Analysis of the FXLMS algorithm with norm-constant time-varying primary path

Norihiro Ishibushi*, Yoshinobu Kajikawa†, and Seiji Miyoshi†
* Graduate School of Science and Engineering, Kansai University, Japan.
†Faculty of Engineering Science, Kansai University, Japan.
E-mail: miyoshi@kansai-u.ac.jp

***Abstract*—We analyze the behaviors of active noise control with a time-varying primary path using a statistical-mechanical method. The principal assumption used in the analysis is that the impulse responses of the primary path and adaptive filter are sufficiently long. We analyze a novel model in which the reference signal is not necessarily white and the primary path is time-varying while its norm is kept constant in the mean sense. We show the existence of macroscopic steady states and the optimal step size.**

I. Introduction

Active noise control (ANC), which has been practically realized owing to the progress of digital signal processing, is implemented using an adaptive filter [1], [2]. The least-mean-square (LMS) algorithm is the most commonly used algorithm for adaptive filters [3], [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.

Various methods have been proposed for theoretically analyzing the LMS algorithm. The principal method is to use the independence assumption [6]. The FXLMS algorithm has also been analyzed on the basis of the independence assumption [7], [8]. In this assumption, the 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 [4], [7]. In addition, analyses of cases where the step size is small [9], [10] and the periodic reference signal is assumed [11] have been reported.

In our previous paper reported at APSIPA ASC 2014 [12], we analyzed the behaviors of the FXLMS algorithm when the reference signal is white and the primary path is naively time-varying by applying a statistical-mechanical method . In the previous naive model, the norm of the primary path and the mean square error (MSE) diverged and the optimal step size disappeared, indicating that the previous model was rather unrealistic.

In this paper, we propose an analytically solvable model, where the primary path is time-varying and the norm of the coefficient vector of the primary path is kept constant in the mean sense. The model is analyzed by applying a statistical-mechanical method [13], [14], [15], [16]. In addition, the model is generalized to nonwhite reference signal. The analysis gives meaningful results.

II. Analytical model of FXLMS algorithm

Figure 1 shows a block diagram of the ANC system considered in this paper.

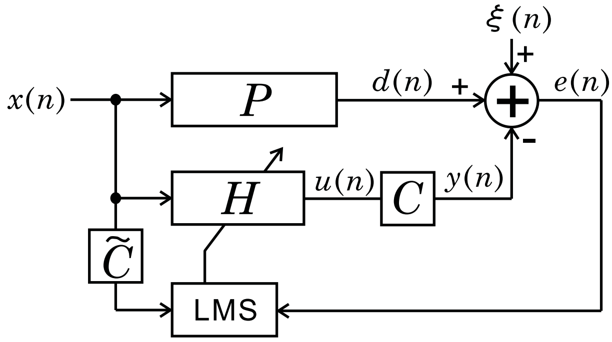

Fig. 1. Block diagram of the ANC system.

The primary path P is represented by an N-tap FIR filter. Its coefficient vector is $\boldsymbol{p}(n) = [p_1(n), p_2(n), \ldots, p_N(n)]^\top$. Each coefficient $p_i(0)$ is independently generated from a distribution with a mean of zero and a variance of unity. Here, $^\top$ denotes transposition and n denotes the time step. The primary path is time-varying, that is,

$$\boldsymbol{p}(n+1) = a^{\frac{1}{N}}\boldsymbol{p}(n) + \sqrt{1-a^{\frac{2}{N}}}\boldsymbol{w}(n), \quad (0 \le a \le 1), \quad (1)$$

where $\boldsymbol{w}(n)$ is an N-dimensional vector. Each coefficient $w_i(n)$ is independently generated from a distribution with a mean of zero and a variance of unity at every time step. a is a parameter that controls the rate of time variation of the primary path. Here, $a = 1$ corresponds to the time-invariant primary path. Note that Eq. (1) means that the norm of the coefficient vector of the primary path is kept constant in the mean sense although the primary path itself is time-varying.

The adaptive filter H is also an N-tap FIR filter. Its coefficient vector is $\boldsymbol{h}(n) = [h_1(n), h_2(n), \ldots, h_N(n)]^\top$. The initial value $h_i(0)$ of each coefficient is zero. The reference signal $x(n)$ is drawn from a distribution with

$$\langle x(n)\rangle = 0, \qquad \langle x(n)x(n-k)\rangle = r_k/N, \quad (2)$$

where $\langle\cdot\rangle$ denotes expectation. The reference signal is shifted through the tapped delay line in the FIR filter. Therefore, the tap input vector is $\boldsymbol{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^\top$. The output of the primary path P is $d(n) = \boldsymbol{p}^\top(n)\boldsymbol{x}(n)$. On the other hand, the output of the adaptive filter H is $u(n) = \boldsymbol{h}^\top(n)\boldsymbol{x}(n)$.

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

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

An 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.

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 using some other 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 $\tilde{\boldsymbol{c}} = [\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_K]^\top$, the update obtained by applying the FXLMS algorithm is

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

where μ is the step-size parameter.

III. Theory

The mean square error (MSE) is

$$\begin{aligned}\left\langle e^2(n)\right\rangle &= \left\langle (d(n) - y(n) + \xi(n))^2 \right\rangle \qquad (6)\\ &= \left\langle d^2(n)\right\rangle - 2\sum_{k=1}^{K} c_k \left\langle d(n)u(n-k+1)\right\rangle + \sigma_\xi^2 \\ &\quad + \sum_{k=1}^{K}\sum_{k'=1}^{K} c_k c_{k'} \left\langle u(n-k+1)u(n-k'+1)\right\rangle. \qquad (7)\end{aligned}$$

Equation (7) includes products of d and u and products of u and u including cases where their time steps are different. To calculate these products, we introduce the N-dimensional vectors $\boldsymbol{k}_j(n) = [k_{j,1}(n), k_{j,2}(n), \ldots, k_{j,N}(n)]^\top$, whose elements are $k_{j,i}(n) = h_{i+j}(n)$. 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 $N \to \infty$ is considered. Note that this long-impulse-response assumption or long-filter assumption is reasonable, considering actual acoustic systems. When the shift number j is $O(1)$, we obtain

$$\boldsymbol{h}^\top(n)\boldsymbol{x}(n) \simeq \boldsymbol{k}_j^\top(n)\boldsymbol{x}(n-j). \tag{8}$$

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-dimensional, i.e., infinitely long, vectors. Equation (8) implies that the gap j in the time direction can be replaced by the subscript of vector $\boldsymbol{k}$. In addition, we introduce macroscopic variables defined by

$$R_j = \frac{1}{N}\sum_{i=1}^{N} p_i(n) k_{j,i}(n), \tag{9}$$

$$Q_j = \frac{1}{N}\sum_{i=1}^{N} h_i(n) k_{j,i}(n). \tag{10}$$

Equations (9) and (10) indicate that R_j and Q_j are the cross-correlation between $\boldsymbol{p}(n)$ and $\boldsymbol{h}(n)$ and the autocorrelation of $\boldsymbol{h}(n)$, respectively. Then, we obtain

$$\langle d(n-j)u(n)\rangle = \sum_{i=-M}^{M} R_i r_{i-j}, \tag{11}$$

$$\langle u(n-j)u(n)\rangle = \sum_{i=-M}^{M} Q_i r_{i-j}, \tag{12}$$

$$\langle d(n-j)d(n)\rangle = r_j. \tag{13}$$

We have omitted the time steps of the microscopic variables since they do not change by $O(1)$ in the $O(1)$ time updates in the model treated in this paper, as described later. We can express the MSE (7) in terms of R_j and Q_j as

$$\begin{aligned}\left\langle e^2(n)\right\rangle = \sum_{k=1}^{K} c_k \sum_{i=-M}^{M} &\left(\sum_{k'=1}^{K} c_{k'} Q_i r_{i-k+k'} \right.\\ &\left. - 2R_i r_{i+k-1} \right) + r_0 + \sigma_\xi^2. \end{aligned} \tag{14}$$

This formula shows that the MSE is a function of the macroscopic variables R_i and Q_i. 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 $\boldsymbol{h}(n)$ of the adaptive filter is updated, the j-shifted vector $\boldsymbol{k}_j(n)$ is also changed. This change can be described as

$$\boldsymbol{k}_j(n+1) = \boldsymbol{k}_j(n) + \mu e(n) \sum_{k=1}^{K} \tilde{c}_k \boldsymbol{x}(n-k+1-j). \tag{15}$$

Note that the time step of the tap input vector $\boldsymbol{x}$ is shifted by j compared with that in (5). Multiplying both sides of (1) by (15), we obtain

$$\begin{aligned}NR_j(n+1) &= a^{\frac{1}{N}} NR_j(n) + a^{\frac{1}{N}} \mu e(n) \sum_{k=1}^{K} \tilde{c}_k d(n-k+1-j) \\ &+ \sqrt{1-a^{\frac{2}{N}}}\, \boldsymbol{w}^\top(n) \left(\boldsymbol{k}_j(n) + \mu e(n) \sum_{k=1}^{K} \tilde{c}_k \boldsymbol{x}(n-k+1-j) \right). \end{aligned} \tag{16}$$

In (16), the l.h.s. and the first term on the r.h.s. are $O(N)$ and the other terms are $O(1)$. This means that the coefficient vector

$\boldsymbol{h}(n)$ 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 n normalized by the tap length N, and use it to represent the adaptive process. If the adaptive filter is updated Ndt times in an infinitely small time dt, we can obtain Ndt equations as follows:

$$\begin{aligned} NR_j(n+1) &= a^{\frac{1}{N}} NR_j(n) \\ &\quad + a^{\frac{1}{N}} \mu e(n) \sum_{k=1}^{K} \tilde{c}_k d(n-k+1-j), \end{aligned} \tag{17}$$

$$\begin{aligned} a^{-\frac{1}{N}} NR_j(n+2) &= NR_j(n+1) \\ &\quad + \mu e(n) \sum_{k=1}^{K} \tilde{c}_k d(n-k+2-j), \end{aligned} \tag{18}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\begin{aligned} a^{-\frac{Ndt-1}{N}} NR_j(n+Ndt) &= a^{-\frac{Ndt-2}{N}} NR_j(n+Ndt-1) \\ &\quad + a^{-\frac{Ndt-2}{N}} \mu e(n+Ndt-1) \\ &\quad \times \sum_{k=1}^{K} \tilde{c}_k d(n-k+Ndt-j). \end{aligned} \tag{19}$$

Here, terms including $\boldsymbol{w}$ are omitted in (17) – (19) since their expectations are zero. Note that each equation is multiplied by the power of a. Summing all these equations, we obtain

$$N(R_j+dR_j) = a^{dt} NR_j + \sum_{i=1}^{Ndt} a^{\frac{i}{N}} Ndt \mu \left\langle e(m) \sum_{k=1}^{K} \tilde{c}_k d(m-k+1-j) \right\rangle. \tag{20}$$

Here, we can represent the effect of the probabilistic variables by their means since the updates are executed Ndt times, that is, many times, to change R_j by dR_j. m is an auxiliary time step that is introduced to represent the difference in time steps. We see that $\left\langle e(m) \sum_{k=1}^{K} \tilde{c}_k d(m-k+1-j) \right\rangle$ in (20) includes many products of d and u from (3) and (4). Since we can represent this expectation using the macroscopic variables, we obtain differential equations that describe the dynamical behaviors of R_j in a deterministic form as

$$\frac{dR_j}{dt} = (\ln a) R_j + \mu \sum_{k'=1}^{K} \tilde{c}_{k'} \Big(r_{-k'+1-j} - \sum_{k=1}^{K} \sum_{i=-M}^{M} c_k r_{i+k-k'-j} R_i \Big), \tag{21}$$

where δ denotes the Kronecker delta. Here, we also used l'Hôpital's rule.

Next, multiplying (5) by (15) 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

$$\begin{aligned} \frac{dQ_j}{dt} &= \mu \sum_{i=-M}^{M} \sum_{k'=1}^{K} \tilde{c}_{k'} \Bigg[(r_{i-\gamma} + r_{i-\epsilon}) R_i \\ &- \sum_{k=1}^{K} c_k (r_{i-(k-k'+j)} + r_{i-(k-k'-j)}) Q_{|i|} \\ &- \mu \Bigg\{ \mathrm{sgn}(\gamma) \sum_{k''=1}^{|\gamma|} \sum_{k''''=1}^{K} \tilde{c}_{k''''} r_{\zeta-k''''} \Bigg(r_\alpha + \sigma_\xi^2 \delta_{\alpha,0} \\ &\quad - \sum_{k=1}^{K} c_k \Bigg((r_{i-(1-k+\alpha)} + r_{i-(1-k-\alpha)}) R_i \\ &\quad\quad - \sum_{k'''=1}^{K} c_{k'''} r_{i-(\alpha-k+k''')} Q_{|i|} \Bigg) \Bigg) \Bigg\} \\ &- \mu \Bigg\{ \mathrm{sgn}(\epsilon) \sum_{k''=1}^{|\epsilon|} \sum_{k''''=1}^{K} \tilde{c}_{k''''} r_{\eta-k''''} \Bigg(r_\beta + \sigma_\xi^2 \delta_{\beta,0} \\ &\quad - \sum_{k=1}^{K} c_k \Bigg((r_{i-(1-k+\beta)} + r_{i-(1-k-\beta)}) R_i \\ &\quad\quad - \sum_{k'''=1}^{K} c_{k'''} r_{i-(\beta-k+k''')} Q_{|i|} \Bigg) \Bigg) \Bigg\} \Bigg] \\ &+ \mu^2 \sum_{i=-M}^{M} \Bigg\{ r_0 + \sum_{k=1}^{K} c_k \Bigg[\sum_{k'=1}^{K} c_{k'} r_{i-k'+k} Q_{|i|} - 2 r_{i-k+1} R_i \Bigg] + \sigma_\xi^2 \Bigg\} \\ &\quad \times \sum_{k''=1}^{K} \sum_{k'''=1}^{K} \tilde{c}_{k''} \tilde{c}_{k'''} r_{k''-k'''+j}, \end{aligned} \tag{22}$$

where

$$\begin{aligned} &\alpha \equiv \Theta(\gamma)\gamma - k'', && \beta \equiv \Theta(\epsilon)\epsilon - k'', \\ &\gamma \equiv j+1-k', && \epsilon \equiv -j+1-k', \\ &\zeta \equiv -\Theta(-\gamma)\gamma - k'' + 1, && \eta \equiv -\Theta(-\epsilon)\epsilon - k'' + 1. \end{aligned}$$

Here, $\Theta(\cdot)$ and $\mathrm{sgn}(\cdot)$ denote the step function and sign function, respectively.

Considering the $2M+1$ vectors $\boldsymbol{k}_j, j=-M,\dots,M$, and assuming $R_j = Q_j = 0$ when $|j| > M$, then (21) and (22) are first-order ordinary differential equations with $3M+2$ variables, that is,

$$\frac{d}{dt}\boldsymbol{z} = \boldsymbol{G}\boldsymbol{z} + \boldsymbol{b}, \tag{23}$$

where $\boldsymbol{z} = [Q_0, \cdots, Q_M, R_{-M}, \cdots, R_0, \cdots, R_M]^\top$. The matrix $\boldsymbol{G}$ and vector $\boldsymbol{b}$ are determined by (21) and (22). Using the fact that $\boldsymbol{z}$ at $t=0$ is a zero vector as the initial condition, we can analytically solve (23) to obtain

$$\boldsymbol{z}(t) = e^{\boldsymbol{G}t}(\boldsymbol{G}^{-1}e^{\boldsymbol{G}t}\boldsymbol{b} + \boldsymbol{G}^{-1}\boldsymbol{b}). \tag{24}$$

IV. Results and Discussion

We confirm the validity of the theory obtained in the previous section by comparison with simulation results. Figure 2 shows the learning curves obtained theoretically and by simulation. The conditions are $\mu = 0.1$, $r_k = \delta_{k,0}$, that is, the reference signal is white, $\sigma_\xi^2 = 0$, $K = 2$, and $c_1 = c_2 = \tilde{c}_1 = \tilde{c}_2 = 1$. In the theoretical calculation, the results are obtained by substituting R_j and Q_j, which are obtained by solving (23), into (14) in the case of $M = 100$. In the computer simulations, $N = 200$ and ensemble means for 10^4 trials are plotted. Figure 2 shows that the theoretical results agree with the simulation results. The primary path does not have the time-varying property when $a = 1$. When a is relatively small, we see that the MSE increases and approaches a steady value after taking the minimum value. These phenomena are different from those reported in [12].

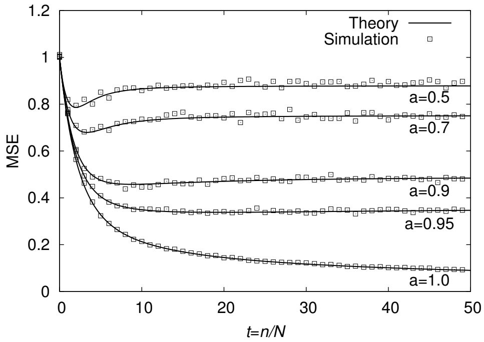

Fig. 2. Learning curves obtained theoretically and by simulation.

Figure 3 shows the relationship between the step size μ and the steady-state MSE. The steady-state MSE can be calculated by letting the l.h.s. of (23) be a zero vector. Our previous naive time-varying model [12] was also analytically solvable but rather unrealistic since the MSE diverged and there were no steady states. In contrast, the time-varying model analyzed in this paper has steady states. Therefore, this novel model based on Eq. (1) is interesting and its analysis gives meaningful results.

The conditions are $a = 0.99, 0.95, 0.9, 0.8$, and 0.6, and $\sigma_\xi^2 = 0$. The correlation functions of the reference signal are $r_k = \delta_{k,0}$ (white), and $r_0 = 1, r_{\pm 1} = 0.5, r_{\pm k} = 0$ when $k \geq 2$ (nonwhite). In the theoretical calculation, $M = 50$. In the computer simulations, $N = 10^3$ and the ensemble means from $t = 90$ to $t = 100$ for 10^3 trials are plotted. The theoretical results agree with the simulation results reasonably well. As the step size μ increases, the noise misadjustment increases gradually and the MSE increases. The upper limit of μ for the convergence of the MSE is $\mu = 0.5$ for the white reference signal and $\mu = 0.3317$ for this nonwhite reference signal. Taking a closer look at Fig. 3, when the step size μ is small, as μ decreases, the MSE also increases. This is due to the lag misadjustment for small μ. We see that an optimal step size μ exists owing to the trade-off between the noise misadjustment and the lag misadjustment. As a decreases, the optimal value of μ decreases. When $a = 0.99, 0.95, 0.9, 0.8$, and 0.6, the optimum step sizes are $\mu_{\text{opt}} = 0.427, 0.383, 0.356, 0.319$, and 0.256 for the white reference signal and $\mu_{\text{opt}} = 0.242, 0.218, 0.203, 0.184$, and 0.150 for this nonwhite reference signal, respectively. Note that the optimum step sizes do not disappear. These phenomena are different from those reported in [12] in which there were not the optimum step sizes since the MSE diverged.

V. Conclusion

We have analyzed the behaviors of the FXLMS algorithm using a statistical-mechanics approach. In particular, we have analyzed the case where the reference signal is not necessarily

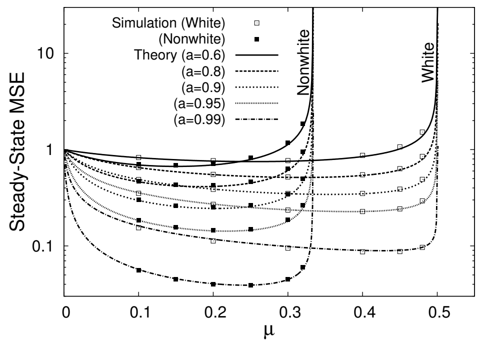

Fig. 3. Relationship between step size μ and steady-state MSE obtained theoretically and by simulation.

white and the primary path is time-varying while its norm is kept constant in the mean sense. The obtained theory quantitatively agrees with the results of computer simulations. We have shown the existence of macroscopic steady states and that the optimal step size does not disappear in this model. These results are in contrast to those of our previous naive model reported at APSIPA ASC 2014.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Numbers 24360152 and 15K00256.

References

[1] S. M. Kuo and D. R. Morgan, "Active noise control: a tutorial review," *Proc. of IEEE*, vol.87, no.6, pp.943–973, June 1999.
[2] Y. Kajikawa, W. S. Gan, and S. M. Kuo, "Recent advances on active noise control: open issues and innovative applications," *APSIPA Trans. Signal and Information Processing*, e3, vol.1, 2012.
[3] B. Widrow and M. E. Hoff, Jr., "Adaptive switching circuits," *IRE WESCON Conv. Rec.*, Pt.4, pp.96–104, 1960.
[4] S. Haykin, *Adaptive Filter Theory, Fourth ed.*, Prentice Hall, New Jersey, 2002.
[5] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, Hoboken, NJ, 2003.
[6] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. of IEEE*, vol.64, no.8, pp.1151–1162, Aug. 1976.
[7] E. Bjarnason, "Analysis of the filtered-X LMS algorithm," *IEEE Trans. Speech Audio Process.*, vol.3, no.6, pp.504–514, Nov. 1995.
[8] I. T. Ardekani and W. Abdulla, "Theoretical convergence analysis of FxLMS algorithm," *Signal Process.*, vol.90, pp.3046–3055, 2010.
[9] H. J. Butterweck, "The independence assumption: a dispensable tool in adaptive filter theory." *Signal Process.*, vol.57, pp.305–310, 1997.
[10] 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.
[11] 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.
[12] N. Egawa, Y. Kajikawa, and S. Miyoshi, "Statistical-mechanical analysis of the FXLMS algorithm with time-varying primary path," *Proc. APSIPA ASC 2014*, 1088 (5 pages), Dec. 2014.
[13] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction*, Oxford University Press, New York, 2001.
[14] A. Engel and C. V. Broeck, *Statistical Mechanics of Learning*, Cambridge University Press, Cambridge, 2001.
[15] 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.
[16] S. Miyoshi and Y. Kajikawa, "Statistical-mechanical analysis of the FXLMS algorithm with actual primary path," *Proc. ICASSP2015*, pp.3502–3506, Apr. 2015.