An Efficient Adaptive Digital Filtering Algorithm for Identification of Second Order Volterra Systems

이차 볼테라 시스템 인식을 위한 효율적인 적응 디지탈 필터링 알고리즘

Y. S. Hwang*, V. J. Mathews**, I. W. Cha*, D. H. Youn*

황 영 수, V. J. Mathews, 차 일 환, 윤 대 희

ABSTRACT

This paper introduces an adaptive nonlinear digital filtering algorithm that uses the sequential regression (SER) method to update the second order Volterra filter coefficients in a recursive way. Conventionally, the SER method has been used to invert large matrices which result from direct application of Wiener filter theory to the Volterra filter. However, the algorithm proposed in this paper uses the SER approach to update the least squares solution which is derived for Gaussian input signals. In such an algortihm, the size of the matrix to be inverted is smaller than that of conventional approaches, and hence the proposed method is computationally simpler than conventional nonlinear system identification techniques. Simulation results are presented to demonstrate the performance of the proposed algorithm.

요 약

본 연구는 이차 볼테라 필터 계수를 연속적으로 변화시키기 위하여, sequential regression(SER) 방법을 이용한 적응 비선형 디지탈 필터링 알고리즘에 대하여 서술하였다. 일반적으로, SER 방법은 Wiener 필터 이론을 볼테라필터에 직접 적용시킬때 생기는 큰 행렬을 역변환시키기 위하여 사용되었다. 그러나, 본 연구에서는 입력신호가 가우시안일 경우, 최소 자승해를 구하기 위하여 SER 방법을 이용하였다. 이 알고리즘에서, 역변환시킬 행렬의 크기는 일반적 접근 방법보다 작게 되기때문에, 일반적 비선형 시스템 인식 기술보다 본 연구에서 제시한 방법의 계산량이 적다. 본 연구에서 제안한 알고리즘의 성능을 검토하기 위하여 시뮬레이션 결과를 구했다.

*ASSP Lab. Dept. of Electronic Engineering Yonsei University, SEOUL, KOREA.

**Dept of Electrical Engineering, The University of Utah Salt Lake City, Utah 84112 U.S.A

I. INTRODUCTION

The Volterra series representation of systems, which is an extension of linear systems theory, has attracted considerable interest lately [1-7]. The output y(k) of a (causal) discrete-time, nonlinear system can be represented as a function of the input sequence x(k) using the Volterra series expansion as

$$y(k)=h_0+\sum_{m_1=0}^{\infty} h_1(m_1)\,x(k-m_1)$$
$$+\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty} h_2(m_1,m_2)\,x(k-m_1)\,x(k-m_2)+\cdots$$
$$+\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}\cdots\sum_{m_p=0}^{\infty} h_p(m_1,m_2,\cdots,m_p)$$
$$x(k-m_1)\,x(k-m_2)\cdots x(k-m_p)+\cdots, \quad (1)$$

where $h_p(m_1,m_2,\cdots,m_p)$ is the p-th order Volterra kernal [1] of the system. In this paper, we consider the identification of weakly nonlinear systems [3] that can be represented using a second order Volterra series. System analysis using smaller order Volterra series has been applied to nonlinear transistor circuits [3,8], study of nonlinear drift oscillations of moored vessels subject to random sea waves [6,7] and a variety of applications involving weakly nonlinear systems.

Conventional approaches to nonlinear system identification involve identifying an equivalent linear system whose inputs have been augmented by second and higher order products of the input sequence to take care of the nonlinearities [3,9]. For example, when the relevant signals are stationary, applying the Wiener filter theory to the second order system described by the difference equation

$$y(k)=\sum_{i=0}^{N-1} a_i\,x(k-i)+\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} b_{i,j}\,x(k-i)\,x(k-j) \quad (2)$$

yields the following optimal minimum mean squared solution [9]:

$$\tilde{H}=\tilde{R}_{xx}^{-1}\,\tilde{R}_{yx} \quad (3)$$

where

$$\tilde{H}=[a_0,\,a_1,\cdots,\,a_{N-1},\,b_{0,0},\,b_{0,1},\cdots b_{N-1,N-1}]^T, \quad (4)$$

$$\tilde{R}_{xx}=E\{\tilde{X}(k)\,\tilde{X}^T(k)\}, \quad (5)$$

$$\tilde{R}_{yx}=E\{y(k)\,\tilde{X}(k)\}, \quad (6)$$

and

$$\tilde{X}(k)=[x(k),\,x(k-1),\cdots,\,x(k-N+1),\,x^2(k)$$
$$-r_{xx}(0),\,x(k)x(k-1)-r_{xx}(1),$$
$$\cdots,\,x^2(k-N+1)-r_{xx}(0)]^T. \quad (7)$$

In the above equations and the rest of the paper $E\{\cdot\}$ denotes the statistical expectation of $\{\cdot\}$, superscript T denotes matrix (vector) transpose and

$$r_{xy}(m)=E\{x(k)y(k-m)\} \quad (8)$$

From (3) it can be seen that computing the optimal solution requires inverting an $(N^2+N)\times(N^2+N)$ matrix, which may be computationally very difficult for even moderatly large values of N. Recently, Koh and Powers have shown that when the relevant signals are Gaussian and the optimal minimum mean squared solution is sought, the order of the matrix to be inverted is the same as that of the linear system (i.e., $(N\times N)$) [6,7,11]. In this paper, we present a least squares (LS) solution to the nonlinear system identification problem, when the input signals are Gaussian. The size of the matrix to be inverted is the same as that of Koh and Power's approach. We also present an adaptive algorithm that uses the sequential regression (SER) method [9,10] to compute the LS solution in a recursive way. Thus the algorithm can be applied to the problems with nonstationary signals or time-varying nonlinear systems by defining an appropriate window function.

The rest of the paper is organized as follows: A formal statement of the problem is made in

Section II. The LS solution is also derived in this section. The adaptive Volterra filter with SER method is introduced in Section III. The effectiveness of the proposed algorithm is demonstrated in Section IV using a simulation example. Finally, we make the concluding remarks in Section V.

II. PROBLEM STATEMENT AND OPTIMUM SOLUTION

Let H in Fig. 1 represent an unknown nonlinear system that can be represented as a causal second order Volterra filter whose output y(k) of the system can be expressed in terms of the input sequence x(k) as

$$y(k)=\sum_{i=0}^{N-1} a_i x(k-i)+\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} b_{i,j}\{x(k-i)x(k-j)-E\{x(k-i)x(k-j)\}\} \quad (9)$$

where a_i's and $b_{i,j}$'s are the first and second order Volterra kernals of the system. We will assume that the input sequence x(k) is zero mean and Gaussian. The last term $E\{x(k-i)x(k-j)\}$ in (9) is included without loss of generality, so that the output sequence y(k) is also zero mean. The output y(k) in (9) can be equivalently expressed in matrix form as

$$y(k)=A^T X(k)+tr\{B[X(k)X^T(k)-R_{xx}]\} \quad (10)$$

where

$$X(k)=[x(k),\ x(k-1),\ \cdots,\ x(k-N+1)]^T \quad (11)$$

$$A=[a_0,\ a_1,\ \cdots,\ a_{N-1}]^T \quad (12a)$$

$$B=\begin{bmatrix} b_{0,0} & b_{0,1} & \cdots & b_{0,N-1} \\ b_{1,0} & b_{1,1} & \cdots & b_{1,N-1} \\ \vdots & & & \vdots \\ b_{N-1,0} & b_{N-1,1} & \cdots & b_{N-1,N-1} \end{bmatrix} \quad (12b)$$

is a symmetric matrix so that $b_{i,j}=b_{j,i}$, $R_{XX}=E\{X(k)X^T(k)\}$ is the autocorrelation matrix of X(k), and $tr\{\cdot\}$ denotes the trace of the matrix $\{\cdot\}$.

The problem here is to derive an adaptive filtering algorithm that uses the SER method to track the (possibly time varying) parameters A and B of the Volterra filter in (10) so as to minimize the cost functional

$$C(\hat{A}(k),\hat{B}(k))=\sum_{i=0}^{k} q(i,k)\{y(i)-[\hat{A}^T(k)X(i)+tr\{\hat{B}(k)(X(i)X^T(i)-\hat{R}_{xx}(k))\}]\}^2 \quad (13)$$

where carets ( ) denote estimated quantities. Also, q(i,k) is a weighting function for the squared estimation error $(y(i)-\hat{y}_k(i))^2$, where

$$\hat{y}_k(i)=\hat{A}^T(k)X(i)+tr\{\hat{B}(k)[X(i)X^T(i)-\hat{R}_{xx}(k)]\} \quad (14)$$

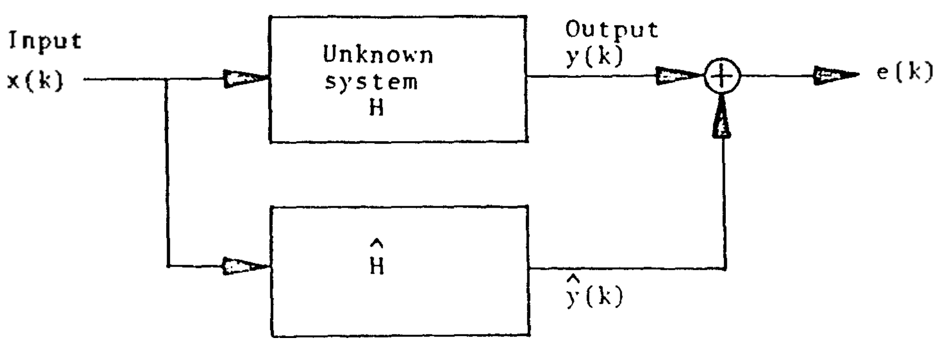


Fig. 1. Block diagram for the system identification problem

is an estimate of the output at time i based on the parameter estimates $\hat{A}(k)$, $\hat{B}(k)$ and $\hat{R}_{XX}(k)$ at time k. (Note that the time index k has been used for the parameter estimates). For analytical tractability, we will assume that the autocorrelation matrix R_{xx} is estimated at time k as

$$\hat{R}_{xx}(k) = \sum_{i=0}^{k} q(i,k)\, X(i)\, X^T(i) \tag{15}$$

Substituting (15) in (13), the cost functional to be minimized becomes

$$C(\hat{A}(k), \hat{B}(k)) = \sum_{i=0}^{k} q(i,k)\, [\, y(i) - [\hat{A}^T(k) X(i) + tr\{\hat{B}(k)\,(X(i)X^T(i) - \sum_{j=0}^{k} q(j,k)\, X(j) X^T(j))\}]\,]^2 \tag{16}$$

We can minimize the above cost functional with respect to $\hat{A}(k)$ by setting the gradient of $C(\hat{A}(k), \hat{B}(k))$ with respect to $\hat{A}(k)$ to zero (i.e., $\nabla_{\hat{A}(k)} C(\hat{A}(k),\hat{B}(k)) = 0$). After some straightforward computations, we obtain

$$\hat{R}_{yx}(k) - \hat{R}_{xx}(k)\, \mathring{A}(k) + \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \hat{b}_{m,n}(k) \{ \sum_{i=0}^{k} q(i,k)\, x(i-m)\, x(i-n)\, X(i) \}$$
$$- [\sum_{j=0}^{k} q(j,k)\, X^T(j)\, \hat{B}(k)\, X(j)\,]\, [\sum_{i=0}^{k} q(i,k)\, X(i)] \tag{17}$$

where

$$\hat{R}_{yx}(k) = \sum_{i=0}^{k} q(i,k)\, y(i)\, X(i) \tag{18}$$

and $\mathring{A}(k)$ yields the minimum $C(\hat{A}(k),\hat{B}(k))$ for a given $\hat{B}(k)$.

If the weighting function q(i,k) for $0 \le i \le k$ represents the impulse response function of a lowpass filter with unit gain at zero frequency (i.e., $\sum_{i=0}^{k} q(i,k) = 1$), $\sum_{i=0}^{k} q(i,k)x(i-j)$ and $\sum_{i=0}^{k} q(i,k)x(i-m)x(i-n)x(i-j)$ approximate the mean and third order moment of x(k), respectively [12,13]. Now, since the input signal was assumed to be zero mean and Gaussian, the odd order moments of the signal are zero and therefore, the second and third terms of (17) are approximately zero. Substitution of this in (17) results in the simplified expression for $\mathring{A}(k)$

$$\mathring{A}(k) = \hat{R}_{xx}^{-1}(k)\, \hat{R}_{yx}(k) \tag{19}$$

To compute $\mathring{B}(k)$, we once again set the gradient of $C(\hat{A}(k),\hat{B}(k))$ with respect to $\hat{B}(k)$ to zero and obtain the LS solution

$$\mathring{B}(k) = \frac{1}{2}\, \hat{R}_{xx}^{-1}(k)\, \hat{T}_{yx}(k)\, \hat{R}_{xx}^{-1}(k) \tag{20}$$

where

$$\hat{T}_{yx}(k) = \sum_{i=0}^{k} q(i,k)\, y(i)\, X(i)\, X^T(i). \tag{21}$$

The derivation of (20) is given in Appendix A.

It may be pointed out here that the LS solution in (19) and (20) have the same functional form as the optimum minimum mean squared estimates for A and B derived in [6,7,11]. In the next section, we will develop a recursive algorithm for updating the least squares solution at each time instant, when the weighting function q(i,k) is exponential.

III. THE EXPONENTIALLY WEIGHTED SER ALGORITHM

The SER algorithm has been used to update the optimum linear [10] and nonlinear [9] filter coefficients in a recursive manner. In this section, we adopt the LS solution in (19) and (22) and apply the SER algorithm to update the Volterra filter weights that minimize the cost functional given by (16) with the weighting function q(i,k) selected as

$$q(i,k) = (1-\beta)\,\beta^{k-i} \tag{22}$$

where $0 < \beta < 1$. This weighting function penalizes the current estimation errors more than the past ones. Smaller value of β makes the weighting function sharper implying that only a shorter history of the squared errors are effectively used. The choice of β depends on prior knowledge of the stationarity of the signals and systems involved. If system parameters change very slowly, we will use β close to 1. On the other hand, we will choose β to be smaller if the parameters are known to change relatively fast. It may be noted here that for large values of k

$$\sum_{i=0}^{k} q(i,k) = 1-\beta^{k+1} \approx 1 \tag{23}$$

and therefore the results in Section II are applicable here. Thus rewriting the LS solutions in (19) and (20), we have

$$\mathring{A}(k) = \hat{R}_{xx}^{-1}(k)\,\hat{R}_{yx}(k) \tag{24}$$

and

$$\mathring{B}(k) = \frac{1}{2}\hat{R}_{xx}^{-1}(k)\,\hat{T}_{yx}(k)\,\hat{R}_{xx}^{-1}(k) \tag{25}$$

where

$$\hat{R}_{y}(k) = (1-\beta)\sum_{i=0}^{k}\beta^{k-i}\,y(i)\,X(i) \tag{26}$$

$$\hat{R}_{xx}(k) = (1-\beta)\sum_{i=0}^{k}\beta^{k-i}\,X(i)\,X^T(i) \tag{27}$$

and

$$\hat{T}_{yx}(k) = (1-\beta)\sum_{i=0}^{k}\beta^{k-i}\,y(i)\,X(i)\,X^T(i) \tag{28}$$

To obtain the recursive relationships, we proceed as follows: Substituting (26) and (27) in (24) and replacing k by k-1 and multiplying both sides by β, we get

$$\left[(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}\,X(i)\,X^T(i)\right]\hat{A}(k-1)$$
$$= (1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}\,y(i)\,X(i) \tag{29}$$

Note that in (29) we have used $\hat{A}(k)$ instead of $\mathring{A}(k)$, just to emphasize the fact that this is an estimated quantity, even though it does represent the LS solution for A given by (24). Similarly, we will use $\hat{B}(k)$ instead of $\mathring{B}(k)$ from now on.

Substitution of (26) in (24) yields

$$\hat{R}_{xx}(k)\,\hat{A}(k) = (1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}\,y(i)\,X(i) + (1-\beta)$$
$$y(k)\,X(k) \tag{30}$$

Substituting (30) in (29) and manipulating the resulting equation, we obtain the recursive relationship for $\hat{A}(k)$ as

$$\hat{A}(k) = \hat{A}(k-1) + (1-\beta)\,\hat{R}_{xx}^{-1}(k)\,X(k)$$
$$[y(k) - X^T(k)\,\hat{A}(k-1)]. \tag{31}$$

The derivation of the recursive relationship for the quadratic weights is more involved, even though straightforward. $\hat{B}(k)$ is related to $\hat{B}(k-1)$ as

$$\hat{B}(k) = \frac{1}{\beta}[I-(1-\beta)\,\hat{R}_{xx}^{-1}(k)\,X(k)\,X^T(k)]\,\hat{B}(k-1)$$
$$[I-(1-\beta)\,X(k)\,X^T(k)\,\hat{R}_{xx}^{-1}(k)]$$
$$+\frac{1}{2}(1-\beta)\,y(k)\,\hat{R}_{xx}^{-1}(k)\,X(k)\,X^T(k)\,\hat{R}_{xx}^{-1}(k) \tag{32}$$

where I denotes the $N \times N$ identity matrix. Now, from (31) and (32), we can see that the Volterra filter coefficients can be updated in a recursive way if $\hat{R}_{xx}^{-1}(k)$ can be obtained in a recurstive manner. $\hat{R}_{xx}^{-1}(k)$ is computed using the relationship

$$\hat{R}_{xx}^{-1}(k) = \frac{1}{\beta}\hat{R}_{xx}^{-1}(k-1) - \frac{1}{d(k)}$$
$$\left[\frac{1-\beta}{\beta^2}\hat{R}_{xx}^{-1}(k-1)\,X(k)\,X^T(k)\,\hat{R}_{xx}^{-1}(k-1)\right] \tag{33}$$

where

$$d(k) = 1 + \frac{1-\beta}{\beta} X^T(k)\,\hat{R}_{xx}^{-1}(k-1)\,X(k) \qquad (34)$$

and d(k) is assumed to be nonzero. Derivation of (32) and (33) are given in Appendices B and C, respectively. The recursive algorithm is summarized in Fig. 2.

Koh and Powers have developed an adaptive implementation of the optimal minimum mean squared solution [6,7,11] using the LMS (Least Mean Square) algorithm [14]. However, this adaptive implementation results in convergence properties that depend on the eigenvalue spread of the input correlation matrix. Since the method proposed here is a recursive least squares solution, the convergence speed is independent of the signal statistics and depends only on the weighting function employed.

It may be pointed out that convergence of the LS solution to the optimal minimum mean squared estimate can be proved under assumptions of uncorrelated input vectors. Interested

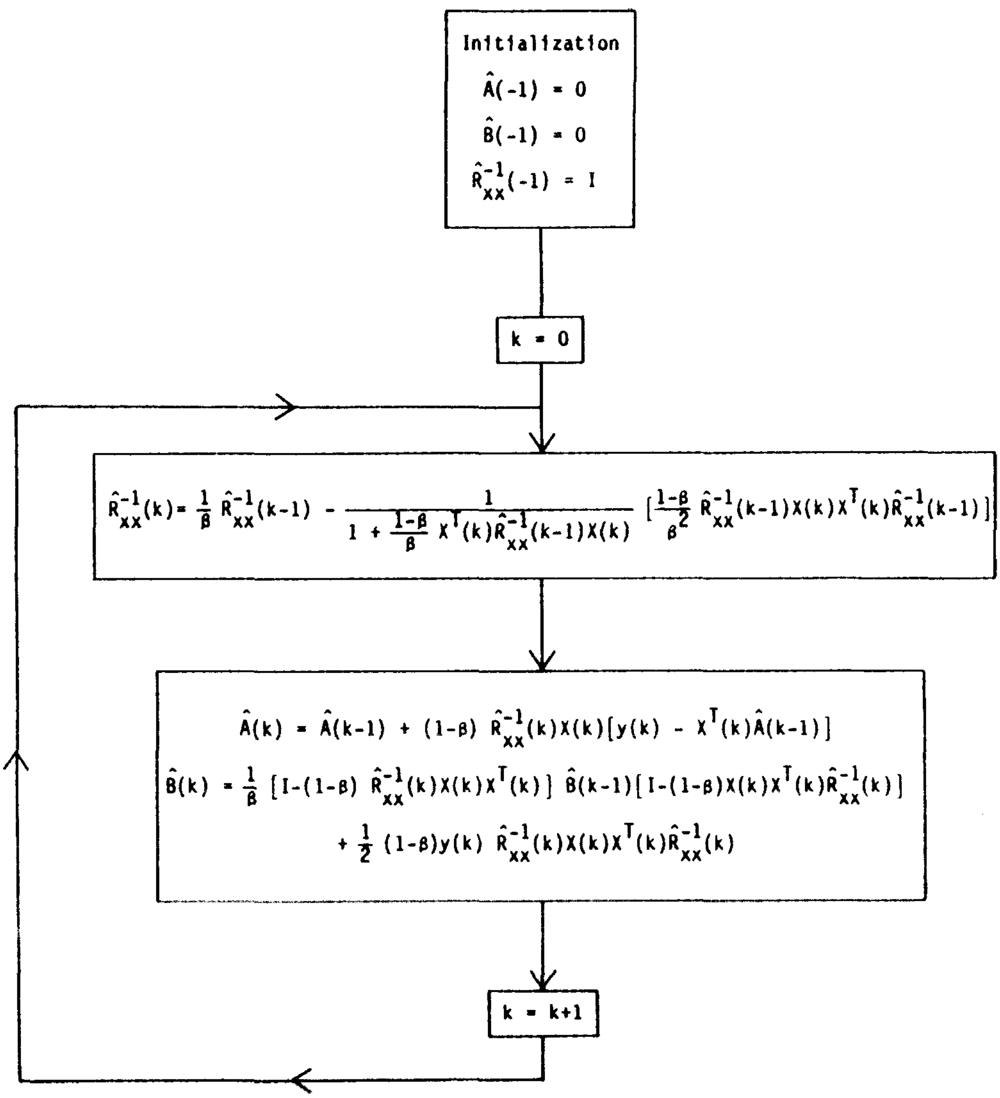


Fig. 2. Second Order Adaptive Volterra Filter With SER Algorithm

readers are referred to [17].

Due to improper model order selection or due to zero input to the system for arbitrary lengths of time, it is possible that the auto-correlation matrix $\hat{R}_{xx}(k)$ is singular or ill-conditioned. To avoid problems that arise from inverting such matrices, we may modify $\hat{R}_{xx}(k)$ as

$$\hat{\hat{R}}_{xx}(k) = \hat{R}_{xx}(k) + uI \tag{26}$$

where u is a small positive number. The effect of this modification is that of adding a small amount of white noise to the input sequence which will ensure that $\hat{\hat{R}}_{xx}(k)$ is nonsingular [10].

In the next section, we will demonstrate the usefulness of the proposed algorithm with a simulation example.

IV. A SIMULATION EXAMPLE

To study the performance of the proposed algorithm, we consider the system identification problem for a second order Volterra filter whose coefficients A and B are given by

$$A = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0.6 \\ -0.2 \end{bmatrix} \tag{35}$$

and

$$B = \begin{bmatrix} b_{0,0} & b_{0,1} \\ b_{1,0} & b_{1,1} \end{bmatrix} = \begin{bmatrix} 0.3 & 0.1 \\ -0.1 & 0.15 \end{bmatrix} \tag{36}$$

Note that B is symmetric. The system identification problem is schematically dipicted in Fig. 1, where the reference input x(k) is white, Gaussian with zero mean and unit variance and y(k) is obtained from x(k) using (10) where A and B are given by (35) and (36), respectively.

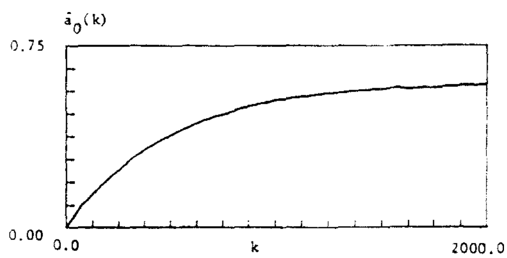
Thirty independent simulations were run using 2000 data samples each with $\beta = 0.998$, and the results presented are averaged over the thirty runs. After 2000 iterations, the ensemble averages of the parameter estimates were

$$\hat{A}(2000) = \begin{bmatrix} 0.5891 \\ -0.1977 \end{bmatrix} \tag{37}$$

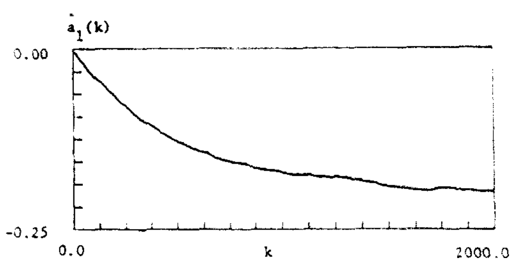
and

$$\hat{B}(2000) = \begin{bmatrix} 0.2799 & -0.0988 \\ -0.0988 & -0.1352 \end{bmatrix} \tag{38}$$

Figures 3a-e display plots of ensemble averages of $\hat{a}_0(k)$, $\hat{a}_1(k)$, $\hat{b}_{0,0}(k)$, $\hat{b}_{0,1}(k)$, $\hat{b}_{1,0}(k)$ and $\hat{b}_{1,1}(k)$, respectively. From (37) and (38) and also these figures, we can see that the estimates converge to the true values. Note that in (37) and (38), the differences among the true and estimated parameter values at the 2000-th iteration are within ten percent of the correct values and are well within the ranges expected from statistical variability. Also, the ensemble average of the squared estimation errors $((\hat{y}(k)-y(k))^2)$ are displayed in Fig. 3f. This result demonstrates that the estimation error decreases exponentially with time.



(a) $\hat{a}_0(k)$



(b) $\hat{a}_1(k)$

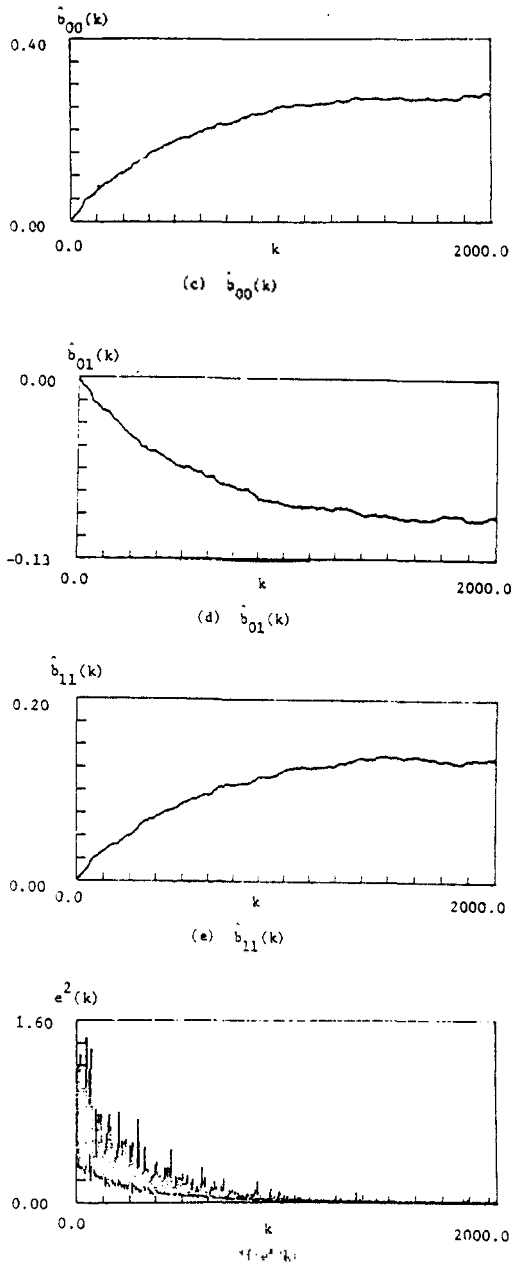


Fig. 3. Ensemble averages of the system identification problem in the simulation example

V. CONCLUSION

The adaptive second order Volterra filter employing the sequential regression method was developed in this paper. The recursive algorithm updates the least squares solution that minimizes the cost functional in (16), when the weighting function q(i,k) is exponential. The method developed in this paper differred from conventional nonlinear system identification techniques in that the order of the matrix to be inverted was the same as that for linear system identification problems. The Gaussian assumption for the input signal sequence was crucial in this large reduction in computational complexity.

The method developed in this paper differred from Koh and Power's least mean square approach in that the convergence of our algorithm was independent of the eigenvalue spread of the input autocorrelation matrix. Instead, the convergence speed depends only on the weighting function used.

The effectiveness of the adaptive system identification algorithm was demonstrated using a simulation example. The computational simplicity of the proposed method over conventional nonlinear system identification techniques along with its convergence properties should make it a prime candidate for system identification of weakly nonlinear systems that can be represented/approximated as second order Volterra series.

APPENDIX A
LEAST SQUARES SOLUTION FOR $\hat{B}(k)$

We minimize the cost functional in (16) with respect to $\hat{B}(k)$ by setting the gradient of $C(\hat{A}(k),\hat{B}(k))$ with respect to $\hat{B}(k)$ to zero (i.e., $\nabla_{\hat{B}(k)} C[\hat{A}(k), \hat{B}(k)] = 0$). After some computations and using the following relationship

$$\nabla_{\hat{B}(k)} \{ tr[\hat{B}(k) X(i) X^T(i)] \} = X(i) X^T(i), \quad \text{(A1)}$$

We obtain

$$\sum_{i=0}^{k} q(i,k) y(i) X(i) X^T(i) = [\sum_{i=0}^{k} q(i,k) y(i)]$$

$$\sum_{j=0}^{k} q(j,k) X(j) X^T(j) + \sum_{i=0}^{k} q(i,k) \hat{A}^T(k) X(i) X(i) X^T(i) -$$

$$[\sum_{i=0}^{k} q(i,k) \hat{A}^T(k) X(i)] [\sum_{j=0}^{k} q(j,k) X(j) X^T(j)]$$

$$[\sum_{i=0}^{k} (q(i,k) X^T(i) \hat{B}(k) X(i)] [\sum_{j=0}^{k} (q(j,k) X(j) X^T(j)]$$

$$[1-\sum_{i=0}^{k} q(i,k)] + \sum_{i=0}^{k} \{ q(i,k) [\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \hat{b}_{m,n}(k)$$

$$(x(i-m)x(i-n) - \sum_{j=0}^{k} q(j,k)x(j-m)x(j-n))]$$

$$X(i)X^T(i)\} \quad (A2)$$

As discussed in Section II, $\sum_{i=0}^{k} q(i,k)x(i-j)$ and $\sum_{i=0}^{k} q(i,k)x(i-m)x(i-n)x(i-j)$ approximate the mean and third order moment of x(k), respectively, and the following approximations also hold:

$$\sum_{i=0}^{k} q(i,k)y(i) \approx E\{y(i)\} = 0 \quad (A3)$$

$$\sum_{i=0}^{k} q(i,k) A^T(k) X(i) X(i) X^T(i)$$

$$= \sum_{i=0}^{k} q(i,k) [a_0(k)x(i) + a_1(k)x(i-1) + \cdots$$

$$+ a_{N-1}(k)x(i-N+1)] X(i)X^T(i) \approx 0 \quad (A4)$$

and

$$\sum_{i=0}^{k} q(i,k) A^T(k) X(i) = a_0(k) [\sum_{i=0}^{k} q(i,k)x(i)]$$

$$+ a_1(k) [\sum_{i=0}^{k} q(i,k)x(i-1)] + \cdots$$

$$+ a_{N-1} [\sum_{i=0}^{k} q(i,k)x(i-N+1)] \approx 0 \quad (A5)$$

Using (A3) - (A4), (A2) can be simplified as

$$\sum_{i=0}^{k} q(i,k)y(i)X(i)X^T(i)$$

$$= \sum_{i=0}^{k} \{ q(i,k) [\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \hat{b}_{m,n}(k)(x(i-m)x(i-n)$$

$$- \sum_{j=0}^{k} q(j,k)x(j-m)x(j-n))] X(i)X^T(i)\}. \quad (A6)$$

The right hand term of (A6) is an $N \times N$ matrix whose $(g+1, l+1)$th element $t_{g+1,l+1}(k)$ is

$$t_{g+1,l+1}(k) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \hat{b}_{m,n}(k) \{ \sum_{i=0}^{k} q(i,k)x(i-m)$$

$$x(i-n)x(i-g)x(i-l)$$

$$-[\sum_{j=0}^{k} q(j,k)x(j-m)x(j-n)] [\sum_{j=0}^{k} q(j,k)x(j-g)$$

$$x(j-l)]\} \quad (A7)$$

As discussed before for sums of smaller order products of the input sequence x(k), the weighted sum of the fourth order products of x(k) in (A7) approximates the fourth order moment of x(k). That is

$$\sum_{i=0}^{k} q(i,k)x(i-m)x(i-n)x(i-g)x(i-l)$$

$$\approx E\{x(i-m)x(i-n)x(i-g)x(i-l)\} \quad (A8)$$

Since the fourth order expection of Gaussian signals can be written as sum of product of second order expectations as [15]

$$E\{x(i-m)x(i-n)x(i-g)x(i-l)\}$$

$$= E\{x(i-m)x(i-n)\} E\{x(i-g)x(i-l)\}$$

$$+ E\{x(i-m)x(i-g)\} E\{x(i-n)x(i-l)\}$$

$$+ E\{x(i-m)x(i-l)\} E\{x(i-n)x(i-g)\} \quad (A9)$$

we can approximate the left hand side of (A8) as

$$[\sum_{i=0}^{k} q(i,k)x(i-m)x(i-n)] [\sum_{i=0}^{k} q(i,k)x(i-g)x(i-l)]$$

$$+ [\sum_{i=0}^{k} q(i,k)x(i-m)x(i-g)] [\sum_{i=0}^{k} q(i,k)x(i-n)x(i-l)]$$

$$+ [\sum_{i=0}^{k} q(i,k)x(i-m)x(i-l)]$$

$$[\sum_{i=0}^{k} q(i,k)x(i-n)x(i-g)] \quad (A10)$$

Substituting the above in (A7), we obtain

$$t_{g+1,l+1}(k) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \hat{b}_{m,n}(k) \{ [\sum_{i=0}^{k} q(i,k)x(i-m)$$

$$x(i-g)] [\sum_{i=0}^{k} q(i,k)x(i-n)x(i-l)$$

$$+[\sum_{i=0}^{k} q(i,k)x(i-m)x(i-l)]$$

$$[\sum_{i=0}^{k} q(i,k)x(i-n)x(i-g)]\,|. \qquad (A11)$$

Noting that $\hat{b}_{m,n}(k) = \hat{b}_{n,m}(k)$, it can be easily seen that (A11) is exactly the same as the (g+1, l+1)th element of the $N \times N$ matrix $2\hat{R}_{xx}(k)\hat{B}(k)\hat{R}_{xx}(k)$. Then (A6) can be simplified to the LS solution for $\hat{B}(k)$ as

$$\hat{B}(k) = \frac{1}{2}\hat{R}_{xx}^{-1}(k)\,\hat{T}_{yx}(k)\,\hat{R}_{xx}^{-1}(k) \qquad (A12)$$

where

$$\hat{T}_{yx}(k) = \sum_{i=0}^{k} q(i,k)\,y(i)\,X(i)X^T(i) \qquad (A13)$$

APPENDIX B
RECURSIVE COMPUTATION OF QUADRATIC FILTER COEFFICIENTS

To obtain the recursive relationship for the quadratic filter coefficients, we proceed as follows: Substituting (27) and (28) in (25) and replacing k by k-1, we get

$$(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}\,y(i)\,X(i)X^T(i)$$

$$=\frac{2}{\beta}[(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}X(i)X^T(i)]\,\hat{B}(k-1)$$

$$[(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}X(i)X^T(i)] \qquad (B1)$$

where $\hat{B}(k)$ is used to instead of $\mathring{B}(k)$. Also, substitution of (28) into (25) and a simplification yield

$$2\hat{R}_{xx}(k)\hat{B}(k)\hat{R}_{xx}(k) = (1-\beta)\sum_{i=0}^{k}\beta^{k-i}\,y(i)X(i)X^T(i)$$

$$+(1-\beta)\,y(k)\,X(k)X^T(k) \qquad (B2)$$

Substituting (B1) in (B2), we obtain

$$2\hat{R}_{xx}(k)\hat{B}(k)\hat{R}_{xx}(k) = \frac{2}{\beta}[(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}X(i)$$

$$X^T(i)]\hat{B}(k-1)[(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}X(i)X^T(i)]$$

$$+(1-\beta)\,y(k)\,X(k)X^T(k). \qquad (B3)$$

We will use the following simple matrix equality on (B3):

$$ABA = (A+C)B(A+C) - DBA - ABC - CBC. \qquad (B4)$$

Letting $(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}X(i)X^T(i)$, B(k-1) and $(1-\beta)X(k)X^T(k)$ to be matrices A, B and C in (B4) and substituting (B4) in (B3), we get

$$\hat{R}_{xx}(k)\hat{B}(k)\hat{R}_{xx}(k) = \frac{1}{\beta}[(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}X(i)X^T(i)$$

$$+(1-\beta)X(k)X^T(k)]\,\hat{B}(k-1)$$

$$[(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}X(i)X^T(i)+(1-\beta)X(k)X^T(k)]$$

$$-\frac{1}{\beta}[(1-\beta)X(k)X^T(k)]\,\hat{B}(k-1)[(1-\beta)$$

$$\sum_{i=0}^{k-1}\beta^{k-i}X(i)X^T(i)]$$

$$-\frac{1}{\beta}[(1-\beta)\sum_{i=0}^{k-1}\beta^{k-i}X(i)X^T(i)]\,\hat{B}(k-1)$$

$$[(1-\beta)X(k)X^T(k)]$$

$$-\frac{1}{\beta}[(1-\beta)X(k)X^T(k)]B(k-1)[(1-\beta)X(k)X^T(k)]$$

$$+\frac{1}{2}(1-\beta)\,y(k)X(k)X^T(k) \qquad (B5)$$

Substitution (27) in (B5) and a rearrangement results in the recursive relationship for $\hat{B}(k)$ as

$$\hat{B}(k) = \frac{1}{\beta}[I - (1-\beta)\hat{R}_{xx}(k)X(k)X^T(k)]$$

$$\hat{B}(k-1)[I-(1-\beta)X(k)X^T(k)R_{xx}^{-1}(k)]$$

$$+\frac{1}{2}(1-\beta)\,y(k)\,\hat{R}_{xx}^{-1}(k)X(k)X^T(k)\,\hat{R}_{xx}^{-1}(k) \quad (B6)$$

where I denotes the $N \times N$ identity matrix.

APPENDIX C
RECURSIVE COMPUTATION OF INVERSE AUTOCORRELATION MATRIX

From (18), it follows that

$$\hat{R}_{xx}(k) = (1-\beta)\sum_{i=0}^{k-1} \beta^{k-i} X(i) X^T(i)$$

$$+ (1-\beta) X(k) X^T(k) \qquad (C1)$$

$$= \beta \hat{R}_{xx}(k-1) + (1-\beta) X(k) X^T(k) \qquad (C2)$$

To compute $\hat{R}_{xx}^{-1}(k)$ we can use the matrix inverse lemma [16] which says that if A and B are nonsingular square matrices such that

$$A = B + CD \text{ and } 1 + DB^{-1}C \neq 0,$$

$$A^{-1} = B^{-1} - \frac{1}{1+DB^{-1}C} B^{-1}CDB^{-1} \qquad (C3)$$

Applying the matrix inverse lemma to (C2) with $\hat{R}_{xx}(k)$, $\beta\hat{R}_{xx}(k-1)$, $(1-\beta)X(k)$ and $X^T(k)$ being matrices A, B, C and D, respectively, results in the recursive relationship

$$\hat{R}_{xx}^{-1}(k) = \frac{1}{\beta}\hat{R}_{xx}^{-1}(k-1) - \frac{1}{d(k)}\left[\frac{1-\beta}{\beta^2}\hat{R}_{xx}^{-1}\right.$$

$$(k-1) X(k) X^T(k) \hat{R}_{xx}^{-1}(k-1) \qquad (C4)$$

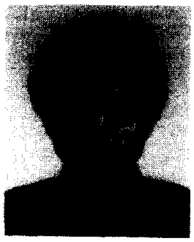
where

$$d(k) = 1 + \frac{1-\beta}{\beta} X^T(k) \hat{R}_{xx}^{-1}(k-1) X(k)$$

REFERENCES

1. M. Schetzen, *The Volterra and Wiener Theory of the Nonlinear Systems*, Wiley, New York, 1980.
2. M. Schetzen, "Nonlinear System Modeling Based on the Wiener Theory", *Proc. IEEE,* Vol. 69, No. 12, pp.1557-1573, December 1981.
3. E.J. Ewen and D.D. Weiner, "Identification of Weakly Nonlinear Systems Using Input and Output Measurements", *IEEE Trans. Circuits and Systems*, Vol. CAS-27, No. 12, pp.1255-1261, December 1980.
4. S. Boyd, Y.S. Tang and L.O. Chua, "Measuring Volterra Kernels", *IEEE Trans. Circuits and Systems*, Vol. CAS-30, No. 8, pp.571-577, August 1983.
5. H.K. Thapar and B.J. Leon, "Transform Domain and Time Domain Characterization of Nonlinear Systems with Volterra Series", *IEEE Trans. Circits and Systems*, Vol. CAS-31, No. 10, pp.906-912, October 1984.
6. T. Koh and E.J. Powers, "Second Order Volterra Filtering and its Application", to be published.
7. T. Koh, E.J., Powers and R.W. Miksad, "An Approach to Time Domain Modelling of Nonlinear Drift Oscillations in Random Seas", *Proc. Int. Symp. Offshore Engineering*, Pis de Janeiro, Brazil, 1983.
8. S. Narayanan, "Application of Volterra Series to Intermodulation Distortion of Transister Feedback Amplifiers", *IEEE Trans. Circuit Theory*, Vol. CT-17, pp.518-527, November 1970.
9. D. Graupe, *Identification of Systems*, Van Nostrand, 1972.
10. N. Ahmed, et. al., "A Short-Term Sequential Regression Algorithm", *IEEE Trans. Acoust. Speech and Signal Proc.*, Vol. ASSP-27, No. 5, pp.453-457, October 1979.
11. T. Koh and E.J. Powers, "An Adaptive Nonlinear Digital Filter With Lattice Orthogonalization", *Proc. ICASSP 83*, pp.37-40, Boston, 1983.
12. N. Ahmed and S. Vijayendra, "An Algorithm for Line Enhancement", *Proc. IEEE*, Vol. 70, No. 12, pp.1459-1460, December 1982.
13. D. Johnstone and O. Frost, "High Resolution Differential Time-of-Arrival and Differential Doppler Estimation", ARGO Systems, Inc., 1969 East Meadors Circle, Palo Alto, CA 94303.
14. B. Widrow, et. al., "Adaptive Noise Cancelling: Principles and Applications", *Proc. IEEE*, Vol. 63, No. 12, pp.1692-1716, December 1975.
15. D. Middleton, *An Introduction to Statistical Communication Theory*, McGraw-Hill, New York, pp.343, 1960.
16. D.K. Faddeev and V.N. Faddeeva, *Computational Methods of Linear Algebra*, W.H. Freeman and Company, San Francisco, pp.174, 1963.
27. K.K. Yu, "Nonlinear Adaptive Digital Filter with Sequential Regression (SER) Algorithm", *M.S. Thesis*, Department of Eectrical and Computer Engineering, University of Iowa, Iowa City, Iowa 52242, December 1984.

▲ **Y. S. Hwang**

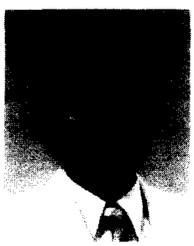


1982. 2 : Department of Electronic Engineering, Yonsei University (B.S)
1984. 2 : Department of Electronic Engineering Yonsei University (M.S)
1984～ . Department of Electronic Engineering, Yonsei University Ph.D course

▲ **V. John Mathews**

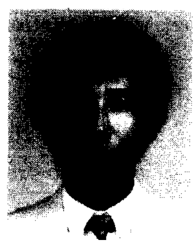
1980 : Electronics & Communication Engineering from University of Uradras, India (B.E.)
1981 : Electrical & computer Engineering from University of Iowa, Iowa City (M.S.)
1984 : Electrical & computer Engineering from University of Iowa, Iowa City (Ph.D)
1980～1984 : Teaching-Research Fellowship at the University of Iowa
1984～1985 : Visiting Assistant Professor with the Department of Electrical and Computer Engineering
1985～ : Assistance Professor with the Department of Electrical Engineering, University of Utah, Salt Lake City

▲ **I. W. Cha**



1959. 2 : Department of Electrical Engineering, Yonsei University (B.S)
1983 Department of Electronic Engineering, Yonsei University (Ph.D)
1969～1970 : Institute of Sound and Vibration Research, University of Southampton, England
1959～1972 : Department of Electrical Engineering, Yonsei University
1981～1987 : The prêsident of the Acoustical Society of Korea
1986. 9～1987. 7 : Department of Electrical and Computer Engineering, University of Iowa, Visiting Professor
1973～ : Department of Electronic Engineering, Yonsei University

▲ **D.H. Youn**



1977. 2 : Department of Electronic Engineering, Yonsei University (B.S)
1979. 5 : Department of Electrical Engineering, Kansas State University (M.S)
1982. 5 : Department of Electrical Engineering, Kansas State University (Ph.D)
1982. 8～1985. 6 : Department of Electrical and Computer Engineering, University of Iowa
1985. 9～ : Department of Electronic Engineering, Yonsei University