VARIABLE STEP SIZE ALGORITHMS FOR NETWORK ECHO CANCELLATION

O.O. Oyerinde and S.H. Mneney
School of Electrical, Electronic and Computer Engineering, University of KwaZulu-Natal, King
George V Avenue, Glenwood, Durban, 4041, South Africa
oyerinde@ukzn.ac.za and Mneneys@ukzn.ac.za

ABSTRACT

Convergence rate of an algorithm is an important factor that determines the deployment of such algorithm in a real time application. In this paper, we propose improved versions of normalized least mean square (NLMS) algorithm: single and multiple -variable step size normalized least mean square (VSSNLMS) algorithms for echo cancellation. The presented algorithms exhibit faster convergence rate in comparison to NLMS algorithm. Simulation results employing standard figure of merits show how the algorithms perform better than NLMS algorithm based echo canceller. The good performance exhibit by these algorithms in terms of convergence rate as indicated by Means Squared Error (MSE) and Echo Return Loss Enhancement (ERLE) will lend them to deployment in the real-time network echo cancellation applications.
.

Keywords: Echo cancellation, double talk, normalized least mean square (NLMS), single variable step size normalized least mean square (SVSSNLMS), multiple variable step size normalized least mean square (MVSSNLMS).

1 INTRODUCTION

Echo cancellation in communication system has been deployed in telephone networks for voice quality enhancement for several decades. Echo, a delayed or distorted version of the transmitted signal reflected back to the source is caused by the four-wire to two-wire impedance mismatch in telephone networks. Distinct echoes are perceived when an un-attenuated reflection's round-trip delay exceeds few tens of a millisecond. If the echo's round-trip delay approaches a quarter of a second and there is little or no attenuation of the echo, most people cannot carry on with a conversation without stuttering. Consequently, there is a need for network echo cancellers for echo paths with long impulse responses such as 32ms or more.

In [1, 2] Adaptive Electrical Echo canceller for Telephone Network based on a combination of a Normalized Least Mean Square (NLMS) and Geigel double-talk detector (DTD) algorithms was presented. The improvement of the canceller as a result of the combination of the speech detector algorithm with NLMS algorithm was obvious in the results presented, but this was with a penalty of a slow convergence rate for longer impulse responses. In [3] a new NLMS adaptation scheme for echo cancellation was presented. The scheme combines the advantages of the Geigel algorithm with some initiative ideas. A new architecture that was introduced makes the Geigel DTD algorithm to be more sensitive to the double talk condition, thus improving the echo canceller performance during the double talk condition but the problem of slow convergence rate was not addressed. In a bid to address the convergence rate exhibited by the echo canceller based on NLMS algorithm, various algorithms have been proposed with varied performances. Among these algorithms are proportionate normalized least mean squares (PNLMS) and PNLMS++ proposed in [4] and [5] respectively.

This paper focuses on improving the convergence rate of the echo canceller based on NLMS algorithm by employing variable step size instead of fixed step size for NLMS adaptive algorithm. This work is an improvement on the work presented in [1, 2].

Throughout this paper bold small letters such as x denote column vectors and dependency on time index n are denoted as x_n. $E\{x\}$ is the expectation of x. Superscript T denotes transpose.

The paper is organized as follows. The system model is described in Section II. In Section III, the NLMS adaptive algorithm is presented while in Section IV the proposed Single and Multiple-VSSNLMS, and DTD algorithms are presented. Figure of merits used to establish the performance of the algorithms are discussed in section V and the simulation processes are discussed in Section VI,

while the conclusion is drawn in Section VII.

2 SYSTEM MODEL

The system model for echo canceller and double-talk detector considered in this paper is illustrated in Fig.1. The echo path impulse response vector is represented by vector $\boldsymbol{h}_{ep} = [h_0,...,h_{L-1}]^T$ and its model in the canceller is represented by the vector, $\hat{\boldsymbol{h}}_n = [h_{0,n},...,h_{L-1,n}]^T$, where L is the adaptive filter length. The signal x_n is the sampled far-end signal. The response of the model $\hat{y}_n$ is subtracted from the combination of the echo and the speech of the near-end speaker y_n leaving only the sampled speech of the near-end speaker v_n to be sent to the far-end user. The problem, of course, is in building (and maintaining) the model and, to some extent, in obtaining the response of the model to the excitation signal.

Echo cancellers as in Fig.1 are predominantly used to terminate long-distance 4-wire circuits on a per call basis, each circuit having a different impulse response. Also, during the call, variations in the echo path may occur. Therefore, the echo path model $\hat{h}_n$ must have the ability to learn and adapt to the new echo path impulse response at the beginning of each call. To accomplish this, the echo canceller uses an adaptive filter to construct the echo impulse response model. The adaptive filter is generally based on mathematical algorithm(s). The adaptive filter attempts to build the echo impulse response model by adjusting its filter coefficients (or tap-weights) in such a way as to drive e_n to zero. This is fine if y_n consists only of the echo of the far-end speech. In that case, the correlation of x_n and y_n contains valuable information about the echo impulse response. If, on the other hand, y_n also contains significant amounts of the summation of near-end signal, v_n and background noise, then the echo impulse response information is corrupted by any extraneous correlation between x_n and v_n. For this reason, practical echo cancellers need to inhibit adaptation of the filter taps when significant near-end signal is present and this is made possible by the presence of DTD.

3 NLMS ADAPTIVE ALGORITHM

The simplest and most popular adaptive iterative algorithm is the list mean square (LMS) algorithm given by the following equation [6]:
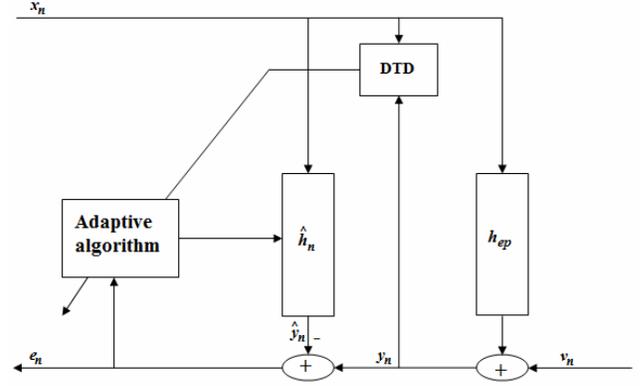

Figure 1: System model for echo canceller and double-talk detector

$$\hat{\boldsymbol{h}}_{n+1} = \hat{\boldsymbol{h}}_n + \mu e_n \boldsymbol{x}_n , \qquad (1)$$

$$e_n = y_n - \hat{\boldsymbol{h}}_n^T \boldsymbol{x}_n , \qquad (2)$$

where μ is the fixed step-size.

LMS algorithm adjust the estimated impulse response so as to minimize the cost function, $E\{|e_n|^2\}$, i.e., the mean square error. Each iteration updates the current estimate of $\hat{\boldsymbol{h}}_n$ by $\mu e_n \boldsymbol{x}_n$, which is a step in the direction of a stochastic approximation to the gradient $E\{|e_n|^2\}$. The algorithm, though widely used because of its simplicity of implementation, suffers from relatively slow and data-dependent convergence behaviour. In order to make LMS algorithm insensitive to changes of the level of input signal, x_n, the fixed step-size μ is normalized, resulting in the NLMS adaptive algorithm given as [6]:

$$\hat{\boldsymbol{h}}_{n+1} = \hat{\boldsymbol{h}}_n + \mu e_n \frac{\boldsymbol{x}_n}{\|\boldsymbol{x}_n\|^2} , \qquad (3)$$

where $\|\boldsymbol{x}_n\|^2$ denote the Euclidean norm of the input vector $\boldsymbol{x}_n$.

4 PROPOSED VARIABLE STEP SIZE NLMS (VSSNLMS) ALGORITHMS AND DTD

4.1 Single-VSSNLMS Algorithm

The NLMS algorithm is given more attention in real-time applications because it exhibits a good balance between computational cost and performance. However, a very serious problem associated with both the LMS and NLMS algorithms is the choice of the step-size (μ) parameters. A small step size (small compared to the reciprocal of the input signal strength) will ensure small

misadjustments in the steady state, but the algorithms will converge slowly and may not track the nonstationary behaviour of the operating environment very well. On the other hand a large step size will in general provide faster convergence and better tracking capabilities at the cost of higher misadjustment. Any selection of the step-size must therefore be a trade-off between the steady-state misadjustment and the speed of adaptation.

Several studies [7, 8, 9] have thus presented the idea of variable step-size LMS algorithms in order to eliminate the "guesswork" involved in selection of the step-size parameter and at the same time ensuring that the speed of convergence is fast. When operating in stationary environment, the steady-state misadjustment values is very small, and when operating in non-stationary environment the algorithm should be able to sense the rate at which the optimal coefficients are changing and select a step-size that can result in estimates that are close to the best possible in the mean-squared-error sense.

The variable step-size expression for Single-VSSNLMS algorithm employed in this paper is obtained by extending the approach used in [7] to derive similar variable step-size expression for the LMS algorithm. This is done by adapting the step-size sequence using a gradient descent algorithm so as to reduce the squared-estimation error at each time index. The Single-VSSNLMS algorithm is then given as:

$$\hat{\boldsymbol{h}}_{n+1} = \hat{\boldsymbol{h}}_n + \mu_n e_n \frac{\boldsymbol{x}_n}{\left\| \boldsymbol{x}_n \right\|^2} \quad . \tag{4}$$

The variable step-size μ_n is updated as [10]:

$$\hat{\mu}_n = \mu_{n-1} - \frac{\rho}{2} \frac{\partial e_n^2}{\partial \mu_{n-1}} \tag{5a}$$

$$= \mu_{n-1} - \frac{\rho}{2} \frac{\partial^T e_n^2}{\partial \hat{\boldsymbol{h}}_n} \cdot \frac{\partial \hat{\boldsymbol{h}}_n}{\partial \mu_{n-1}} \tag{5b}$$

$$= \mu_{n-1} + \frac{\rho e_n e_{n-1} \boldsymbol{x}_n^T \boldsymbol{x}_{n-1}}{\left\| \boldsymbol{x}_{n-1} \right\|^2} \quad . \tag{6}$$

In Eq. (6), ρ is a small positive constant that controls the adaptive behavior of the step-size sequence μ_n. Deriving conditions on ρ so that convergence of the adaptive system can be guaranteed appears to be a difficult task. However, the convergence of the adaptive filter can be guaranteed by restricting μ_n to always stay within the range that would ensure convergence. Therefore the step size obtained from Eq. (6) would not be used for coefficient adaptation at any particular time index if it falls outside the values that guarantee convergence of the NLMS algorithm with a fixed step-size. As a result the step-

size sequence μ_n will be restricted to within the range $0 < \mu_n < 2$ [11]. The variable step size μ_n is then restricted as follow:

$$\mu_n = \begin{cases} \mu_{\max} & if \quad \hat{\mu}_n > \mu_{\max} \\ \mu_{\min} & if \quad \hat{\mu}_n < \mu_{\min} \\ \hat{\mu}_n & otherwise \end{cases} \tag{7}$$

where $0 < \mu_{\min} < \mu_{\max} < 2$.

In [12] the order of coefficient update of NLMS is given as $O(ML)$ where L is the filter length and M is the echo path maximum delay. However, the VSSNLMS algorithm only requires L extra additions and $(L+4)$ extra multiplications (divisions) compared with NLMS algorithm, the value which is more or less negligible.

4.2   Multiple-VSSNLMS Algorithm

In Multiple-VSSNLMS algorithm rather than using a single variable step size for the adaptation of all the echo canceller's coefficients in the coefficient vector, $\hat{\boldsymbol{h}}_n$, each coefficient is adapted with unique variable step size resulting in multiple- VSSNLMS algorithm. As a result, the variable step-size μ_n in Eq. (4) becomes a vector $\boldsymbol{\mu}_n = \left[\mu_{0,n}, \ldots, \mu_{L-1,n} \right]^T$ and is derived following Eq. (5) and Eq. (6) as :

$$\hat{\boldsymbol{\mu}}_n = \hat{\boldsymbol{\mu}}_{n-1} + \frac{\rho e_n e_{n-1} \boldsymbol{x}_n^T \boldsymbol{x}_{n-1}}{\left\| \boldsymbol{x}_{n-1} \right\|^2} \quad . \tag{8}$$

Similarly, each of the variable step size, μ_n in the multiple-variable step size vector $\boldsymbol{\mu}_n$ is restricted within the range as given in Eq. (7).

4.3   Geigel Double Talk Detection (DTD)

During double talk, the period where there is presence of the far- and near- end speech simultaneously, double-talk detector is needed to inhibit taps adaptation. A very efficient and simple way of detecting double-talk is to compare the magnitude of the far-end and near-end signals and declare double -talk if the near-end signal is lager in magnitude than a value set by the far-end speech. Geigel DTD algorithm [13], attributed to A. A. Geigel is a proven algorithm in general use for this purpose and is given by Eq. (9) through which a double talk is declared if

$$\left| y_n \right| \geq \xi \max \left\{ \left| x_n \right|, \left| x_{n-1} \right|, \ldots, \left| x_{n-L+1} \right| \right\}, \tag{9}$$

where ξ is the detector threshold factor normally set to 0.5 if the network hybrid attenuation, Echo Return Loss (ERL), is assumed to be 6dB and to 0.71 if the ERL is assumed to be 3dB. Beside this threshold factor, a hangover time, τ_{hold}, is also specified such

that if double-talk is detected, then the adaptation is inhibited for this specified duration beyond the detected end of double-talk.

5 FIGURE OF MERITS

There are two figures of merit employed in this simulation. One of these figure of merit used to establish the performances of the proposed echo canceller algorithms is a quantity called Echo Return Loss Enhancement (ERLE). This is a comparison of the echoes before and after cancellation. It is calculated as:

$$ERLE = 10\log_{10}\left(\frac{power\ of\ the\ echo\ signal}{power\ of\ the\ residual\ echo}\right)\ dB$$

,

$$= 10\log_{10}\left(\frac{E\{y_n^2\}}{E\{(y_n - \hat{y}_n)^2\}}\right)\ dB\ ,$$

$$= 10\log_{10}\left(\frac{E\{y_n^2\}}{E\{e_n^2\}}\right)\ dB\ .(10)$$

The ERLE therefore is the amount of attenuation of the echo signal introduced by the echo canceller. It does not include any further reduction in the residual echo by any extra nonlinear processing after the basic echo cancellation. The ERLE provides a figure of merit for determining how effective the echo cancellation process is. It assumes that there is always a certain amount of loss incurred by the echo and then shows the rate of improvement after echo cancellation. It reflects both the convergence rate and the steady-state residual echo. The plot of ERLE versus time or sample shows the rate of change in the enhancement: it shows the rate of convergence of the algorithm to the steady-state error value. The ERLE gives a good indication of the performance of the echo canceller. Over time the ERLE changes, initially it may be quite small but as the algorithm converges towards the optimum tap-weight values it increases. Theoretically the steady state ERLE could be very large and an ideal echo canceller with a perfectly linear echo signal would output an infinite ERLE in a very short period of time. Practically however, there are limiting factors to this result; the echo path always contains some non-linearities introduced by various components in the transmission path; the devices that generate the echo produce a certain amount of echo loss that little can be done about and the use of finite-precision devices limit the accuracy of the computations. Therefore the ERLE will not reach its theoretical steady-state

maximum value. Nevertheless, a good performing echo canceller will output a very large steady-state ERLE in a very short convergence time [14].

Another important figure of merit used is the MSE which shows the adaptation curves of the algorithm employed. It is given mathematically as the expectation of the norms of the square error as follow:

$$\text{MSE} = 10\times\log_{10}\left(E\{\|e(n)\|^2\}\right)\quad\text{(dB)}$$

$$= 10\times\log_{10}\left(E\{e(n)e^*(n)\}\right)\ \text{(dB)}\qquad(11)$$

6 SIMULATION

The performances of both Single and Multiple-VSSNLMS algorithms have been compared with that of NLMS algorithm with fixed step size. The echo path is modeled with an impulse response, g(n) of a linear digital filter. In order to account for the delay experienced by the echo signal and the ERL of hybrid transformer in a telephone network, g(n) is chosen as a delayed and attenuated version of the excitation signal according to ITU G.168 standard for testing network echo canceller performance [15]. The mathematical expression for g(n) is given as:

$$g(n) = 10\exp(-\frac{ERL}{20})\times M_i(n-\delta),\qquad(12)$$

where the sequence $M_i(n)$ denotes the echo paths with varied time-delay, and δ represents the total delay experienced by the echo signal. For all the results presented in this paper, ERL value of 6dB is used because this is a typical worst case value encountered for most networks, and most current networks even have typical ERL values better than 6dB.

Two types of excitation signals are employed for the simulation of the results presented in this paper: the type of random signal used in [1, 2], as well as sampled speech signal as shown in Fig. 2 and Fig. 7 respectively. For each of the excitation signals, maximum echo delay was set to 16ms (128 samples) and 32ms (256 samples), while the echo canceller length (adaptive filter length) was set to 256(32ms) and 512 (64ms) respectively. For effective performance of any echo canceller the length of the echo canceller is always selected such as to be longer than the maximum possible echo delay in the network. The following other parameters were used in the simulation:μ =0.02 for NLMS algorithm and also to initialize the Single-VSSNLMS and Multiple-VSSNLMS algorithms, ρ = 2×10^{-4} . The performances of the proposed algorithms based on the figure of merits discussed in section V are as shown in Fig. 3 to Fig. 6, and Fig. 8 to Fig. 11 for random signal and sampled speech signal as excitation signals respectively.
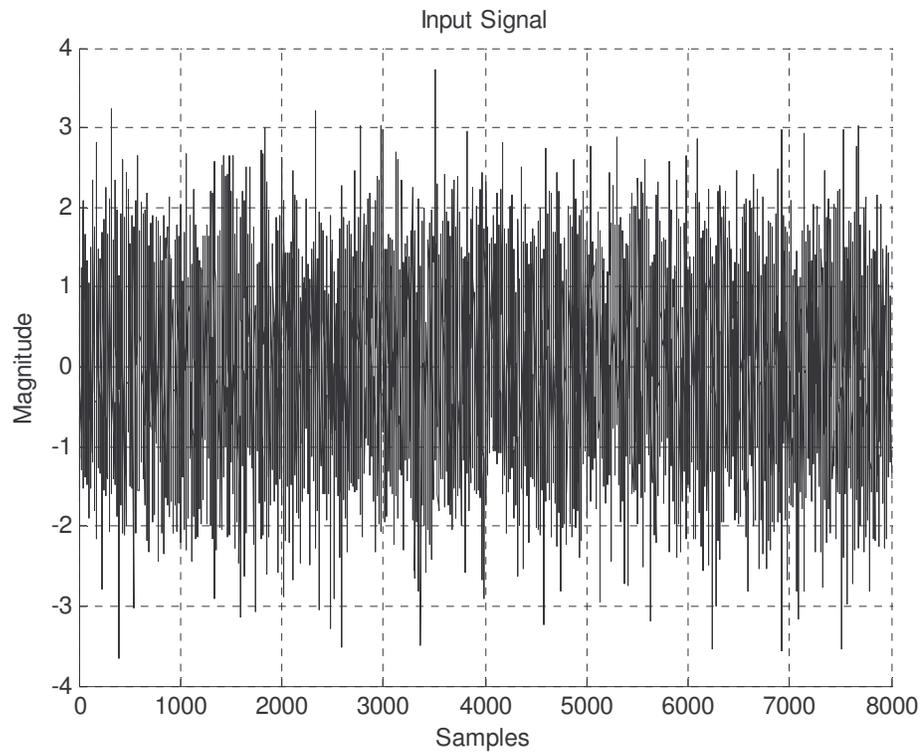
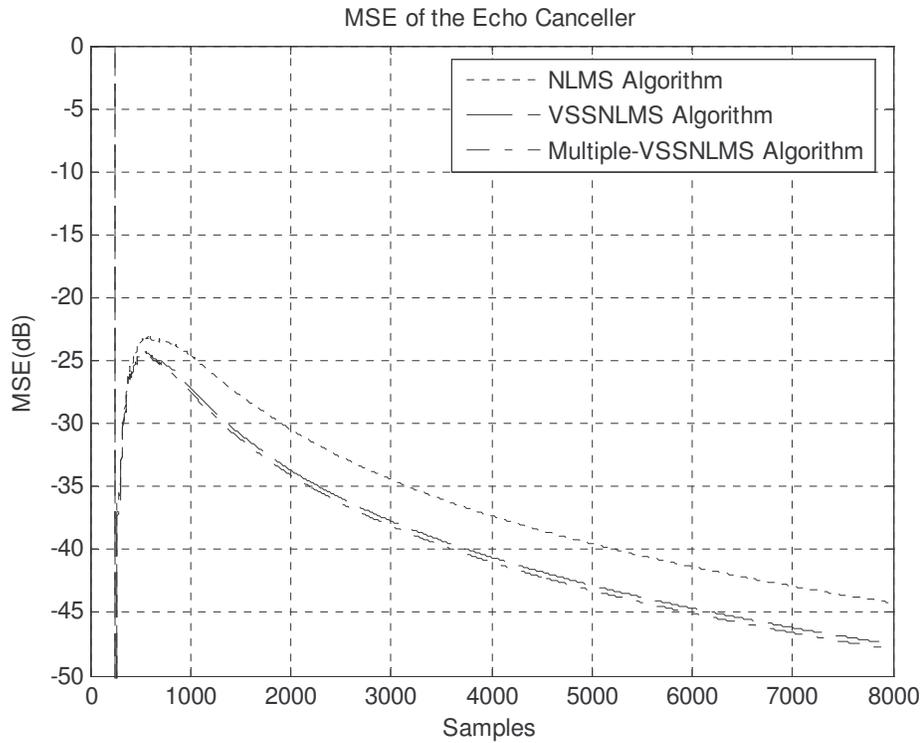**Figure 2:** Random signal as the excitation signal

Figure 3: MSE for the algorithms with random signal as the excitation signal, $L = 256$(32ms), $M = 128$ (16ms).
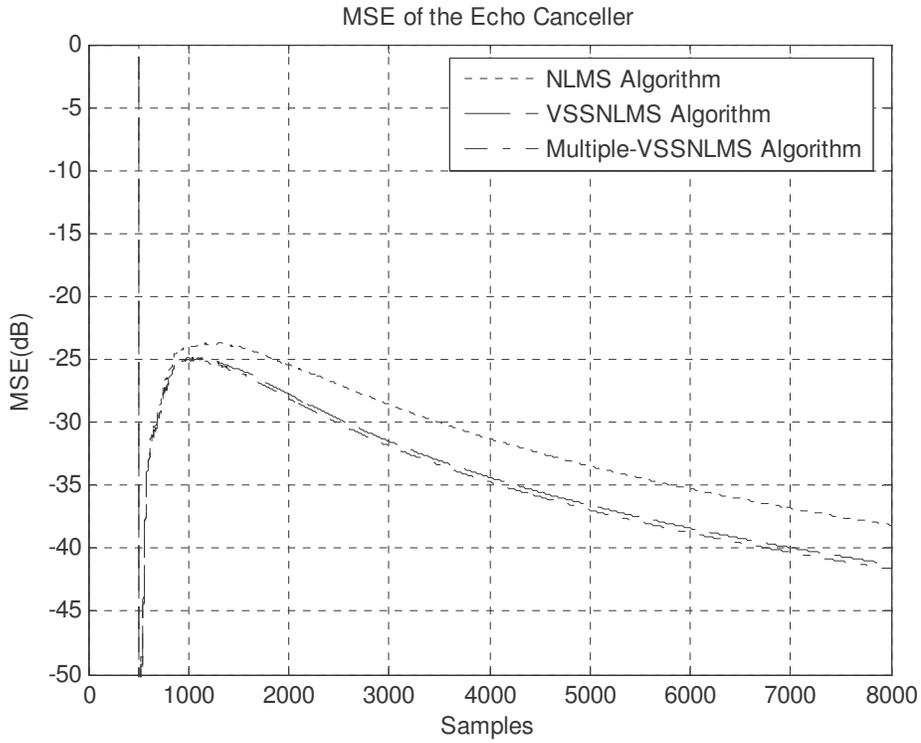
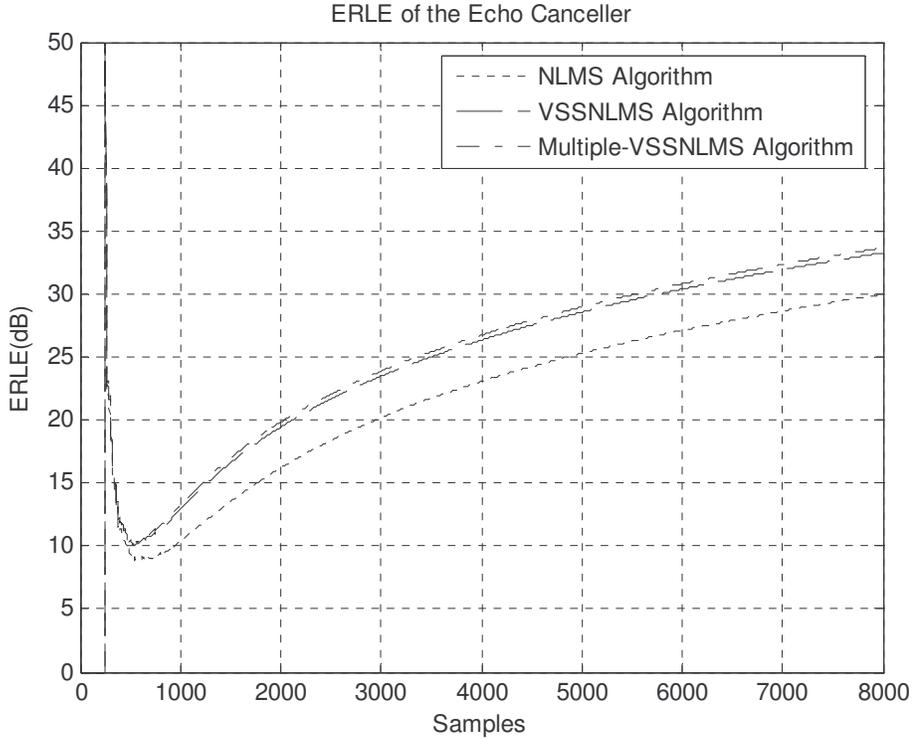**Figure 4:** MSE for the algorithms with random signal as the excitation signal, L =512(64ms), M =256 (32ms).

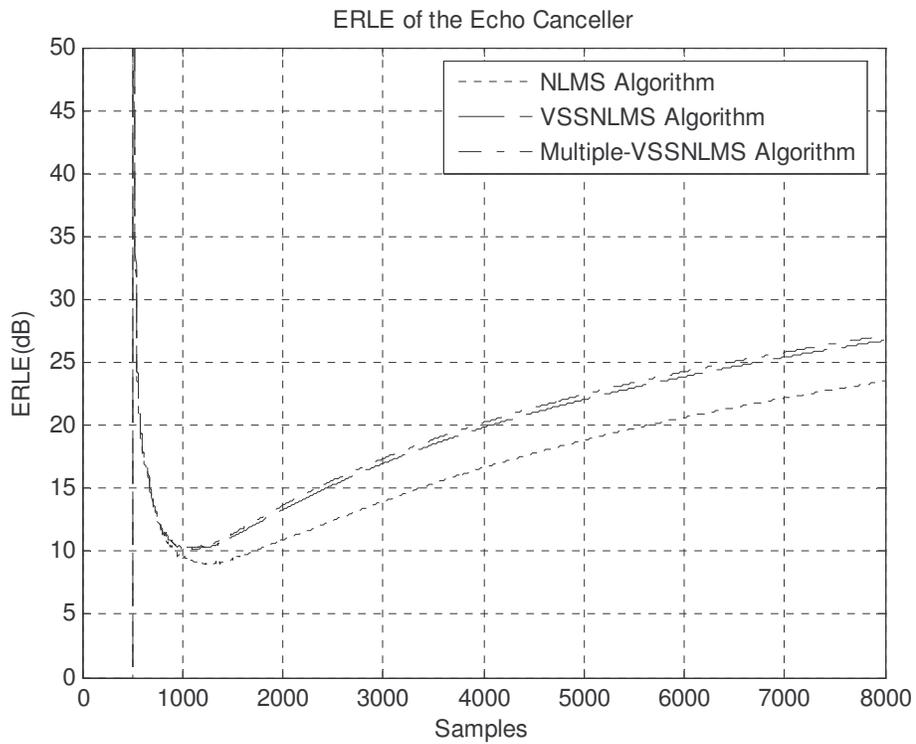**Figure 5:** ERLE for the algorithms with random signal as the excitation signal, L =256(32ms), M =128 (16ms).

Figure 6: ERLE for the algorithms with random signal as the excitation signal, L =512(64ms), M =256 (32ms).
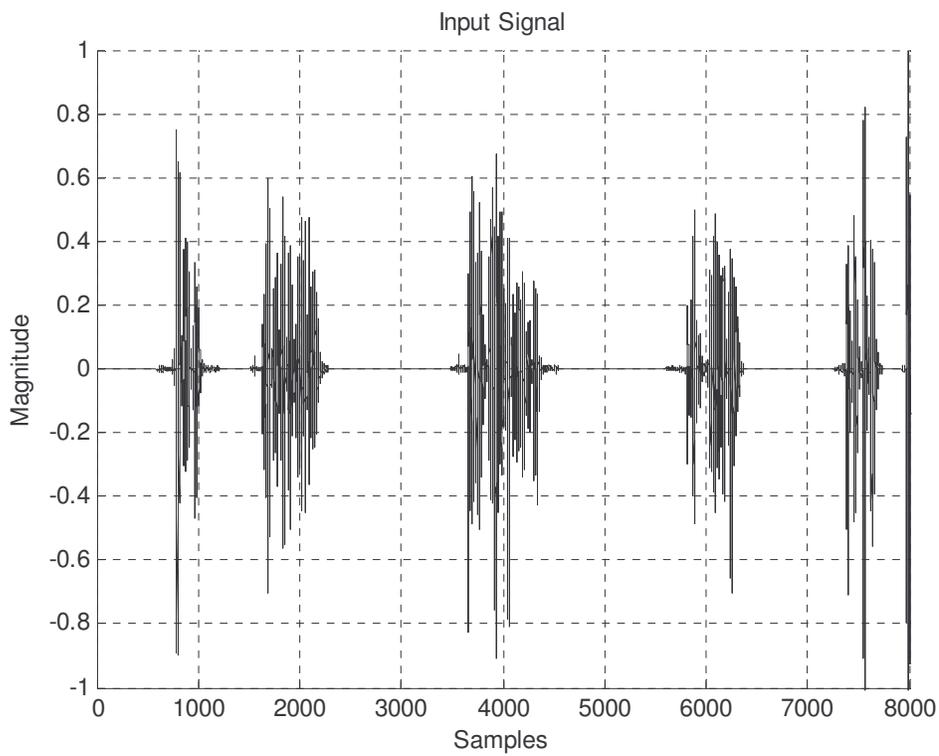
Figure 7: Sampled speech signal as the excitation signal
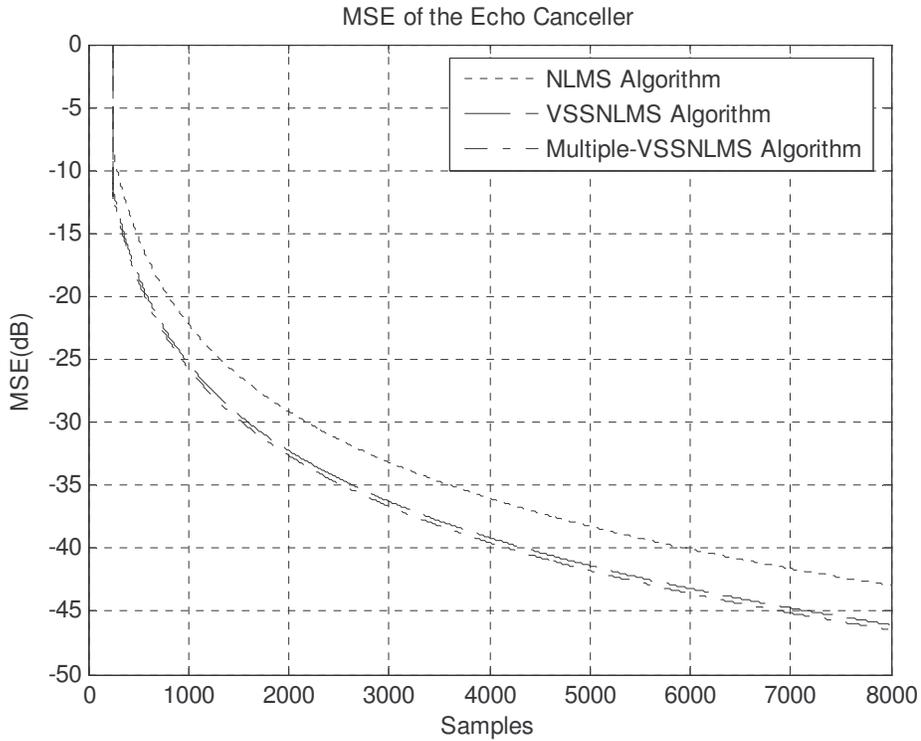
Figure 8: MSE for the algorithms with sampled speech signal as the excitation signal, L =256(32ms), M =128 (16ms).
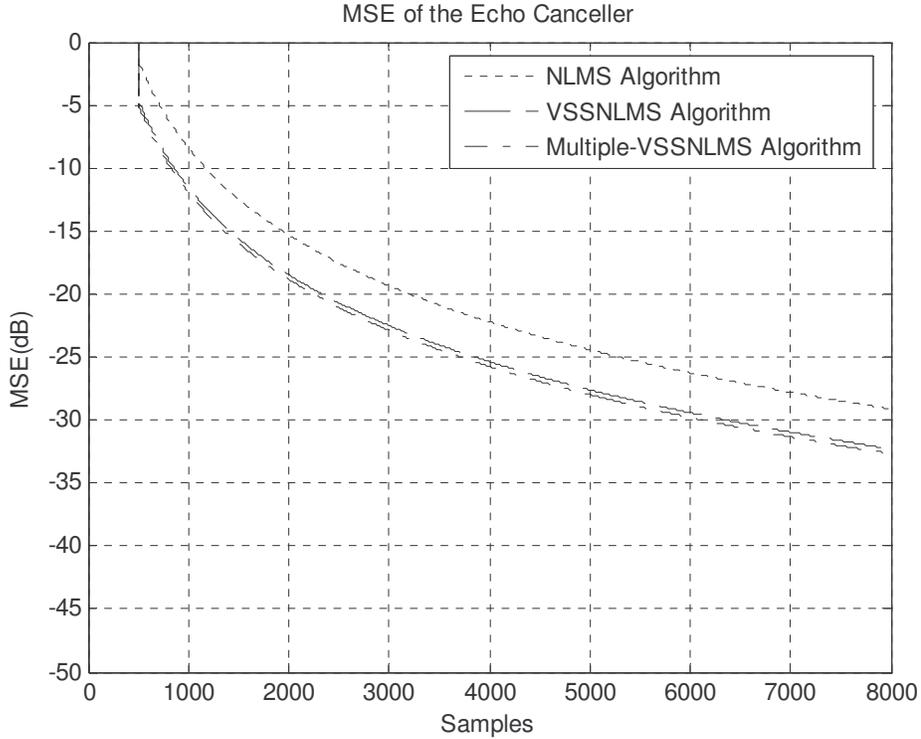

Figure 9: MSE for the algorithms with sampled speech signal as the excitation signal, L =512(64ms), M =256 (32ms).

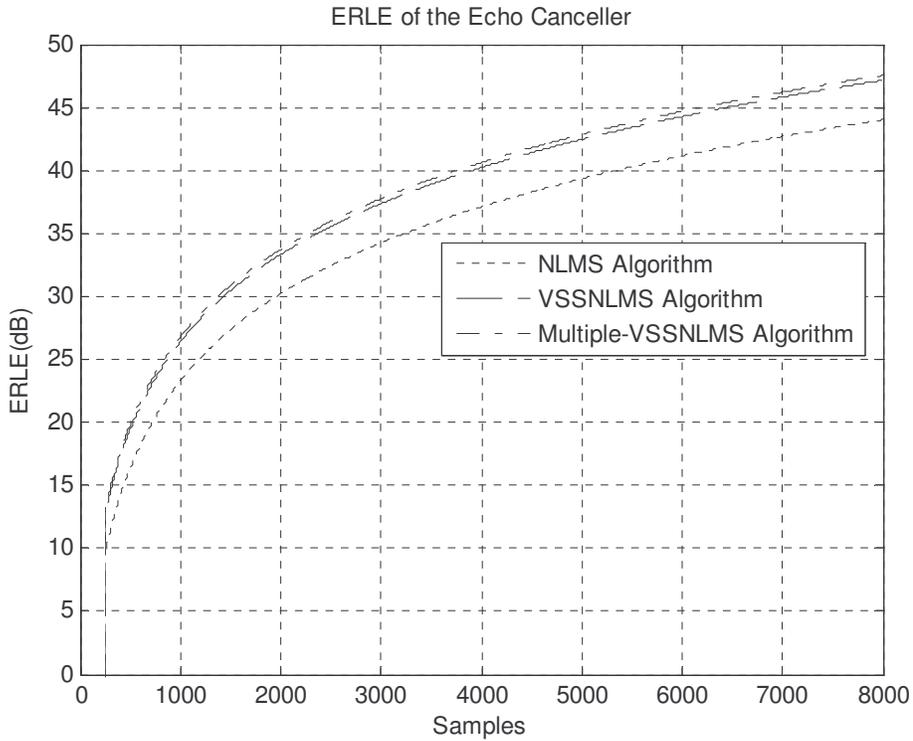**ERLE of the Echo Canceller**

Figure 10: ERLE for the algorithms with sampled speech signal as the excitation signal, $L =256(32ms)$, $M =128$ (16ms).
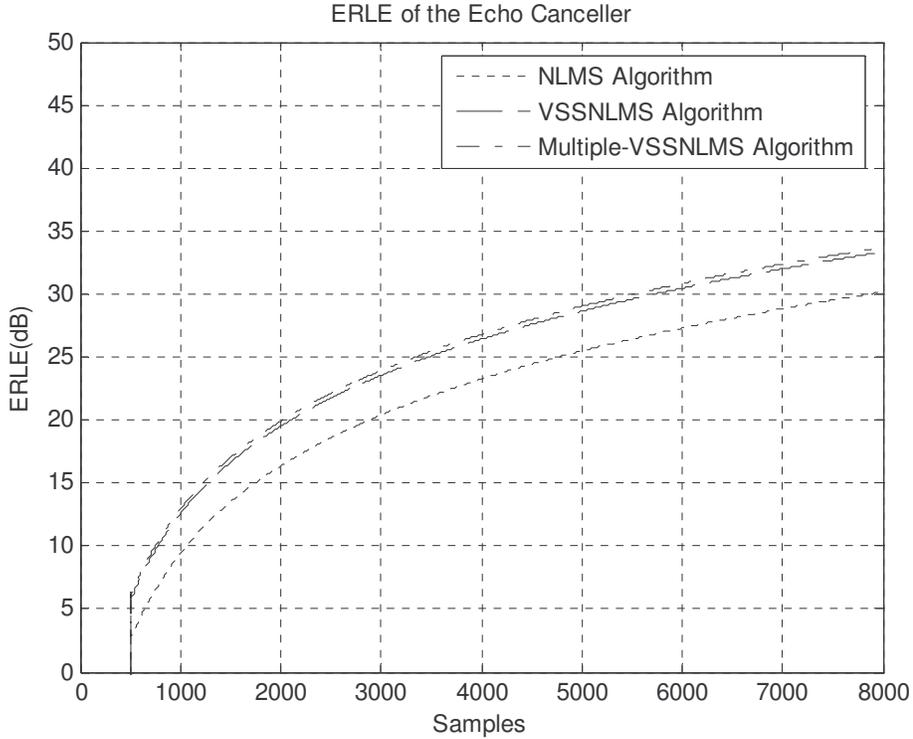
ERLE of the Echo Canceller

Figure 11: ERLE for the algorithms with sampled speech signal as the excitation signal, $L =512(64ms)$, $M =256$ (32ms).
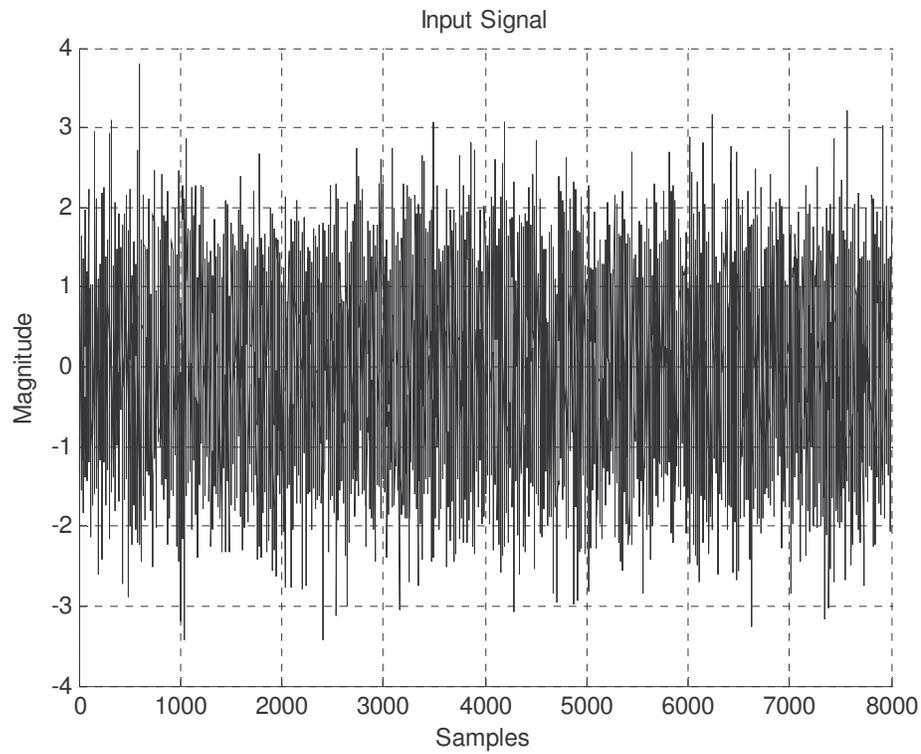
Input Signal

Figure 12: Reference signal (far-end signal) for double-talk condition testing

Near-end Signal

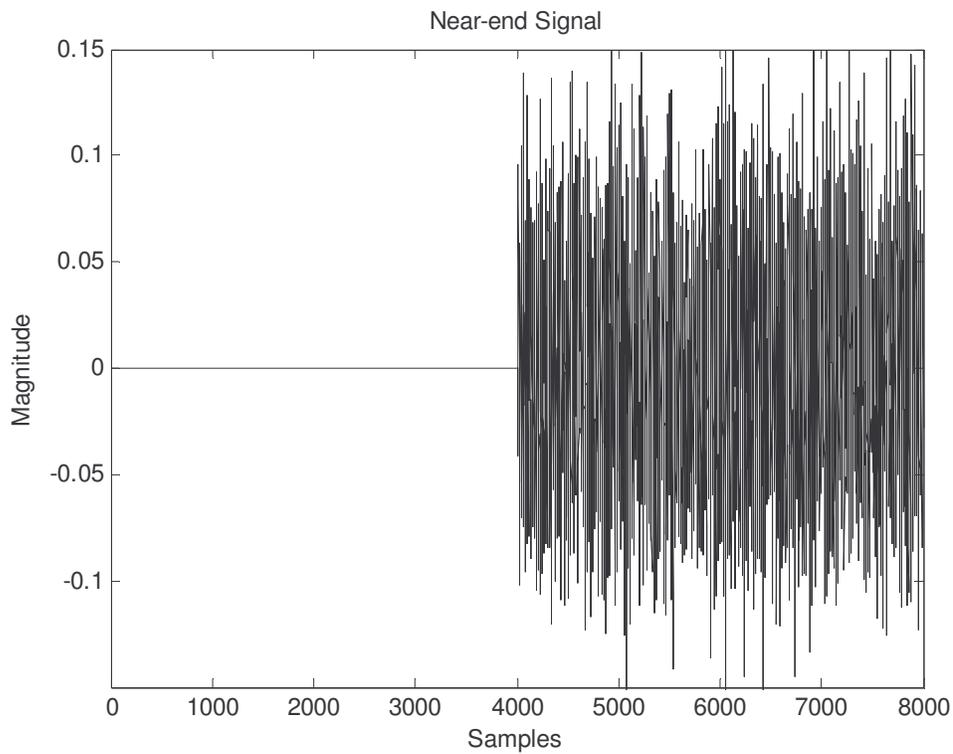**Figure 13:** Near-end signal for double-talk condition testing

MSE of the Echo Canceller during double-talk condition

Figure 14: MSE for the combination DTD and proposed algorithms during the double-talk condition, L =512(64ms), M =256 (32ms).

ERLE of the Echo Canceller during double-talk condition

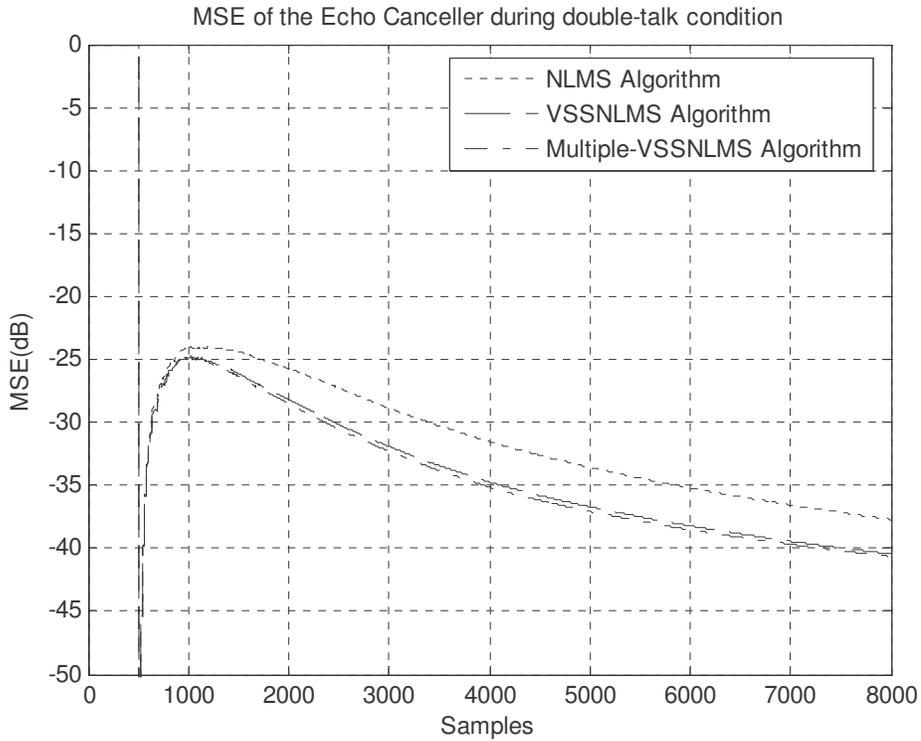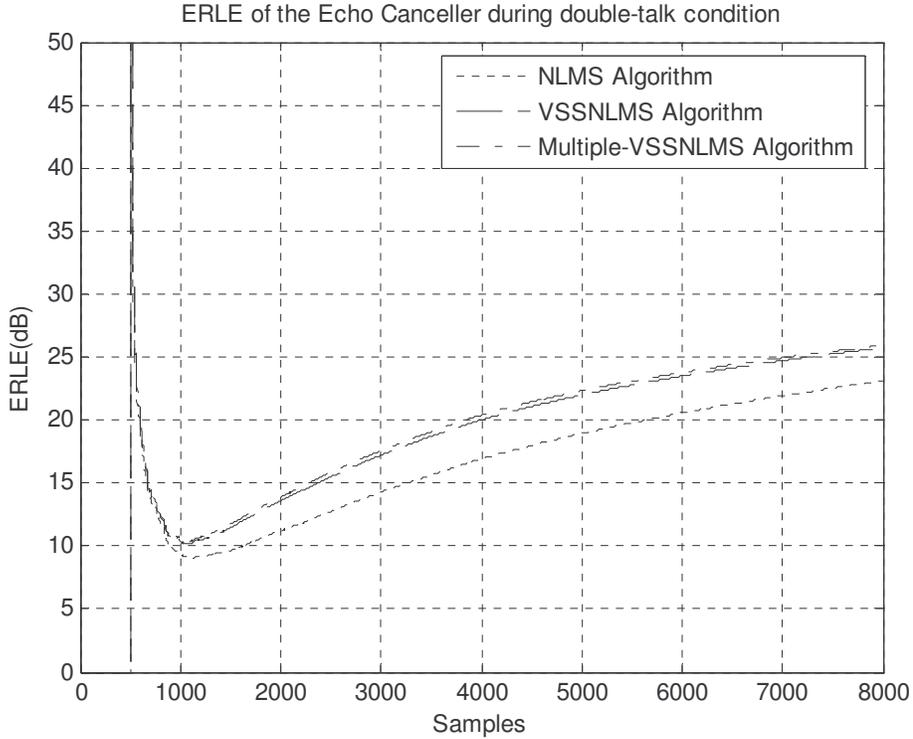**Figure 10:** ERLE for the combination DTD and proposed algorithms during the double-talk condition, L =512(64ms), M =256 (32ms).

In order to establish the robustness of the combination of Geigel DTD algorithm with the proposed echo canceller algorithms during double talk condition, random signal of different magnitude compared with the reference signal was added with the echo signal to serve as the near-end signal after about half of the period of the cancellation process has elapsed. The simulation was run for the echo canceller of length 512. Fig. 14 and Fig.15 show how effective the combination of Geigel DTD and the proposed algorithms performed in comparison with the combination with NLMS algorithm during the double talk condition. Although the performances of the algorithms were reduced compared with the situation where there was no double-talk, the results still show that there was effective cancellation during this condition with the help of Geigel DTD algorithm.

It could be observed from these results that both Single-VSSNLMS and Multiple-VSSNLMS algorithms outperformed the NLMS algorithm as a result of the variability of the step-size, but the differences in the performance of Single-VSSNLMS and Multiple-VSSNLMS algorithms are negligible. This shows that assignment of a unique variable step size for the adaptation of each of the coefficients of the echo canceller makes no or little difference compared with adapting all the coefficients with the same variable step size.

7   CONCLUSION

In this paper we have presented Single-VSSNLMS and Multiple-VSSNLMS algorithms for the network echo cancellation. These algorithms use variable step sizes instead of fixed step size used by NLMS algorithm. As a result, the convergence rates of these algorithms are significantly faster than that of NLMS algorithm. These algorithms also exhibit high performance during double-talk condition. As a result of the negligible difference in the performance of the Single-VSSNLMS and Multiple-VSSNLMS algorithms, it could be concluded that Single-VSSNLMS algorithm which is less complex than Multiple-VSSNLMS algorithm should be employed in the real-time network echo cancellation applications.

8   REFERENCES

[1] O. O. Oyerinde, and T. K. Yesufu: Adaptive Electrical Echo Canceller for Telephone Networks, CD-ROM Proc. IEEE Military Communication Conference, MILCOM 2005, Atlantic City NJ, USA, Vol. xxii+3341, pp.1-5, Oct. 17-20, (2005).

[2] O. O. Oyerinde, and T. K. Yesufu: Adaptive Electrical Echo Canceller Algorithm, Proc. Intelligent Engineering System through Artificial Neural Network Conference, ANNIE 2005, Missouri-Rolla , USA, Vol. 15, pp.613-622, Nov. 6-9, 2005.

[3] J. F. Liu: A Novel Adaptation Scheme in the NLMS Algorithm for Echo Cancellation, *IEEE Signal Processing Letter.*, Vol. 8, No. 1 pp. 20– 22, January, (2001).

[4] D. L. Duttweiler: Proportionate normalized least mean square adaptation in echo cancellers, IEEE Trans. Speech Audio Processing, Vol. 8, pp. 508–518, Sept.., (2002).

[5] S. L. Gay: An efficient fast convergence adaptive filter for network echo cancellation, Proc. Assilomar Conf., Nov., (1998).

[6] B. Widrow, J. R. Glover, J.M. McCool Jr., J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler Jr., E. Dong, R. C. Goodlin: Adaptive noise canceling: principles and applications, Proc. IEEE 63 (12), pp. 1692-1716, Dec. (1975).

[7] V.J. Mathews, Z. Xie: A stochastic gradient adaptive filter with gradient adaptive step-size, IEEE Trans. Signal Process., Vol.41, no.6, pp. 2075–2087 June (1993).

[8] T. Aboulnasr: A Robust variable step-size LMS-Type Algorithm: Analysis and Simulation, IEEE Trans. Signal Process. Vol.45, no.3, pp. 631–639 March, (1997).

[9] D.I. Pazaitis, A.G. Constantinides: A novel kurtosis driven variable step-size adaptive algorithm, IEEE Trans. Signal Proc. Vol.47, no.3 pp.864–872 March (1999).

[10] Y.K. Shin, J.G. Lee: A study on the fast convergence algorithm for the LMS adaptive 3lter design, Proc. KIEE, Vol.19, no. 5, pp. 12–19, October (1985).

[11] M. Tarrab, A. Feuer: Convergence and performance analysis of the normalized LMS algorithm with uncorrelated Gaussian data, IEEE Trans. Inform. Theory, Vol.34, no.4, pp.680– 691, July (1988).

[12] Dieter Schafhuber, and Gerald Matz: MMSE and Adaptive Prediction of Time- Varying Channels for OFDM Systems, IEEE Transactions on Wireless Communications, vol. 4, no. 2, pp. 593-602, March (2005).

[13] D. L. Duttweiler: A twelve-channel digital echo canceller, IEEE Trans. Commun., Vol. COM-26, pp. 647-653, May (1978).

[14] Y. Lu and J.M. Morris: Gabor Expansion for Adaptive Echo Cancellation, IEEE Signal Processing Magazine, pp. 68-80, March (1999).

[15] ITU G.168, Recommendations: Digital Echo Canceller, (2002)