $\sigma$ be estimated and good must be its estimate $\hat{\sigma}$ ? (ii) How can (4) be implemented on an actual DSP system ?

Regarding the estimation of $\sigma$, note that

$$\lim_{y\to\infty}\left[g(y)\right] = \sigma\sqrt{\frac{\pi}{2}} = \sigma_{\hat{d}} \tag{5}$$

Thus, assuming the nonlinearity can be approximated by $g(y)$, measuring the maximum output signal $\hat{d}(n)$ in Fig. 1 in a controlled experiment can provide a reasonable estimate $\hat{\sigma}$. The effect of the error in this estimate will be addressed in the following sections.

Regarding a practical implementation, the exponential term in the weight update equation (4) can be implemented using a Taylor series expansion for $e^x$ ($x \geq 0$) [9]. For an *m*-term expansion, the additional computational effort relative to the standard LMS is 2*m*-1 multiplications, *m*-1 additions and one division (to obtain $e^{-x}$ from $e^x$ for better accuracy). Excellent accuracy can be obtained using $m = 6$. Thus, $N \gg m$ for most typical applications, and the required computational complexity is a small price to be paid for the improved cancellation performance.

## 3. MEAN WEIGHT BEHAVIOR

Considering the accuracy of the series implementation, the analysis of the algorithm behavior is based on (4). Note, however, that the effect of the estimate $\hat{\sigma}$ is considered. Taking the expected value of (4) conditioned on $\mathbf{W}(n)$ and noting that $z(n)$ and $x(n)$ are uncorrelated yields:

$$\begin{aligned} E\left\{\mathbf{W}(n+1)\middle|\mathbf{W}(n)\right\} &= \mu E\left\{ e^{-\frac{\left(\mathbf{W}^T(n)\mathbf{X}(n)\right)^2}{2\hat{\sigma}^2}}\mathbf{X}(n)\mathbf{X}^T(n) \middle| \mathbf{W}(n)\right\}\mathbf{W}^{\mathbf{o}} \\ &- \mu\, E\left\{ g\left[\mathbf{W}^T(n)\mathbf{X}(n)\right] e^{-\frac{\left(\mathbf{W}^T(n)\mathbf{X}(n)\right)^2}{2\hat{\sigma}^2}}\mathbf{X}(n)\middle|\mathbf{W}(n)\right\} + \mathbf{W}(n) \end{aligned} \tag{6}$$

Stating: $y = \mathbf{W}^T(n)\mathbf{X}(n)$, and $f(y) = e^{-\frac{y^2}{2\hat{\sigma}^2}}$ the first expectation can be evaluated using [7,Eq. A.13]:

$$\begin{aligned} E\left\{f(y)\mathbf{X}(n)\mathbf{X}^T(n)\right\} &= E\left\{y\mathbf{X}(n)\right\}E\left\{y\mathbf{X}^T(n)\right\}E\left\{y^2\right\}^{-2}E\left\{f(y)y^2\right\} \\ &+ \left[E\left\{\mathbf{X}(n)\mathbf{X}^T(n)\right\} - E\left\{y\mathbf{X}(n)\right\}E\left\{y\mathbf{X}^T(n)\right\}E\left\{y^2\right\}^{-1}\right]E\left\{f(y)\right\} \end{aligned} \tag{7}$$

By integration we obtain:

$$E\left\{f(y)\right\} = \left(\frac{1}{\hat{\sigma}^2}E\left\{y^2\right\}+1\right)^{-1/2} \tag{8}$$

and:

$$E\left\{f(y)y^2\right\} = E\left\{y^2\right\}\left(\frac{E\left\{y^2\right\}}{\hat{\sigma}^2}+1\right)^{-3/2} \tag{9}$$

Substituting (8) and (9) in (7) yields:

$$E\left\{f(y)\mathbf{X}(n)\mathbf{X}^T(n)\right\} = \frac{E\left\{\mathbf{X}(n)\mathbf{X}^T(n)\right\}}{\left(\frac{E\left\{y^2\right\}}{\hat{\sigma}^2}+1\right)^{1/2}} - \frac{E\left\{y\mathbf{X}(n)\right\}E\left\{y\mathbf{X}^T(n)\right\}}{\hat{\sigma}^2\left(\frac{E\left\{y^2\right\}}{\hat{\sigma}^2}+1\right)^{3/2}} \tag{10}$$

Neglecting the correlation between $\mathbf{W}(n)$ and $\mathbf{X}(n)$, the linear expectations in (10) yields:

$$\begin{cases} E\left\{y^2\right\} = \mathbf{W}^T(n)\mathbf{R}_{\mathbf{xx}}\mathbf{W}(n) \\ E\left\{y\mathbf{X}(n)\right\} = \mathbf{R}_{\mathbf{xx}}\mathbf{W}(n) \;\; ; \; E\left\{\mathbf{X}(n)\mathbf{X}^T(n)\right\} = \mathbf{R}_{\mathbf{xx}} \end{cases} \tag{11}$$

Using (11) in (10) we obtain the solution for the first expectation in (5). The second expectation in (5) is obtained by expanding $\mathbf{X}(n)$ in an orthonormal series about $y$, in the same way as (7):

$$E\left\{h(y)\mathbf{X}(n)\right\} = E\left\{h(y)y\right\}E\left\{y\mathbf{X}(n)\right\}E\left\{y^2\right\}^{-1} \tag{12}$$

where $h(y) = g(y)e^{-\frac{y^2}{2\hat{\sigma}^2}}$. Integrating by parts we can obtain:

$$E\left\{h(y)y\right\} = \frac{E\left\{y^2\right\}}{\left(\frac{E\left\{y^2\right\}}{\hat{\sigma}^2}+1\right)\left(\frac{E\left\{y^2\right\}}{\sigma^2}+\frac{E\left\{y^2\right\}}{\hat{\sigma}^2}+1\right)^{1/2}} \tag{13}$$

Using (11) and (13) in (12) results in the solution to the second expectation of (6). Substituting (10), (11) and (13) in (6) and 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 [8]. Thus, the final result for (6) can be approximated by:

$$
\begin{aligned}
E\{\mathbf{W}(n+1)\} &= E\{\mathbf{W}(n)\} + \frac{\mu \mathbf{R}_{\mathbf{xx}} \mathbf{W}^{\mathbf{o}}}{\left(\frac{1}{\hat{\sigma}^2} E\{\mathbf{W}^T(n)\}\mathbf{R}_{\mathbf{xx}} E\{\mathbf{W}(n)\}+1\right)^{1/2}} \\
&- \frac{\mu \mathbf{R}_{\mathbf{xx}} E\{\mathbf{W}(n)\} E\{\mathbf{W}^T(n)\}\mathbf{R}_{\mathbf{xx}} \mathbf{W}^{\mathbf{o}}}{\hat{\sigma}^2\left(\frac{1}{\hat{\sigma}^2} E\{\mathbf{W}^T(n)\}\mathbf{R}_{\mathbf{xx}} E\{\mathbf{W}(n)\}+1\right)^{3/2}} \\
&- \frac{\mu\left(\frac{1}{\hat{\sigma}^2} E\{\mathbf{W}^T(n)\}\mathbf{R}_{\mathbf{xx}} E\{\mathbf{W}(n)\}+1\right)^{-1} \mathbf{R}_{\mathbf{xx}} E\{\mathbf{W}(n)\}}{\left[\left(\frac{1}{\sigma^2}+\frac{1}{\hat{\sigma}^2}\right) E\{\mathbf{W}^T(n)\}\mathbf{R}_{\mathbf{xx}} E\{\mathbf{W}(n)\}+1\right]^{1/2}}
\end{aligned}
\tag{14}
$$

### 3.1 Mean Weight Steady-State Behavior

Assuming algorithm's convergence as $n \to \infty$, and defining $\mathbf{W}_\infty = \lim_{n\to\infty} E\{\mathbf{W}(n)\}$, (14) yields, after some manipulations:

$$
\mathbf{W}^{\mathbf{o}} = \frac{\left(\frac{T}{\hat{\sigma}^2}+1\right)^{1/2} + \frac{T}{\hat{\sigma}^2}\left[\left(\frac{1}{\sigma^2}+\frac{1}{\hat{\sigma}^2}\right)T+1\right]^{1/2}}{\left(\frac{T}{\hat{\sigma}^2}+1\right)\left[\left(\frac{1}{\sigma^2}+\frac{1}{\hat{\sigma}^2}\right)T+1\right]^{1/2}} \mathbf{W}_\infty \tag{15}
$$

where: $T = \mathbf{W}_\infty^T \mathbf{R}_{\mathbf{xx}} \mathbf{W}_\infty$. Since $\mathbf{R}_{\mathbf{xx}}$ is positive semi-definite, (15) shows that $\mathbf{W}_\infty$ is collinear to $\mathbf{W}^{\mathbf{o}}$, *i.e.*: $\mathbf{W}_\infty = p \cdot \mathbf{W}^{\mathbf{o}}$ with $p \in \mathbb{R}^+$. Substituting $p \cdot \mathbf{W}^{\mathbf{o}}$ for $\mathbf{W}_\infty$ in (15) and solving for $p$ leads to:

$$
p^4 + \left(\frac{1}{\hat{\eta}^2} - \frac{\eta^2}{\hat{\eta}^2} - 1\right)p^2 - \frac{1}{\hat{\eta}^2} = 0 \tag{16}
$$

where:

$$
\eta^2 = \frac{1}{\sigma^2}\mathbf{W}^{\mathbf{o}T}\mathbf{R}_{\mathbf{xx}}\mathbf{W}^{\mathbf{o}} \quad \text{and} \quad \hat{\eta}^2 = \frac{1}{\hat{\sigma}^2}\mathbf{W}^{\mathbf{o}T}\mathbf{R}_{\mathbf{xx}}\mathbf{W}^{\mathbf{o}} \tag{17}
$$

$\eta^2$ in (17) is defined as the degree of nonlinearity. Thus, $\hat{\eta}^2$ is the estimate of $\eta^2$ given $\hat{\sigma}^2$. A physical interpretation of $\eta^2$ can be obtained by noting that $\mathbf{W}^{\mathbf{o}T}\mathbf{R}_{\mathbf{xx}}\mathbf{W}^{\mathbf{o}} = \sigma_d^2$ is the power of the signal $d(n)$ in Fig. 1. Also, (5) shows that $\sigma^2$ determines the maximum power at the output of the nonlinearity. Thus,

$$
\eta^2 = \frac{\pi}{2}\frac{\sigma_d^2}{\max\{\sigma_{\hat{d}}^2\}} \tag{18}
$$

which indicates how much the nonlinearity impairs the ability of the adaptive system to provide the necessary cancellation power.

Eq. (16) has only one solution such that $p \in \mathbb{R}^+$. Evaluating this solution yields:

$$
\lim_{n\to\infty} E\{\mathbf{W}(n)\} = \left[\sqrt{\frac{1}{2} + \frac{\eta^2}{2\hat{\eta}^2} - \frac{1}{2\hat{\eta}^2} + \sqrt{\frac{(\eta^2-1)^2}{4\hat{\eta}^4} + \frac{(\eta^2+1)}{2\hat{\eta}^2} + \frac{1}{4}}}\right]\mathbf{W}^{\mathbf{o}} \tag{19}
$$

Note that for $\hat{\eta}^2 = \eta^2$ (perfect estimation), (19) reduces to:

$$
\mathbf{W}_\infty = \left[\sqrt{1 - \frac{1}{2\eta^2} + \sqrt{1 + \frac{1}{4\eta^4}}}\right]\mathbf{W}^{\mathbf{o}} \tag{20}
$$

which corresponds to the optimum weight vector (minimum of the MSE surface) determined in [6]. Thus, the new algorithm leads to an unbiased mean weight vector for perfect estimation of $\sigma$.

## 4. MSE BEHAVIOR

For sufficiently small μ, a simplified model for the MSE behavior can be obtained from the results derived in [5] for the MSE surface of the problem in Fig. 1. Generalizing [5,Eq.14] for correlated signals and considering $\mathbf{W}(n) \approx E\{\mathbf{W}(n)\}$ (small fluctuations) leads to:

$$
\begin{aligned}
\xi(n) &= E\{\varepsilon^2(n)\} = \mathbf{W}^{\mathbf{o}T}\mathbf{R}_{\mathbf{xx}}\mathbf{W}^{\mathbf{o}} + \sigma_z^2 \\
&- \frac{2}{\sqrt{\frac{1}{\sigma^2}E\{\mathbf{W}^T(n)\}\mathbf{R}_{\mathbf{xx}}E\{\mathbf{W}(n)\}+1}}\mathbf{W}^{\mathbf{o}T}\mathbf{R}_{\mathbf{xx}}E\{\mathbf{W}(n)\} \\
&+ \sigma^2 \cdot \operatorname{arcsen}\left(\frac{E\{\mathbf{W}^T(n)\}\mathbf{R}_{\mathbf{xx}}E\{\mathbf{W}(n)\}}{E\{\mathbf{W}^T(n)\}\mathbf{R}_{\mathbf{xx}}E\{\mathbf{W}(n)\}+\sigma^2}\right)
\end{aligned}
\tag{21}
$$

where $E\{\mathbf{W}(n)\}$ is obtained from (14).

### 4.1 MSE Steady-State Behavior

Assuming weight convergence and using $\mathbf{W}_\infty = p \cdot \mathbf{W}^{\mathbf{o}}$ in (21):

$$
\lim_{n\to\infty}\xi(n) = \mathbf{W}^{\mathbf{o}T}\mathbf{R}_{\mathbf{xx}}\mathbf{W}^{\mathbf{o}}\left[1 - \frac{2p}{\sqrt{p^2\eta^2+1}} + \frac{1}{\eta^2}\arcsin\left(\frac{p^2\eta^2}{p^2\eta^2+1}\right)\right] + \sigma_z^2 \tag{22}
$$

## 5. SIMULATIONS

This section presents simulations to verify the accuracy of the analytical model given by (14), (19), (21) and (22). Consider the system in Fig. 1 with the following parameters: $x(n)$ with $\sigma_x^2 = 1$. The eigenvalue spread ($\lambda_{\max}/\lambda_{\min}$) of $\mathbf{R}_{\mathbf{xx}}$ is equal to 10.34. $\sigma_z^2 = 10^{-6}$, $\mu = 0.01$, $\mathbf{W}^{\mathbf{o}T}\mathbf{W}^{\mathbf{o}} = 1$, $\mathbf{W}(0) = \mathbf{0}$ and $\mathbf{W}^{\mathbf{o}} = [0.4130 \quad 0.4627 \quad 0.4803 \quad 0.4627 \quad 0.4130]^T$. The exponential function in the weight update equation is implemented in the simulation by a 6-term Taylor expansion [9].

Fig. 2 shows the mean behavior of the third weight for $\eta^2 = 0.0001$, 0.1, 0.3, 0.5 and 2 with respective estimates $\hat{\eta}^2 = 0.0002$, 0.05, 0.3, 0.4 and 2.4. These cases range from a nearly linear system to a large degree of nonlinearity. The estimates $\hat{\eta}^2$ present different levels of inaccuracy. The theoretical curves using (14) (continuous curves) and Monte Carlo simulations (1000 runs) (ragged curves) show excellent agreement. The behaviors of the other weights are similar.

Fig. 3 shows the MSE behavior. Again, there is excellent agreement between theory (Eq. (21)) and simulation results (averaged over 1000 runs).

In both cases, coefficients and MSE, the errors due to the exponential function approximation are negligible.

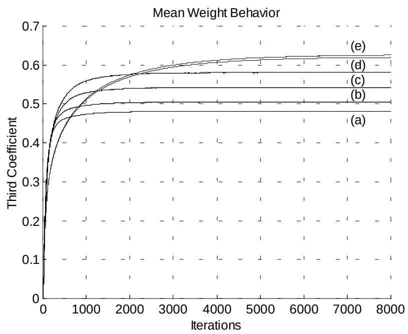


**Fig. 2**. Mean behavior of the third coefficient. (a) $\eta^2 = 0.0001$, $\hat{\eta}^2 = 0.0002$; (b) $\eta^2 = 0.1$, $\hat{\eta}^2 = 0.05$; (c) $\hat{\eta}^2 = \eta^2 = 0.3$; (d) $\eta^2 = 0.5$, $\hat{\eta}^2 = 0.4$; (e) $\eta^2 = 2$, $\hat{\eta}^2 = 2.4$.

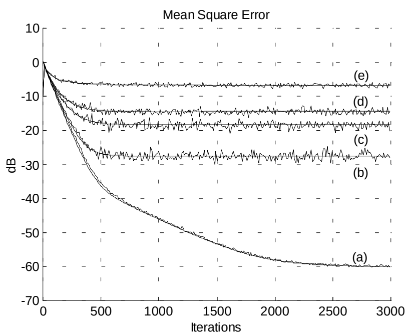


**Fig. 3**. Mean square error. (a) $\eta^2 = 0.0001$, $\hat{\eta}^2 = 0.0002$; (b) $\eta^2 = 0.1$, $\hat{\eta}^2 = 0.05$; (c) $\hat{\eta}^2 = \eta^2 = 0.3$; (d) $\eta^2 = 0.5$, $\hat{\eta}^2 = 0.4$; (e) $\eta^2 = 2$, $\hat{\eta}^2 = 2.4$.

Fig. 4 compares steady-state misadjustments achieved by the LMS (horizontal line) and by the new algorithm (curved line) as a function of the error in estimating $\eta^2$. The title of each plot corresponds to the value of $\eta^2$. The vertical axes give the misadjustment and the horizontal axes the ratio $\hat{\eta}^2/\eta^2$. The LMS misadjustment as a function of $\eta^2$ was obtained from [5].

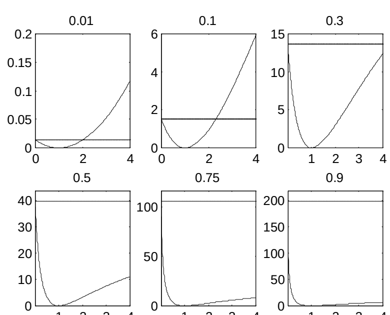


**Fig. 4**. Comparisons between the steady-state misadjustments of the LMS and the proposed algorithm.

The true stochastic gradient implementation occurs for $\hat{\eta}^2/\eta^2 = 1$. As expected, the new algorithm always outperforms LMS for $\hat{\eta}^2/\eta^2 = 1$. It also outperforms LMS whenever the curved line is below the horizontal line. As $\eta^2$ increases, the new algorithm performs much better than LMS even for quite bad estimates of $\eta^2$. The results in Fig. 4 reflect the best steady-state cancellation level achievable by each algorithm with slow learning and given enough time. A theoretical study comparing convergence rates and steady-state results as a function of μ is under way.

## 6. SUMMARY

A new adaptive algorithm has been proposed for applications in which intrinsic saturation nonlinearities exist at the adaptive filter output. The new algorithm implements the true stochastic gradient in its weight update equation. An analytical model has been derived which is able to predict the mean weight and MSE behaviors for Gaussian inputs and slow learning. Practical issues regarding the estimation of the nonlinearity parameter and the algorithm implementation using dedicated hardware were addressed. The new algorithm was shown to outperform the LMS algorithm even for considerable errors in estimating the nonlinearity.

## 7. REFERENCES[1]


[1] Kuo, S. and Morgan, D. *Active Noise Control Systems: Algorithms and DSP Implementations*, Wiley, 1996.

[2] Widrow, B. and Stearns, S.D. *Adaptive Signal Processing*, Prentice-Hall, 1985.

[3] Hansen, C. "Active Noise Control – From Laboratory to Industrial Implementation," *Proceedings of Noise-Con 97*, USA, Pennsylvania, 1997, pages 3-38.

[4] Heinle, F., Rabenstein, R. and Stenger, A. "A Measurement Method for the Linear and Nonlinear Properties of Electro-Acoustic Transmission Systems," *Signal Processing*, vol. 64, No. 1, pages 49-60, 1998.

[5] Costa, M.H., Bermudez, J.C.M. and Bershad N.J. "Statistical Analysis of the LMS Algorithm with a Zero-Memory Nonlinearity after the Adaptive Filter," *ICASSP 1999*, USA, Phoenix, 1999.

[6] Costa, M.H., Bermudez, J.C.M. and Bershad N.J. "The Performance Surface in Nonlinear Mean Square Estimation: Application to the Active Noise Control Problem," *ICASSP 2000*, Turkey, Istanbul, 2000.

[7] Bershad, N.J., Celka, P. and Vesin, J.M. "Stochastic Analysis of Gradient Adaptive Identification of Nonlinear Systems with Memory for Gaussian Data and Noisy Input and Output Measurements," *IEEE Trans. on Signal Processing*, vol. 47, No. 3, pages 675-689, March, 1999.

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

[9] SPRA480 (1998), TMS320C54x DSPLIB User's Guide: "Optimized DSP Library for C Programmers on the TMS320C54x", Texas Instruments, pages 49-50.


[1] The authors would like to thank Prof. Neil J. Bershad for his constructive comments and suggestions.