

# Error Signal Difference Based Variable Step Size LMS Algorithm\*

Min Zhang, Baoqi Huang, Liqian Wang

State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing, China

mzhang@bupt.edu.cn, baoqi\_huang@foxmail.com, wang@bupt.edu.cn.

**Abstract** - A variable step size Least Mean Square (LMS) algorithm based on the difference of the error signal (ESD-VSS) is proposed. The convergence and steady state behavior of the ESD-VSS is analyzed and the key parameters are given. Simulations are conducted and the results show that, compared with the Modified Variable Step Size (MVSS) algorithm, modified LMS algorithm and the standard LMS algorithm, the ESD-VSS provides faster convergence speed and smaller steady state misadjustment.

**Index Terms** - LMS algorithm, variable step size, adaptive filter, MSE

## I. Introduction

The LMS algorithm proposed by Widrow and Hoff has been widely used in channel equalization [1] and system identification [2], due to its robustness and simplicity. However the tradeoff between the convergence speed and the steady state misadjustment is a challenging problem. To achieve faster convergence speed and small misadjustment, a number of modified algorithms have been proposed. Ref. [3] utilized the self-correlation to control the step size, and suppress the interference of irrelevant noise effectively, but the convergence speed and the tracking capability are still expected to be improved. Ref. [4] gave a variable step size LMS algorithm using estimate signal error and thus it is of relatively faster tracking speed and smaller steady state error. However, the computational complexity of the algorithm is larger due to the tangent and exponent arithmetic, which makes it difficult to implement.

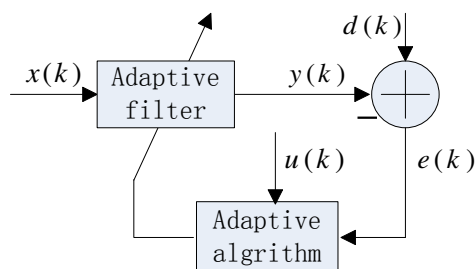


Fig.1 Structure of standard adaptive filter

In this paper, we propose a variable step size LMS algorithm which updates its step size according to the difference of the error signals. The intention is that a large

difference contributes to a big step size and provides a faster convergence speed while a small difference is capable of realizing small steady state misadjustment. The updating equation is simple and easy to implement compared with the reported algorithms in [5], and [6].

Condition for coefficients vector, convergence behavior of the mean square error (MSE) and expression for steady state misadjustment are provided. In addition, simulations of channel equalization and system identification with different assumptions are demonstrated and the performance of convergence speed and steady state misadjustment are evaluated in comparison of those of the known algorithms.

## II. The Principle of the Proposed Algorithm

In this section, we propose the new variable step size algorithm and give the equation for the update of the step size and the tap weights then analyze its computational complexity.

To acquire faster convergence speed with relatively small computational complexity, the ESD-VSS updates its coefficients and step size according to the following equations:

$$\mu(k) = \mu_0 [1 + \beta |e(k-1) - e(k)|] \quad (1)$$

$$w(k+1) = w(k) + u(k)e(k)x(k) \quad (2)$$

where  $\mu_0$  is the original step size,  $k$  is the iteration number,  $\beta$  is the adaptive parameter of the step size,  $w(k)$  is the vector of adaptive filter weights,  $x(k)$  is the adaptive filter input vector and  $e(k)$  represents the error signal.

From (1), the step size  $\mu(k)$  is controlled by both  $\beta$  and the difference between the error signals of the consecutive iteration. If the adaptive system has not reach steady state, the difference between the error signals will be big and gives a big step size while when the system reaches steady state, the difference will be small and gives a small misadjustment. To make the algorithm robust to noise, parameter  $\beta$ , which is correlated to the choice original step size  $\mu_0$ , must be set in a small range.  $\beta$  is designed to be within (0, 3) to guarantee that steady state is available.

From (1) and (2), we can conclude that there are  $N+1$

\* This work is partially supported by NSF Grant #61072008, 863 Program #2012AA011302 and Doctoral Scientific Fund Project of the ministry of Education of China#20120005110010 to M. Zhang.

multiplication operations and 1 addition to update tap weights, 2 multiplications and 2 additions to update the variable step size.

### III. Property analysis

In this section, we analyze the performance of the ESD-VSS algorithm. All the analyses are based on the assumptions given as follows:

- 1) The input signal  $x(k)$  and the tap weight  $w(k)$  are statistically independent.
- 2) The input signal  $x(k)$  is a zero-mean stochastic process.
- 3) The additive noise  $n(k)$  is a zero-mean Gaussian process

#### A. Convergence Behavior of the coefficient vector

Assume an unknown adaptive filter with coefficient vector given by  $w_0$ , employing the ESD-VSS variable step size algorithm, with white noise  $n(k)$  that has zero mean and variance  $\sigma_n^2$  added to the unknown filter.

The error in the adaptive filter coefficients as related to the ideal coefficient vector  $w_0$  is described as

$$\Delta w(k) = w(k) - w_0 \quad (3)$$

Equation (2) then can be rewritten as:

$$\begin{aligned} \Delta w(k+1) &= \Delta w(k) + u(k)e(k)X(k) \\ &= \Delta w(k) + u(k)X(k)[e_0(k) - X^T(k)\Delta w(k)] \\ &= [I - u(k)X(k)X^T(k)]\Delta w(k) + u(k)e_0(k)X(k) \end{aligned} \quad (4)$$

where  $e_0(k) = d(k) - X^T(k)W(k)$  is the optimal output error.

Substituting (1) into (4), we get

$$\Delta w(k+1) = [I - u(k)X(k)X^T(k)]\Delta w(k) + u(k)e_0(k)X(k) \quad (5)$$

With the assumption that the elements of  $X(k)$  is statistically independent of  $\Delta w(k)$  and  $e_0(k)$ , the expected error in the coefficient vector of the system can be simplified as:

$$E[\Delta w(k+1)] = \{I - u(k)R\}E[\Delta w(k)] \quad (6)$$

According to [7], (6) has the same form compared with  $E[\Delta V_{k+1}] = (I - uE[f'(n_k)]R)E[\Delta V_k]$ , so an effective step size  $u_{eff}$  is defined as

$$u_{eff} = uE[f'(n_k)] = u(k) \quad (7)$$

Thus we may conclude that the nonlinear error algorithm converges in the mean for

$$0 < u_0(1 + \beta M) < 2/\lambda_{\max} \quad (8)$$

where  $\lambda_{\max}$  is the maximum value among the eigenvalue of  $R$  (auto correlation of the input signal vector) and  $M$  is the

largest value of  $|e(k-1) - e(k)|$ , which means  $|e(k-1) - e(k)| \leq |e(k)| + |e(k-1)| \leq M$ .

#### B. Convergence Behavior of MSE

As mean value convergence of the coefficient vector  $w(k)$  is not the sufficient condition of the mean square error convergence. We expand the MSE with  $\Delta w(k)$  and  $\varepsilon_{\min}$  as follow:

$$MSE = E[e^2(k)] = \varepsilon_{\min} + E[\Delta w(k)\Delta w^T(k)] \quad (9)$$

where  $\varepsilon_{\min} = E[w^2(k)]$ . Hence, the convergence of MSE is equal to the convergence of covariance matrix of coefficient  $E[\Delta w(k)\Delta w^T(k)]$ . So according to (5), the covariance matrix of coefficient can be written as

$$\begin{aligned} \text{cov}[\Delta W(k+1)] &= E[W(k)W^T(k)] - E[u(k)]E[W(k)W^T(k)]R \\ &\quad - E[u(k)]RE[W(k)W^T(k)] + E[u^2(k)]\varepsilon_{\min}R \\ &\quad + E[u^2(k)]\{2RE[W(k)W^T(k)]R + \text{Rtr}\{RE[W(k)W^T(k)]\}\} \end{aligned} \quad (10)$$

where  $R = E[X(k)X^T(k)]$  is the auto correlation of the input signal vector. According to [8], (10) can be derived as:

$$\begin{aligned} \text{cov}[\Delta W'(k+1)] &= \text{cov}[\Delta W'(k)] - E[u(k)]\text{cov}[\Delta W'(k)]\Lambda \\ &\quad - E[u(k)]\Lambda \text{cov}[\Delta W'(k)] + E[u^2(k)]\sigma_n^2\Lambda \\ &\quad + E[u^2(k)]\{2\Lambda \text{cov}[\Delta W'(k)]\Lambda \\ &\quad + \Lambda \text{tr}\{\Lambda \text{cov}[\Delta W'(k)]\}\} \end{aligned} \quad (11)$$

where  $\text{cov}[\Delta W'(k)] = E[Q^T \Delta w(k)\Delta w^T(k)Q]$  and  $Q^T Q = Q Q^T = I$  ( $Q$  is the unitary matrix that diagonalizes  $R$  to  $\Lambda$  through a similarity transformation). Therefore MSE convergence is equal to the sum of the elements along diagonal line convergence. The following relationship can be derived

$$\text{cov}[\Delta W'(k+1)] = B \text{cov}[\Delta W'(k)] + E[u^2(k)]\sigma_n^2\Lambda \quad (12)$$

The elements of  $B$  are given below

$$b_{ij} = \begin{cases} E[u^2(k)]\sigma_n^2\lambda_i\lambda_j & \text{for } i \neq j \\ 1 - 2E[u(k)]\lambda_i + 2E[u^2(k)]\lambda_i^2 + E^2[u(k)]\lambda_i^2 & \text{for } i = j \end{cases} \quad (13)$$

To guarantee the convergence of MSE, the sum of the elements in any row of  $B$  must satisfy  $0 < \sum b_{ij} < 1$ . So the stable condition is

$$0 < 1 - 2E[u(k)]\lambda_i + 2E[u^2(k)]\lambda_i^2 + E^2[u(k)]\lambda_i^2 \sum_{j=0}^N \lambda_j < 1 \quad (14)$$

$$0 < u_0(1 + \beta M) < 1/(\lambda_{\max} + \text{tr}\{R\}) \quad (15)$$

where

$|e(k-1) - e(k)| \leq |e(k)| + |e(k-1)| \leq M = M_{\text{determination}}$ , is the determining level of signal '1'.

#### C. Misadjustment of the algorithm

The output signal error of adaptive filter is

$$e(k) = d(k) - w_0^T X(k) - \Delta w^T X(k) = \varepsilon(k) - \Delta W^T X(k) \quad (16)$$

With independence assumption, expected value of  $e^2(k)$  is

$$\begin{aligned} \varepsilon(k) &= E[e^2(k)] \\ &= \varepsilon_{\min} - 2E[\Delta W^T(k)]E[e_0(k)X(k)] \\ &\quad + E\{tr[X(k)X^T(k)\Delta W(k)\Delta W^T(k)]\} \end{aligned} \quad (17)$$

With equation  $tr(A \cdot B) = tr(B \cdot A)$  and  $R = E[X(k)X^T(k)]$ , (17) can be simplified as

$$\varepsilon(k) = \varepsilon_{\min} + E[\Delta w^T(k)Rw(k)] \quad (18)$$

Then the excess MSE of  $k$  th iteration is

$$\begin{aligned} \varepsilon(k) - \varepsilon_{\min} &= tr\{E[R\Delta W(k)\Delta W^T(k)]\} \\ &= tr\{E[\Lambda cov[\Delta W(k)]]\} \end{aligned}$$

According to (12) and (13),

$$\varepsilon(k) - \varepsilon_{\min} = \frac{u(k)\sigma_n^2 tr[R]}{1 - u(k)tr[R]} \quad (19)$$

So the excess MSE can be given by

$$\varepsilon(\infty) - \varepsilon_{\min} = \lim_{k \rightarrow \infty} [\varepsilon(k) - \varepsilon_{\min}] = \frac{u_0 \sigma_n^2 tr[R]}{1 - u_0 tr[R]} \quad (20)$$

If  $k$  is big enough, (20) can be written as

$$\varepsilon(\infty) - \varepsilon_{\min} = \lim_{k \rightarrow \infty} [\varepsilon(k) - \varepsilon_{\min}] = \frac{u_0 \sigma_n^2 tr[R]}{1 - u_0 tr[R]} \quad (21)$$

Then the misadjustment is obtained as

$$M = \frac{\varepsilon(\infty) - \varepsilon_{\min}}{\varepsilon_{\min}} = \frac{u(k)tr[R]}{1 - u(k)tr[R]} \quad (22)$$

Equation (22) is of the same form as that in Ref. [9].

#### IV. Simulation and Discussions

Simulations are conducted to evaluate the performance of the ESD-VSS LMS algorithm.

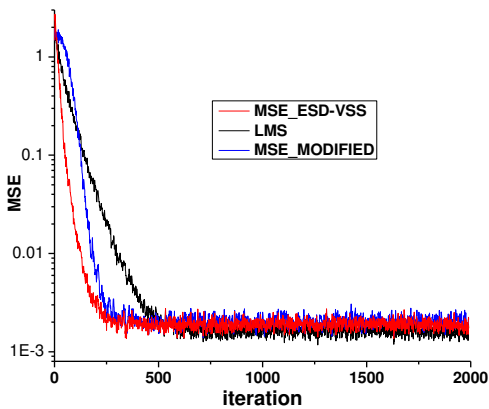


Fig. 2 Comparison of MSE behavior when  $W = 2.9$

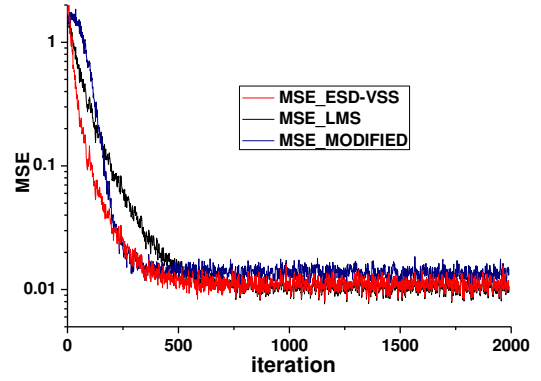


Fig. 3 Comparison of MSE behavior when  $W = 3.1$

#### A. Channel equalization

We test and compare the performance of the ESD-VSS algorithm with the standard LMS and the modified VSS [5] in adaptive channel equalization environment, of which the channel is described as

$$h(k) = 0.5 \left( 1 + \cos \frac{2\pi(k-2)}{W} \right) \quad k = 1, 2, 3$$

The parameter  $W$  is used to control the amount of amplitude distortion. Parameters of these algorithms are selected to produce a comparable level of misadjustment. Signal in this channel is corrupted by an additive white Gaussian noise  $noise(k)$ , which has zero mean and variance of  $\sigma^2 = 0.001$ . So the input signal of the adaptive filter can be described as  $X(k) = X_0(k) * h(k) + noise(k)$ . Fig.2 shows the result of MSE behaviors of the three algorithms with 11 taps and  $W = 2.9$ . The parameters for the ESD-VSS used are  $\mu = 0.02$ ,  $\beta = 2.3$  and  $\alpha = 0.97$ ,  $\beta = 2.3$ ,  $\gamma = 0.025$  (which are used in its original paper) are used for the MVSS algorithm. Besides, the step-size for standard LMS algorithm is  $\mu_0 = 0.02$ .

Figure 3 shows MSE curve of the three algorithms with distortion control parameter  $W = 3.1$  and the noise variance  $\sigma^2 = 0.005$ . Parameters for the modified LMS algorithm are  $\alpha = 0.958$ ,  $\beta = 0.99$ ,  $\gamma = 0.025$ . Parameter  $\beta$  of the ESD-VSS algorithm is 2.3. The other parameters remain unchanged.

According to Fig. 2 and Fig. 3, the ESD-VSS algorithm provides much faster convergence speed at comparable level of misadjustment in both cases, and realizes a faster convergence speed with much less complexity compared to modified LMS algorithm.

#### B. System identification

The ESD-VSS variable step-size LMS algorithm, ESD-VSS, is implemented in a system identification environment. The performance of the algorithm is compared with the standard LMS algorithm and MVSS algorithm [6]. Parameters of these algorithms are selected according to the recommended values in their publication and also with a consideration to

produce a comparable misadjustment. The unknown system is assumed to be finite impulse response (FIR) system and the length of coefficient vector is 11.

### 1) Uncorrelated signal

Signal applied to the system is a pseudorandom bit sequence, with zero mean value and unit variance. Additive noise  $noise(k)$  with the same mean value and variance is also applied to the simulation. Parameters of the ESD-VSS algorithm are  $\alpha = 2.6$   $u = 0.034$ . To obtain comparable misadjustment, for the modified algorithm we set parameters  $\alpha = 0.997$ ,  $\beta = 0.99$  and  $\gamma = 0.025$ . The step size of standard LMS algorithm is 0.01.

Fig 4 shows the behavior of the ESD-VSS algorithm, standard LMS algorithm and MVSS algorithm. Notice that the small misadjustment level is achieved by the algorithms, but the convergence rate of the standard LMS is not satisfying. The ESD-VSS and MVSS algorithm reduce the tradeoff between misadjustment and convergence rate. The ESD-VSS reaches steady state after about 200 iterations while the MVSS reaches steady state about 50 iterations later, which proves the efficiency of ESD-VSS algorithm.

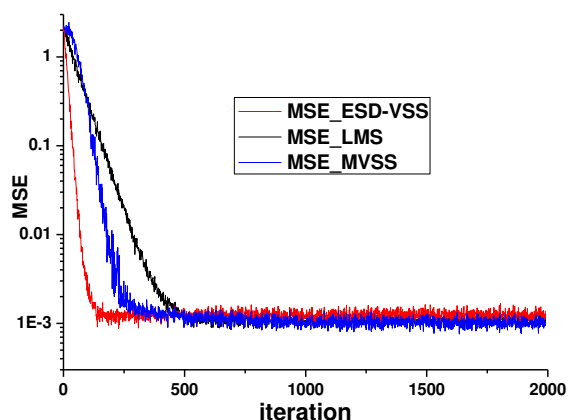


Fig.4 Comparison of MSE behavior for uncorrelated signal and low SNR

### 2) Uncorrelated signal with abrupt change in system coefficients

To check the robustness of the ESD-VSS algorithm, we set an abrupt coefficient change at the 1000<sup>th</sup> iteration. All the other related parameters are kept the same as simulation shown in Fig.4. Fig.5 compares the behaviors of the algorithms, which shows that the ESD-VSS algorithm converge faster and reaches steady state first with smaller computational complexity compared with the MVSS algorithm. The simulation demonstrated the ESD-VSS algorithm has better tracking performance after the abrupt coefficients change.

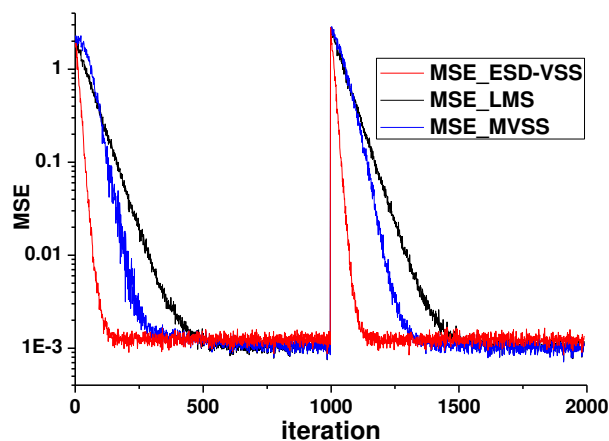


Fig. 5 MSE behavior for uncorrelated signal with abrupt change in system coefficients

## V. Conclusion

A new variable step size LMS algorithm (ESD-VSS) has been proposed, with a comprehensive consideration of both convergence speed and misadjustment. The variable step size is updated using the difference between the error signals. The simulation results show that the ESD-VSS.

ESD-VSS algorithm provides better performance over MVSS algorithm and the modified LMS algorithm, as well as standard LMS algorithm in the environment of channel equalization and system identification.

## References

- [1] Kwon Soon-Won, Bae Hyeon Min, "Variable precision distributed arithmetic (VPDA) MIMO equalization for power and area efficient 112 Gb/s optical DP-QPSK system", Journal of Lightwave Technology, vol. 31, issue.2, pp.282-294, 2013
- [2] Wu F.Y, F. Tong, "Non Uniform Norm Constraint LMS algorithm for Sparse System Identification", IEEE communications letters. vol.17, no.2, pp. 385-388, 2013.
- [3] YU RongGui, YE ShuiSheng. "A Modified Variable Step Size LMS Adaptive Filtering Algorithm and I Application in Noise Cancellation". Computer engineering and science, vol.31, no.4,2009.
- [4] Kwong, R.H, "A Variable Step Size LMS Algorithm", IEEE Trans. Signal Process. vol. 40, no. 7, pp. 1633-1642, 1992.
- [5] QU Yan-bin, MENG Fan-gang ect. "A new variable step size LMS adaptive filtering algorithm", IEEE International Symposium on Industrial Electronics, Vigo, Spain, pp.1601-1605, 2007
- [6] Zhao Shengkui, Man Zhihong, Khoo Suiyang, "A New Variable Step-Size Transform Domain LMS Algorithm with System Identification", IEEE International Conference on, Control and Automation, 2007. ICCA 2007. pp.2721 - 2724, 2007
- [7] Douglas, S.C, Meng, T.H. "optimum error nonlinearities for LMS adaptive, international conference on Acoustics,Speech,and Signal Processing", Albuquerque, NM, pp.1421-1424,1990
- [8] S. Haykin, "Adaptive Filter Theory", 4th ed. Englewood Cliffs, Nj, Prentice Hall, 2002.
- [9] Paulo S.R.Diniz, "adaptive filtering: algorithms and practical implementation", second edition. Kluwer Academic Publishers, 2002.