ANALYSIS OF THE ADAPTIVE FILTER ALGORITHM FOR FEEDBACK-TYPE ACTIVE NOISE CONTROL

Hideaki Sakai and Shigeyuki Miyagi

Dept. of Systems Science, Graduate School of Informatics, Kyoto University
Yoshida, Sakyou-ku, Kyoto, 606-8501 Japan
E-mail: hsakai@i.kyoto-u.ac.jp

ABSTRACT

The feedback-type active noise control (ANC) system uses only one microphone to provide necessary signals for adjusting the adaptive filter. Due to the complicated nature of the whole adaptive filter structure there have been no theoretical results about its convergence properties. In this paper, first a stationary point of the adaptive filter using the filtered-X LMS algorithm is obtained by the averaging method combined with the frequency domain technique. Then the local convergence condition is derived. This is a counterpart of the well-known 90° condition for the feedforward-type ANC. Finally, the convergence condition is explicitly given for a simple example and its validity is shown by some simulations.

1. INTRODUCTION

Recently there have been growing interests in the feedback-type active noise control (ANC) system in Fig.1 where only one microphone is used to provide necessary signals for adjusting the adaptive filter in ANC [1]. This is in contrast with the conventional feedforward-type ANC system where two microphones are used to pick up the reference signal to the primary path and the error signal at the end of the secondary path [2]. The convergence condition of the latter type ANC is well-known. The so-called 90° condition says that the phase difference between the transfer functions of the secondary path and its estimate should lie in the interval $(-\pi/2, \pi/2)$ [2]. However, as far as the authors are aware, there have been no theoretical results about the convergence properties of the feedback-type ANC algorithms. In this paper, we present some results about the stationary point of the adaptive filter and the convergence condition using the averaging method in [3] with the frequency domain technique developed in [4]. This technique converts adaptive algorithms into those in the discrete frequency domain by the discrete Fourier transform (DFT) and is successfully applied to the analysis of rather complicated adaptive systems such as the delayless subband adaptive filter where fixed filters, decimators and upsamples for rate conversion are included.

The configuration we are treating is due to [5] and is shown in Fig.2 where only the error signal picked up by the microphone is available. To generate a "reference" signal to the adaptive filter, the "feedback control" filter is inserted to recover the original external noise $w(n)$. But its transfer function $\hat{B}(z)$ may be different from that of the (physical) feedback path $B(z)$ so that in general the artificially generated reference signal $x(n)$ is not exactly equal to $w(n)$. The weights in the adaptive filter are updated according to the filtered-X LMS algorithm. The idea behind this is that if $w(n)$ is available, the original problem becomes the optimal prediction of $w(n)$ by the output of the cascade of the adaptive filter and the physical feedback path $B(z)$ with the input $w(n)$. Interchanging the filters in the cascade and replacing $B(z)$, $w(n)$ by $\hat{B}(z)$, $x(n)$, respectively, we have the filtered-X LMS algorithm.

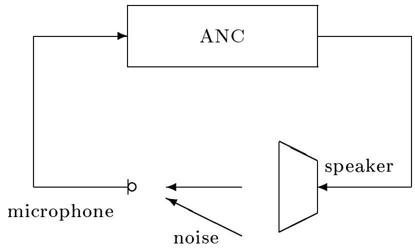

Fig. 1. Feedback-type active noise control system

2. DERIVATION OF THE STATIONARY POINT OF THE ADAPTIVE FILTER

Since we are essentially dealing with the prediction problem of a zero mean stationary process $w(n)$, as stated in [6] some cares are needed to insure the causality of the steady state transfer function of the adaptive filter when the analysis is performed in the discrete frequency domain. We use the technique in [4] to analyze the scheme in Fig.2. The relations of the signals in Fig.2 are written as

$$e(n) = w(n) - \sum_{i=0}^{N_b-1} b_i x'(n-i) \quad (1)$$

$$x'(n) = \sum_{i=0}^{N-1} h_i(n) x(n-i) \quad (2)$$

$$x(n) = e(n) + \sum_{i=0}^{N_b-1} \hat{b}_i x'(n-i) \quad (3)$$

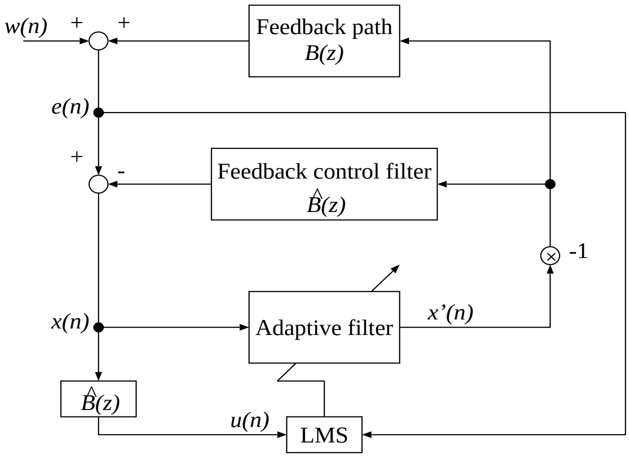

Fig. 2. Block diagram of the adaptive filter for feedback-type ANC

where N is the number of the tap coefficients $\{h_i(n)\}$ of the adaptive filter, $N_b - 1$ is the order of the transfer functions of $B(z)$ and $\hat{B}(z)$ whose impulse responses are $\{b_i\}$ and $\{\hat{b}_i\}$, respectively. We also assume that $N_b \ll N$. Then the error signal $e(n)$ is given by

$$e(n) = w(n) - \sum_{i=0}^{N_b-1} b_i \sum_{k=0}^{N-1} h_k(n-i)x(n-i-k). \quad (4)$$

Since the tap weight $h_i(n)$ $(i = 0, 1, \cdots, N-1)$ is updated by the filtered-X LMS algorithm

$$h_i(n+1) = h_i(n) + \mu u(n-i)e(n) \quad (5)$$

with a small positive gain μ, the difference between $h_k(n)$ and $h_k(n-i)$ is of $O(\mu)$. So its effect through $e(n)$ in (5) is of $O(\mu^2)$ and can be discarded. Hence, the second term in (4) can be regarded as the output of the cascade filter of the adaptive filter with the fixed $\{h_i(n)\}$ and the feedback path where the input is $x(n)$. So $e(n)$ is approximately expressed as

$$e(n) \simeq w(n) - \sum_{l=0}^{N-1} \left(\sum_{i=0}^{N_b-1} b_i h_{l-i}(n) \right) x(n-l). \quad (6)$$

Next, we define the following L-dimensional vectors as

$$\begin{aligned} \boldsymbol{h}(n) &= [\ h_0(n),\ \cdots,\ h_{N-1}(n),\ \mathbf{0}^T\ ]^T \\ \boldsymbol{x}(n) &= [\ x(n),\ \cdots,\ x(n-N+1),\ \mathbf{0}^T\ ]^T \end{aligned} \quad (7)$$

where "$\mathbf{0}$" denotes the $(L-N)$-dimensional zero vector and $L \geq 2N$. The reason why $L - N$ zeros are padded in (7) is that the N-point DFT of $\{h_i(n)\}$ $(i = 0, \cdots, N-1)$ results in the N-period sequence in the discrete frequency domain and this in turn introduces the N-periodicity in $\{h_i(n)\}$ through the inverse DFT. Thus $h_i(n)$ for negative i becomes non-zero. To avoid this and insure the causality zero padding is introduced [6]. Similarly we define the L-dimensional vectors for other signals and the vectors $\boldsymbol{b}$, $\hat{\boldsymbol{b}}$ with $(L - N_b)$ zeros padded for tap coefficients of $B(z)$ and $\hat{B}(z)$, respectively. Then (6) can be written as

$$e(n) \simeq w(n) - (\boldsymbol{b} \otimes \boldsymbol{h}(n))^\dagger \boldsymbol{x}(n) \quad (8)$$

where "$\dagger$" and "$\otimes$" denote the complex conjugate transpose and the convolution, respectively. Also, $(\boldsymbol{b} \otimes \boldsymbol{h}(n))$ is made L-dimensional by deleting the extra zeros. Then the adaptive rule can be written as

$$\boldsymbol{h}(n+1) = \boldsymbol{h}(n) + \mu \boldsymbol{u}(n)e(n) \quad (9)$$

Also, we define the L-point DFT matrix by

$$\boldsymbol{F} = \left[\exp\left(-\mathrm{j}\frac{2\pi lm}{L} \right) \right] \quad l, m = 0, 1, \cdots, L-1$$

and the L-point DFT of $\boldsymbol{w}$, $\boldsymbol{e}$, $\boldsymbol{x}$, $\boldsymbol{x}'$, $\boldsymbol{u}$, $\boldsymbol{h}$, $\boldsymbol{b}$, $\hat{\boldsymbol{b}}$ are denoted by the corresponding capital letters as $\boldsymbol{W}$, $\boldsymbol{E}$, $\boldsymbol{X}$, $\boldsymbol{X}'$, $\boldsymbol{U}$, $\boldsymbol{H}$, $\boldsymbol{B}$, $\hat{\boldsymbol{B}}$, respectively. Also the following diagonal matrix is defined for $\boldsymbol{H}(n) = (\ H_0(n),\ H_1(n),\ \cdots,\ H_{L-1}(n)\ )^T$ as

$$\boldsymbol{\Lambda}_{\boldsymbol{H}(n)} = \mathrm{diag}[\ H_0(n),\ H_1(n),\ \cdots,\ H_{L-1}(n)]$$

and similarly for $\boldsymbol{B}$, $\hat{\boldsymbol{B}}$ as $\boldsymbol{\Lambda}_B$, $\boldsymbol{\Lambda}_{\hat{B}}$. Noting that

$$\boldsymbol{F}^\dagger \boldsymbol{F} = L\boldsymbol{I}$$

where $\boldsymbol{I}$ denotes the $L \times L$ identity matrix and using this in (8) we have

$$\begin{aligned} e(n) &\simeq w(n) - \frac{1}{L}\boldsymbol{x}^\dagger(n)\boldsymbol{F}^\dagger \boldsymbol{F}(\boldsymbol{b} * \boldsymbol{h}(n)) \\ &\simeq w(n) - \frac{1}{L}\boldsymbol{X}^\dagger(n)\boldsymbol{\Lambda}_{\boldsymbol{B}}\boldsymbol{H}(n) \end{aligned} \quad (10)$$

Applying $\boldsymbol{F}$ to (9) and using (10), we have

$$\boldsymbol{H}(n+1) = \boldsymbol{H}(n) + \mu \boldsymbol{U}(n) \left[w(n) - \frac{1}{L}\boldsymbol{X}^\dagger(n)\boldsymbol{\Lambda}_B \boldsymbol{H}(n) \right]. \quad (11)$$

Since

$$u(n) = \sum_{i=0}^{N_b-1} \hat{b}_i x(n-i),$$

the l-th element of $\boldsymbol{U}(n)$ can be written as

$$(\boldsymbol{U}(n))_l = \sum_{i=0}^{N_b-1} \hat{b}_i \mathrm{e}^{\mathrm{j}\frac{2\pi li}{L}} \sum_{k=i}^{L-1+i} x(n-k)\mathrm{e}^{-\mathrm{j}\frac{2\pi lk}{L}}.$$

But the index i moves from 0 to $N_b - 1$ and $N_b \ll L$ so that we can replace the range of the second summation with $0 \leq k \leq L-1$ by neglecting the "end effects". Thus we have the approximate relation

$$\boldsymbol{U}(n) \simeq \boldsymbol{\Lambda}_{\hat{B}}^* \boldsymbol{X}(n), \quad (12)$$

where "$*$" denotes the complex conjugate. Similarly the corresponding approximate expressions for (1)–(3) are

$$\begin{aligned} \boldsymbol{E}(n) &\simeq \boldsymbol{W}(n) - \boldsymbol{\Lambda}_B^* \boldsymbol{X}'(n) \\ \boldsymbol{X}'(n) &\simeq \boldsymbol{\Lambda}_{H(n)}^* \boldsymbol{X}(n) \\ \boldsymbol{X}(n) &\simeq \boldsymbol{E}(n) + \boldsymbol{\Lambda}_{\hat{B}}^* \boldsymbol{X}'(n). \end{aligned} \quad (13)$$

Hence by eliminating $\boldsymbol{E}(n)$ and $\boldsymbol{X}'(n)$ in (13) we have

$$\boldsymbol{X}(n) \simeq \boldsymbol{Q}(n)\boldsymbol{W}(n) \tag{14}$$

with

$$\boldsymbol{Q}(n) = \left[\ \boldsymbol{I} + \boldsymbol{\Lambda}^*_{H(n)}\left(\boldsymbol{\Lambda}^*_B - \boldsymbol{\Lambda}^*_{\hat{B}}\right)\ \right]^{-1}. \tag{15}$$

Substituting (12) and (14) into (11) we have the discrete frequency domain expression of (9) as

$$\begin{aligned}\boldsymbol{H}(n+1) \simeq \boldsymbol{H}(n) + \mu &\left[\boldsymbol{\Lambda}^*_{\hat{B}}\boldsymbol{Q}(n)\boldsymbol{W}(n)w(n)\right.\\ &\left.-\frac{1}{L}\boldsymbol{\Lambda}^*_{\hat{B}}\boldsymbol{Q}(n)\boldsymbol{W}(n)\boldsymbol{W}^\dagger(n)\boldsymbol{Q}^\dagger(n)\boldsymbol{\Lambda}_B\boldsymbol{H}(n)\right]. \end{aligned}\tag{16}$$

Since L is large and $w(n)$ is a zero-mean stationary process, the element of $\boldsymbol{W}(n)$, that is, the DFT of $w(n)$ is uncorrelated with each other. Hence,

$$\mathrm{E}\left[\boldsymbol{W}(n)\boldsymbol{W}^\dagger(n)\right] \simeq L\mathrm{diag}\left[\ S_0,\ S_1,\ \cdots,\ S_{N-1}\ \right] \equiv L\boldsymbol{\Lambda}_S \tag{17}$$

where $S(\mathrm{e}^{\mathrm{j}\omega})$ is the spectral density of $w(n)$ and $S_l = S(\mathrm{e}^{\mathrm{j}\frac{2\pi l}{L}})$. Also,

$$w(n) = \frac{1}{L}\sum_{l=0}^{L-1}\sum_{k=0}^{L-1} w(n-k)\mathrm{e}^{\mathrm{j}\frac{2\pi l}{L}k} = \frac{1}{L}\boldsymbol{W}^\dagger(n)\boldsymbol{\pi} \tag{18}$$

where $\boldsymbol{\pi}$ is an L-dimensional vector whose elements are all 1. So from (17) and (18)

$$\mathrm{E}\left[\boldsymbol{W}(n)w(n)\right] \simeq \boldsymbol{\Lambda}_S\boldsymbol{\pi} = \boldsymbol{S}. \tag{19}$$

We use the averaging method in [3] to analyze (16). By taking the average with respect to $\boldsymbol{W}(n)$ and $w(n)$ in the right hand side of (16), replacing $\boldsymbol{H}(n)$ with the corresponding deterministic quantity $\bar{\boldsymbol{H}}(n)$ and using (17), (19), the averaged system is given by

$$\begin{aligned}\bar{\boldsymbol{H}}(n+1) = \bar{\boldsymbol{H}}(n) + \mu &\left[\boldsymbol{\Lambda}^*_{\hat{B}}\bar{\boldsymbol{Q}}(n)\boldsymbol{S}\right.\\ &\left.-\boldsymbol{\Lambda}^*_{\hat{B}}\bar{\boldsymbol{Q}}(n)\boldsymbol{\Lambda}_S\bar{\boldsymbol{Q}}^\dagger(n)\boldsymbol{\Lambda}_B\bar{\boldsymbol{H}}(n)\right]_+\end{aligned}\tag{20}$$

where $\bar{\boldsymbol{Q}}(n)$ is given by (15) with $\boldsymbol{H}(n)$ replaced by $\bar{\boldsymbol{H}}(n)$ and $[\ ]_+$ indicates that the causal part is taken from the inverse transform of the quantity in a square bracket [6]. This operation is necessary to keep $\bar{\boldsymbol{H}}(n)$ to be causal. Since all the matrices in (20) are diagonal, the l-th element of (20) is written as the following scalar nonlinear difference equation

$$\begin{aligned}&\bar{H}_l(n+1) = \bar{H}_l(n) + \\ &\mu\left[\frac{\hat{B}^*_l S_l}{1+\bar{H}_l{}^*(n)\epsilon^*_l} - \frac{\hat{B}^*_l S_l B_l \bar{H}_l(n)}{\left(1+\bar{H}_l(n)\epsilon_l\right)\left(1+\bar{H}_l{}^*(n)\epsilon^*_l\right)}\right]_+\end{aligned}\tag{21}$$

where $B_l = B(\mathrm{e}^{\mathrm{j}\frac{2\pi l}{L}})$, $\hat{B}_l = \hat{B}(\mathrm{e}^{\mathrm{j}\frac{2\pi l}{L}})$, and $\epsilon_l = B_l - \hat{B}_l$.

Thus the stationary point of the original filtered-X LMS algorithm in (5) is obtained by solving (21) with $\bar{H}_l(n+1) = \bar{H}_l(n) = H_l$. When $N \to \infty$ $(L \to \infty)$, we can replace the discrete frequencies with the continuous ones so that instead of H_l we use $H(z)$ where $z = \mathrm{e}^{\mathrm{j}\omega}$. Hence, it follows that the stationary point is determined by

$$\left[\hat{B}(z^{-1})S(z) - \frac{\hat{B}(z^{-1})S(z)B(z)H(z)}{1+H(z)\epsilon(z)}\right]_+ = 0 \tag{22}$$

where $\epsilon(z) = B(z) - \hat{B}(z)$ and $1 + H(z^{-1})\epsilon(z^{-1})$ is purely non-causal except the constant term so that from the denominator we can get rid of this. In general, it is very difficult to solve this "generalized Wiener-Hopf" equation. Let the spectral factorization of $S(z)$ be $S(z) = G(z)G(z^{-1})$ where $G(z)$ is of minimum phase. Then $G(z^{-1})$ can be factored out from the left hand side of (22).

Here we present two cases where we can have explicit solutions. The first case is that $B(z) = \hat{B}(z)$ $(\epsilon(z) = 0)$. Then we immediately have the solution

$$H_{\mathrm{opt}}(z) = \frac{1}{B_{\min}(z)G(z)}\left[\frac{B(z^{-1})G(z)}{B_{\min}(z^{-1})}\right]_+$$

where $B(z)B(z^{-1}) = B_{\min}(z)B_{\min}(z^{-1})$ and $B_{\min}(z)$ is a stable polynomial. A more interesting case is for $\epsilon(z) \neq 0$. Assume that $B(z) = z^{-d}C(z)$, $\hat{B}(z) = z^{-d}\hat{C}(z)$ where d is a positive integer denoting the delay and $C(z)$, $\hat{C}(z)$ are stable polynomials. Further assume for the moment that $1 + H(z)\epsilon(z)$ is of minimum phase. Then (22) can be simplified as

$$\left[z^d G(z)\right]_+ - \frac{G(z)C(z)H(z)}{1+H(z)\epsilon(z)} = 0$$

so that if we set $A(z) = \left[z^d G(z)\right]_+ / G(z)$, we have the solution

$$H_{\mathrm{opt}}(z) = \frac{A(z)}{C(z) - z^{-d}(C(z) - \hat{C}(z))A(z)}, \tag{23}$$

provided that this is a stable transfer function. For (23) $1 + H(z)\epsilon(z)$ is of minimum phase. We also note that $A(z)$ is the transfer function of the optimal d-step ahead linear predictor of $w(n+d)$.

3. THE CONVERGENCE CONDITION OF THE ADAPTIVE ALGORITHM

The local stability of (21) around the stationary point is examined by calculating the derivative of (21). We use a special definition of the derivative with respect to a complex variable in [7] where we note that the following property $\partial\bar{H}^*_l(n)/\partial\bar{H}_l(n) = 0$ holds. Since $\bar{\boldsymbol{H}}(n+1)$ and $\bar{\boldsymbol{H}}(n)$ in (20) are causal, we can get rid of the operation $[\ ]_+$ from the right hand side of (20) by adding some purely noncausal vector which may be dependent on $\bar{\boldsymbol{H}}^*(n)$. Hence in calculating the derivative we discard the operation $[\ ]_+$ in (21) and obtain

$$\frac{\partial\bar{H}_l(n+1)}{\partial\bar{H}_l(n)} = 1 - \frac{\mu\hat{B}^*_l S_l B_l}{(1+\bar{H}_l(n)\epsilon_l)^2(1+\bar{H}^*_l(n)\epsilon^*_l)}. \tag{24}$$

Substituting the stationary point (23) of the second case we have

$$\left.\frac{\partial\bar{H}_l(n+1)}{\partial\bar{H}_l(n)}\right|_{\bar{H}_l(n)=H_{\mathrm{opt},l}} = 1 - \frac{\mu S_l\hat{C}^*_l\left|C_l - A_l\epsilon_l\right|^2(C_l - A_l\epsilon_l)}{|C_l|^2}. \tag{25}$$

For the (local) stability we require that the absolute value of the right hand side of (25) is less than 1. Since $0 < \mu \ll 1$ and $S_l > 0$, we have the condition

$$\mathrm{Re}\left[\hat{C}^*_l(C_l - A_l\epsilon_l)\right] > 0 \tag{26}$$

where $\epsilon_l = \mathrm{e}^{-\mathrm{j}\frac{2\pi ld}{L}}(C_l - \hat{C}_l)$. This is a counter-part of the so-called 90° condition in the feedforward-type ANC . However in the feedback-type ANC the condition (26) depends on the property of $w(n)$. Under (26) the transfer function of the adaptive filter converges to $H_{\rm opt}(z)$ in (23) at least locally for large N and in this steady state from (1), (2), (3) and (23) the error signal $e(n)$ is symbolically expressed as

$$\begin{aligned} e(n) &= \frac{1 - \hat{B}(z)H_{\rm opt}(z)}{1 + (B(z) - \hat{B}(z))H_{\rm opt}(z)} w(n) \\ &= (1 - z^{-d}A(z))w(n) \end{aligned}$$

where z^{-1} means the unit delay operator. This is the minimum variance d-step ahead prediction error of $w(n)$. That is, even if $\hat{B}(z)$ differs from $B(z)$ within some range, the scheme in Fig. 2 gives the optimal steady state performance.

4. A SIMPLE EXAMPLE

Here we consider a simple case where $B(z) = \gamma z^{-1}$, $\hat{B}(z) = \hat{\gamma} z^{-1}$ and $w(n)$ is a first order (lowpass) AR process with the innovation variance 1, that is, $G(z) = (1 - az^{-1})^{-1}$ with $0 < a < 1$. Then (23) becomes

$$H_{\rm opt}(z) = \frac{a}{\gamma\{1 - z^{-1}(1-\beta)a\}} \quad (27)$$

with $\beta = \hat{\gamma}/\gamma$ and the corresponding impulse response

$$h_{{\rm opt},l} = a^{l+1}(1-\beta)^l/\gamma \quad (l \geq 0). \quad (28)$$

The condition that (27) is the stationary point of the adaptive algorithm is $|(1-\beta)a| < 1$, that is, $1 - 1/a < \beta < 1 + 1/a$. Under this condition we consider (26) in this case, that is,

$$\beta\left\{1 - (1-\beta)a\cos\frac{2\pi l}{L}\right\} > 0 \quad (l = 0, \cdots, L-1).$$

But the quantity in the bracket is positive so that we have the overall local convergence condition as

$$0 < \beta < 1 + \frac{1}{a}. \quad (29)$$

Some simulations have been made to check the theoretical findings. In Fig.3 the learning curve showing the empirical variance of the squared error $e^2(n)$ is presented for the case where $a = 0.9$, $\gamma = 1$, $\hat{\gamma} = 0.5$ ($\beta = 0.5$) with $N = 8$, and $\mu = 0.01$. The variance is obtained by averaging over 50 data sets and the steady state variance is 1.02879 which is very close to the minimum variance 1.0. In Table 1, the steady state first 4 impulse responses of the adaptive filter are presented together with the theoretical ones in (28). The agreements are good. Finally we have observed that for $\beta = -0.05$ and 2.2 the adaptive algorithm diverges. This coincide well with (29).

5. CONCLUSION

We have presented the analysis of the adaptive filter algorithm for feedback-type ANC concerning its stationary point and the local convergence condition using the averaging method combined with the frequency domain expression of the adaptive algorithm. The obtained theoretical results coincide well with the simulation results. A further study is needed about the property of the generalized Wiener-Hopf equation describing the stationary point.

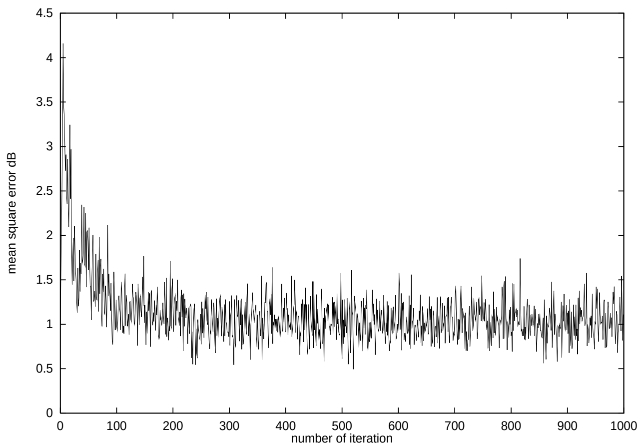

Fig. 3. Learning curve under the condition $a = 0.9$, $\gamma = 1$ and $\hat{\gamma} = 0.5$.

	h_0	h_1	h_2	h_3
empirical	0.870165	0.414190	0.201025	0.093168
theoretical	0.900000	0.405000	0.182250	0.082013

Table 1. Theoretical and estimated impulse responses.

6. REFERENCES

[1] S. J. Elliott and P. A. Nelson, "Active noise control," *IEEE Signal Processing Magazine*, pp.12-35, Oct. 1993.

[2] S. M. Kuo and D. R. Morgan, *Active Noise Control Systems*, John Wiley & Sons Inc., 1996.

[3] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms –Stability and Performance–*, Prentice Hall,1995.

[4] S. Miyagi and H. Sakai, "Convergence analysis of the delayless subband adaptive filter based on the frequency domain expression", *Proc. ICASSP*, pp.368–371, Istanbul, June 2000.

[5] K. Fujii, "Filter structure for the feedback control of the noise," *Proc. Conf. Acoust. Soc. Japan*, pp. 513–516, March 1999. (in Japanese)

[6] S. J. Elliott and B. Rafaely, "Frequency-domain adaptation of causal digital filters," *IEEE Trans. Signal Processing*, pp.1354-1364, May 2000.

[7] S. Haykin, *Adaptive Filter Theory*, Third Edition, Prentice Hall, 1996.