STATISTICAL-MECHANICAL ANALYSIS OF THE FXLMS ALGORITHM WITH ACTUAL PRIMARY PATH

Seiji Miyoshi and Yoshinobu Kajikawa

Faculty of Engineering Science, Kansai University
3-3-35 Yamate-cho, Suita-shi, Osaka, 564-8680 Japan

ABSTRACT

A theory that predicts the behaviors of the Filtered-X LMS algorithm was derived by using a statistical-mechanical method. In this paper, the theory is generalized to explain the system behaviors in the case of an actual primary path. In the theory, cross-correlations between the element of a primary path and that of an adaptive filter and autocorrelations of the elements of the adaptive filter are treated as macroscopic variables. Simultaneous differential equations that describe the dynamical behaviors of the macroscopic variables are obtained under conditions in which the tapped-delay line is sufficiently long. The equations are analytically solved to obtain the correlations and finally compute the mean-square error. In order to generalize the theory to the case of an actual primary path, the correlations of the elements of the primary path are absorbed. The generalized theory quantitatively predict the behaviors in the case of an actual primary path.

Index Terms— Filtered-X LMS algorithm, adaptive filter, active noise control, statistical-mechanical method, actual primary path

1. INTRODUCTION

In recent years, active noise control (ANC) has been practically realized and applied to various fields[1, 2, 3]. ANC is divided into two types, feedforward and feedback ANC[3]. The feedforward ANC is considered in this paper.

The Filtered-X LMS (FXLMS) algorithm, which is the generalized procedure of the least-mean-square (LMS) algorithm[4, 5, 6, 7] considering the impulse response of the secondary path, is commonly used algorithm for the feedforward ANC[8].

Various methods have been proposed to theoretically analyze the FXLMS algorithm. The principal method is to use the independence assumption[9, 10, 11, 12, 13, 14, 15]. In the independence assumption, successive tap input vectors of the tapped-delay line are assumed to be independently generated at each time step. However, the actual elements of the tap input vector are merely shifted to the next position. Hence, each tap input vector is related to the previous one and the vectors are thus, not independent. Owing to this fact, analyses based on the independence assumption involve essential and potential problems[5].

There are various methods based on assumptions other than the independence assumption. In [16, 17, 18], another form of independence is assumed. That is, the correlation between the tap input vectors is assumed to be more dominant than the correlation between the weight vector of the adaptive filter and the tap input vectors. However, analytical results based on such assumptions cannot precisely explain experimental results, particularly when there is little or no background noise. In [15, 19, 20, 21], the step size is assumed to be small. In [22, 23], it is assumed that the reference signal is sinusoidal. In [24], it is assumed that both the unknown system and the adaptive filter have a small number of taps. Thus, a general theory for the FXLMS algorithm has not been given in the literature even though this algorithm is widely used.

In [25, 26], the dynamical and steady-state properties of the FXLMS algorithm were theoretically analyzed by applying a statistical-mechanical method. Although the theory does not use the conditions assumed in the previous studies described above, it was assumed that the elements of the primary path are independently generated in the theory. However, the actual elements have strong correlations. Therefore in this paper, the theory is generalized to apply it to the case of an actual primary path by absorbing the correlations of the elements.

2. ANALYTICAL MODEL OF FXLMS ALGORITHM

Figure 1 shows a block diagram of the ANC system considered in this paper. The primary path P is represented by an N_p-tap FIR filter. Its coefficient vector is $\boldsymbol{p} = [p_1, p_2, \ldots, p_{N_p}]^T$. Each coefficient p_i is generated from the stochastic process expressed as

$$\langle p_i \rangle = 0, \qquad \langle p_i p_{i-j} \rangle = \kappa_j, \tag{1}$$

and is time-invariant. Here, $\langle \cdot \rangle$ denotes expectation. If $\kappa_j = 0$ when $j \neq 0$, the elements of the impulse response

This work was partially supported by JSPS KAKENHI Grant Numbers 24360152, 24560280, and the Nonverbal Communication Technology Research Group of Organization for Research and Development of Innovative Science and Technology, Kansai University.

of the primary path are uncorrelated with each other. In this paper, the autocorrelated primary path is considered by introducing the covariance function κ. The adaptive filter H is an N_h-tap FIR filter. Its coefficient vector is $\boldsymbol{h}(n) = [h_1(n), h_2(n), \ldots, h_{N_h}(n)]^T$, where n denotes the time step. The reference signal $x(n)$ is drawn from a distribution with

$$\langle x(n) \rangle = 0, \qquad \langle x(n)x(n-k) \rangle = r_k/N_h. \tag{2}$$

The correlation function (2) implies that the reference signal is white if $r_k = 0$ ($k \neq 0$) and that the model includes the case of nonwhite reference signals. The reference signal is shifted through the tapped-delay line. Therefore, the tap input vectors of the primary path and adaptive filter are $\boldsymbol{x}_p(n) = [x(n), x(n-1), \ldots, x(n-N_p+1)]^T$ and $\boldsymbol{x}_h(n) = [x(n), x(n-1), \ldots, x(n-N_h+1)]^T$, respectively. The output of the primary path P is $d(n) = \boldsymbol{p}^T\boldsymbol{x}_p(n)$. On the other hand, the output of the adaptive filter H is $u(n) = \boldsymbol{h}(n)^T\boldsymbol{x}_h(n)$.

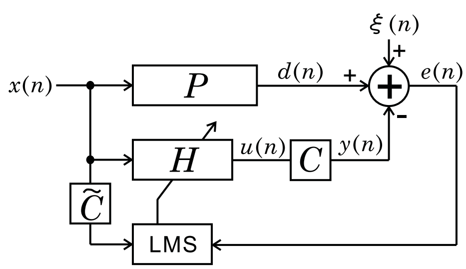

Fig. 1. Block diagram of ANC system.

The secondary path C is modeled by a K-tap FIR filter. Its coefficient vector is $\boldsymbol{c} = [c_1, c_2, \ldots, c_K]^T$ and is time-invariant. The output $y(n)$ of the secondary path is

$$y(n) = \sum_{k=1}^{K} c_k u(n-k+1). \tag{3}$$

The error signal $e(n)$ is generated by adding an independent background noise $\xi(n)$ to the difference between $d(n)$ and $y(n)$. That is,

$$e(n) = d(n) - y(n) + \xi(n). \tag{4}$$

Here, the mean and variance of $\xi(n)$ are zero and σ_ξ^2, respectively.

In the FXLMS algorithm, the estimated secondary path $\tilde{C}$ is used which has been estimated in advance by a certain method, since the true secondary path is generally unknown. When the estimated secondary path $\tilde{C}$ is a K-tap FIR filter and its coefficient vector is $\tilde{\boldsymbol{c}} = [\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_K]^T$, the update procedure obtained by the FXLMS algorithm is

$$\boldsymbol{h}(n+1) = \boldsymbol{h}(n) + \mu e(n) \sum_{k=1}^{K} \tilde{c}_k \boldsymbol{x}_h(n-k+1), \tag{5}$$

where μ is the step size.

3. THEORY

From (3) and (4), the MSE is expressed as

$$\begin{aligned}\langle e^2(n) \rangle &= \langle d^2(n) \rangle \\ &+ \sum_{k=1}^{K}\sum_{k'=1}^{K} c_k c_{k'} \langle u(n-k+1)u(n-k'+1) \rangle \\ &- 2\sum_{k=1}^{K} c_k \langle d(n)u(n-k+1) \rangle + \sigma_\xi^2. \end{aligned} \tag{6}$$

Equation (6) includes many products of u and u and products of d and u, including cases where their time steps are different. To calculate these products, the N_h-dimensional vectors

$$\boldsymbol{k}_j(n) = [k_{j,1}(n), k_{j,2}(n), \ldots, k_{j,N_h}(n)]^T, \tag{7}$$

are introduced where $j = -M, \ldots, M$ and $k_{j,i}(n) = h_{\mathrm{mod}(i+j-1, N_h)+1}(n)$. That is, $\boldsymbol{k}_j(n)$ is the j-shifted vector of the coefficient vector $\boldsymbol{h}(n)$ of the adaptive filter. Note that $\boldsymbol{k}_0(n) = \boldsymbol{h}(n)$.

In the following, the limit $N_p, N_h \to \infty$ is considered. Here, $a = N_p/N_h$ is kept constant. When the shift number j is $O(1)$, we can obtain

$$\boldsymbol{h}(n)^T\boldsymbol{x}_h(n) = \boldsymbol{k}_j(n)^T\boldsymbol{x}_h(n-j). \tag{8}$$

Equation (8) is based on the fact that the shift of the tap input vector is canceled by the shift of the elements of the adaptive filter. Here, the effect of the edge of the adaptive filter can be ignored since both $\boldsymbol{h}(n)$ and $\boldsymbol{k}_j(n)$ are N_h-dimensional, i.e., infinitely long, vectors. Equation (8) implies that the gap j in the time direction can be replaced by the subscript of the vector $\boldsymbol{k}$. In addition, we introduce two macroscopic variables defined by $R_j(n) \equiv \frac{1}{\bar{a}N_h}\sum_{i=1}^{\bar{a}N_h} p_i k_{j,i}(n)$ and $Q_j(n) \equiv \frac{1}{N_h}\sum_{i=1}^{N_h} h_i(n)k_{j,i}(n)$. $R_j(n)$ and $Q_j(n)$ are the cross-correlation between $\boldsymbol{p}$ and $\boldsymbol{h}(n)$ and the autocorrelation of $\boldsymbol{h}(n)$, respectively. Here, $\bar{a} = \min(a, 1)$.

Then, $\langle d(n-j)u(n) \rangle = \bar{a}\sum_{i=-M}^{M} R_i r_{i-j}$, $\langle u(n-j)u(n) \rangle = \sum_{i=-M}^{M} Q_i r_{i-j}$, and $\langle d(n-j)d(n) \rangle = a\sum_{i=-L}^{L} \kappa_i r_{i-j}$ are obtained. Here, the time steps of the macroscopic variables have been omitted, since they do not change by $O(1)$ in the $O(1)$ time updates in the model considered in this paper. The MSE (6) can be expressed in terms of the cross-correlation R_j and autocorrelation Q_j as

$$\begin{aligned}\langle e(n)^2 \rangle = &\sum_{k=1}^{K} c_k \sum_{i=-M}^{M} \left(\sum_{k'=1}^{K} c_{k'} Q_i r_{i-k+k'} - 2aR_i r_{i+k-1} \right) \\ &+ ar_0 + \sigma_\xi^2. \end{aligned} \tag{9}$$

This formula shows that the MSE is a function of the macroscopic variables R and Q. Therefore, we derive differential equations that describe the dynamical behaviors of these variables in the following.

$$
\begin{aligned}
\frac{dQ_j}{dt} =& \mu \sum_{k'=1}^{K} \tilde{c}_{k'} \Bigg\{ \sum_{i=-M}^{M} \left[\bar{a} R_i \left(r_{i-\gamma} + r_{i-\epsilon} \right) - \sum_{k=1}^{K} c_k Q_i \left(r_{i-k'+k+j} + r_{i-k'+k-j} \right) \right] \\
& -\mu \Bigg[\mathrm{sgn}(\gamma) \sum_{k''=1}^{|\gamma|} \Bigg(\delta_{\alpha,0} \sigma_\xi^2 + a \sum_{i=-L}^{L} \kappa_i r_{i+\alpha} - \sum_{k=1}^{K} c_k \sum_{i=-M}^{M} \Bigg(\bar{a} R_i r_{i+k-1-\alpha} + \bar{a} R_i r_{i+k-1+\alpha} - \sum_{k'''=1}^{K} c_{k'''} Q_i r_{i-k+k'''-\alpha} \Bigg) \Bigg) \sum_{i=1}^{K} \tilde{c}_i r_{k'-i-j+\alpha} \\
& + \mathrm{sgn}(\epsilon) \sum_{k''=1}^{|\epsilon|} \Bigg(\delta_{\beta,0} \sigma_\xi^2 + a \sum_{i=-L}^{L} \kappa_i r_{i+\beta} - \sum_{k=1}^{K} c_k \sum_{i=-M}^{M} \Bigg(\bar{a} R_i r_{i+k-1-\beta} + \bar{a} R_i r_{i+k-1+\beta} - \sum_{k'''=1}^{K} c_{k'''} Q_i r_{i-k+k'''-\beta} \Bigg) \Bigg) \sum_{i=1}^{K} \tilde{c}_i r_{k'-i+j+\beta} \Bigg] \Bigg\} \\
& + \mu^2 \left[\sum_{k=1}^{K} c_k \sum_{i=-M}^{M} \left(\sum_{k'=1}^{K} c_{k'} Q_i r_{i-k+k'} - 2 \bar{a} R_i r_{i+k-1} \right) + a \sum_{i=-L}^{L} \kappa_i r_i + \sigma_\xi^2 \right] \sum_{k''=1}^{K} \sum_{k'''=1}^{K} \tilde{c}_{k''} \tilde{c}_{k'''} r_{k''-k'''-j},
\end{aligned}
\tag{10}
$$

First, a differential equation for R_j is derived. When the coefficient vector $\boldsymbol{h}$ of the adaptive filter is updated, the j-shifted vector $\boldsymbol{k}_j$ is also changed. This change can be described as

$$
\boldsymbol{k}_j(n+1) = \boldsymbol{k}_j(n) + \mu e(n) \sum_{k=1}^{K} \tilde{c}_k \boldsymbol{x}_h(n-k+1-j). \tag{11}
$$

Multiplying both sides of (11) on the left by the N_h-dimensional vector $\bar{\boldsymbol{p}} = [p_1, p_2, \ldots, p_{\bar{a}N_h}, \underbrace{0, \ldots, 0}_{(1-\bar{a})N_h}]^T$, then

$$
\begin{aligned}
\bar{a} N_h R_j(n+1) =& \bar{a} N_h R_j(n) \\
& + \mu e(n) \sum_{k=1}^{K} \tilde{c}_k \bar{d}(n-k+1-j)
\end{aligned}
\tag{12}
$$

is obtained. Here, $\bar{d}(n) = \bar{\boldsymbol{p}}^T \boldsymbol{x}_h(n)$. Then, $\left\langle \bar{d}(n-j) u(n) \right\rangle = \langle d(n-j) u(n) \rangle$ and $\left\langle \bar{d}(n-j) d(n) \right\rangle = \frac{\bar{a}}{a} \langle d(n-j) d(n) \rangle$ are obtained.

If the adaptive filter is updated $N_h dt$ times in an infinitely small time dt, we can obtain $N_h dt$ equations that are similar to (12). Here, the continuous time t is defined by the time step m normalized by the tap length N_h [25, 26]. Summing all these equations, we obtain a differential equation that describes the dynamical behavior of R_j in a deterministic form as follows:

$$
\frac{dR_j}{dt} = \mu \sum_{k'=1}^{K} \tilde{c}_{k'} \left(\sum_{i=-L}^{L} \kappa_i r_{i+k'+j-1} - \sum_{k=1}^{K} c_k \sum_{i=-M}^{M} R_i r_{i+k-k'-j} \right). \tag{13}
$$

Next, multiplying (5) by (11) and proceeding in the same manner as for the derivation of the above differential equation for R_j, we can derive a differential equation for Q_j, which is given by (10), where sgn$(\cdot)$ and $\Theta(\cdot)$ are the sign and step functions, respectively. In addition, $\alpha \equiv \Theta(\gamma)\gamma - k''$, $\beta \equiv \Theta(\epsilon)\epsilon - k''$, $\gamma \equiv j + 1 - k'$, and $\epsilon \equiv -j + 1 - k'$.

The correlations for up to M shifts are considered. Therefore, the $2M+1$ vectors $\{\boldsymbol{k}_j\}$, $j = -M, \ldots, M$ are considered and it is assumed that $R_j = Q_j = 0$ when $|j| > M$. Thus, (10) and (13) are first-order ordinary differential equations with $3M+2$ variables, that is,

$$
\frac{d}{dt} \boldsymbol{z} = \boldsymbol{G} \boldsymbol{z} + \boldsymbol{b}, \tag{14}
$$

where $\boldsymbol{z} = [Q_0, \ldots, Q_M, R_{-M}, \ldots, R_0, \ldots, R_M]^T$ and the matrix $\boldsymbol{G}$ and vector $\boldsymbol{b}$ are determined by (10) and (13). All initial values of $Q_j (j \neq 0)$ and R_j are equal to zero because p_i and $h_i(0)$ are independently generated. Therefore, $\boldsymbol{z}$ at $t = 0$ is $\boldsymbol{z}_0 = [1, \underbrace{0, \ldots, 0}_{3M+1}]^T$. Using this as the initial condition, we can analytically solve (14) to obtain

$$
\boldsymbol{z}(t) = e^{\boldsymbol{G}t} \left(\boldsymbol{z}_0 - \boldsymbol{G}^{-1} e^{-\boldsymbol{G}t} \boldsymbol{b} + \boldsymbol{G}^{-1} \boldsymbol{b} \right). \tag{15}
$$

Whether the MSE converges or diverges depends on the step size. Therefore, knowing the upper bound of the step size is very important. In the case of the analysis in this paper, the MSE converges if R and Q converge from (9). It is necessary that all eigenvalues of the matrix $\boldsymbol{G}$ in (14) are negative for the convergence of R and Q. Therefore, the upper bound of the step size can be obtained by solving the closed-form equation $\lambda_{\max} = 0$, where $\lambda_{\max}$ is the maximum eigenvalue of $\boldsymbol{G}$.

4. RESULTS AND DISCUSSION

Actual elements of the impulse response of a primary path are generally strongly correlated with each other. For example, Fig. 2 shows an example of the impulse response of a primary path obtained experimentally. While the variance of the elements of the impulse response of the primary path is $\hat{\kappa}_0 \equiv \frac{1}{N_P} \sum_{i=1}^{N_P} p_i^2 = 1.21081 \times 10^{-3}$, the correlation between the neighboring elements is $\hat{\kappa}_1 \equiv \frac{1}{N_P - 1} \sum_{i=2}^{N_P} p_i p_{i-1} = 0.963143 \times 10^{-3}$. This indicates that there are strong correlations between the elements of the impulse response of the actual primary path. In our previous studies[25, 26], the elements were assumed to be generated independently. In this paper, the theory has been generalized in the previous section to explain the system behaviors in the case of the actual primary path. That is, the theory absorbs the characteristics of the primary path by setting $\kappa_j = \hat{\kappa}_j$, where κ_j is defined by (1).

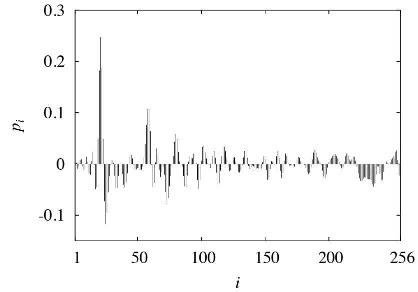

Fig. 2. Example of impulse response of an actual primary path.

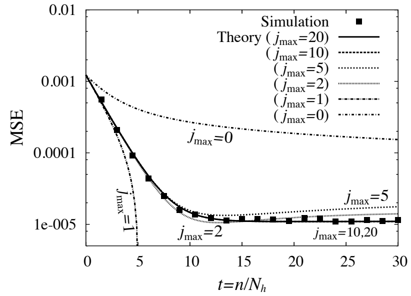

Fig. 3. Learning curves in the case of the actual primary path.

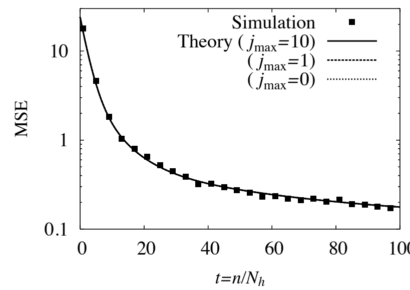

Fig. 4. Learning curves in the case of the actual primary path and realistic secondary path.

Figure 3 shows the learning curves obtained theoretically along with the corresponding simulation results in the case of the actual primary path. The conditions are $a = N_p/N_h = 1$; $\mu = 0.1$; $r_k = \delta_{k,0}$, i.e., the reference signal is white; the variance of the background noise is $\sigma_\xi^2 = 1 \times 10^{-5}$; the secondary path C of a two-tap FIR filter, i.e., $K = 2$, with coefficients of $c_1 = c_2 = 1$; and the estimated secondary path has no error, in other words, $\tilde{c}_1 = \tilde{c}_2 = 1$.

In the theoretical calculations, the value of M in (14) is 10. The results of six cases are plotted, that is, $j_{\text{max}} = 0, 1, 2, 5, 10$, and 20. Here, j_{max} is the maximum value of the subscript j of κ_j defined by (1). For example, $j_{\text{max}} = 0$ means that $\kappa_j = 0, |j| > 0$, that is, the theory only considers the variance of the elements of the impulse response of the primary path. The $j_{\text{max}} = 2$ means that $\kappa_j = 0, |j| > 2$, that is, the theory considers $\kappa_j, j = -2, -1, 0, 1, 2$.

In the computer simulations, $N_h = N_p = 256$ and ensemble means for 1000 trials are plotted. The impulse response of the primary path in the computer simulation is that shown in Fig. 2, that is, the actual impulse response measured experimentally.

Figure 3 shows the following. The theory for $j_{\text{max}} = 0$, that is, the theory only absorbing the variance, cannot predict the behaviors of the computer simulation. These results correspond to our previous theory [25, 26]. The theory for $j_{\text{max}} = 1$ also cannot predict the behaviors. However, the larger the value of j_{max}, the more closely the theoretical results agree with the simulation results. The theoretical results for $j_{\text{max}} = 10$ and 20 almost overlap and agree with the simulation results reasonably well.

Figure 4 shows the learning curves obtained theoretically along with the corresponding simulation results when the impulse response of the primary path is that shown in Fig. 2, and the coefficient vector $\boldsymbol{c}$ of the secondary path is that reported in [2]. This coefficient vector $\boldsymbol{c}$ is highly realistic. The value of K is 25.

The variance of the background noise is $\sigma_\xi^2 = 0.1$. The step size μ is 0.001. The other conditions are the same as those for Fig. 3, that is, $a = N_p/N_h = 1$; $r_k = \delta_{k,0}$, i.e., the reference signal is white; and the estimated secondary path has no error, in other words, $\tilde{\boldsymbol{c}} = \boldsymbol{c}$.

In the theoretical calculations, the results of three cases are plotted, $j_{\text{max}} = 0, 1$, and 10. The value of M in (14) is 10. In the computer simulations, ensemble means for 1000 trials are plotted.

Figure 4 shows that the theoretical results agree with the simulation results reasonably well. The three theoretical learning curves almost coincide with each other, a strong contrast to Fig. 3. This indicates that the variance of the elements of the primary path is sufficient, and that the other correlations do not affect the behaviors of the system when the secondary path is realistic. The reason for the difference between Figs. 3 and 4 can be explained that the effect of the correlation of the primary path is weakened by the randomness of the elements of the secondary path.

As described in this section, by absorbing the correlations of the elements of the impulse response of the primary path by using a certain j_{max}, the theory derived in this paper can predict the behaviors in the case of an actual primary path. This shows that even if concrete values of the actual elements, that are inherently deterministic, of the impulse response of the primary path are not known, the theory can predict the behaviors in the case of an actual primary path through the statistic $\hat{\kappa}$. Furthermore, the effect of the correlation of the primary path is decreased by that of the realistic secondary path. The most important point is that the generalized theory derived in this paper can predict the behaviors in the case of actual primary and secondary paths.

5. CONCLUSIONS

A theory that predicts the behaviors of the FXLMS algorithm was derived by using a statistical-mechanical method. Especially, in this paper, we have generalized the theory to predict the system behaviors in the case of an actual primary path of which elements have strong correlations.

6. REFERENCES

[1] P. A. Nelson and S. J. Elliott, *Active Control of Sound*, Academic, San Diego, CA, 1992.

[2] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems — Algorithms and DSP Implementations*, Wiley, New York, 1996.

[3] S. M. Kuo and D. R. Morgan, "Active noise control: a tutorial review," *Proc. IEEE*, vol.87, no.6, pp.943–973, June 1999.

[4] B. Widrow and M. E. Hoff, Jr., "Adaptive switching circuits," *IRE WESCON Conv. Rec.*, Pt.4, pp.96–104, 1960.

[5] S. Haykin, *Adaptive Filter Theory, Fourth ed.*, Prentice Hall, NJ, 2002.

[6] H. Kushner and G. G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, Springer, New York, 2003.

[7] A. Benveniste, M. Metivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*, Springer, New York, 2002.

[8] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice Hall, NJ, 1985.

[9] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol.64, no.8, pp.1151–1162, Aug. 1976.

[10] J. E. Mazo, "On the independence theory of equalizer convergence," *Bell Syst. Tech. J.*, vol.58, pp.963–993, 1979.

[11] W. A. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms: a general study, analysis, and critique," *Signal Process.*, vol.6, pp.113–133, 1984.

[12] E. Bjarnason, "Analysis of the Filtered-X LMS algorithm," *IEEE Trans. Speech Audio Process.*, vol.3, no.6, pp.504–514, Nov. 1995.

[13] I. T. Ardekani and W. Abdulla, "Theoretical convergence analysis of FxLMS algorithm," *Signal Process.*, vol.90, pp.3046–3055, 2010.

[14] I. T. Ardekani and W. Abdulla, "On the convergence of real-time active noise control systems," *Signal Process.*, vol.91, pp.1262–1274, 2011.

[15] S. Chan and Y. Chu, "Performance analysis and design of FxLMS algorithm in broadband ANC system with online secondary-path modeling," *IEEE Trans. Audio, Speech, Lang., Process.*, vol.20, no.3, pp.982–993, Mar. 2012.

[16] M. H. Costa, J. C. M. Bermudez, and N. J. Bershad, "Stochastic analysis of the Filtered-x LMS algorithm in systems with nonlinear secondary paths," *IEEE Trans. Signal Process.*, vol.50, no.4, pp.1327–1342, June 2002.

[17] O. J. Tobias, J. C. M. Bermudez, R. Seara, and N. Bershad, "An improved model for the second moment behavior of the Filtered-X LMS algorithm," *Proc. IEEE Adaptive Syst. Signal Process., Commun., Contr. Symp.*, Lake Louise, AB, Canada, pp.337–341, Oct. 2000.

[18] L. S. Resende and J. C. M. Bermudez, "An efficient model for the convergence behavior of the FXLMS algorithm with Gaussian inputs," *Proc. IEEE/SP 13th Workshop on Statistical Signal Processing*, pp.97–102, 2005.

[19] H. J. Butterweck, "The independence assumption: a dispensable tool in adaptive filter theory," *Signal Process.*, vol.57, pp.305–310, 1997.

[20] O. J. Tobias, J. C. M. Bermudez, and N. J. Bershad, "Mean weight behavior of the Filtered-X LMS algorithm," *IEEE Trans. Signal Process.*, vol.48, no.4, pp.1061–1075, Apr. 2000.

[21] S. Miyagi and H. Sakai, "Mean-square performance of the filtered-reference/filtered-error LMS algorithm," *IEEE Trans. Circuits Syst. I*, vol.52, no.11, pp.2454–2463, Nov. 2005.

[22] L. Vicente and E. Masgrau, "Novel FxLMS convergence condition with deterministic reference," *IEEE Trans. Signal Process.*, vol.54, no.10, pp.3768–3774, Oct. 2006.

[23] Y. Hinamoto and H. Sakai, "Analysis of the Filtered-X LMS algorithm and a related new algorithm for active control of multitonal noise," *IEEE Trans. Speech Audio Process.*, vol.14, no.1, pp.123–130, Jan. 2006.

[24] S. C. Douglas and W. Pan, "Exact expectation analysis of the LMS adaptive filter," *IEEE Trans. Signal Process.*, vol.43, no.12, pp.2863–2871, Dec. 1995.

[25] S. Miyoshi and Y. Kajikawa, "Statistical-mechanics approach to the Filtered-X LMS algorithm," *Electron. Lett.*, vol.47, no.17, pp.997–999, Aug. 2011.

[26] S. Miyoshi and Y. Kajikawa, "Statistical-mechanical analysis of the FXLMS algorithm with nonwhite reference signals," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP2013)*, pp.5652–5656, May 2013.