Performance Monitoring In Nonlinear Adaptive Noise Cancellation

S. A. BILLINGS

F. A. ALTURKI

Department Of Control Engineering

University Of Sheffield

Research Report No.(409)

November, 1990

Abstract

Correlation based model validity tests are introduced to monitor the operation of nonlinear adaptive noise cancellation filters and to detect if the filters are operating correctly or incorrectly. The tests are derived for a NARMAX (Nonlinear Auto Regressive Moving Average model with eXogenous inputs) filter design based on a SubOptimal Least Squares (SOLS) estimation algorithm. Simulation studies are included to illustrate the performance of the new tests.

1 Introduction

The recent developments in microchip technology have increased the interest in signal processing methods. One of the most important fields of signal processing is adaptive noise cancellation in which a signal is reconstructed from noisy measurements. The basic idea of noise cancellation is to use the noise corrupted signal with the aid of a reference input to regenerate the signal of interest [21]. This approach is computationally cheap compared with the so-called Kalman filter design. The adaptive noise cancellation approach has been widely used in many applications such as speech processing, echo cancellation and antenna side-lobe interference etc. [21][16][14][15].

The early work in adaptive noise cancellation was developed by Howells and Applebaum in 1957. In 1959, Widrow and Hoff developed the LMS algorithm, which has been extensively studied by many authors [22].

In applications where the nonlinear effect is mild linear filters provide an acceptable performance, but in general, linear filters are not adequate when applied to systems where nonlinear terms dominate. To overcome this problem some authors have considered nonlinear designs and a popular approach has been to use the Volterra series and the LMS algorithm [9][19]. The disadvantage of this approach is that it is computationally expensive since it requires a large parameter set and an alternative method based on the NARMAX model coupled with a suboptimal least squares (SOLS) algorithm was introduced in an attempt to alleviate this problem [6]. The NARMAX model provides a concise description for a wide class of nonlinear systems with a relatively small number of parameters and the new suboptimal least squares (SOLS) algorithm has an excellent convergence rate compared with the LMS algorithm. Irrespective of which particular algorithm is used there is a need to monitor the performance of the noise cancellation mechanism. This is easy to achieve in simulation where the true filter parameters and signal statistics are known but in practice this information will not be available to the experimenter. There are many situations which can cause the filter to operate suboptimally including, underspecification of the filter order, poor convergence, incorrect software implementation, a sudden change in the system structure, the use of a linear filter when the system is nonlinear etc. Ideally we require a metric of filter performance which is relatively easy to compute, simple to interpret and which will detect most of these fault conditions. The present study addresses this problem and introduces correlation validation tests to provide one method of performance monitoring for both linear and nonlinear noise cancellation.

2 Adaptive noise cancellation and system identification

The general structure of an adaptive noise canceller is illustrated in Figure(1) where $d(t)$ represents the measured signal which is composed of the signal of interest $s(t)$ and noise $n(t)$. It is assumed that $s(t)$ and $n(t)$ are uncorrelated and that the noise $n(t)$ passes through an unknown path designated by the transfer function $T(.)$ to produce the reference input $x(t)$. The concept is to fit a filter $F(.)$ operating on $x(t)$ to generate $y(t)$ which is an estimate of $n(t)$. The design incorporates on-line estimation of the filter parameters by minimizing the mean square error $E[\epsilon^2(t)]$. To ensure the causality of the filter $F(.)$ a delay t_d is inserted in the primary channel.

FIR (Finite Impulse Response) filters have been widely used in signal processing applications often in association with the LMS algorithm. Some authors have considered the design of FIR filters using the RLS (Recursive Least Squares) algorithm and demonstrated the superior convergence rate of this algorithm compared with LMS [7][10]. IIR (Infinite Impulse Response) filters have recently been interpreted in terms of a system identification problem

[11][6][13][17] which has been studied extensively in the last two decades [17]. The main advantage of IIR filters is that they can be used to model a wide class of linear systems with a smaller parameter set compared with FIR filters.

The noise cancellation problem can be viewed as a system identification problem by reorganizing the formulation in Figure(1). From Figure(1), the optimal filter $F(.)$ will be the inverse of $T(.)$, assuming that $T^{-1}(.)$ exists and is stable, such that

$$y(t) = n(t - t_d) = T^{-1}(x(t)) = F(x(t)) \tag{2..1}$$

But since $d(t - t_d) = n(t - t_d) + s(t - t_d)$ then

$$d(t - t_d) = T^{-1}(x(t)) + s(t - t_d) \tag{2..2}$$

Now $T^{-1}(x(t))$ can be considered as the system to be identified with input $x(t)$ and output $d(t - t_d)$ where $s(t - t_d)$ is interpreted as coloured noise as depicted in Figure(2). The model structure adopted for the filter $F(.)$ and the estimation algorithm used in the current investigation are described in the next section.

3 Nonlinear Filter Representation

The model that represents the filter $F(.)$ plays a key rule in the noise cancellation method. It is desirable that the model is linear-in-the-parameters to simplify the implmentation of the estimation algorithms and to reduce the computational load. Ideally the model should also be general enough to represent a wide range of nonlinear systems but should reduce to a linear filter if this is appropriate. There is no point in using a nonlinear filter if a linear filter design is adequate.

Although the finite Volterra series model has been widely used in both nonlinear system identification and noise cancellation the main disadvantage of this representation is that it requires a very large number of parameters to characterize even simple nonlinear systems. The problem arises because in the Volterra model the output is expressed in terms of past inputs only. The NARMAX representation avoids these problems by expanding the present output in terms of past inputs and outputs and can therefore be considered as a nonlinear IIR filter.

4 The NARMAX Model

The NARMAX model was initially introduced by Billings and Leontaritis in 1981 and studied further in (1985). A NARMAX description for the nonlinear noise cancellation problem can be derived by considering input-output maps based on conditional probability density function to yield

$$\begin{aligned} d(t-t_d) &= dc + f\{d(t-t_d-1), \ldots, d(t-t_d-n_d), x(t), \ldots, x(t-n_x), \\ &\quad s(t-t_d-1), \ldots, s(t-t_d)\} + s(t-t_d) \qquad (4..3) \end{aligned}$$

Expanding Equation(4..3) as a polynomial NARMAX and regrouping terms gives

$$
\begin{aligned}
d(t-t_d) &= dc + G^{dx}[d(t-t_d-1),\ldots,d(t-t_d-n_d),x(t),\ldots,x(t-n_x)] \\
&\quad + G^{dxs}[d(t-t_d-1),\ldots,d(t-t_d-n_d),x(t) \\
&\quad ,\ldots,x(t-n_x),s(t-t_d-1),\ldots,s(t-t_d-n_d)] \\
&\quad + G^{s}[s(t-t_d-1),\ldots,s(t-t_d-n_d)] + s(t-t_d) \qquad (4..4)
\end{aligned}
$$

where G[.] is a polynomial function. Equation(4..4) can be rewritten as

$$
\begin{aligned}
d(t-t_d) &= \Phi^T(t)\theta + s(t-t_d) \qquad (4..5) \\
&= \begin{bmatrix} \Phi^T_{dx}(t) & \Phi^T_{dxs}(t) & \Phi^T_{s}(t) \end{bmatrix} \begin{pmatrix} \theta_{dx} \\ \theta_{dxs} \\ \theta_{s} \end{pmatrix} + s(t-t_d) \qquad (4..6)
\end{aligned}
$$

where

$$
\begin{aligned}
G^{dx}[.] &= \Phi^T_{dx}(t)\theta_{dx} \\
G^{dxs}[.] &= \Phi^T_{dxs}(t)\theta_{dxs} \\
G^{s}[.] &= \Phi^T_{s}(t)\theta_{s}
\end{aligned}
$$

The SOLS (Sub-Optimal Least Squares) algorithm introduced in [4] for nonlinear systems gives unbiased estimates of the parameter vector θ. The number of parameters to be estimated would be reduced if $y(t)$ could be monitored such that Equation(4..3) could be expressed as

$$
d(t-t_d) = dc + f\{y(t-1),\ldots,y(t-n_d),x(t),\ldots,x(t-n_x)\} + s(t-t_d) \quad (4..7)
$$

to eliminate all cross-product noise terms. In practice $y(t)$ can not be measured directly and $y(t)$ in Equation(4..7) is therefore replaced by the predicted values $\hat{y}(t)$to define a SOLS algorithm for adaptive nonlinear noise cancellation.

The algorithm is computationally simple and can be presented in the following unified algorithm [20][6]:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)\epsilon(t) \tag{4..8}$$

$$K(t) = \frac{P(t-1)zz(t)}{\lambda(t) + \Phi^T P(t-1)zz(t)} \tag{4..9}$$

$$P(t) = [P(t-1) - \frac{P(t-1)zz(t)\Phi^T(t)P(t-1)}{\lambda(t) + \Phi^T(t)P(t-1)zz(t)}]/\lambda(t) \tag{4..10}$$

$$\epsilon(t) = d(t-t_d) - \Phi^T(t)\hat{\theta}(t-1) \tag{4..11}$$

where

$$\Phi(t) = \Phi_{\hat{y}x}(t) \tag{4..12}$$

$$\hat{\theta}(t) = \hat{\theta}_{dx}(t) \tag{4..13}$$

$$zz(t) = \Phi(t) \tag{4..14}$$

$$\lambda(t) \doteq \lambda_0\lambda(t-1) + (1-\lambda_0) \tag{4..15}$$

The structure of the model is determined by n_d, n_x and the degree of f in Equation(4..7) and an appropriate assignment of these terms should give optimum filter performance. However the structure of the system is unknown apriori and whilst various trial values can be used for n_d, n_x and the degree of f a method of detecting correct filter operation is needed to aid the experimenter in selecting values for these parameters.

5 Performance Monitoring

There are many situations in noise cancellation applications where performance monitoring would be valuable. If, for example, there is a noticeable nonlinearity in the system then there might be a danger of ignoring the nonlinear terms and using an inappropriate linear design. Alternatively there may be situations where the filter parameters do not converge to the optimal values leading to poor filter performance. Or the filter parameter set may not include all the necessary terms required to effectively characterize the system. In each case some sort of monitoring scheme which ideally can both detect and discriminate between these effects and which is independent of the precise filter structure would be invaluable. Most of the performance monitoring measures used in signal processing are a function of $\epsilon(t)$ such as the mean squared error, the expected signal-to-noise ratio or measures such as $E(y^2)/E(e^2)$ [22][12]. These methods can not be readily applied on-line since they depend on an apriori knowledge of the signal $s(t)$ or on a comparison of different filter realizations which can only be done in simulation. In general these methods can not distinguish between linear and nonlinear effects. Correlation functions which may offer a solution to these problems are considered in the next section.

5.1 Correlation Tests

The linear cross-correlation between the input and the error signal was suggested in [18] as an automatic gain control for FIR filters. Correlation functions were also used in [8] to monitor the operation of linear self-tuners. In the present study this concept is applied to noise cancellation and extended to work for both linear and nonlinear cases.

The problem of detecting a deficiency in the filter operation can initially be studied by considering the minimization of the cost function. Hence

$$\begin{aligned} minE(\epsilon^2) &= minE([d(t) - y(t)]^2) && (5.1.1) \\ &= minE([n(t) + s(t) - y(t)]^2) && (5.1.2) \\ &= minE(n^2(t) + 2n(t)s(t) - 2n(t)y(t) && (5.1.3) \\ &\quad + s^2(t) - 2s(t)y(t) + y^2(t)) \end{aligned}$$

Since $n(t)$ and $s(t)$ are uncorrelated, $E[n(t)s(t)] = 0$ and $E[s(t)y(t)] = 0$, then

$$\begin{aligned} minE(\epsilon^2) &= minE(n^2(t) - 2n(t)y(t) + s^2(t) + y^2(t)) && (5.1.4) \\ &= minE([n(t) - y(t)]^2 + s^2(t)) && (5.1.5) \end{aligned}$$

The minimum occurs when $y(t) \longrightarrow n(t)$ and consequently $\epsilon(t) \longrightarrow s(t)$ in the mean squared sense. This is the condition for the canceller to perform optimally.

When $\epsilon(t)$ fails to converge to $s(t)$, then in general for any linear-in-the-parameters model including the linear (FIR or IIR), Volterra or NARMAX models the cause may be one of or a combination of one of the following conditions:

(i) the parameter estimates have not converged to the proper values, that is

$$\hat{\theta}^x = \theta_0^x + \Delta\theta^x \qquad\qquad \Delta\theta^x \neq 0$$

Then $\epsilon(t)$ will take the form

$$\epsilon(t) = s(t) + \Phi_x^T(t)\Delta\theta^x$$

(ii) the fitted model has erroneous structure, that is, some linear or nonlinear terms are missing. For example if

$$n(t) = a_1x(t) + a_2x(t-1) + a_3x(t)x(t-1)$$

and the model

$$y(t) = \theta_1x(t) + \theta_2x(t-1)$$

is fitted then $\epsilon(t)$ will be

$$\epsilon(t) = s(t) + a_3x(t)x(t-1) + \Delta\theta_1x(t) + \Delta\theta_2x(t-1)$$

where $\Delta\theta_1$ and $\Delta\theta_2$ may or may not equal zero.

In all cases if the filter is operating correctly then the residuals will be unpredictable from all linear and nonlinear combination of input and output terms. The validity of this statement can be established by using previous results by the authors [3][5] derived for system identification and adapting these to the filter monitoring situation.

It is well known that for linear systems [17], the estimated system parameters will be unbiased if

$$E(\epsilon(t)\epsilon(t-\tau)) = \Phi_{\epsilon\epsilon}(\tau) = \delta(\tau) \tag{5.1.6}$$

$$E(x(t)\epsilon(t-\tau)) = \Phi_{x\epsilon}(\tau) = 0 \tag{5.1.7}$$

If the system is nonlinear however these tests are inadequate [3][5]. In the system identification case where a process and noise model are estimated, it has been proved [3][5], that the residuals $\epsilon(t)$ will be unpredictable from all linear and nonlinear combinations of past inputs $x(t)$ and outputs iff

$$\begin{aligned}
\Phi_{\epsilon\epsilon}(\tau) &= \delta(\tau) \;\; \forall\tau && (5.1.8)\\
\Phi_{x\epsilon}(\tau) &= 0 \;\; \forall\tau && (5.1.9)\\
\Phi_{\epsilon\epsilon x}(\tau) &= 0 \;\; \tau \geq 0 && (5.1.10)\\
\Phi_{x^{2'}\epsilon}(\tau) &= 0 \;\; \forall\tau && (5.1.11)\\
\Phi_{x^{2'}\epsilon^2}(\tau) &= 0 \;\; \forall\tau && (5.1.12)
\end{aligned}$$

Whenever the noise is purely additive at the output alternative algorithms such as SOLS and IV can be used. This latter case corresponds to the noise cancellation problem and is appropriate for models of the form

$$d(t) = G^x[x(t)] + G^{xs}[x(t), s(t)] + G^s[s(t)]$$

where $G^s[s(t)]$ represents the additive noise which may be coloured. Notice that if only the first term were present in the above equation then the model would be a Volterra series.

If either a SOLS or IV algorithm is used then the presence of any term of $G^x[x(t)]$ or $G^{xs}[x(t), s(t)]$, but not $G^s[s(t)]$, in the residuals would induce biased estimates. Only a subset of the tests in Equations(5.1.8-5.1.12) are therefore required in this case and it can be shown that [5] unbiased estimates will be obtained iff

$$\Phi_{x\epsilon}(\tau) = E[x(t)\epsilon(t+\tau)] = 0 \tag{5.1.13}$$

$$\Phi_{x^{2'}\epsilon}(\tau) = E\{[x^2(t) - \overline{x^2(t)}]\epsilon(t+\tau)\} = 0 \tag{5.1.14}$$

$$\Phi_{x^{2'}\epsilon^2}(\tau) = E\{[x^2(t) - \overline{x^2(t)}]\epsilon^2(t+\tau)\} = 0 \tag{5.1.15}$$

where $(.)'$ indicates that the mean value has been removed.

Although these tests were derived for the system identification case [5], the equivalence of this problem and noise cancellation established in section(2) shows that they are appropriate in the present application. The correlation tests of Equations(5.1.13-5.1.15) can therefore be used in performance monitoring for noise cancellation. A summary of the interpretation of the tests is given in Table(1) subject to the conditions defined in [5].

Notice that the tests of Equations(5.1.13-5.1.15) are standard linear correlations of nonlinear functions of x and ϵ. They can therefore be computed using standard correlation algorithms or correlators.

$\Phi_{x^{2'}\epsilon^2}(\tau)$	$\Phi_{x^{2'}\epsilon}(\tau)$	$\Phi_{x\epsilon}(\tau)$	Comments
$= 0$	$= 0$	$= 0$	Filter model unbiased
$\neq 0$	$\neq 0$	$\neq 0$	Wrong filter model structure
$\neq 0$	$\neq 0$	$= 0$	Even power of $x(t)$ omitted from the model if odd order moment of $x(t)$ is zero
$\neq 0$	$= 0$	$\neq 0$	Odd power of $x(t)$ omitted from the model if odd order moment of $x(t)$ is zero
$= 0$	$= 0$	$\neq 0$	No nonlinearities. Linear term missing

Table 1: Interpretation of the Correlation tests subject to the conditions specified in (Billings and Voon 1986).

6 Performance tests for Nonlinear Adaptive Noise Cancellation

The discussion in the previous section demonstrated that the correlation functions of Equations(5.1.13-5.1.15) are appropriate tests for linear and nonlinear noise cancellation. The only problem that remains is that in the identification application these tests were used off-line whereas in noise cancellation an on-line implementation would be preferable. One solution to this problem is to measure the values of $\epsilon(t)$ and $x(t)$ on-line for a reasonable number of iterations. Equations(5.1.13-5.1.15) can then be calculated over this window of iterations parallel with cancellation. This can easily be accomplished using parallel processing techniques. Although the correlation tests work well in the standard parameter estimation case [5], it is likely that the sensitivity of detection will be reduced when used on-line and very small windows will clearly increase the probability of false alarms.

Ideally the tests should provide an indication of filter operation and should indicate if some modification either to the structure of the filter (i.e., include nonlinear terms) or to other factors such as increasing the gain of the algorithm is required in order to achieve satisfactory filter performance.

6.1 On-line Performance tests

One way of implementing the performance tests on-line is to recursively estimate the correlation functions of Equations(5.1.13-5.1.15). Define the cross-correlation function of the discrete-valued vectors a and b as

$$\phi_{ab}(\tau) = \frac{1}{N+1}\sum_{t=0}^{N} a(t-\tau)b(t). \tag{6.1.1}$$

Equation(6.1.1) can be updated recursively using the expression

$$\phi_{ab}(\tau, t) = \eta\phi_{ab}(\tau, t-1) + (1-\eta)a(t-\tau)b(t) \tag{6.1.2}$$

where $\eta = N/(N+1)$. All the tests of Equations(5.1.13-5.1.15) can be expressed in the form of Equation(6.1.2) with appropriate definition of a and b. The range of τ is related to the expected lags in the model which should be small for IIR filters. A good initial range for τ would be $0 \le \tau \le 5$ but this value can readily be increased if appropriate.

Alternatively, consider the function

$$\hat{\rho}_{ab}(t) = \eta\hat{\rho}_{ab}(t-1) + (1-\eta)\bar{a}_t b_t \tag{6.1.3}$$

where

$$\bar{a}_t = 1/m \sum_{i=1}^{m} a(t-i)$$

and η in these equations can be considered as a forgetting factor. So for example for $\eta = .995$, a data length of $N = .995/(1-.995) = 199$ will be remembered. A forgetting factor will be helpful in the transient period where the error $\epsilon(t)$ is nonstationary. The choice of η is often a compromise between slow detection when η is large (or less sensitive for large windows) and an increase in the possibility of false alarms when η is small.

It can be shown that $\hat{\rho}_{ab}$ of Equation(6.1.3) is related to ϕ_{ab} of Equation(6.1.1) by

$$\hat{\rho}_{ab}(t) = 1/m \sum_{\tau=0}^{m} \phi_{ab}(\tau, t). \tag{6.1.4}$$

Clearly Equations 5.1.14 and 5.1.15 for $\Phi_{x^{2'}\epsilon}(\tau)$ and $\Phi_{x^{2'}\epsilon^2}(\tau)$, can also be approximated as

$$\hat{\rho}_{x^{2'}\epsilon}(t) = \eta\hat{\rho}_{x^{2'}\epsilon}(t-1) + (1-\eta)\bar{c}_t\epsilon_t \tag{6.1.5}$$

$$\hat{\rho}_{x^{2'}\epsilon^2}(t) = \eta\hat{\rho}_{x_{2'}\epsilon^2}(t-1) + (1-\eta)\bar{c}_t\epsilon_t^2 \tag{6.1.6}$$

where

$$\bar{c}_t = 1/m \sum_{i=1}^{m} (x^2(t-i) - \bar{x^2}_m)$$

and $\bar{x^2}_m$ is the estimated mean value of $x^2(t)$ which can be calculated recursively. Any drift of the mean of $\hat{\rho}(t)$ away from zero or any sudden increase in $\hat{\rho}(t)$ should be interpreted as an increase in the corresponding correlation function implying poor filter performance. Although it is difficult to choose the threshold values in this case, the computations are kept to a minimum. Alternatively

Equations(5.1.13-5.1.15) can be computed over a window of data and used to interpret why the filter is performing suboptimally.

7 Computational Aspects

In all simulation studies the Suboptimal Least Square (SOLS) algorithm is implemented using the numerically stable factorisation method of Bieman (1977) [1]. Also the correlation functions described above are calculated according to the formula

$$\hat{\Phi}_{xy}(k) = \frac{\frac{1}{N}\sum_{t=1}^{N-k}(x(t)-\bar{x})(y(t+k)-\bar{y})}{\sqrt{\Phi_{xx}(0)\Phi_{yy}(0)}} \qquad -1 \leq \Phi_{xy}(k) \leq 1 \tag{7..7}$$

Confidence intervals plotted on the graphs indicate if the correlation between variables is significant or not. If N is large the standard deviation of the correlation estimate is $1/\sqrt{N}$ and the 95% confidence limits are therefore approximately $\pm 1.96/\sqrt{N}$.

8 Simulation Results

8.1 Simulation S1

The system S1

$$x(t) = n(t) - .5x(t-1) - .8x^2(t-1)$$

was simulated with a uniformly distributed random noise $n(t)$ between -0.5 and +0.5 and a triangular wave signal for $s(t)$ of magnitude 0.5 and a period 52. Initially, the noise canceller was operated with a filter defined by the structure

$$y(t) = a_1 x(t) + a_2 x(t-1)$$

for 500 iterations. The cross correlation functions were calculated to examine the performance of the filter and these are illustrated in Figure(3). In all cases the correlation functions have been calculated using Equations(5.1.13-5.1.15) over a window of length 500. Since the values of the linear cross correlation function $\Phi_{x\epsilon}(\tau)$ are within the 95% confidence intervals, a purely linear analysis would suggest that there is no more information in the residuals. The missing nonlinear term would therefore not be detected and a poor filter performance would be incorrectly attributed to a poor signal to noise ratio or some other false effect. Measurements of $E(\epsilon^2(t))$ can not detect that terms have been omitted because the actual minimum value is not known apriori. The sum of the errors may converge but this will be a biased value and it is impossible to know if this corresponds to the optimum filter design without an exhaustive search over all possible filter structures. However calculating the nonlinear functions $\Phi_{x^{2'}\epsilon}(\tau)$ shows a high correlation value at lag one as shown in Figure(3a). This suggests that the model is inadequate and a nonlinear term of lag one maybe missing from the filter model. Running the canceller for another 500 iterations, Figure(3b), confirmed this suggestion. It is always prefereable to apply the correlation monitor after the transient period since the initial stationarity as the parameters adjust may provide false results.

When the nonlinear term $x^2(t-1)$ is inserted in the model so the filter has the exact inverse of S1

$$y(t) = a_1x(t) + a_2x(t-1) + a_3x^2(t-1)$$

and the simulation is repeated, the values of all the correlation functions, Figure(3d), are inside the 95% confidence intervals indicating that the filter is performing correctly.

The on-line approximation of the correlation functions were tested on this example for the detection of sudden changes with $\eta = .995$ and $m = 10$ in Equations(6.1.3-6.1.6). The canceller was started with the correct filter structure and run for 1000 iterations. The filter was then replaced by the reduced order structure to simulate a sudden change in the system structure. The on-line tests which are illustrated in Figure(4) show that after the initial transient period (0-500) the values of $\hat{\rho}(t)$ settle down to small magnitudes around zero. But within just a few iterations following iteration 1000 the magnitudes rise significantly correctly indicating the change.

8.2 Simulation S2

A second nonlinear system S2 defined by

$$x(t) = n(t) - .2x(t-1) - x(t-1)n(t-1) + .1n(t-1) + .4x(t-2)$$

was simulated with a uniformly distributed random noise $n(t)$ with values between -0.5 and +0.5 and a triangular wave signal for $s(t)$ of magnitude 0.25 and period 52.

A linear filter of the structure

$$y(t) = a_1x(t) + a_2x(t-1) + a_3y(t-1)$$

was initially used for the model. The correlation functions are again calculated for two successive windows. The values of the linear and nonlinear correlation functions dipicted in Figure(5a) and (5b) are outside the confidence intervals for some lags mainly lag two in $\Phi_{x\epsilon}(\tau)$ and lag one in $\Phi_{x^{2'}\epsilon}(\tau)$ and $\phi_{x^{2'}\epsilon^2}(\tau)$. This indicates that some terms have been ommited from the model.

When a linear term at lag two and a nonlinear term at lag one are added to give the exact inverse of s2

$$y(t) = a_1x(t) + a_2x(t-1) + a_3y(t-1) + a_4x(t-2) + a_5x(t-1)y(t-1)$$

all the corrrelation functions are satisfied, Figure(5d), suggesting that the canceller is performing adequately and a very high noise suppression is obtained as demonstrated in Figure(6).

Again, the simple on-line tests were repeated in this case. For the first 1000 iterations the filter structure was the exact inverse of the system. At iteration 1000, the two terms $x(t-2)$ and $x(t-1)y(t-1)$ were removed from the filter structure. The on-line plots of $\hat{\rho}(t)$ in Figure(7) correctly show a considerable increase following this rapid change.

8.3 Simulation S3

A third system S3 defined by

$$x(t) = .3x(t-1) + .1x(t-2) + n(t-1) + .4n^2(t-2) + .1n(t-2)x(t-1)$$

was simulated with a uniformly distributed random noise $n(t)$ with values between -1 and +1. The signal $s(t)$ was a triangular wave with magnitude of 0.2 and a period of 52, and t_d was set to be $t_d = 1$.

Initially a filter structure of the form

$$y(t) = a_1x(t) + a_2x(t-1) + a_3x(t-2) + a_4y(t-1)x(t-1)$$

was used. The correlation functions are plotted in Figure(8b). Clearly the values of the correlation functions are outside the permitted intervals indicating that the is not operating correctly. Adding one more nonlinear term to the filter gave a filter structure of the form

$$y(t) = a_1x(t) + a_2x(t-1) + a_3x(t-2) + a_4y^2(t-1) + a_5x(t-1)y(t-1)$$

which is the exact inverse of S3. The filter performance is improved but some values of the correlations are still slightly high in Figure(8c). This is probably due to the transient effects which occur as the filter parameters are tuned. Computing over the next 500 samples, calculated between iteration 500 and iteration 1000, Figure(8d), shows that the correlation values are now well inside the confidence intervals.

These tests will also be valid for the case of model overfitting providing the true model is a subset of the fitted model. In simulations the tests were shown to work for a window length of 500, however it can work for much less than this value and maybe down to 100.

9 Conclusions

In this article three simple correlation tests have been proposed to monitor the performance of linear and nonlinear adaptive noise cancellers. Although the canceller design was based on the NARMAX representation and the SOLS algorithm, the results are not restricted to these cases and apply to all linear and analytic nonlinear systems using either LMS or RLS based estimators. The tests perform well for simulated systems and were shown to detect and distinguish between missing linear and nonlinear terms in the model. Simple on-line measures have also been designed to monitor filter performance. These computationally simple tests were shown to be useful especially for detecting sudden structural changes in the system, but the derivation of confidence intervals for these tests requires further study.

References

[1] G. BIERMAN 1977 *Factorisation Methods For Discrete Seqential Estimation.* New York Academic Press.

[2] S. A. BILLINGS AND LEONTARITIES 1981 *Institute of Electrical Engineers Conference*, 183-190. Identification of nonlinear systems using parameter estimation techniques.

[3] S. A. BILLINGS AND VOON 1983 *Institute of Electrical Engineers Proceedings*, 193-199. Structure detection and model validity tests in the identification of nonlinear systems.

[4] S. A. BILLINGS AND VOON 1984 *International Journal of Systems Science*, 601-615. Least square parameter estimation algorithms for nonlinear systems.

[5] S. A. BILLINGS AND VOON 1986 *International Journal of Control*, 235-244. Correlation based model validity tests for nonlinear models.

[6] S. A. BILLINGS, M. ELHITMY AND A. CURTIS 1990 *Journal of Sound and Vibration* 140(3), 457-473. Adaptive noise cancellation for a class of NIIR filters.

[7] J. CIOFFI AND T. KAILATH 1984 *Institute of Electrical and Electronic Engineers Transactions* ASSP-32, 304-337. Fast Recursive-Least-Squares transversal filters for adaptive filtering.

[8] D. CLARKE, P. GATHROP 1975 *Institute of Electrical Engineers Proceedings*, 929-934. Self-tuning controller.

[9] M. COKER AND D. SIMKINS 1980 *Institute of Electrical and Electronic Engineers Proceedings*, 470-473. A nonlinear adaptive noise canceller.

[10] C. COWAN 1987 *Institute of Electrical Engineers Proceedings* 134, 211-215. Performance comparison of finite linear adaptive filters.

[11] H. FAN AND W. JENKINS 1986 *Institute of Electrical and Electronic Engineers Transactions* CAS-33, 939-947. A new adaptive IIR filter.

[12] H. FAN AND W. JENKINS 1988 *Institute of Electrical and Electronic Engineers Transactions* ASSP-36, 1820. An investigation of an adaptive IIR echo canceller: advantages and problems.

[13] G. GOODWIN AND K. SIN 1984 *Adaptive Filtering, Prediction And Control.* Prentice-Hall.

[14] W. HARRISON, J. LIM AND E. SINGER 1986 *Institute of Electrical and Electronic Engineers Transactions* ASSP-3, 21-22. A new application of adaptive noise cancellation.

[15] P. HILL 1985 *MSEE Thesis, West Virginia University.* Real time real life adaptive noise cancellation using the homogenious agorithm.

[16] G. KANG AND L. FRANSEN 1986 *Institute of Electrical and Electronic Engineers Transactions* CAS-34, 753-758. Experimentation with an adaptive noise cancellation filter.

[17] L. LJUNG, T. SODERSTROM 1983 *Theory And Practice Of Recursive Identification.* MIT Press.

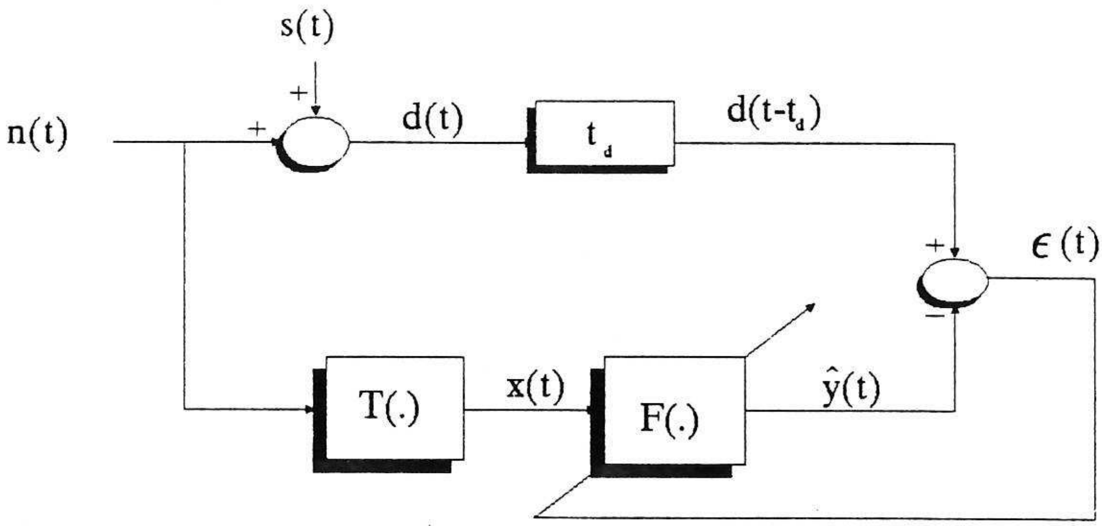
[18] T. J. SHAN AND T. KAILATH 1988 *Institute of Electrical and Electronic Engineers Transactions*, 122-127. Adaptive algorithms with automatic gain control feature.

[19] G. SICURANZA, A. BUCCONI AND P. MITRI 1984 *Institute of Electrical and Electronic Engineers Proceedings*, 3101-3104. Adaptive echo cancellation with nonlinear digital filters.

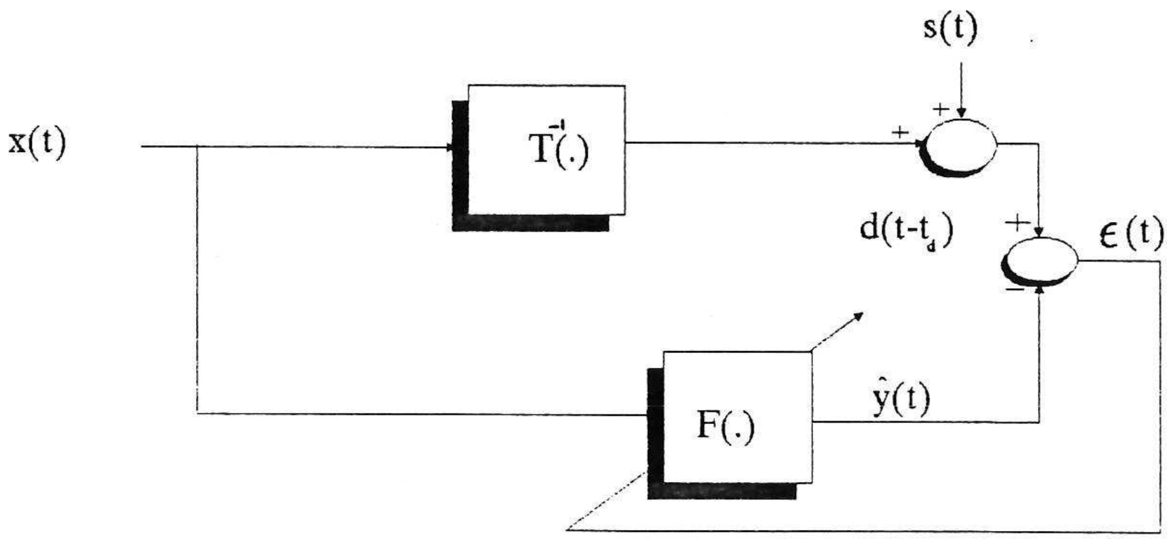
[20] T. SODERSTROM, L. LJUNG AND I. GUSTAVSSON 1978 *Automatica* 14, 231-244. A theoretical analysis of recursive identification methods.

[21] B. WIDROW, J. GLOVER, J. MCCOOL, J. KAUNITZ, C. WILLIAMS, R. HEARNS, J. ZEIDLER, E. DONG AND R. GOODLIN 1975 *Institute of Electrical and Electronic Engineers Proceedings* 63, 1692-1716. Adaptive noise cancelling: principles and applications.

[22] B. WIDROW AND S. STEARNS 1985 *Adaptive Signal Processing.* Englewood Cliffs, New Jersey: Prentice Hall.



Figure(1) Typical Noise Canceller.



Figure(2) Output Error Identification

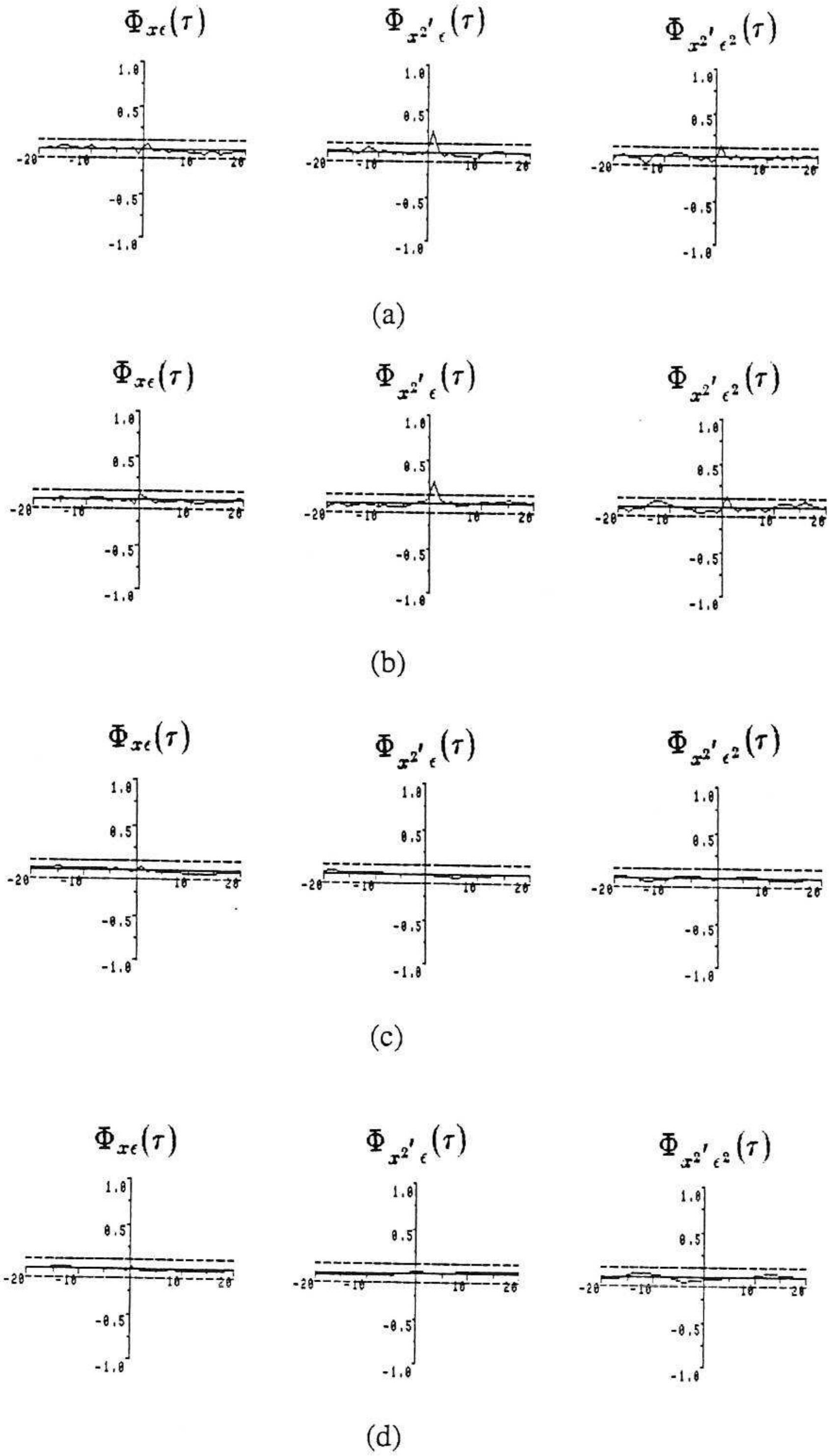


Figure (3) Correlation Functions for S1 and
a: Reduced Order (iterations 0-500)
b: Reduced Order (iterations 500-1000)
c: Full Order (iterations 0-500)
d: Full Order (iterations 500-1000)

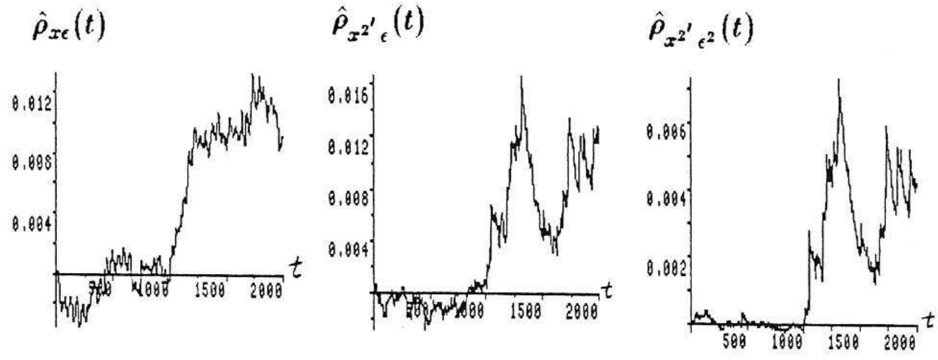


Figure (4) The Estimates of The Correlation Functions For S1

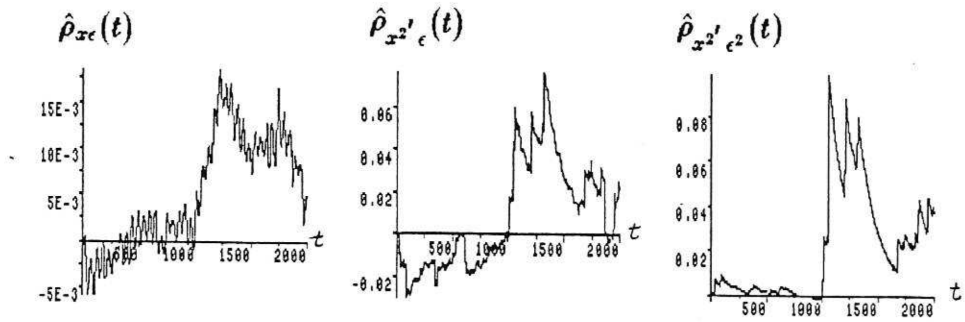


Figure (7) The Estimates of The Correlation Functions For S2

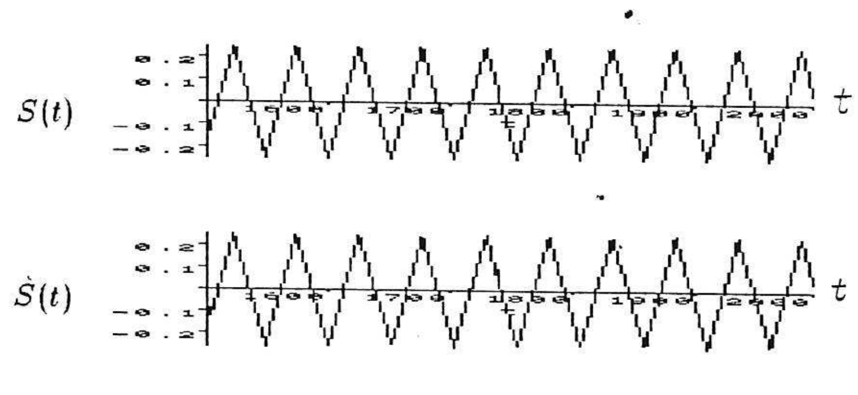
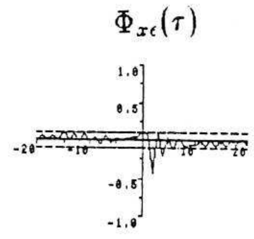
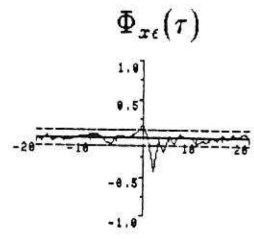


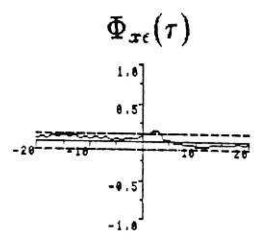
Figure (6) S(t) and Estimation of S(t)



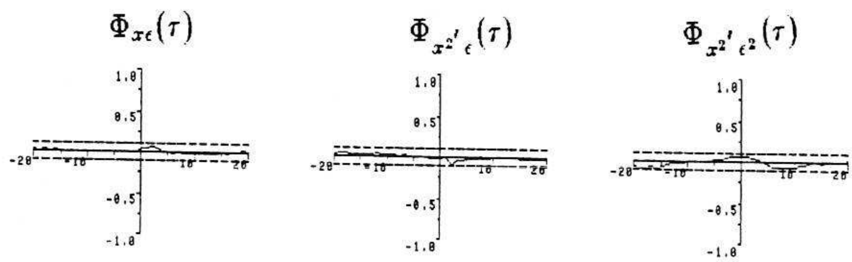
(a)



(b)

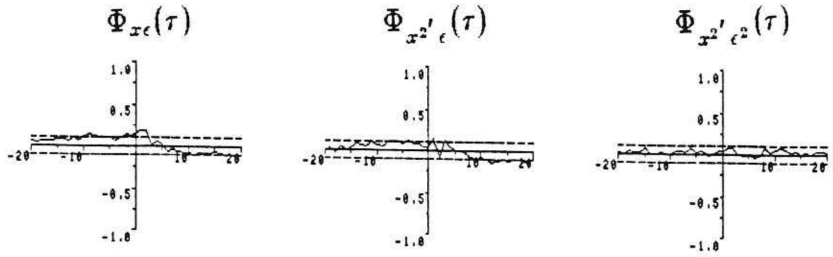


(c)

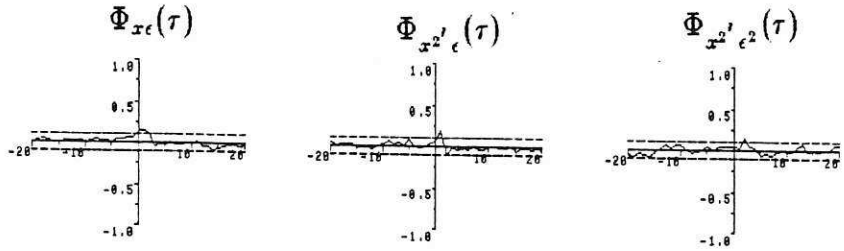


(d)

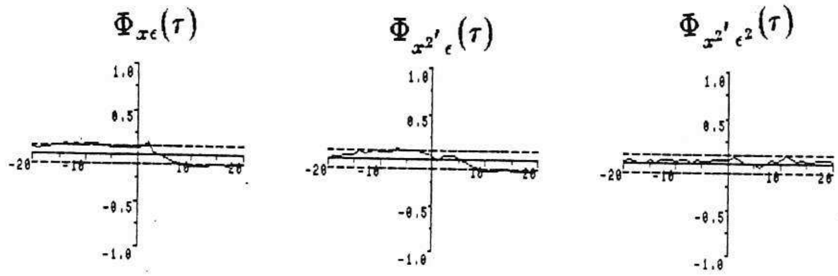
Figure (5) Correlation Functions for S2 and
a: Reduced Order (iterations 0-500)
b: Reduced Order (iterations 500-1000)
c: Full Order (iterations 0-500)
d: Full Order (iterations 500-1000)



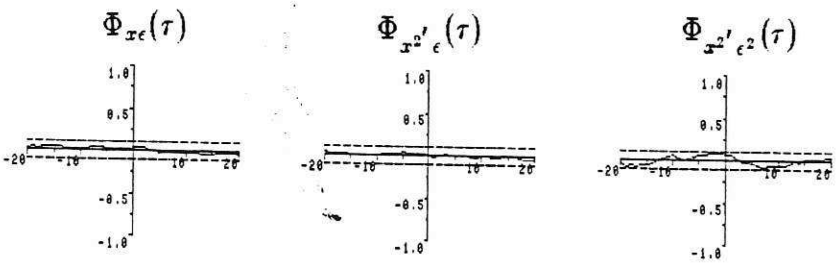
(a)



(b)



(c)



(d)

Figure (8) Correlation Functions for S3 and
a: Reduced Order (iterations 0-500)
b: Reduced Order (iterations 500-1000)
c: Full Order (iterations 0-500)
d: Full Order (iterations 500-1000)