ELSEVIER

Signal Processing 71 (1998) 29–44

SIGNAL PROCESSING

A new adaptive constrained LMS time delay estimation algorithm

Shyh-Neng Lin, Shiunn-Jang Chern*

Department of Electrical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 80424, ROC

Received 24 October 1996; received in revised form 27 April 1998

Abstract

In this paper, a new adaptive constrained LMS time delay estimation (TDE) algorithm is devised. It is known that in the TDE problem, the time differences between relevant sensors can be modeled as a finite impulse response (FIR) filter whose weight coefficients are samples of a sinc function. Moreover, in case of non-integer TDE, the performance of estimation result is highly dependent upon the convergence rate of weight coefficients of the FIR filter. To speed up the convergence rate of the weight coefficients, in this paper, we propose a new constrained LMS TDE algorithm by making use of the constraint that the sum of the squares of the weight coefficients of the FIR filter equals unity. Here, we show that the constrained optimum solution is identical to the true weights solution which is the error free optimum solution. Also, to document the advantage of the proposed algorithm, the statistical analysis of the steady-state weight-error vector as well as the mean square error of the estimator, using the proposed algorithm, are derived. As confirmed by the theoretical and simulation results, the new proposed algorithm for non-integer TDE outperform the conventional LMS TDE algorithm. © 1998 Elsevier Science B.V. All rights reserved.

Zusammenfassung

In diesem Artikel wird ein neuer adaptiver LMS-Algorithmus mit Nebenbedingungen zur Zeitverzögerungsschätzung (TDE) konstruiert. Bekanntlich können beim TDE-Problem die Zietdifferenzen zwischen zwei relevanten Sensoren durch ein Transversalfilter (FIR-Filter) modelliert werden, dessen Gewichtsfaktoren Abtastwerte einer sinc-Funktion sind. Darüber hinaus hängt die Güte des Schätzergebnisses im Falle nicht-ganzzahliger TDE stark von der Konvergenzgeschwindigkeit der Gewichtsfaktoren des FIR-Filters ab. Um die Konvergenzgeschwindigkeit der Gewichtsfaktoren zu beschleunigen, schlagen wir in diesem Artikel einen neuen LMS-TDE-Algorithmus mit Nebenbedingungen vor, indem die Bedingung ausgenutzt wird, daß die Summe der Quadrate der Gewichtsfaktoren des FIR-Filters den Wert eins ergeben muß. Wir zeigen hier, daß die optimale Lösung unter dieser Bedingung identisch ist mit der Lösung mit den wahren Gewichten, die die fehlerfreie optimale Lösung darstellt. Um die Vorteile des vorgeschlagenen Algorithmus zu belegen, werden sowohl die statistische Analyse des Vektors der Gewichtsfehler im stationären Zustand wie auch der erwartete quadratische Fehler des Schätzers unter Verwendung des vorgeschlagenen Algorithmus hergeleitet. Wie durch die theoretischen und simulierten Ergebnisse bestätigt wird, übertrifft der neue vorgeschlagene Algorithmus für nicht-ganzzahlige TDE den herkömmlichen LMS-TDE-Algorithmus. © 1998 Elsevier Science B.V. All rights reserved.

Résumé

Dans cet article, nous présentons un nouvel algorithme d'estimation du délai temporel (EDT), par les moindres carrés. Il est bien connu que dans un problème d'EDT, les différences temporelles entre des senseurs pertinents peut être

* Corresponding author. Fax: 886-7-5254199; e-mail: chern@ee.nsysu.edu.tw.

PII: S0165-1684(98)00132-7

modélisée comme un filtre à réponse impulsionnelle finie (FIR) dont les coefficients de poids sont des échantillons d'une fonction sinc. De plus, dans le cas d'EDT non entier, la performance de l'estimation dépend fortement du taux de convergence des poids du filtre FIR. Pour accélérer cette convergence, nous proposons dans cet article un nouvel algorithme d'EDT par les moindres carrés, en faisant usage de la contrainte selon laquelle la somme des carrés des coefficients du filtre FIR doit être égale à l'unité. Ici, nous montrons que la solution optimale sous contrainte est identique à la solution exacte, optimale sans erreur. Pour mettre en évidence l'avantage de cet algorithme, nous présentons également une analyse statistique du vecteur d'erreur sur les poids à l'état stable, de même que l'erreur quadratique moyenne de l'estimateur. Comme la théorie et les résultats de simulations le confirment, le nouvel algorithme d'EDT non entier proposé ici dépasse les performances de l'algorithme d'EDT par les moindres carrés conventionnel.

Keywords: Time delay estimation; Constrained LMS TDE algorithm

Notation

D	time delay
D_i	the decimal part of D
$e(n)$	error signal
$e_o(n)$	estimation error produced in the constrained optimum solution
f	smoothing factor
g_i	ith diagonal term of $\boldsymbol{h}_o\boldsymbol{h}_o^T$
h_i	weight coefficients of adaptive filter
$h_m(n)$	weight coefficient with the largest amplitude
h_m^*	optimum weight coefficient of $h_m(n)$
h_m^t	true weight coefficient of $h_m(n)$
$\boldsymbol{h}$	weight vector
$\boldsymbol{h}_o$	optimum weight vector
$\boldsymbol{h}_s$	true weight vector of the true delay
$J_{min}(n)$	minimum mean-square error (MMSE)
$J_{ex}(n)$	excess mean-square error
$\boldsymbol{K}(n)$	weight-error correlation matrix
$k_i(n)$	diagonal terms of $\boldsymbol{K}(n)$
m	the integer part of D
$M(n)$	mean-square difference (MSD)
$M(\infty)$	steady-state MSD
$2p+1$	number of tap weights
$\underline{\boldsymbol{r}}_{yx}$	cross-correlation vector of $y(n)$ and $\boldsymbol{x}(n)$
$\underline{\boldsymbol{R}}_{xx}$	autocorrelation matrix of $\boldsymbol{x}(n)$
$s(n)$	source signal
$s(n-D)$	delayed signal
$w_1(n)$	noise of the first sensor
$w_2(n)$	noise of the second sensor
$\boldsymbol{x}(n)$	vector of $x(n)$
$x(n)$	input signal of the first sensor
$y(n)$	input signal of the second sensor
$z(n)$	filtered output signal
λ	Lagrange multiplier
γ	steady-state mean-square error (MSE) ratio
$\underline{\nabla}(n)$	gradient vector
$\overline{\mu}(n)$	step size
$\boldsymbol{\varepsilon}(n)$	weight-error vector
σ_s^2	source signal power
σ_x^2	power of $x(n)$
σ_w^2	noise power

1. Introduction

The need to estimate time delay between signals received at two spatially separated sensors arises in many applications such as target localization by sonar systems and position estimation by radio navigation systems [1,2]. In the parametric estimation approach, time delay between two sensors can be modeled as an FIR filter whose weight coefficients are samples of a sinc function [3,4]. Generally, the estimate of time delay can be obtained by interpolating on the weights in the filter to select the point in the tapped delay line that corresponds to the peak weight [7] provided that the time delay is an integer multiple of the sampling interval. However, when this is not the case, the estimated weight coefficients of the FIR filter have to be applied to the interpolation formula, implemented by various numerical methods [3,4,7], for extracting the time delay which is not the integer multiple of the sampling interval.

For non-integer time delay estimation (TDE), to reduce the computation requirement, Ching and Chan [5] proposed the look-up table method, where the weight coefficients of the FIR filter are adjusted adaptively by using the constraint that the weight coefficients are samples of the sinc function. During the adaptation processes, an adaptive LMS filtering algorithm is employed but only the weight with the largest amplitude is adapted, which involves a look-up table. The result is a faster adaptation and the elimination of interpolation needed in [3,4,7] for non-integer TDE cases. Since the performance of non-integer TDE by this method is directly related to the convergence behavior of the weight coefficient with the largest amplitude. Therefore, in the low signal-to-noise ratio (SNR) case, the weight coefficient with the largest amplitude, obtained by this method, may not converge to its true value, yielding a wrong result.

To circumvent the drawback described above, recently, Chern and Lin [6] proposed an efficient and simple scheme for non-integer TDE, which is referred to as the direct delay estimation (DDE) formula. Here, the DDE formula does not involve the interpolation formula and the look-up table proposed by Ching and Chan [5]. In [6], it was shown that, for non-integer TDE, the DDE formula approach is superior to the look-up table method in both performance and computation complexity. However, the performance of the DDE formula highly depends on the relative values of the weight coefficient with the largest amplitude and its two adjacent weight coefficients (in both sides).

As described earlier, the performance of non-integer TDE is highly related to the convergence behavior of the estimated weight coefficients. Under general conditions, the weight coefficients of the FIR filter estimated by the conventional LMS TDE algorithm [5–7], are not sufficient enough to converge to their true values. To improve the convergence speed of the weight coefficients to the true values, in this paper, a new constrained LMS TDE algorithm is developed by using the constraint that the sum of the squares of the weight coefficients of the FIR filter equals unity. Also, the weight coefficients, estimated by this new adaptation algorithm, will be applied to the DDE formula for non-integer TDE. To document the advantage of the proposed algorithm, the performance of non-integer TDE is examined and compared to the conventional LMS TDE algorithm with the DDE formula. Moreover, to further investigate the statistical property of this new algorithm with application to TDE, theoretical analysis under certain conditions is developed.

2. The constrained LMS TDE algorithm with DDE formula

To proceed with the development of the new constrained LMS TDE algorithm, we first consider the signal model where the discrete signals, $x(n)$ and $y(n)$, are received at two spatially separated sensors, i.e.,

$$x(n) = s(n) + w_1(n) \tag{1}$$

and

$$y(n) = s(n-D) + w_2(n), \tag{2}$$

respectively. Here, $s(n)$ is the desired source signal, whose delayed version is $s(n-D)$. Also, for convenience, we assume that the noises $w_1(n)$ and $w_2(n)$ are stationary white Gaussian random processes with zero-mean and uncorrelated with each other. Here, parameter D is the time difference between sensor outputs, thus, the problem of TDE is simply to determine D from $x(n)$ and $y(n)$. Ideally, the delayed signal $s(n-D)$ can be represented by [3]

$$s(n-D) = \sum_{i=-\infty}^{\infty} \operatorname{sinc}(i-D)s(n-i), \tag{3}$$

in consequence, Eq. (2) can be expressed as

$$y(n) = \sum_{i=-\infty}^{\infty} h_i x(n-i) - \sum_{i=-\infty}^{\infty} h_i w_1(n-i) + w_2(n), \tag{4}$$

with

$$h_i = \operatorname{sinc}(i-D) = \frac{\sin \pi(i-D)}{\pi(i-D)}, \tag{5}$$

where the largest value of h_i will occur at $i = D$, provided that D is an integer. For convenience to discuss, first, we consider the case without noises in both sensors, such that the last two terms on the right-hand side of Eq. (4) are null. Using the parametric approach [3], the TDE problem can be modeled as depicted in Fig. 1, where the filtered

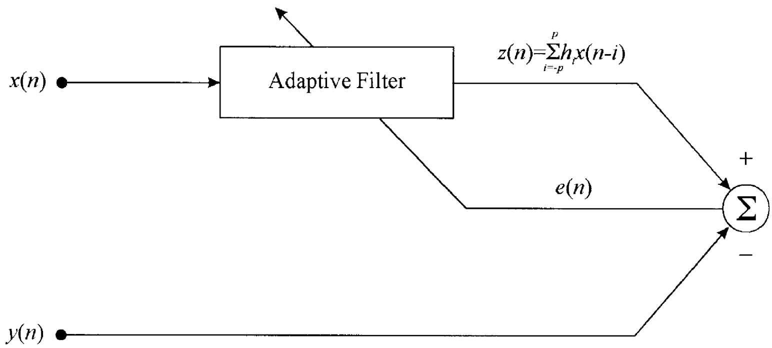


Fig. 1. The configuration of the adaptive time delay estimation scheme.

output is given by

$$z(n) = \sum_{i=-p}^{p} h_i x(n-i). \tag{6}$$

In Eq. (6), h_i, $i = -p, \dots, 0, \dots, p$, are the weight coefficients of the adaptive filter. From Fig. 1, the error signal is defined by

$$e(n) = y(n) - z(n) = y(n) - \boldsymbol{h}^{\mathrm{T}}\boldsymbol{x}(n), \tag{7}$$

where the weight and the input signal vectors are designated by $\boldsymbol{h} = [h_{-p}, \dots, h_0, \dots, h_p]^{\mathrm{T}}$ and $\boldsymbol{x}(n) = [x(n+p), \dots, x(n), \dots, x(n-p)]^{\mathrm{T}}$, respectively, and the superscript T denotes the transpose operation. Indeed, the error signal $e(n)$ defined in Eq. (7) may not be null due to the effects of noise and the value of p which is related to the length of weight coefficients. The sensitivity of choosing the value of p was examined in [4] and will not be discussed here. The weight coefficients h_i in Eq. (6) are adjusted to minimize the mean-square value of the error signal, $E[e^2(n)]$. The mean square error (MSE) is designated by

$$E[e^2(n)] = E[y^2(n)] - 2\boldsymbol{h}^{\mathrm{T}}\underline{\boldsymbol{r}}_{yx} + \boldsymbol{h}^{\mathrm{T}}\boldsymbol{R}_{xx}\boldsymbol{h}, \tag{8}$$

where $\underline{\boldsymbol{r}}_{yx} = E[y(n)\boldsymbol{x}(n)]$ is denoted as the cross-correlation vector between the input vector $\boldsymbol{x}(n)$ and desired response $y(n)$, and $\boldsymbol{R}_{xx} = E[\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)]$ is the autocorrelation matrix of $\boldsymbol{x}(n)$.

The optimum weight coefficients can be obtained by simply minimizing the MSE with respect to weight coefficients. Moreover, it is known that the performance of estimating the weight coefficients can be improved, when the weight coefficients are constrained in a natural manner. Indeed, in the TDE problem, the weight coefficients have the desired characteristics as described above. To see this, we recall from Eq. (5) that $h_i = \mathrm{sinc}(i - D)$, the weights h_i, $i = -p, \dots, 0, \dots, p$, are the samples of a sinc function. Theoretically, the sum of the squares of the samples of a sinc function equals unity, e.g., $\sum_{i=-\infty}^{\infty} h_i^2 = \sum_{i=-\infty}^{\infty} \mathrm{sinc}^2(i - D) = 1$. For example, for $D = 0.2$ and $p = 15$, we have $\sum_{i=-p}^{+p} h_i^2 = 0.995$. Thus, for p to be large enough we may simply assume that $\sum_{i=-p}^{+p} h_i^2 = 1$, or

$$\boldsymbol{h}^{\mathrm{T}}\boldsymbol{h} = 1. \tag{9}$$

In consequence, to minimize Eq. (8) subject to the constraint of Eq. (9), we have

$$J(n) = E[y^2(n)] - 2\boldsymbol{h}^{\mathrm{T}}\underline{\boldsymbol{r}}_{yx} + \boldsymbol{h}^{\mathrm{T}}\boldsymbol{R}_{xx}\boldsymbol{h} + \lambda(\boldsymbol{h}^{\mathrm{T}}\boldsymbol{h} - 1), \tag{10}$$

where λ is the Lagrange multiplier [8]. Taking the derivative of $J(n)$ with respect to $\boldsymbol{h}$, we have the gradient vector

$$\underline{\nabla}(n) = \frac{\partial J(n)}{\partial \boldsymbol{h}} = -2\underline{\boldsymbol{r}}_{yx} + 2\boldsymbol{R}_{xx}\boldsymbol{h} + 2\lambda\boldsymbol{h}. \tag{11}$$

The constrained optimum weight vector, $\boldsymbol{h}_{\mathrm{o}}$, can be obtained by setting Eq. (11) to null, i.e.,

$$\boldsymbol{h}_{\mathrm{o}} = \frac{\underline{\boldsymbol{r}}_{yx}}{\boldsymbol{R}_{xx} + \lambda\boldsymbol{I}}, \tag{12}$$

where $\boldsymbol{I}$ is an identity matrix. For convenience, under the assumption that the desired source signal and noises are white random processes and uncorrelated with each other, we have $\boldsymbol{R}_{xx} = \sigma_x^2\boldsymbol{I} =$

$(\sigma_s^2 + \sigma_w^2)\boldsymbol{I}, \underline{\boldsymbol{r}}_{yx} = \sigma_s^2 \boldsymbol{h}_s$, with $\boldsymbol{h}_s = [\mathrm{sinc}(-p-D), \ldots, \mathrm{sinc}(-D), \ldots, \mathrm{sinc}(p-D)]^{\mathrm{T}}$. Here, $\boldsymbol{h}_s$ is the true weight vector corresponding to the true delay D. Also, σ_x^2, σ_s^2 and σ_w^2 are defined as the average powers of the input signal, desired signal component and noises ($w_1(n)$ or $w_2(n)$), respectively. Consequently, Eq. (12) can be rewritten as

$$\boldsymbol{h}_{\mathrm{o}} = \frac{\sigma_s^2 \boldsymbol{h}_s}{\sigma_s^2 + \sigma_w^2 + \lambda}. \tag{13}$$

It is noted that $\boldsymbol{h}_{\mathrm{o}}$ must satisfy the constraint in Eq. (9), in consequence, the Lagrange multipler, λ , can be solved by substituting the constraint, $\boldsymbol{h}_{\mathrm{o}}^{\mathrm{T}} \boldsymbol{h}_{\mathrm{o}} = 1$, into Eq. (13), yielding

$$\left(\frac{\sigma_s^2}{\sigma_s^2 + \sigma_w^2 + \lambda} \right)^2 \boldsymbol{h}_s^{\mathrm{T}} \boldsymbol{h}_s = 1. \tag{14}$$

Moreover, by using the fact that $\boldsymbol{h}_s^{\mathrm{T}} \boldsymbol{h}_s = 1$, we get

$$\lambda = -\sigma_w^2. \tag{15}$$

Finally, we can easily show that $\boldsymbol{h}_{\mathrm{o}} = \boldsymbol{h}_s$. The implication of this result means that the optimum solution of the constrained optimization problem could achieve the true weight vector $\boldsymbol{h}_s$. However, this is not the case when the conventional optimization approach is adopted and can be viewed as the special case of Eq. (13) with $\lambda = 0$.

Proceeding in a similar way as the conventional LMS adaptation algorithm was derived, we have the new constrained adaptive LMS TDE algorithm, i.e.,

$$\begin{aligned} \boldsymbol{h}(n+1) &= \boldsymbol{h}(n) + \mu(n)[e(n)\boldsymbol{x}(n) - \lambda \boldsymbol{h}(n)] \\ &= [1 + \mu(n)\sigma_w^2]\boldsymbol{h}(n) + \mu(n)e(n)\boldsymbol{x}(n). \end{aligned} \tag{16}$$

To assure the convergence of Eq. (16), the step-size $\mu(n)$ should be in the range $0 < \mu(n) < 2/\mathrm{power}(n)$, where power($n$) is designated as the sum of the square-values of $x(n)$ for all the tap inputs of the FIR filter at time n [6,8]. Again, for $\lambda = 0$, Eq. (16) will reduce to the conventional adaptive LMS TDE algorithm [5–7]

$$\boldsymbol{h}(n+1) = \boldsymbol{h}(n) + \mu(n)e(n)\boldsymbol{x}(n). \tag{17}$$

Thus, if we can show that the steady-state mean weight vector of Eq. (16) could converge to $\boldsymbol{h}_s$, we can expect that the result of non-integer TDE evaluated by this new contrained algorithm with DDE formula will be more accurate than the conventional LMS TDE algorithm with DDE formula [6].

Since in a practical application the noise power, σ_w^2, of Eq. (16), is not available and needs to be estimated. In what follows, we will introduce a simple scheme for estimating the noise power. First, from Eqs. (2) and (6), the cross-correlation between $y(n)$ and $z(n)$, by definition, is given by

$$\begin{aligned} \sigma_{yz}^2 &= E[y(n)z(n)] \\ &= \sigma_s^2[\mathrm{sinc}(-p-D)h_{-p} + \cdots + \mathrm{sinc}(-D)h_0 \\ &\qquad + \cdots + \mathrm{sinc}(p-D)h_p] \\ &= \sigma_s^2 \boldsymbol{h}_s^{\mathrm{T}} \boldsymbol{h} = \sigma_s^2 Q, \end{aligned} \tag{18}$$

with $Q = \boldsymbol{h}_s^{\mathrm{T}} \boldsymbol{h}$. Since the power of the input signal $x(n)$ was defined by $\sigma_x^2 = E[x(n)x(n)] = \sigma_s^2 + \sigma_w^2$, we have

$$\sigma_x^2 - \sigma_{yz}^2 = \sigma_s^2(1-Q) + \sigma_w^2. \tag{19}$$

It is noted that, ideally, when $\boldsymbol{h}$ converges to the true weight vector, $\boldsymbol{h}_s$, the parameter Q will approach unity. Hence, we may use Eq. (19) to estimate the noise power, σ_w^2, and apply it to Eq. (16) during the adaptation processes. From Eq. (19), we learn that the accuracy of estimating the σ_w^2 depends on the estimation results of parameters, σ_x^2 and σ_{yz}^2. One possibility of estimating both σ_x^2 and σ_{yz}^2 can be the one as discussed in [9]:

$$\hat{\sigma}_x^2(n) = f\hat{\sigma}_x^2(n-1) + (1-f)x^2(n) \tag{20}$$

and

$$\hat{\sigma}_{yz}^2(n) = f\hat{\sigma}_{yz}^2(n-1) + (1-f)y(n)z(n), \tag{21}$$

with f being the smoothing factor, $0 < f \leqslant 1$.

Finally, after the weight coefficients of the FIR filter being estimated by the constrained LMS TDE algorithm of Eq. (16), the DDE formula [6] can be applied to estimate the time delay (integer or non-integer). For convenience, we let $h_m(n)$ be the weight of the FIR filter with the largest amplitude at nth iteration, for $i = -p, \ldots, 0, \ldots, +p$, with m being one of the integer index numbers of i. Accordingly, the weights $h_{m+1}(n)$ and $h_{m-1}(n)$ can be determined when $h_m(n)$ is identified. As discussed in [6], if $h_{m+1}(n) \geqslant 0$, in order to achieve better performance for TDE, the DDE formula

$$\hat{D}(n) = m(n) + \frac{h_{m+1}(n)}{h_{m+1}(n) + h_m(n)} \tag{22}$$

is used, otherwise, we have

$$\hat{D}(n) = m(n) - \frac{h_{m-1}(n)}{h_{m-1}(n) + h_m(n)} \tag{23}$$

instead, where $\hat{D}(n)$ is the estimated time delay at the nth iteration. The implication of the DDE formula shown in Eqs. (22) and (23) is that the time delay (integer or non-integer) can be directly estimated if three weight coefficients are available during the adaptation processes.

3. Statistical analysis of the constrained LMS TDE algorithm

In this section, the steady-state characteristic of the constrained LMS TDE algorithm with the DDE formula for TDE is examined.

3.1. Mean square difference of weights

We first evaluate the sum of the weight-error variances (i.e., mean-square difference (MSD)) in steady state. For convenience, the weight-error vector is denoted as $\boldsymbol{\varepsilon}(n) = \boldsymbol{h}(n) - \boldsymbol{h}_o$; from Eq. (16), we get

$$\begin{aligned}\boldsymbol{\varepsilon}(n+1) = {} & [\boldsymbol{I} + \mu(\sigma_w^2 \boldsymbol{I} - \boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n))]\boldsymbol{\varepsilon}(n) \\ & + \mu[\boldsymbol{x}(n)e_o(n) + \sigma_w^2 \boldsymbol{h}_o],\end{aligned} \tag{24}$$

where $\boldsymbol{I}$ is the identity matrix, and $e_o(n) = y(n) - \boldsymbol{h}_o^{\mathrm{T}}\boldsymbol{x}(n)$ is the estimation error when the constrained optimum solution is achieved. By definition, the weight-error correlation matrix is designated by $\boldsymbol{K}(n+1) = E[\boldsymbol{\varepsilon}(n+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(n+1)]$, from Appendix A, we have

$$\begin{aligned}\boldsymbol{K}(n+1) = {} & (1 + 2\mu\sigma_w^2 + \mu^2\sigma_w^4)\boldsymbol{K}(n) \\ & - (\mu + \mu^2\sigma_w^2)[\boldsymbol{K}(n)\boldsymbol{R}_{xx} + \boldsymbol{R}_{xx}\boldsymbol{K}(n)] \\ & + 2\mu^2 \boldsymbol{R}_{xx}\boldsymbol{K}(n)\boldsymbol{R}_{xx} + \mu^2 \boldsymbol{R}_{xx}\mathrm{tr}[\boldsymbol{R}_{xx}\boldsymbol{K}(n)] \\ & + \mu^2 \boldsymbol{R}_{xx} J_{\min} + \mu^2\sigma_w^4 \boldsymbol{h}_o\boldsymbol{h}_o^{\mathrm{T}},\end{aligned} \tag{25}$$

where $J_{\min} = E[e_o^2(n)]$ is the minimum mean-square error (MMSE) and $\mathrm{tr}(\cdot)$ denotes the trace operation. Using the assumption we made earlier, e.g., $\boldsymbol{R}_{xx} = (\sigma_w^2 + \sigma_s^2)\boldsymbol{I} = \sigma_x^2\boldsymbol{I}$ and $\underline{\boldsymbol{r}}_{yx} = \sigma_s^2\boldsymbol{h}_s$, Eq. (25) can be simplified

$$\begin{aligned}\boldsymbol{K}(n+1) = {} & (1 + 2\mu\sigma_w^2 + \mu^2\sigma_w^4)\boldsymbol{K}(n) \\ & - 2(\mu\sigma_x^2 + \mu^2\sigma_x^2\sigma_w^2)\boldsymbol{K}(n) \\ & + 2\mu^2\sigma_x^4\boldsymbol{K}(n) + \mu^2\sigma_x^2 J_{\mathrm{ex}}(n)\boldsymbol{I} \\ & + \mu^2\sigma_x^2 J_{\min}\boldsymbol{I} + \mu^2\sigma_w^4\boldsymbol{h}_o\boldsymbol{h}_o^{\mathrm{T}},\end{aligned} \tag{26}$$

where $J_{\mathrm{ex}}(n) = \mathrm{tr}[\boldsymbol{R}_{xx}\boldsymbol{K}(n)] = \mathrm{tr}[\sigma_x^2\boldsymbol{K}(n)]$ is defined as the excess MSE of the constrained adaptive LMS TDE algorithm. To use the fact that $\underline{\boldsymbol{r}}_{yx} = \sigma_s^2\boldsymbol{h}_s$, $\boldsymbol{h}_o^{\mathrm{T}}\boldsymbol{h}_o = 1$ and $\boldsymbol{h}_o = \boldsymbol{h}_s$, we get $J_{\min} = E[e_o^2(n)] = 2\sigma_w^2$.

To investigate the steady-state performance of the estimated weight coefficients, we let $M(n)$ be the mean-square difference (MSD) which equals to $\mathrm{tr}[\boldsymbol{K}(n)]$. Thus, to evaluate the mean-square difference (MSD), we need to obtain the diagonal terms of $\boldsymbol{K}(n)$, $k_i(n)$, $i = 1, \dots, 2p+1$. From Eq. (26), $k_i(n+1)$ is given by

$$\begin{aligned}k_i(n+1) = {} & k_i(n) - 2\mu\sigma_s^2 k_i(n) + 2\mu^2\sigma_x^2\sigma_s^2 k_i(n) \\ & + \mu^2\sigma_w^4 k_i(n) + \mu^2\sigma_x^2 J_{\mathrm{ex}}(n) \\ & + \mu^2\sigma_x^2 J_{\min} + \mu^2\sigma_w^4 g_i,\end{aligned} \tag{27}$$

where g_i denotes the ith diagonal term of $\boldsymbol{h}_o\boldsymbol{h}_o^{\mathrm{T}}$. In the steady state, Eq. (27) can be further simplifed under certain conditions. In our case, if the step size, μ, can be chosen properly, in general, excess MSE (i.e., J_{ex}) will be relatively smaller than MMSE (i.e., $J_{\min}$). For example, for $\mu = 0.001$ and SNR = 5 dB, we have $J_{\mathrm{ex}}(\infty)/J_{\min} \approx 1/37$. So in Eq. (27), $\sigma_x^2 J_{\mathrm{ex}}(\infty)$ can be neglected comparing with $\sigma_x^2 J_{\min}$. Thus, in the steady state, we may obtain the closed-form expression of $k_i(\infty)$,

$$k_i(\infty) = \frac{\mu[(1 + \sigma_w^2/\sigma_s^2)J_{\min} + (\sigma_w^4/\sigma_s^2)g_i]}{2[1 - \mu\sigma_s^2(1 + (\sigma_w^2/\sigma_s^2) + \frac{1}{2}(\sigma_w^2/\sigma_s^2)^2)]}. \tag{28}$$

In consequence, the steady-state solution of MSD is given by

$$M(\infty)=\mathrm{tr}[\boldsymbol{K}(\infty)]=\sum_{i=1}^{2p+1}k_i(\infty)$$
$$=\frac{\mu[(1+\sigma_w^2/\sigma_s^2)(2p+1)J_{\min}+(\sigma_w^4/\sigma_s^2)\boldsymbol{h}_o^{\mathrm{T}}\boldsymbol{h}_o]}{2[1-\mu\sigma_s^2(1+(\sigma_w^2/\sigma_s^2)+\frac{1}{2}(\sigma_w^2/\sigma_s^2)^2)]}$$
$$=\frac{\mu[(1+\sigma_w^2/\sigma_s^2)(2p+1)J_{\min}+(\sigma_w^4/\sigma_s^2)]}{2[1-\mu\sigma_s^2(1+(\sigma_w^2/\sigma_s^2)+\frac{1}{2}(\sigma_w^2/\sigma_s^2)^2)]}. \tag{29}$$

Since $\sigma_s^2/\sigma_w^2=\mathrm{SNR}$ is the signal-to-noise ratio, Eq. (29) can be expressed as

$$M(\infty)=\frac{\mu[(1+1/\mathrm{SNR})(2p+1)J_{\min}+(1/\mathrm{SNR})\sigma_w^2]}{2[1-\mu\sigma_s^2(1+(1/\mathrm{SNR})+\frac{1}{2}(1/\mathrm{SNR})^2)]}. \tag{30}$$

From Eq. (30), we see that the steady-state MSD is controlled by the parameters of the SNR, the step size, μ, and the value of p which is related to the length of the FIR filter.

3.2. Mean square error of estimated time delay

To evaluate the quality of estimated time delay (or estimator), using the proposed algorithm with the DDE formula, in this subsection, a theoretical analysis is performed. Recall that, at the nth iteration, after having obtained the weight vector by Eq. (16), using the proposed algorithm, the estimated time delay, $\hat{D}(n)$, can be evaluated by the DDE formula defined in Eq. (22) or Eq. (23). The mean square error (MSE) of the estimated and the true time delay, at the nth iteration, is defined by $E[(\hat{D}(n)-D)^2]$. From Eq. (B.7) of Appendix B, we have the steady-state MSE to be

$$\lim_{n\to\infty}E[(\hat{D}(n)-D)^2]$$
$$\approx\frac{0.5\mu\delta_1}{(h_m^{\mathrm{t}}+h_{m+1}^{\mathrm{t}})^2}\left\{(1-D_i)^2\left[\left(1+\frac{\sigma_w^2}{\sigma_s^2}\right)J_{\min}+\frac{\sigma_w^4}{\sigma_s^2}g_{m+1}\right]+D_i^2\left[\left(1+\frac{\sigma_w^2}{\sigma_s^2}\right)J_{\min}+\frac{\sigma_w^4}{\sigma_s^2}g_{m+1}\right]\right\}, \tag{31}$$

where D_i is the decimal part of D, $\delta_1=1/(1-\mu\sigma_s^2(1+(\sigma_w^2/\sigma_s^2)+\frac{1}{2}(\sigma_w^2/\sigma_s^2)^2))$, $J_{\min}=E[e_o^2(n)]=2\sigma_w^2$, $g_m=(h_m^{\mathrm{t}})^2$ and $g_{m+1}=(h_{m+1}^{\mathrm{t}})^2$. Here, h_m^{t} and h_{m+1}^{t} represent the true weight coefficients of $h_m(n)$ and $h_{m+1}(n)$, respectively.

For comparison, the steady-state MSE of the estimated and the true time delay, using the conventional LMS TDE algorithm with DDE formula, can be obtained in a similar way as we did for Eq. (31) (see Eq. (B.10) of Appendix B),

$$\lim_{n\to\infty}E[(\hat{D}(n)-D)^2]$$
$$\approx\frac{0.5\mu\delta_2}{(h_m^{\mathrm{t}}+h_{m+1}^{\mathrm{t}})^2}\left(1+\frac{\sigma_w^2}{\sigma_s^2}\right)^2[(1-D_i)^2J'_{\min}+D_i^2J'_{\min}], \tag{32}$$

with $\delta_2=1/(1-\mu\sigma_s^2(1+\sigma_w^2/\sigma_s^2))$ and $J'_{\min}=\sigma_s^2+\sigma_w^2-\sigma_s^4/(\sigma_s^2+\sigma_w^2)$.

At this moment, it is interesting to compare the steady-state MSE estimators of the proposed algorithm with the conventional algorithm. To do so, we define the steady-state MSE ratio, γ, to be

$$\gamma=\frac{\text{Eq. (32)}}{\text{Eq. (31)}}. \tag{33}$$

In this study, since the value of $\mu\sigma_s^2\ll 1$, Eq. (33) can be simplified to get, i.e.,

$$\gamma=\frac{(1-D_i)^2(1+\frac{1}{2}\sigma_w^2/\sigma_s^2)+D_i^2(1+\frac{1}{2}\sigma_w^2/\sigma_s^2)}{(1-D_i)^2(1+\frac{1}{2}(\sigma_w^2/(\sigma_s^2+\sigma_w^2))g_{m+1})+D_i^2(1+\frac{1}{2}(\sigma_w^2/(\sigma_s^2+\sigma_w^2))g_m)}$$
$$=\frac{(1-D_i)^2(1+\frac{1}{2}(1/\mathrm{SNR}))+D_i^2(1+\frac{1}{2}(1/\mathrm{SNR}))}{(1-D_i)^2(1+\frac{1}{2}(1/(\mathrm{SNR}+1))g_{m+1})+D_i^2(1+\frac{1}{2}(1/(\mathrm{SNR}+1))g_m)}, \tag{34}$$

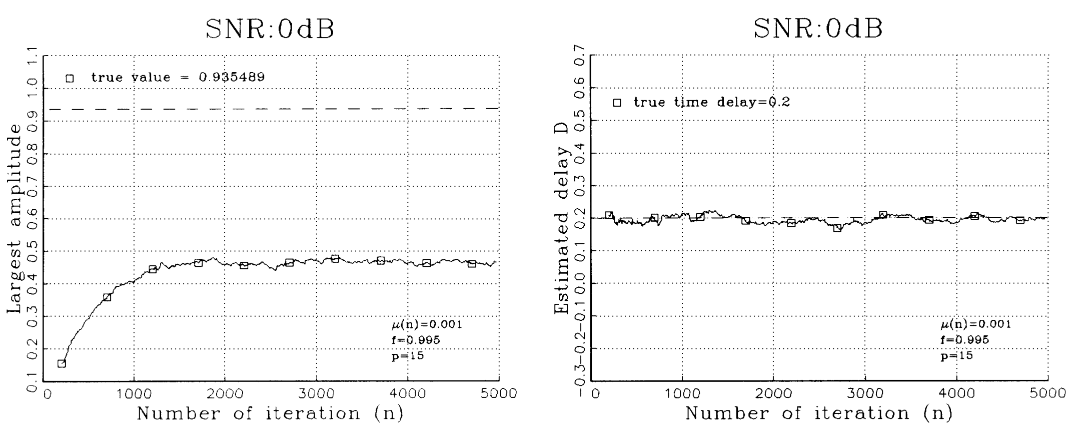
where $\mathrm{SNR} = \sigma_s^2/\sigma_w^2$ is the signal-to-noise ratio. Note that if the value of γ is greater than unity, we may have an improved performance by the proposed algorithm. Also, observed from Eq. (34), when SNR becomes lower, the value of γ becomes relatively large. To see this, we consider the case with true time delay $D = 0.2$ (i.e., $m = 0$ and $D_i = 0.2$) and the step-size $\mu = 0.001$, in consequence, we have the corresponding parameters $g_m = g_0 = (h_0^{\mathrm{t}})^2 = \mathrm{sinc}^2(0 - 0.2) = 0.8751$ and $g_{m+1} = g_1 = (h_1^{\mathrm{t}})^2 = \mathrm{sinc}^2(1 - 0.2) = 0.0547$. Substituting these parameters in Eq. (34), we get $\gamma = 1.15$ for $\mathrm{SNR} = 5\,\mathrm{dB}$, $\gamma = 1.23$ for $\mathrm{SNR} = 3\,\mathrm{dB}$, and $\gamma = 1.46$ for $\mathrm{SNR} = 0\,\mathrm{dB}$, respectively. This means that the performance improvement is relevant to the SNR, that is, the lower the value of SNR, more improvement with the proposed algorithm can be obtained with respect to the conventional algorithm.

4. Computer simulation results

To demonstrate the merits of our method, computer simulation is carried out to evaluate the performance of non-integer TDE, where both stationary and nonstationary time delay environments are considered. Also, the accuracy of the theoretical analysis in terms of steady-state mean square difference (MSD) of weights and steady-state mean square error (MSE) of the estimator, are examined. In computer simulation the source signal $s(n)$ and noises, $w_1(n)$ and $w_2(n)$, are generated from three separated white Gaussian random signal generators with zero mean. The delayed signal, $s(n-D)$, is obtained by passing $s(n)$ through an FIR filter of order 41 (should be greater than the weight coefficients, $2p+1$, of the adaptive filter), with its impulse response being the samples of a sinc function. On the other hand, the number of weight coefficients of the adaptive filter is chosen to be 31 (e.g. $p = 15$). Moreover, the simulation results are the average of 100 independent realizations (or runs).

4.1. Stationary time delay case

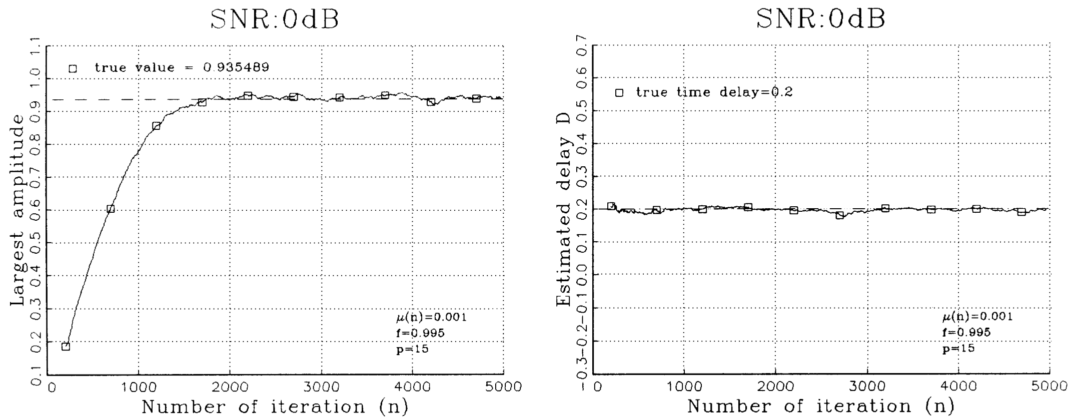
First, we consider the case with constant time delay, $D = 0.2$, and to have fair comparison the step size is chosen to be $\mu(n) = 0.001$ for all methods. Also, the smoothing factor of Eqs. (20) and (21) is chosen to be $f = 0.995$. It is noted that for $D = 0.2$, the value of the true weight coefficient with the largest amplitude, by definition, should be $h_0 = \mathrm{sinc}(0 - 0.2) = 0.935489$, e.g., $m = 0$. For $\mathrm{SNR} = 0\,\mathrm{dB}$, as depicted in Fig. 2(a), we see that the weight coefficient with the largest amplitude estimated by the conventional LMS TDE algorithm could not converge to the true value ($h_0 = 0.935489$).



(a) the value of the weight coefficient with the largest amplitude

(b) time delay estimation

Fig. 2. The performance of TDE using the LMS TDE algorithm with DDE formula, for $D = 0.2$ and $\mathrm{SNR} = 0\,\mathrm{dB}$.



(a) the value of the weight coefficient with the largest amplitude

(b) time delay estimation

Fig. 3. The performance of TDE using the constrained LMS TDE algorithm with DDE formula for $D = 0.2$ and SNR $= 0$ dB.

Table 1
The comparison of TDE performance between the conventional method and the proposed method, for constant time delay $D = 0.2$

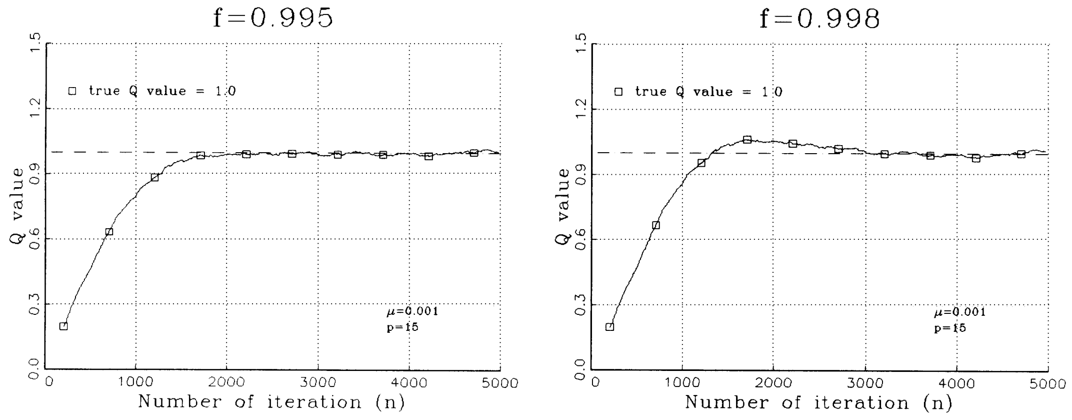
	5 dB		3 dB		0 dB	
	Mean	MSE	Mean	MSE	Mean	MSE
Conventional method	0.196990	0.000250	0.196134	0.000484	0.194526	0.001547
Proposed method	0.198163	0.000216	0.197884	0.000383	0.196759	0.001015

But, the performance of non-integer TDE evaluated by the DDE formula is still satisfied as evident from Fig. 2(b). On the other hand, as depicted in Fig. 3(a), under the same condition as in the proposed algorithm, the weight coefficient with the largest amplitude estimated by our method converges much faster and closer to the true value. Also, from Fig. 3(b) we learn that the new proposed algorithm with the DDE formula for non-integer TDE is superior to the one shown in Fig. 2(b). For comparison, the statistical characteristics of TDE results, in terms of the steady-state mean and mean-square error (MSE) of the estimated time delay, with different SNR are listed in Table 1 as reference. From Table 1, we learn that the proposed method (i.e., constrained LMS TDE algorithm with DDE formula) has better mean and smaller MSE than the conventional method (i.e., LMS TDE algorithm with DDE formula). This is, especially, true when SNR becomes lower.

Although, ideally, the new constrained LMS TDE algorithm could reach the true weight vector, it depends on the accuracy of the estimated noise power of Eq. (19), or, equivalently, it is related to the convergent property of parameter Q (of Eq. (18)).Therefore, it is interesting to see the effect of parameter Q. For SNR $= 0$ dB, Fig. 4(a) and Fig. 4(b) show the convergence behavior of parameter Q with the smoothing factor to be $f = 0.995$ and $f = 0.998$, respectively. From Fig. 4(a) and Fig. 4(b) it can be observed that the parameter Q could converge near to unity (desired value), which implies that the proposed scheme for estimating the noise power is reasonable.

4.2. The theoretical analysis results

In this section, the accuracy of the theoretical analysis of the steady-state MSD (see Eq. (30)) as



(a) the smoothing factor $f = 0.995$ (b) the smoothing factor $f = 0.998$

Fig. 4. The convergence behavior of Q value for SNR = 0 dB and different smoothing factor.

well as the steady-state MSE (see Eqs. (31) and (32)) for the estimator of TDE is investigated. First, to verify the accuracy of the steady-state MSD for weight coefficient, we consider the case with $D = 0.2$ for SNR = 0 dB and the parameters $\mu = 0.0007, f = 0.995$ and $p = 15$. From Fig. 5, the experimental result of steady-state MSD agrees with the theoretical analysis result (dot line) very well. Some other results for SNR = 5 and 10 dB with $\mu = 0.001$, are listed in Table 2 as reference. To further examine the accuracy of the theoretical results for SNR = 5 dB with different values of step-size, μ, is illustrated in Table 3. From Table 3, again, we found that the theoretical results of

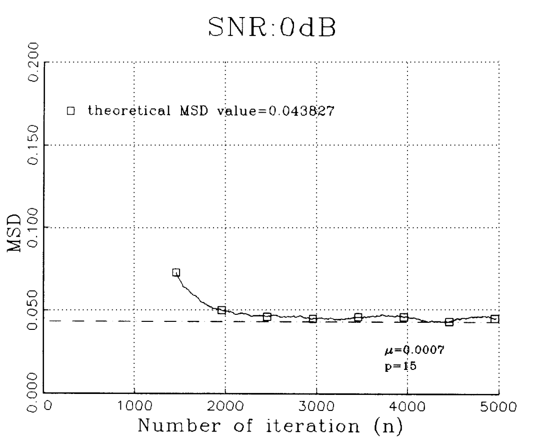


Fig. 5. The experimental result of MSD for SNR = 0 dB.

Table 2
The accuracy of the analysis results of the steady-state MSD by Eq. (30), for different values of SNR, with parameters $D = 0.2$, $f = 0.995$ and $p = 15$

SNR (dB)	Experimental MSD value	Theoretical MSD value
0	0.045167	0.043827
5	0.013058	0.012919
10	0.003434	0.003413

Table 3
The accuracy of the analysis results of the steady-state MSD by Eq. (30), for different values of μ, with parameters SNR = 5 dB, $D = 0.2, f = 0.995$ and $p = 15$

μ	Experimental MSD value	Theoretical MSD value	Error ratio (%)	$\frac{J_{ex}(\infty)}{J_{min}}$
0.0010	0.013058	0.012919	1.06	0.027262
0.0013	0.017063	0.016800	1.56	0.035608
0.0016	0.021135	0.020704	2.08	0.044105
0.0019	0.025269	0.024568	2.85	0.052752
0.0022	0.029466	0.028459	3.54	0.061555

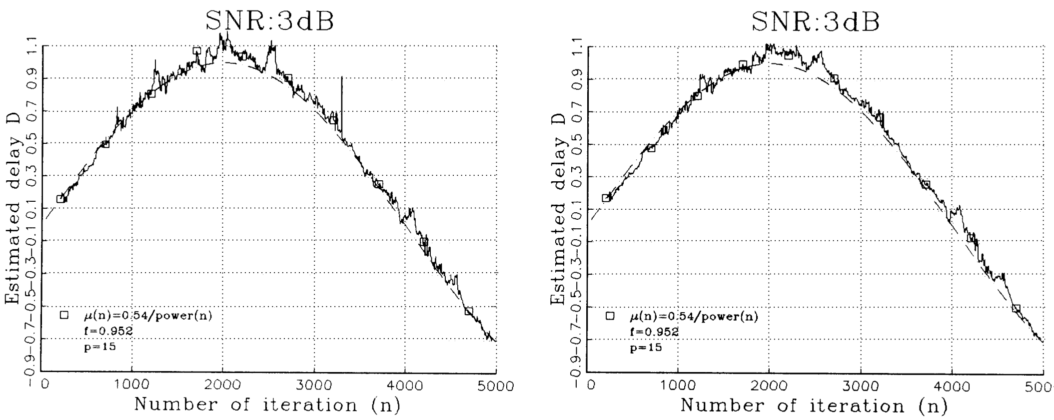
steady-state MSD agreed quite well with the experimental results. Especially, when the parameter μ becomes smaller, the error ratio between the theoretical and experimental values is relatively small. This is because the derivation of Eq. (30) was performed under the assumption that the excess MSE (i.e., J_{ex}) is much smaller than the MMSE (i.e., J_{min}).

Next, we would like to verify the accuracy of the steady-state MSE of the estimator of TDE derived in Eqs. (31) and (32) for the constrained and unconstrained adaptive LMS TDE algorithms with DDE formula, respectively. The results are listed in Table 4. From Table 4, we learn that the improved ratio of performance, in terms of steady-state MSE of the estimator, using the theoretical expression (see Eq. (34)) is consistent with the one of experimental results. Also, we found that the performance improvement is relevant to the SNR, e.g., $\gamma = 1.15$ for SNR $= 5$ dB, $\gamma = 1.23$ for SNR $= 3$ dB, and $\gamma = 1.46$ for SNR $= 0$ dB. This means that the lower the value of SNR the more improvement with the proposed algorithm can be obtained with respect to the conventional adaptive LMS TDE algorithm. We may conclude that the proposed algorithm is superior to the conventional algorithm in terms of non-integer TDE.

4.3. Nonstationary time delay case

Next, to investigate the tracking capability of the proposed method, we consider the case with nonstationary time delay,

$$D(n) = \sin\left(\frac{n\pi}{4000}\right). \tag{35}$$



(a) the LMS TDE algorithm with DDE formula

(b) the constrained LMS TDE algorithm with DDE formula

Fig. 6. The performance comparison of TDE using the LMS TDE algorithms with and without constraint, for $D(n) = \sin(n\pi/4000)$ and SNR $= 3$ dB.

Table 4
The accuracy of the analysis results of the steady-state MSE based on Eqs. (31) and (32), with parameters $D = 0.2$ (i.e., $D_i = 0.2$), $\mu = 0.001, f = 0.995$ and $p = 15$

	5 dB		3 dB		0 dB	
	Exp. MSE	Theo. MSE	Exp. MSE	Theo. MSE	Exp. MSE	Theo. MSE
Conventional method: Eq. (32)	0.000250	0.000240	0.000484	0.000467	0.001547	0.001496
Proposed method: Eq. (31)	0.000216	0.000209	0.000383	0.000380	0.001015	0.001022
Steady-state MSE ratio: $\gamma =$ Eq. (32)/Eq. (31)	1.16	1.15	1.26	1.23	1.52	1.46

Here, $\mu(n) = 0.54/\text{power}(n)$ and the smoothing factor $f = 0.952$ are chosen. The results of estimated delay, $D(n)$, with number of iterations, are shown in Fig. 6(a) and Fig.6 (b), for SNR $= 3$ dB. From Fig. 6(a) and Fig. 6(b), we learn that the result of TDE obtained by the constrained LMS TDE algorithm with DDE formula is superior to that of the conventional LMS TDE algorithm with DDE formula. Specifically, during the adaptation processes, when the integer index $(m(n))$ is changed from one $(m(n) = 1)$ to zero $(m(n) = 0)$, the spurious peak occurred in Fig. 6(a) for the conventional LMS TDE algorithm with DDE formula. However, this is not the case when the new constrained LMS TDE algorithm with DDE formula is used.

To further examine the tracking capability, we consider another case with

$$D(n) = 0.2 + \frac{n}{500} \times 0.06. \tag{36}$$

The results are shown in Fig. 7(a) and Fig. 7(b), for SNR $= 0$ dB with $\mu(n) = 0.25/\text{power}(n)$ and $f = 0.952$. Similarly, the spurious peak also occurs in Fig. 7(a) for the conventional LMS TDE algorithm with DDE formula, when $m(n)$ is changed from zero $(m(n) = 0)$ to one $(m(n) = 1)$.

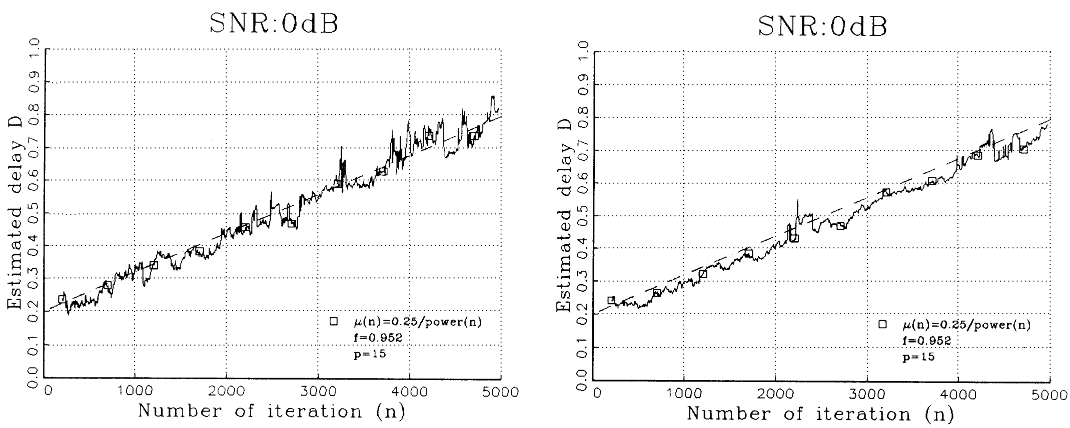
5. Conclusions

In this paper, a new constrained LMS TDE algorithm has been developed in a noisy environment for speeding up the convergence rate of the weight coefficients of the FIR filter. This will result in having better performance for non-integer time delay estimation (TDE) with the DDE formula, compared with the conventional adaptive LMS TDE algorithm.

Moreover, the closed form expressions of the steady-state MSE of the estimators for TDE, using the proposed algorithm, were derived. Based on these theoretical expressions, we also showed that the proposed algorithm did perform better than the conventional one. Especially, in Section 4.2, we showed that the lower the value of SNR the more the improvement with the proposed algorithm was obtained, compared with the conventional adaptive LMS TDE algorithm.

Acknowledgements

The authors wish to thank the reviewers for their valuable suggestions and comments. Also, the financial support of this study by the National Science Council, Republic of China, under contract number NSC-86-2213-E-110-016, is greatly acknowledged.



(a) the LMS TDE algorithm with DDE formula

(b) the constrained LMS TDE algorithm with DDE formula

Fig. 7. The performance comparison of TDE using the LMS TDE algorithms with and without constraint, for $D(n) = 0.2 + n/500 \times 0.06$ and SNR $= 0$ dB.

Appendix A.

In this appendix, the derivation of Eq. (25) is given in what follows. By definition, the weight-error correlation matrix is designated by $\boldsymbol{K}(n+1) = E[\boldsymbol{\varepsilon}(n+1)\boldsymbol{\varepsilon}^{\mathrm{T}}(n+1)]$. From Eq. (24), we have

$$\boldsymbol{K}(n+1) = (1 + 2\mu\sigma_w^2 + \mu^2\sigma_w^4)\boldsymbol{K}(n) - (\mu + \mu^2\sigma_w^2)[\boldsymbol{K}(n)\boldsymbol{R}_{xx} + \boldsymbol{R}_{xx}\boldsymbol{K}(n)] + \mu^2 E[\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)\boldsymbol{\varepsilon}(n)\boldsymbol{\varepsilon}^{\mathrm{T}}(n)\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)] + \mu^2\boldsymbol{R}_{xx}J_{\min} + \mu^2\sigma_w^4\boldsymbol{h}_{\mathrm{o}}\boldsymbol{h}_{\mathrm{o}}^{\mathrm{T}}, \tag{A.1}$$

where $J_{\min} = E[e_{\mathrm{o}}^2(n)]$ denotes the minimum mean square error (MMSE). The third term on the right-hand side of Eq. (A.1) involves fourth-order moments of sample vectors of the input process. These high-order moments can be evaluated by using the Gaussian moment factoring theorem [8]. Let z_1, z_2, z_3 and z_4 denote four samples of a real Gaussian process with zero mean. By the Gaussian moment factoring theorem [8], we have

$$E[z_1z_2z_3z_4] = E[z_1z_2]E[z_3z_4] + E[z_1z_3]E[z_2z_4] + E[z_1z_4]E[z_2z_3]. \tag{A.2}$$

To use Eq. (A.2), we express the $(2p+1)$-by-$(2p+1)$ matrix representing the multiple product $\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)\boldsymbol{\varepsilon}(n)\boldsymbol{\varepsilon}^{\mathrm{T}}(n)\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)$ as a multiple sum of the elements of the component vectors. Let the brace notation $\{a_{ij}(n)\}$ denote this matrix having elements $a_{ij}(n)$, with $i,j = 0,1,\dots,2p$. We may then write

$$\{a_{ij}(n)\} = \boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)\boldsymbol{\varepsilon}(n)\boldsymbol{\varepsilon}^{\mathrm{T}}(n)\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n) = \left\{\sum_{l=1}^{2p+1}\sum_{m=1}^{2p+1} x(n-i)x(n-l)\varepsilon_l(n)\varepsilon_m(n)x(n-m)x(n-j)\right\}, \tag{A.3}$$

where $x(n-i)x(n-l)$ denotes the element on the ith row and lth column of $\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)$, $\varepsilon_l(n)\varepsilon_m(n)$ denotes the element on the lth row and mth column of $\boldsymbol{\varepsilon}(n)\boldsymbol{\varepsilon}^{\mathrm{T}}(n)$, and $x(n-m)x(n-j)$ denotes the element on the mth row and jth column of $\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)$. Thus, we have

$$\begin{aligned}
E[\{a_{ij}(n)\}] &= E[\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)\boldsymbol{\varepsilon}(n)\boldsymbol{\varepsilon}^{\mathrm{T}}(n)\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)]\\
&= \sum_{l=1}^{2p+1}\sum_{m=1}^{2p+1} E[x(n-i)x(n-l)\varepsilon_l(n)\varepsilon_m(n)x(n-m)x(n-j)]\\
&= \sum_{l=1}^{2p+1}\sum_{m=1}^{2p+1} E[x(n-i)x(n-l)x(n-m)x(n-j)]E[\varepsilon_l(n)\varepsilon_m(n)]\\
&= \sum_{l=1}^{2p+1}\sum_{m=1}^{2p+1} E[x(n-i)x(n-l)]E[x(n-m)x(n-j)]E[\varepsilon_l(n)\varepsilon_m(n)]\\
&\quad + \sum_{l=1}^{2p+1}\sum_{m=1}^{2p+1} E[x(n-i)x(n-m)]E[x(n-l)x(n-j)]E[\varepsilon_l(n)\varepsilon_m(n)]\\
&\quad + E[x(n-i)x(n-j)]\sum_{l=1}^{2p+1}\sum_{m=1}^{2p+1} E[x(n-l)x(n-m)]E[\varepsilon_l(n)\varepsilon_m(n)]\\
&= 2\boldsymbol{R}_{xx}\boldsymbol{K}(n)\boldsymbol{R}_{xx} + \boldsymbol{R}_{xx}\,\mathrm{tr}[\boldsymbol{R}_{xx}\boldsymbol{K}(n)],
\end{aligned} \tag{A.4}$$

where $\mathrm{tr}(\cdot)$ denotes the trace operation. Finally, substituting Eq. (A.4) into Eq. (A.1), we obtain

$$\boldsymbol{K}(n+1) = (1 + 2\mu\sigma_w^2 + \mu^2\sigma_w^4)\boldsymbol{K}(n) - (\mu + \mu^2\sigma_w^2)[\boldsymbol{K}(n)\boldsymbol{R}_{xx} + \boldsymbol{R}_{xx}\boldsymbol{K}(n)] + 2\mu^2\boldsymbol{R}_{xx}\boldsymbol{K}(n)\boldsymbol{R}_{xx} + \mu^2\boldsymbol{R}_{xx}\,\mathrm{tr}[\boldsymbol{R}_{xx}\boldsymbol{K}(n)] + \mu^2\boldsymbol{R}_{xx}J_{\min} + \mu^2\sigma_w^4\boldsymbol{h}_{\mathrm{o}}\boldsymbol{h}_{\mathrm{o}}^{\mathrm{T}}. \tag{A.5}$$

Appendix B.

The steady-state theoretical expression of the mean square error (MSE) of the estimator for time delay estimation is derived, in this appendix, under certain conditions. To proceed with the analysis, we recall that, at the nth iteration, after having obtained the weight vector by Eq. (16), using the new constrained LMS TDE algorithm, the estimated time delay, $\hat{D}(n)$, can be evaluated by the DDE formula defined in Eq. (22) or Eq. (23). Also, as described before the value of true time delay, D (or $D(n)$) was defined by $D = m + D_i$, where m denotes the integer part of D and D_i is the decimal part of D.

In our cases, Eq. (22) (i.e., $\hat{D}(n) = m + h_{m+1}(n)/(h_m(n) + h_{m+1}(n))$) is adopted, the error between the estimated and true time delay, at the nth iteration, is given by

$$\hat{D}(n) - D = \left(m + \frac{h_{m+1}(n)}{h_m(n) + h_{m+1}(n)}\right) - (m + D_i) = \frac{h_{m+1}(n)}{h_m(n) + h_{m+1}(n)} - D_i. \tag{B.1}$$

For convenience, we denote the estimated weight errors, $\varepsilon_m(n)$ and $\varepsilon_{m+1}(n)$, to be $\varepsilon_m(n) = h_m(n) - h_m^*$ and $\varepsilon_{m+1}(n) = h_{m+1}(n) - h_{m+1}^*$, respectively. Here, h_m^* and h_{m+1}^* represent the optimum weight coefficients of $h_m(n)$ and $h_{m+1}(n)$, respectively. Now, by using the fact that $D_i = h_{m+1}^*/(h_m^* + h_{m+1}^*)$, Eq. (B.1) can be rewritten as

$$\begin{aligned}\hat{D}(n) - D &= \frac{h_{m+1}^* + \varepsilon_{m+1}(n)}{h_m^* + h_{m+1}^* + \varepsilon_m(n) + \varepsilon_{m+1}(n)} - D_i \\ &= \frac{[h_{m+1}^* - D_i(h_m^* + h_{m+1}^*)] + \varepsilon_{m+1}(n) - D_i[\varepsilon_m(n) + \varepsilon_{m+1}(n)]}{h_m^* + h_{m+1}^* + \varepsilon_m(n) + \varepsilon_{m+1}(n)} \\ &= \frac{\varepsilon_{m+1}(n) - D_i[\varepsilon_m(n) + \varepsilon_{m+1}(n)]}{h_m^* + h_{m+1}^* + \varepsilon_m(n) + \varepsilon_{m+1}(n)}.\end{aligned} \tag{B.2}$$

In the steady state, since we have $\varepsilon_m(n) \ll h_m^*$ and $\varepsilon_{m+1}(n) \ll h_{m+1}^*$ (the errors are smaller compared to the optimum values), we may simplify Eq. (B.2),

$$\hat{D}(n) - D \approx \frac{\varepsilon_{m+1}(n) - D_i[\varepsilon_m(n) + \varepsilon_{m+1}(n)]}{h_m^* + h_{m+1}^*} = \frac{(1 - D_i)\varepsilon_{m+1}(n) - D_i\varepsilon_{m+1}(n)}{h_m^* + h_{m+1}^*}. \tag{B.3}$$

In consequence, the steady-state MSE of the estimator, can be obtained as

$$\begin{aligned}\lim_{n\to\infty} E[(\hat{D}(n) - D)^2] &\approx \frac{1}{(h_m^* + h_{m+1}^*)^2} \lim_{n\to\infty} E[((1 - D_i)\varepsilon_{m+1}(n) - D_i\varepsilon_{m+1}(n))^2] \\ &= \frac{1}{(h_m^* + h_{m+1}^*)^2} \lim_{n\to\infty} E[(1 - D_i)^2\varepsilon_{m+1}^2(n) - 2D_i(1 - D_i)\varepsilon_{m+1}(n)\varepsilon_m(n) + D_i^2\varepsilon_m^2(n)] \\ &= \frac{1}{(h_m^* + h_{m+1}^*)^2}\left\{(1 - D_i)^2 \lim_{n\to\infty} E[\varepsilon_{m+1}^2(n)] + D_i^2 \lim_{n\to\infty} E[\varepsilon_m^2(n)]\right\}.\end{aligned} \tag{B.4}$$

It is noted that both $\varepsilon_m(n)$ and $\varepsilon_{m+1}(n)$ are two components of weight-error vector $\boldsymbol{\varepsilon}(n)$ defined in Eq. (24). The last two terms, $\lim_{n\to\infty} E[\varepsilon_{m+1}^2(n)]$ and $\lim_{n\to\infty} E[\varepsilon_m^2(n)]$, on the right-hand side of Eq. (B.4), can be easily obtained, from $k_i(\infty)$ of Eq. (28), that is,

$$\lim_{n\to\infty} E[\varepsilon_{m+1}^2(n)] = k_{m+1}(\infty) = \frac{\mu[(1 + \sigma_w^2/\sigma_s^2)J_{\min} + (\sigma_w^4/\sigma_s^2)g_{m+1}]}{2[1 - \mu\sigma_s^2(1 + (\sigma_w^2/\sigma_s^2) + \frac{1}{2}(\sigma_w^2/\sigma_s^2)^2)]} \tag{B.5}$$

and

$$\lim_{n\to\infty} E[\varepsilon_m^2(n)] = k_m(\infty) = \frac{\mu[(1+\sigma_w^2/\sigma_s^2)J_{\min} + (\sigma_w^4/\sigma_s^2)g_m]}{2[1-\mu\sigma_s^2(1+(\sigma_w^2/\sigma_s^2)+\frac{1}{2}(\sigma_w^2/\sigma_s^2)^2)]}, \tag{B.6}$$

with $J_{\min} = 2\sigma_w^2$, $g_m = (h_m^*)^2$ and $g_{m+1} = (h_{m+1}^*)^2$. Again, from Eqs. (13) and (15), the constrained optimum weight vector, $\boldsymbol{h}_o$, (i.e., $\boldsymbol{h}_o = [h_{-p}^*, \ldots, h_0^*, \ldots, h_p^*]^{\mathrm{T}}$), is shown to be identical to the true weight vector, $\boldsymbol{h}_s$ (i.e., $\boldsymbol{h}_s = [h_{-p}^{\mathrm{t}}, \ldots, h_0^{\mathrm{t}}, \ldots, h_{\mathrm{p}}^{\mathrm{t}}]^{\mathrm{T}}$), we then have $h_m^* = h_m^{\mathrm{t}}$ and $h_{m+1}^* = h_{m+1}^{\mathrm{t}}$. Finally, by substituting Eqs. (B.5) and (B.6) into Eq. (B.4), we get the steady-state MSE of the estimator, by the constrained LMS TDE algorithm with DDE formula, i.e.,

$$\lim_{n\to\infty} E[(\hat{D}(n)-D)^2] \approx \frac{0.5\mu\delta_1}{(h_m^{\mathrm{t}}+h_{m+1}^{\mathrm{t}})^2}\left\{(1-D_i)^2\left[\left(1+\frac{\sigma_w^2}{\sigma_s^2}\right)J_{\min} + \frac{\sigma_w^4}{\sigma_s^2}g_{m+1}\right] + D_i^2\left[\left(1+\frac{\sigma_w^2}{\sigma_s^2}\right)J_{\min} + \frac{\sigma_w^4}{\sigma_s^2}g_{m+1}\right]\right\}, \tag{B.7}$$

with $\delta_1 = 1/(1-\mu\sigma_s^2(1+(\sigma_w^2/\sigma_s^2)+\frac{1}{2}w(\sigma_w^2/\sigma_s^2)^2))$.

For comparison, we need also to derive the steady-state MSE of the estimator, by the conventional LMS TDE algorithm with DDE formula. This, in fact, can be performed in a similar way as we derived Eq. (28). First, $k_i(\infty)$ of the conventional LMS TDE algorithm can be obtained as we did for Eq. (28), that is, $k_i(\infty) = \mu J'_{\min}/(2[1-\mu\sigma_s^2(1+\sigma_w^2/\sigma_s^2)])$, with $J'_{\min} = \sigma_s^2+\sigma_w^2-\sigma_s^4/(\sigma_s^2+\sigma_w^2)$. In consequence, we have the corresponding expressions of $\lim_{n\to\infty} E[\varepsilon_{m+1}^2(n)]$ and $\lim_{n\to\infty} E[\varepsilon_m^2(n)]$, i.e.,

$$\lim_{n\to\infty} E[\varepsilon_{m+1}^2(n)] = k_{m+1}(\infty) = \frac{\mu J'_{\min}}{2[1-\mu\sigma_s^2(1+\sigma_w^2/\sigma_s^2)]} \tag{B.8}$$

and

$$\lim_{n\to\infty} E[\varepsilon_m^2(n)] = k_m(\infty) = \frac{\mu J'_{\min}}{2[1-\mu\sigma_s^2(1+\sigma_w^2/\sigma_s^2)]}. \tag{B.9}$$

Next, from Eq. (13) with $\lambda = 0$, we have the optimum weight coefficients, h_m^* and h_{m+1}^*, of the conventional LMS TDE algorithm, to be $h_m^* = (\sigma_s^2/(\sigma_s^2+\sigma_w^2))h_m^{\mathrm{t}}$ and $h_{m+1}^* = (\sigma_s^2/(\sigma_s^2+\sigma_w^2))h_{m+1}^{\mathrm{t}}$, respectively. Finally, we have the steady-state MSE of the estimator, by the conventional LMS TDE algorithm with DDE formula, i.e.,

$$\lim_{n\to\infty} E[(\hat{D}(n)-D)^2] \approx \frac{0.5\mu\delta_2}{(h_m^{\mathrm{t}}+h_{m+1}^{\mathrm{t}})^2}\left(1+\frac{\sigma_w^2}{\sigma_s^2}\right)^2[(1-D_i)^2J'_{\min} + D_i^2J'_{\min}], \tag{B.10}$$

with $\delta_2 = 1/(1-\mu\sigma_s^2(1+\sigma_w^2/\sigma_s^2))$. This completes the derivation.

References

[1] G.C. Carter, Time delay estimation for passive sonar signal processing, IEEE Trans. Acoust. Speech Signal Process. ASSP-29 (1981) 463–470.

[2] G.C. Carter, Coherence and time delay estimation, Proc. IEEE 75 (2) (1987) 236–255.

[3] Y.T. Chan, J.M. Riley, T.B. Plant, A parameter estimation approach to time delay estimation and signal detection, IEEE Trans. Acoust. Speech Signal Process. ASSP-28 (1) (1980) 8–16.

[4] Y.T. Chan, J.M. Riley, T.B. Plant, Modeling of time delay and its application to estimation of nonstationary delays, IEEE Trans. Acoust. Speech Signal Process. ASSP-29 (3) (1981) 577–581.

[5] P.C. Ching, Y.T. Chan, Adaptive time delay estimation with constraints, IEEE Trans. Acoust. Speech Signal Process. ASSP-36 (4) (1988) 599–602.

[6] S.J. Chern, S.N. Lin, An adaptive time delay estimation with direct computation formula, J. Acoust. Soc. Amer. 96 (2) (August 1994) 811–820.
[7] P.L. Feintuch, N.J. Bershad, F.A. Reed, Time delay estimation using the LMS adaptive filter-Dynamic behavior, IEEE Trans. Acoust. Speech Signal Process. ASSP-29 (3) (1981) 571–576.
[8] S. Haykin, Adaptive Filter Theory, 2nd Ed., Prentice-Hall, Englewood Cliffs, NJ USA, 1991.
[9] D.H. Youh, N. Ahmed, G.C. Carter, An adaptive approach for time delay estimation of band-limited signals, IEEE Trans. Acoust. Speech Signal Process. ASSP-31 (3) (1983) 780–784.