A NEW UNBIASED EQUATION ERROR ALGORITHM FOR IIR ADF AND ITS APPLICATION TO ALE

J. Okello †, Y. Kinugasa ††, Y. Itoh †, Y. Fukui †, M. Kobayashi †††

† Faculty of Engineering, Tottori University, Japan
4-101 Koyama-Minami, Tottori 680-8552 Japan
†† Department of Electrical Engineering, Matsue College, Japan
††† Faculty of Engineering, Chubu University, Japan

ABSTRACT

In this paper, a new online algorithm for updating equation error IIR ADF is proposed. The proposed algorithm, which involves maintaining a constant power of the desired signal, is independent of the white disturbance signal, and hence there is no bias in the coefficient's estimate of the ADF. We also provide the analysis and simulation results which verify this kind of performance. Application of the proposed algorithm to adaptive line enhancer (ALE) is also provided. When compared with the method which uses cascaded notch filter, we observe a considerable improvement in performance due to the complete elimination of effect of white noise under mean sense condition.

1. INTRODUCTION

In adaptive signal processing, finite impulse response (FIR) ADF has been widely used due to its simplicity and stable convergence characteristics, which are well known. However, FIR ADF has a shortcoming in systems where good performance can only be attained if the filter order is made very large. Such an occurrence is possible in acoustic echo cancellation applications. Consequently, there is has been a lot of interest in IIR ADF, which can achieve the same level of performance as the FIR type, but at an advantage of using only a fewer number of coefficients.

Several approaches have made in order to estimate an unknown system by the use of IIR ADF [1][2]. In the case of using the output error, it has been shown that there is a possibility of convergence to a local minimum, with no guarantee on system stability [2]. Based on the problems associated with the output error mode of identification, the equation error formulation of IIR ADF has been actively researched on [3]-[6]. In its simpler form, where the mean square of the equation error is directly minimized using a gradient based algorithm, estimated parameters contain a bias if there is a disturbance signal [1][2].

In this paper, a new gradient-based algorithm for an equation error type of system identification is proposed. The proposed method ensures that the power of the desired signal is always kept constant throughout the system identification. This condition then enables coefficient estimation without bias, if the disturbance signal is white. We begin with the general introduction to equation error type of IIR ADF. In section 3, we present the proposed algorithm, while in section 4 we propose a system for using the proposed algorithm in ALE. Finally, we provide the simulation results, which verify the performance of the system.

2. EQUATION ERROR IIR ADF

Consider an IIR ADF with the transfer function $H(z)$, which is given by

$$H(z) = \frac{\hat{B}(z)}{\hat{A}(z)} \tag{1}$$

where $B(z) = \hat{b}_0 + \hat{b}_1 z^{-1} + \ldots + \hat{b}_M z^{-M}$ and

$A(z) = \hat{a}_0 + \hat{a}_1 z^{-1} + \ldots + \hat{a}_M z^{-M}$.

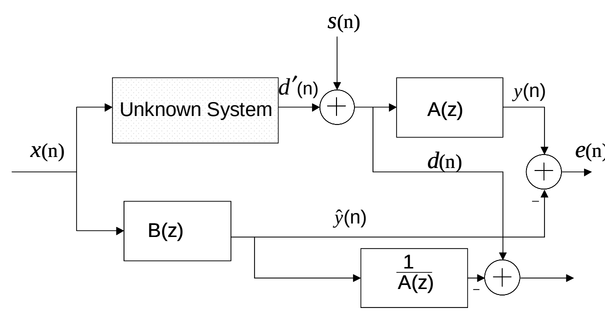

Fig.1. IIR system Identification model using the equation error formulation.

Without loss of generality we have assumed that the order of the numerator and denominator polynomial of $H(z)$ are the same. The desired signal, which has been corrupted by noise is given by

$$d(n) = d'(n) + s(n) \tag{2}$$

where $d'(n)$ is the output of the unknown system and $s(n)$ is the disturbance signal, which is assumed to be statistically independent from the input signal $x(n)$.

The disturbance signal is also assumed to be white with a variance of σ^2. The error signal used for updating the ADF based on the equation error approach is shown in Fig.1. It can also be expressed as

$$e(n) = \mathbf{d}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) - \mathbf{x}^{\mathsf{T}}(n)\hat{\mathbf{b}}(n)$$
$$= \mathbf{d}'^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) + \mathbf{s}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) - \mathbf{x}^{\mathsf{T}}(n)\hat{\mathbf{b}}(n) \qquad (3)$$

where, $\mathbf{d}(n) = [d(n)\ d(n-1)\ldots d(n-M)]^{\mathsf{T}}$

$\mathbf{x}(n) = [x(n)\ x(n-1)\ldots x(n-M)]^{\mathsf{T}}$,
$\mathbf{s}(n) = [s(n)\ s(n-1)\ldots s(n-M)]^{\mathsf{T}}$

$\hat{\mathbf{a}}(n) = [\hat{a}_0(n)\ \hat{a}_1(n)\ldots \hat{a}_M(n)]^{\mathsf{T}}$

$\hat{\mathbf{b}}(n) = [\hat{b}_0(n)\ \hat{b}_1(n)\ldots \hat{b}_M(n)]^{\mathsf{T}}$

The mean square of the estimation error $E[e^2(n)]$ is given by

$$E[e^2(n)] = E\left[\left(\mathbf{d}'^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) - \mathbf{x}^{\mathsf{T}}(n)\hat{\mathbf{b}}(n)\right)^2\right] + \hat{\mathbf{a}}^{\mathsf{T}}(n)E[\mathbf{s}(n)\mathbf{s}^{\mathsf{T}}(n)]\hat{\mathbf{a}}(n)$$
$$= E\left[\left(\mathbf{d}'^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) - \mathbf{x}^{\mathsf{T}}(n)\hat{\mathbf{b}}(n)\right)^2\right] + \sigma^2\hat{\mathbf{a}}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) \qquad (4)$$

This function is a unimodal function of both $\hat{\mathbf{a}}(n)$ and $\hat{\mathbf{b}}(n)$ [2]. However, when it is minimized directly using a gradient-based algorithm, the second term will contribute to the bias of parameter estimation.

Several methods have been proposed in order to eliminate the effect of the second term onto the parameter estimation. In one of the methods known as the unit norm, all the coefficients of are updated, while keeping the value of $\hat{\mathbf{a}}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n)$ to be unity. The offline algorithm, which results, essential eliminates the effect of noise albeit with complex computation [4].

Another approach would be to minimize $E[e^2(n)] - \lambda\hat{\mathbf{a}}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n)$, by indirectly estimating the variance of the noise such that $\lambda = \sigma^2$ [6].

3. PROPOSED ADAPTIVE ALGORITHM

In this paper we propose a new simple online algorithm, which ensures that the power of the desired signal is constant and independent of the coefficients of the ADF. The proposed algorithm is obtained by minimizing $\dfrac{E[e^2(n)]}{\hat{\mathbf{a}}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n)}$ and $E[e^2(n)]$ by updating $\hat{\mathbf{a}}(n)$ and $\hat{\mathbf{b}}(n)$, respectively, based on a gradient descent algorithm. The proposed adaptive algorithm is therefore be given,

$$\hat{\mathbf{a}}(n+1) = \hat{\mathbf{a}}(n) - \mu\frac{\partial}{\partial\hat{\mathbf{a}}(n)}\left\{\frac{e^2(n)}{\hat{\mathbf{a}}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n)}\right\}$$

$$= \hat{\mathbf{a}}(n) - 2\mu\left\{\mathbf{d}(n) - \frac{e(n)\hat{\mathbf{a}}(n)}{\|\hat{\mathbf{a}}(n)\|^2}\right\}\frac{e(n)}{\|\hat{\mathbf{a}}(n)\|^2}$$

$$\hat{\mathbf{b}}(n+1) = \hat{\mathbf{b}}(n) + \mu\mathbf{x}(n)e(n) \qquad (5)$$

where $\|\hat{\mathbf{a}}(n)\|^2 = \hat{\mathbf{a}}^T(n)\hat{\mathbf{a}}(n)$ and $\hat{a}_0(n) = 1.0$.

3.1 Effect of disturbance signal on convergence

In this section, we prove that the effect of a white disturbance signal does not contribute to the convergence of the coefficients in the mean sense. We shall assume that the disturbance noise $s(n)$ and the input signal $x(n)$ are statistically independent from each other. Thus,

$$E\left[\left\{\mathbf{d}(n) - \frac{e(n)\hat{\mathbf{a}}(n)}{\|\hat{\mathbf{a}}(n)\|^2}\right\}\frac{e(n)}{\|\hat{\mathbf{a}}(n)\|^2}\right]$$

$$= E\left[\left\{\mathbf{d}'(n) + \mathbf{s}(n) - \frac{\{\mathbf{s}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) + \mathbf{d}'^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) - \mathbf{x}^{\mathsf{T}}(n)\hat{\mathbf{b}}(n)\}}{\|\hat{\mathbf{a}}(n)\|^2}\right.\right.$$
$$\left.\left.\times\hat{\mathbf{a}}(n)\right\}\frac{\{\mathbf{s}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) + \mathbf{d}'^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) - \mathbf{x}^{\mathsf{T}}(n)\hat{\mathbf{b}}(n)\}}{\|\hat{\mathbf{a}}(n)\|^2}\right]$$

$$= E\left[\left\{\mathbf{d}'(n) - \frac{\{\mathbf{d}'^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) - \mathbf{x}^{\mathsf{T}}(n)\hat{\mathbf{b}}(n)\}\hat{\mathbf{a}}(n)}{\|\hat{\mathbf{a}}(n)\|^2}\right\}\right.$$
$$\left.\times\frac{\{\mathbf{d}'^{\mathsf{T}}(n)\hat{\mathbf{a}}(n) - \mathbf{x}^{\mathsf{T}}(n)\hat{\mathbf{b}}(n)\}}{\|\hat{\mathbf{a}}(n)\|^2}\right]$$
$$+ E\left[\left\{\mathbf{s}(n) - \frac{\mathbf{s}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n)\hat{\mathbf{a}}(n)}{\|\hat{\mathbf{a}}(n)\|^2}\right\}\frac{\mathbf{s}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n)}{\|\hat{\mathbf{a}}(n)\|^2}\right] \qquad (6)$$

The last term in Eq.(6) can be expressed as

$$\frac{\sigma^2\hat{\mathbf{a}}(n)}{\|\hat{\mathbf{a}}(n)\|^2} - \frac{\sigma^2\hat{\mathbf{a}}^{\mathsf{T}}(n)\hat{\mathbf{a}}(n)\hat{\mathbf{a}}(n)}{\|\hat{\mathbf{a}}(n)\|^4} = \mathbf{0} \qquad (7)$$

In other words, in the mean sense, the white disturbance signal does not have any effect in the convergence of the coefficients of the ADF.

3.2 Effect of normalization on the error surface

It is well known that the function $f(\hat{\mathbf{a}}(n)) = E[e_s^2(n)]$ is a unimodal function [2], where $e_s(n)$ is the equation error without the disturbance signal. For a given value of $\hat{\mathbf{b}}(n)$, say

$\hat{\mathbf{b}}(n) = \mathbf{b_1}$, the gradient of $f(\hat{\mathbf{a}}(n))$ has a single unique point at which its value is zero. At this particular point, the gradient of this function can be expressed as

$$\frac{\partial}{\partial \hat{\mathbf{a}}(n)} f(\hat{\mathbf{a}}(n)) = k\mathbf{R}(\hat{\mathbf{a}}(n) - \mathbf{a_1}) \tag{8}$$

where $\mathbf{a_1}$ is the value of $\hat{\mathbf{a}}(n)$ corresponding to the minimum point of the error surface for $\hat{\mathbf{b}}(n) = \mathbf{b_1}$, k is some constant and $\mathbf{R}$ is a symmetrical matrix whose Eigen vector is not $(\hat{\mathbf{a}}(n) - \mathbf{a_1})$ unless $(\hat{\mathbf{a}}(n) - \mathbf{a_1}) = 0$. Furthermore, since $f(\hat{\mathbf{a}}(n))$ can be considered to be a quadratic function of $\hat{\mathbf{a}}(n)$, then in the case of an exact modeled system we can consider that $f(\hat{\mathbf{a}}(n)) = k(\hat{\mathbf{a}}(n) - \mathbf{a_1})^T \mathbf{R}(\hat{\mathbf{a}}(n) - \mathbf{a_1})$, such that $f(\mathbf{a_1}) = 0$. Taking this into account, the gradient of the normalized equation error $f'(\hat{\mathbf{a}}(n)) = E[e^2(n)] / \|\hat{\mathbf{a}}(n)\|^2$ for the same value of $\hat{\mathbf{b}}(n) = \mathbf{b_1}$, will be given by

$$\begin{aligned}\frac{\partial}{\partial \hat{\mathbf{a}}(n)} f'(\hat{\mathbf{a}}(n)) &= \frac{2k\mathbf{R}(\hat{\mathbf{a}}(n) - \mathbf{a_1})}{\|\hat{\mathbf{a}}(n)\|^2} \\ &\quad - \frac{2k(\hat{\mathbf{a}}(n) - \mathbf{a_1})^T \mathbf{R}(\hat{\mathbf{a}}(n) - \mathbf{a_1})}{\|\hat{\mathbf{a}}(n)\|^4} \hat{\mathbf{a}}(n) \\ &= 2k\left\{\mathbf{I} - \frac{\hat{\mathbf{a}}(n)(\hat{\mathbf{a}}(n) - \mathbf{a_1})^T}{\|\hat{\mathbf{a}}(n)\|^2}\right\} \frac{\mathbf{R}(\hat{\mathbf{a}}(n) - \mathbf{a_1})}{\|\hat{\mathbf{a}}(n)\|^2} \\ &= 2k\left\{\mathbf{I} - \frac{\mathbf{M}}{\|\hat{\mathbf{a}}(n)\|^2}\right\} \frac{\mathbf{R}(\hat{\mathbf{a}}(n) - \mathbf{a_1})}{\|\hat{\mathbf{a}}(n)\|^2}\end{aligned} \tag{9}$$

From this equation, we can say that one solution of the minimum of the normalized equation error occurs at a point where $(\hat{\mathbf{a}}(n) - \mathbf{a_1}) = 0$. The other solution occurs at a point where $\mathbf{R}(\hat{\mathbf{a}}(n) - \mathbf{a_1})$ is an Eigen vector of matrix $\mathbf{M}$, with a corresponding Eigen value $\|\hat{\mathbf{a}}(n)\|^2$. However, since $\|\hat{\mathbf{a}}(n)\|^2$ is not an Eigen value of $\mathbf{M}$, it follows that $\mathbf{R}(\hat{\mathbf{a}}(n) - \mathbf{a_1})$ is not an Eigen vector of matrix $\mathbf{M}$, and $(\hat{\mathbf{a}}(n) - \mathbf{a_1}) = 0$. Thus the normalization of the equation error does not change the minimum of the error surface.

3.3 Application of the algorithm to adaptive line enhancer (ALE)

ALE has been used as way of enhancing sinusoidal signal, which has been mixed with noise [2][7]. This kind of system amounts to k^{th} step prediction, where proper choice of the value k is required [2]. When updated using LMS algorithm the effect of the white noise will however cause a frequency estimation bias. An alternative approach has been the use of the IIR adaptive notch filter. Unfortunately, IIR adaptive notch filter suffer from the problem of slow convergence especially if the notch bandwidth is very narrow.

In this section, we propose a new method of enhancing sinusoidal signals mixed with noise. The proposed method incorporates the algorithm, which has been proposed for the equation error IIR ADF. Figure 2 illustrates the proposed method, where in this case $x_s(n)$ is the sinusoidal signal, $s(n)$ is a white noise.

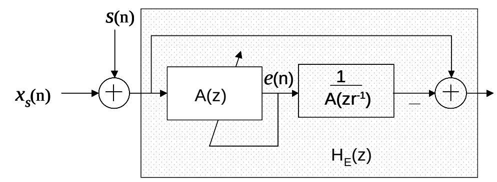

Fig.2 Proposed adaptive line enhancer.

The adaptive algorithm for the ALE is given by

$$\hat{a}_i(n+1) = \hat{a}_i(n) - 2\mu\left\{d(n-i) - \frac{e(n)\hat{a}_i(n)}{\|\hat{\mathbf{a}}(n)\|^2}\right\}\frac{e(n)}{\|\hat{\mathbf{a}}(n)\|^2} \tag{11}$$

for $i = 1, \ldots, M/2$, where the order M of filter $A(z)$ is even, $\hat{a}_0(n) = 1.0$ and $\hat{a}_i(n) = \hat{a}_{N-1-i}(n)$. With this kind of setting, the zeros of $A(z)$ will always be constrained to lie on the unit circle of the z-plane, such that $A(z)/A(zr^{-1})$ is a notch filter, with the normalized notch frequency $(f_{N1}, f_{N2}, \ldots)$ determined by the roots of $A(z)$, while the value of r determines the notch bandwidth.

4. SIMULATION AND RESULTS

In this section, we present the simulation results of the proposed equation error algorithm, and its application to ALE. In the case of IIR ADF formulated using the equation error, the unknown system was a fourth order moving average auto-regressive filter. An exact modeled ADF was considered with the signal to noise ratio (S/N) set to 10dB. S/N is defined as

$$S/N = 10\log_{10} E[d'^2(n)] / E[s^2(n)] \text{ dB} \tag{12}$$

The measure of performance was the echo return loss enhancement (ERLE), which is given by

$$ERLE = 10\log_{10} \frac{E[(\hat{\mathbf{a}}^T(n)\mathbf{d}'(n))^2]}{E[(\hat{\mathbf{a}}^T(n)\mathbf{d}'(n) - \hat{y}(n))^2]} \text{ (dB)} \tag{13}$$

Figure 3 shows the result obtained. In this figure, the proposed algorithm was compared with the algorithm in reference [6], and the equation error algorithm, which directly minimizes the equation error. From this result we observe that the proposed algorithm had roughly the same performance characteristic as the algorithm in [6], which also has roughly the same performance as [3].

Next, we investigated the performance of the proposed line enhancer. We considered an exact modeled system (*M*=4), where the sinusoid signals $(\cos(0.2\pi n)+\cos(0.7\pi n))$ were mixed with a white noise such that signal to noise ratio was 25dB. The result was compared with the cascaded notch filter, which was updated using the algorithm that minimizes the bias [8]. The value of r was set to 0.95. The measure of performance Ec was based on the mean square deviation of the notch frequency from those of the sinusoids.

$$Ec = 10\log_{10}\frac{0.2^2+0.7^2}{E\left[\left(f_{N1}-0.2\right)^2+\left(f_{N2}-0.7\right)^2\right]}\ \text{dB} \quad (14)$$

where after convergence, f_{N1} and f_{N2} correspond to 0.2 and 0.7, respectively. Figure 4 shows the convergence characteristic, while Fig.5 shows the absolute value of the transfer function after convergence of the algorithm. From this result, we observe that the proposed exact modeled line enhancer had a faster convergence speed, and better estimation accuracy in comparison to the cascaded notch filter.

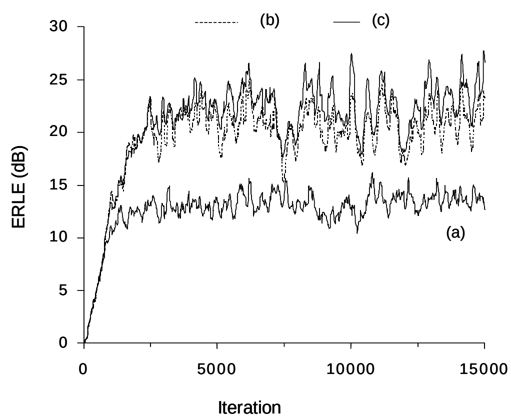

Fig.3. Convergence characteristic of the proposed equation error algorithm in comparison to b algorithm in [6] and (a) the algorithm in [1].

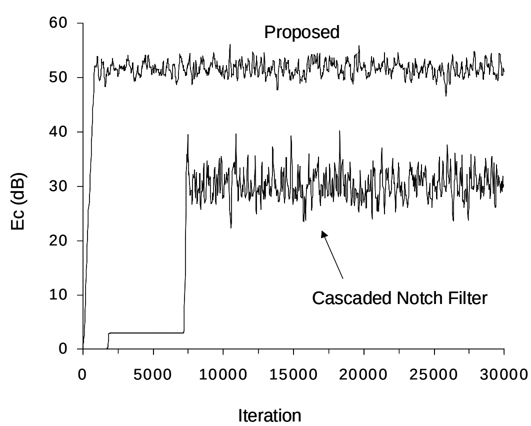

Fig.4. Convergence characteristic of the proposed ALE in comparison to the cascaded notch filter.

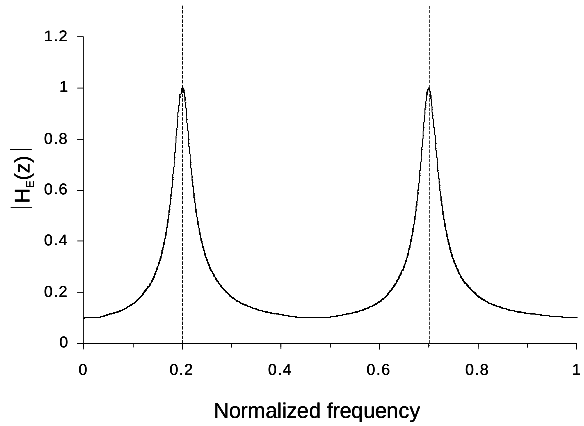

Fig.5. Band pass characteristic after convergence of the proposed ALE.

5. SUMMARY

We have proposed a new algorithm for an IIR ADF formulated using the equation error mode of estimation. The proposed algorithm is independent of the white disturbance signal and hence it eliminates the bias under such a condition. From the simulation result provided, we observe that there is comparable but very slight improvement in performance in comparison to similar algorithm with the same computational level. Further analytical comparison is however recommended.

We also proposed a structure for ALE, which employs the proposed algorithm. In comparison to the cascaded notch filter, the proposed algorithm had a faster convergence speed and a better estimation accuracy.

6. REFERENCES

[1] Shynk J.J. "Adaptive IIR Filtering". *IEEE Accoust., Speech and Signal Processing Magazine*. Vol.6, pp.4-21, 1989.

[2] Regalia P.R. "Adaptive IIR Filtering in Signal Processing and Control". Mercel Dekker, Inc., 1995.

[3] Ho. K.C. and Chan Y.T. "Bias Removal in Equation-Error IIR Filters". *IEEE Trans. on Signal Processing*, vol.43. No.1, pp.51-62, 1995.

[4] Regalia P.R. "An unbiased equation error identifier and reduced order approximations". *IEEE Trans. on Signal Processing*, vol.42. No.6, pp.1397-1412, 1994.

[5] Kim H. and Song W. "Unbiased Equation-Error Adaptive IIR filtering Based on Monic Normalization". *IEEE on Signal Processing letter*, vol.6, No.2, pp. 35-37, 1999.

[6] Song W. and Shin H. "Bias Free Adaptive IIR Filtering". *IEEE ICCASP Istanbul, Turkey*, vol.6, No.2, 2000.

[7] Widrow B. and Stearns S.D. "Adaptive Signal Processing". Prentice-Hall, Inc., 1985.

[8] Okello J., Arita S., Itoh Y., Fukui Y., and Kobayashi M. "An Adaptive Notch Filter for Eliminating Multiple Sinusoids with Reduced Bias". *Proceedings of IEEE ISCAS Geneva*, vol.3, pp. 551-554, May 2000.