SHARF: An Algorithm for Adapting IIR Digital Filters

MICHAEL G. LARIMORE, MEMBER, IEEE, JOHN R. TREICHLER, MEMBER, IEEE, AND C. RICHARD JOHNSON, JR., MEMBER, IEEE

Abstract—The concept of adaptation in digital filtering has proven to be a powerful and versatile means of signal processing in applications where precise *a priori* filter design is impractical. Adaptive filters have traditionally been implemented with FIR structures, making their analysis fairly straightforward but leading to high computation cost in many cases of practical interest (e.g., sinusoid enhancement). This paper introduces a class of adaptive algorithms designed for use with IIR digital filters which offer a much reduced computational load for basically the same performance. These algorithms have their basis in the theory of hyperstability, a concept historically associated with the analysis of closed-loop nonlinear time-varying control systems. Exploiting this theory yields HARF, a hyperstable adaptive recursive filtering algorithm which has provable convergence properties. A simplified version of the algorithm, called SHARF, is then developed which retains provable convergence at low convergence rates and is well suited to real-time applications. In this paper both HARF and SHARF are described and some background into the meaning and utility of hyperstability is given. In addition, computer simulations are presented for two practical applications of IIR adaptive filters: noise and multipath cancellation.

I. Introduction

THE CONCEPT of adaptation in digital filtering has proven to be a powerful and versatile means of signal processing in applications where precise *a priori* filter design is impractical. Self-adjusting or adaptive filters have been successfully applied to a wide spectrum of problems, ranging from channel equalization to antenna beam-forming and noise cancelling [1]-[4]. For the most part, such signal processing applications have relied upon the well-known adaptive finite impulse response (FIR) filter configuration. Yet, in practice, situations commonly arise wherein the nonrecursive nature of this adaptive filter results in a heavy computational load. Consequently, in recent years, active research has attempted to extend the adaptive FIR filter algorithms to the more general feedback or infinite impulse response (IIR) configuration. The immediate reward lies in the substantial decrease in computation that a feedback filter can offer over a nonrecursive filter.

Several authors [5]-[9] have recently suggested various procedures for adjusting the parameters of an IIR filter based on a data-dependent gradient search. Primarily, these algorithms represent a generalization of the gradient-estimation procedures serving as the foundation for FIR adaptation. However, due to the feedback configuration, compounded by the nonlinear nature of the adaptation process, a rigorous analysis of convergence properties has not been established, and consequently a simple and broadly applicable algorithm has not emerged.

This paper presents an alternate approach for adapting the parameters of the IIR filter that is based on the theory of hyperstability [10], [11], a powerful concept that was developed for the stability analysis of time-varying nonlinear systems [12]. As will be shown, the adaptive IIR process can be viewed as a linear system having time-varying nonlinear feedback. The updating procedure can be chosen to assure that the resulting closed-loop configuration is hyperstable, and hence convergent. The resulting algorithm will be derived in two stages. The first technique, which has been designated the hyperstable adaptive recursive filter (HARF) [10], has been proven convergent for a broad class of circumstances. The second technique [11] represents a simplified version of the first, hence the acronym SHARF, and is hyperstably convergent only for slow rates of adaptation. However, the SHARF algorithm is more amenable to real-time implementation, as its required computation is on the order of that of the LMS algorithm [13] used in adapting FIR filters.

To develop the adaptive algorithm Section II of this paper reformulates the IIR filtering problem in a framework that allows use of results from output error identification [14], [15]; the HARF algorithm is a direct consequence. Since the concept of hyperstability has not previously been applied to adaptive signal processing, Section III discusses its requirements and implications in the context of adaptive filtering. Section IV presents the simplified algorithm (SHARF) and demonstrates the resulting performance. In comparison with another adaptive IIR filter algorithm [7], it is seen that convergence is not only guaranteed, but actually accelerated. The remaining two sections describe two typical applications where adaptive IIR filtering using SHARF proves to be of significant value: adaptive noise cancelling [16] and adaptive multipath cancelling [17].

II. Adaptive IIR Filtering

A. Gradient-Based Formulations

The basic notation and structure of the adaptive filtering problem is shown in Fig. 1. A discrete input sequence $x(k)$

Manuscript received July 3, 1979; revised February 4, 1980. The work of C. R. Johnson, Jr. was supported by the Engineering Foundation under Grant RC-A-77-1D and by the National Science Foundation under Grant ECS-79-08695.

M. G. Larimore is with the Department of Electrical Engineering and Computer Science, University of Colorado, Colorado Springs, CO 80907.

J. R. Treichler is with ARGO Systems, Inc., Sunnyvale, CA 94086.

C. R. Johnson, Jr. is with the Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.

0096-3518/80/0800-0424$00.75 © 1980 IEEE

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

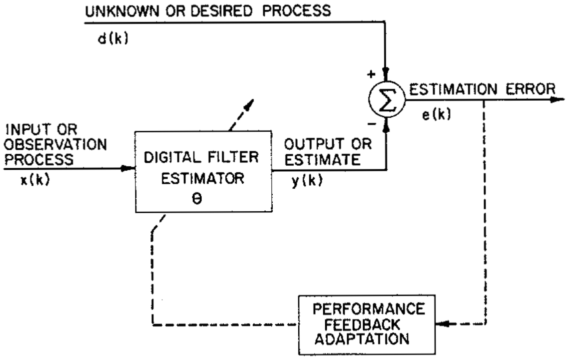


Fig. 1. Adaptive filter structure and notation.

is applied to a digital filter having a parameter vector $\boldsymbol{\theta}$, to produce output sequence $y(k)$. It is desired that the filter's parameter vector be designed in such a manner that the output closely approximates some external signal $d(k)$, and that the error between the two,

$$e(k) = d(k) - y(k), \tag{1}$$

be small. Typically, the squared-error cost function is called upon as a design performance measure due to its mathematical simplicity:

$$J \triangleq \frac{1}{2} \sum_{l=0}^{L} e^2(k-l). \tag{2}$$

To alleviate the necessity for careful *a priori* measurements of $x(k)$ and $d(k)$, the filter can be made self-adjusting as shown in Fig. 1. By allowing variations to the internal filter parameters in $\boldsymbol{\theta}$ determined by a performance feedback algorithm, the filter ideally converges to optimal design $\boldsymbol{\theta}^*$ minimizing J. For FIR filters, such convergent algorithms are readily available and are robust under a wide class of input environments.

One such family of adaptive procedures [13] is based on a gradient search of the performance surface. If the parameter vector at time k is denoted as $\boldsymbol{\theta}(k)$, then the updating algorithm is given by

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) - \mu(k)\, \nabla_{\boldsymbol{\theta}} J(\boldsymbol{\theta}(k)) \tag{3}$$

where $\mu(k)$ is a scalar sequence and $\nabla_{\boldsymbol{\theta}} J$ is the gradient of the cost function with respect to the parameter vector $\boldsymbol{\theta}$. For proper choice of sequence $\mu(k)$ convergence can be assured; however, only when the cost function is unimodal with respect to $\boldsymbol{\theta}$ will convergence be to the globally optimal design $\boldsymbol{\theta}^*$ for arbitrary initial choice of $\boldsymbol{\theta}(0)$. When the true gradient is unavailable, unbiased estimates will suffice in certain cases. For adaptive FIR filters, gradient-based algorithms owe their usefulness to 1) the ease of computation of the gradient and 2) the hyperparabolic nature of the squared-error cost function in the filter parameter space.

For adaptive recursive filters, these two features no longer apply. The output of an adaptive IIR filter can be expressed as an autoregressive moving average (ARMA) process driven by $x(k)$,

$$y(k) = \sum_{j=0}^{M} \hat{b}_j(k) x(k-j) + \sum_{i=1}^{N} \hat{a}_i(k) y(k-i). \tag{4}$$

Note that the N parameters $\hat{a}_i(k)$ are feedback coefficients, and $M+1$ parameters $\hat{b}_j(k)$ are feedforward coefficients. The time-varying nature of each adjustable parameter is implied by the argument k. The adapted parameter vector is then

$$\boldsymbol{\theta}(k) = [\hat{b}_0(k)\ \hat{b}_1(k) \cdots \hat{b}_M(k) \vdots \hat{a}_1(k)\ \hat{a}_2(k) \cdots \hat{a}_N(k)]^T. \tag{5}$$

The optimal IIR filter occurs for $\boldsymbol{\theta}(k) \equiv \boldsymbol{\theta}^*$, where

$$\nabla_{\boldsymbol{\theta}} J \big|_{\boldsymbol{\theta}^*} = \sum_{l=0}^{L} e(l)\, \nabla_{\boldsymbol{\theta}} e(l) \big|_{\boldsymbol{\theta}^*} = 0. \tag{6}$$

This necessary condition reduces to a set of scalar equations,

$$\sum_{l=0}^{L} e(l) \frac{\partial e(l)}{\partial \hat{a}_i}\bigg|_{\boldsymbol{\theta}^*} = -\sum_{l=0}^{L} e(l) \left\{ y(l-i) + \sum_{j=1}^{N} \hat{a}_j \frac{\partial y(l-j)}{\partial \hat{a}_i} \right\}\bigg|_{\boldsymbol{\theta}^*} = 0, \quad 1 \leqslant i \leqslant N \tag{7a}$$

$$\sum_{l=0}^{L} e(l) \frac{\partial e(l)}{\partial \hat{b}_j}\bigg|_{\boldsymbol{\theta}^*} = -\sum_{l=0}^{L} e(l) \left\{ x(l-j) + \sum_{i=1}^{N} \hat{a}_i \frac{\partial y(l-i)}{\partial \hat{b}_j} \right\}\bigg|_{\boldsymbol{\theta}^*} = 0, \quad 0 \leqslant j \leqslant M. \tag{7b}$$

Note that these equations dismiss interdependence of the filter parameter set; as such, (7a) and (7b) represent a greatly simplified version of the true gradient formulas that would result in the presence of parameter adaptation. The effect of parameter adjustment in a feedback filter manifests itself as additional filter dynamics, making unbiased gradient estimation a function of filter time constants.

Due to the feedback characteristics of the IIR configuration, this simplified set of gradient equations is nonlinear, and in general it can be satisfied by multiple extrema. For a specific numerical example, refer to [18]. Consequently, seeking the optimal design $\boldsymbol{\theta}^*$ by a gradient search of the performance function will not necessarily be successful. Convergence to local minima will invariably occur for certain initial values of parameters. Furthermore, the computation of the gradient, as indicated by (7b), is itself a recursive process, and can represent significant computation [19].

Summarizing, the general problem of designing IIR filters, adaptive or not, lacks the mathematical advantages that have made least-squares performance useful for estimation. Also, the general gradient search framework is less suitable to develop an adaptation strategy to seek an IIR filter design.

B. Output Error Identification Approach

Consider a less general, more structured problem. Let $d(k)$ itself be a bounded ARMA process driven by $x(k)$,

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

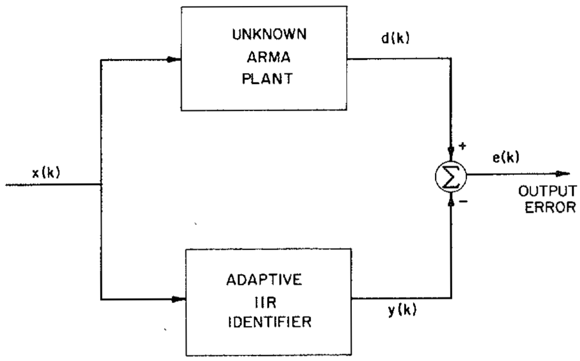


Fig. 2. Output error parameter identification.

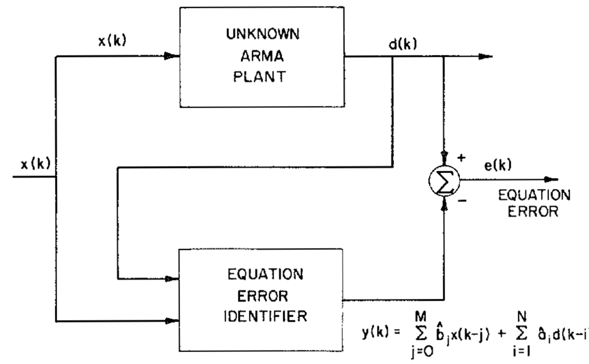


Fig. 3. Equation error parameter identification.

$$d(k) = \sum_{j=0}^{M_d} b_j x(k-j) + \sum_{i=1}^{N_d} a_i d(k-i) \tag{8}$$

where the generating parameters a_i and b_j are constant. Assume further that sufficient variables are provided in the adaptive filter to span the parameter space generating $d(k)$, i.e., $M \geqslant M_d$ and $N \geqslant N_d$. Without loss of generality the assumption that $M = M_d$ and $N = N_d$ is permitted, with any excess generating parameters equaling zero. The error process $e(k)$ becomes

$$e(k) = \sum_{j=0}^{M} \{b_j - \hat{b}_j(k)\}\, x(k-j) + \sum_{i=1}^{N} \{a_i d(k-i) - \hat{a}_i(k) y(k-i)\}. \tag{9}$$

It is sufficient to choose $\hat{b}_j \equiv b_j$ and $\hat{a}_i \equiv a_i$ to minimize J. If so, then

$$e(k) = \sum_{i=1}^{N} a_i e(k-i), \tag{10}$$

and due to the bounded-input bounded-output (BIBO) stability of (8),

$$\lim_{k \to \infty} e(k) = 0. \tag{11}$$

Once steady state has occurred, this choice of parameters, as would be expected, yields a minimal squared error.

With the filter design problem stated in this fashion, it is a restatement of the output error identification problem [14], [15] shown in Fig. 2. In this situation, an unknown ARMA plant has input $x(k)$ and output $d(k)$, perhaps measured in the presence of noise. It is desired to estimate or identify its internal parameters a_i and b_j in an unbiased fashion. This can be done by using a parallel model [20], driven by the same input, the output $y(k)$ of which is compared with the plant output $d(k)$. On the basis of this output error, the parameter estimates are formed. However, there is an important distinction: In the identification problem a performance measure based on the error is used only as a means of attaining small parameter error; in the filtering problem, a small output error is instead the desired end. In certain cases, a filter can tolerate substantial parameter error and still perform satisfactorily.

Many of the same computational problems that burden adaptive IIR filtering also plague output error identification, i.e., parameter evaluation by gradient search remains a nontrivial task. Consequently, approaches to identification other than output error, e.g., equation error [21], [22], shown in Fig. 3, have been more widely accepted at the cost of potential parameter bias. Rather than modeling with the feedback structure of (4), the actual plant output $d(k)$ is substituted for the model output on the right-hand side, giving rise to the series/parallel model [20]

$$y(k) = \sum_{j=0}^{M} \hat{b}_j(k) x(k-j) + \sum_{i=1}^{N} \hat{a}_i(k) d(k-i). \tag{12}$$

With sufficient degrees of freedom, simple gradient-based adaptation of the parameters assures consistent convergence, with one important exception. Whenever $d(k)$ is measured in the presence of observation noise, the parameter estimates will, in general, be biased [23].

C. The Hyperstable Adaptive Recursive Filter

Landau [14] has recently introduced an unbiased output error procedure for seeking parameter estimates of an ARMA plant, the *hyperstable output error identifier*. Presented here is an IIR adaptive filtering algorithm based on this hyperstable output error identifier. This technique, the hyperstable adaptive recursive filter (HARF) [10], embodies two modifications of Landau's algorithm making it more suitable for the signal processing context. First, Landau's technique calls for diminishing adaptation gain factors, ultimately converging to zero. When identifying a plant with constant parameters, such an algorithm is acceptable; however, it has been recognized that for most adaptive filtering problems, adaptation must remain active to track changes in the signal environment [13]. Secondly, the hyperstable identifier is not strictly causal, requiring the current output sample $d(k)$ to form the current parameter values $\hat{b}_j(k)$ and $\hat{a}_i(k)$. Again, while this condition is reasonable for parameter estimation, it is undesirable for real-time filtering.

The HARF algorithm represents the first technique proposed for adaptive IIR filtering which has provable convergence properties. Despite its moderate computational complexity, a careful study of its behavior allows simplifications to be made, preserving most of HARF's desirable properties while reducing the required computation.

The HARF implementation is shown in Fig. 4 for use in the following discussion. Notice from the figure that in addition

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

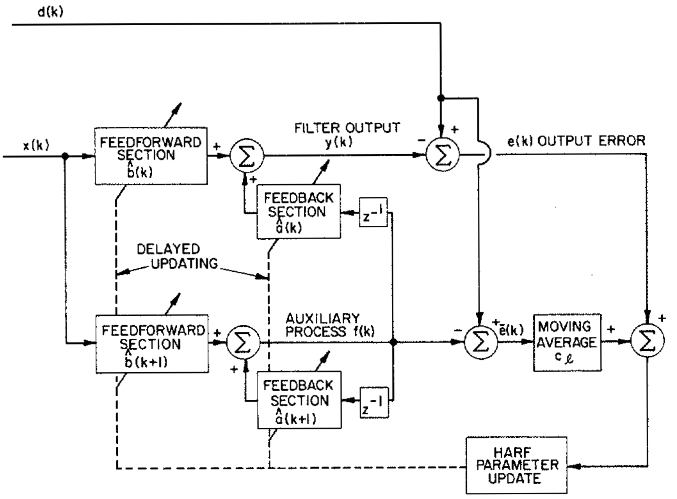


Fig. 4. Implementation of HARF.

to the principal adaptive filter that forms the output $y(k)$, there is an auxiliary process generated:

$$f(k) = \sum_{i=1}^{N} \hat{a}_i(k+1) f(k-i) + \sum_{j=0}^{M} \hat{b}_j(k+1) x(k-j). \quad (13)$$

This ARMA process is used both in forming the output

$$y(k) = \sum_{i=1}^{N} \hat{a}_i(k) f(k-i) + \sum_{j=0}^{M} \hat{b}_j(k) x(k-j), \quad (14)$$

and in the adaptive algorithm, to appear shortly. The parameters in these two equations are separated in time by one sample; that is, weighting coefficients used in (13) have undergone one additional update versus those in (14). If convergence should occur, then $\hat{a}_i(k+1) = \hat{a}_i(k)$ and $\hat{b}_j(k+1) = \hat{b}_j(k)$, and $y(k)$ asymptotically converges to $f(k)$. However, in the transient stages of adaptation, the distinction between $y(k)$ and $f(k)$ proves necessary.

Suppose at each sample the adaptive filter coefficients are updated according to the formulas

$$\hat{a}_i(k) = \hat{a}_i(k-1) + \frac{\mu_i}{q(k)} f(k-i-1) \left(\{d(k-1) - y(k-1)\} + \sum_{l=1}^{P} c_l \{d(k-l-1) - f(k-l-1)\} \right) \quad 1 \leq i \leq N, \quad (15a)$$

$$\hat{b}_j(k) = \hat{b}_j(k-1) + \frac{\rho_j}{q(k)} x(k-j-1) \left(\{d(k-1) - y(k-1)\} + \sum_{l=1}^{P} c_l \{d(k-l-1) - f(k-l-1)\} \right), \quad 0 \leq j \leq M, \quad (15b)$$

where $q(k)$ is a normalizing factor greater than unity,

$$q(k) = 1 + \sum_{l=1}^{N} \mu_l f^2(k-l-1) + \sum_{l=0}^{M} \rho_l x^2(k-l-1), \quad (16)$$

and μ_i and ρ_j are arbitrary positive constants. In addition, the P constants c_l are chosen by the designer so that the discrete transfer function

$$G(z) = \frac{1 + \sum_{l=1}^{P} c_l z^{-l}}{1 - \sum_{i=1}^{N} a_i z^{-i}} \quad (17)$$

is strictly positive real (SPR) [26]. The implications of this requirement are discussed in Section III.

Under these conditions, proof can be given [10] that the moving average quantity

$$v(k) = (d(k) - f(k)) + \sum_{l=1}^{P} c_l (d(k-l) - f(k-l)) \quad (18)$$

converges to zero and as a result

$$y(k) \to f(k) \to d(k), \quad (19)$$

which is the desired performance. For further details see [10].

Before proceeding to a discussion of the SPR assumption necessary for this hyperstable formulation, consider briefly a heuristic description of the HARF updating algorithm. In (15) it can be seen that, aside from assorted positive scaling factors, the update to each coefficient is basically a product of two components. First is the value of the signal corresponding to the given weight in the output equation (14). For example, the update to $\hat{a}_l(k)$ depends on $f(k-l)$, and their product $\hat{a}_l(k) f(k-l)$ appears in (14). The second factor [appearing in brackets in (15)] is dependent on the instantaneous performance of the filter, disguised by a moving average expression. Therefore, for a given quality of performance, largest adjustment is made to the coefficients contributing the most to the output via (14). These features are shared by a family of adaptation procedures. Readers familiar with gradient-based FIR adaptation procedures will recognize a similarity. However, as noted, the complexity of an IIR structure results in specific differences that cannot presently be accounted for by means of a gradient descent approach.

III. Hyperstability and Adaptive Filtering

The concept of system hyperstability, upon which this adaptive IIR filtering algorithm is based, was developed by V. M. Popov [12], and provides a generalized description of output stability in time-varying or nonlinear cases. Use of this analysis has occurred primarily in the control theory literature. For example, it has recently been applied to output error identification via the model reference adaptive system structure [14]. However, the signal processing community has not satisfactorily benefited by this analysis technique. In this section the hyperstability theorem is stated and a heuristic description of its conditions and implications is given. In addition, its relationship to adaptive filtering algorithms is shown.

Hyperstability is defined for the discrete-time case as follows [24], [25]. Let $G(z)$ be a rational scalar transfer function for a linear time-invariant system driven by $u(k)$ and responding

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

with $y(k)$. The system is said to be *hyperstable* if its state vector remains bounded over time for all driving sequences $u(k)$ satisfying jointly with the output

$$\sum_{l=0}^{K_0} u(l)\,y(l) < K^2 \qquad \forall K_0. \tag{20}$$

The present discussion deals with a stronger variation of this definition, rigorously denoted as *asymptotic hyperstability*, which requires that the state vector decay to zero for the same class of input sequences; references to hyperstability in this section will actually imply this latter definition. Note that if $G(z)$ has a proper rational form, the output $y(k)$ will likewise decay to zero.

The hyperstability theorem [12], [24] is a simple statement: The system described above is hyperstable if and only if its transfer function $G(z)$ is strictly positive real (SPR) [26], i.e.,

$$\mathrm{Re}\,[G(z)] > 0, \qquad z = e^{j\theta}. \tag{21}$$

That is, a SPR system will have a bounded output when driven by any input (including certain divergent sequences) satisfying (20).

The hyperstability theorem has an interesting interpretation in terms of system passivity [24], [25]. A familiar physical example arises in network theory, in a continuous-time context. It is well known that the driving-point impedance $Z(s)$ of a passive network is SPR, and relates driving current to response voltage at a network port. Consider a state realization where each state variable corresponds to an energy storage component. For any current such that the energy delivered into the network is bounded, i.e.,

$$E = \int_0^T v(t)\,i(t)\,dt < K^2 \qquad \forall T, \tag{22}$$

then the energy stored internally must be dissipated, i.e., $\|x(t)\| \to 0$. By analogy, any system that is SPR can be thought of as dissipative in a mathematical sense.

It is in the closed-loop configuration that system hyperstability becomes useful in adaptive filtering; parameter adaptation of digital filters can be restated in such a configuration [25]. Let $u(k)$ be a sequence derived as a nonlinear time-varying function of the output, denoted as a general feedback element $\mathfrak{F}$ in Fig. 5(a). Had the feedback been linear, $F(z)$ as indicated in Fig. 5(b), BIBO stability is easily checked in the frequency domain, by locating the zeros of

$$1 + F(z)\,G(z). \tag{23}$$

A zero not inside the unit circle implies instability. A physical interpretation, of course, requires that loop gain never be -1 at any frequency, i.e., give rise to 180° phase shift. A sufficient but unnecessary condition would be to restrict $F(z)$ and $G(z)$ each to contribute less than ±90° at all frequencies. That is, if both are SPR, the closed loop is guaranteed stable.

Still, such a condition is not useful when analyzing cases involving nonlinear time-varying feedback. Instead, a condition must be stated in terms of the time domain behavior of $\mathfrak{F}$. Note from Fig. 5(a) that the feedback element $\mathfrak{F}$ is driven by $v(k)$, and responds with $w(k) = -u(k)$. Then if

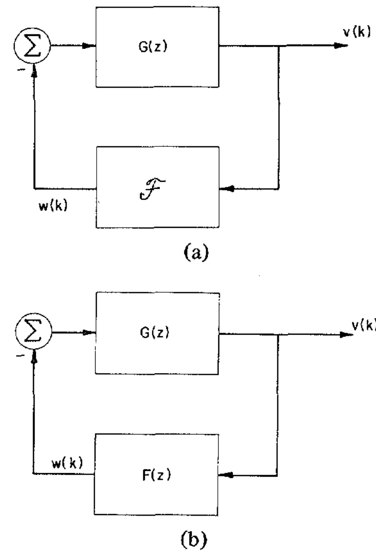


Fig. 5. Undriven closed-loop system. (a) General feedback element. (b) Linear feedback element.

$$\sum_{l=0}^{K} v(l)\,w(l) \geq -\gamma_0^2, \qquad \forall K > 0, \tag{24}$$

it follows that

$$\sum_{l=0}^{K} u(l)\,y(l) \leq \gamma_0^2, \qquad \forall K > 0. \tag{25}$$

Consequently, if $G(z)$ is SPR then the closed loop is hyperstable. This represents a means of generalizing the positive reality concept to the nonlinear time-varying feedback element $\mathfrak{F}$. For linear elements, it can be shown by using eigenfunction analysis that this condition indeed assures that phase response at all frequencies not exceed ±90°. From a more physical standpoint, if the energy delivered into the feedback element $\mathfrak{F}$

$$\sum_{l=0}^{K_0} v(l)\,w(l)$$

is bounded below as in (24), then $\mathfrak{F}$ is dissipative feedback. (This is analogous to the positive restriction on the physical energy delivered to a passive network.) Alternatively, (24) requires that the "sign" of the feedback element "on the average," implied by the summation, should be bounded below.

The hyperstability theorem guarantees stability for a class of intrinsically nonlinear time-varying systems. Clearly, the requirement given by (20) represents a sufficiency condition, and as such is overly restrictive. In the simple linear case cited above, both the strict positive reality of the forward path, and the positive reality of the feedback path are unnecessary for stability. Consequently, in the general case, one would expect the conditions to be unnecessary. While a less restrictive criterion for nonlinear systems is of interest, the formulation given here is satisfactory for use in analysis of a class of adaptive filtering algorithms.

To demonstrate the relationship of hyperstability to adaptive filtering, first define an auxiliary error quantity

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

$$\bar{e}(k) \triangleq d(k) - f(k)$$

where $d(k)$ and $f(k)$ are given by (8) and (13). Note that $\bar{e}(k)$ is closely related to $e(k)$ of (1); as noted, $f(k) \to y(k)$ as parameter convergence progresses, implying that $\bar{e}(k) \to e(k)$. For practical convergence rates, the error quantities are virtually identical; the distinction is necessary in the interest of rigor. An equation for the auxiliary error can be formed in the same manner as (9),

$$\bar{e}(k) = \sum_{j=0}^{M} [b_j - \hat{b}_j(k+1)]\, x(k-j) + \sum_{i=1}^{N} [a_i d(k-i) - \hat{a}_i(k+1) f(k-i)]. \tag{26}$$

By adding and subtracting the term

$$\sum_{i=1}^{N} a_i f(k-i) \tag{27}$$

and rearranging,

$$\begin{aligned}\bar{e}(k) &= \sum_{i=1}^{N} a_i \{d(k-i) - f(k-i)\} \\ &\quad + \sum_{i=1}^{N} \{a_i - \hat{a}_i(k+1)\} f(k-i) \\ &\quad + \sum_{j=0}^{M} \{b_j - \hat{b}_j(k+1)\} x(k-j) \\ &= \sum_{i=1}^{N} a_i \bar{e}(k-i) - w(k)\end{aligned} \tag{28}$$

where

$$w(k) = -\left\{ \sum_{i=1}^{N} [a_i - \hat{a}_i(k+1)]\, f(k-i) + \sum_{j=0}^{M} [b_j - \hat{b}_j(k+1)]\, x(k-j) \right\}. \tag{29}$$

Thus, the auxiliary error $\bar{e}(k)$ is an Nth order autoregressive process whose poles are identical to those of the unknown ARMA plant. The driving function $w(k)$ is a function of the parameter errors $(b_j - \hat{b}_j(k+1))$ and $(a_i - \hat{a}_i(k+1))$. This relation is shown diagrammatically in Fig. 6(a), where

$$A(z) = 1 - \sum_{i=1}^{N} a_i z^{-i}. \tag{30}$$

Note that the filter input $x(k)$ enters into the computation of $w(k)$ as a time-varying factor.

In the adaptive case, the parameter estimates $\hat{a}_i(k)$ and $\hat{b}_j(k)$ are updated using performance feedback. This effectively closes the loop, producing Fig. 6(b), where the update algorithm determines the functional form of $\mathcal{F}$. Note that the feedback is in general nonlinear/time-varying; therefore, choice of an algorithm that satisfies the conditions of the hyperstability theorem

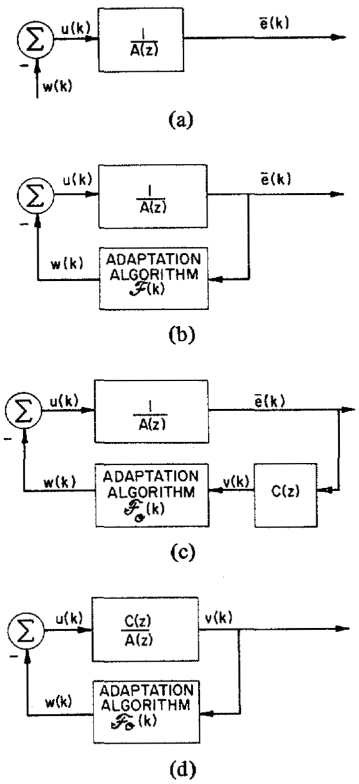


Fig. 6. Adaptive filter in hyperstability context. (a) Open-loop. (b) Adaptation feedback. (c) Linear preprocessor. (d) Hyperstable adaptation.

embodied in (24) is suitable to assure convergence of the auxiliary error.

To meet the hyperstability conditions, the forward element must be SPR. In general, the simple autoregressive form $1/A(z)$ will fail. A means of augmenting it to force SPR can be achieved by separation of $\mathcal{F}$ into a linear preprocessor $C(z)$ followed by a general element, as in Fig. 6(c). The output of the $C(z)$ is an auxiliary process $v(k)$, i.e., (18),

$$v(k) = \bar{e}(k) + \sum_{l=1}^{P} c_l \bar{e}(k-l). \tag{31}$$

In this case, $v(k)$ is simply a weighted average or smoothed version of the output error. Rearranging the system gives Fig. 6(d), with a forward composite linear element

$$G(z) = \frac{C(z)}{A(z)} = \frac{1 + \sum_{l=1}^{P} c_l z^{-l}}{1 - \sum_{i=1}^{N} a_i z^{-i}}. \tag{32}$$

In this form, the closed-loop output becomes $v(k)$, a moving-averaged version of the auxiliary error. This error enters into the function $\mathcal{F}_0$ for updating the adaptive filter weights $\hat{a}_i(k)$ and $\hat{b}_j(k)$, and generating the driving sequence $w(k)$. If the values of c_l are chosen to assure the SPR of $G(z)$, and the algorithm using $v(k)$ satisfies the relation (24), the closed-loop system is hyperstable and $v(k) \to 0$. The error quantity $\bar{e}(k)$, expressible as an internal state variable of $G(z)$, must likewise converge to zero. The HARF update algorithm described in

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

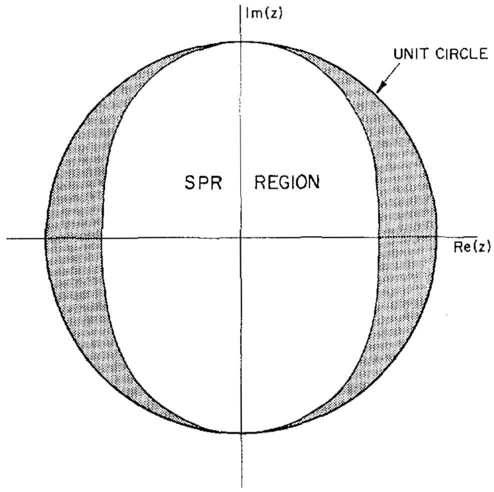


Fig. 7. Second-order system, region of SPR.

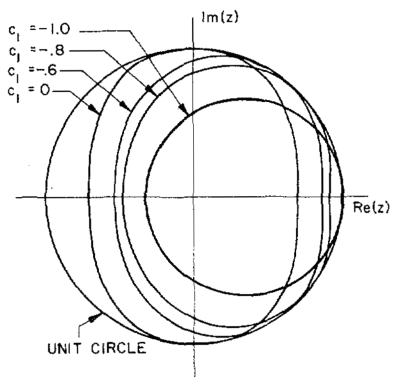


Fig. 8. Second-order system, region of SPR with error smoothing.

the previous section is shown in [10] to satisfy the hyperstability conditions.

The coefficients c_l that form a moving average of the output error represent a set of P design parameters, chosen to assure the SPR of $G(z)$ in (32). The denominator, determined by the unknown ARMA process $d(k)$, contributes a phase that must be tempered by zero placement, i.e., choice of c_l, to bound the net phase by $\pm 90°$. As an example, consider Fig. 7. Assume that $d(k)$ is produced by a second-order filter having complex poles. If the transfer function

$$G(z) = \frac{1}{1 - a_1 z^{-1} - a_2 z^{-2}} \tag{33}$$

were analyzed, for only certain conjugate pole-pairs would it be SPR; the SPR region is shown in relation to the unit circle as an unshaded oval. Thus, by effectively eliminating the flexibility of a numerator for $G(z)$, only certain pole-pairs allow the use of the hyperstability formulation. It is interesting to note that this excludes the vicinity near $z = 1$, the region where poles of an oversampled continuous process will cluster.

Once the numerator of (32) is introduced, the region of SPR pole location can be purposely deformed to encompass parts of the unit circle within which the unknown poles are likely to be found. Fig. 8 demonstrates this effect for several values of a single smoothing coefficient c_1. Note that $c_1 = 0$ produces the oval region as before; as c_1 becomes negative, and $G(z)$ gives a zero on the positive real axis, the SPR region is deformed toward $z = 1$. Finally, for $c_1 = -1$ the region becomes circular, tangential with the unit circle at $z = 1$, and encompasses the low-frequency pole locations. The introduction of the c_l parameters not only tailors the region of SPR, but also influences convergence behavior [11], as shown in a later section.

The most serious practical consideration in choosing the c_l coefficients to guarantee SPR of (32) is that the denominator is unknown. Given some *a priori* knowledge about the dynamics of the process $d(k)$, a reasonable choice of a numerator for $G(z)$ is simply an estimate of the denominator. Clearly, a perfect estimate produces a total cancellation of numerator and denominator, giving a degenerate SPR result, and a deterministic algorithm equivalent to an equation error variant, e.g., LMS. Inaccurate estimates, although not cancelling the dynamics of $G(z)$, may serve to contain the phase angle. For example, a biased estimate for the denominator parameters derived via the equation error technique has been found to give strict positive reality of (32) for certain SNR levels [27]. As a rule, placement of a zero in the vicinity of each pole provides a reasonable set of coefficients c_l. Currently, investigation is proceeding [28]-[30] on a joint self-adaptive algorithm for adjusting the c_l coefficients, parallelling recent suggestions in output error identification [31].

IV. The Simple Hyperstable Recursive Filter

While the hyperstability formulation of the adaptive IIR filtering problem provides a useful perspective, the resulting HARF algorithm suffers from two significant sources of computational complexity. First, examination of (15) indicates that an auxiliary ARMA process $f(k)$ is necessary to compute not only the filter output, but the weight updates as well. Secondly, the HARF algorithm includes a normalizing scale factor $q(k)$, computed for each iteration. Both of these components of the algorithm substantially increase algorithm cost in terms of hardware and/or sampling rate.

To make the adaptive IIR filtering algorithm more amenable to real-time processing, certain reasonable simplifications can be made. By specifying the rate constants μ and ρ to be sufficiently small as in successful gradient approximation procedures [19], the weights change very little from iteration to iteration; therefore,

$$\hat{a}_i(k+1) \cong \hat{a}_i(k)$$
$$\hat{b}_j(k+1) \cong \hat{b}_j(k). \tag{34}$$

A comparison of (13) and (14) indicates that

$$f(k) \cong y(k). \tag{35}$$

The output equation (14) becomes

$$y(k) = \sum_{i=1}^{N} \hat{a}_i(k)\, y(k-i) + \sum_{j=0}^{M} \hat{b}_j(k)\, x(k-j) \tag{36}$$

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

and the moving average process $v(k)$ in (18) becomes

$$v(k) \cong \{d(k) - y(k)\} + \sum_{l=1}^{P} c_l \{d(k-l) - y(k-l)\}$$

$$= e(k) + \sum_{l=1}^{P} c_l e(k-l), \qquad (37)$$

a simple moving average of the output error. Finally, note that $q(k)$ in (16) is a simple normalizing factor that controls the instantaneous adaptation rate, reducing the effective step size for large values of filter input and output. Once again, assuming that μ and ρ are sufficiently small,

$$q(k) \cong 1.$$

Using these approximations, (15) becomes

$$\hat{a}_i(k) \cong \hat{a}_i(k-1) + \mu_i y(k-1-i)\, v(k-1), \quad 1 \leqslant i \leqslant N \qquad (38a)$$

$$\hat{b}_j(k) \cong \hat{b}_j(k-1) + \rho_j x(k-1-j)\, v(k-1), \quad 0 \leqslant j \leqslant M. \qquad (38b)$$

The set (36)–(38) has been denoted the simple hyperstable adaptive recursive filter, SHARF.

Note that significant reduction in computation and storage has been realized; the update to each weight requires only the knowledge of the smoothed output error process $v(k)$. This computational savings was accomplished at the cost of no longer rigorously satisfying the hyperstability condition (24), so that convergence is no longer guaranteed for arbitrary positive μ and ρ. For practical purposes, however, slow adaptation maintains close approximation to a hyperstable structure.

It is interesting to note that certain earlier attempts at adaptive IIR filtering, notably the recursive LMS algorithm [7], are clearly special cases of (38). Its update equations, extrapolating from similar equations used in adaptive FIR filtering, are

$$\hat{a}_i(k+1) = \hat{a}_i(k) + \mu e(k)\, y(k-i) \qquad (39a)$$

$$\hat{b}_j(k+1) = \hat{b}_j(k) + \rho e(k)\, x(k-j) \qquad (39b)$$

where

$$e(k) = d(k) - y(k). \qquad (40)$$

Note that this is equivalent to the constraint that $c_l = 0$ for $l = 1$ to P in (37), i.e.,

$$v(k) = e(k). \qquad (41)$$

According to the hyperstability analysis, convergence is assured only if μ and ρ are small and the autoregressive function

$$G(z) = \frac{1}{1 - \sum_{i=1}^{N} a_i z^{-i}} \qquad (42)$$

is SPR. As shown in Section III, in general, zero placement is necessary for SPR satisfaction.

To demonstrate the behavior of the SHARF, a series of simulations were conducted. The desired process $d(k)$ was second order, generated by filtering white noise with

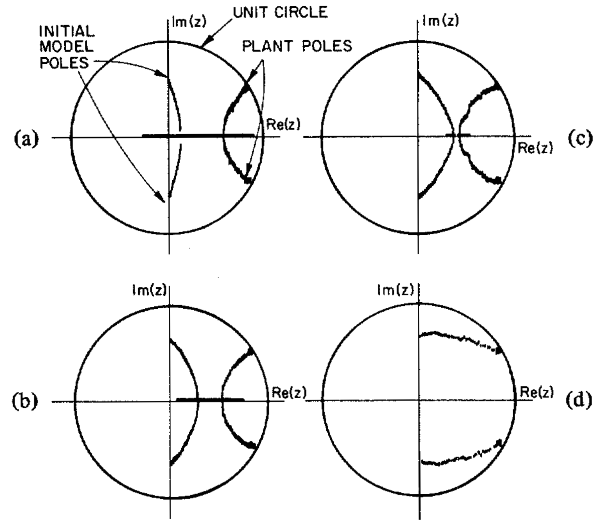


Fig. 9. Pole trajectories for simulation of the SHARF algorithm. (a) $c_1 = 0$, 160K iterations. (b) $c_1 = -0.8$, 160K iterations. (c) $c_1 = -1.0$, 120K iterations. (d) $c_1 = -2.5$, 40K iterations.

$$\frac{0.057}{1 - 1.645\, z^{-1} + 0.9025\, z^{-2}}. \qquad (43)$$

The migrations of the adaptive filter's two poles are shown in Fig. 9, starting at the imaginary poles, $0.6e^{\pm j90^\circ}$, converging to the poles of (43), $0.95e^{\pm j30^\circ}$. The four trajectories show the effects of a single smoothing coefficient c_1, beginning with $c_1 = 0$ and ranging through $c_1 = -2.5$. Pole migration, of course, is a complicated transformation of the adaptive weight locus, and as such provides a distorted view of filter behavior. However, on a qualitative level, it can be seen that the variation of the smoothing parameter not only reduces meanderings but also speeds convergence. This is partly due to the effect that smoothing has on the strength of the error process $v(k)$, in turn reducing the average size of the algorithm's update terms in (38).

In this example, it is worth noting that despite violation of the SPR requirement, for example, when $c_1 = 0$, convergence did still occur. This simply indicates that the sufficiency condition was overly restrictive in the example. Consider a second case, involving a second-order process generated by a filter

$$\frac{1}{1 - 1.7z^{-1} + 0.7225\, z^{-2}}, \qquad (44)$$

having a pair of real poles at 0.85. In this case, the SHARF algorithm was simulated using $c_1 = 0$ and $c_1 \gtrsim -1$; the first case does not satisfy the SPR requirement, while the second case does. To eliminate the ambiguities of convergence rate, the adaptive filter was initialized to have its poles at 0.845, within 0.005 of the true location. Despite the proximity of the adaptive filter to its desired parameter set, for $c_l = 0$, i.e., recursive LMS, (in this nonpositive real case) the weights quickly readjusted to an alternative configuration, involving a single low frequency pole, effectively discarding the second

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

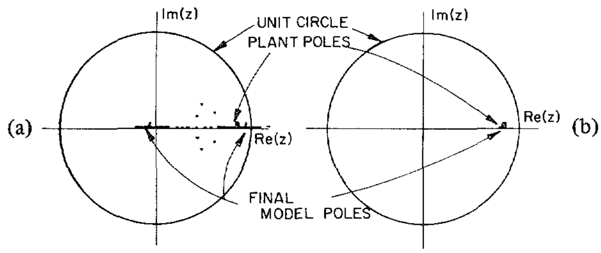


Fig. 10. Pole trajectories for the recursive LMS and SHARF algorithms. (a) Recursive LMS ($c_1 = c_2 = 0$). (b) SHARF ($c_1 = -1.0, c_2 = 0$).

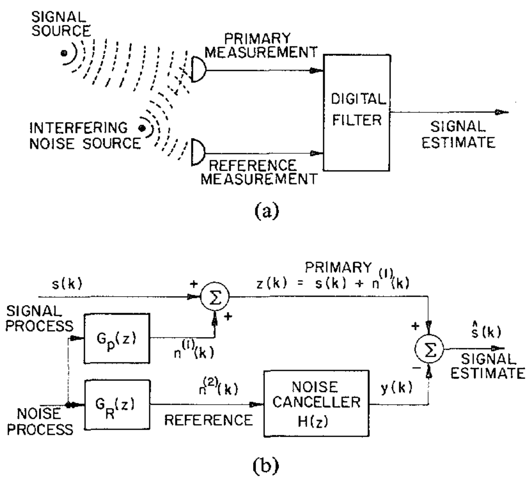


Fig. 11. Noise cancelling signal model. (a) Physical model. (b) Lumped, linear model.

degree of freedom. In the second case, convergence was consistent to the unbiased pole estimates (see Fig. 10).

V. SHARF Applied to Adaptive Noise Cancelling

The next two sections discuss some applications of the SHARF algorithm which clearly benefit from IIR filtering [16], [17]. Previous work in these areas has been done successfully with adaptive FIR filters; however, situations can exist where the IIR configuration is almost a requirement for real-time processing.

The use of adaptive filters in signal enhancement and noise-suppression has been the focus of much research in past years [1]-[4], [32]-[34], and can be found in applications ranging from biomedical measurements to antenna beam-forming. In particular, [4] presents the basis for adaptive noise cancelling (ANC), and describes several potential applications. Fig. 11 depicts a model for the ANC situation. It is desired to estimate the signal component $s(k)$, measurable only in the presence of an additive uncorrelated noise process $n^{(1)}(k)$; this observed process is called the *primary input*

$$z(k) = s(k) + n^{(1)}(k). \tag{45}$$

By virtue of the geometry, a second sensor is able to provide a *reference* measurement of a related noise process

$$x(k) = n^{(2)}(k). \tag{46}$$

Proper filtering of the reference process, as shown in Fig. 11(b), can provide a substantial reduction in the interfering noise in the primary, thereby improving the estimate of $s(k)$.

From Fig. 11(b), the lumped, linear model of transmission path characteristics between noise source and sensors, it is clear that the ideal noise cancelling filter

$$H^*(z) = \frac{G_P(z)}{G_R(z)} \tag{47}$$

would result in perfect cancellation of noise in the signal estimate. Two possible situations are noteworthy: 1) the reference path, modeled as $G_R(z)$, is characterized by a spectral region of low energy and 2) the primary noise path $G_P(z)$ possesses a resonance. In either case, the ideal filter $H^*(z)$ contains poles and requires an IIR configuration.

Because *a priori* design of the noise canceller is impractical in real applications involving unknown, nonstationary noise statistics, the noise cancelling filter is normally implemented adaptively; in particular, [4] uses the FIR LMS adaptive algorithm. To achieve suitable noise reduction with an FIR filter, its impulse response must be long enough to span the significant portions of the ideal impulse response. In cases where $H^*(z)$ possesses dominant pole behavior, this means that an FIR approximation may require an unwieldy number of adjustable parameters, resulting in 1) a heavy computational load and 2) a convergence rate slowed as necessary to control performance degradation due to parameter "misadjustment" [13]. The latter effect sacrifices a principal advantage of adaptive filtering, specifically, the ability of the processor to track changes in the noise environment.

If an adaptive IIR filter were used instead, the potential exists for improving noise suppression while requiring less computation. Examination of Fig. 12 indicates that a rearrangement of the original noise cancelling configuration yields the same output error identification problem described in Section II. The only addition is an independent "noise" $s(k)$ present at the output of $H^*(z)$. Heuristic arguments can be made as to why this "observation noise" has no effect on the converged model, i.e., the noise cancelling filter [16].

For the purposes of demonstration, the results of a simple simulation are given, where the assumption of matching order is met. The noise source emits a unity power white Gaussian noise process. The transmission paths are explicitly

$$G_P(z) = \frac{0.1630}{1 - 1.739 z^{-1} + 0.81 z^{-2}} \tag{48}$$

$$G_R(z) = 1.0,$$

implying that

$$H^*(z) = \frac{0.1630}{1 - 1.739 z^{-1} + 0.81 z^{-2}}. \tag{49}$$

In the time domain, this optimal impulse response is given by

$$h^*(k) = 0.1630(0.9)^k \cdot \{\cos(0.2618k) + 3.732 \sin(0.2618k)\}$$

whose envelope decays to under 10 percent only after 21

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

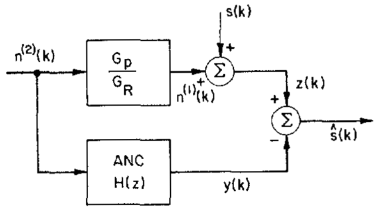


Fig. 12. Noise canceller as output error identifier.

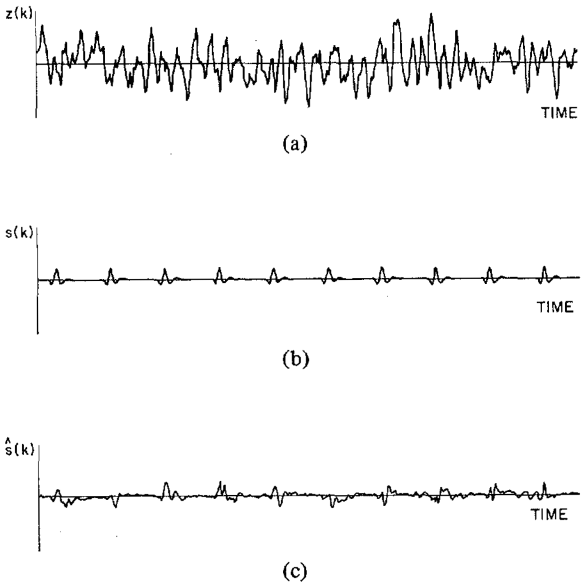


Fig. 13. Performance of SHARF noise canceller. (a) Noise-corrupted primary input. (b) Signal component. (c) Signal estimate, the noise canceller output.

terms. For this simulation, the interfering noise $n^{(1)}(k)$ is a narrow-band process of center frequency $f_s/24$ where f_s is the sampling frequency. Refer to Fig. 13(a), which shows a typical trace of 1000 samples of the primary measurement. The signal component, which is present in this primary trace is seen alone to scale in Fig. 13(b). Note that it is a simple periodic, impulsive waveform, completely masked by the strong noise component $n^{(1)}(k)$. If an IIR filter with one numerator parameter and two denominator parameters were updated using the SHARF algorithm with $\rho = 0.005$ and $\mu = 0.1$, in about 2000 adaptations a reasonable degree of noise suppression can be detected, around 13 dB improvement. In Fig. 13(c), the corresponding output trace is shown. Qualitatively speaking, the signal spikes have become easily detectable in the residual noise. For this simulation, a single c_i coefficient was used, $c_1 \gtrsim -1.0$. In Fig. 8 it is indicated that such a value assures the SPR of (32) in this case.

For comparison, a conventional FIR noise cancelling filter is applied to the same data. The filter was allowed six adaptive weights, i.e., requiring 50 percent more computation than the adaptive IIR filter previously simulated. The adaptive rate was chosen to provide convergence in about 1500 iterations, requiring approximately a comparable total computational effort. At convergence, only about 5 dB noise suppression was achieved. For the same rate of convergence, increased number of adaptive weights will improve performance, but only at a severe increase in computation:

9 weights	10 dB
12 weights	11 dB
15 weights	12 dB.

Note that the improvements are limited by misadjustment noise effects as the number of weights increases.

As a demonstration of the robustness of the SHARF ANC processor, another simulation was formulated. Two major additions were incorporated, both reflecting practical considerations. First, an independent noise component was added to the primary input process. This would be typical for most applications, where multiple noise sources could be responsible for signal interference. Naturally, such an uncorrelated noise component provides a bound on the achievable suppression, since it cannot be removed by noise cancelling. Secondly, the adaptive filter was constrained to an insufficient number of parameters. In practice, this might occur if the exact order were unknown, or if the transmission paths were of a nonrational nature, due to distributed system effects. For this case, the ideal noise cancelling filter would require four numerator coefficients and three denominator coefficients, whereas the SHARF noise canceller was allowed only four parameters equally divided between numerator and denominator. Independent white Gaussian noise was added at the primary input of relative power level -10 dB. At convergence, roughly 7.5 dB of noise reduction was achieved. By no means does this adaptive processor provide optimal noise suppression in the least-squares sense. Interestingly, however, noise cancelling based on the SHARF IIR filter is quite successful, despite violations of the basic input assumptions.

VI. SHARF Applied to Adaptive Multipath Cancelling

The received signal in a communication link may be subject to many forms of degradation, among them are the effects of *multipath propagation.* In such cases, the received energy usually results from the convergence of distinct reflections. Associated with each scatterer is a path length and corresponding propagation time, and a scaling factor determined by spatial losses, reflective cross section, and energy absorption by the reflector. Ideally, the received signal is a linear combination of delayed versions of the transmitted signal. Since each reflective path contributes a portion of the transmitted signal at a given time delay, the result in high bit-rate digital transmission can be gross intersymbol interference; the path having the largest delay time defines the necessary intersymbol guardband and hence the available bandwidth and allowable bit rate [35].

In many communication systems the data rate cannot be reduced to accommodate intersymbol interference. In such cases, the effects of multipath on a single element receiver must rely upon a form of bandwidth equalization, similar to techniques used to overcome spectral distortion due to receiver characteristics, modulation schemes, or channel response. Recently, a multipath equalization technique was presented based on the principle of adaptive linear prediction,

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

implemented using the FIR configuration [35]. The adaptive nature of the processor eliminates the requirements for certain *a priori* design information, while allowing flexibility to track changes in the constellation of reflectors.

As discussed in [17], [35], in its simplest form ideal multipath can be characterized in discrete time by a received signal

$$x(k) = s(k) + \alpha s(k - \Delta), \qquad |\alpha| < 1 \tag{50}$$

where $s(k)$ is the transmitted signal. Note that only a single reflection is considered significant. This so-called "two-ray" case is of fundamental importance in point-to-point communications. The channel's equivalent transfer function is then

$$G(z) = 1 + \alpha z^{-\Delta} \tag{51}$$

and requires an ideal equalizer

$$H_{eq}(z) = \frac{1}{1 + \alpha z^{-\Delta}} \tag{52}$$

to remove the accompanying amplitude and phase distortion.

While (51) characterizes multipath effects in the frequency domain, a more meaningful interpretation can be had by viewing the received signal's autocorrelation function,

$$R_x(l) = (1 + \alpha^2) R_s(l) + \alpha R_s(l - \Delta) + \alpha R_s(l + \Delta). \tag{53}$$

If the channel suffers from a *resolvable multipath* with respect to the bandwidth requirements, i.e.,

$$R_s(l) R_s(l - \Delta) \cong 0, \tag{54}$$

then equalizer given by (52) can be implemented adaptively in the linear predictor configuration shown in Fig. 14. The filter $H(z)$ predicts an estimate of $x(k)$ from observations that are at least L samples removed. It can be shown that the predictor provides an output whose autocorrelation ideally has been eliminated beyond lag L.

From the configuration of Fig. 14, it can be shown [17] that ideal equalization will be achieved if the predictor's filter is chosen as the IIR expression

$$H(z) = \frac{\alpha z^{-\Delta}}{1 + \alpha z^{-\Delta}}. \tag{55}$$

The FIR implementation described in [35] provides only an approximation to the ideal predictor and its performance can be seriously degraded as α approaches unity.

On the other hand, use of an adaptive IIR filter as a linear predictor, driven by the SHARF algorithm, can provide ideal multipath cancellation. To demonstrate, the results of a simple experiment will be presented. A stochastic signal $s(k)$ was generated by filtering bipolar white noise samples with a moving average filter of length ten. The signal's autocorrelation was then triangular,

$$R_s(l) = \begin{cases} 1 - |l|/10 & l \leq 10 \\ 0 & l > 10. \end{cases} \tag{56}$$

The multipath was given strong amplitude, $\alpha = 0.9$, and a sizable delay, $\Delta = 30$,

$$x(k) = s(k) + 0.9 s(k - 30). \tag{57}$$

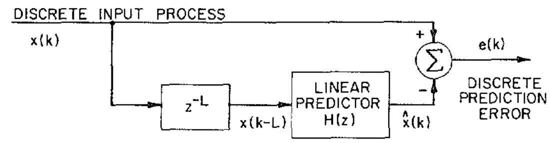


Fig. 14. Discrete linear predictor.

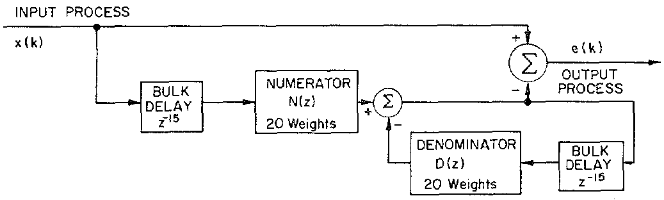


Fig. 15. SHARF multipath canceller.

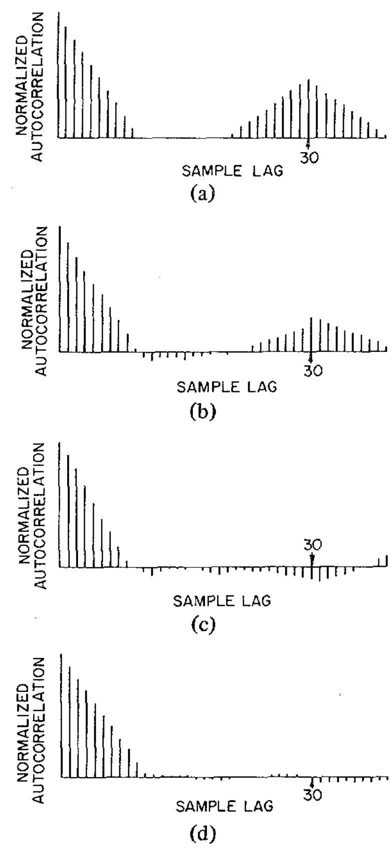


Fig. 16. Normalized autocorrelation of processed data. (a) Prior to adaptation. (b) 1000 iterations. (c) 10 000 iterations. (d) 25 000 iterations.

Note that the multipath is resolvable, since the correlation time of the transmitted signal, ten samples, is less than one-half of the multipath delay, $\Delta = 30$. As the samples were taken at the receiver, they were processed by an adaptive multipath canceller, having the configuration shown in Fig. 15. Note that the bulk delay was 15 samples, exceeding the correlation time of $s(k)$. The 40 adaptive weights were initialized to zero.

Since the goal of the processor is to reduce the intersymbol interference caused by multipath transmission, evaluation of performance can be made by examining the autocorrelation of the prediction error. Fig. 16 shows this autocorrelation,

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

normalized to unity power, as the adaptive processor evolves over time. Fig. 16(a) shows the autocorrelation of the received data prior to processing. Notice the two resolvable peaks, one corresponding to the direct path and the second due to the reflected path. After 1000 samples have passed through the adaptive filter, the error process is characterized by the correlation of Fig. 16(b). Note that the primary peak is relatively untouched, while a significant reduction of the multipath effects has begun. At 10 000 iterations, Fig. 16(c) shows that further improvement has been realized. Finally, Fig. 16(d) shows the correlation achieved after 25 000 sampling intervals. At this point, the intersymbol interference has been virtually eliminated.

Of course, the fact that the final correlation is very close to that of the transmitted signal does not necessitate good fidelity. However, in this simulation the adaptive coefficients converged to values very close to those given by (55); consequently, the error process is indeed a good estimate of $s(k)$. This particular simulation, while dramatic, was actually bordering on worst case conditions. The choice of large α meant that the poles of the adaptive prediction filter must approach the unit circle, thereby resulting in high sensitivity to small random variations in the coefficients.

VII. Conclusions

The work described in this article has produced a class of algorithms particularly suited to adaptive IIR filtering. Two members of this class here highlighted here, HARF, a moderately complex algorithm with provable convergence properties, and SHARF, a much simplified version of HARF which retains provable convergence for small values of the adaptation constants. It is interesting to note that the LMS algorithm, used for both FIR and IIR adaptive filtering, is a special (albeit very important) case of the SHARF algorithm.

This paper has outlined the basis of the HARF algorithm, discussed the applicability of the hyperstability concept to the adaptive filtering problem, and has presented the development of the SHARF algorithm. In the process several examples of its convergence characteristics were demonstrated which show significant performance improvement over various gradient schemes which have been suggested. In addition, two practical signal processing examples have been presented which demonstrate the utility of an IIR adaptive filter. In short we have developed a class of adaptive algorithms which make it possible to obtain the often dramatic performance/computation advantages offered in many applications by IIR filters. Even so, many questions remain to be answered before this class is fully characterized. Some of them are as follows.

1) How should the vector of auxiliary coefficients $\{c_i\}_{i=1}^{P}$ be chosen and can this be done adaptively?

2) Can the convergence characteristics of the algorithm be predicted?

3) What are practical upper limits on the choice of the adaptation constants, and hence the convergence rates, for the SHARF algorithm?

4) How do observation noise and additive input noise affect the performance and convergence properties of SHARF and HARF?

5) What happens when the filter has insufficient complexity to achieve zero error?

Further understanding in these areas will only enhance the attractiveness of adaptive IIR filtering for signal processing applications.

Acknowledgment

The authors wish to thank P. Thompson and T. Saxe of Stanford University for their computer assistance and helpful discussions during the initial stages of this work. The occasional use of ARGOSystems' conference call capability is also appreciated. The authors express their apologies to L. L. Scharf of Colorado State University for their unfortunate choice of acronym for the simplified algorithm.

References

[1] L. J. Griffiths, "A simple adaptive algorithm for real-time processing in antenna arrays," *Proc. IEEE*, vol. 57, pp. 1696–1704, Oct. 1969.

[2] O. L. Frost, III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926–935, Aug. 1972.

[3] A. Gersho, "Adaptive equalization of highly dispersive channels for data transmission," *Bell Syst. Tech. J.*, vol. 48, pp. 55–70, Jan. 1969.

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

[5] S. D. Stearns, G. R. Elliott, and N. Ahmed, "On adaptive recursive filtering," in *Proc. 10th Asilomar Conf. Circuits, Syst., Comput.*, Pacific Grove, CA, Nov. 1976, pp. 5–10.

[6] S. A. White, "An adaptive recursive filter," in *Proc. 9th Asilomar Conf. Circuits, Syst., Comput.*, Pacific Grove, CA, Nov. 1975, pp. 21–25.

[7] P. L. Feintuch, "An adaptive recursive LMS filter," *Proc. IEEE*, vol. 64, pp. 1622–1624, Nov. 1976.

[8] D. Parikh, N. Ahmed, and S. D. Stearns, "An adaptive lattice algorithm for recursive filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 110–111, Feb. 1980.

[9] P. A. Thompson, "A constrained recursive adaptive filter for enhancement of narrowband signals in white noise," in *Proc. 12th Asilomar Conf. Circuits, Syst., Comput.*, Pacific Grove, CA, Nov. 1978, pp. 214–218.

[10] C. R. Johnson, Jr., "A convergence proof for a hyperstable adaptive recursive filter," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 745–749, Nov. 1979.

[11] J. R. Treichler, M. G. Larimore, and C. R. Johnson, Jr., "Simple adaptive IIR filtering," in *Proc. 1978 Int. Conf. Acoust., Speech, Signal Processing*, Tulsa, OK, Apr. 1978, pp. 118–122.

[12] V. M. Popov, *Hyperstability of Control Systems.* Berlin: Springer-Verlag, 1973.

[13] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.

[14] I. D. Landau, "Unbiased recursive identification using model reference adaptive techniques," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 194–202, Apr. 1976.

[15] C. R. Johnson, Jr., M. G. Larimore, and J. R. Treichler, "Simple adaptive output error identification," in *Proc. 10th Southeastern Symp. Syst. Theory*, Mississippi State University, Mar. 1978, pp. III-A-1-8.

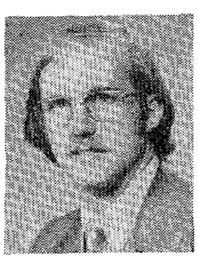
[16] M. G. Larimore, C. R. Johnson, Jr., and J. R. Treichler, "Adaptive cancelling using SHARF," in *Proc. 21st Midwest Symp. Circuits Syst.*, Iowa State University, Ames, Aug. 1978, pp. 124–129.

[17] M. G. Larimore, J. R. Treichler, and C. R. Johnson, Jr., "Multipath cancellation by adaptive recursive filtering," in *Proc. 12th Asilomar Conf. Circuits, Syst., Comput.*, Pacific Grove, CA, Nov. 1978, pp. 219–223.

[18] C. R. Johnson, Jr. and M. G. Larimore, "Comments on and additions to 'An adaptive recursive LMS filter,'" *Proc. IEEE*, vol. 65, pp. 1399–1401, Sept. 1977.

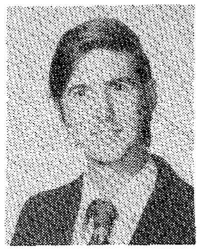
Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.

[19] D. Parikh and N. Ahmed, "On an adaptive algorithm for IIR filters," *Proc. IEEE*, vol. 65, pp. 585-587, May 1978.
[20] I. D. Landau, "A survey of model reference adaptive techniques–Theory and applications," *Automatica*, vol. 10, pp. 353-379, July 1974.
[21] J. M. Mendel, *Discrete Techniques of Parameter Estimation.* New York: Marcel Dekker, 1973.
[22] C. R. Johnson, Jr., "Adaptive parameter matrix and vector output estimation via an equation error formulation," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, pp. 392-397, July 1979.
[23] C. R. Johnson, Jr. and A. L. Hamm, "Debiasing equation error frequency detection of a noisy sinusoid without noise variance prespecification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 654-656, Dec. 1979.
[24] B. D. O. Anderson, "A simplified viewpoint of hyperstability," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 292-294, June 1968.
[25] M. G. Larimore, "Hyperstability and adaptive filtering," in *Proc. 13th Asilomar Conf. Circuits, Syst., Comput.*, Pacific Grove, CA, Nov. 1979, pp. 412-415.
[26] L. Hitz and B. D. O. Anderson, "Discrete positive real functions and their application to system stability," *Proc. Inst. Elec. Eng.*, vol. 116, pp. 153-155, Jan. 1969.
[27] C. R. Johnson, Jr., J. R. Treichler, and M. G. Larimore, "Remarks on the use of SHARF as an output error identifier," in *Proc. 17th IEEE Conf. Decision Contr.*, San Diego, CA, Jan. 1979, pp. 1094-1095.
[28] C. R. Johnson, Jr., "SCHARF," VPI & SU Dep. Elec. Eng., Tech. Rep. EE-7910, June 1979.
[29] D. Parikh, S. C. Sinha, and N. Ahmed, "On a modification of the SHARF algorithm," in *Proc. 22nd Midwest Symp. Circuits Syst.*, Philadelphia, PA, June 1979, pp. 362-366.
[30] C. R. Johnson, Jr. and T. Taylor, "CHARF convergence studies," in *Proc. 13th Asilomar Conf. Circuits, Syst., Comput.*, Pacific Grove, CA, Nov. 1979, pp. 403-407.
[31] I. D. Landau, "Elimination of the real positivity condition in the design of parallel MRAS," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 1015-1020, Dec. 1978.
[32] J. R. Treichler, "Transient and convergent behavior of the adaptive line enhancer," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 53-62, Feb. 1979.
[33] J. R. Glover, Jr., "Adaptive noise cancelling applied to sinusoidal interferences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 484-491, Dec. 1977.
[34] R. J. Keeler and L. J. Griffiths, "Acoustic Doppler extraction of adaptive linear prediction filtering," *J. Acoust. Soc. Amer.*, vol. 61, pp. 1218-1227, May 1977.
[35] D. R. Morgan, "Adaptive multipath cancellation for digital data communications," *IEEE Trans. Commun.*, vol. COM-26, pp. 1380-1390, Sept. 1978.



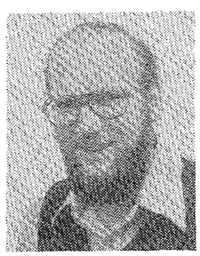
Michael G. Larimore (S'69-M'78) was born in Cleveland Heights, OH, on January 7, 1949. He received B.S. and M.S. degrees in electrical engineering from the University of Colorado, Boulder, in 1971 and 1972, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1977.

Since 1977 he has been an Assistant Professor at the University of Colorado, Colorado Springs. During the summer of 1979 he was employed by IBM Research, San Jose, CA, studying communication and signal processing techniques applied to high density magnetic recording. His current research activities involve applications of signal processing techniques to spectrum analysis, system identification, multipath and noise suppression, and transmission and magnetic storage of digital information.



John R. Treichler (S'69-M'78) was born in Velasco, TX, on September 22, 1947. He received the B.A. and Masters degrees in electrical engineering from Rice University, Houston, TX, in 1970, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA in 1977.

From 1970-1974 he was in the U.S. Navy where he attained the rank of Lieutenant and the billet of Operations Officer. He spent the next three years at Stanford University, performing doctoral research in the area of adaptive signal processing. Since 1977 he has been a member of the Technical Staff, ARGOSystems, Inc., Sunnyvale, CA. In addition, he serves as a lecturer at Stanford University. His current interests include spectrum analysis, estimation theory, communication theory, and adaptive signal processing, particularly as applied to sonar and radar systems. He is currently working on the application of high-speed digital hardware to radio communications problems.



C. Richard Johnson, Jr. (S'74-M'77) was born in Macon, GA, on May 27, 1950. He received the B.E.E. degree with high honors in 1973 from the Georgia Institute of Technology, Atlanta, and the M.S.E.E. degree in 1975 and the Ph.D. degree in electrical engineering with minors in engineering-economic systems and art history in 1977, both from Stanford University, Stanford, CA.

For the 1972-1973 academic year, he received a World Student Fund scholarship for study at the Technische Universität, Hannover, Germany. At Stanford University he was a Teaching Assistant for courses in control theory and digital signal processing from 1973-1975 and from 1976-1977. During the 1975-1976 academic year he was a Research Assistant in the Information Systems Laboratory at Stanford University. As a 1978 NASA-ASEE Summer Faculty Fellow at the NASA Langley Research Center, he investigated the adaptive control of large space structures. In 1977 he joined the Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, as an Assistant Professor. He currently teaches and conducts research in the areas of adaptive systems theory and control theory.

Dr. Johnson is a member of the American Institute of Aeronautics and Astronautics, Sigma Xi, Eta Kappa Nu, Phi Kappa Phi, and Tau Beta Pi.

Authorized licensed use limited to: Cornell University Library. Downloaded on August 01,2020 at 22:29:40 UTC from IEEE Xplore. Restrictions apply.