STATISTICAL-MECHANICAL ANALYSIS OF THE FXLMS ALGORITHM WITH NONWHITE REFERENCE SIGNALS

Seiji Miyoshi and Yoshinobu Kajikawa

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

ABSTRACT

We analyze the learning curves of the FXLMS algorithm using a statistical-mechanical method when the reference signal is not necessarily white. We treat the nonwhite reference signal by introducing the correlation function of the signal to the method proposed in our previous study. 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. We obtain simultaneous differential equations that describe the dynamical behaviors of the macroscopic variables under the conditions in which the tapped-delay line is long. We analytically solve the equations to obtain the correlations and finally compute the mean-square error. The obtained theory quantitatively agrees with the results of computer simulations. The theory also gives the upper limit of the step size in the FXLMS algorithm.

Index Terms— Filtered-X LMS algorithm, adaptive filter, active noise control, statistical-mechanical informatics, nonwhite reference signals

1. INTRODUCTION

Recently, active noise control (ANC) has been practically realized owing to the progress of digital signal processing technology[1, 2, 3]. ANC is divided into two types, feedforward and feedback ANC[3]. In this paper, feedforward ANC is considered.

The most commonly used algorithm in adaptive filters is the least-mean-square (LMS) algorithm, which was proposed more than half a century ago[4, 5]. When we apply the LMS algorithm to ANC, we should estimate the secondary path beforehand and use inputs that have passed through the estimated secondary path. This procedure is called the Filtered-X LMS (FXLMS) algorithm[6].

Various methods have been proposed to theoretically analyze the LMS algorithm. The principal method is to use the independence assumption[7, 8, 9]. The FXLMS algorithm has also been analyzed on the basis of the independence assumption[10, 11, 12, 13]. In this assumption, successive input vectors of the tapped-delay line are assumed to be independently generated at each time step. However, the actual input vector components are merely shifted to the next position. Hence, each input vector is strongly related to the previous one and the vectors are thus not independent. Owing to this fact, analytical results based on the independence assumption cannot precisely and generally explain experimental results[5].

There are various methods based on assumptions other than the independence assumption. In [13, 14, 15, 16], the step size is assumed to be small. In [17, 18, 19], it is assumed that the correlation between the input signal vectors is more dominant than the correlation between the weight vector of the adaptive filter and the input signal vectors. In [20, 21, 22], it is assumed that the input signal is sinusoidal. In [23], 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.

Meanwhile, numerous powerful analytical and numerical methods have been developed in statistical mechanics. The field in which these methods are used to solve problems in information technology or information science is called statistical-mechanical informatics[24], which is producing significant results in many fields, such as associative memory models, error-correcting codes, wireless communications, image processing, statistical learning, and so forth. In this paper, we theoretically analyze the learning curves of the FXLMS algorithm by applying a statistical-mechanical method when the reference signal is not necessarily white.

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 given by $\langle p_i \rangle = 0, \langle p_i p_j \rangle = \delta_{i,j}$ and is time-invariant. Here, $\langle \cdot \rangle$ denotes expectation and δ denotes the Kronecker delta.

This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan, with Grants-in-Aid for Scientific Research (24360152, 24560280).

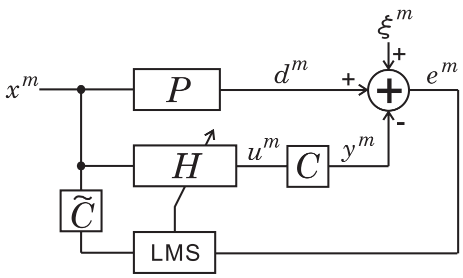

Fig. 1. Block diagram of the ANC system.

The adaptive filter H is an N-tap FIR filter. Its coefficient vector is $\boldsymbol{h}^m = [h_1^m, h_2^m, \ldots, h_N^m]^T$, where m denotes the time step and $N \geq N_P$. The initial value h_i^0 of each coefficient is generated from the stochastic process given by $\langle h_i^0 \rangle = 0, \langle h_i^0 h_j^0 \rangle = \delta_{i,j}$.

The input signal x^m is drawn from a distribution with

$$\langle x^m \rangle = 0, \qquad \langle x^m x^{m-k} \rangle = \sigma_k^2 / N. \tag{1}$$

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

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^m of the secondary path is

$$y^m = \sum_{k=1}^{K} c_k u^{m-k+1}. \tag{2}$$

The error signal e^m is generated by adding an independent background noise ξ^m to the difference between d^m and y^m. That is,

$$e^m = d^m - y^m + \xi^m. \tag{3}$$

Here, the mean and variance of ξ^m are zero and σ_ξ^2, respectively.

The LMS algorithm is used to update the adaptive filter. Here, the coefficient vector $\boldsymbol{c}$ of the secondary path is unknown in general. Therefore, the estimated secondary path $\tilde{C}$, which has been estimated in advance by a certain method, is used to update the adaptive filter. This procedure is called the FXLMS algorithm. 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 obtained by the FXLMS algorithm is

$$\boldsymbol{h}^{m+1} = \boldsymbol{h}^m + \mu e^m \sum_{k=1}^{K} \tilde{c}_k \boldsymbol{x}^{m-k+1}, \tag{4}$$

where μ is the step size.

3. THEORY

From (2) and (3), the MSE is expressed as

$$\left\langle (e^m)^2 \right\rangle = \left\langle (d^m)^2 \right\rangle + \sum_{k=1}^{K} \sum_{k'=1}^{K} c_k c_{k'} \left\langle u^{m-k+1} u^{m-k'+1} \right\rangle - 2 \sum_{k=1}^{K} \left\langle d^m u^{m-k+1} \right\rangle + \sigma_\xi^2. \tag{5}$$

Equation (5) includes many products of d and u including cases where their time steps are different. To calculate these products, we introduce the N-dimensional vectors

$$\boldsymbol{k}_j^m = [k_{j,1}^m, k_{j,2}^m, \ldots, k_{j,N}^m]^T, \quad j = -M, \ldots, M \tag{6}$$

whose elements are $k_{j,i}^m = h_{\mathrm{mod}(i+j-1,N)+1}^m$. That is, $\boldsymbol{k}_j^m$ is the j-shifted vector of the coefficient vector $\boldsymbol{h}^m$ of the adaptive filter. Note that $\boldsymbol{k}_0^m = \boldsymbol{h}^m$.

In the following, the limit $N_P, N \to \infty$ is considered. This condition is called the thermodynamic limit in statistical mechanics. Here, $a = N_P/N$ is kept constant. When the shift number j is $O(1)$, we can obtain

$$(\boldsymbol{h}^m)^T \boldsymbol{x}^m = (\boldsymbol{k}_j^m)^T \boldsymbol{x}^{m-j}. \tag{7}$$

Equation (7) is based on the fact that the shift of the tap input vector is canceled with the shift of the elements of the adaptive filter. Here, the effect of the edge of the adaptive filter can be ignored since both $\boldsymbol{h}^m$ and $\boldsymbol{k}_j^m$ are N-dimensional, i.e., infinitely long, vectors. Equation (7) 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 $R_j^m \equiv \frac{1}{aN} \sum_{i=1}^{aN} p_i k_{j,i}^m$ and $Q_j^m \equiv \frac{1}{N} \sum_{i=1}^{N} h_i^m k_{j,i}^m$. R_j^m and Q_j^m are the cross-correlation between $\boldsymbol{p}$ and $\boldsymbol{h}^m$ and the autocorrelation of $\boldsymbol{h}^m$, respectively. Here, note that both correlations are not functions of the time-direction shift but functions of the element-direction shift.

Then, we obtain $\left\langle d^{m-j} u^m \right\rangle = a \sum_{i=-M}^{M} R_i \sigma_{i-j}^2$, $\left\langle u^{m-j} u^m \right\rangle = \sigma^2 Q_j$, and $\left\langle d^{m-j} d^m \right\rangle = a\sigma_j^2$. Here, we have omitted the time steps of the macroscopic variables since they do not change by $O(1)$ in the $O(1)$ time updates in the model considered in this paper, as described later. We can express the MSE (5) in terms of the cross-correlation R_j and autocorrelation Q_j as

$$\left\langle (e^m)^2 \right\rangle = \sum_{k=1}^{K} c_k \sum_{-M}^{M} \left(\sum_{k'=1}^{K} c_{k'} Q_i \sigma_{i-k+k'}^2 - 2aR_i \sigma_{i+k-1}^2 \right) + a\sigma_0^2 + \sigma_\xi^2. \tag{8}$$

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.

We first derive a differential equation for R_j. 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

$$\boldsymbol{k}_j^{m+1} = \boldsymbol{k}_j^m + \mu e^m \sum_{k=1}^{K} \tilde{c}_k \boldsymbol{x}^{m-k+1-j}. \tag{9}$$

Note that the time step of the tap input vector $\boldsymbol{x}$ is shifted by j compared with that in (4). Multiplying both sides of (9) on the left by the N-dimensional vector $[p_1, p_2, \ldots, p_{aN}, \underbrace{0, \ldots, 0}_{(1-a)N}]^T$, we obtain

$$aNR_j^{m+1} = aNR_j^m + \mu e^m \sum_{k=1}^{K} \tilde{c}_k d^{m-k+1-j}. \tag{10}$$

In (10), the left-hand side and the first term on the right-hand side are $O(N)$ and the other terms are $O(1)$. This means that the coefficient vector $\boldsymbol{h}^m$ of the adaptive filter should be updated $O(N)$ times to change R_j by $O(1)$. Therefore, we introduce the continuous time t, which is the time step m normalized by the tap length N, and use it to represent the adaptive process[25]. If the adaptive filter is updated Ndt times in an infinitely small time dt, we can obtain Ndt equations that are similar to (10). Summing all these equations, we obtain a differential equation that describes the dynamical behavior of R_j in a deterministic form as follows:

$$\frac{dR_j}{dt} = \mu \sum_{k'=1}^{K} \tilde{c}_{k'} \left(\sigma_{k'+j-1}^2 - \sum_{k=1}^{K} c_k \sum_{i=-M}^{M} R_i \sigma_{i+k-k'-j}^2 \right). \tag{11}$$

Next, multiplying (4) by (9) 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 (12), where $\mathrm{sgn}(\cdot)$ is the sign function. In addition, $\alpha \equiv \Theta(\gamma)\gamma - k''$, $\beta \equiv \Theta(\epsilon)\epsilon - k''$, $\gamma \equiv j + 1 - k'$, and $\epsilon \equiv -j + 1 - k'$, where $\Theta(\cdot)$ is the step function.

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 $|j| > M$. Then (11) and (12) are first-order ordinary differential equations with $3M+2$ variables, that is,

$$\frac{d}{dt}\boldsymbol{z} = \boldsymbol{G}\boldsymbol{z} + \boldsymbol{b}, \tag{13}$$

where $[Q_0, \ldots, Q_M, R_{-M}, \ldots, R_0, \ldots, R_M]^T$ and the matrix $\boldsymbol{G}$ and vector $\boldsymbol{b}$ are determined by (11) and (12). All initial values of $Q_j (j \neq 0)$ and R_j are equal to zero because p_i, h_i^0, and $k_{j,i}^0 (j \neq 0)$ 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 (13) to obtain

$$\boldsymbol{z}(t) = e^{\boldsymbol{G}t} \left(\boldsymbol{z}_0 - \boldsymbol{G}^{-1} e^{-\boldsymbol{G}t} \boldsymbol{b} + \boldsymbol{G}^{-1} \boldsymbol{b} \right) \tag{14}$$

[26], where $e^{\boldsymbol{D}}$ is the matrix exponential function defined by $e^{\boldsymbol{D}} \equiv \sum_{k=0}^{\infty} \frac{1}{k!} \boldsymbol{D}^k = \boldsymbol{I} + \frac{1}{1!}\boldsymbol{D} + \frac{1}{2!}\boldsymbol{D}^2 + \cdots$.

4. RESULTS AND DISCUSSION

We first investigate the validity of the theory by comparison with simulation results regarding the dynamical behaviors of the MSE, that is, the learning curves. Figure 2 shows the learning curves obtained theoretically in the previous section, along with the corresponding simulation results. The correlation function of the input signal is $\sigma_0^2 = 1, \sigma_k^2 = 0$ when $k > 0$ (White), and $\sigma_0^2 = 1, \sigma_1^2 = 0.5, \sigma_k^2 = 0$ when $k > 1$ (Nonwhite). There is no background noise, that is, $\sigma_\xi^2 = 0$. The numbers of taps of the primary path P and the adaptive filter H are equal, i.e., $a = N_P/N = 1$. The secondary path C is a two-tap FIR filter, that is, $K = 2$, and its coefficients are $c_1 = c_2 = 1$. The estimated secondary path has no error, in other words, $\tilde{c}_1 = \tilde{c}_2 = 1$.

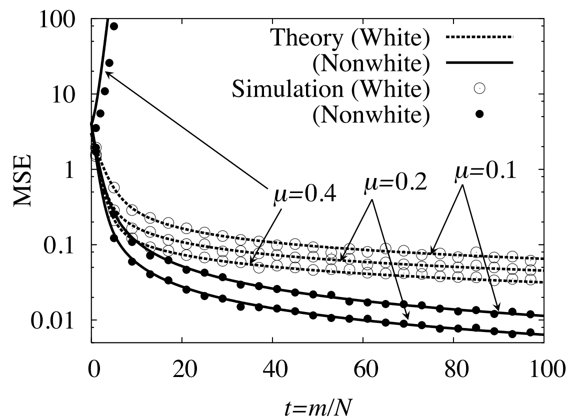

Fig. 2. Learning curves obtained theoretically and by simulation.

In Fig. 2, the curves represent theoretical results and the symbols represent simulation results. In the theoretical calculation, the results are obtained by substituting R_j and Q_j, which are obtained by solving (13), into (8) in the case where the range of the correlations considered is $M = 20$. In the computer simulations, the numbers of taps of the primary path and the adaptive filter are $N_P = N = 500$. Ensemble means for 1000 trials are plotted. Figure 2 shows that the theoretical results agree with the simulation results including the difference between the behaviors for white and nonwhite input signals. It is also shown that the upper limit of the step size

$$
\begin{aligned}
\frac{dQ_j}{dt} = \mu \sum_{k'=1}^{K} \tilde{c}_{k'} \Bigg\{ & \sum_{i=-M}^{M} \left[aR_i \left(\sigma_{i-\gamma}^2 + \sigma_{i-\epsilon}^2 \right) - \sum_{k=1}^{K} c_k Q_i \left(\sigma_{i-k'+k+j}^2 + \sigma_{i-k'+k-j}^2 \right) \right] \\
& - \mu \Bigg[\mathrm{sgn}(\gamma) \sum_{k''=1}^{|\gamma|} \left(\delta_{\alpha,0}\sigma_\xi^2 + a\sigma_\alpha^2 - \sum_{k=1}^{K} c_k \sum_{i=-M}^{M} \left(aR_i \sigma_{i+k-1-\alpha}^2 + aR_i \sigma_{i+k-1+\alpha}^2 - \sum_{k'''=1}^{K} c_{k'''} Q_i \sigma_{i-k+k'''-\alpha}^2 \right) \right) \sum_{i=1}^{K} \tilde{c}_i \sigma_{k'-i-j+\alpha}^2 \\
& + \mathrm{sgn}(\epsilon) \sum_{k''=1}^{|\epsilon|} \left(\delta_{\beta,0}\sigma_\xi^2 + a\sigma_\beta^2 - \sum_{k=1}^{K} c_k \sum_{i=-M}^{M} \left(aR_i \sigma_{i+k-1-\beta}^2 + aR_i \sigma_{i+k-1+\beta}^2 - \sum_{k'''=1}^{K} c_{k'''} Q_i \sigma_{i-k+k'''-\beta}^2 \right) \right) \sum_{i=1}^{K} \tilde{c}_i \sigma_{k'-i+j+\beta}^2 \Bigg] \Bigg\} \\
+ \mu^2 \Bigg[& \sum_{k=1}^{K} c_k \sum_{i=-M}^{M} \left(\sum_{k'=1}^{K} c_{k'} Q_i \sigma_{i-k+k'}^2 - 2aR_i \sigma_{i+k-1}^2 \right) + a\sigma_0^2 + \sigma_\xi^2 \Bigg] \sum_{k''=1}^{K} \sum_{k'''=1}^{K} \tilde{c}_{k''} \tilde{c}_{k'''} \sigma_{k''-k'''-j}^2,
\end{aligned}
\tag{12}
$$

μ for white input signals is larger than 0.4, whereas that for nonwhite input signals is smaller than 0.4.

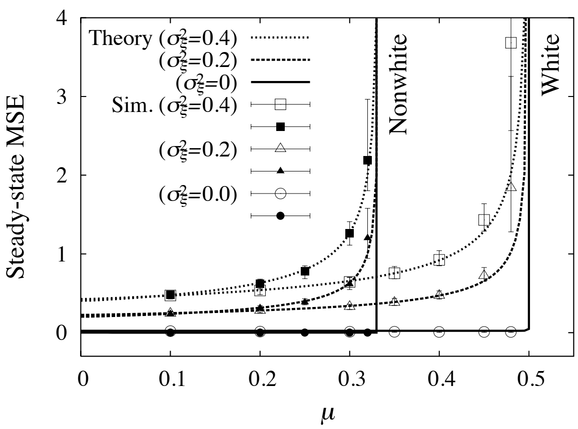

Fig. 3. Relationship between step size μ and steady-state MSE obtained theoretically and by simulation.

We next investigate the upper limit of the step size μ in detail. This is very important from the practical viewpoint. We compare the derived theory and simulation results in terms of the relationship between the step size μ and the steady-state MSE. The number of taps and the coefficient vector of the secondary path C are $K = 2$ and $\boldsymbol{c} = [1, 1]^T$, respectively. The secondary path estimation has no error, that is, $\tilde{\boldsymbol{c}} = \boldsymbol{c}$. The correlation function of the input signal is $\sigma_0^2 = 1, \sigma_k^2 = 0$ when $k > 0$ (White), and $\sigma_0^2 = 1, \sigma_1^2 = 0.5, \sigma_k^2 = 0$ when $k > 1$ (Nonwhite). The variances of the background noise are $\sigma_\xi^2 = 0, 0.2$, and 0.4. Figure 3 shows the results obtained theoretically and by simulation. In the theoretical calculation, the ratio of the number of taps of the primary path to that of the adaptive filter is $a = 1$. The range of the correlations considered is $M = 20$ and MSEs at $t = 10^5$ are plotted. In the computer simulations, the numbers of taps of the primary path and the adaptive filter are $N_P = N = 200$ and the means of the 200 squared errors from $t = 900$ to $t = 1100$ are plotted. Symbols and error bars represent medians and standard deviations, respectively. The theoretical results agree with the simulation results reasonably well in Fig. 3. This agreement indicates that the derived theory can explain the simulation results regarding the upper limit of the step size μ for which the MSE converges when the input signals are both white and nonwhite.

Although the case where the secondary path C is a two-tap filter was investigated, the obtained theory is effective when C has a larger number of taps.

5. CONCLUSIONS

We have analyzed the learning curves of the FXLMS algorithm using a statistical-mechanical method when the reference signal is not necessarily white. The obtained theory quantitatively agrees with the results of computer simulations. The theory also gives the upper limit of the step size. The independence assumption and other assumptions that have often been used in the literature are not used in the theory derived in this paper. The principal assumption used is that N is large.

The investigation of the case where the estimated secondary path has some error, especially the relationship with the 90 degree condition, is an important future work.

6. RELATION TO PRIOR WORK

Our previous theory[27, 28] treated only white reference signals. This assumption was a severe limitation since the reference signal to be canceled is not necessarily white. Therefore, in this paper, we have generalized the previous theory to the case where the reference signal is not necessarily white by introducing the correlation function of the signal.

7. REFERENCES

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

[2] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems — Algorithms and DSP Implementations*, New York: Wiley, 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] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice Hall, 1985.

[7] 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.

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

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

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

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

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

[13] 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.

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

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] Y. Hinamoto and H. Sakai, "A Filtered-X LMS algorithm for sinusoidal reference signals —Effects of frequency mismatch—," *IEEE Signal Process. Lett.*, vol.14, no.4, pp.259–262, Apr. 2007.

[23] 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.

[24] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction*, Oxford University Press, New York, 2001.

[25] A. Engel and C. V. Broeck, *Statistical Mechanics of Learning*, Cambridge University Press, Cambridge, 2001.

[26] K. Shimizu and M. Suzuki, *Notes on Ordinary and Partial Differential Equations*, Corona, Tokyo, 1995 (in Japanese).

[27] 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.

[28] S. Miyoshi and Y. Kajikawa, "Theoretical discussion of the Filtered-X LMS algorithm based on statistical mechanical analysis," *Proc. IEEE Statistical Signal Processing Workshop (SSP2012)*, pp.341–344, Aug. 2012.