CrossMark

ORIGINAL PAPER

Optimal step size of least mean absolute third algorithm

Sihai Guan[1] · **Zhi Li**[1]

© Springer-Verlag London 2017

Abstract This paper proposes an optimized least mean absolute third (OPLMAT) algorithm to improve the capability of the adaptive filtering algorithm against Gaussian and non-Gaussian noises when the unknown system is a time-varying parameter system under low signal–noise rate. The optimal step size of the OPLMAT is obtained based on minimizing the mean-square deviation at the current time. In addition, the mean convergence and steady-state error of the OPLMAT algorithm are derived theoretically, and the computational complexity of OPLMAT is analyzed. Furthermore, the simulation experimental results of system identification presented illustrate the principle and efficiency of the OPLMAT algorithm. Simulation results demonstrate that the proposed algorithm performs much better than the LMAT and NLMAT algorithms.

Keywords LMAT · Optimal step size · Low signal–noise rate · Most of the noise densities · System identification

1 Introduction

Adaptive filter (AF) algorithms are frequently employed in equalization, active noise control, acoustic echo cancelation, biomedical engineering and many other fields [1]. The step size is a critical parameter in addressing the issue of obtaining either fast convergence rate or low excess mean-square error. Although a large step size responds quickly to plant changes,

it may lead to a large MSD and even cause loss of convergence, while a small step size may degrade the tracking speed of the AF algorithm. To address this issue, a relatively large step size can be selected during the initial convergence of an AF algorithm and then a small step size used as the algorithm approaches its steady state. Selection of the step size should balance low steady-state error with a fast convergence rate. So the variable step size method attracts considerable attentions from scholars. There are various methods that can be adopted to get the step size formula. Shin et al. [2] made use of a special function to update step size. Kwong and Johnston [3] proposed a method based on the input data to obtain the step size. Ang [4] presented an approach based on squared instantaneous estimation errors to update the step size. step size parameter was selected in [5] such that the sum of the squares of the measured estimation errors was minimized at the current instant in time. Liu et al. [6] and Benesty et al. [7] utilized the various nonparametric variance estimates and proposed the nonparametric variable step size NLMS algorithms. In [8], Huang and Lee presented a method based on the mean-square error and the estimated system noise power to obtain the step size. Many of these algorithms were developed in a system identification context, and the reference signal is considered as the output of a time-invariant system, usually corrupted by additive noise. However, in practical engineering applications, the system is often non-stationary. Moreover, it makes more sense to minimize the system misalignment, instead of the classical error-based cost function. The approximating optimal step size is selected in [9] such that the MSD was minimized. Lee and Park [10] made use of mean-square deviation to update the step size of the APA algorithm. Zhi et al. [11] proposed the optimal step size of the PAP algorithm. Ciochină et al. [12] proposed an optimized NLMS algorithm which takes advantage of the joint optimization problem with both the

✉ Zhi Li
15107739422@163.com

Sihai Guan
gcihey@sina.cn

[1] School of Electro-Mechanical Engineering Xidian University, No. 2 South Taibai Road, Xi'an 710071, Shaanxi, China

normalized step size and regularization parameters. More recently, research has focused on AF algorithms based on high-order error power (HOEP) conditions [13–15]. Among HOEP algorithms, the LMAT algorithm can achieve the most robust against the unknown noise of several different probability densities [16]. For example, in [17,18], impulsive noise is modeled by an α-stable random process. Besides, some algorithms for impulsive noise environments are introduced. The LMAT algorithm is based on the minimization of the mean of the absolute error value of the third power. The error function is a perfect convex function with respect to filter coefficients, so there is no local minimum for the LMAT algorithm. The feature of unknown system, the characteristics of the additive noise, SNR and the input excitation govern the effectiveness of the LMAT-type algorithm. However, in practical engineering applications, the measurement noise of an unknown system is non-Gaussian, and the system is often non-stationary with low SNR. So on the basis of work presented in [9–12,16], the optimal step size rather than approximating optimal step size is selected in this paper such that the MSD at the current in time is minimum. The purpose of the OPLMAT algorithm is to deal with the Gaussian, uniform, Rayleigh and exponential noise distributions, which is different from the literature [17,18]. The mean convergence and MSD of the OPLMAT algorithm are also derived. The computational complexity of the OPLMAT algorithm is analyzed theoretically. Finally, we have carried out four system identification simulation experiments to illustrate significant superiorities of the OPLMAT algorithm over the LMAT and NLMAT algorithms.

The main contributions of this work are as follows: (a) an OPLMAT algorithm based on minimizing the MSD at the current in time; (b) the step size of the OPLMAT algorithm can achieve minimum steady-state error; (c) the stability of the proposed algorithm is studied; (d) the steady-state errors are derived from both Gaussian and non-Gaussian noises; and (e) the obtained results may also have some potential value in practical applications.

Briefly, this paper is organized as follows. The proposed OPLMAT algorithm is introduced in Sect. 2. The performance of the OPLMAT algorithm is studied in Sect. 3. The numerical simulations are carried out in Sect. 4, and conclusions are presented in Sect. 5.

2 Proposed OPLMAT algorithm

The coefficient vector of the unknown system is defined as $\mathbf{W}_{\text{opt}} = [w_0, w_1, \ldots, w_{L-1}]^{\text{T}}$. L is the filter length. $\mathbf{X}(n) = [x(n), x(n+1), \ldots, x(n+L-1)]^{\text{T}}$ denotes the input data vector of the unknown system at times instant n, and $d(n)$ denotes the observed output signals, respectively.

$$d(n) = \mathbf{W}_{\text{opt}}^{\text{T}}\mathbf{X}(n) + \xi(n) \tag{1}$$

where $\xi(n)$ is a stationary additive noise with zero mean and variance of σ_{ξ}^2. In addition, $\xi(n)$ is assumed to be uncorrelated with any other signal. $\mathbf{X}(n)$ is also stationary with zero mean, a variance of σ_x^2, and $\mathbf{X}(n)$ is Gaussian with a definite positive autocorrelation matrix.

The cost function used for obtaining the OPLMAT algorithm is given by

$$\begin{aligned} \mathbf{W}(n) &= \arg\ \min\ J\left(\mathbf{W}(n)\right) \\ &= \arg\ \min\ \frac{1}{3}\mathrm{E}\left[|e(n)|\right]^3 \end{aligned} \tag{2}$$

where

$$e(n) = d(n) - y(n) \tag{3}$$

The corresponding filter output is

$$y(n) = \mathbf{W}^{\text{T}}(n)\mathbf{X}(n) \tag{4}$$

Assuming

$$\mathbf{V}(n) = \mathbf{W}_{\text{opt}} - \mathbf{W}(n) \tag{5}$$

So

$$e(n) = \mathbf{V}^{\text{T}}(n)\mathbf{X}(n) + \xi(n) \tag{6}$$

The updated recursion of coefficients vectors can be derived as Eq. (7):

$$\begin{aligned} \mathbf{W}(n+1) &= \mathbf{W}(n) - \mu(n)\frac{\partial J\left(\mathbf{W}(n)\right)}{\partial \mathbf{W}(n)} \\ &= \mathbf{W}(n) + \mu(n)\mathbf{X}(n)e^2(n)\mathrm{sgn}\left(e(n)\right) \end{aligned} \tag{7}$$

In Eq. (7), $\mathrm{sgn}\left(e(n)\right)$ denotes the sign function of the variable $e(n)$ [13].

Combine Eqs. (5) and (7),

$$\mathbf{V}(n+1) = \mathbf{V}(n) - \mu(n)\mathbf{X}(n)e^2(n)\mathrm{sgn}\left(e(n)\right) \tag{8}$$

The following assumptions are used in the subsequent analysis [15,19].

Assumption A1 $\mathbf{W}(n)$ is independent of $\mathbf{X}(n)$.

Assumption A2 the error sequence $e(n)$ conditioned on the weight vector $\mathbf{V}(n)$ is a zero mean and Gaussian.

Based on *Assumption* A2, the distribution function of $Y = e^2(n)$ is shown as Eq. (9).

$$F(y) = p\left(e^2(n) \le y\right)$$
$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}\sigma_{e|V(n)}} \exp\left(-\frac{e^2(n)}{2\sigma_{e|V(n)}^2}\right) d\,e(n)$$
$$= erf\left(\frac{\sqrt{y}}{\sqrt{2}\sigma_{e|V(n)}}\right) \tag{9}$$

Based on Eq. (9), the probability density function of $Y = e^2(n)$ is shown as Eq. (10).

$$p(y) = \frac{\partial F(y)}{\partial y}$$
$$= \frac{1}{\sqrt{2\pi}\sigma_{e|V(n)}e(n)} \exp\left(-\frac{e^2(n)}{2\sigma_{e|V(n)}^2}\right) \tag{10}$$

So

$$E[V(n+1)]$$
$$= E[V(n)] - \mu(n)E\left[X(n)e^2(n)\text{sgn}(e(n))\right]$$
$$= E[V(n)] - \mu(n)E\left\{E\left[X(n)e^2(n)\text{sgn}(e(n))|V(n)\right]\right\}$$
$$= E[V(n)] - \mu(n)E\left\{E\left[e^2(n)|V(n)\right]\right.$$
$$\left. \times E\left[X(n)\text{sgn}(e(n))|V(n)\right]\right\}$$
$$= E[V(n)] - 2\sqrt{\frac{2}{\pi}}\mu(n)\sigma_e R_{xx} E[V(n)] \tag{11}$$

where $R_{xx} = E\left[X(n)X^T(n)\right]$ and $R_{xx} = \sigma_x^2 I$. I denotes the identity matrix of proper dimension.

Ultimately, it is easy to show that the mean behavior of the weight vector converge to the optimal weight vector W_{opt} if $\mu(n)$ is bounded by:

$$0 < \mu(n) < \sqrt{\frac{\pi}{2}}\frac{1}{\lambda_{\max}\sigma_e} \tag{12}$$

where $\lambda_{\max}$ represents the maximum eigenvalue of the regressor covariance matrix R_{xx}.

As seen from Eq. (11), we will get $E[V(\infty)]$ when $n \to \infty$.

$$E[V(\infty)] = 0 \tag{13}$$

Now we will derive the optimal $\mu(n)$.

$$V^T(n+1)V(n+1)$$
$$= \left[V(n) - \mu(n)X(n)e^2(n)\text{sgn}(e(n))\right]^T$$
$$\times \left[V(n) - \mu(n)X(n)e^2(n)\text{sgn}(e(n))\right]$$
$$= V^T(n)V(n) - \mu(n)V^T(n)X(n)e^2(n)\text{sgn}(e(n))$$
$$- \mu(n)\text{sgn}(e(n))e^2(n)X^T(n)V(n)$$
$$+ \mu^2(n)X^T(n)X(n)e^4(n) \tag{14}$$

Assuming tr(R_{xx}) is the trace of the R_{xx} and when $L \gg 1$, then tr$(R_{xx}) = L\sigma_x^2$ [7].

Taking the statistical expectation of both sides with (14), Eq. (15) can be obtained.

$$E\left[V^T(n+1)V(n+1)\right]$$
$$= E\left[V^T(n)V(n)\right] + \mu^2(n)L\sigma_x^2 E\left[e^4(n)\right]$$
$$- 2\mu(n)E\left[V^T(n)X(n)e^2(n)\text{sgn}(e(n))\right] \tag{15}$$

Based on Eqs. (9), (10) and Eqs. (11), (16) can be obtained.

$$E\left[V^T(n)X(n)e^2(n)\text{sgn}(e(n))\right]$$
$$= E\left\{E\left[V^T(n)X(n)e^2(n)\text{sgn}(e(n))|V(n)\right]\right\}$$
$$\approx E\left\{2\sigma_e^2\sqrt{\frac{2}{\pi}}\frac{1}{\sigma_e}E\left[\left(V^T(n)X(n)X^T(n)V(n)\right.\right.\right.$$
$$\left.\left.\left. + V^T(n)X(n)\xi(n)\right)|V(n)\right]\right\}$$
$$= 2\sqrt{\frac{2}{\pi}}\sigma_e\sigma_x^2 E\left[V^T(n)V(n)\right] \tag{16}$$

Based on Eq. (6),

$$E\left[e^4(n)\right]$$
$$= E\left\{\left[V^T(n)X(n) + \xi(n)\right]^4\right\}$$
$$= E\left\{\left[V^T(n)X(n)\right]^4\right\} + 4E\left\{\left[V^T(n)X(n)\right]^3\xi(n)\right\}$$
$$+ 6E\left\{\left[V^T(n)X(n)\right]^2[\xi(n)]^2\right\}$$
$$+ 4E\left\{\left[V^T(n)X(n)\right][\xi(n)]^3\right\} + E\left\{[\xi(n)]^4\right\} \tag{17}$$

Combining *Assumption* A1–*Assumption* A2 with Eq. (17), we get Eq. (18).

$$
\mathrm{E}\left[e^4(n)\right] = \mathrm{E}\left\{\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{X}(n)\right]^4\right\}
$$
$$
+ 6\mathrm{E}\left\{\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{X}(n)\right]^2\right\}\sigma_\xi^2 + \mathrm{E}\left\{[\xi(n)]^4\right\} \tag{18}
$$

Substituting Eqs. (16) and (18) into Eq. (15), we get Eq. (19).

$$
\mathrm{E}\left[\mathbf{V}^{\mathrm{T}}(n+1)\mathbf{V}(n+1)\right]
$$
$$
= \mathrm{E}\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{V}(n)\right] - 4\sqrt{\frac{2}{\pi}}\mu(n)\sigma_e\sigma_x^2\,\mathrm{E}\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{V}(n)\right]
$$
$$
+ \mu^2(n)L\sigma_x^2\,\mathrm{E}\left\{\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{X}(n)\right]^4\right\}
$$
$$
+ \mu^2(n)L\sigma_x^2\,\mathrm{E}\left\{[\xi(n)]^4\right\}
$$
$$
+ 6\mu^2(n)L\sigma_x^2\,\mathrm{E}\left\{\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{X}(n)\right]^2\right\}\sigma_\xi^2 \tag{19}
$$

Based on Eq. (8) in [16],

$$
\begin{cases}
\mathrm{E}\left\{\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{X}(n)\right]^4\right\} = 3\sigma_x^4\,\mathrm{E}^2\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{V}(n)\right] \\
\mathrm{E}\left\{\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{X}(n)\right]^2\right\} = \sigma_x^2\,\mathrm{E}\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{V}(n)\right]
\end{cases} \tag{20}
$$

Substituting Eq. (20) into Eq. (19), we obtain Eq. (21).

$$
\mathrm{E}\left[\mathbf{V}^{\mathrm{T}}(n+1)\mathbf{V}(n+1)\right]
$$
$$
= \mathrm{E}\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{V}(n)\right] - 4\sqrt{\frac{2}{\pi}}\mu(n)\sigma_e\sigma_x^2\,\mathrm{E}\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{V}(n)\right]
$$
$$
+ 6\mu^2(n)L\sigma_x^4\,\mathrm{E}\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{V}(n)\right]\sigma_\xi^2
$$
$$
+ 3\mu^2(n)L\sigma_x^6\,\mathrm{E}^2\left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{V}(n)\right]
$$
$$
+ \mu^2(n)L\sigma_x^2\,\mathrm{E}\left\{[\xi(n)]^4\right\} \tag{21}
$$

The MSD at times instant n is defined as $\mathrm{MSD}(n) = \mathrm{E}\left[\|\mathbf{V}(n)\|^2\right]$. So Eq. (21) can be derived as Eq. (22).

$$
\mathrm{MSD}(n+1)
$$
$$
= \left[1 - 4\sqrt{\frac{2}{\pi}}\mu(n)\sigma_e\sigma_x^2 + 6\mu^2(n)L\sigma_x^4\sigma_\xi^2\right]\mathrm{MSD}(n)
$$
$$
+ \mu^2(n)L\sigma_x^2\,\mathrm{E}\left[\xi^4(n)\right] + 3\mu^2(n)L\sigma_x^6\,\mathrm{MSD}^2(n) \tag{22}
$$

When the algorithm tends to be stable, MSD(n) is very small. So, Eq. (22) can be rewritten as Eq. (23) when discarding $\mathrm{MSD}^2(n)$.

$$
\mathrm{MSD}(n+1)
$$
$$
= f\left(L, \mu(n), \sigma_x^2, \sigma_e, \sigma_\xi^2\right)\mathrm{MSD}(n)
$$
$$
+ g\left(L, \mu(n), \sigma_x^2, \xi(n)\right) \tag{23}
$$

where

$$
f\left(\alpha(n), L, \mu(n), \sigma_x^2, \sigma_e, \sigma_\xi^2\right)
$$
$$
= 1 - 4\sqrt{\frac{2}{\pi}}\mu(n)\sigma_e\sigma_x^2 + 6\mu^2(n)L\sigma_x^4\sigma_\xi^2 \tag{24}
$$
$$
g\left(L, \mu(n), \sigma_x^2, \xi(n)\right) = L\sigma_x^2\mu^2(n)\mathrm{E}\left[\xi^4(n)\right] \tag{25}
$$

In the real practical world, $\sigma_e^2(n)$ and $\sigma_x^2(n)$ are unknown; however, we can estimate $\sigma_e^2(n)$ and $\sigma_x^2(n)$ by using Eq. (26) [8].

$$
\begin{cases}
\sigma_e^2(n) = \frac{1}{\rho + \sigma_x^2(n)}\mathbf{p}^{\mathrm{T}}(n)\mathbf{p}(n) \\
\sigma_x^2(n) = \chi\sigma_x^2(n-1) + (1-\chi)\mathbf{X}^{\mathrm{T}}(n)\mathbf{X}(n)
\end{cases} \tag{26}
$$

where $\mathbf{p}(n) = \mathrm{E}\left[e(n)\mathbf{X}(n)\right]$, ρ is a small positive number to guarantee that the denominator of Eq. (26) remains finite when $\sigma_x^2(n) = 0$.

Although in Eq. (26), $\mathbf{p}(n)$ is also unknown, we can estimate this parameter by using Eq. (27).

$$
\mathbf{p}(n) = \chi\mathbf{p}(n-1) + (1-\chi)e(n)\mathbf{X}(n) \tag{27}
$$

Based on NPVSS-NLMS algorithm [7], we can get $(1 - \frac{1}{2L}) \leq \chi < 1$.

The result from Eq. (23) illustrates a "separation" between the convergence and misadjustment components. Therefore, the term $f\left(L, \mu(n), \sigma_x^2, \sigma_e, \sigma_\xi^2\right)$ influences the convergence rate of the algorithm. It can be noticed that the fastest convergence mode is obtained when the function of Eq. (24) reaches its minimum.

$$
\mu(n) = \sqrt{\frac{2}{\pi}}\frac{\sigma_e}{3L\sigma_x^2\sigma_\xi^2} \tag{28}
$$

The stability condition can be found by imposing $\left|f\left(L, \mu(n), \sigma_x^2, \sigma_e, \sigma_\xi^2\right)\right| < 1$ which leads to Eq. (29).

$$
0 < \mu(n) < \sqrt{\frac{2}{\pi}}\frac{2\sigma_e}{3L\sigma_x^2\sigma_\xi^2} \tag{29}
$$

There are two important issues to be considered: (1) in the context of system identification, it is reasonable to follow a minimization problem in terms of the system misalignment and (2) we have main parameters $\mu(n)$ which influences the overall performance of the OPLMAT algorithm.

Thus, based on Eq. (23),

$$\frac{\partial \mathrm{MSD}(n)}{\partial \mu(n)} = 0 \tag{30}$$

Substituting Eqs. (23), (24) and (25) into Eq. (30), the optimal step size is then given by

$$\mu(n) = \frac{2}{L} \sqrt{\frac{2}{\pi}} \frac{\sigma_e \, \mathrm{MSD}(n)}{6\sigma_x^2 \sigma_\xi^2 \, \mathrm{MSD}(n) + \mathrm{E}\left[\xi^4(n)\right]} \tag{31}$$

Using Eq. (31) in Eq. (23), followed by several straight forward computations, it results in

$$
\begin{aligned}
\mathrm{MSD}(n+1) = {} & \mathrm{MSD}(n) \\
& - \frac{16}{L\pi} \frac{\sigma_e^2 \sigma_x^2 \, \mathrm{MSD}^2(n)}{6\sigma_x^2 \sigma_\xi^2 \, \mathrm{MSD}(n) + \mathrm{E}\left[\xi^4(n)\right]} \\
& + \frac{8}{L\pi} \frac{\sigma_e^2 \sigma_x^2 \, \mathrm{MSD}^2(n) \, \mathrm{E}\left[\xi^4(n)\right]}{\left\{6\sigma_x^2 \sigma_\xi^2 \, \mathrm{MSD}(n) + \mathrm{E}\left[\xi^4(n)\right]\right\}^2} \\
& + \frac{48}{L\pi} \frac{\sigma_e^2 \sigma_x^4 \sigma_\xi^2 \, \mathrm{MSD}^3(n)}{\left\{6\sigma_x^2 \sigma_\xi^2 \, \mathrm{MSD}(n) + \mathrm{E}\left[\xi^4(n)\right]\right\}^2}
\end{aligned}
\tag{32}
$$

Denoting $\lim_{n\to\infty} \mathrm{MSD}(n) = \mathrm{MSD}(\infty)$ and developing Eq. (32), we obtain Eq. (33).

$$\sigma_e^2 \sigma_x^2 \left\{ 6\sigma_x^2 \sigma_\xi^2 \, \mathrm{MSD}(\infty) + \mathrm{E}\left[\xi^4(\infty)\right] \right\} \mathrm{MSD}^2(\infty) = 0 \tag{33}$$

Based on Eq. (33), we know $\mathrm{MSD}(\infty) = 0$. It means that our aforementioned step size adjusting mechanism can achieve zero steady-state mean-square deviation. For a Gaussian desired signal, $\mathrm{E}\left[\xi^4(n)\right] = 3\sigma_\xi^4$ [15]. For a uniform desired signal, $\mathrm{E}\left[\xi^4(n)\right] = 9\sigma_\xi^4/5$ [15]. For a Rayleigh desired signal, $\mathrm{E}\left[\xi^4(n)\right] = 8\sigma_\xi^4$ [20]. For an exponential desired signal, $\mathrm{E}\left[\xi^4(n)\right] = 3\sigma_\xi^4$ [20].

The excess mean-square error (EMSE) at times instant n is given by Eq. (34).

$$
\begin{aligned}
\mathrm{EMSE}(n) &= \mathrm{E}\left[e^2(n)\right] \\
&= \mathrm{E}\left\{ \left[\mathbf{V}^{\mathrm{T}}(n)\mathbf{X}(n) + \xi(n)\right] \right. \\
&\quad \left. \times \left[\mathbf{X}^{\mathrm{T}}(n)\mathbf{V}(n) + \xi(n)\right] \right\} \\
&= \sigma_\xi^2 + \sigma_x^2 \, \mathrm{MSD}(n)
\end{aligned}
\tag{34}
$$

Table 1 OPLMAT algorithm summary

Initialization: $\mathbf{W}(0) = \mathbf{0}$, ρ, χ
Desire output: $d(n) = \mathbf{X}^{\mathrm{T}}(n)\mathbf{W}_{\mathrm{opt}} + \xi(n)$
Filter output: $y(n) = \mathbf{X}^{\mathrm{T}}(n)\mathbf{W}(n)$
Error: $e(n) = d(n) - y(n)$
Update:
$\mathbf{p}(n) = \chi \mathbf{p}(n-1) + (1-\chi)e(n)\mathbf{X}(n)$
$\sigma_x^2(n) = \chi \sigma_x^2(n-1) + (1-\chi)\mathbf{X}^{\mathrm{T}}(n)\mathbf{X}(n)$
$\sigma_e^2(n) = \frac{1}{\rho + \sigma_x^2(n)} \mathbf{p}^{\mathrm{T}}(n)\mathbf{p}(n)$
$\mu(n) = \frac{2}{L}\sqrt{\frac{2}{\pi}} \frac{\sigma_e \, \mathrm{MSD}(n)}{6\sigma_x^2 \sigma_\xi^2 \, \mathrm{MSD}(n) + \mathrm{E}[\xi^4(n)]}$
$f\left(L, \mu(n), \sigma_x^2, \sigma_e, \sigma_\xi^2\right) = 1 - 4\sqrt{\frac{2}{\pi}}\mu(n)\sigma_e \sigma_x^2 + 6L\sigma_x^4 \sigma_\xi^2 \mu^2(n)$
$g\left(L, \mu(n), \sigma_x^2, \xi(n)\right) = L\sigma_x^2 \mu^2(n)\mathrm{E}\left[\xi^4(n)\right]$
$\mathrm{MSD}(n+1) = f\left(L, \mu(n), \sigma_x^2, \sigma_e, \sigma_\xi^2\right) \mathrm{MSD}(n) +$
$\quad g\left(L, \mu(n), \sigma_x^2, \xi(n)\right)$
$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu(n)\mathbf{X}(n)e^2(n)\mathrm{sgn}\left(e(n)\right)$

Finally, the steady-state $\mathrm{EMSE}_{\mathrm{EMSE}(\infty)}$ is given by using Eqs. (33) and (34).

$$\mathrm{EMSE}(\infty) = \sigma_\xi^2 \tag{35}$$

A summary of the procedure for the OPLMAT algorithm based on the analysis presented above is given in Table 1.

3 Computational complexity

The updated step size and σ_e formulas are added to the OPLMAT algorithm compared to the LMAT algorithm, meaning that the computational complexity of OPLMAT algorithm is greater than that of the LMAT algorithm. Furthermore, unlike the NLMAT algorithm, there is also no recursion required to compute $\mathbf{X}^{\mathrm{T}}(n)\mathbf{X}(n)$ in OPLMAT algorithm. The computational complexity of the OPLMAT algorithm is slightly greater than that of the LMAT algorithm. However, the computational complexity of the OPLMAT algorithm is less than the complexity of the NLMAT algorithm when L is large. From Table 2, the OPLMAT algorithm has a considerable complexity advantage over NLMAT when $L > 4$. Besides, the NLMAT algorithm also needs to estimate σ_e. So compared to the NLMAT algorithm, there is no recursion to compute $\mathbf{X}^{\mathrm{T}}(n)\mathbf{X}(n)$ and there are no comparisons. For convenience, the computational complexity of the OPLMAT algorithm and that of other existing LMAT-type algorithms are listed in Table 2.

4 Simulation results

This section presents the results of simulations in the context of system identification using various noise distributions

Table 2 Computational complexity of OPLMAT, LMAT and NLMAT algorithms

Algorithm	$\times$	$\div$	$\sqrt{}$	Comparisons	$+$
LMAT	$2L+2$	1	1	1	$2L+1$
	$4L+1$	2	1	$L\ln L+2$	$4L+2$
NLMAT	$3L+4$	1	1	2	$3L+3$
OPLMAT	$2L+8$	1	1	0	$2L+3$

Where $\times$ multiplications, $\div$ divisions, $+$ additions, $\sqrt{}$ square root

Table 3 MSD values

	MSD/dB		
	LMAT	NLMAT	OPLMAT
Experiment 1	-14.95	-18.65	-29.27
Experiment 2	-12.38	-16.43	-25.54
Experiment 3	-4.62	-9.09	-21.35
Experiment 4	-9.25	-12.99	-20.70

Fig. 1 MSD comparisons under this condition of Experiment 1

Fig. 2 Learning curves of MSD error under this condition of Experiment 1

Fig. 3 Evolution of $\mu(n)$ as a function of iteration number under the condition of Experiment 1

of both stationary and non-stationary systems to illustrate the accuracy of the OPLMAT algorithm. The length of the unknown coefficient vector $\mathbf{W}_{opt}$ is L. The input signal $x(n)$ is a Gaussian white noise with zero mean and $\sigma_x^2 = 1$. The correlated input signal $y(n)$ is calculated by using $y(n) = 0.5y(n-1) + x(n)$. In all of our experiments, the coefficient vectors are initialized as zero vectors. Four different noise distributions (Gaussian, uniform, Rayleigh and exponential) are used in the experiments. $\text{MSD}(n) = 10\log_{10}\left(\|\mathbf{W}(n) - \mathbf{W}_{opt}\|_2^2\right)$ is used to measure the performance of algorithms. In addition, MSD error equals the absolute of simulation values and theory value (Eq. 23). The results are obtained via Monte Carlo simulation using 20 independent run sets and an iteration number of 5000. The values of steady-state error of the three algorithms are recognized in Table 3.

Experiment 1

The system noise is Gaussian white noise, and the input signal is $x(n)$ with SNR $= 3$ dB. A time-varying system is modeled, and its coefficients are varied from a random walk process defined by $\mathbf{W}_{opt}(n) = \mathbf{W}_{opt} + \boldsymbol{v}(n)$, where $\boldsymbol{v}(n)$ is an independent identically distributed (i.i.d.) Gaussian sequence with $m_v = 0$ and $\sigma_v^2 = 0.01$. $\mathbf{W}_{opt} = [0.8, 0.2, 0.7, 0.2, 0.1]^T$. The MSD curves for the LMAT ($\mu = 0.02$), NLMAT ($\mu = 0.02$), and OPLMAT ($\chi = 0.98$) algorithms with the uncorrelated input signal are shown in Fig. 1. Figure 2 depicts the MSD error curves for the OPLMAT algorithm. In order to make case of the OPLMAT algorithm more persuasive, we provide a plot elucidating the evolution of $\mu(n)$ as a function of the number of iterations in Fig. 3.
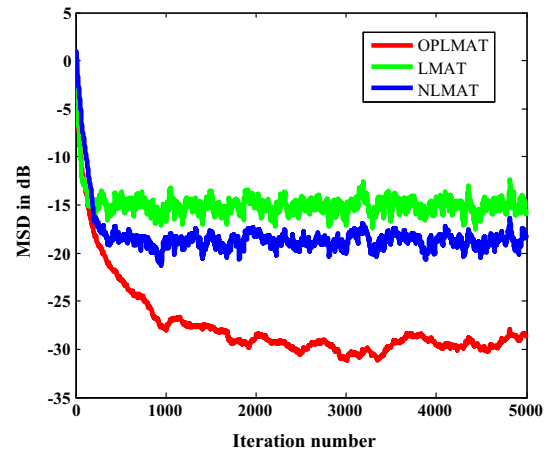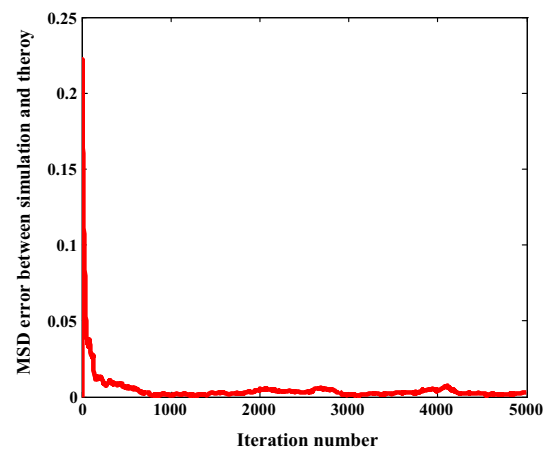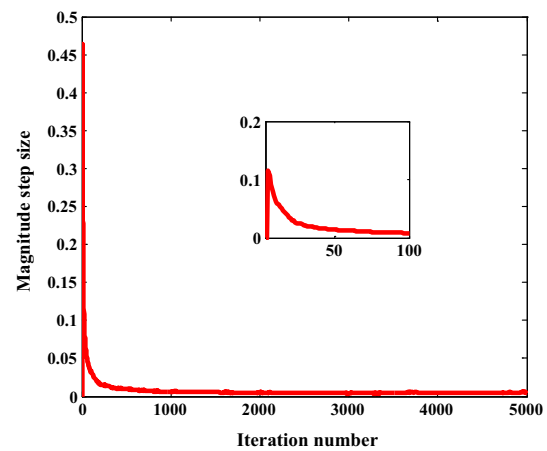
Experiment 2

The system noise is a uniformly distributed noise over the interval $(-3, 3)$ and the input signal is correlated input sig-
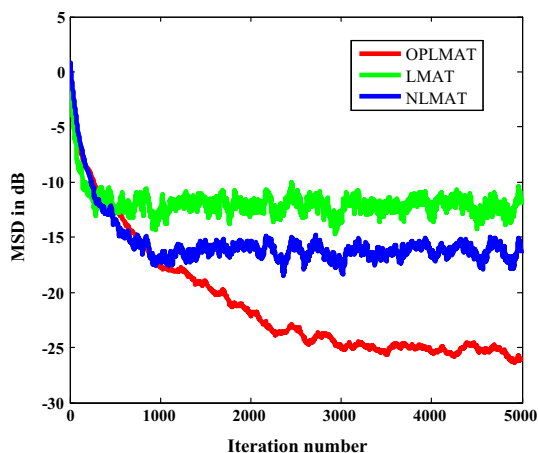
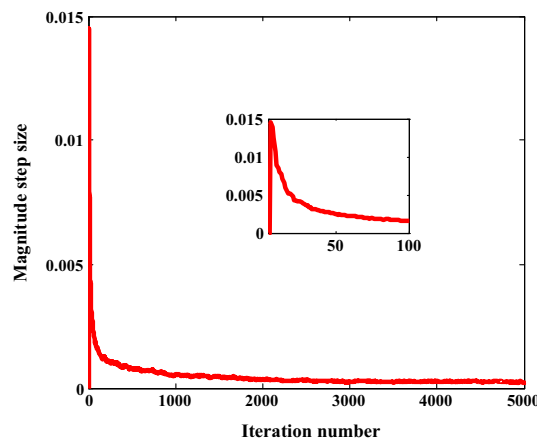**Fig. 4** MSD comparisons under this condition of Experiment 2

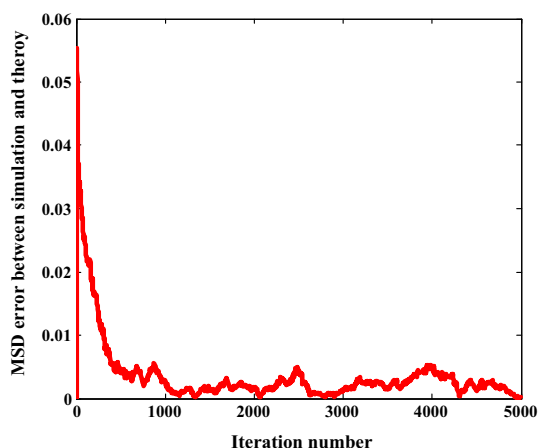**Fig. 6** Evolution of $\mu(n)$ as a function of iteration number under the condition of Experiment 2

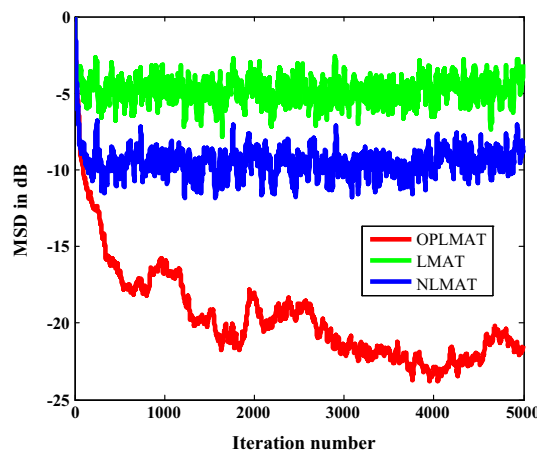**Fig. 5** Learning curves of MSD error under this condition of Experiment 2

Fig. 7 MSD comparisons under this condition of Experiment 3

nal $y(n)$ with SNR $= -5$ dB. A time-unvarying system is modeled, and its coefficients vector of the unknown system is $\mathbf{W}_{opt} = [0.8, 0.2, 0.7, 0.2, 0.1]^T$. The MSD curves for the LMAT ($\mu = 0.05$), NLMAT ($\mu = 0.05$), and OPLMAT ($\chi = 0.98$) algorithms with the uncorrelated input signal are shown in Fig. 4. Figure 5 shows the MSD error curves for the OPLMAT algorithm. In order to make case of the OPLMAT algorithm more persuasive, we provide a plot describing the evolution of $\mu(n)$ as a function of the number of iterations in Fig. 6.

Experiment 3

The system noise is a Rayleigh distribution with 3, and the input signal is correlated input signal $y(n)$ with SNR $= 14$ dB. A time-varying system is modeled, and its coefficients are varied from a random walk process defined by $\mathbf{W}_{opt}(n) = \mathbf{W}_{opt} + \boldsymbol{v}(n)$, where $\boldsymbol{v}(n)$ is an i.i.d. Gaussian sequence with $m_v = 0$ and $\sigma_v^2 = 0.01$. $\mathbf{W}_{opt} = [0.8, 0.2, 0.7, 0.2, 0.1]^T$. The MSD curves for the LMAT ($\mu = 0.01$), NLMAT

($\mu = 0.01$) and OPLMAT ($\chi = 0.98$) algorithms with the uncorrelated input signal are shown in Fig. 7. Figure 8 plots the MSD error curves for the OPLMAT algorithm. In order to make case of the OPLMAT algorithm more persuasive, we provide a plot showing the evolution of $\mu(n)$ as a function of the number of iterations in Fig. 9.

Experiment 4

The system noise is an exponential distribution with 2, and the input signal is correlated input signal $y(n)$ with SNR $= 14$ dB. A time-varying system is modeled, and its coefficients are varied from a random walk process defined by $\mathbf{W}_{opt}(n) = \mathbf{W}_{opt} + \boldsymbol{v}(n)$, where $\boldsymbol{v}(n)$ is an i.i.d. Gaussian sequence with $m_v = 0$ and $\sigma_v^2 = 0.01$. $\mathbf{W}_{opt} = [0.8, 0.2, 0.7, 0.2, 0.1]^T$. The MSD curves for the LMAT ($\mu = 0.005$), NLMAT ($\mu = 0.005$) and OPLMAT ($\chi = 0.98$) algorithms with the uncorrelated input signal are shown in Fig. 10. Figure 11 shows the MSD error curves for the OPLMAT algorithm. In order to make case of the OPLMAT algorithm more persua-
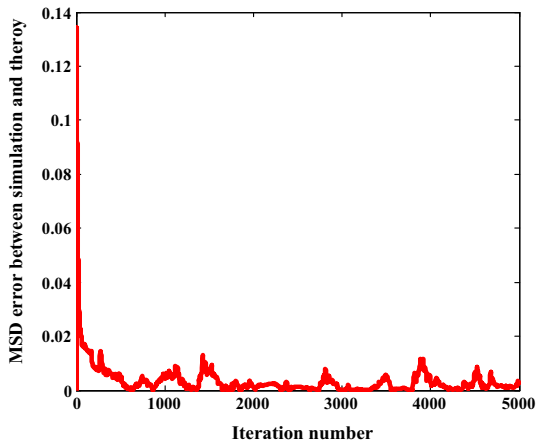
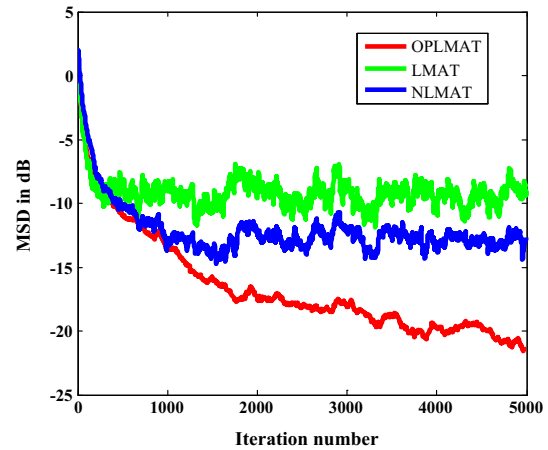**Fig. 8** Learning curves of MSD error under this condition of Experiment 3

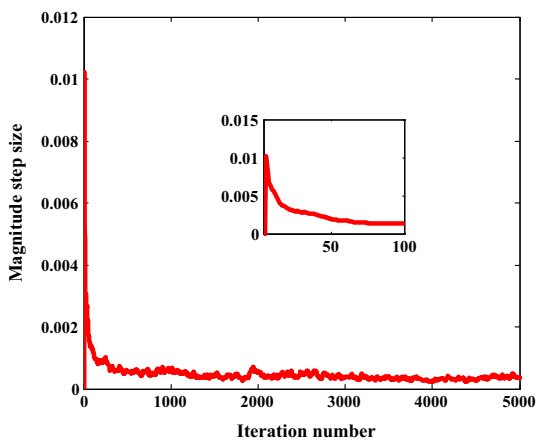**Fig. 10** MSD comparisons under this condition of Experiment 4

Fig. 9 Evolution of $\mu(n)$ as a function of iteration number under the condition of Experiment 3
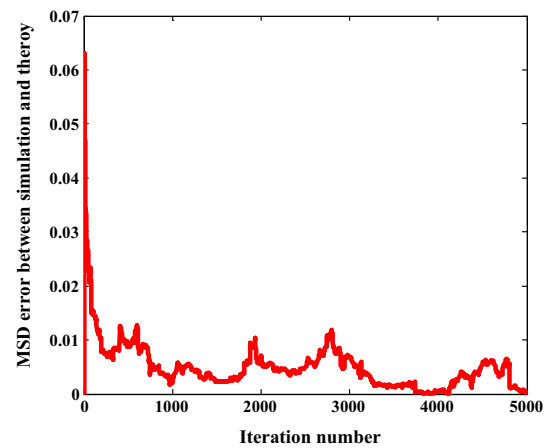

Fig. 11 Learning curves of MSD error under this condition of Experiment 4

sive, we provide a plot showing the evolution of $\mu(n)$ as a function of the number of iterations in Fig. 12.

Figures 1, 4, 7 and 10 show the OPLMAT algorithm has a smaller misalignment than the LMAT and NLMAT algorithms in steady-state stage. The reason for this observation is small $\mu(n)$ in steady-state stage. From Fig. 1 and Table 3, the improvement due to implementation of the LMAT and NLMAT algorithms can nearly approach 14.32 and 10.62 dB, respectively. Thus, compared to the LMAT and NLMAT algorithms, the OPLMAT algorithm can perform better in identifying the unknown coefficients under this condition. From the other experimental results, the same conclusion can be obtained.

Figures 2, 5, 8 and 11 show the MSD error. We observe an excellent match between predictions provided by our newly designed algorithm and results given by Monte Carlo simulations.

From Figs. 3, 6, 9 and 12, we know that $\mu(n)$ has a large value in the initial stage, which results in a high convergence
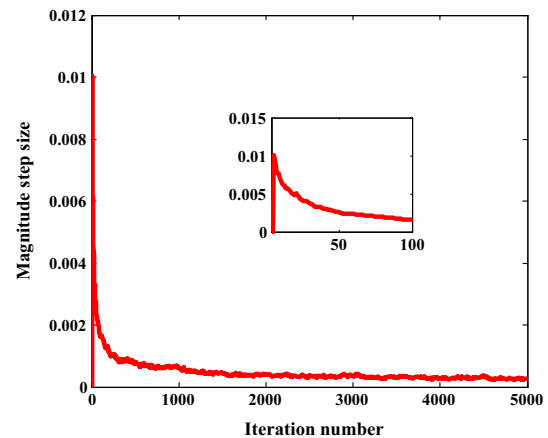

Fig. 12 Evolution of $\mu(n)$ as a function of iteration number under the condition of Experiment 4

rate as expected. After the algorithm converges on the point where a low misalignment is desired, $\mu(n)$ automatically decreases.

5 Conclusions

In the context of system identification using the LMAT-type algorithm, this paper described a way to derive the optimal variable step size of the LMAT based on mean-square deviation analysis. The OPLMAT algorithm is developed in this paper in order to addresses both stationary and non-stationary unknown systems in the presence of several types of noise under low SNR. The optimal step size leads the proposed algorithm to achieve the lowest steady-state error theoretically. The step size can be updated based on the number of iteration. In addition, the computational complexity of this algorithm is less when $L > 4$. Simulation results showed that the proposed algorithm performed better than the LMAT and NLMAT algorithms. The analytical result corroborated the simulations.

References

1. Sayed, A.H.: Adaptive Filters. Wiley, Hoboken (2008)
2. Shin, H.C., Sayed, A.H., Song, W.J.: Variable step size NLMS and affine projection algorithms. IEEE Signal Process. Lett. **11**(2), 132–135 (2004)
3. Kwong, R.H., Johnston, E.W.: A variable step size LMS algorithm. IEEE Trans. Signal Process. **40**(7), 1633–1642 (1992)
4. Ang, W.P., Farhang-Boroujeny, B.: A new class of gradient adaptive step size LMS algorithms. IEEE Trans. Signal Process. **49**(4), 805–810 (2001)
5. Wang, P., Kam, P.-Y.: An automatic step size adjustment algorithm for LMS adaptive filters and an application to channel estimation. Phys. Commun. **5**, 280–286 (2012)
6. Liu, J.C., Yu, X., Li, H.R.: A nonparametric variable step size NLMS algorithm for transversal filters. Appl. Math. Comput. **217**(17), 7365–7371 (2011)
7. Benesty, J., Rey, H., Vega, L.R., et al.: A nonparametric VSSNLMS algorithm. IEEE Signal Process. Lett. **13**(10), 581–584 (2006)
8. Huang, H.-C., Lee, J.: A new variable step size NLMS algorithm and its performance analysis. IEEE Trans. Signal Process. **60**(4), 2055–2060 (2012)
9. Zhao, S., Jones, D.L., Khoo, S., et al.: New variable step sizes minimizing mean-square deviation for the LMS-type algorithms. Circuits Syst. Signal Process. **33**(7), 2251–2265 (2014)
10. Lee, C.H., Park, P.G.: Optimal step size affine projection algorithm. IEEE Signal Process. Lett. **19**(7), 431–434 (2012)
11. Zhi, Y.F., Shang, F.F., Zhang, J., et al.: Optimal step size of pseudo affine projection algorithm. Appl. Math. Comput. **273**, 82–88 (2016)
12. Ciochină, S., Paleologu, C., Benesty, J.: An optimized NLMS algorithm for system identification. Signal Process. **118**, 115–121 (2016)
13. Lee, Y.H., Mok, J.D., Kim, S.D., Cho, S.H.: Performance of least mean absolute third (LMAT) adaptive algorithm in various noise environments. IET Electron. Lett. **34**(2), 241–243 (1998)
14. Walach, E., Widrow, B.: The least mean fourth (LMF) adaptive algorithm and its family. IEEE Trans. Inf. Theory **30**(2), 275–283 (1984)
15. Eweda, E., Bershad, N.J.: Stochastic analysis of a stable normalized least mean fourth algorithm for adaptive noise canceling with a white Gaussian reference. IEEE Trans. Signal Process. **60**(12), 6235–6244 (2012)
16. Zhao, H., Yu, Y., Gao, S., et al.: A new normalized LMAT algorithm and its performance analysis. Signal Process. **105**(12), 399–409 (2014)
17. Arikan, O., Cetin, A.E., Erzin, E.: Adaptive filtering for non-Gaussian stable processes. IEEE Signal Process. Lett. **2**(11), 1400–1403 (1995)
18. Aydin, G., Arikan, O., Cetin, A.E.: Robust adaptive filtering algorithms for α-stable random processes. IEEE Trans. Circuits Syst. II Analog Digit. Signal Process. **46**(2), 198–202 (1999)
19. Price, R.: A useful theorem for nonlinear devices having Gaussian inputs. IRE Trans. Inf. Theory **4**(2), 69–72 (1958)
20. Spiegel, M.R.: Mathematical Handbook of Formulas and Tables. McGraw-Hill, New York (2012)