ENHANCEMENT OF NOISY SPEECH BY FORWARD/BACKWARD ADAPTIVE DIGITAL FILTERING

J. W. Kim and C. K. Un

Communications Research Laboratory
Department of Electrical Engineering
Korea Advanced Institute of Science and Technology
P.O. Box 150, Chongyangni
Seoul, Korea

ABSTRACT

In this paper, a forward/backward adaptive digital filtering method is studied for the enhancement of noisy speech. The enhancing algorithm uses future samples as well as previous samples to estimate a current sample, and utilizes the correlational property of speech. It is effective for the enhancement of narrow-band as well as wide-band noisy speech. To improve the enhancing capability further, its modified version is also considered. This modified algorithm, of which structure is relatively simple, works for any kind of additive noise, and results in signal-to-noise ratio gain over existing enhancement algorithms by about 2 dB.

I. INTRODUCTION

In many cases, speech signal is liable to be corrupted by background noise which may be of wide-band, narrow-band, or both. To clean up this noise from the corrupted speech, various algorithms for the enhancement of noisy speech have been studied [1]-[9]. These algorithms include several classes of algorithms which are based on short-time spectral estimation of speech [2], parameter estimation of a speech model from degraded speech [3]-[5], and adaptive filtering [6]-[9].

In this work, we study the enhancement of noisy speech by forward/backward adaptive digital filtering, and compare the results to those of other algorithms such as the Wiener filtering method [1] and the adaptive digital filtering method proposed by Sambur [7]. This method utilizes the property of short time autocorrelation of speech. Estimation of a correlated signal can be done by previous or future samples, or both. Estimation by previous samples is called 'forward prediction', and estimation by future samples is called 'backward prediction'. In both cases, the predictor yields almost the same performance. However, the former results in a causal system, and the latter a noncausal system. In this work, we consider forward/backward prediction. With this scheme, one can expect more accurate prediction of a correlated signal than with either forward or backward prediction scheme alone, but the resulting system is noncausal. Nevertheless, this noncausal prediction system may be used for the enhancement of noisy speech.

II. ADAPTIVE DIGITAL FILTERING ALGORITHM BASED ON FORWARD/BACKWARD PREDICTION

A block diagram of the forward/backward adaptive digital filter (F/B ADF) is shown in Fig. 1.

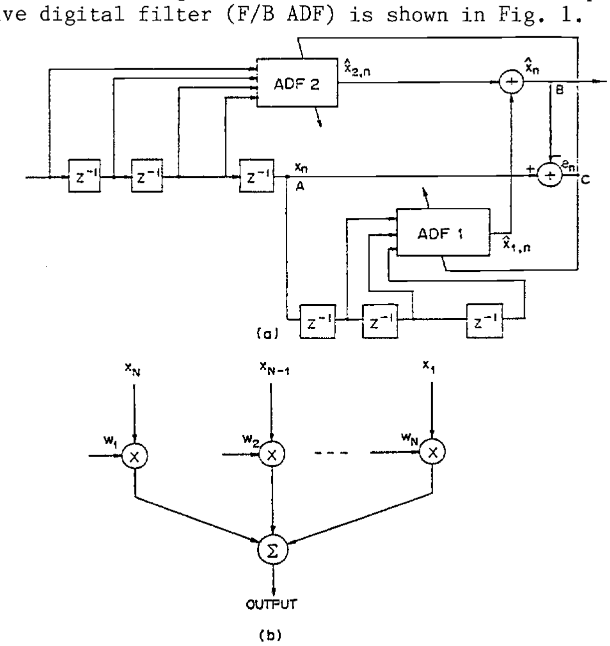

Fig. 1 (a) Block diagram of a forward/backward ADF. (b) The structure of an ADF.

The forward adaptive digital filter, ADF1, has M weighting coefficients $\{W_{1,1}, W_{1,2}, \ldots, W_{1,M}\}$, and the backward adaptive digital filter, ADF2, has N weighting coefficients $\{W_{2,1}, W_{2,2}, \ldots, W_{2,N}\}$. These coefficients are adjusted to minimize the mean-square error (MSE) defined as

$$\varepsilon \triangleq E[e_n^2] \tag{1}$$

where e_n is the difference signal between the reference input x_n and the filter output $\hat{x}_n$ which are assumed to be real, that is,

$$e_n = x_n - \hat{x}_n. \tag{2}$$

The filter output sinal $\hat{x}_n$ is the sum of $\hat{x}_{1,n}$ and $\hat{x}_{2,n}$ which are the outputs of the two filters, ADF1 and ADF2, respectively. The two outputs of ADF1 and ADF2, $\hat{x}_{1,n}$ and $\hat{x}_{2,n}$, are defined, respectively, as

$$\hat{x}_{1,n} \triangleq W_1^T X_1 = X_1^T W_1$$

and

$$\hat{x}_{2,n} \triangleq W_1^T X_2 = X_2^T W_2 \tag{3}$$

3. 3. 1

CH2243-4/86/0000-0089 $1.00

where 'T' denotes the transpose of a vector or matrix, and the weight vectors and input vectors are given, respectively, as

$$\mathbf{W}_1 = [W_{1,1}, W_{1,2}, \cdots, W_{1,M}]^T,$$

$$\mathbf{W}_2 = [W_{2,1}, W_{2,2}, \cdots, W_{2,N}]^T,$$

$$\mathbf{X}_1 = [x_{n-1}, x_{n-2}, \cdots, x_{n-M}]^T,$$

and $\mathbf{X}_2 = [x_{n+1}, x_{n+2}, \cdots, x_{n+N}]^T.$

From (1) and (2), we have the MSE ε as

$$\varepsilon = E[x_n^2] - 2\mathbf{W}_1^T\mathbf{P}_{x1} - 2\mathbf{W}_2^T\mathbf{P}_{x2} + \mathbf{W}_1^T\mathbf{R}_{x1,x2}\mathbf{W}_2 + \mathbf{W}_2^T\mathbf{R}_{x2,x1}\mathbf{W} + \mathbf{W}_1^T\mathbf{R}_{x1,x1}\mathbf{W}_1 + \mathbf{W}_2^T\mathbf{R}_{x2,x2}\mathbf{W}_2 \quad (4)$$

where the autocorrelation matrices, $\mathbf{R}_{x1,x2}$, $\mathbf{R}_{x2,x1}$, are defined, respectively, as

$$\mathbf{R}_{x1,x1} \triangleq E[\mathbf{X}_1\mathbf{X}_1^T]$$

and $\mathbf{R}_{x2,x2} \triangleq E[\mathbf{X}_2\mathbf{X}_2^T],$

and the cross-correlation matrices, $\mathbf{R}_{x1,x2}$, $\mathbf{R}_{x2,x1}$, $\mathbf{P}_{x1}$, and $\mathbf{P}_{x2}$ are defined, respectively, as

$$\mathbf{R}_{x1,x2} \triangleq E[\mathbf{X}_1\mathbf{X}_2^T]$$

$$\mathbf{R}_{x2,x1} \triangleq E[\mathbf{X}_2\mathbf{X}_1^T]$$

$$\mathbf{P}_{x1} \triangleq E[x_n\mathbf{X}_1]$$

and $\mathbf{P}_{x2} \triangleq E[x_n\mathbf{X}_2]. \quad (5)$

We note that the cross-correlation matrices, $\mathbf{R}_{x1,x2}$ and $\mathbf{R}_{x2,x1}$, satisfy the relation

$$\mathbf{R}_{x1,x2}^T = \mathbf{R}_{x2,x1}. \quad (6)$$

We can see from (4) that the MSE ε is an elliptic parabolic function of coefficients $\mathbf{W}_1$ and $\mathbf{W}_2$, and thus it has a unique global minimum for some $\mathbf{W}_1$ and $\mathbf{W}_2$. Therefore, the optimum $\mathbf{W}_{1,OPT}$ and $\mathbf{W}_{2,OPT}$ that minimize the MSE can be obtained. By differentiating ε (given by (4)) with respect to $\mathbf{W}_1$ and $\mathbf{W}_2$, we have the following equations:

$$\nabla\varepsilon(\mathbf{W}_1) \triangleq \frac{\partial\varepsilon}{\partial\mathbf{W}_1} = -2(\mathbf{P}_{x1}-\mathbf{R}_{x1,x2}\mathbf{W}_2-\mathbf{R}_{x1,x1}\mathbf{W}_1) \quad (7)$$

and $\nabla\varepsilon(\mathbf{W}_2) \triangleq \frac{\partial\varepsilon}{\partial\mathbf{W}_2} = -2(\mathbf{P}_{x2}-\mathbf{R}_{x2,x1}\mathbf{W}_1-\mathbf{R}_{x2,x2}\mathbf{W}_2). \quad (8)$

Setting these to zero, we obtain the optimum coefficients vectors $\mathbf{W}_{1,OPT}$ and $\mathbf{W}_{2,OPT}$ as

$$\mathbf{W}_{1,OPT} = \mathbf{R}_{x1,x1}^{-1}(\mathbf{P}_{x1}-\mathbf{R}_{x1,x2}\mathbf{W}_{2,OPT}) \quad (9)$$

and $\mathbf{W}_{2,OPT} = \mathbf{R}_{x2,x2}^{-1}(\mathbf{P}_{x2}-\mathbf{R}_{x2,x1}\mathbf{W}_{1,OPT}). \quad (10)$

Alternatively, (9) and (10) may be written as

$$\mathbf{W}_{1,OPT} = [\mathbf{R}_{x1,x1}-\mathbf{R}_{x1,x2}\mathbf{R}_{x2,x2}^{-1}\mathbf{R}_{x2,x1}]^{-1} \cdot[\mathbf{P}_{x1}-\mathbf{R}_{x1,x2}\mathbf{R}_{x2,x2}^{-1}\mathbf{P}_{x2}] \quad (11)$$

and $\mathbf{W}_{2,OPT} = [\mathbf{R}_{x2,x2}-\mathbf{R}_{x2,x1}\mathbf{R}_{x1,x1}^{-1}\mathbf{R}_{x1,x2}]^{-1} \cdot[\mathbf{P}_{x2}-\mathbf{R}_{x2,x1}\mathbf{R}_{x1,x1}^{-1}\mathbf{P}_{x1}]. \quad (12)$

Substituting (11) and (12) into (4) yields the minimum MSE $\varepsilon_{min,FB}$ of the F/B ADF as

$$\begin{aligned}\varepsilon_{min,FB} = E[x_n^2] &- \{[\mathbf{R}_{x1,x1}-\mathbf{R}_{x1,x2}\mathbf{R}_{x2,x2}^{-1}\mathbf{R}_{x2,x1}]^{-1} \cdot[\mathbf{P}_{x1}-\mathbf{R}_{x1,x2}\mathbf{R}_{x2,x2}^{-1}\mathbf{P}_{x2}]\}^T\cdot\mathbf{P}_{x1} \\ &- \{[\mathbf{R}_{x2,x2}-\mathbf{R}_{x2,x2}\mathbf{R}_{x1,x1}^{-1}\mathbf{R}_{x1,x2}]^{-1} \cdot[\mathbf{P}_{x2}-\mathbf{R}_{x2,x1}\mathbf{R}_{x1,x1}^{-1}\mathbf{P}_{x1}]\}^T\cdot\mathbf{P}_{x2}.\end{aligned} \quad (13)$$

Since the transpose and inverse operators commute when the matrices are symmetric, (13) may be written as

$$\begin{aligned}\varepsilon_{min,FB} = E[x_n^2] &- [\mathbf{P}_{x1}^T-\mathbf{P}_{x2}^T\mathbf{R}_{x2,x2}^{-1}\mathbf{R}_{x2,x1}] \cdot[\mathbf{R}_{x1,x1}-\mathbf{R}_{x1,x2}\mathbf{R}_{x2,x2}^{-1}\mathbf{R}_{x2,x1}]^{-1}\mathbf{P}_{x1} \\ &- [\mathbf{P}_{x2}^T-\mathbf{P}_{x1}^T\mathbf{R}_{x1,x1}^{-1}\mathbf{R}_{x1,x2}] \cdot[\mathbf{R}_{x2,x2}-\mathbf{R}_{x2,x1}\mathbf{R}_{x1,x1}^{-1}\mathbf{R}_{x1,x2}]^{-1}\mathbf{P}_{x2}.\end{aligned} \quad (14)$$

Also, when there is a forward prediction filter only, its minimum MSE, $\varepsilon_{min,F}$, becomes

$$\varepsilon_{min,F} = E[x_n^2] - \mathbf{P}_{x1}^T\mathbf{R}_{x1,x1}^{-1}\mathbf{P}_{x1}. \quad (15)$$

It may be shown that the minimum MSE of the F/B ADF is always smaller than or equal to that of the forward prediction filter.

To obtain the coefficients $\mathbf{W}_1$ and $\mathbf{W}_2$ by the adaptive algorithm, we use the steepest descent method. The results are given as

$$\mathbf{W}_{1,n+1} = \mathbf{W}_{1,n} + \frac{\mu_1}{2}\{-\nabla\varepsilon(\mathbf{W}_{1,n})\} \quad (16)$$

and $\mathbf{W}_{2,n+1} = \mathbf{W}_{2,n} + \frac{\mu_2}{2}\{-\nabla\varepsilon(\mathbf{W}_{2,n})\} \quad (17)$

where μ_1 and μ_2 are convergence factors for ADF1 and ADF2, respectively. It may be shown that $\mathbf{W}_1$ converges to $\mathbf{W}_{1,OPT}$ if the following condition is satisfied [10]:

$$0 < \mu_1 < \frac{2}{\lambda_{1,max}} \quad (18)$$

where $\lambda_{1,max}$ is the maximum eigenvalue of $\mathbf{R}_{x1,x1}$. By the LMS algorithm of Widrow and Hoff, and from (7) and (16) the coefficient vector $\mathbf{W}_1$ is updated as

$$\mathbf{W}_{1,n+1} = \mathbf{W}_{1,n} + \mu_1\mathbf{X}_1e_n. \quad (19)$$

Similar results may be obtained with respect to $\mathbf{W}_2$. It may also be shown that $\mathbf{W}_2$ converges to $\mathbf{W}_{2,OPT}$ if the following condition is satisfied:

$$0 < \mu_2 < \frac{2}{\lambda_{2,max}} \quad (20)$$

where $\lambda_{2,max}$ is the maximum eigenvalue of $\mathbf{R}_{x2,x2}$. Also, $\mathbf{W}_2$ is updated as the following:

$$\mathbf{W}_{2,n+1} = \mathbf{W}_{2,n} + \mu_2\mathbf{X}_2e_n. \quad (21)$$

With the above results, we see that the F/B ADF can be used to estimate a correlated signal.

III. APPLICATION OF THE F/B ADF TO THE ENHANCEMENT OF NOISY SPEECH

We now show the applications of an F/B ADF and a modified F/B ADF to the enhancement of speech corrupted by wide-band or narrow-band noise. For the enhancement of noisy speech corrupted by white noise, we use an F/B ADF through which the noisy speech is filtered. It is noted that in this scheme, the requirements normally needed in many existing enhancement algorithms, such as the stationarity of noise, discrimination of silence from speech, or exact pitch information, are not needed.

The structure of the enhancing system for noisy speech corrupted by narrow-band noise is different from that for the wide-band noise case. Its block diagram is shown in Fig. 2. One can note that for the case of narrow-band noise, the enhancement system with the F/B ADF requires the knowledge of discrimination between silence and speech as in other enhancement algorithms. In Fig. 2 the speech detector decides whether the current block is silence or speech. If it is silence, the on/off switch activates the F/B ADF to adjust its coefficients. When a speech block is detected, the adjustment of filter coefficients stops. The enhanced speech is the difference signal between the input signal and the filter output signal. This scheme is similar to that introduced by Hoy et al. [8].

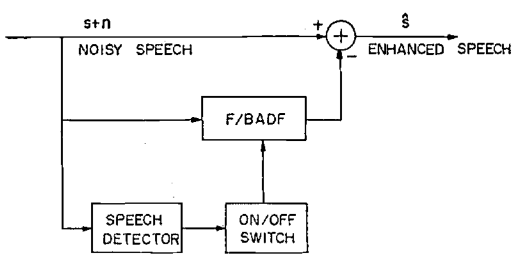

Fig. 2 Enhancement of narrow-band noisy speech with a forward/backward ADF.

Although the enhancement scheme using the F/B ADF is simple, and provides quite good results for wide-band noisy speech with low signal-to-noise ratio (SNR) and narrow-band noisy speech with any SNR level, it has limitations for real applications. Because of the limited prediction performance of the F/B ADF, it hardly enhances the wide-band noisy speech of 10 dB or highter in SNR. And the enhancement scheme of narrow-band noisy speech using the F/B ADF requires discrimination of silence from speech, a quite difficult job for noisy speech. Furthermore, in a real environment, noisy signals are often composed of both wide-band and narrow-band noise varying with time in magnitude and frequency bands. In this case, neither of the proposed schemes would work well. To overcome these difficulties, we introduce a modified F/B ADF which is shown in Fig. 3. Comparing Figs. 1 and 3, we see that the latter does not have 2 delayers. With this structure the modified F/B ADF is not a prediction filter. However, if we take its output as the predicted value of the reference signal, the modified F/B ADF can be considered to be a prediction filter. Since the prediction performance of the modified F/B ADF is much better than that of the F/B ADF, enhancement for wide-band noisy speech can be improved. Also, since the modified F/B ADF has a smoothing effect, it can

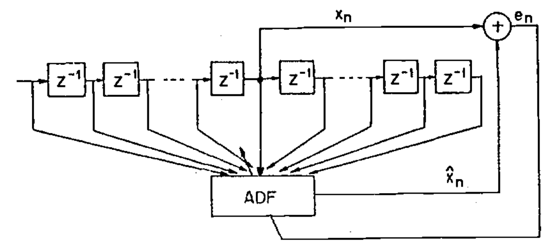

Fig. 3 Block diagram of a modified forward/backward ADF.

be used for the enhancement of narrow-band noisy speech with the same structure used for the enhancement of wide-band noisy speech. When silence and speech are discriminated, the modified F/B ADF can also be used for subtraction of narrow-band noise from noisy speech with improved results as compared to the F/B ADF with its structure shown in Fig. 2.

IV. SIMULATION RESULTS AND DISCUSSION

In this section we investigate the effectiveness of the F/B ADF and the modified F/B ADF by computer simulation. For computer simulations, we used real speech sampled at 8 kHz, and Gaussian random noise with zero mean for noise source. Wide-band noisy speech was generated by adding Gaussian random noise to the sampled speech. Narrow-band noisy speech was generated by adding the sampled speech and Gaussian random noise filtered by a band-pass filter. In our simulation, we set the values of convergence factors of both ADF's in the F/B ADF to be equal, and used the same numbers of coefficients (N=M) for both filters.

To show the effectiveness of the F/B ADF and the modified F/B ADF methods, we present in Table I the results for the enhancement of wide-band noisy speech. In this table the performances of the F/B ADF and modified F/B ADF schemes and that of the Wiener filtering method [1] with all-pole parameters of order 10 are shown. From these results, we see that the modified F/B ADF algorithm outperforms the F/B ADF or Wiener filter method regardless of the SNR level. Comparing the F/B ADF and Wiener filtering methods, the former performs better than the latter when SNR is low (i.e., 5 dB or less). When SNR of noisy speech is high (i.e., 10 dB or higher), however, the opposite is true. Nevertheless, considering the difficulty of discrimination of silence from speech, the nonstationarity of noise and the complexity required for the wiener filtering, the

Table I. Enhancement of wide-band noisy speech by Wiener filter, forward/backward ADF and modified forward/backward ADF.

Input SNR(dB) / Filter	0.71 (2.67)*	5.73 (4.91)	10.75 (7.81)
F/B ADF (M=N=8)	6.40 (5.41)	9.0 (7.2)	10.87 (8.9)
Modified F/B ADF (N=7)	6.5 (5.3)	10.41 (7.66)	13.4 (9.9)
Wiener filter	3.88 (4.08)	8.09 (6.27)	11.84 (8.68)

* The values in parentheses are SNR_{seg}(dB).

3. 3. 3

F/B ADF method appears to be preferable. Also, we have compared the performance of the proposed algorithm to that of Sambur's approach [7]. It has been found that the performance of Sambur's method is about the same as that of the F/B ADF. However, the requirement of accurate knowledge of pitch period in Sambur's method makes his method inferior.

In addition, we used the modified F/B ADF filtering method for the enhancement of both wide-band and narrow-band noisy speech with the same values of the convergence factor and the number of coefficients N being equal to 7. Tables II and III show the results for wide-band and narrow-band noisy speech, respectively. Figs. 4 and 5 show the output waveforms of this scheme applied to wide-band noisy speech and to narrow-band noisy speech, respectively. We see from these results that the modified F/B ADF method which has a relatively simple structure is effective for any kind of noise. Furthermore, it can also be used in speech waveform quantization to reduce granular noise that is not correlated with speech.

Table II. Enhancement of wide-band noisy speech by modified forward/backward ADF with constant convergence factor (5×10^{-6}).

Input SNR(dB)	0.71 (2.67)*	5.73 (4.92)	10.75 (7.81)
Output SNR(dB)	5.75 (4.94)	10.11 (7.47)	13.17 (9.84)

* The values in parentheses are SNR_{seg}(dB)

Table III. Enhancement of narrow-band noisy speech by modified forward/backward ADF with constant convergence factor (5×10^{-6}).

Input SNR(dB)	2.53 (3.37)*	6.17 (5.28)	10.63 (7.98)
Output SNR(dB)	8.73 (6.38)	12.79 (9.54)	14.92 (12.27)

* The values in parentheses are SNR_{seg}(dB)

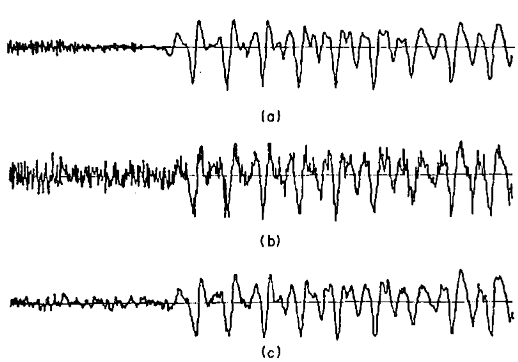

Fig. 4 Enhancement of wide-band noisy speech (5 dB) by a modified forward/backward ADF.
(a) Clean speech
(b) Noisy speech
(c) Enhanced speech

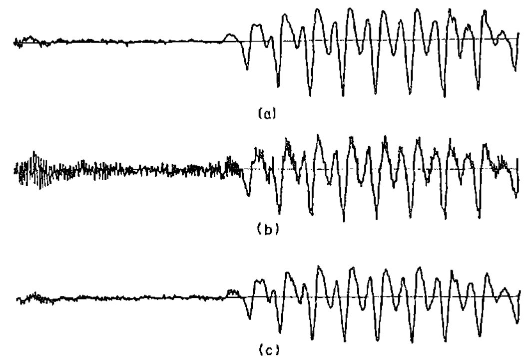

Fig. 5 Enhancement of narrow-band noisy speech (5 dB) by a modified forward/backward ADF.
(a) Clean speech
(b) Noisy speech
(c) Enhanced speech

REFERENCES

1. J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," Proc. IEEE, vol. 67, No. 12, Dec. 1979.
2. S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust., Speech, Signal Processsing, vol. ASSP-29, pp. 113-120, Apr. 1979.
3. J. S. Lim and A. V. Oppenheim, "All-pole modeling of degraded speech," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-26, pp. 197-210, June 1978.
4. K. Steiglitz, "On the simultaneous estimation of poles and zeros in speech analysis," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-25, pp. 229-234, June 1979.
5. K. H. Song and C. K. Un, "Pole-zero modeling of speech based on high-order pole model fitting and decomposition method," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-31, pp. 1556-1565, Dec. 1983.
6. B. Widrow et al., "Adaptive noise cancelling: Principles and applications," Proc. IEEE, vol. 63, pp. 1692-1716, Dec. 1975.
7. M. R. Sambur, "Adaptive noise cancelling for speech signals," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-26, pp. 419-423, Oct. 1978.
8. L. Hoy et al., "Noise suppression methods for speech applications," in Proc. 1983 IEEE Int. Acoust., Speech, Signal Processing, pp. 1133-1136, Apr. 1983.
9. L. W. Varner, T. A. Miller, and T. E. Eger, "A simple adaptive filtering technique for speech enhancement," in Proc. 1983 IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 1126-1128, Apr. 1983.
10. B. Widrow et al., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," Proc. IEEE, vol. 64, No. 8, pp. 1151-1162, Aug. 1976.

3. 3. 4