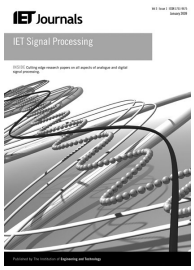
Published in IET Signal Processing
Received on 9th May 2012
Revised on 9th December 2012
Accepted on 19th January 2013
doi: 10.1049/iet-spr.2012.0090



ISSN 1751-9675

Stochastic modelling and analysis of filtered-x least-mean-square adaptation algorithm

Iman Tabatabaei Ardekani, Waleed H. Abdulla

Electrical Engineering Department, The University of Auckland, Private Bag 92019, Auckland 1142, Auckland, New Zealand
E-mail: itab001@aucklanduni.ac.nz

Abstract: This study represents a stochastic model for the adaptation process performed on adaptive control systems by the filtered-x least-mean-square (FxLMS) algorithm. The main distinction of this model is that it is derived without using conventional simplifying assumptions regarding the physical plant to be controlled. This model is then used to derive a set of closed-form mathematical expressions for formulating steady-state performance, stability condition and learning rate of the FxLMS adaptation process. These expressions are the most general expressions, which have been proposed so far. It is shown that some previously derived expressions can be obtained from the proposed expressions as special and simplified cases. In addition to computer simulations, different experiments with a real-time control setup confirm the validity of the theoretical findings.

1 Introduction

When an adaptive control system is followed by an unwanted signal channel (secondary path), it cannot be adjusted by standard adaptive algorithms (update equations). Alternatively, filtered reference (Fx) adaptive algorithms can be used for such cases [1]. In these algorithms, the training data sequence (reference signal) is filtered by an estimate model of the secondary path (secondary path model) before being used in the main update equation. The most popular algorithm of this type is the filtered-x least-mean-square (FxLMS) algorithm [2] which has been derived from the standard LMS algorithm. The implementation of the FxLMS algorithm is very convenient but its theoretical analysis has been a challenge over the last two decades.

There have been several contributions in modelling and analysis of the FxLMS adaptation process [3–12]. However, only a few have intended to find closed-form mathematical expressions for formulating this process. Even if such expressions were derived, simplified cases had to be considered. This is mainly because of the high level of complexity associated with the mathematical modelling of the FxLMS adaptation process. Long summarised early work on this subject [5], whereas deriving closed-form expressions for the stability bound ($\mu_{\max}$) and steady-state performance (J_{ss}) of this process. However, he derived these expressions for a pure delay secondary path. Elliott derived another expression for $\mu_{\max}$, which has become more popular than Long's expression [7]. Bjarnason conducted a comprehensive analysis on the FxLMS adaptation process [8]; however, once he intended to derive closed-form expressions for $\mu_{\max}$ and J_{ss}, he had to simplify his model by assuming a pure delay secondary path, a perfectly accurate secondary path model, and an ideal reference signal (broad-band white). Also, Vicente and Masgrau derived another expression for $\mu_{\max}$ assuming that the reference signal is deterministic [13].

All the aforementioned closed-form expressions for $\mu_{\max}$ were derived for a pure delay secondary path, a perfectly accurate secondary path model and a broad-band white reference signal. These assumption are not valid in real-life applications. For example, practical results reported in [14] show that a reliable $\mu_{\max}$ is different with those proposed in available literature. The authors have investigated behaviours of the FxLMS adaptation process in practical conditions. This investigation resulted in a set of closed-form mathematical expressions for formulating behaviours of the FxLMS adaptation process. The results have been published in a series of research papers [9–12]. In [9], basic closed-form expressions for formulating behaviours of the FxLMS adaptation process with a general secondary path were derived. In [10], these expressions were generalised by considering a practical reference signal. In [12] the expressions derived in [9] were generalised by taking influences of imperfect secondary path models into account. The relative drawback of the analysis conducted in [10] is that it assumes a perfect secondary path model. Also, the relative drawback of the analysis conducted in [12] is that it assumes an ideal reference signal. This paper intends to overcome the relative drawbacks of the previously published papers [9–12]. In fact, the main distinction between this paper and the authors' previous works is that this paper simultaneously considers a general secondary path, an arbitrary secondary path model, and a realistic reference signal, resulting in the most comprehensive stochastic analysis which have been conducted on the FxLMS adaptation process.

The rest of the paper is organised as follows. Sections 2 describes the FxLMS algorithm and the adaptation process performed by this algorithm. Section 3 develops a stochastic model for the FxLMS adaptation process. Section 4 uses the developed model to derive closed-form expressions for formulating behaviours of the FxLMS adaptation process. Also, this section discusses the theoretical results and use them to derive the conventional expressions for $\mu_{\max}$ and J_{∞} as special and simplified cases. Section 5 represents computer simulation results. Section 6 introduces the fully implemented embedded setup used in this research and show the agreement between theoretical and experimental results. Finally, Section 7 gives the concluding remarks.

2 Mathematical description

Fig. 1 demonstrates the functional block diagram of a general FxLMS-based adaptive control system, where it is desired to control a dual-channel physical plant by adaptive adjustment of a digital filter. The physical plant consists of two separate signal channels p and s, called primary and secondary paths, respectively. The plant includes an stochastic process, generating the reference signal $x(n)$. The error signal $e(n)$ has to be controlled by the control signal $y(n)$. The response of the secondary path to $y(n)$ is then combined with that of the primary path, resulting in $e(n)$. Here, $y(n)$ is estimated by an adaptive filter, w, and the FxLMS algorithm is responsible for the adjustment of w in such a way that $e(n)$ becomes minimal. Note that the primary path output, $d(n)$, cannot be measured here but both of $x(n)$ and $e(n)$ can be measured.

2.1 FxLMS algorithm

Assuming that w is a finite impulse response (FIR) filter of length L, $y(n)$ can be represented by

$$y(n) = \boldsymbol{w}^{\mathrm{T}}(n)\boldsymbol{x}(n) \tag{1}$$

$$\boldsymbol{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \dots & x(n-L+1) \end{bmatrix}^{\mathrm{T}} \tag{2}$$

$$\boldsymbol{w}(n) = \begin{bmatrix} w_0(n) & w_1(n) & \dots & w_{L-1}(n) \end{bmatrix}^{\mathrm{T}} \tag{3}$$

where $\boldsymbol{x}(n)$ and $\boldsymbol{w}(n)$ are the reference and weight vectors, respectively. The secondary path is assumed to be a FIR system of length Q with an unknown impulse response given by

$$\boldsymbol{s} \triangleq \begin{bmatrix} s_0 & s_1 & \cdots & s_{Q-1} \end{bmatrix}^{\mathrm{T}} \tag{4}$$

In this case, $e(n)$ can be formulated from Fig. 1 by

$$e(n) = d(n) - \sum_{q=0}^{Q-1} s_q \boldsymbol{w}^{\mathrm{T}}(n-q)\boldsymbol{x}(n-q) \tag{5}$$

Aiming at minimising the power of $e(n)$, the FxLMS algorithm updates $\boldsymbol{w}(n)$ by

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu e(n)\hat{\boldsymbol{f}}(n) \tag{6}$$

where μ is the step-size and $\hat{\boldsymbol{f}}(n)$, called the filtered-reference vector, is obtained by filtering $\boldsymbol{x}(n)$ by an estimate model of the secondary path. The secondary path estimate model is assumed to be another FIR system with the impulse response given by a $Q \times 1$ vector

$$\hat{\boldsymbol{s}} \triangleq \begin{bmatrix} \hat{s}_0 & \hat{s}_1 & \cdots & \hat{s}_{M-1} & 0_{\text{white } Q} & \cdots & 0 \end{bmatrix}^{\mathrm{T}} \tag{7}$$

where M is the length of the impulse response ($M < Q$) . In this case, $\hat{\boldsymbol{f}}(n)$ can be computed by

$$\hat{\boldsymbol{f}}(n) = \sum_{m=0}^{Q-1} \hat{s}_m \boldsymbol{x}(n-m) \tag{8}$$

In [2], it is proved that the FxLMS algorithm causes the weight vector $\boldsymbol{w}(n)$ to converge to the optimal Wiener-Hopf filter $\boldsymbol{w}_{\mathrm{o}}$; thus, the optimal error signal can be found from (5) as

$$e_{\mathrm{o}}(n) = d(n) - \boldsymbol{w}_{\mathrm{o}}^{\mathrm{T}} \sum_{q=0}^{Q-1} s_q \boldsymbol{x}(n-q) \tag{9}$$

2.2 Alternative expression for FxLMS algorithm

In the analysis of the FxLMS algorithm, it is more convenient to use the rotated-reference vector and rotated-weight misalignment vector, instead of the original reference and weight vectors [9]. This is because the auto-correlation matrix of the rotated-reference vector is diagonal, and the equilibrium point of the rotated-weight misalignment vector is the origin. Here, a rotation matrix, $\boldsymbol{F}$, is produced by

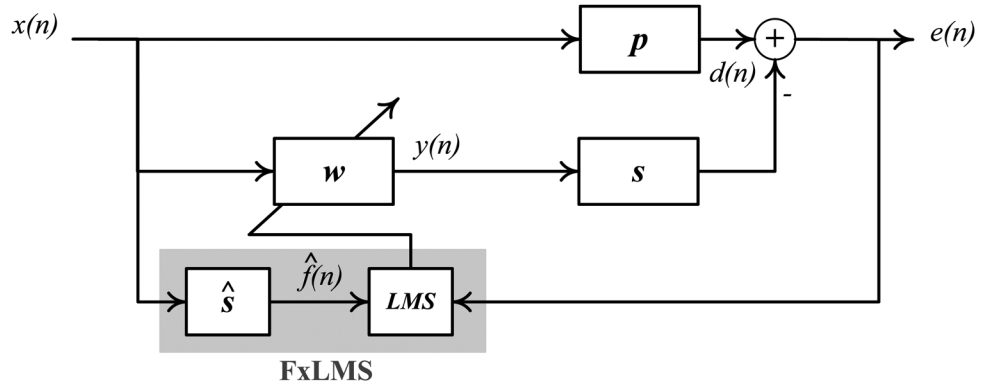


Fig. 1 *Functional block diagram for FxLMS-based adaptive control*

doi: 10.1049/iet-spr.2012.0090

diagonalising the Toeplitz matrix $\boldsymbol{R} = E\{\boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{T}}(n)\}$ as

$$\boldsymbol{R} = \boldsymbol{F}\boldsymbol{\Lambda}\boldsymbol{F}^{\mathrm{T}} \tag{10}$$

$\boldsymbol{F}$ is a modal matrix, formed by the Eigenvectors of $\boldsymbol{R}$ as

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_0 & \boldsymbol{F}_1 & \dots & \boldsymbol{F}_{L-1} \end{bmatrix} \tag{11}$$

and $\boldsymbol{\Lambda}$ is a diagonal matrix, formed by the Eigenvalues of $\boldsymbol{R}$ as

$$\boldsymbol{\Lambda} = \mathrm{diag}\left(\lambda_0, \lambda_1, \dots, \lambda_{L-1}\right) \tag{12}$$

Now, the rotated-reference vector $\boldsymbol{z}(n)$ and the rotated-weight misalignment vector $\boldsymbol{c}(n)$ are defined as

$$\boldsymbol{z}(n) = \boldsymbol{F}^{\mathrm{T}}\boldsymbol{x}(n) \tag{13}$$

$$\boldsymbol{c}(n) = \boldsymbol{F}^{\mathrm{T}}\left(\boldsymbol{w}(n) - \boldsymbol{w}_{\mathrm{o}}\right) \tag{14}$$

By using (5), (13) and (14), the following expression for $e(n)$ can be obtained

$$e(n) = e_{\mathrm{o}}(n) - \sum_{q=0}^{Q-1} s_q \boldsymbol{c}^{\mathrm{T}}(n-q)\boldsymbol{z}(n-q) \tag{15}$$

Also, the FxLMS update equation given in (6) can be reexpressed as

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) + \mu e(n)\hat{\boldsymbol{g}}(n) \tag{16}$$

where the rotated-filtered-reference vector is defined as

$$\hat{\boldsymbol{g}}(n) = \boldsymbol{F}^{\mathrm{T}}\hat{\boldsymbol{f}}(n) \tag{17}$$

2.3 Independence assumptions

The analysis of the FxLMS algorithm with stochastic reference signals is usually performed by using a set of independence assumptions. As stated in [15], 'the independence assumptions apparently cannot be analytically justified for practical cases, but this is perhaps the best that can be done from the pragmatic point of view of obtaining a good trade-off between model realism and model tractability'. The primary independence assumption states that any sequential vectors of a Gaussian signal, can be considered as an independent identically distributed (i.i.d) sequence with zero mean [15]. Based on this assumption, it can be shown that ($\forall n, m, p \in \mathbb{N}$)

$$E\left\{\boldsymbol{x}(n-m)\boldsymbol{x}^{\mathrm{T}}(n-p)\right\} = \delta_{m,p}\boldsymbol{R} \tag{18}$$

where $\boldsymbol{x}(n)$ denotes a tap vector of a Gaussian signal. According to the secondary independence assumption, for the problem of adaptive identification of an unknown system with FIR, the optimal error $e_{\mathrm{o}}(n)$ is independent of both of the input signal and weight vectors [15]. Also, it is usually assumed that elements of $\boldsymbol{w}(n)$ and samples of the reference signal are statistically independent [1, 2].

3 Stochastic modelling of MSE and excess-MSE functions

The FxLMS adaptation process can be studied by analysing the variation of the mean square error (MSE), defined as

$$J(n) = E\left\{e^2(n)\right\} \tag{19}$$

In this paper, $J(n)$ is initially expressed by using the expression given in [9]

$$J(n) = J_{\mathrm{o}} + \sum_{q=0}^{Q-1} s_q^2 E\left\{\boldsymbol{c}^{\mathrm{T}}(n-q)\boldsymbol{\Lambda}\boldsymbol{c}(n-q)\right\} \tag{20}$$

where $J_{\mathrm{o}} = E\left\{e_{\mathrm{o}}^2(n)\right\}$ is the minimal MSE level. Now, the excess-MSE function $J_{\mathrm{ex}}(n)$ is defined as a dynamic measure, determining the deviation of the MSE function from its minimal level; thus

$$J(n) = J_{\mathrm{o}} + J_{\mathrm{ex}}(n) \tag{21}$$

From (20) and (21), $J_{\mathrm{ex}}(n)$ can be formulated by

$$J_{\mathrm{ex}}(n) \triangleq \sum_{q=0}^{Q-1} s_q^2 E\left\{\boldsymbol{c}^{\mathrm{T}}(n-q)\boldsymbol{\Lambda}\boldsymbol{c}(n-q)\right\} \tag{22}$$

Substituting (12) and (14) into (22) results in

$$J_{\mathrm{ex}}(n) = \sum_{q=0}^{Q-1}\sum_{l=0}^{L-1} \lambda_l s_q^2 m_l(n-q) \tag{23}$$

where $m_l(n)$ denoted the second-order moments of the lth adaptive weight

$$m_l(n) \triangleq E\left\{c_l^2(n)\right\} \quad l = 0, 1, \dots, L-1 \tag{24}$$

In (24), c_l denoted the lth element of $\boldsymbol{c}$. In order to investigate the variation of $J_{\mathrm{ex}}(n)$ in the FxLMS algorithm, its time difference is defined as

$$\Delta J_{\mathrm{ex}}(n) = J_{\mathrm{ex}}(n+1) - J_{\mathrm{ex}}(n) \tag{25}$$

From (23) and (25), $\Delta J_{\mathrm{ex}}(n)$ can be expressed as

$$\Delta J_{\mathrm{ex}}(n) = \sum_{q=0}^{Q-1}\sum_{l=0}^{L-1} \lambda_l s_q^2 \Delta m_l(n-q) \tag{26}$$

where $\Delta m_l(n)$ is given by

$$\Delta m_l(n) = E\left\{c_l^2(n+1)\right\} - E\left\{c_l^2(n)\right\} \tag{27}$$

On the other hand, from (16), it can be shown that

$$c_l(n+1) = c_l(n) + \mu\hat{g}_l(n)e(n) \tag{28}$$

where $\hat{g}_l(n)$ is the lth element of $\hat{\boldsymbol{g}}(n)$. Now, by combining (27) and (28), $\Delta m_l(n)$ is formulated by

$$\Delta m_l(n) = \mu^2 E\left\{\hat{g}_l^2(n)e^2(n)\right\} + 2\mu E\left\{c_l(n)\hat{g}_l(n)e(n)\right\} \tag{29}$$

From (17), it can be shown that

$$\hat{g}_l(n) = \boldsymbol{F}_l^{\mathrm{T}} \hat{\boldsymbol{f}}(n) \tag{30}$$

By substituting the above expression for $\hat{g}_l(n)$ into (29), $\Delta m_l(n)$ can be formulated as

$$\Delta m_l(n) = A_l(n) + B_l(n) \tag{31}$$

$$A_l(n) = \mu^2 \boldsymbol{F}_l^{\mathrm{T}} E\left\{\hat{\boldsymbol{f}}(n)\hat{\boldsymbol{f}}^{\mathrm{T}}(n)e^2(n)\right\}\boldsymbol{F}_l \tag{32}$$

$$B_l(n) = 2\mu \boldsymbol{F}_l^{\mathrm{T}} E\left\{\hat{\boldsymbol{f}}(n)c_l(n)e(n)\right\} \tag{33}$$

Appendices 1 and 2 calculate $A_l(n)$ and $B_l(n)$ as

$$\begin{aligned} A_l(n) = {} & \mu^2 \lambda_l \|\hat{\boldsymbol{s}}\|^2 J_{\mathrm{o}} \\ & + \mu^2 \lambda_l \|\hat{\boldsymbol{s}}\|^2 \sum_{p=0}^{Q-1} \sum_{k=0}^{L-1} \lambda_k s_p^2 m_k(n-p) \end{aligned} \tag{34}$$

$$\begin{aligned} B_l(n) = {} & -2\mu\lambda_l \sum_{p=0}^{Q-1} s_p \hat{s}_p m_l(n-p) \\ & + 2\mu^2 \lambda_l^2 \sum_{p,r=0}^{Q-1} r s_p \hat{s}_p s_r \hat{s}_r m_l(n-p-r) \end{aligned} \tag{35}$$

Now, substituting (34) and (35) into (31) results in

$$\begin{aligned} \Delta m_l(n) = {} & \mu^2 \lambda_l \|\hat{\boldsymbol{s}}\|^2 J_0 - 2\mu\lambda_l \sum_{p=0}^{Q-1} s_p \hat{s}_p m_l(n-p) \\ & \mu^2 \lambda_l \|\hat{\boldsymbol{s}}\|^2 \sum_{p=0}^{Q-1} \sum_{k=0}^{L-1} \lambda_k s_p^2 m_k(n-p) \\ & + 2\mu^2 \lambda_l^2 \sum_{p,r=0}^{Q-1} r s_p \hat{s}_p s_r \hat{s}_r m_l(n-p-r) \end{aligned} \tag{36}$$

Subsequently, substituting the above expression for $\Delta m_l(n)$ into (26) gives

$$\begin{aligned} \Delta J_{\mathrm{ex}}(n) = {} & \mu^2 \lambda_{\mathrm{rms}}^2 \|\hat{\boldsymbol{s}}\|^2 ||\boldsymbol{s}||^2 L J_{\mathrm{o}} \\ & - 2\mu \sum_{p,q=0}^{Q-1} \sum_{l=0}^{L-1} \lambda_l^2 s_q^2 s_p \hat{s}_p m_l(n-p-q) \\ & + \mu^2 \|\hat{\boldsymbol{s}}\|^2 L \lambda_{\mathrm{rms}}^2 \sum_{q,p=0}^{Q-1} \sum_{k=0}^{L-1} \lambda_k s_q^2 s_p^2 m_k(n-p-q) \\ & + 2\mu^2 \sum_{q,p,r=0}^{Q-1} \sum_{l=0}^{L-1} r \lambda_l^3 s_q^2 s_p \hat{s}_p s_r \hat{s}_r m_l(n-p-r-q) \end{aligned} \tag{37}$$

where λ_{rms} is the rms value of the Eignevalues. In a slow adaptation process, the second-order moments are updated slowly: $m_l(n-p-r-q) \simeq m_l(n-p-q)$. By using this assumption, (37) is simplified to

$$\begin{aligned} \Delta J_{\mathrm{ex}}(n) = {} & \mu^2 \lambda_{rms}^2 \|\hat{\boldsymbol{s}}\|^2 ||\boldsymbol{s}||^2 L J_{\mathrm{o}} \\ & - \mu \sum_{q,p=0}^{Q-1} \sum_{l=0}^{L-1} \gamma_{l,p,q} m_l(n-p-q) \end{aligned} \tag{38}$$

where

$$\gamma_{l,p,q} = \lambda_l s_q^2 \left(2\lambda_l s_p \hat{s}_p - \mu \lambda_{rms}^2 \|\hat{\boldsymbol{s}}\|^2 \Delta\right) \tag{39}$$

$$\Delta = L s_p^2 + 2\left(\frac{\lambda_l}{\lambda_{rms}}\right)^2 s_p \hat{s}_p \frac{\boldsymbol{s}^{\mathrm{T}} \boldsymbol{\Psi} \hat{\boldsymbol{s}}}{\|\hat{\boldsymbol{s}}\|^2} \tag{40}$$

In the above formulation, the diagonal matrix $\boldsymbol{\Psi}$ is defined as

$$\boldsymbol{\Psi} = \mathrm{diag}(0, 1, \ldots, Q-1) \tag{41}$$

Equations (38)–(41) give a general stochastic model for the FxLMS adaptation process.

3.1 Modelling influences of input signal band-width

The analysis of the FxLMS adaptation process can be simplified by assuming a broad-band white reference signal; however, actual reference signals can be only white over a limited frequency range (band-limited white). Considering such a realistic signal is one of the distinction of the analysis performed in this paper. The authors proved that for a band-limited white signal with band-width of B_{w} and power of σ_x^2, the Eigenvalues of the auto-correlation matrix can only be equal to zero or $\lambda_l = B_{\mathrm{w}}^{-1}\sigma_x^2$ [10]. Also, they proved that $\lambda_{\mathrm{rms}} = B_{\mathrm{w}}^{-0.5}\sigma_x^2$ [10]. Using these equalities, (39) and (40) are simplified to

$$\gamma_{l,p,q} = \lambda_l \frac{\sigma_x^2}{B_{\mathrm{w}}} s_q^2 \left(2 s_p \hat{s}_p - \mu \sigma_x^2 \|\hat{\boldsymbol{s}}\|^2 \Delta\right) \tag{42}$$

$$\Delta = L s_p^2 + \frac{2}{B_{\mathrm{w}}} s_p \hat{s}_p \frac{\boldsymbol{s}^{\mathrm{T}} \boldsymbol{\Psi} \hat{\boldsymbol{s}}}{\|\hat{\boldsymbol{s}}\|^2} \tag{43}$$

Therefore substituting (42) into (38) results in

$$\begin{aligned} \Delta J_{\mathrm{ex}}(n) = {} & \mu^2 \frac{\sigma_x^4}{B_{\mathrm{w}}} \|\hat{\boldsymbol{s}}\|^2 ||\boldsymbol{s}||^2 L J_{\mathrm{o}} \\ & - \mu \frac{\sigma_x^2}{B_{\mathrm{w}}} \sum_{q,p=0}^{Q-1} \sum_{l=0}^{L-1} s_q^2 \lambda_l \\ & \times \left(2 s_p \hat{s}_p - \mu \sigma_x^2 \|\hat{\boldsymbol{s}}\|^2 \Delta\right) m_l(n-p-q) \end{aligned} \tag{44}$$

doi: 10.1049/iet-spr.2012.0090

Now, combining (23), (43) and (44) results in

$$\Delta J_{\mathrm{ex}}(n) = \mu^2 \frac{\sigma_x^4}{B_{\mathrm{w}}} \|\hat{\boldsymbol{s}}\|^2 ||\boldsymbol{s}||^2 L J_{\mathrm{o}} + \mu^2 \frac{\sigma_x^4}{B_{\mathrm{w}}} \|\hat{\boldsymbol{s}}\|^2 L \sum_{p=0}^{Q-1} s_p^2 J_{\mathrm{ex}}(n-p) - 2\mu \frac{\sigma_x^2}{B_{\mathrm{w}}} \left(1 - \mu \frac{\sigma_x^2}{B_{\mathrm{w}}} \boldsymbol{s}^{\mathrm{T}} \boldsymbol{\Psi} \hat{\boldsymbol{s}}\right) \sum_{p=0}^{Q-1} s_p \hat{s}_p J_{\mathrm{ex}}(n-p) \tag{45}$$

3.2 Modelling influences of imperfect secondary path models

Now, functions $\mathcal{J}(n)$ and $\hat{\mathcal{J}}(n)$ are defined as

$$\mathcal{J}(n) = \sum_{p=0}^{Q-1} s_p^2 J_{\mathrm{ex}}(n-p) \tag{46}$$

and

$$\hat{\mathcal{J}}(n) = \sum_{p=0}^{Q-1} s_p \hat{s}_p J_{\mathrm{ex}}(n-p) \tag{47}$$

For a slow adaptation process, it is shown in [12] that

$$\hat{\mathcal{J}}(n) = \frac{\boldsymbol{s}^{\mathrm{T}} \hat{\boldsymbol{s}}}{||\boldsymbol{s}||^2} \mathcal{J}(n) \tag{48}$$

Substituting (46)–(48) into (45) results in

$$\Delta J_{\mathrm{ex}}(n) = \mu^2 \frac{\sigma_x^4}{B_{\mathrm{w}}} \|\hat{\boldsymbol{s}}\|^2 ||\boldsymbol{s}||^2 L J_{\mathrm{o}} + \left[\mu^2 \frac{\sigma_x^4}{B_{\mathrm{w}}} \|\hat{\boldsymbol{s}}\|^2 L - 2\mu \frac{\sigma_x^2}{B_{\mathrm{w}}} \frac{\boldsymbol{s}^{\mathrm{T}} \hat{\boldsymbol{s}}}{||\boldsymbol{s}||^2} \left(1 - \mu \frac{\sigma_x^2}{B_w} \boldsymbol{s}^{\mathrm{T}} \boldsymbol{\Psi} \hat{\boldsymbol{s}}\right)\right] \mathcal{J}(n) \tag{49}$$

Herein, perfectness ratios ρ_1, ρ_2 and ρ_3 are defined as

$$\rho_1 = \frac{\|\hat{\boldsymbol{s}}\|^2}{||\boldsymbol{s}||^2} \tag{50}$$

and

$$\rho_2 = \frac{\boldsymbol{s}^{\mathrm{T}} \hat{\boldsymbol{s}}}{\|\hat{\boldsymbol{s}}\|^2} \tag{51}$$

and

$$\rho_3 = \frac{\boldsymbol{s}^{\mathrm{T}} \boldsymbol{\Psi} \hat{\boldsymbol{s}}}{\boldsymbol{s}^{\mathrm{T}} \boldsymbol{\Psi} \boldsymbol{s}} \tag{52}$$

Obviously, for a perfect model ($\hat{\boldsymbol{s}} = \boldsymbol{s}$) all the perfectness ratios are equal to 1. Now, by using the above definitions for ρ_1, ρ_2 and ρ_3, (49) can be expressed by

$$\Delta J_{\mathrm{ex}}(n) = \alpha J_{\mathrm{o}} - \beta \mathcal{J}(n) \tag{53}$$

where α and β are given by

$$\alpha = \mu^2 \rho_1 ||\boldsymbol{s}||^4 \frac{\sigma_x^4}{B_{\mathrm{w}}} L \tag{54}$$

$$\beta = \mu \rho_1 \frac{\sigma_x^2}{B_{\mathrm{w}}} \left[2\rho_2 - \mu \sigma_x^2 ||\boldsymbol{s}||^2 \left(L + \frac{2}{B_{\mathrm{w}}} \rho_2 \rho_3 D_{\mathrm{eq}}\right)\right] \tag{55}$$

Here, D_{eq} is defined as the equivalent delay of the secondary path

$$D_{\mathrm{eq}} = \frac{\boldsymbol{s}^{\mathrm{T}} \boldsymbol{\Psi} \boldsymbol{s}}{||\boldsymbol{s}||^2} \tag{56}$$

Now combining (25), (46) and (53) results in the following model for the FxLMS adaptation process

$$J_{\mathrm{ex}}(n+1) = \alpha J_{\mathrm{o}} + J_{\mathrm{ex}}(n) - \beta \sum_{p=0}^{Q-1} s_p^2 J_{\mathrm{ex}}(n-p) \tag{57}$$

Note that the above model considers a general secondary path, a general reference signal and an arbitrary secondary path model.

4 Formulating behaviours of FxLMS adaptation process

Based on the model developed in the previous section, a set of closed-form expressions for formulating behaviours of the FxLMS adaptation process is derived in the following.

4.1 Steady-state performance

Assuming a stable adaptation process, it can be shown that $J_{\mathrm{ex}}(n+1) = J_{\mathrm{ex}}(n) = J_{\mathrm{ex}}(n-p) = J_{\mathrm{ex}}(\infty)$ in steady-state conditions. In this case (57) results in

$$J_{\mathrm{ex}}(\infty) = \frac{\alpha}{\beta ||\boldsymbol{s}||^2} J_{\mathrm{o}} \tag{58}$$

Substituting (58) into (21), the MSE function level in steady-state conditions can be expressed as

$$J_{\mathrm{ss}} = J_{\mathrm{o}} + \frac{\alpha}{\beta ||\boldsymbol{s}||^2} J_{\mathrm{o}} \tag{59}$$

Now substituting (54) and (55) into (59) results in

$$J_{\mathrm{ss}} = J_{\mathrm{o}} + \frac{\mu \sigma_x^2 ||\boldsymbol{s}||^2 L}{2\rho_2 - \mu \sigma_x^2 ||\boldsymbol{s}||^2 \left(L + (2/B_{\mathrm{w}}) \rho_2 \rho_3 D_{\mathrm{eq}}\right)} J_{\mathrm{o}} \tag{60}$$

Equation (60) represents a closed-form expression for the steady-state MSE level in the FxLMS adaptation process, considering a general secondary path, a realistic reference signal and a general secondary path model. In fact, this equation is the most general closed-form expression which has been derived so far for formulating the steady-state MSE level in the FxLMS adaptation process with practical conditions.

4.2 Stability condition

As mentioned in Section 2, J_o is a minimal level for $J(n)$. Also from (22), it can be shown that $J(n)$ is positive definite. In this case, $J_{ex}(n)$, which is the difference between $J(n)$ and J_o, should be always positive. On the other hand, the stability of the FxLMS adaptation process requires $J_{ex}(n)$ to be finite in steady-state conditions. Considering these conditions and from (58), it can be deduced that the FxLMS adaptation process is stable only for $\beta > 0$ (other terms of (58) are always positive). Therefore a stability condition can be derived from (55) as

$$\mu\rho_1\frac{\sigma_x^2}{B_w}\left[2\rho_2 - \mu\sigma_x^2||\boldsymbol{s}||^2\left(L + \frac{2}{B_w}\rho_2\rho_3 D_{eq}\right)\right] > 0 \quad (61)$$

Satisfying this condition requires μ to be positive and smaller than an upper-bound, μ_{max}, given by

$$\mu_{max} = \frac{2\rho_2}{\sigma_x^2||\boldsymbol{s}||^2\left(L + (2/B_w)\rho_2\rho_3 D_{eq}\right)} \quad (62)$$

This closed-form expression for μ_{max} can apply to a general case, where the actual secondary path and secondary path model are two arbitrary and general systems, and the reference signal is a stochastic white signal with an arbitrary band-width.

4.3 Learning rate

According to (22), $J_{ex}(n)$ is a positive definite function of system variables. Hence, this function can be considered as a Lyapunov function. According to the Lyapunov stability theory, a system is stable if the time difference of its Lyapunov function (which is $\Delta J_{ex}(n)$ here) is negative and the process convergence rate (or learning rate) is directly related to the absolute value of the time difference of its Lyapunov function. Hence, the absolute value of $\Delta J_{ex}(n)$ represents the instantaneous learning rate of the FxLMS adaptation process. In transient conditions, the first term in (53) is not comparable with the second term. In this case, the absolute value of $\Delta J_{ex}(n)$ is directly related to β. Accordingly, the learning rate of the FxLMS adaptation process (ω) can be defined as $\omega = \beta$. From (55)

$$\omega = \mu\rho_1\frac{\sigma_x^2}{B_w}\left[2\rho_2 - \mu\sigma_x^2||\boldsymbol{s}||^2\left(L + \frac{2}{B_w}\rho_2\rho_3 D_{eq}\right)\right] \quad (63)$$

The step-size leading to the fastest learning rate (μ_f) can be obtained by maximising ω with respect to μ as

$$\left.\frac{d\omega}{d\mu}\right|_{\mu_f} = 0 \quad (64)$$

Combining (63) and (64) results in

$$\mu_f = 0.5\mu_{max} \quad (65)$$

It means that the FxLMS algorithm has its highest possible learning rate if the step-size is set to the half of its upper-bound. This result, which is obtained for a general case, is in a perfect agreement with the theoretical results obtained in [5, 8–10, 16] for some simplified cases.

4.4 Simplified cases

Equations (60) and (62) represent two general expressions for the steady-state MSE level (J_{ss}) and the step-size upper-bound (μ_{max}) of the FxLMS adaptation process, respectively. These expressions can apply to the case with a realistic secondary path, an arbitrary secondary path model, and a realistic reference signal. Other expressions for J_{ss} and μ_{max} can be found in the well-known references [8, 16, 17] or in the authors previous papers [9, 10, 12]. However, all of these previously derived expressions can be derived from (60) and (62) as special and simplified cases.

For example, the well-known expression derived for μ_{max} by Elliott and Elliott *et al.* in [16, 17] or the expression derived for J_{ss} by Bjarnason in [8] can be obtained from the proposed formulations by assuming a pure delay secondary path with the time delay of D samples ($D_{eq} = D$), a broad-band white reference signal ($B_w = 1$), and a perfect secondary path model ($\rho_1 = \rho_2 = \rho_3 = 1$). Also, the relatively more general expression derived by the authors in [9] can be obtained by setting $B_w = 1$ and $\rho_1 = \rho_2 = \rho_3 = 1$. The expression derived in [10], can be obtained by setting $\rho_1 = \rho_2 = \rho_3 = 1$ and the one derived in [12] can be obtained by setting $B_w = 1$.

The expression for μ_{max}, derived by the authors in [11], can also be obtained from (62) by setting $B_w = (1/L)$ and $\rho_1 = \rho_2 = \rho_3 = 1$. This expression corresponds to the case with a tonal reference signal, where there is only a single non-zero Eignevalue in the auto-correlation matrix of the input signal. The normalised bandwidth B_w is set to $(1/L)$ because there are L eigenvalues in total and each of them corresponds to the power of the signal in a certain frequency beam [18]. Accordingly, the proposed expressions for J_{ss} and μ_{max} can cover many of the previously derived expressions.

4.5 Effects of secondary path models on stability

According to (62), it is possible that the secondary path model causes μ_{max} to become negative. In this case, there is no positive μ for which the FxLMS algorithm becomes stable. Based on this logic, secondary path models causing the FxLMS algorithm to become unstable can be determined.

Usually, L is set to a large number, therefore it can be deduced that even if ρ_2 and ρ_2 have opposite signs, the denominator of (62) remains positive. Accordingly, the sign of μ_{max} can be determined only by its numerator. Based on this logic, μ_{max} is positive only when $\rho_2 > 0$. From (51), ρ_2 can be expressed as

$$\rho_2 = \frac{||\boldsymbol{s}||}{||\hat{\boldsymbol{s}}||}\cos\phi \quad (66)$$

where ϕ is the angle between $\boldsymbol{s}$ and $\hat{\boldsymbol{s}}$ in Q-dimensional space

$$\cos\phi = \frac{\boldsymbol{s}^T\hat{\boldsymbol{s}}}{||\boldsymbol{s}||||\hat{\boldsymbol{s}}||} \quad (67)$$

Since both $||\boldsymbol{s}||$ and $||\hat{\boldsymbol{s}}||$ are positive, it can be shown that for having $\rho_2 > 0$, the following inequality should hold

$$\cos\phi > 0 \quad (68)$$

As an elegant result, when the angle between the actual secondary path vector and its model vector is greater than

doi: 10.1049/iet-spr.2012.0090

90°, then the FxLMS algorithm cannot become stable. This result is in an excellent agreement with the 90° condition, derived by Boucher for pure delay secondary paths and tonal input signal [19] or by Morgan for an identity secondary path model [20]. For these simplified cases, in which only one element in vectors $\boldsymbol{s}$ and $\hat{\boldsymbol{s}}$ are non-zero, Boucher showed that if the secondary path model is 90° out-of-phase, then the FxLMS algorithm cannot become stable. However, in the derivation of (68), a general secondary path, its arbitrary model, and an arbitrary input signal bandwidth are considered. This generalisation is one of the contributions of this paper.

5 Simulation results

Fig. 2 shows the primary and secondary paths impulse responses of the simulated system. Also, this figure shows the impulse response of the imperfect secondary path model used in the computer simulation. The reference signal is a band-limited white signal of bandwidth $B_w = 0.8$ and power $\sigma_x^2 = 1$. Two simulation experiments for relatively small step-size of $\mu = 0.005$ and relatively large step-size of $\mu = 0.05$ are conducted. Each simulation experiment includes 100 simulation runs with independent reference signals. The variation of the square of the residual error, obtained from each run, is stored in the computer memory. The MSE function is then estimated by averaging over the stored data as

$$J(n) \simeq \frac{1}{100}\sum_{i=1}^{100} e_i^2(n) \tag{69}$$

where $e_i(n)$ is the residual acoustic noise, obtained from the ith simulation run. The excess-MSE function and its difference can be obtained from (21) and (25). For this purpose, an estimate of J_o should be computed by using (9) and (19), which, in turn, requires an estimate of $\boldsymbol{w}_o$ (Wiener-Hopf optimal solution). This optimal solution can be obtained from the statistical parameters of the reference signal and by using the standard methods given in the adaptive filter theory [1].

Based on the process discussed above, the variation of the $\Delta J_{ex}(n)$ can be calculated for each computer simulation

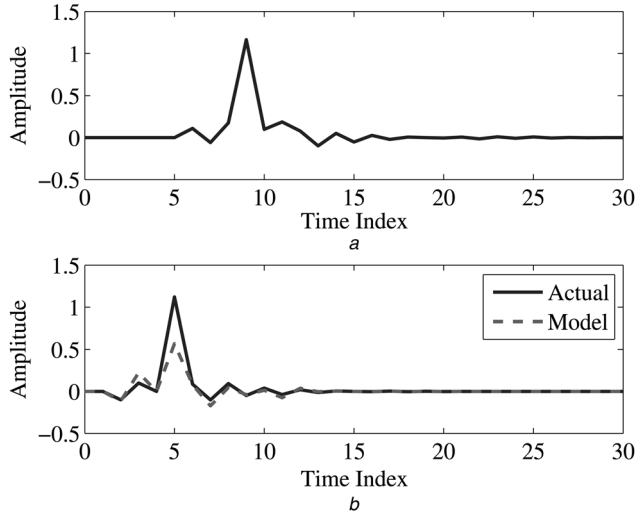


Fig. 2 *Impulse responses of primary path, secondary path and secondary path model in computer simulation*

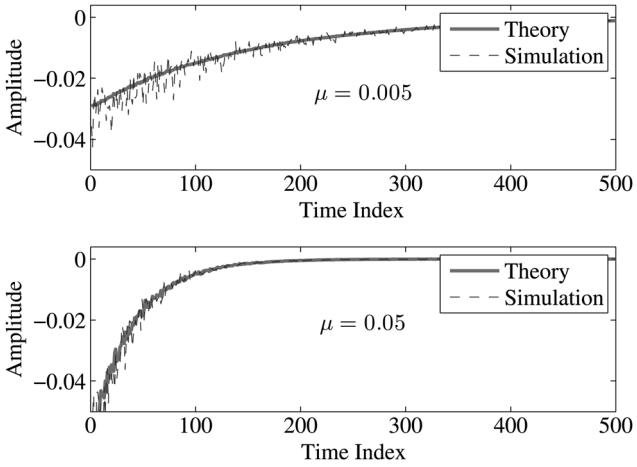


Fig. 3 *Variations of $J_{ex}(n)$ for different step-sizes and in different working conditions, red lines: theoretical results, blue lines: computer simulation*

experiments. The results can be plotted, as shown in Fig. 3. Also, the theoretical variation of $\Delta J_{ex}(n)$, obtained by using (53), is shown in this figure. The agreement between the theoretical and simulation results is evident.

The above simulation experiments can be repeated for different cases with different step-sizes, reference signal band-widths, and secondary path models. However, in all of the cases, the proposed theoretical model can efficiently describe system behaviours. In fact, the agreement between the theoretical and simulation results takes away the ambiguity of the independence assumptions used in the derivation of the theoretical model . The verification of this model is important at this stage because this model is used in Section 4 to derive theoretical expressions for formulating behaviours of the FxLMS adaptation process. Verifying these expressions using practical results is left to the next section. Note that, this verification can be also shown by using simulation results; however, in order to use the available space efficiently, and to include the experimental results into the paper, this is shown through experimental results.

6 Experimental results with an FxLMS-based active noise control system

Active noise control (ANC) is one of the most well-known applications of the FxLMS algorithm. In ANC, an adaptive control system is responsible for the estimation of a control signal for feeding a loudspeaker. The loudspeaker produces an anti-noise (in acoustic domain) which has to propagate in an unwanted signal channel (secondary path). If the anti-noise signal is estimated and produced properly, its combination with the environmental noise at a desired zone of silence becomes minimal. Note that the anti-noise has to travel across the secondary path to reach the desired zone of silence. The adaptive controller, which produces the control signal, is adjusted by the FxLMS algorithm. Hence, a reference signal $x(n)$ and an error signal $e(n)$ should be provided for the FxLMS algorithm. These signals are picked up by using two microphones, called the reference and error microphones, respectively. In ANC terminology, the reference and error signals are referred to as the reference and residual noise, respectively. More information about ANC theory can be found in [21].

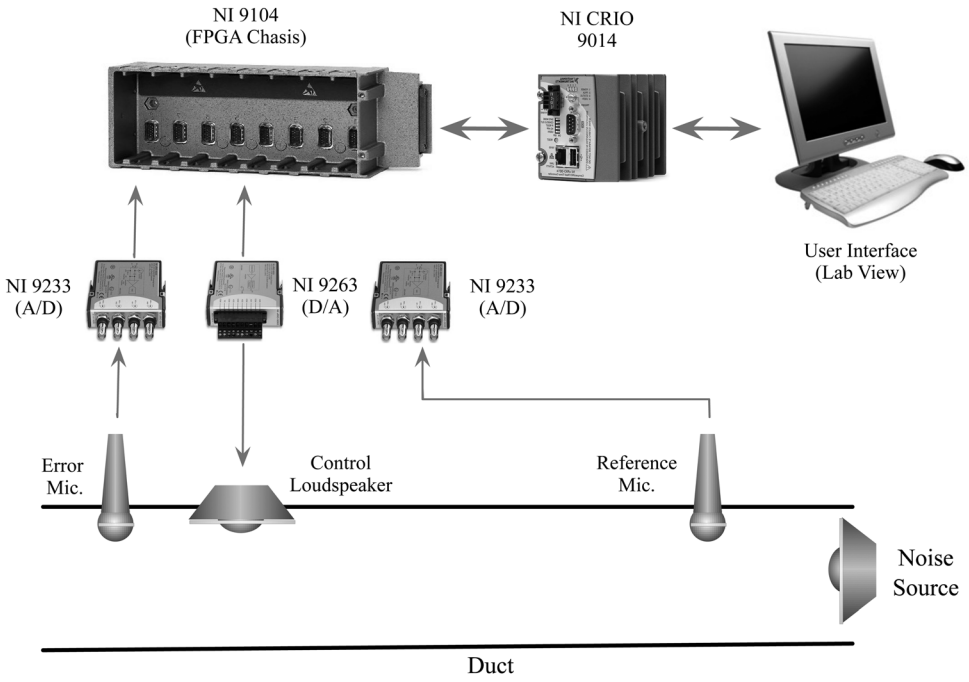


Fig. 4 *Schematic diagram of experimental ANC setup*

Fig. 4 shows the schematic diagram of the implemented ANC setup. This setup is developed to create a silence-zone at the end of an acoustic duct. The length, width and height of this duct are 150, 30 and 25 cm, respectively. The main hardware component used in this setup is a Compact Reconfigurable Input/Output (CRIO) 9014 which is a Field Programmable Gate Array (FPGA)-based real-time embedded controller made by National Instrument Company (NI). More information about this setup can be found in [10].

Using the real-time secondary path identification system, integrated in the experimental setup, the secondary path impulse response of the acoustic duct can be estimated precisely. After estimating this impulse response, its coefficients are downloaded from the CRIO memory into LabVIEW environment, where they can be exported to an ASCII text file. This file is then used by a MATLAB function to compute $\|\boldsymbol{s}\|^2$ and D_{eq} from (56): $\|\boldsymbol{s}\|^2 = 0.4684$ and $D_{\text{eq}} = 42.39$. The acoustic noise, injected into the duct, is a band-limited white signal with $B_{\text{w}} = 0.18$ and $\sigma_x^2 = 6\,\text{W}$. The imperfect secondary path model, shown in Fig. 5 is uploaded into the CRIO memory to be used by the FxLMS algorithm. The perfectness ratios of this model are $\rho_1 = 0.3597$, $\rho_2 = 0.9135$ and $\rho_3 = 0.2009$. In this situation,

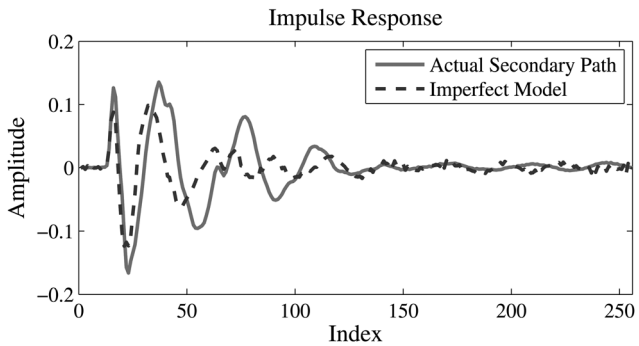


Fig. 5 *Secondary path impulse response and its imperfect model in experiments*

several experiments with different values of μ are conducted. For each experiment, the output of the error microphone (error signal) is monitored, recorded and analysed in LabVIEW. The obtained results can be then compared with the theoretical results.

Different experiments show that when $\mu < 0.0017$, the implemented ANC system is always stable and when $0.0017 < \mu < 0.0020$ the system starts diverging after a short time; and when $\mu > 0.0020$ the system starts diverging from the beginning. Hence, a practical $\mu_{\max}$ is located between 0.0017 and 0.0020. However, it is not technically possible to estimate any specific value for $\mu_{\max}$ form the experimental results. This is because of experimental conditions uncertainties, for example, changing characteristics of the surrounding environment, non-stationary behaviours of the background noise and uncertainties associated with physical plants, control systems and measurement devices. Now, by using (62), the theoretical value of $\mu_{\max}$ can be obtained as $\mu_{\max} = 0.0019$. It means that, in theory, for $\mu < 0.0019$ the system always becomes stable and for $\mu > 0.0019$, the system always diverges. This result is in a good agreement with the experimental results described above.

Now, μ is set to a relatively small number (about 0.1 of its upper-bound). After a long time, when the system reaches its steady-state conditions, the power of the residual noise is computed in LabVIEW. The step-size μ is incrementally increased and the system is restarted again. For each μ, the above experiment is repeated and the steady-state power of the residual noise is measured. The measured data can be plotted as a function of μ, as shown in Fig. 6. For each μ, the theoretical value of the steady-state residual acoustic noise (J_{ss}) can be also computed by using the closed-form expression given in (60). The results is plotted as another function of μ in Fig. 6. The agreement between the experimental and theoretical results is apparent in this figure.

Finally, it is desired to investigate the validity of the expression given for the learning rate ω. For this purpose, the transient convergence speed of the implemented ANC

doi: 10.1049/iet-spr.2012.0090

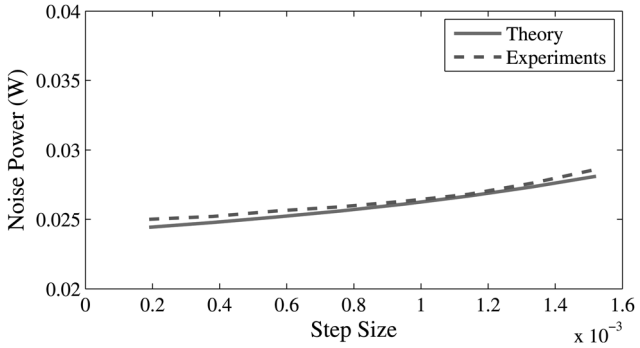


Fig. 6 *Steady-state residual acoustic noise power in the implemented ANC setup*

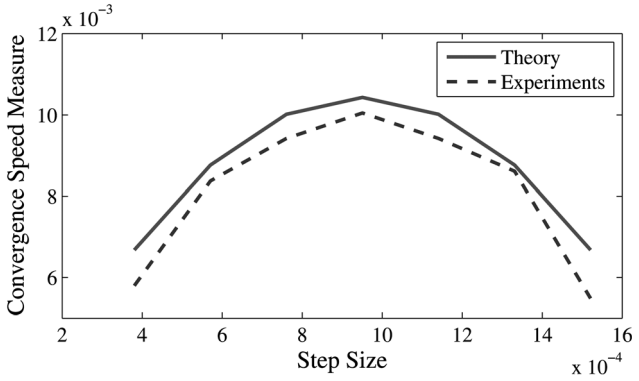


Fig. 7 *Learning rate of the implemented ANC setup*

system should be evaluated. This measure can be interpreted as the slop of the residual noise power once the FxLMS algorithm begins operating. Therefore in order to estimate this slop in practice, the average speed on which the residual noise power reduces for -6 dB is measured. This speed can be evaluated by using the following formulation

$$\omega \simeq \frac{1.5071}{N_{6\,\mathrm{dB}}} \frac{W}{\mathrm{Sample}} \tag{70}$$

where 1.5071 is equivalent to -6 dB reduction in the acoustic power when the reference power is $\sigma_x^2 = 6$ W, and $N_{6\,\mathrm{dB}}$ is the time index at which the steady-state residual noise power is attenuated for 6 dB. For each experiment with a particular value of μ, $N_{6\,\mathrm{dB}}$ can be measured and, then, it can be recorded in LabVIEW. Finally, a practical ω for each case can be evaluated from (70) by using the recorded data. The obtained results is plotted as a function of μ, as shown in Fig. 6. Also, for each experiment, the theoretical value of ω is computed by using (63) and plotted in Fig. 7. The agreement between the experimental and theoretical results is apparent in this figure. Also, in both the experimental and theoretical curves shown in this figure, it can be seen that the highest learning rate (or convergence speed) occurs when μ is set to $0.5\mu_{\max}$. This is in a perfect agreement with the theoretical results obtained in Section 6.4.

7 Conclusions

Modelling and analysis of the FxLMS adaptation process is an interesting research subject associated with a high level of mathematical complexity. There have been several contributions made by different researchers on this subject; however, only a few have intended to derive general analytical formulations for modelling this process in practical conditions. Even if they intended to do so, they had to simplify their analytical model by using unrealistic assumptions. A relatively comprehensive analysis on the FxLMS adaptation process in practical conditions is performed in this paper. The theoretical findings form a set of closed-form expressions for formulating the steady-state performance, stability condition and learning rate of the FxLMS adaptation process.. These expressions are found to be in an excellent agreement with simulation and experimental results. Also, all the previously derived expressions can be derived from these expressions as special and simplified cases.

8 Acknowledgment

The authors would like to acknowledge The University of Auckland for supporting this research by a FRDF grant; Project number: 3700501.

9 References

1 Widrow, B., Stearns, S.D.: 'Adaptive signal processing' (Prentice-Hall, 1985)
2 Widrow, B., Shur, D., Shaffer, S.: 'On adaptive inverse control'. Proc. 15th Asilomar Conf. on Circuits, Systems and Computers, November 1981, pp. 185–189
3 Resende, L., Bermudez, J.: 'Statistical analysis of the FxLMS algorithm about the steady-state solution'. Telecommunications Symp., 2006 Int., September 2006, pp. 306–309
4 Tobias, O., Bermudez, J., Bershad, N.: 'Mean weight behavior of the filtered-x lms algorithm', *Signal Process., IEEE Trans.*, 2000, **48**, (4), pp. 1061–1075
5 Long, G., Ling, F., Proakis, J.: 'The LMS algorithm with delayed coefficient adaptation', *Acoust., Speech Signal Process., IEEE Trans.*, 1989, **37**, (9), pp. 1397–1405
6 Long, G., Ling, F., Proakis, J.: 'Corrections to 'The LMS algorithm with delayed coefficient adaptation', *Signal Process., IEEE Trans.*, 1992, **40**, (1), pp. 230–232
7 Elliott, S.J., Nelson, P.A.: 'Multiple-point equalization in a room using adaptive digital filters', *J. Audio Eng. Soc.*, 1989, **37**, (11), pp. 899–907
8 Bjarnason, E.: 'Analysis of the filtered-x LMS algorithm', *Speech Audio Process., IEEE Trans.*, 1995, **3**, (6), pp. 504–514
9 Tabatabaei Ardekani, I., Abdulla, W.H.: 'Theoretical convergence analysis of FxLMS algorithm', *Signal Process.*, 2010, **90**, (12), pp. 3046–3055
10 Tabatabaei Ardekani, I., Abdulla, W.H.: 'On the convergence of real-time active noise control systems', *Signal Process.*, 2011, **91**, (5), pp. 1262–1274
11 Tabatabaei Ardekani, I., Abdulla, W.H.: 'On the stability of adaptation process in active noise control systems', *J. Acoust. Soc. Am.*, 2011, **129**, (1), pp. 173–184
12 Tabatabaei Ardekani, I., Abdulla, W.H.: 'Effects of imperfect secondary path modeling on adaptive active noise control systems', *Control Syst. Technol., IEEE Trans.*, 2012, **20**, (5), pp. 1252–1262
13 Vicente, L., Masgrau, E.: 'Novel FxLMS convergence condition with deterministic reference', *Signal Process., IEEE Trans.*, 2006, **54**, (10), pp. 3768–3774
14 Takenouchi, Y., Sizuhi, H., Omoto, A.: 'Behavior of the practically implemented filtered reference LMS algorithm in an active noise control system', *J. Acoust. Sci. Technol.*, 2006, **27**, pp. 20–27
15 Gardner, W.A.: 'Learning characteristics of stochastic gradient descent algorithm: a general study, analysis and critique', *Signal Process.*, 1984, **6**, pp. 113–133
16 Elliott, S.J.: 'Signal processing for active control' (Academic Press, San Diego, CA, 2001)
17 Elliott, S., Stothers, I.M., Neslson, P.A.: 'A multiple error LMS algorithm and its applications to active control of sound and vibration', *Acoust., Speech Signal Process., IEEE Trans.*, 1987, **35**, pp. 1423–1434

18 Thomson, D.: 'Spectrum estimation and harmonic analysis', *Proc. IEEE*, 1982, **70**, (9), pp. 1055–1096
19 Boucher, C., Elliott, S., Nelson, P.: 'Effect of errors in the plant model on the performance of algorithms for adaptive feedforward control', *Radar Signal Process., IEE Proc. F*, 1991, **138**, (4), pp. 313–319
20 Morgan, D.R.: 'An analysis of multiple correlation cancellation loops with a filter in the auxiliary path'. Acoustics, Speech, and Signal Processing, IEEE Int. Conf. ICASSP'80., April 1980, pp. 457–461
21 Kajikawa, Y., Gan, W.-S., Kuo, S.M.: 'Recent advances on active noise control: open issues and innovative applications', *APSIPA Trans. Signal Inf. Process.*, 2012, **1**, pp. 1–21

10 Appendix

10.1 Appendix 1: Derivation of (34)

$J(n)$ is independent of the stochastic process $\boldsymbol{x}(n)$ [12]. Accordingly, $\hat{\boldsymbol{f}}(n)$ is independent of $J(n)$. In this case, (32) can be simplified to

$$A_l(n) = \mu^2 \boldsymbol{F}_l^{\mathrm{T}} E\{\hat{\boldsymbol{f}}(n)\hat{\boldsymbol{f}}^{\mathrm{T}}(n)\}\boldsymbol{F}_l J(n) \tag{71}$$

On the other hand, from (8) and (18), it can be shown that

$$E\{\hat{\boldsymbol{f}}(n)\hat{\boldsymbol{f}}^{\mathrm{T}}(n)\} = \|\hat{\boldsymbol{s}}\|^2 \boldsymbol{R} \tag{72}$$

where $\|.\|$ denotes the Euclidean vector norm. Now, by substituting (72) into (71), $A_l(n)$ is simplified to

$$A_l(n) = \mu^2 \|\hat{\boldsymbol{s}}\|^2 \boldsymbol{F}_l^{\mathrm{T}} \boldsymbol{R} \boldsymbol{F}_l J(n) \tag{73}$$

From (10) and considering $\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F} = \boldsymbol{I}$, it can be shown that $\boldsymbol{F}_l^{\mathrm{T}} \boldsymbol{R} \boldsymbol{F}_l = \lambda_l$; therefore (73) can be simplified to

$$A_l(n) = \mu^2 \lambda_l \|\hat{\boldsymbol{s}}\|^2 J(n) \tag{74}$$

Now, substituting (21) into (74) results in

$$A_l(n) = \mu^2 \lambda_l \|\hat{\boldsymbol{s}}\|^2 J_{\mathrm{o}} + \mu^2 \lambda_l \|\hat{\boldsymbol{s}}\|^2 J_{\mathrm{ex}}(n) \tag{75}$$

Finally, combining (23) and (75) results in (34).

10.2 Appendix 2: Derivation of (35)

By substituting (15) into (33), $B_l(n)$ is expanded to

$$\begin{aligned} B_l(n) = {} & 2\mu \boldsymbol{F}_l^{\mathrm{T}} E\{\hat{\boldsymbol{f}}(n) c_l(n) e_{\mathrm{o}}(n)\} \\ & - 2\mu \boldsymbol{F}_l^{\mathrm{T}} \sum_{p=0}^{Q-1} s_p E\{c_l(n)\hat{\boldsymbol{f}}(n)\boldsymbol{c}^{\mathrm{T}}(n-p)\boldsymbol{z}(n-p)\} \end{aligned} \tag{76}$$

where c_l denotes the lth element of $\boldsymbol{c}$. Considering the independence of $e_{\mathrm{o}}(n)$ from the adaptive weights and input signal, it can be shown that the first term in (76) is zero. Also, the second term can be expanded by the expression given for $\hat{\boldsymbol{f}}(n)$ in (8); therefore

$$\begin{aligned} B_l(n) = {} & -2\mu \boldsymbol{F}_l^{\mathrm{T}} \\ & \times \sum_{p,m=0}^{Q-1} s_p \hat{s}_m E\{c_l(n)\boldsymbol{x}(n-m)\boldsymbol{c}^{\mathrm{T}}(n-p)\boldsymbol{z}(n-p)\} \end{aligned} \tag{77}$$

Note that, according to (7), for $m \geq M$, $\hat{s}_m = 0$. Substituting (13) and (14) into (77) results in

$$\begin{aligned} B_l(n) = {} & -2\mu \boldsymbol{F}_l^{\mathrm{T}} \\ & \times \sum_{p,m=0}^{Q-1} \sum_{i=0}^{L-1} s_p \hat{s}_m E\{c_l(n) c_i(n-p)\boldsymbol{x}(n-m) z_i(n-p)\} \end{aligned} \tag{78}$$

where z_i denotes the ith element of $\boldsymbol{z}$. (13) results in $z_i(n-p) = \boldsymbol{x}^{\mathrm{T}}(n-p)\boldsymbol{F}_i$; therefore (78) is modified to

$$\begin{aligned} B_l(n) = {} & -2\mu \boldsymbol{F}_l^{\mathrm{T}} \\ & \times \sum_{p,m=0}^{Q-1} \sum_{i=0}^{L-1} s_p \hat{s}_m E\{c_l(n) c_i(n-p)\boldsymbol{x}(n-m)\boldsymbol{x}^{\mathrm{T}}(n-p)\}\boldsymbol{F}_i \end{aligned} \tag{79}$$

Since $\boldsymbol{x}(n)$ and $\boldsymbol{c}(n)$ are independent, (79) results in

$$\begin{aligned} B_l(n) = {} & -2\mu \boldsymbol{F}_l^{\mathrm{T}} \sum_{p,m=0}^{Q-1} \sum_{i=0}^{L-1} s_p \hat{s}_m E\{c_l(n) c_i(n-p)\} \\ & \times E\{\boldsymbol{x}(n-m)\boldsymbol{x}^{\mathrm{T}}(n-p)\}\boldsymbol{F}_i \end{aligned} \tag{80}$$

Now, substituting (18) into (80) results in

$$B_l(n) = -2\mu \sum_{p=0}^{Q-1} \sum_{i=0}^{L-1} s_p \hat{s}_p E\{c_l(n) c_i(n-p)\}\boldsymbol{F}_l^{\mathrm{T}} \boldsymbol{R} \boldsymbol{F}_i \tag{81}$$

From (10) and considering $\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F} = \boldsymbol{I}$, it can be shown that $\boldsymbol{F}_l^{\mathrm{T}} \boldsymbol{R} \boldsymbol{F}_i = \lambda_l \delta_{l,i}$; therefore (81) can be simplified to

$$B_l(n) = -2\mu\lambda_l \sum_{p=0}^{Q-1} s_p \hat{s}_p E\{c_l(n) c_l(n-p)\} \tag{82}$$

From the FxLMS update equation, given in (16), it can be shown that

$$\boldsymbol{c}(n) = \boldsymbol{c}(n-p) + \mu \sum_{k=1}^{p} \hat{\boldsymbol{g}}(n-k) e(n-k) \tag{83}$$

When the adaptation process is slow, (83) can be approximated by

$$\boldsymbol{c}(n) \simeq \boldsymbol{c}(n-p) + \mu p \hat{\boldsymbol{g}}(n-p) e(n-p) \tag{84}$$

Therefore the following equation can be derived for the variation of the lth adaptive weight

$$c_l(n) \simeq c_l(n-p) + \mu p \hat{g}_l(n-p) e(n-p) \tag{85}$$

doi: 10.1049/iet-spr.2012.0090

Now, combining (82) and (85) results in

$$B_l(n) = -2\mu\lambda_l \sum_{p=0}^{Q-1} s_p \hat{s}_p E\{c_l^2(n-p)\} - 2\mu^2\lambda_l \sum_{p=0}^{Q-1} p s_p \hat{s}_p E\{c_l(n-p)\hat{g}_l(n-p)e(n-p)\} \quad (86)$$

By using (24), (30) and (33), (86) is modified to

$$B_l(n) = -2\mu\lambda_l \sum_{p=0}^{Q-1} s_p \hat{s}_p m_l(n-p) - \mu\lambda_l \sum_{r=0}^{Q-1} r s_r \hat{s}_r B_l(n-r) \quad (87)$$

Equation (87), can be expanded to

$$B_l(n) = -2\mu\lambda_l \sum_{p=0}^{Q-1} s_p \hat{s}_p m_l(n-p) + 2\mu^2\lambda_l^2 \sum_{p=0}^{Q-1} s_p \hat{s}_p \sum_{r=0}^{Q-1} r s_r \hat{s}_r m_l(n-p-r) + \cdots \quad (88)$$

For $\mu \ll 1$, (88) can be approximated by its first two terms as shown in (35).