the recursive filter does not have enough degrees of freedom to represent the Wiener filter) then $\hat{a}$ and $\hat{b}$ will not, in general, converge to a and b since the assumptions in (3) and (4) are that the correlation functions correspond to those of the Wiener filter. If, for example $H(z^{-1})$, the filter to be modeled, is given instead by

$$H(z^{-1}) = \frac{a}{1 - bz^{-1}}$$

then

$$\hat{a} = a$$

$$\hat{b} = \frac{(1-b^2)b}{a} \frac{a}{1-b^2} = b.$$

That is, the equilibrium values for $\hat{a}$ and $\hat{b}$ are the Wiener filter settings. Thus the recursive algorithm, (7) and (8), converges to the Wiener filter.

Now let us return to (7) and (8). Because none of the statistics are known *a priori*, (9) and (10) replace them by estimates. The estimates of R_{XX}, R_{Xd} are unbiased. The estimates of R_{XY}, R_{YY}, and R_{dY} are possibly biased but more importantly, are functions of A and B.

Johnson and Larimore assume, and are supported by simulation results, that the steady state behavior of the mean weights of the system are statistically independent of the present data. This assumption is implicit in being able to solve equation (D) explicitly for the stationary point of the algorithm in [1]. Thus by their two examples they have supported the hypothesis that (5) and (6) give the values to which the means of (9) and (10) (the recursive LMS algorithm) converge. The recursive LMS algorithm does not converge to the least mean-square error filter in this example since (7) and (8) do not converge to the MMSE weights for a and b. This does not imply that the algorithm is faulty since the examples of the above letter do not legitimately consider unconstrained LMS filters. The problem of finding the best filter within a given structural class is not the basis for deriving a Wiener filter. The Wiener filter can be derived by using the orthogonality principle, making the error orthogonal to all of the data. Equation (D) is not the orthogonality principle since we should require

$$E[\epsilon(k)x(j)] = 0, \qquad \text{for all } j \leqslant k.$$

It is suspected that the second condition

$$E[\epsilon(k)y(j)] = 0, \qquad \text{for all } j \leqslant k - 1$$

implies the first whenever the feedback structure has sufficient degrees of freedom to represent the Wiener filter.

It is important to note that seeking the Wiener filter implemented in a specified recursive structure is the motivation in [1]. The resulting filter may not provide the minimum mean square error in all cases. The object was to obtain a filter that reduces mean-square error. The resulting algorithm is of interest, since in many cases it does this, and it, rather than its motivation, should be the subject of discussion. The point here is that for the case presented in the above letter, the Wiener filter cannot be implemented for the severe set of constraints on the problem. This is evidenced by the minimum-mean-square error occurring at a point where the error is *not* orthogonal to the data. Since the recursive adaptive algorithm attempts to set the error orthogonal to the data, for this example it provides a higher mean-square error.

Among the features of the adaptive recursive LMS filter is that when given sufficient degrees of freedom, i.e., two feedforward and two feedbackward adaptive taps to model the two pole-two zero fixed parameter network, then the system converges to the proper tap values. This is documented in Table I. In addition, if the adaptive algorithm is given three feedforward and three feedback taps, then it converges more rapidly to a solution with smaller mean square error, but with different tap values (Table II). Interestingly, the transfer function of the recursive adaptive filter matches the network being modeled. That is, given additional degrees of freedom, the algorithm produces a redundant pole-zero pair which cancels. The same property holds for higher order filters. On the other hand, as the number of taps in the adaptive system is reduced below that required to provide a solution, the degradation in performance is not the dramatic threshold behavior implied in the above letter. Instead, for higher order systems, it is gradual, becoming more severe as fewer taps are used, as simulations have shown.

In most adaptive system applications, the order of the system required is not known. The designer allocates more taps to the problem than the minimum number sufficient to provide a solution. The limitation pointed out does not present a restriction in practice and should not detract from the large class of problems which the processor can handle. Determination of that class is an area of active research.

TABLE I

MODELING $H(Z) = (0.05 - 0.40Z^{-1})/(1 - 1.1314Z^{-1} + 0.25Z^{-2})$ WITH THE RECURSIVE ADAPTIVE FILTER WITH TWO FEEDFORWARD AND TWO FEEDBACK TAPS; $k_1 = k_2 = -4.3 \times 10^{-4}$; WEIGHTS INITIALLY ALL ZERO

Number of Iterations, n	$a_0(n)$	$a_1(n)$	$a_2(n)$	$b_1(n)$	$b_2(n)$	$b_3(n)$	Normalized rms error
8192	.0456	-.3633	-.2968	-.3816	-.3845	-.0573	.2096
16384	.0520	-.3649	-.2701	-.4497	-.4431	.0713	.0939
24576	.0503	-.3676	-.2599	-.4805	-.4530	.1273	.0378
32768	.0503	-.3679	-.2552	-.4926	-.4613	.1466	.0143
40960	.0502	-.3682	-.2538	-.4967	-.4639	.1539	.0049
49152	.0500	-.3683	-.2533	-.4982	-.4648	.1566	.0018
57344	.0500	-.3684	-.2531	-.4988	-.4652	.1576	.0007
65536	.0500	-.3684	-.2529	-.4989	-.4653	.1580	.0003

TABLE II

MODELING $H(Z) = (0.05 - 0.40Z^{-1})/(1 - 1.134Z^{-1} + 0.25Z^{-2})$ WITH THE RECURSIVE ADAPTIVE FILTER WITH THREE FEEDFORWARD AND THREE FEEDBACK TAPS; $k_1 = k_2 = -4.3 \times 10^{-4}$; WEIGHTS INITIALLY ALL ZERO

Number of Iterations, n	$a_0(n)$	$a_1(n)$	$b_1(n)$	$b_2(n)$	Normalized Error rms
8192	.0460	-.374	-.636	-.295	.4984
16384	.0568	-.381	-.817	-.095	.3652
24576	.0503	-.393	-.945	.056	.2508
32768	.0518	-.393	-1.035	.141	.1545
40960	.5024	-.398	-1.082	.194	.0807
49152	.0497	-.398	-1.107	.223	.0416
57344	.0497	-.400	-1.121	.238	.0205
65536	.0502	-.400	-1.126	.244	.0090
"correct value"	.05	-.40	-1.131	.25	

The two simulations presented in [1] were representative of a large number of runs under diverse inputs and initial conditions. The adaptive network was given a sufficient number of taps capable to handle the problems. Significantly, it did provide effective solutions. Previous work in this area [2] indicated that an adaptive recursive filter was not technically feasible. Our results have demonstrated otherwise.

REFERENCES

[1] P. L. Feintuch, "An adaptive recursive LMS filter," *Proc. IEEE*, vol. 64, no. 11, pp. 1622–1624, Nov. 1976.

[2] P. E. Mantey, "Convergent automatic synthesis procedures for sampled data networks with feedback," Stanford Electronics Lab., Stanford University, Stanford, CA, Tech. Rep. 6773-1, Oct. 1964.

Comments on "An Adaptive Recursive LMS Filter"

BERNARD WIDROW AND JOHN M. MCCOOL

The usual application of the "least mean-square" or LMS algorithm of Widrow and Hoff [1] is to nonrecursive or feedback-free adaptive systems. An example of such a system is the adaptive transversal filter, which has been shown to be capable, when its operation is governed by the LMS algorithm, of adjusting itself to minimize mean-square error, [2], [3] where "error" is defined as the difference between the filter's

Manuscript received January 31, 1977.

B. Widrow is with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305.

J. McCool is with the Fleet Engineering Department, Naval Undersea Center, San Diego, CA 92132

output and a "desired response" or externally supplied training signal. The drawback of the nonrecursive LMS filter is that it has a finite impulse response and can realize only zeros of a digital filter transfer function.[1]

In the above letter,[2] a recursive adaptive filter based on the LMS algorithm has been described. This particular filter is structurally capable of realizing both zeros and poles of a transfer function and of having an infinite impulse response. It thus promises to be a useful and powerful tool in certain practical applications, indicated below. We must regrettably report, however, that the mathematical derivation of the recursive LMS algorithm presented in the letter is incorrect and that, contrary to Feintuch's claim, the algorithm does not in general minimize mean-square error.[3]

Feintuch specifies the recursive filter by the equation

$$y(n) = \sum_{k=0}^{N_f} a_k s(n-k) + \sum_{k=1}^{N_B} b_k y(n-k). \tag{1}$$

In vector notation he obtains

$$y(n) = A^T X(n) + B^T Y(n). \tag{2}$$

He defines the error as the difference between the desired response $d(n)$ and the actual response $y(n)$

$$\epsilon(n) = d(n) - y(n) = d(n) - A^T X(n) - B^T Y(n). \tag{3}$$

He then squares and takes expected values to obtain the mean-square error

$$E[\epsilon^2(n)] = E[d^2(n)] + A^T R_{XX} A + B^T R_{YY} B - 2A^T R_{dX} - 2B^T R_{dY} + 2A^T R_{XY} B \tag{4}$$

where the covariance terms are defined as

$$R_{XX} = E[X(n)X^T(n)],\quad R_{YY} = E[Y(n)Y^T(n)],\quad R_{dX} = E[d(n)X(n)],\quad R_{dY} = E[d(n)Y(n)],\quad \text{and } R_{XY} = E[X(n)Y^T(n)].$$

Since all algorithms in the LMS family [4]-[19] are based on optimization by the method of steepest descent, the next step required is differentiation of (4) to obtain the gradient. In taking this step, however, Feintuch argues that the covariance terms R_{XY}, R_{dY}, R_{YY} are constants when differentiated with respect to the feed-forward and feedback weights A and B. This argument is incorrect because these terms are functions of A and B. The gradient expressions given by his equations (5) and (6) are thus also incorrect and the derivation of his remaining (7)-(11) invalid.

Let us examine the recursive LMS filter from another point of view. Fig. 1 shows a nonrecursive filter comprising an adaptive transversal filter whose impulse response is controlled by adjusting its weighting coefficients. This filter is an "LMS filter" when the coefficients are adjusted through the LMS algorithm. Fig. 2 shows a recursive filter comprising two adaptive transversal filters, one providing a feed-forward network and implementing zeros and the other providing a feedback network and implementing poles.[4] When the LMS algorithm is used to adjust the weights of both filters, the result is a "recursive LMS filter" identical to the one described by Feintuch. If the input signal and

[1] The terms "zero," "pole," and "transfer function" belong to the domain of of fixed filters; they are nevertheless useful in the analysis of adaptive filters, though their meaning in this context cannot yet be precisely defined.

[2] P. L. Feintuch, *Proc. IEEE*, vol. 64, pp, 1622-1624, Nov. 1976.

[3] As well as is known at the present time, the algorithm minimizes some function of the error. Experimental evidence indicates that this function may in general be *unimodal*, but what it is remains a subject of research. The mean-square error function is in general *multimodal*. When using the recursive LMS algorithm, in some cases mean-square error apparently is minimized but in others it clearly is not. Under certain conditions we have observed the recursive LMS algorithm, initially set at the minimum mean-square-error solution, to cause the weights to vary from this solution and stabilize elsewhere.

[4] The unit delay in the feedback network eliminates instantaneous feedback, whose only function would be gain control, which is accomplished by the feed-forward weights.

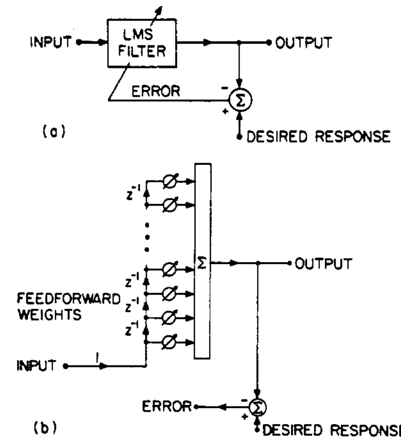

Fig. 1. Nonrecursive LMS Filter. (a) Simplified representation. (b) Schematic diagram.

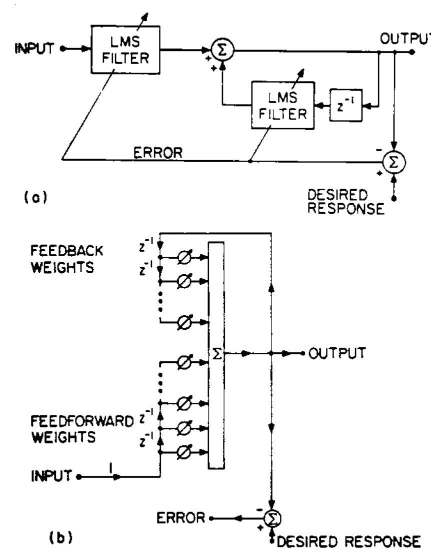

Fig. 2. Recursive LMS Filter. (a) Simplified representation. (b) Schematic diagram.

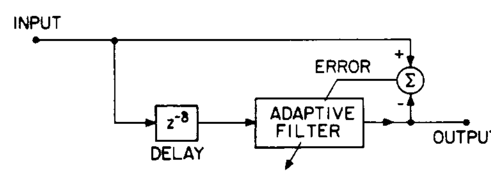

Fig. 3. The adaptive line enhancer.

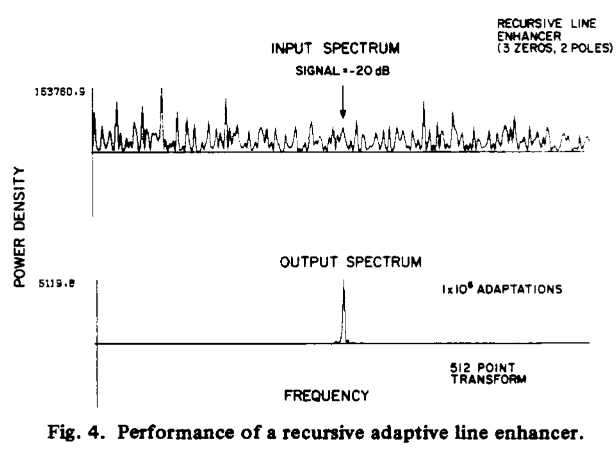

Fig. 4. Performance of a recursive adaptive line enhancer.

desired response of this filter are statistically stationary, it is clear that the covariance matrix of the inputs to the weights of the feed-forward filter is fixed and independent of the weight values, while the covariance matrix of the inputs to the feedback filter is dependent on the values of both the feed-forward and feedback filter weights. The latter dependence is characteristic of adaptive feedback systems.

The recursive adaptive filter was studied more than ten years ago by P. E. Mantey as a doctoral student at Stanford. He showed that the mean-square-error function was not quadratic and was sometimes multimodal. Mantey did not pursue his work because of the unpredictability of this filter and the difficulty of understanding its behavior. Instead he devised a recursive adaptive process using the desired response as feedback signal rather than the filter output [17]. His goal was to achieve constancy in the covariance terms and to obtain a quadratic mean-square-error function with a linear gradient. Feintuch's mathematical derivation corresponds to Mantey's second algorithm rather than to the recursive LMS algorithm.

Despite the foregoing qualifications Feintuch's work is an important contribution. He has stimulated new interest in the recursive LMS filter and has shown experimentally that it performs well as a substitute for the nonrecursive transversal filter in the self-tuning adaptive filter or "adaptive line enhancer" described by Widrow *et al.* [18] and shown in Fig. 3. In our work with the line enhancer, we have confirmed experimentally that low-level narrowband signals in noise can be effectively detected by a recursive LMS filter with poles close to the unit circle.[5] Fig. 4 shows, for example, plotted on linear scales, the input and output power spectral densities of a signal before and after processing by a filter with three zeros and two poles. The measured improvement in signal-to-noise ratio, with only five adaptive weights, is approximately 40 dB. We are thus confident that the recursive LMS filter has important potential applications in the fields of signal detection, instantaneous frequency estimation [19], and spectral analysis.

References

[1] B. Widrow and M. E. Hoff, Jr., "Adaptive switching circuits," in *IRE WESCON Conv. Rec.*, pt. 4, 1960, pp. 96-104.

[2] B. Widrow, "Adaptive filters," in *Aspects of Network and System Theory*, R. Kalman and N. DeClaris, Eds. New York: Holt, Rinehart, and Winston, 1971, pp. 563-587.

[3] B. Widrow, J. McCool *et al.*, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151-1162, Aug. 1976.

[4] J. Koford and G. Groner, "The use of an adaptive threshold element to design a linear optimal pattern classifier," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 42-50, Jan. 1966.

[5] N. Nilsson, *Learning Machines.* New York: McGraw-Hill, 1965.

[6] B. Widrow, P. Mantey, L. Griffiths, and B. Goode, "Adaptive antenna systems," *Proc. IEEE*, vol. 55, pp. 2143-2159, Dec. 1967.

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

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

[9] L. J. Griffiths and P. E. Mantey, "Iterative least-squares algorithm for signal extraction," in *1969 Proc. Second Hawaii Int. Conf. System Sciences*, Western Periodicals Co., pp. 767-770.

[10] R. Lucky, "Automatic equalization for digital communication," *Bell Syst. Tech. J.*, vol. 44, pp. 547-588, Apr. 1965.

[11] R. Lucky *et al.*, *Principles of Data Communication.* New York: McGraw-Hill, 1968.

[12] R. Riegler and R. Compton, Jr., "An adaptive array for interference rejection," *Proc. IEEE*, vol. 61, pp. 748-758, June 1973.

[13] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proc. IEEE*, vol. 63, pp. 719-720, Apr. 1975.

[14] W. F. Gabriel, "Adaptive arrays—An introduction," *Proc. IEEE*, vol. 64, pp. 239-272, Feb. 1976.

[15] B. Widrow and J. McCool, "Comparison of adaptive algorithms based on the methods of steepest descent and random search," *IEEE. Trans. Antennas Propagat.*, vol. AP-24, pp. 615-637, Sept. 1976.

[16] J. E. Timm, "The Delta-LMS adaptive algorithm," *Proc. of the 3rd IFAC Symposium*, New York: American Elsevier, June 1973, pp. 721-728.

[17] P. E. Mantey, "Convergent automatic synthesis procedures for sampled-data networks with feedback," Stanford Electronics Lab., Stanford University, Tech. Rep. 6773-1, Oct. 1964 (Ph.D. Dissertation).

[18] B. Widrow, J. Glover, J. McCool *et al.* "Adaptive noise cancelling: Principles and applications," *Proc. IEEE*, vol. 63, pp. 1692-1716, Dec. 1975.

[19] L. J. Griffiths, "Rapid measurement of instantaneous frequency," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 209-222, Apr. 1975.

[5] By experiment it appears that the recursive LMS algorithm has the extraordinary property of remaining stable even though noise in the feedback weights may occasionally push the poles outside the unit circle. The feedback of the adaptive process interacts favorably with the feedback of the filter itself to produce a "superstability" that will pull the poles back in from beyond the brink of instability.

Reply[6] by Paul L. Feintuch[7]

I should like to thank Professor Widrow and Dr. McCool for their comments and simulation examples which agree with mine [1]. I too have found that the device is stable and has potential for signal detection and spectral analysis applications. However, I must disagree with their claim that the analysis is incorrect. A heuristic derivation was presented for which the key assumptions were clearly pointed out. What is appearing here is a differing viewpoint rather than mathematical error. There is no question that when viewed as in the above letter, the output correlation statistics are functions of the weights. I was aware of this by referencing White [2]. However, the approach, instead, was that of a Wiener filter for which *a priori* statistical information is used to dictate the filter parameters rather than the reverse. For the fixed parameter case, it is certainly valid to view the problem in this way. The procedure was to assume that we have the Wiener filter and are using its input and output statistics to determine the parameters in the recursive digital filter structure. At this stage the correlation matrices are not functions of the filter weights and are thus constants when forming the gradient vectors. The gradient search procedure removes the need to invert matrices, and the problem reduces to one of obtaining the statistics, just as it did in the transversal LMS case. The open question, as was noted, is the validity of replacing the output correlations, which the Wiener filter would produce, by estimates using the instantaneous output values. These estimates are biased at the outset but have asymptotic properties which, though not yet understood, must be desirable to produce the simulation results that we have both been observing. The heuristic derivation was presented to show that a logical procedure suggested the processor structure, rather than its being an ad-hoc hookup. I suspect that when the properties of the assumed estimates are understood, the behavior of the entire filter will be understood as well.

I have also carefully examined Mantey's results [3] and do not see how his work could lead to my algorithm. Instead, his results indicate that an adaptive feedback structure could *not* have desirable stability or steady-state properties. He thus pursued network modeling in a feedforward manner only, terminating further research on a recursive adaptive filter.

The processor is by no means completely understood and its analysis is complicated. However, the device produces exciting results.

References

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

[2] S. A. White, "An adaptive versive digital filter," presented at Proc. 9th Asilomar Conf. Circuits, Systems, and Computers, Nov. 3, 1975.

[3] P. E. Mantey, "Convergent automatic synthesis procedures for sampled data networks with feedback," Stanford Electronics Lab., Stanford University, Tech. Rep. 6773-1, Oct. 1964.

[6] Manuscript received March 16, 1977.
[7] The author is with Hughes Aircraft Company, Fullerton, CA.

The Analysis of a Third-Order System

VIMAL SINGH

***Abstract*—An example of third-order nonlinear feedback system is considered. Its previously known sector for global asymptotic stability is (0, 1]. In the present letter, several sectors for global asymptotic stability are obtained.**

We consider a third-order nonlinear feedback system whose linear part $\tilde{G}(s)$ is given by

$$\tilde{G}(s) = \frac{s^2}{(s^2+1)(s+1)}. \tag{1a}$$

Manuscript received March 21, 1977.
The author is with the Department of Electrical Engineering, M.N.R. Engineering College, Allahabad 211004, India.