An Adaptive Filter for Noise Cancelling

AHMED S. ABUTALEB

***Abstract*—This paper introduces a new nonlinear filter that is used for adaptive noise cancelling. The derivation and convergence properties of the filter are presented. The performance, as measured by the signal to noise ratio between the signal and its estimate, is compared to that of the commonly used least mean square (LMS) algorithm. It is shown, through simulation, that the proposed nonlinear noise canceller has, on the average, better performance than the LMS canceller. The proposed adaptive noise canceller is based on Pontryagin minimum principle and the method of invariant imbedding. The computational time for the proposed method is about 10 percent that of the LMS, in the studied cases which is a substantial improvement.**

***Key Words*—Adaptive noise cancelling, Pontryagin minimum principle, invariant imbedding, nonlinear estimation, LMS algorithm, adaptive filtering.**

I. Introduction

NOISE CANCELLING is a variation of optimal filtering that is highly advantageous in many applications [1]–[3]. It makes use of an auxiliary or reference input derived from one or more sensors located at points in the noise field where the signal is weak or undetectable. This input is filtered and subtracted from a primary input containing both signal and noise (which is a filtered version of the reference input). As a result the primary noise is attenuated or eliminated by cancellation and the desired signal becomes the output of the process.

In typical implementations of adaptive noise cancelling (ANC) the filtering of the reference input is achieved using a finite impulse filter (FIR). The filter coefficients are then adjusted using a gradient search algorithm such as the least mean square (LMS) [2], [4] that minimizes a specified cost function. Another approach [5], [6] is to model the interference (in the primary input) as the output of an infinite impulse response (IIR) filter (with unknown coefficients) driven by the reference input and white noise. The IIR filter coefficients are adjusted using, e.g., recursive prediction error algorithm [7]. Hence an estimate of the interference signal is obtained and subtracted from the primary input to yield an estimate of the signal. This process is equivalent to passing the primary input through an IIR filter to obtain the signal. These implementations of the ANC perform well in many situations. Their performance is, however, suboptimal due to the following: 1) the FIR or IIR filter structure (both are linear time invariant) are only approximation of the optimal filter (which may be nonlinear time varying). The quality of the approximation depends on the filter order as well as the signal and interference properties; 2) the search algorithms sometimes converge very slowly when the interference covariance matrix has large spread of its eigenvalues [8]; 3) the use of high-order filters is not uncommon to achieve good performance which results in increased computation time; 4) the performance of FIR or IIR filters degrades rapidly as the SNR decreases, and 5) the performances degrades as the nonstationarity in the desired signal and/or the reference input increases. These drawbacks were the motivation to find a difference ANC that must have the following properties: a) In modelling the signal or interference, the model that uses the minimum number of states must be used, thus, avoiding the problem of having to estimate a larger number of coefficients as in IIR- or FIR-based methods. The model must be more general than FIR or IIR. These objectives can be achieved if the process is modelled as the output of a linear time-varying filter driven by white noise. b) The time-varying filter (which models the process) has an unknown structure which should not be estimated explicitly, in the interest of reducing computation time. This can be achieved through a transformation to be described shortly. c) The resulting adaptive filter should not be assumed to have FIR or IIR structure. An approach to the problem using optimal control concepts, or more specifically the Pontryagin minimum principle [9], yields a nonlinear filter.

This report presents a new approach for cancelling interference from signal measurements that possess the described properties. In the proposed method, both the signal and the interference are modeled as nonstationary processes. The major contributions of this report are: 1) modeling of the signals under study as outputs of a linear time-varying filter driven by the reference input and white noise; 2) the introduction and derivation of a new nonlinear filter that is well suited to this kind of model; 3) the application of these concepts to the problem of noise cancelling.

In the proposed method, the interference is modeled as a nonstationary process. The Pontryagin minimum principle [9] is used, in a novel way, to derive a set of nonlinear equations that relate the interference in the primary to the reference input. This approach resulted in a two-point boundary value problem which cannot be solved in real time. Using the method of invariant imbedding [10], the two point boundary value problem can be transformed into one point boundary value problem, and the resulted

Manuscript received December 31, 1986; revised February 18, 1988. This paper was recommended by Associate Editor K. W. Martin.

The author is with the Department of Electrical Engineering, College of Engineering and Architecture, Temple University, Philadelphia, PA 19122.

IEEE Log Number 8822595.

equations can be solved in real time. The solution of these equations yield the estimates of the signal and interference.

The proposed method was compared to the LMS-based algorithm of Widrow [2], [4], where it was shown to have superior performance. The chosen criterion is measured by the error between the real signal and its estimated value. One example is given: The desired signal is sinusoidal with time-varying random phase, while the interference is sinusoidal with zero phase and with different frequency. The primary interference is a linearly filtered version of the reference. The computer time needed to arrive at the estimates using the proposed method, was approximately 10 percent of the time for the LMS algorithm.

In Section II, the proposed Pontryagin ANC is introduced. In Section III, the properties of the derived filter or Pontryagin ANC are developed. In Section IV, simulation and results are provided and discussed. Finally in Section V, discussion and conclusions are presented.

II. Pontryagin ANC

A. The General ANC Concept

Fig. 1 shows the basic problem and the general ANC.

A signal $s(t)$ is transmitted over a channel to a noisy sensor that receives the signal and adds an independent noise component $v_2(t)$. The sensor also receives an independent interference, $n_1(t)$, which is passed through a linear filter $h(t)$, and added to the signal. The combined signal, interference, and receiver noise form the primary input, $r_2(t)$, to the ANC. Another sensor receives only the interference $n_1(t)$ in addition to an independent noise component $v_1(t)$. The combined interference and receiver noise form the reference input to the ANC. The adaptive filter operates on the reference input to generate an estimate of the interference, $\hat{x}_1(t)$, which is then subtracted from the primary input, $r_2(t)$, to yield an estimate of the signal $\hat{s}(t)$. Since the characteristics of the filter, $h(t)$, is usually unknown, an adaptive filtering approach proved to be superior to any fixed filter method. The conventional adaptive filter is basically an FIR filter [4] or an IIR [5]. The filter parameters can be adjusted using one of many algorithms, e.g., least mean square (LMS) [4]. By minimizing a cost function, $E\{[r_2(t)-\hat{x}_1(t)]^2\}$, an estimate of the desired signal, $\hat{s}(t)$, can be obtained. The method is explained in detail in [2]. The procedure used to minimize the cost function and the signal model determines the shape of the adaptive filter.

B. ANC Using Pontrygain Minimum Principle

In this subsection we present an alternative approach to ANC that is based on optimal control concepts. The improvement in performance using this approach is due mainly to the introduction of a new nonlinear filter which is called Pontryagin filter. The interference in the primary channel $x_1(t)$, is modeled as the output of a linear time varying filter driven by the reference input, $r_1(t)$, and a white noise component. By minimizing a particular cost function that is subject to some dynamic constraints, the new ANC will be derived. First the linear time varying model is introduced.

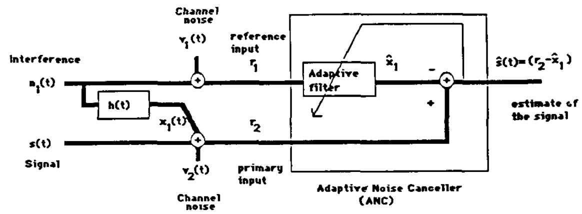

Fig. 1. The general adaptive noise canceller.

The signals will be modeled as outputs of a linear time varying filter driven by white noise. This is completely different from the autoregressive (AR) or autoregressive moving average (ARMA) models which are commonly used. In the present approach, only two time-varying functions have to be estimated, which reduces the computation time drastically. The accuracy of the filtering operation, however, will not be degraded. Actually it will improve due to the fact that a time-varying model may better represent the real situation in many cases.

Let the interference in the primary channel, $x_1(t)$, be the output of a linear time-varying filter that is described by the following differential equation:

$$\dot{x}_1(t)=a_1(t)x_1(t)+a_2(t)r_1(t)+w_1(t) \tag{1}$$

where

$x_1(t)$	interference at the primary channel at time t,
$r_1(t)$	reference input or received signal at the interference channel at time t,
"˙"	derivative with respect to time,
$w_1(t)$	white Gaussian noise with energy level 1,
$a_1(t), a_2(t)$	unknown time functions that represent the time-varying filter.

The observation available at the receivers are

$$r_1(t)=n_1(t)+v_1(t) \tag{2}$$

and

$$r_2(t)=s(t)+x_1(t)+v_2(t) \tag{3}$$

where

$r_1(t)$	reference input at time t,
$n_1(t)$	interference,
$v_1(t)$	observation noise of the interference sensor and channel that is independent of $x_1(t)$, $n_1(t)$, $w_1(t)$ and $v_2(t)$. The energy level is dependent upon the interference to noise ratio of the channel,
$r_2(t)$	primary input at time t,
$s(t)$	desired weak signal,
$x_1(t)$	interference at primary channel ($=n_1(t)*h(t)$),

where

- "$*$" convolution operation,
- $h(t)$ transfer function between the interference, at the reference channel and the interference at the primary channel,
- $v_2(t)$ observation noise of the primary receiver and channel, it is independent of $x_1(t)$. The energy level is determined by the signal-to-noise ratio of the channel.

The ANC problem has been reformulated using a state space representation. The objective is to find the signal estimate, $\hat{s}(t)$. It is clear that estimates of the filter time varying parameters are needed ($a_1(t)$ and $a_2(t)$). These estimates will be obtained implicitly such that no extra calculations are needed, which can be achieved through minimization of the cost function.

C. Estimation of the Interference

The objective is to estimate (or filter) the interference, $x_1(t)$, that is governed by (1) using the observations $r_1(t)$ and $r_2(t)$. The approach presented here is based on Pontryagin minimum principle and the method of invariant imbedding [9], [10], [12]. The basic idea is to interpret $a_1(t)$, $a_2(t)$, and $w_1(t)$ as unknown controllers that attempt to drive the system state $x_1(t)$ to follow (or track) a trajectory that is given by the noisy observation $r_2(t)$ of (3). Thus the stochastic estimation problem is transformed into a deterministic optimization (or tracking) problem with dynamic constraints, which lends itself easily to application of the Pontryagin minimum principle. It is pointed out that a duality between estimation and optimization can be exploited [12] (for example, the Kalman filter was derived using Pontryagin minimum principle [13]).

A detailed derivation is given because we are presenting a new nonlinear filter. Define a cost function (that is actually the squared error):
Let

$$J=1/2\int_0^T d\delta\{[r_2(\delta)-x_1(\delta)]^2+\mu_{a_1}a_1^2(\delta)+\mu_{a_2}a_2^2(\delta)+\mu_{w_1}w_1^2(\delta)\} \quad (4)$$

where T is the running current time, μ_{a_1}, μ_{a_2}, and μ_{w_2} are weights that represent the *a priori* knowledge about the range of variations of $a_1(t)$, $a_2(t)$, and $w_1(t)$, respectively. These weights affect speed of convergence and accuracy of the estimation procedures, they are similar to the constant gain parameter, α, of the LMS algorithm. Their values are chosen according to some approximate rule that was derived based on convergence study of the Pontryagin filter. This rule is stated in the results section. Notice that the cost function is similar to that used in conventional ANC.

The problem now can be stated as follows: Find the values of $a_1(t)$, $a_2(t)$, $w_1(t)$, and $x_1(t)$ that minimize the cost function J in (4), subject to the dynamic constraints of (1).

Using the Pontryagin minimum principle [13], the solution is obtained as follows:

i) Form Pontryagin H Function:

$$\begin{aligned}H(x_1,r_1,r_2,a_1,a_2,w_1,\lambda)&=H=1/2[r_2(t)-x_1(t)]^2\\&+1/2\mu_{a_1}a_1^2(t)+1/2\mu_{a_2}a_2^2(t)+1/2\mu_{w_1}w_1^2(t)\\&+\lambda(t)[a_1(t)x_1(t)+a_2(t)r_1(t)+w_1(t)] \quad (5)\end{aligned}$$

where $\lambda(t)$ is known as the Lagrange multiplier or the costate.

ii) Minimize H with Respect to the Unknown Controllers $a_1(t)$, $a_2(t)$ and $w_1(t)$:

$$\partial H/\partial a_1(t)=0=\mu_{a_1}a_1(t)+\lambda(t)x_1(t) \quad (6)$$

which yields

$$a_1(t)=-\lambda(t)x_1(t)/\mu_{a_1} \quad (7)$$

for $a_2(t)$ we get

$$\partial H/\partial a_2(t)=0=\mu_{a_2}a(t)+\lambda(t)r_1(t) \quad (8)$$

which yields

$$a_2(t)=-\lambda(t)r_1(t)/\mu_{a_2} \quad (9)$$

and for $w_1(t)$ we get

$$\partial H/\partial w_1(t)=0=\mu_{w_1w_1}(t)+\lambda(t) \quad (10)$$

which yields

$$w_1(t)=-\lambda(t)/\mu_{w_1}. \quad (11)$$

iii) Substitute the Values obtained for $a_1(t)$, $a_2(t)$, and $w_1(t)$ to obtain the minimum H function (H^*):

$$\begin{aligned}H^*&=1/2[r_2(t)-x_1(t)]^2+1/2\mu_{a_1}[\lambda(t)x_1(t)/\mu_{a_1}]^2\\&+1/2\mu_{a_2}[\lambda(t)r_1(t)/\mu_{a_2}]^2+1/2\mu_{w_1}[\lambda(t)/\mu_{w_1}]^2\\&+\lambda(t)[-\lambda(t)x_1^2(t)/\mu_{a_1}\\&-\lambda(t)r_1^2(t)/\mu_{a_2}-\lambda(t)/\mu_{w_1}] \quad (12)\end{aligned}$$

which is reduced to

$$\left(H^*=1/2[r_2(t)-x_1(t)]^2-(\lambda^2(t)/2)\cdot[x_1^2(t)/\mu_{a_1}+r_1^2(t)/\mu_{a_2}+1/\mu_{w_1}]. \quad (13)\right.$$

iv) Solve the following two Differential Equations:

$$\begin{aligned}\dot{x}_1(t)&=\partial H^*/\partial\lambda\\&=-\lambda(t)[x_1^2(t)/\mu_{a_1}^2+r_1(t)/\mu_{a_2}+1/\mu_{w_1}] \quad (14)\end{aligned}$$

and

$$\begin{aligned}\dot{\lambda}(t)&=-\partial H^*/\partial x_1(t)\\&=[r_2(t)-x_1(t)]+\lambda^2(t)x_1(t)/\mu_{a_1}. \quad (15)\end{aligned}$$

The initial condition $x_1(0)$ is assumed to be known with some uncertainty. The other initial condition $\lambda(0)$, however, is unknown. The final condition $\lambda(T)$, at the current time T, is known and equal 0 ($\lambda(T)=0$ because $x_1(T)$ is unknown), this result comes from the derivation of the minimum principle [13]. Thus, we have a two-point boundary value problem (TPBVP) that cannot be solved in real time. The solution of this problem will yield the desired estimate of $x_1(t)$. The method of invariant imbedding [10] is used to transform the two-point boundary value problem to one point boundary value problem that can be solved in real time. This is discussed in the next step.

v) Invariant Imbedding:

Rewrite (14) and (15) in a compact form as follows:

$$\dot{x}_1(t) = \Upsilon(x_1, \lambda) \tag{16}$$

and

$$\dot{\lambda}(t) = \eta(x_1, \lambda) \tag{17}$$

where

$$\Upsilon(x_1, \lambda) = -\lambda(t)\left[x_1^2(t)/\mu_{a_1} + r_1^2(t)/\mu_{a_2} + 1/\mu_{w_1}\right] \tag{18}$$

and

$$\eta(x_1, \lambda) = \left[r_2(t) - x_1(t)\right] + \lambda^2(t)x_1(t)/\mu_{a_1}. \tag{19}$$

The method of invariant imbedding [10], [12] yields the following differential equations with known initial conditions:

$$\dot{\hat{x}}_1(t) = \Upsilon(\hat{x}_1) + p_1(t)\eta(\hat{x}) \tag{20}$$

and

$$\dot{p}_1(t) = \frac{\partial \Upsilon}{\partial \hat{x}_1}p_1(t) - p_1(t)\frac{\partial \eta}{\partial \lambda} + p_1(t)\frac{\partial \eta}{\partial \hat{x}_1}p_1(t) - \frac{\partial \Upsilon}{\partial \lambda} \tag{21}$$

where

$\hat{x}_1(t)$ is the estimate of $x_1(t)$

$$\Upsilon(\hat{x}_1) = \Upsilon(\hat{x}_1, \lambda)|_{\lambda=0} \tag{22}$$

$$\eta(\hat{x}_1) = \eta(x_1, \lambda)|_{\lambda=0} \tag{23}$$

$$\frac{\partial \Upsilon}{\partial \hat{x}_1} = \left.\frac{\partial \Upsilon(x_1, \lambda)}{\partial x_1}\right|_{\lambda=0} \tag{24}$$

$$\frac{\partial \eta}{\partial \lambda} = \left.\frac{\partial \eta(x_1, \lambda)}{\partial \lambda}\right|_{\lambda=0} \tag{25}$$

$$\frac{\partial \eta}{\partial \hat{x}_1} = \left.\frac{\partial \eta(x_1, \lambda)}{\partial x_1}\right|_{\lambda=0} \tag{26}$$

$$\frac{\partial \Upsilon}{\partial \lambda} = \left.\frac{\partial \Upsilon(x_1, \lambda)}{\partial \lambda}\right|_{\lambda=0}. \tag{27}$$

Substituting (22)–(27) into (20) and (21), one obtains

$$\dot{\hat{x}}_1(t) = p_1(t)\left[r_2(t) - \hat{x}_1(t)\right] \tag{28}$$

which is the state estimation equation, and

$$\dot{p}_1(t) = -p_1^2(t) + \hat{x}_1^2(t)/\mu_{a_1} + r_1^2(t)/\mu_{a_2} + 1/\mu_{w_1}. \tag{29}$$

Equation (29) is similar to, but not the same as, the error variance equation.

The initial conditions are assumed to be known. A simple rule for choosing the values of the initial condition is explained in the results section. This rule seemed to be working well for a wide range of signals and under different SNR conditions.

D. Estimation of the Signal

The equations used to estimate the interference, $\hat{x}_1(t)$, in the primary channel were derived and they are given by (28) and (29). Their form is surprisingly simple, in spite of the fact that they are nonlinear. These equations represent the adaptive filter block in Fig. 1. The signal estimate is given by

$$\hat{s}(t) = r_2(t) - \hat{x}_1(t). \tag{30}$$

In some situations, $\hat{s}(t)$ is noisy and postfiltering might be needed. This can be achieved by a different Pontryagin filter that has similar structure to that derived previously. The filter is derived in [12].

III. Properties of Pontryagin Filter

Properties of the Pontryagin filter (or adaptive noise canceller) are derived in the next section.

A. Stability

One consideration of both practical and theoretical interest is the stability of the Pontryagin filter. Stability refers to the behavior of the states estimates when measurements and external inputs are suppressed [14] i.e., $r_1(t) = 0 = r_2(t)$. The unforced filter equations take the form

$$\dot{\hat{x}}_1(t) = -p_1(t)\hat{x}_1(t) \tag{31}$$

and

$$\dot{p}_1(t) = -p_1^2(t) + \hat{x}_1^2(t)/\mu_{a_1} + /\mu_{w_1}. \tag{32}$$

Assuming steady state,

$$\dot{p}_1 = 0 = -p_\infty^2 + \hat{x}^2(t)/\mu_{a_1} + 1/\mu_{w_1} \tag{33}$$

where p_∞ is the steady-state value of $p_1(t)$.

In all of the simulation results presented, $\mu_{w_1} \approx 10^{-4}$ while $\mu_{a_1} \approx 1$ (see the results section). Also the system state $x_1(t)$ and its estimate $\hat{x}_1(t)$ of order ± 3.0. Hence, (33) can be rearranged as follows:

$$p_\infty^2 = \hat{x}_1^2/\mu_{a_1} + 1/\mu_{w_1} \approx 1/\mu_{w_1} \tag{34}$$

or

$$p_\infty \approx 1/\sqrt{\mu_{w_1}} \tag{35}$$

where $\approx$ means approximately equal to. Only the positive value of p_∞ is used in (35), since the negative value will yield an unstable filter as will be shown below.

Substituting (35) into (31), one obtains

$$\dot{\hat{x}}_1(t) \approx -p_\infty \hat{x}_1(t). \tag{36}$$

An eigenvalue at $-p_\infty$ is present which is negative, which indicates that the filter equations are asymptotically stable.

B. Bias

The filter equations of $\hat{x}_1(t)$ were derived earlier and will be repeated here for convenience

$$\dot{\hat{x}}_1(t) = p_1(t)[r_2(t) - \hat{x}_1(t)] \tag{37}$$

and

$$\dot{p}_1(t) = -p_1^2(t) + \hat{x}_1^2(t)/\mu_{a_1} + r_1^2(t)/\mu_{a_2} + 1/\mu_{w_1}. \tag{38}$$

In studying the bias, it is assumed that a steady state has been reached. This approach reduces the complexity and allows us to isolate other important features of the problem. It would be possible, in theory, to find an exact expression for the bias that is, in general, complicated. The resulting expression would be cumbersome, with bias effects tending to mask the problems of interest without yielding further insight.

Using the steady-state assumption, an expression for $p_1(t)$ is obtained. A differential equation for the error can then be derived. Another differential equation for the expected value of the error (bias) can be obtained. The forcing input of this equation is approximately zero and this will reduce the bias problem to that of the stability of the expected value of the error. The eigenvalues are shown to be negative, indicating asymptotic stability (i.e., the convergence of the expected value of the error to zero is achieved.

The system dynamics are repeated here as follows:

$$\dot{x}_1(t) = a_1(t)x(t) + a_2(t)r_1(t) + w_1(t). \tag{39}$$

The error $e(t)$ is defined as

$$e(t) = \hat{x}_1(t) - x_1(t). \tag{40}$$

Taking the derivative with respect to time, one obtains

$$\dot{e}(t) = \dot{\hat{x}}_1(t) - \dot{x}_1(t) \tag{41}$$

$$= p_1(t)[r_2(t) - \hat{x}_1(t)] - a_1(t)x_1(t) - a_2(t)r_1(t) - w_1(t). \tag{42}$$

Substituting for the expression of $r_2(t)$ rearranging, one obtains

$$\dot{e}(t) = e(t)[-p_1(t) + a_1(t)] - a_1(t)\hat{x}_1(t) - a_2(t)r_1(t) - w_1(t) + p_1(t)[s(t) + v_2(t)] \tag{43}$$

and

$$\dot{\hat{x}}_1(t) = -p_1(t)e(t) + p_1(t)[s(t) + v_2(t)]. \tag{44}$$

In steady state we substitute p_∞ for $p_1(t)$. We also assume that $E\{\dot{e}\} \approx d/dt E\{e(t)\}$ where $E\{\cdot\}$ is the expected value of "$\cdot$". Substitute in (43) and taking the expectation (notice that $a_1(t)$, and $a_2(t)$ are deterministic quantities) it follows that

$$d/dt E\{e(t)\} \approx E\{e(t)\}[-p_\infty + a_1(t)] - a_1(t)E\{\hat{x}_1(t)\}. \tag{45}$$

In (45), it was assumed that $r_1(t)$ and $v_2(t)$ have zero mean.

To study the stability of this equation, the equilibrium point is first obtained by setting the left-hand side of (45) to zero, and then by substracting the equilibrium point from the expected value of the error [15], [16]; viz:

$$0 = E\{e_{eq}(t)\}[-p_\infty + a_1(t)] - a_1(t)E\{\hat{x}_1(t)\}$$

which yields

$$E\{e_{eq}(t)\} = a_1(t)E\{\hat{x}_1(t)\}/[-p_\infty + a_1(t)]. \tag{46}$$

$E\{e_{eq}(t)$ is the equilibrium value of the $E\{e(t)\}$. Assume that $p_\infty \gg a_1(t)$ so that,

$$E\{e_{eq}(t)\} \approx -a_1(t)E\{\hat{x}_1(t)\}/p_\infty \approx \text{constant} = 0. \tag{47}$$

Thus

$$d/dt[E\{e(t)\} - E\{e_{eq}(t)\}] \approx d/dt E\{e(t)\}. \tag{48}$$

Substituting (46)–(48) in (45) one obtains

$$\begin{aligned} d/dt&[E\{e(t)\} - E\{e_{eq}(t)\}] \\ &\approx d/dt E\{e(t)\} \\ &\approx [E\{e(t)\} - E\{e_{eq}(t)\}][-p_\infty + a_1(t)] \\ &\quad - a_1(t)E\{\hat{x}(t)\} \\ &\approx E\{e(t)\}[-p_\infty + a_1(t)] \\ &\quad + a_1(t)E\{\hat{x}(t)\}[-p_\infty + a_1(t)]/ \\ &\quad [-p_\infty + a_1(t)] - a_1(t)E\{\hat{x}_1(t)\} && (49) \\ &\approx E\{e(t)\}[-p_\infty + a_1(t)] && (50) \\ &\approx -p_\infty E\{e(t)\} \end{aligned}$$

i.e.,

$$d/dt E\{e(t)\} \approx -p_\infty E\{e(t)\}. \tag{51}$$

Since p_∞ is positive, the eigenvalue of (51) is negative and consequently the $E\{e(t)\}$ (or the bias) is asymptotically stable around the equilibrium point (which is approximately zero), i.e., the bias converges to zero. This is the result of interest.

Notice that the establishment of zero bias for the filter estimate is approximate and is dependent on the relative values between $\mu_{a_1}, \mu_{a_2}, \mu_{w_1}$ and the range of variation of the signal. If the signal is known to have a large range of variation, one can always scale down its values to ensure that the estimate is unbiased.

C. Variance and Consistency

Since the estimate, $\hat{x}_1(t)$, is asymptotically unbiased, we shall assume that it is unbiased over a finite processing time in order to reduce the complexity of the analysis. Thus, we define the variance, $\xi(t)$, of the error $e(t)$ as follows:

$$\xi(t) = E\{e^2(t)\}. \tag{52}$$

Properties of the error variance are determined using an approach similar to that used to evaluate the bias. A differential equation is derived for the error variance. With simple approximations, it shall be established that the driving force term is zero and the equilibrium point is also zero. Hence, with analysis of the stability of the error variance (i.e., its eigenvalue), the convergence of the error variance to zero will be demonstrated.

We now derive a differential equation for the error variance [15]:

$$\dot{\xi}(t) = d/dt E\{e^2(t)\} \approx 2E\{e(t)\dot{e}(t)\} \tag{53}$$

substitute the expression for $\dot{e}(t)$ (i.e., (43) into (53)):

$$\begin{aligned}\dot{\xi}(t) &= 2E\{e(t)[e(t)(-p_1(t)+a_1(t)) \\ &\quad - a_1(t)\hat{x}_1(t) - a_2(t)r_1(t) \\ &\quad - w_1(t) + p_1(t)(s(t)+v_2(t))]\} \\ &= 2E\{e^2(t)[-p_1(t)+a_1(t)]\} \\ &\quad -2a_1(t)E\{e(t)\hat{x}_1(t)\} - 2a_2(t)E\{e(t)r_1(t)\} \\ &\quad -2E\{e(t)w_1(t)\} \\ &\quad +2E\{e(t)p_1(t)[s(t)+v_2(t)]\} \\ &= 2\xi(t)[-p_1(t)+a_1(t)] - 2a_1(t)E\{e(t)\hat{x}_1(t)\} \\ &\quad -2a_2(t)E\{e(t)r_1(t)\} \\ &\quad -2E\{e(t)w_1(t)\} \\ &\quad +2E\{e(t)p_1(t)[s(t)+v_2(t)]\}.\end{aligned} \tag{54}$$

Equation (54) is an exact expression for the error variance (assuming that the error is unbiased and Gaussian). One should notice that the second, third, fourth, and fifth terms on the right-hand side of (54) are input to the equation. It can be demonstrated that these terms are bounded. Also, notice that at steady state $-p_1(t)+a(t) \approx -p_\infty$. Hence, the output of the equation ($\xi(t)$) is bounded i.e., $\xi(t)$ converges.

To study the approximate stability (convergence) of the error variance, the equilibrium point is first derived. We then shift the variance to the equilibrium point such that the new equilibrium is zero. Finally, by studying the eigenvalue of the resulting system, one might make the conclusions about the stability of the error variance.

To find the equilibrium point ξ_{eq}, we set $\dot{\xi}(t)$ to zero; viz:

$$\begin{aligned}\dot{\xi}(t) = 0 = 2\xi_{eq}(t)[-p_1(t)+a_1(t)] \\ -2a_1(t)E\{e(t)\hat{x}_1(t)\} \\ -2a_2(t)E\{e(t)r_1(t)\} \\ -2E\{e(t)w_1(t)\} + 2E\{e(t)p_1(t)[s(t)+v_2(t)]\}\end{aligned}$$

rearranging one gets

$$\begin{aligned}\xi_{eq}(t) &= [a_1(t)E\{e(t)\hat{x}_1(t)\} + a_2(t)E\{e(t)r_1(t)\} \\ &\quad + E\{e(t)w_1(t)\} \\ &\quad - E\{e(t)p_1(t)(s(t)+v_2(t))\}]/ \\ &\qquad [-p_1(t)+a_1(t)].\end{aligned} \tag{55}$$

If one assumes the steady state has been reached, it follows that

$$\begin{aligned}\xi_{eq}(t) &\approx -[a_1(t)E\{e(t)\hat{x}_1(t)\} + a_2(t)E\{e(t)r_1(t)\} \\ &\quad + E\{e(t)w_1(t)\} \\ &\quad - E\{e(t)p_\infty(s(t)+v_2(t))\}]/p_\infty.\end{aligned} \tag{56}$$

To reduce the complexity, let $e(t)$, $s(t)$, and $v_2(t)$ be uncorrelated. Thus $E\{e(t)p_\infty[s(t)+v_2(t)]\} \approx 0$. For large values of p_∞, one might assume that the denominator is much larger than the numerator, thus

$$\xi_{eq}(t) \approx \text{constant} = 0 \tag{57}$$

shifting (54) around the equilibrium point one gets

$$\begin{aligned}&d/dt(\xi(t)+\xi_{eq}) \\ &\quad = 2(\xi(t)+\xi_{eq})(-p_\infty) - 2a_1(t)E\{e(t)\hat{x}_1(t)\} \\ &\quad\quad -2a_2(t)E\{e(t)r_1(t)\} - 2E\{e(t)w_1(t)\} \\ &\quad\quad +2E\{e(t)p_\infty[s(t)+v_2(t)]\}.\end{aligned} \tag{58}$$

Using (57) for the left-hand side and (55) for the right-hand side of (58) one can show that

$$\dot{\xi}(t) \approx -2p_\infty \xi(t). \tag{59}$$

Since p_∞ is positive, then the eigenvalue of (59) is negative and the error variance, $\xi(t)$, is asymptotically stable around the zero equilibrium point. Thus, the error variance converges to zero and the filter estimate, $\hat{x}_1(t)$, is as a consequence, both unbiased and consistent, i.e., the estimate converges to the true value $x_1(t)$. It should be realized that the previous analysis is approximate, and as a matter of fact, the error variance, ξ, is dependent on the SNR, the shape of $a_1(t)$, $a_2(t)$, $h(t)$, and many other parameters.

The previous analysis, however, reveals some important points:

1) The bias and the variance are relatively insensitive to the SNR. This has been observed before during some studies in special cases of adaptive noise cancelling [12].

2) The tuning parameters, μ_{a_1}, μ_{a_2}, and μ_{w_1} are required to be small enough to make the value of p_∞ sufficiently large to reduce the effect of the time variation of the functions $h(t)$, $a_1(t)$, $a_2(t)$ and the different expectation operations on the evaluation of the bias and the error variance.

In summary, we have proven, using simple approximations, that the filter equations are stable and the estimate is asymptotically unbiased and converges to the true value. Thus Pontryagin ANC can accurately estimate the interference in the primary. A subsequent subtraction operation can yield the desired signal estimate, $\hat{s}(t)$.

IV. Simulation and Results

One case was simulated for which the interference in the primary is $x_1(t) = 0.5\ [n_1(t)+0.2n_1(t-\Delta)]$ where Δ is

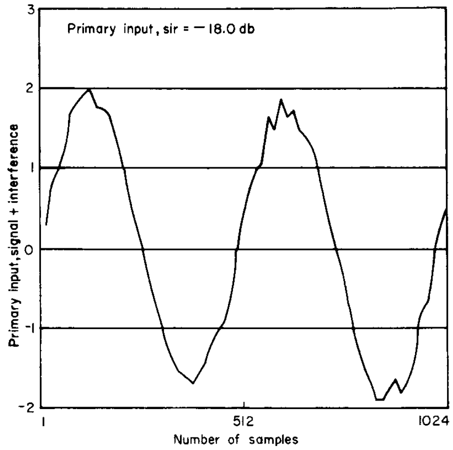

Fig. 2. Primary input $r_2(t)$. Average of 10 Monte Carlo simulations. SIR = −18 dB; SNR of the channel = 43 dB.

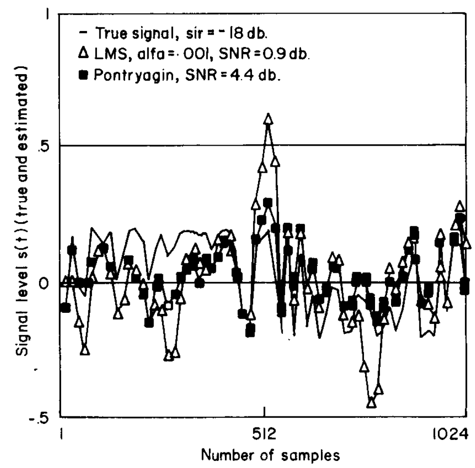

Fig. 3. True signal $s(t)$ and its estimated values Pontryagin-ANC and LMS-ANC. The gain constant (alfa) of LMS = 0.001. The estimate SNR = 4.4 dB for Pontryagin and is 0.9 dB for LMS. Average of 10 Monte Carlo simulations.

the sampling interval. The LMS algorithm (with 30 weights) was used and its results are compared to that of Pontryagin ANC. The number of weights and the gain constant, α, for the LMS algorithm were chosen such that the accuracy and speed of convergence are acceptable, and the computation time is not excessive. One might argue that by increasing the number of weights and a better choice of α [11], one might obtain better performance than that obtained using the Pontryagin ANC. It is pointed out that a more refined choice for the filter tuning parameters μ_{a_1}, μ_{a_2}, and μ_{w_1} will lead to improved performance under any conditions.

The performance is measured by evaluating the SNR value between the true and estimated values of the signal. The simulated signal is $s(t) = 0.2\sin[2\pi f_1 t + \varphi(t)]$ and the interference is $n_1(t) = 3\sin 2\Pi f_2 t$, where $f_1 = 1$ Hz and $f_2 = 2$ Hz, and $\varphi(t)$ is a time varying random phase. $\varphi(t)$ is uniformly distributed between $-\Pi/2$ and $+\Pi/2$. v_1 and v_2 are two independent zero mean Gaussian white noise processes. v_1 is zero mean and zero variance, while v_2 is zero mean and 0.001 variance. These values yielded an SNR of +43 dB at the primary input. The weights and initial conditions of Pontryagin ANC were chosen according to the following rule:

$\mu_{a_1} = \mu_{a_2} = 1$, $\mu_{w_1} = 20\Delta^2$,
Δ = sampling interval = 1 ms;
$\hat{x}$, (0) = 2,
number of samples = 1024;
$p_1(0) = 0.1/\sqrt{\Delta}$;
for LMS, (α = gain constant) was chosen to be 0.001.

The results represented are the average of 10 Monte Carlo simulations. The primary signal, $r_2(t)$, is shown in Fig. 2. $r_2(t)$ is composed of three additive components, i.e.,

$$r_2(t) = x_1(t) + s(t) + v_2(t), \text{ with SNR} = 43 \text{ dB}$$

where

$$x_1(t) = 0.5[n_1(t) + 0.2n_1(t-\Delta)]$$
= interference at the primary input.

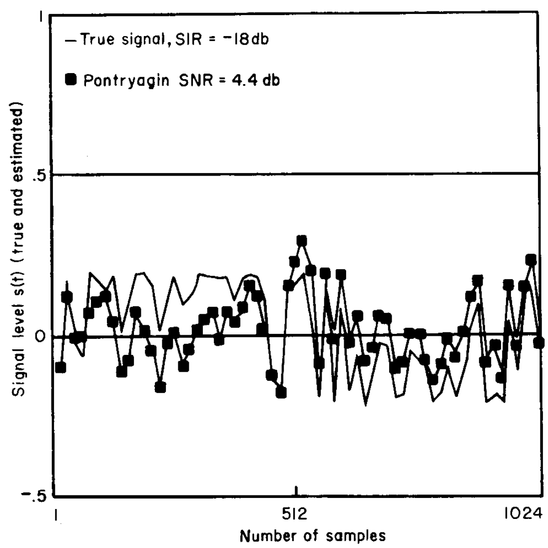

Fig. 4. True signal $s(t)$ and its estimated value using Pontryagin-ANC. Notice convergence starts to occur around the 400 sampling point. The simular The estimated SNR = 4.4 dB. Average of 10 Monte Carlo simulations.

The signal-to-interference ratio (SIR) = −18.0 dB. In Fig. 3, the signal, $s(t)$, and its estimates using LMS and Pontryagin ANC are shown. Notice that convergence starts to occur at about the 400th sampling point. While the Pontryagin ANC appears to approach convergence at this point, the LMS-ANC is still suffering from sudden departures from the true signal, as indicated near the 500 sampling point and the 800 sampling point. Consequently, the SNR of Pontryagin ≈ 4.4 dB, while for LMS was 0.9 dB. These values are calculated after the 500th sampling point. Fig. 4 depicts the true signal, $s(t)$, and its estimate

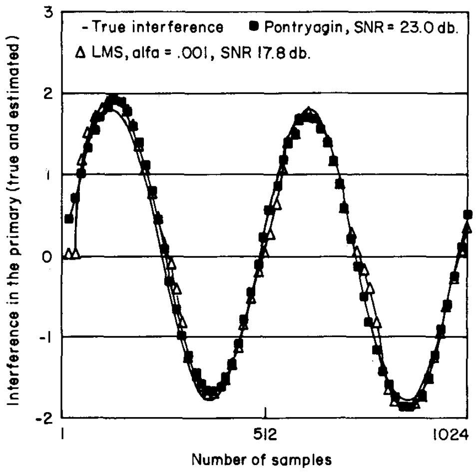

Fig. 5. True interference in the primary $x_1(t)$ and its estimated values using Pontryagin-ANC and LMS-ANC. The gain constant (alfa) of LMS = 0.001. The estimate SNR = 23.0 dB for Pontryagin and is 17.8 dB for LMS. Average of 10 Monte Carlo simulations.

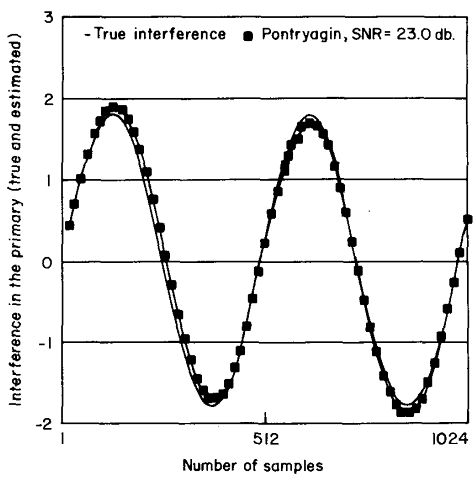

Fig. 6. True interference at the primary $x_1(t)$ and its estimated value using Pontryagin-ANC. The estimate SNR 23.0. dB. Average of 10 Monte Carlo simualtions.

using Pontryagin-ANC. In Fig. 5, the true interference at the primary input, $x_1(t)$, and its estimated values using Pontryagin-ANC and LMS-ANC are presented. Notice that convergence starts to occur around the 400th sampling point. Notice also that LMS-ANC has relatively large departures from the true value around the 500th and 800th sampling points. This results in sudden departures from the true signal (as noted in Fig. 3). The SNR values were calculated after the 500th sampling point. These values were 23.0 dB for Pontryagin-ANC and 17.8 dB for LMS-ANC. In Fig. 6, the true interference and its estimated values using Pontryagin-ANC are presented.

It is known that the LMS algorithm performance degrades rapidly if the nonstationarity and/or a periodicity in the signal is manifest [2]. The results of this study is in agreement. As shown in Fig. 6, the Pontryagin estimate of the interference is unbiased and it converges to the true value (SNR = 23 dB), which agrees with the theoretical studies of the Pontryagin-ANC. The superior performance is very likely due to the fact that Pontryagin-ANC utilized a nonstationary model for the process involved and a nonlinear filter that is suitable for this situation.

V. Discussion and Conclusions

This paper introduces a new approach for adaptive noise cancelling. The results indicated an improvement in performance and speed over the LMS algorithm. The nonlinear filter used in the approach that is based on the Pontryagin minimum principle and the method of invariant imbedding. The filter assumes that the interference at the primary input is the output of a linear time varying filter driven by the reference input, and white noise with two unknown time functions that describe the filter. The method eliminates the need to calculate many constant parameters as opposed to conventional AR or ARMA models. Estimation of the time varying functions was achieved by modelling them as controllers which drive the system to follow a noisy observation. The Pontryagin minimum principle was employed which resulted in a two point boundary value problem. The method of invariant imbedding was then used to transform the problem to initial value problem that can be solved in real time.

The performance of the filter was compared to that of the conventional LMS algorithm using Monte Carlo simulation under the same conditions. The results are in favor of the new scheme.

References

[1] G. S. Furno, and W. J. Tompkins, "A learning filter for removing noise interference," *IEEE Trans. Biomed. Eng.*, vol. BME-30, pp. 234–235, Apr. 1983.

[2] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englweood Cliffs, NJ: Prentice-Hall, 1985.

[3] B. Widrow, K. M. Duvall, R. P. Gooch, and W. C. Newman, "Signal cancellation phenomena in adaptive antennas: Causes and cures," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 469–478, May 1978.

[4] B. Widrow, *et al.*, "Adaptive noise cancelling: principles and applications," *Proc. IEEE*, vol. 63, pp. 1692–1716, Dec. 1975.

[5] B. Friedlander, "System identification for adaptive noise cancelling," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-30, pp. 699–708, Oct. 1982.

[6] P. Heinonen, T. Sarmarki, J. Malmivuo, and Y. Neuvo, "Periodic interference rejection using coherent sampling and waveform estimation," *IEEE Trans. Circuits Syst.*, vol. CAS-31, pp. 438–446, May 1984.

[7] L. Ljung, and T. Soderstrom, *Theory and Practice of Recursive Identification*. Cambridge, MA; M.I.T. Press 1982.

[8] S. S. Narayan, A. M. Peterson, and J. J. Narasimha, "Transform domain LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 609–615, June 1983.

[9] L. S. Pontryagin, V. Boltyanskii, R. Gamkrelidze, and E. Mischenko, "The mathematical theory of optimal processes," New York: Interscience, 1962.

[10] R. Bellman, H. Kagiwada, R. Kakaba, and R. Sridhar, "Invariant imbedding and nonlinear filtering theory," *J. Astronaut. Sci.*, vol. 13, pp. 110–115, May–June 1966.

[11] W. B. Mikhael, F. Wu, G. Kang, and L. Fransen, "Optimum adaptive algorithms with applications to noise cancellation," *IEEE Trans. Circuits Syst.*, vol. CAS-31, pp. 312–315, Mar. 1984.

[12] A. Abutaleb, "Adaptive noise cancelling and nonlinear estimation theory," in *Proc. 22nd Ann. Allerton Conf. Univ. of Illinois at Urbana*, Chicago, IL Oct. 1984.

[13] D. G. Schultz, and J. L. Melsa, *State Functions and Linear Control Systems*. New York: McGraw-Hill, 1967.
[14] A. Gleb, ed., *Applied Optimal Estimation*. Cambridge, MA: M.I.T. Press, 1974.
[15] H. L. Van Trees, *Detection Estimation and Modulation Theory Part I*. New York: Wiley, 1968.
[16] M. Vidyasagar, *Nonlinear Systems Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1978.

Ahmed S. Abutaleb received the B.Sc. degree in electrical engineering from Cairo University, Egypt, in 1976 and the M.Sc. and the Ph.D. degrees from the Moore School of Engineering, University of Pennsylvania, PA, in 1980 and 1984, respectively. He spent 1978–1979, at the M.I.T., Cambridge, MA, studying signal processing techniques and power systems reliability.

Between 1984 and 1986, he was an assistant professor at the Medical University of South Carolina, Department of Biometry, where he has developed a new nonlinear adaptive filter that has been applied successfully to a wide range of problems in 1-D and 2-D signal processing. Since June 1986 he has been an assistant Professor, Department of Electrical Engineering, College of Engineering, Temple University in Pennsylvania. His current research interests are: Statistical signal processing, Radar tracking systems, nonlinear estimation theory, imge processing, econometrics, and neural networks.

Dr. Abutaleb is a member of Sigma Xi.