

# A variable step-size adaptive system identification algorithm based on modified decorrelation principle

Huamin Ge<sup>1, a</sup>, Qing Wan<sup>2, b</sup> and WeiWei Zhao<sup>1</sup>

<sup>1</sup> Nanjing University of Information Science & Technology , Nanjing 210044 , China

<sup>2</sup>Nanjing Automotive Engineering Research Institute , Nanjing 210028 , China

<sup>a</sup>ghm@nuist.edu.cn, <sup>b</sup>wanqing@saicmotor.com

**Keywords:** decorrelation; variable step-size LMS algorithm; adaptive system identification; Labview

**Abstract.** To overcome the defects of traditional adaptive Least-mean-square(LMS) algorithm, figure out the problem that the existing algorithms' system identification abilities would decline while their input signals are correlated with each other, a novel variable step-size and variable parameters adaptive system identification algorithm based on modified decorrelation principle is proposed. The design theories and process, the selection principles of the adjusting parameters and the performance analysis of the algorithm are all given below. In the algorithm, the rate of its convergence is improved by the modified decorrelation principle. Variable step size in the algorithm is based on hyperbolic tangent function, it is adjusted by the mean value of the gradient vector, so the system can modify its step size by the energy of the input signals and the real-time error. The variable step size also introduces the self-correlation of the real-time error to restrain the influence of the independent noises. Theory and simulation results verified the system identification abilities and the superiorities of the algorithm under different circumstances.

## Introduction

Adaptive filter is a special kind of Wiener filter who can automatically adjust its own parameters, this characteristic decides it can be used for adaptive system identification. However, there is a contradiction between the convergence speed and the steady-state MSE in traditional LMS algorithm. To overcome this contradiction, numerous variable step-size LMS algorithms were proposed. The step-size was proportional to the cross-correlation function of instantaneous estimation error and input signal in paper<sup>[2]</sup>, this algorithm improved the performances compared with NLMS with increase in computation cost, and it can't be used in real-time occasions. Paper<sup>[3]</sup> used the square form of instantaneous error to update step-size, this algorithm got a better performance in stationary environment but it was very sensitive to measurement noise. Paper<sup>[5]</sup> was the improvement of paper<sup>[4]</sup>, it showed us a simple form of Sigmoid function, this kind of step-size can get a faster convergence speed and better system tracking ability when there is an abrupt system variation, but its performance would get worse in lower SNR environment. Paper<sup>[7]</sup> was the improvement of paper<sup>[6]</sup>, it employed the self-correlation of instantaneous error signal to reduce the noise disturbance, it also introduced a compensating factor to make sure the step-size varies slowly in steady-state. Actually, the self-correlation of the error signal was very small in high order system, this would make the step-size of the algorithm diminish quickly and get the minimal value before system converged. Paper<sup>[8]</sup> was the improvement of paper<sup>[3]</sup>, it achieved a better performance, but the use of the square form of instantaneous error like paper<sup>[3]</sup> in step-size raised its sensitivity to measurement noise. Paper<sup>[9]</sup> used a quotient form of filtered versions of the quadratic error to adjust the step-size, but it was also very sensitive to measurement noise. Paper<sup>[10]</sup> was inconvenient to use for it proposed two different forms of variable step-size corresponding to stationary and nonstationary environment. Paper<sup>[11]</sup> proposed an algorithm with a gradient-based weighted average, but the weight of the system may fluctuate when the amplitude of input signal was too high. Paper<sup>[12]</sup> used the cost function to get the optimal step-size.

Based on the analysis above, considering the powerful functions of virtual instrument and the benefits of Hyperbolic tangent function, a novel variable step-size and variable parameters LMS adaptive system identification algorithm was proposed based on LabVIEW for real-time applications. A good algorithm must consider the characters of input signal, but paper<sup>[1-7]</sup> were all based on the hypothesis that input signals were all independent to each other. Although paper<sup>[8-12]</sup> had better performance with correlated input signal, they didn't deal with the input signal and their performance would decline quickly when the correlation coefficient greater than 0.7. In practice, signals are correlated to each other, the performance and the convergence speed of a system would decline sharply if there are no necessary measures. Therefore, in this paper, modified decorrelation principle was used to decorrelate the correlation between input vectors before we identify the system. Theory and simulation results verified the superiorities of the algorithm proposed in this paper.

## Algorithm formulation and its performance analysis

### Principle of adaptive system identification

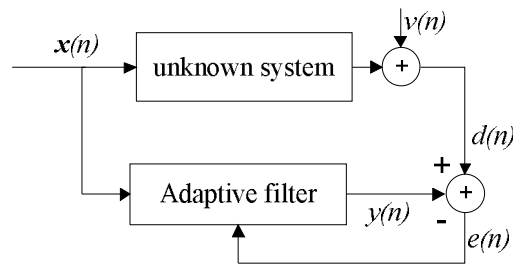


Fig.1. Schematic diagram of adaptive system identification

$x(n)$  in Fig.1 is the tap-delayed input vector at  $n$  moment, it comes from a zero-mean general stationary random process with  $R$  as its correlation matrix,  $x(n) = [x(n), \dots, x(n-N+1)]^H$ .  $w(n) = [w_0(n), w_1(n), \dots, w_{N-1}(n)]^H$  is the corresponding vector consisting of the coefficients of the adaptive filter that models the unknown system.  $d(n)$  represents the desired response consisting of system output plus the measurement noise.  $y(n)$  is the output signal.  $e(n)$  is the instantaneous estimation error.  $v(n)$  denotes measurement noise statistically independent of  $x(n)$ .  $w_o(n)$  is a vector represents the optimal Wiener solution of the unknown system. All signals above can be complex.  $N$  denotes the length of adaptive filter.  $A^H$  denotes conjugate transpose of  $A$ .  $A^T$  denotes the transpose of  $A$ .  $a^*$  denotes complex conjugate of  $a$ . Letters written in bold means that the signal is a vector. From Fig.1, we get:

$$y(n) = w^H(n)x(n) \quad (1)$$

$$e(n) = d(n) - y(n) \quad (2)$$

$$d(n) = w_o^H(n)x(n) + v(n) \quad (3)$$

Adaptive filter algorithm was based on Least-mean-square(LMS) and the Steepest Descent Approach(SDA). Define cost function as  $J = E[e(n)e^*(n)]$ , substituting (2) into  $J$ , we can get the vector form of  $J$  as:

$$J(w) = S_d^2 - w^H P - P^H w + w^H R w \quad (4)$$

Where:

$$P = E[x(n)d^*(n)] \quad (5)$$

$$R = E[x(n)x^H(n)]$$

$R$  is the input correlation matrix.  $P$  is the cross-correlation of  $x(n)$  and  $d(n)$ . The variance of zero-mean desired response  $d(n)$  is  $S_d^2 = E[|d(n)|^2]$ . SDA uses the minus gradient direction of the cost function to update weight vector  $w(n)$ . Represent gradient vector as  $g = \nabla J(w) = \frac{\partial(J(w))}{\partial(w)}$ , then we get the following weight vector update equation of SDA:

$$w(n+1) = w(n) - \frac{1}{2} m g(n) \quad (6)$$

Computing the partial derivatives on both side of (4) yields:

$$\mathbf{g}(n) = \nabla J(n) = -2\mathbf{P} + 2\mathbf{R}\mathbf{w} \quad (7)$$

Accurately, we cannot get  $\mathbf{R}$  and  $\mathbf{P}$  when the system runs in an unknown environment, so we need adaptive filter, who can adjust its own coefficients as its running environment changes. In fact, adaptive filter adopts square error instead of mean-square error, that is, uses the transient estimates of  $\mathbf{R}$  and  $\mathbf{P}$  instead of  $\mathbf{R}$  and  $\mathbf{P}$ . Then (5) can be rewritten as:

$$\begin{aligned} \mathbf{P} &= \mathbf{x}(n)\mathbf{d}^*(n) \\ \mathbf{R} &= \mathbf{x}(n)\mathbf{x}^H(n) \end{aligned} \quad (8)$$

Eq.(7) can be rewritten as:  $\nabla J(n) = -2\mathbf{x}(n)\mathbf{d}^*(n) + 2\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}(n)$ . Then from eq.(6) we have:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + m\mathbf{x}(n)[\mathbf{d}^*(n) - \mathbf{x}^H(n)\mathbf{w}(n)] = \mathbf{w}(n) + m\mathbf{e}^*(n)\mathbf{x}(n) \quad (9)$$

Adaptive system identification algorithm is combined by (1),(2),(3) and (9). This algorithm was wrote in complex form, real form is its special case. LMS is based on general stationary random process, but it also apply to definite or nonstationary environment. Apply this algorithm into system identification, we can get the optimal Wiener solution  $\mathbf{w}(n) = \mathbf{w}_o(n)$  as long as a suitable step-size  $m$  or variable step-size  $m(n)$  was chosen, then the system we identified can be write as  $y(n) = \mathbf{w}^H(n)\mathbf{x}(n)$ .

### **Variable step-size & Variable parameters adaptive system identification algorithm**

According to the principle of variable step-size<sup>[2]</sup>, we take the following step-size adjust scheme:

$$m(n) = b \tanh(a^k \|F(n)\|^m) \quad (10)$$

Where  $\|\cdot\|$  denotes abs.  $a$  controls the shape of step-size.  $b$  controls the range of step-size.  $a$  and  $b$  are all functions vary with  $e(n)$ , that is, controlling factors  $a(n)$  and  $b(n)$ . From (2) and (3), we get:

$$e(n) = v(n) - [\mathbf{w}^H(n) - \mathbf{w}_o^H(n)]\mathbf{x}(n) \quad (11)$$

Considering  $\rho(n) = \mathbf{w}^H(n) - \mathbf{w}_o^H(n)$  and  $S_v^2$  as the power of measurement noise, we get:

$$\begin{aligned} E[\mathbf{x}(n)v(m)] &= 0 \\ E[v(n)v^*(m)] &= S_v^2 \delta(n-m) \end{aligned} \quad (12)$$

then:

$$\begin{aligned} E[e(n)e^*(n)] &= E[\rho^H(n)\mathbf{x}(n)\mathbf{x}^H(n)\rho(n)] + S_v^2 \\ E[e(n)e^*(n-1)] &= E[\rho^H(n)\mathbf{x}(n)\mathbf{x}^H(n-1)\rho(n-1)] \\ E[e(n)e^*(n-1) + e(n)e^*(n)] &= E[e(n)e^*(n-1)] + E[e(n)e^*(n)] \\ E[\mathbf{x}(n)e^*(n)] &= E[-\rho^H(n)\mathbf{x}(n)\mathbf{x}^H(n)] \end{aligned} \quad (13)$$

The convergence in mean sense of a system can be represented by  $E[m(n)]$ . According to Eq.(13), if we introduce  $e(n)$  or the positive integer power of  $e(n)$  in  $m(n)$ , the system would be very sensitive to measurement noise  $v(n)$ . If we introduce  $e(n)e^*(n-1)$  only, the system would not be so sensitive to  $v(n)$ , but the self-correlation of  $e(n)$  would getting smaller as the length of filter getting longer, this would surely smaller the step-size, slow down the convergence speed and deteriorate the performance of the algorithm in limited time. If we introduce  $e(n)e^*(n-1) + e(n)e^*(n)$ , the performance in higher order system would be improved, but it would be still affected by  $v(n)$ .

From Eq.(7) we know that the steepest drop direction of adaptive LMS algorithm is  $-\nabla J(n) = 2\mathbf{x}(n)e^*(n)$ , make  $\nabla J(n) = 0$ , we can get the optimal solution of  $\mathbf{w}(n)$ . Furthermore, we notice that the optimal Wiener solution is  $\mathbf{x}(n)e^*(n)$ , from Eq.(13) we know  $E[\mathbf{x}(n)e^*(n)]$  is insensitive to  $v(n)$ . So we introduce  $\mathbf{x}(n)e^*(n)$  to adjust step-size in this paper. Actually, the cross-correlation of  $\mathbf{x}(n)$  and  $e(n)$  is stronger than the self-correlation of  $e(n)$ , hence, the step-size would not decrease sharply in high order system. Besides, this expression can stronger the tracking ability for it considers input signal in variable step-size formula. In this paper, we adopt the average value of the sum of Gradient vector to update step-size. This formula can avoid the weight vector from fluctuating when the amplitude of input signal is too high or fluctuate significantly, so it can improve the performance of the algorithm.

Based on the analyses above, we can get the formula of  $a(n)$ ,  $b(n)$  and  $F(n)$  in Eq.(10) as follows:  $a(n)$  is the same as paper<sup>[18]</sup>, this is mainly because it is proportional to  $e(n)$  and we should

ensure smaller  $a(n)$  as far as possible to avoid the step-size oversizing along with  $F(n)$ .  $b(n)$  uses the recursive form of the self-correlation of  $e(n)$  to restrain the effect of independent measurement noise furtherly. Besides, when the algorithm converging to optimal value, the self-correlation of  $e(n)$  will getting smaller to smaller the steady-state error:

$$F(n) = \frac{\sum_{i=0}^{N-1} x(n-i)e^*(n)}{N} \quad (14)$$

$$a(n) = \left\| \frac{e(n)}{e(n-1)} \right\|$$

$$b(n) = rb(n-1) + (1-r)\|e(n)e^*(n-1)\|$$

(15)

When  $e(n-1)$  in (15) equals zero,  $a(n)$  approach to  $\infty$ , this would enforce  $\tanh(\infty)$  approach to 1. So in this case, the algorithm would not unsteady but could reflect the real state of the system more accurately.  $r$  is a constant approach to 1, it controls  $b(n)$ ,  $0 < r < 1$ , in this paper,  $r = 0.98$ .

Considering the design principle of the two adjusting factors  $m$  and  $k$  in (10).  $m$  controls the shape of the function, we can make the step-size vary slowly in steady-state by changing its value. When  $\|F(n)\| < 1$ , the step-size would getting larger as the value of  $m$  getting smaller. So, if the value of  $m$  is too large, the step-size might equals zero even if the system have not steady, this would surely bigger the steady-state error. If the value of  $m$  is too small, the step-size would be too big even if  $e(n)$  is tiny, still, this would bigger the steady-state maladjustment error and interfere the convergence performance of the system. When  $\|F(n)\| > 1$ , the functions of  $m$  is opposite to  $\|F(n)\| < 1$ . Factor  $k$  controls the varying speed of the step-size via adjusting  $a(n)$ . The functions of  $k$  when  $a(n) < 1$  and  $a(n) > 1$  are corresponding to the functions of  $m$  when  $\|F(n)\| < 1$  and  $\|F(n)\| > 1$ . So, the choosing of  $m$  and  $k$  should considering the following two occasions: Bigger  $m$  and  $k$  if we need higher convergence speed; Smaller  $m$  and  $k$  if we need higher convergence accuracy.

The variable step-size of this algorithm is combined by (10), (14) and (15). Replace  $m$  in eq.(9) as  $m(n)$  in eq.(10) can finally get our variable step-size LMS adaptive algorithm. Theoretically, the new algorithm has simple structure for it was adjusted by Hyperbolic tangent function;  $e(n)$  and the average value of Gradients vector make the step-size of the algorithm getting smaller along with the convergence progress and would not decrease sharply in high order system, this is corresponding to the principle of variable step-size: speed up the rate of convergence using large step-size at the initial stages, and reduce the estimation error with small step-size in steady-state; The algorithm is able to reduce its sensitivity to the measurement noise  $v(n)$  and keep in good performances in low SNR environment for it was adjusted by factors independent to  $v(n)$ ; The algorithm is able to response the variation of the input signals and the whole system no matter in stationary or nonstationary environment for it was adjusted by Gradients vector. However, our algorithm have not considered correlated input signals yet, this will be discussed in latter sections.

### **Performance of the algorithm**

#### *convergence in mean sense*

The algorithm discussed in the last section was based on general stationary random process, but in nonstationary environment, the weight vector of a system was given by a lowpass filter who was driven by a zero-mean white Gaussian  $\gamma(n)$  with  $s_g^2$  as its variance. When  $s_g^2 = 0$ , the system would turn into stationary, so we would like to learn the convergence in mean sense of the algorithm in nonstationary environment:

$$w_o(n+1) = w_o(n) + \gamma(n) \quad (16)$$

From (11) we get:

$$e(n) = v(n) - \rho(n)x(n) \quad (17)$$

Where  $\rho(n) = \mathbf{w}^H(n) - \mathbf{w}_o^H(n)$  is the weight error vector. From (9),(16)and(17),we can get the transient behavior of  $\rho(n)$ :

$$\begin{aligned}
\rho(n+1) &= \mathbf{w}^H(n+1) - \mathbf{w}_o^H(n+1) \\
&= \mathbf{w}^H(n) + \mathbf{m}(n)e^*(n)\mathbf{x}^H(n) - \mathbf{w}_o^H(n) - \gamma^H(n) \\
&= \rho(n) + \mathbf{m}(n)e^*(n)\mathbf{x}^H(n) - \gamma^H(n) \\
&= \rho(n) + \mathbf{m}(n)[v^*(n) - \rho(n)\mathbf{x}(n)]\mathbf{x}^H(n) - \gamma^H(n) \\
&= \rho(n) + \mathbf{m}(n)v^*(n)\mathbf{x}^H(n) - \mathbf{m}(n)\rho(n)\mathbf{x}(n)\mathbf{x}^H(n) - \gamma^H(n)
\end{aligned} \tag{18}$$

We would like to write the input correlation matrix  $\mathbf{R}$  as  $\mathbf{R} = \mathbf{Q}\mathbf{A}\mathbf{Q}^H$  with  $\mathbf{A}$  the diagonal eigenvalue matrix and  $\mathbf{Q}$  the unitary eigenvector matrix.

Considering  $\mathbf{A} = \mathbf{Q}^H \mathbf{R} \mathbf{Q}$  and  $\mathbf{Q}^H = \mathbf{Q}^{-1}$ , applying the unitary transformation  $\mathbf{Q}^H$  on both sides of (18) yields:

$$\begin{aligned}
\hat{\rho}(n+1) &= \hat{\rho}(n) + \mathbf{m}(n)v^*(n)\hat{\mathbf{x}}^H(n) - \mathbf{m}(n)\hat{\rho}(n)\hat{\mathbf{x}}(n)\hat{\mathbf{x}}^H(n) - \hat{\gamma}^H(n) \\
&= [\mathbf{I} - \mathbf{m}(n)\hat{\mathbf{x}}(n)\hat{\mathbf{x}}^H(n)]\hat{\rho}(n) + \mathbf{m}(n)v^*(n)\hat{\mathbf{x}}^H(n) - \hat{\gamma}^H(n)
\end{aligned} \tag{19}$$

where  $\hat{\rho}(n+1) = \mathbf{Q}^H \rho(n+1)$ ,  $\hat{\rho}(n) = \mathbf{Q}^H \rho(n)$ ,  $\hat{\gamma}(n) = \mathbf{Q}^H \gamma(n)$  and  $\hat{\mathbf{x}}(n) = \mathbf{Q}^H \mathbf{x}(n)$ . Actually,  $\hat{r}_k(n+1)$  is the  $k$  th natural mode of our LMS adaptive filter. When system getting stable, natural modes will disappear, weight vector will approach to optimal Wiener solution  $\mathbf{w}_o(n)$  and  $\hat{\rho}(n+1) \approx \hat{\rho}(n)$ . So the convergence in mean sense can be learned base on the expectation of  $\hat{\rho}(n)$ . Take expectation on both sides of(19) yields:

$$E[\hat{\rho}(n+1)] = E[\mathbf{I} - \mathbf{m}(n)\hat{\mathbf{x}}(n)\hat{\mathbf{x}}^H(n)]E[\hat{\rho}(n)] - E[\hat{\gamma}^H(n)] \tag{20}$$

Eq.(20) is stable in the mean sense only if the eigenvalues of the matrix  $E[\mathbf{I} - \mathbf{m}(n)\hat{\mathbf{x}}(n)\hat{\mathbf{x}}^H(n)]$  are all within the unit circle, suppose  $\mathbf{D}(n) = E[\mathbf{I} - \mathbf{m}(n)\hat{\mathbf{x}}(n)\hat{\mathbf{x}}^H(n)]$ , this problem can be solved as follows:

$$\det[\mathbf{D}(n) - h\mathbf{I}] = \det[\mathbf{I} - E[\mathbf{m}(n)]\mathbf{A} - h\mathbf{I}] = \prod_{i=1}^M (1 - h - E[\mathbf{m}(n)]I_i) = 0 \tag{21}$$

where  $h$  denotes the eigenvalues of  $\mathbf{D}(n)$ ,  $\det$  means get the determinant,  $I_i$  is the  $i$ th eigenvalue of the input correlation matrix  $\mathbf{R}$ . We can ensure all the solutions of (21) within the unit circle by making sure both the smallest and the largest solutions to be within the unit circle. That is:

$$-1 < 1 - E[\mathbf{m}(n)]I_{\max} < 1 - E[\mathbf{m}(n)]I_{\min} < 1 \tag{22}$$

From (22),we can easily find the following bound on the mean step-size:

$$0 < E[\mathbf{m}(n)] < \frac{2}{I_{\max}} \tag{23}$$

Eq.(23) guarantees the convergence of our algorithm in the mean sense for nonstationary environments. Stationary is its special case, that is,  $s_g^2 = 0$ . For the case of a constant step-size  $m$ , Eq.(23) is the same as the condition for the mean convergence of the fixed step-size LMS algorithm.

### Analysis of steady-state performance

$J_{ex}(n)$  denotes the excess mean square error(EMSE) of a system, which is given by paper<sup>[15]</sup> as:

$$J_{ex}(n) = \text{tr}\{\mathbf{R} \cdot \mathbf{K}(n)\} = \sum_{k=1}^N I_k E\left[\left|\hat{r}_k(n+1)\right|^2\right] \tag{24}$$

Where  $\mathbf{K}(n) = E[\rho(n)\rho^T(n)]$ ,  $\lambda$  denotes a column vector consisted by the eigenvalues of  $\mathbf{R}$ ,  $\mathbf{c}(n)$  denotes a matrix consisted by  $E\left[\left|\hat{r}_k(n+1)\right|^2\right]$ ,  $E\left[\left|\hat{r}_k(n+1)\right|^2\right]$  is the secondary moment of the  $k$  th natural mode, then:

$$J_{ex}(n) = \lambda^H \mathbf{c}(n) \tag{25}$$

$$J(n) = J_{ex}(n) + J_{\min} \tag{26}$$

$$J_{\min} = E[v^2(n)] \tag{27}$$

$$\text{Misadjustment: } M = J_{\text{ex}}(\infty)/J_{\min} \quad (28)$$

Next, we shall learn the steady-state performance of the algorithm in terms of the misadjustment.

Considering  $\hat{\rho}(n)$  and  $\hat{\gamma}(n)$ ,  $v(n)$  and  $x(n)$  are statistically independent and

$E[\hat{x}(n)\hat{x}^H(n)] = E[\mathbf{Q}^H x(n)x^H(n)\mathbf{Q}] = E[\mathbf{Q}^H \mathbf{R}\mathbf{Q}] = \mathbf{A}$ , multiplying both sides of (19) by  $\hat{\rho}^H(n+1)$  and taking the expectations yields:

$$\mathbf{C}(n+1) = \mathbf{C}(n) - E[\mathbf{m}(n)]\mathbf{C}(n)\mathbf{A} - E[\mathbf{m}(n)]\mathbf{A}\mathbf{C}(n) + E[\mathbf{m}^2(n)]J_{\min}\mathbf{A} + E[\mathbf{B}(n)] + \mathbf{S}_g^2 \cdot \mathbf{I} \quad (29)$$

$$E[\mathbf{B}(n)] = E[\mathbf{m}(n)\hat{\rho}^H(n)\hat{x}^H(n)\hat{x}(n)\mathbf{m}(n)\hat{\rho}(n)\hat{x}(n)\hat{x}^H(n)] \quad (30)$$

Where  $E[\hat{\rho}^H(n+1)\hat{\rho}(n+1)] = \mathbf{C}(n+1)$  denotes the transformed weight error covariance of  $\rho(n)$ ,  $\mathbf{c}(n)$  is a column vector consisted by the diagonal elements of  $\mathbf{C}(n)$ . From the Gaussian moment factoring theorem we can get:

$$E[\mathbf{B}(n)] = E[\mathbf{m}^2(n)][2\mathbf{A}\mathbf{C}(n)\mathbf{A} + \text{tr}[\mathbf{A}\mathbf{C}(n)]] \quad (31)$$

Substituting (31) into (29) and denote  $c_{ii}(n)$  as the  $i$ th element of the diagonal elements of  $\mathbf{C}(n)$ ,  $\mathbf{I}$  as a column vector of 1's with the same length of  $\lambda$ , then:

$$c_{ii}(n+1) = [1 - 2E[\mathbf{m}(n)]I_i + 2E[\mathbf{m}^2(n)]I_i^2]c_{ii}(n) + E[\mathbf{m}^2(n)]I_i \sum_{j=1}^n I_j c_{jj}(n) + E[\mathbf{m}^2(n)]J_{\min}I_i + \mathbf{S}_g^2 \cdot 1 \quad (32)$$

Note that in (32), the eigenvalue decomposition theorem was used, that is, the trace of a matrix equals the sum of its eigenvalues. In the steady-state, we assume  $\lim_{n \rightarrow \infty} \mathbf{c}(n+1) = \lim_{n \rightarrow \infty} \mathbf{c}(n)$ , so:

$$\begin{aligned} \mathbf{c}(\infty) &= [\mathbf{I} - \mathbf{L}(\infty)]^{-1} \{E[\mathbf{m}^2(\infty)]J_{\min}\boldsymbol{\lambda} + \mathbf{S}_g^2 \cdot \mathbf{I}\} \\ \mathbf{L}(n) &= \text{diag}[l_1, l_2, \dots, l_n] + E[\mathbf{m}^2(n)]\boldsymbol{\lambda}\boldsymbol{\lambda}^H \\ l_i &= 1 - 2E[\mathbf{m}(n)]I_i + 2E[\mathbf{m}^2(n)]I_i^2 \end{aligned} \quad (33)$$

Make  $\mathbf{f} = \text{diag}[1 - l_1, 1 - l_2, \dots, 1 - l_n]$ , then:

$$\mathbf{c}(\infty) = [\mathbf{f} - E[\mathbf{m}^2(\infty)]\boldsymbol{\lambda}\boldsymbol{\lambda}^H]^{-1} \{E[\mathbf{m}^2(\infty)]J_{\min}\boldsymbol{\lambda} + \mathbf{S}_g^2 \cdot \mathbf{I}\} \quad (34)$$

Make use of (25) and the matrix inversion lemma, we get:

$$J_{\text{ex}}(\infty) = \{\mathbf{I} - E[\mathbf{m}^2(\infty)]\boldsymbol{\lambda}^H \mathbf{f}^{-1} \boldsymbol{\lambda}\}^{-1} \cdot \{E[\mathbf{m}^2(\infty)]J_{\min}\boldsymbol{\lambda}^H \mathbf{f}^{-1} \boldsymbol{\lambda} + \mathbf{S}_g^2 \boldsymbol{\lambda}^H \mathbf{f}^{-1} \mathbf{I}\} \quad (35)$$

Where:

$$\begin{aligned} \mathbf{I}^H \mathbf{f}^{-1} \mathbf{I} &= \sum_{i=1}^N \frac{I_i}{2E[\mathbf{m}(\infty)] - 2E[\mathbf{m}^2(\infty)]I_i} \\ \mathbf{I}^H \mathbf{f}^{-1} \mathbf{I} &= \sum_{i=1}^N \frac{1}{2E[\mathbf{m}(\infty)] - 2E[\mathbf{m}^2(\infty)]I_i} \end{aligned} \quad (36)$$

Substituting (36) into (35), neglecting the terms involved with  $E[\mathbf{m}^2(\infty)]$  that are very smaller compared to other terms of misadjustment we have:

$$J_{\text{ex}}(\infty) = J_{\min} \text{tr}[\mathbf{R}] \frac{E[\mathbf{m}^2(\infty)]}{2E[\mathbf{m}(\infty)]} + \frac{N\mathbf{S}_g^2}{2E[\mathbf{m}(\infty)]} \quad (37)$$

Next, we shall compute  $E[\mathbf{m}(\infty)]$  and  $E[\mathbf{m}^2(\infty)]$ , respectively. Taking expectations on both sides of (10) and assuming that  $n$  approaches to infinity yields:

$$\lim_{n \rightarrow \infty} E[\mathbf{m}(n)] = \lim_{n \rightarrow \infty} E[b(n) \tanh(a(n)^k \|F(n)\|^m)] \quad (38)$$

Rewriting (38) into Taylor series and making  $X = a(n)^k \|F(n)\|^m$ , we can approximate the steady-state mean step-size as follows:

$$\lim_{n \rightarrow \infty} E[m(n)] \approx \lim_{n \rightarrow \infty} E[b(n)(X - \frac{X^3}{3} + \frac{2X^5}{15} - \frac{17X^7}{315})] \quad (39)$$

Note that in (39) we only take the first four of the Taylor series for the adjusting functions of  $m$  and  $k$  make  $X < 1$ . From (39), we can approximate the steady-state mean square step-size as follows:

$$\begin{aligned} \lim_{n \rightarrow \infty} E[m^2(n)] &\approx \lim_{n \rightarrow \infty} E[b^2(n)(X - \frac{X^3}{3} + \frac{2X^5}{15} - \frac{17X^7}{315})^2] = \lim_{n \rightarrow \infty} E[b^2(n)X^2 - \frac{2}{3}b^2(n)X^4 + \\ &\frac{51}{135}b^2(n)X^6 - \frac{62}{315}b^2(n)X^8 + \frac{254}{4725}b^2(n)X^{10} - \frac{68}{4725}b^2(n)X^{12} + (\frac{17}{315})^2 b^2(n)X^{14}] \end{aligned} \quad (40)$$

To proceed further, we introduce the assumption:

$$\lim_{n \rightarrow \infty} E[b^p(n)X^q] = \lim_{n \rightarrow \infty} E[b^p(n)a(n)^{q \times k} \|F(n)\|^{q \times m}] \approx \lim_{n \rightarrow \infty} E[b^p(n)] \lim_{n \rightarrow \infty} E[a(n)^{q \times k}] \lim_{n \rightarrow \infty} E[\|F(n)\|^{q \times m}]$$

Where  $p$  and  $q$  are all positive integer. This assumption cannot really hold for the whole algorithm, but approximately true since  $a(n)^k$ ,  $b(n)$  and  $\|F(n)\|^m$  will vary slowly for  $e(n)$  is tiny in steady-state. Using the properties of mathematic expectation and the assumption above, we can figure out (39) and (40):

1) Ensuring  $E[b(\infty)]$  and  $E[b^2(\infty)]$ :

In the steady-state, we can still assume that  $\lim_{n \rightarrow \infty} E[b(n)] = \lim_{n \rightarrow \infty} E[b(n-1)]$  and  $\lim_{n \rightarrow \infty} E[e^*(n)] = \lim_{n \rightarrow \infty} E[e^*(n-1)]$ . Taking the expectation on both sides of (15) and assuming that  $n$  approaches to infinity yields:

$$\lim_{n \rightarrow \infty} E[b(n)] = r \lim_{n \rightarrow \infty} E[b(n-1)] + (1-r) \lim_{n \rightarrow \infty} E[\|e(n)e^*(n-1)\|] \quad (41)$$

So  $E[b(\infty)] = E[e(\infty)e^*(\infty)]$ . From the definition of the cost function we know  $E[e(\infty)e^*(\infty)] = J(\infty) = J_{ex}(\infty) + J_{min}$ , then:

$$E[b(\infty)] = J_{ex}(\infty) + J_{min} \quad (42)$$

On the other hand:

$$\begin{aligned} \lim_{n \rightarrow \infty} E[b^2(n)] &= r^2 \lim_{n \rightarrow \infty} E[b^2(n-1)] + \\ &2r(1-r) \lim_{n \rightarrow \infty} E[b(n-1)\|e(n)e^*(n-1)\|] + (1-r)^2 \lim_{n \rightarrow \infty} E[\|e(n)e^*(n-1)\|^2] \end{aligned} \quad (43)$$

For  $1-r$  is tiny, neglect the last term of (43) we get:

$$E[b^2(\infty)] \approx \frac{2r(1-r)E[b(\infty)e(\infty)e^*(\infty)]}{1-r^2} \quad (44)$$

2) Ensuring  $E[a(\infty)^{q \times k}]$ :

Taking the expectation on both sides of (15) and make  $n \rightarrow \infty$  yields:

$$\lim_{n \rightarrow \infty} E[a^{q \times k}(n)] = \lim_{n \rightarrow \infty} E[\left\| \frac{e(n)}{e(n-1)} \right\|^{q \times k}] \approx 1 \quad (45)$$

3) Ensuring  $E[\|F(\infty)\|^{q \times m}]$ :

In steady-state  $e(n)$  approaches to zero and approximately independent to  $x(n)$ . when  $q \times m = 1$ , form (14) we get:

$$\lim_{n \rightarrow \infty} E[\|F(n)\|] = \lim_{n \rightarrow \infty} E[\frac{\sum_{i=0}^{N-1} x(n-i)e^*(n)}{N}] \approx \lim_{n \rightarrow \infty} E[\frac{\sum_{i=0}^{N-1} x(n-i)}{N}] \lim_{n \rightarrow \infty} E[\|e^*(n)\|] \quad (46)$$

For the length of the adaptive filter is not so big to use central-limit theorem to proceed

$\lim_{n \rightarrow \infty} E[\frac{\sum_{i=0}^{N-1} x(n-i)}{N}]$ , so we would like to figure out  $\lim_{n \rightarrow \infty} E[\|e^*(n)\|]$  first. Input vector is statistically independent to  $v(n)$  and in steady-state independent to  $\rho(n)$ . So from (17) we have:

$$\lim_{n \rightarrow \infty} E[\|e^*(n)\|] = \lim_{n \rightarrow \infty} E[\|-\rho(n)x(n)\|] \approx \lim_{n \rightarrow \infty} \|E[\rho(n)]E[x(n)]\| \quad (47)$$

For  $x(n)$  is zero-mean, that is

$$E[x(n)] = 0 \quad (48)$$

So  $\lim_{n \rightarrow \infty} E[\|e^*(n)\|] \approx 0$ , then  $\lim_{n \rightarrow \infty} E[\|F(n)\|] \approx 0$ . In general case:

$$\lim_{n \rightarrow \infty} E[\|F(n)\|^{q \times m}] \approx \lim_{n \rightarrow \infty} E\left[\frac{\left\|\sum_{i=0}^{N-1} x(n-i)\right\|^{q \times m}}{N^{q \times m}}\right] \lim_{n \rightarrow \infty} E[\|e^*(n)\|^{q \times m}] \quad (49)$$

$$\lim_{n \rightarrow \infty} E[\|e^*(n)\|^{q \times m}] = E[\|-\rho(n)x(n)\|^{q \times m}] \approx \|E[\rho(n)]E[x(n)]\|^{q \times m} = 0 \quad (50) \quad \text{So } \lim_{n \rightarrow \infty} E[\|F(n)\|^{q \times m}] \approx 0.$$

Based on the above analyses, we know  $\lim_{n \rightarrow \infty} E[m(n)]$  is a constant approaches to zero and so does

$\lim_{n \rightarrow \infty} E[m^2(n)]$ . Make  $\lim_{n \rightarrow \infty} E[m(n)] = e$  and  $\lim_{n \rightarrow \infty} E[m^2(n)] = e$ , substituting them into (37) we can get the

misadjustment as follows:

$$M = \frac{J_{ex}(\infty)}{J_{min}} = \text{tr}[\mathbf{R}] \frac{e}{2e} + \frac{NS_g^2}{2eJ_{min}} \quad (51)$$

In steady-state:

$$M = \frac{J_{ex}(\infty)}{J_{min}} = \text{tr}[\mathbf{R}] \frac{e}{2e} \quad (52)$$

Eq.(51) and (52) show that the misadjustment, the steady-state mean step-size and the steady-state mean square step-size are all free from the MSE no matter in stationary or nonstationary environment. This is mainly because our algorithm adopted gradient vector to adjust step-size, so the MSE would approach to zero in steady-state. These analytic results corroborate that our algorithm is very insensitive to measurement noise. The superior properties are also indicated.

## Decorrelation theorem

The elements of the input vector in the above variable step-size adaptive system identification algorithm are the displacement form of  $x(n-N+1)$ , it means that the input signals are all correlated to each other. Simulation results indicated that the performance of the algorithm would decline rapidly if we neglect this problem, and would getting worse as the correlation getting stronger. So we should decorrelate the input signals before we go on into the adaptive algorithm.

Correlation coefficient denotes the correlate degree of signals, which is given by paper<sup>[16]</sup> as:

$$a_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{E[(x - m_x)(y - m_y)]}{\sqrt{E[(x - m_x)]^2 E[(y - m_y)]^2}} \quad (53)$$

Where  $s_{xy}$  is the covariance of random variables  $x$  and  $y$ ,  $m_x$  and  $m_y$  are their mean values,  $s_x$  and  $s_y$  are their standard deviations, respectively. Applying Cauchy-Schwarz we know  $|a_{xy}| \leq 1$ .  $|a_{xy}| = 1$  means  $x$  and  $y$  are ideal linear correlation,  $|a_{xy}| = 0$  means  $x$  and  $y$  are totally independent to each other,  $0 < |a_{xy}| < 1$  means  $x$  and  $y$  are correlated to each other.

Applying this definition into input vector  $x(n)$  and  $x(n-1)$ , paper<sup>[17]</sup> defined the self-correlation coefficient of them as  $a(n) = \frac{x^H(n)x(n-1)}{x^H(n-1)x(n-1)}$ . However, sometimes  $|a(n)| > 1$ , this is not confirm to the definition in (53). So in this paper, we use the theory of transient estimate in section 2.1 and following the definition in (53) to get the self-correlation coefficient of  $x(n)$  and  $x(n-1)$  as follows:

$$b(n) = \frac{E[(x(n) - m_{x(n)})(x(n-1) - m_{x(n-1)})]}{\sqrt{E[(x(n) - m_{x(n)})]^2 E[(x(n-1) - m_{x(n-1)})]^2}} = \frac{(x(n) - m_{x(n)})(x(n-1) - m_{x(n-1)})}{\sqrt{(x(n) - m_{x(n)})^2 (x(n-1) - m_{x(n-1)})^2}} \quad (54)$$

For  $x(n)$  is zero-mean, that is,  $m_x = E[x(n)] = 0$ , (54) equals:



$$b(n) = \frac{\mathbf{x}^H(n)\mathbf{x}(n-1)}{\sqrt{[\mathbf{x}^H(n)\mathbf{x}(n)][\mathbf{x}^H(n-1)\mathbf{x}(n-1)]}} \quad (55)$$

Applying Cauchy-Schwarz we know  $|b(n)| \leq 1$ . Obviously, from the definition of correlation coefficient we know  $b(n)\mathbf{x}(n-1)$  denotes the part correlated to  $\mathbf{x}(n-1)$  in  $\mathbf{x}(n)$ , so subtract the correlated part from  $\mathbf{x}(n)$  means decorrelation. Here define the new input vector:

$$\mathbf{u}(n) = \mathbf{x}(n) - b(n)\mathbf{x}(n-1) \quad (56)$$

$\mathbf{u}(n)$  denotes the part uncorrelated to  $\mathbf{x}(n-1)$  in  $\mathbf{x}(n)$ , it means  $\mathbf{u}(n)$  is independent to  $\mathbf{x}(n-1)$ , that is,  $\mathbf{x}^H(n-1)\mathbf{u}(n) = 0$ . To verify this conclusion, we must introduce the generalization of the mean ergodic theorem:

$x(n)$  is a discrete-time random process denoted by time-series  $x(n), x(n-1), \dots, x(n-N+1)$ , when  $x(n)$  is ergodic, the truth value of its self-correlation function  $r(k)$  is:

$$r(k) = \frac{1}{N} \mathbf{x}^H(n)\mathbf{x}(n-k) \quad (57)$$

In DSP, we always assume that signals are all mean ergodic stationary random signals. For the input vectors in this paper are all one-dimensional, so no matter the vectors are real or complex, we have:

$$\mathbf{x}^H(n)\mathbf{x}(n-1) = \mathbf{x}^H(n-1)\mathbf{x}(n) \quad (58)$$

From (57) and (58), we have the proving process as follows:

$$\begin{aligned} & \mathbf{x}^H(n-1)\mathbf{u}(n) \\ &= \mathbf{x}^H(n-1)\mathbf{x}(n) - \mathbf{x}^H(n-1)b(n)\mathbf{x}(n-1) \\ &= \mathbf{x}^H(n-1)\mathbf{x}(n) - \mathbf{x}^H(n-1) \times \\ & \quad \frac{\mathbf{x}^H(n)\mathbf{x}(n-1)}{\sqrt{[\mathbf{x}^H(n)\mathbf{x}(n)][\mathbf{x}^H(n-1)\mathbf{x}(n-1)]}} \mathbf{x}(n-1) \\ &= \mathbf{x}^H(n-1)\mathbf{x}(n) \left[ 1 - \sqrt{\frac{\mathbf{x}^H(n-1)\mathbf{x}(n-1)}{\mathbf{x}^H(n)\mathbf{x}(n)}} \right] \\ &= \mathbf{x}^H(n-1)\mathbf{x}(n) \left[ 1 - \sqrt{\frac{N \times r(0)}{N \times r(0)}} \right] = 0 \end{aligned} \quad (59)$$

So the new input vector  $\mathbf{u}(n)$  is orthogonal to  $\mathbf{x}(n-1)$ , decorrelate the correlation exist in original input signals. The convergence speed and the performance of the algorithm would be surely improved if we use this new input vector to update the weight vector of the adaptive filter. Next we shall verify that the algorithm could also converge to the optimal Wiener solution when we use  $\mathbf{u}(n)$  instead of  $\mathbf{x}(n)$  in Eq.(9). From (56) we have:

$$\mathbf{x}(n) = \mathbf{u}(n) + b(n)\mathbf{x}(n-1) \quad (60)$$

Then:

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) + m\mathbf{e}^*(n)\mathbf{u}(n) \\ &= \mathbf{w}(n) + m[d^*(n) - \mathbf{x}^H(n)\mathbf{w}(n)]\mathbf{u}(n) \\ &= \mathbf{w}(n) + m\mathbf{d}^*(n)\mathbf{u}(n) - m\mathbf{x}^H(n)\mathbf{w}(n)\mathbf{u}(n) \\ &= \mathbf{w}(n) + m\mathbf{d}^*(n)\mathbf{u}(n) - m[\mathbf{u}(n) + b(n)\mathbf{x}(n-1)]^H \mathbf{w}(n)\mathbf{u}(n) \\ &= \mathbf{w}(n) + m\mathbf{d}^*(n)\mathbf{u}(n) - m\mathbf{u}^H(n)\mathbf{w}(n)\mathbf{u}(n) \\ &= \mathbf{w}(n) + m\mathbf{u}(n)[d^*(n) - \mathbf{u}^H(n)\mathbf{w}(n)] \end{aligned} \quad (61)$$

Make Eq.(7) equals zero, we can get the optimal solution of original weight coefficient as  $\mathbf{w}_o(n) = \mathbf{R}^{-1}\mathbf{P}$ . Compare Eq.(61) and (9), we can also get the optimal Wiener solution  $\mathbf{w}_o(n) = \mathbf{R}^{-1}\mathbf{P}$  as long as we make  $\mathbf{R} = \mathbf{u}(n)\mathbf{u}^H(n)$  and  $\mathbf{P} = \mathbf{u}(n)d^*(n)$  in Eq.(8) and constitute them into (7) then make Eq.(7) equals zero.

## Simulation results in LabVIEW

This section presents three examples to corroborate the analytic results derived in this paper and to illustrate the properties of our algorithm. The algorithm was applied to system identification and it was

simulated in LabVIEW for real-time applications. In the examples, algorithms in paper<sup>[5]</sup> and paper<sup>[18]</sup> were compared to demonstrate that our algorithm can lead to better identification abilities no matter in weakly correlated, strongly correlated, stationary, nonstationary, Low SNR or abrupt system variation environments. Besides, the design principles of adjusting factors  $m$  and  $k$  are also given below.

To begin with, let us describe the experimental setup and parameter settings:  $N=10$ ,  $b(0)=0$ ,  $r=0.98$ ; Initial weight vector  $w(0)=[0,0,\dots,0]^H$ ; Input signal  $x(n)$  is zero-mean, white Gaussian with unit power, take  $x(n)$  as input signal directly and get the correlation coefficient from (53) when we need weakly correlated input signal, when we need strongly correlated input signal, take  $x(n)=lx(n-1)+\theta(n)$  as input signal where  $l$  denotes the correlation coefficient and  $\theta(n)$  is zero-mean, white Gaussian with unit power; The additive measurement noise  $v(n)$  is zero-mean, white Gaussian with  $s_v^2$  as its variance and statistically independent of  $x(n)$ ; The optimal weight vector is  $w_o=[0.1,0.2,0.3,0.4,0.5,-0.5,-0.4,-0.3,-0.2,-0.1]^H$  in stationary environment and  $w_o(n+1)=w_o(n)+\gamma(n)$  in nonstationary environment where  $\gamma(n)$  is zero-mean, white Gaussian with  $s_g^2$  as its variance. The results were obtained by averaging 200 ensemble trials with  $n=1000$  each trial. Our algorithm is called VV-LMS here.

**Example 1.** Design principle of adjusting factors  $m$  and  $k$

Confirming the design principle of adjusting factors  $m$  and  $k$  under stationary environment with weakly correlated input signal and 0.2 as its correlation coefficient,  $s_v^2=0.01$ . The choosing criterion of  $m$  and  $k$  is to minimize the MSE. Under the simulation environment above,  $\|F(n)\|$  is at a level of  $10^{-2}$ , so the step-size would decrease following the increase of  $m$ , but if  $m$  is too large, the step-size might equals zero even if the system have not steady, this would surely bigger the steady-state error as the two curves show in Fig.2 when  $m=5$  and 10. Too small  $m$  would interfere the performance of the system as well. These are consistent with our analytic results. Considering  $b(n)$  is at a level of  $10^{-3}$  under the environment and the maximum value of tanh function is 1, we take  $m=0.05$  to avoid step-size from getting too little.

The functions of different  $k$  to the system can be neglected under the environment above when  $m$  is fixed as Fig.3 shows. Here we take  $k=0.05$  to give consideration to both little and large  $a(n)$ .

**Example 2.** The performances of the algorithms in stationary environment

Fig.4 shows the learning curves of our algorithm and the two other algorithms with weakly correlated input signal and 0.2 as its correlation coefficient,  $s_v^2=0.01$ . It can be seen that our algorithm achieves lower steady-state MSE and faster convergence speed compared to the two other algorithms. And frequently the two other algorithms may diverge and can't converge to optimal weight vector.

Fig.5 shows the comparison of the identification abilities with strongly correlated input signal and 0.9 as its correlation coefficient,  $s_v^2=0.01$ . It can be easily seen that under this circumstance the two other algorithms may diverge even at the beginning of the system whereas our decorrelation algorithm can still converge to the optimal weight vector quickly, and its performance is no worse than weakly correlated case.

Fig.6 shows the learning curves of our algorithm and the two other algorithms with weakly correlated input signal and low SNR, that is,  $s_v^2=0.1$ . It can be easily seen that under this circumstance our algorithm can still converge to the optimal weight vector quickly and its steady-state MSE can still reach to the level of  $10^{-2}$  whereas the algorithm in paper[5] may diverge at the beginning of the system and sometimes the algorithm in paper[18] may also diverge. Keep lowering SNR, our algorithm still have incomparable convergence performance.

Fig.7 shows the learning curves of our algorithm and the two other algorithms with weakly correlated input signal and an abrupt system variation at iteration 500 such that the weight vector that has already identified turn back to zero again and the optimal weight vector changes to  $w_o=[-0.1,-0.2,-0.3,-0.4,-0.5,0.5,0.4,0.3,0.2,0.1]^H$ . It can be seen that our algorithm can still keep in good convergence performance and achieve a faster tracking behavior and lower misadjustment whereas the

algorithm in paper<sup>[5]</sup> is totally divergent and sometimes the algorithm in paper<sup>[18]</sup> may also diverge at iteration 500.

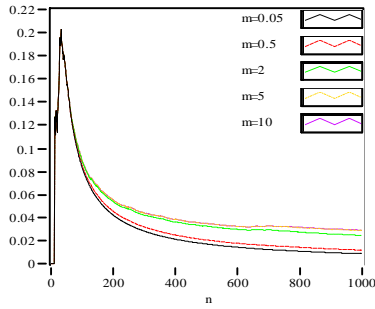


Fig 2 Comparison of MSE with different  $m$

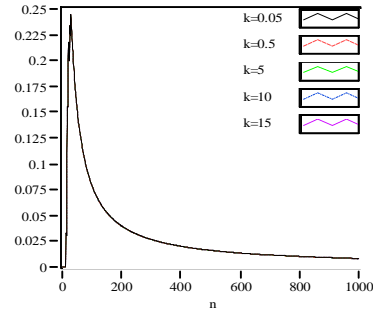


Fig 3 Comparison of MSE with different  $k$

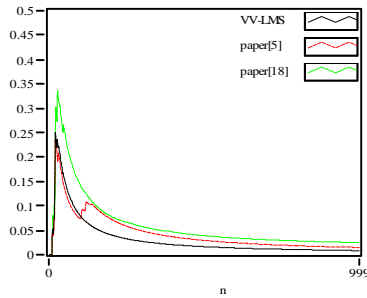


Fig 4 Comparison of MSE with weakly correlated input in stationary plant

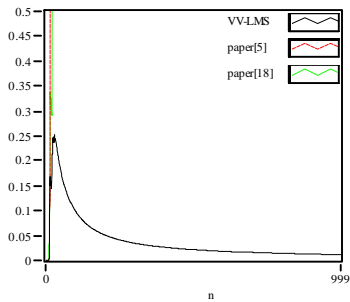


Fig 5 Comparison of MSE with highly correlated input in stationary plant

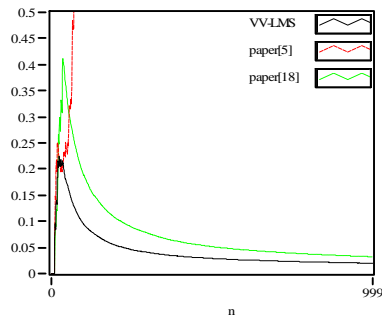


Fig 6 Comparison of MSE with weakly correlated input in stationary&low SNR plant

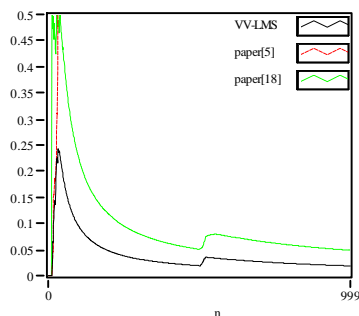


Fig 7 Comparison of MSE with weakly correlated input in stationary plant while system mutated  
**Example 3.**The performances of the algorithms in nonstationary environment

Construct the nonstationary environment by Eq.(16), the input signal is weakly correlated and  $s_v^2 = 0.01$ ,  $s_g^2 = 10^{-10}$ ,  $n = 10000$  here. From Fig.8 we can see that the MSE is little for  $s_g^2$  is tiny for all algorithms. The algorithm in paper[5] may still diverge under these conditions while our algorithm and the algorithm in paper[18] can adaptive converge with the variance of the environment. The performance of our algorithm is obviously better than paper[18].

## Conclusion

This work presented a new variable step-size and variable parameter adaptive system identification algorithm based on modified decorrelation principle. Unlike many existing approaches, we adjust the variable step-size by Hyperbolic tangent function and the mean value of gradient vector, and decorrelate the input signals before go on into the adaptive algorithm. This algorithm can overcome the defects of the existing algorithms that they can't deal with strongly correlated input signals and their performance would decline in nonstationary environment, it is very insensitive to measurement noise and has simple structure. The step-size would adaptively decrease following the convergence progress and can abruptly increase when there is an abrupt system variation. The analytic and simulation results verified that our algorithm has lower steady-state MSE and faster convergence speed no matter the input signals are correlated or noncorrelated, the system is in low SNR, stationary or nonstationary environment and whether there is an abrupt system variation, when compared to the existing representative approaches. The superior properties of our algorithm would be more evident with higher correlated signals. Meanwhile the realization of the algorithm in LabVIEW can satisfy the real-time applications.

## References

- [1] Gelfand S B, Wei Y, Krogmeie J Vr. The stability of variable step-size LMS algorithms[J].IEEE Trans. on Signal Processing(S1053-587X),1999,47(12):3277-3288.
- [2] Ye Hua, Wu Boxiu. The Study on a Variable Step Size Adaptive Filtering Algorithm[J].Acta Electronica Sinica,1990, 18 (4):63-69.
- [3] R.Kwong, E.W.Johnston. A variable step size LMS algorithm[J].IEEE Trans. Signal Process,1992,40 (7):1633-1642.
- [4] Qin Jingfan, Ouyang Jingzheng. A Novel Variable Step Size LMS Adaptive Filtering Algorithm Based on Sigmoid Function[J].Journal of data acquisition&processing,1997,12(3):171-174.
- [5] GAO Ying, XIE Sheng-li. A Variable Step Size LMS Adaptive Filtering Algorithm and Its Analysis[J].Acta Electronica Sinica,2001,29(8):1094-1097.
- [6] CHEN Kai, ZHANG Ping. A Novel Variable Step Size LMS Adaptive Filtering Algorithm[J].Modern Electronic Technique,2004,23:56,57,62.
- [7] LI Lijun, WANG Huakui. An Improved Variable Step Size LMS Adaptive Filtering Algorithm[J].Modern Electronic Technique,2006,7:7-9/12.
- [8] M.H.Costa, J.C.M.Bermudez. A noise resilient variable step-size LMS algorithm[J]. Signal Processing,2008,88: 733-748.
- [9] Shengkui Zhao, Zhihong Man, etc. Variable step-size LMS algorithm with a quotient form[J].Signal Processing, 2009,89:67-76.
- [10] Yonggang Zhang, Ning Li, etc. New Gradient- Based Variable Step Size LMS Algorithms[J].EUR- ASIP Journal on Advances in Signal Processing,2008,Article ID 529480:1-9.
- [11] Jeng-Kuang Hwang, etc. Variable Step-Size LMS Algorithm With a Gradient-Based Weighted Average[J]. IEEE signal processing letters,2009,16(12):1043-1046.
- [12] K. Mayyas, F. Momani. An LMS adaptive algorithm with a new step-size control equation[J].Journal of the Franklin Institute,2011,348:589-605.

- [13] ZHANG Zhong-hua, ZHANG Duan-jin. New variable step size LMS adaptive filtering algorithm and its performance analysis[J].Systems Engineering and Electronics,2009,31(9):2238-2241.
- [14] DOHERTYJF,PORAYATH R.A robust echo canceler for acoustic environments[J].IEEE Transactions on circuits and systems-2:Analog and digital signalprocessing,1997,44(5):389-369.
- [15] S. Haykin, Adaptive Filter Theory[M],fourth ed. Prentice-Hall, Englewood Cliffs,NJ,2002.
- [16] SHENG zhou, Probability Theory and Mathematical Statistics[M],third ed. Beijing: Higher Education Press, 2001 :130-131(in Chinese).
- [17] WANG Zhen-li, ZHANG Xiong-wei, YANG Ji-bin. Study of a new fast adaptive filtering algorithm[J].Journal of China Institute of Communications,2005,26(17):1-6.
- [18] SUN En-chang, LI Yu-heng, ZHANG Dong-ying. Adaptive Variable-step Size LMS Filtering Algorithm and Its Analysis[J].Journal of System Simulation,2007,19(14):3172-3175.