

BLOCK IMPLEMENTATION OF IIR ADAPTIVE FILTERS  
VIA THE EQUATION ERROR APPROACH

EMMANUEL FERNANDEZ-GAUCHERAND AND JOHN Y. CHEUNG\*

School of Electrical Engineering and Computer Science,  
University of Oklahoma, 202 W. Boyd, CEC 219, Norman, OK 73019.

ABSTRACT

The block implementation of adaptive filters is studied in this paper, for the case when both the reference and the adaptive processes are modeled as infinite impulse response (IIR) processes. The performance criterion of the adaptive filter is selected as the square of an average of the output error values over a block of iterations. Minimization of this index is indirectly sought via an equation-error-like approach, which leads to a simple updating equation for the vector of adaptive coefficients. Asymptotic consistency of the obtained block adaptive algorithm is shown, for the deterministic case, by performing a Lyapunov stability analysis. Comparisons are made between the block implementation and its counterpart for the scalar case, through a simple computational complexity analysis. These comparisons favor the scalar implementation over the block implementation. However, the possible efficient use of parallel processors make the block implementation promising.

I. INTRODUCTION

Block filtering involves the computation of a block, i.e. a finite set of consecutive values in time, of filter outputs from a block of input values and past nonoverlapping blocks of output values. For the adaptive case, the parameters of the filter are modified once per block, and the updated values are used to generate the next block of outputs.

Block IIR filtering has been thoroughly studied [1], [2], and more recently comprehensive studies of the block implementation of adaptive finite impulse response (FIR) filters have appeared in the literature [3]. Similar studies for the IIR case do not seem to be available, most probably due to the difficulty of extrapolating scalar output error techniques to the block implementation case. However, by making use of an equation-error-like concept for the corresponding iterative optimization, a computationally efficient block adaptive algorithm

is readily obtained as it has been shown to be the case for the scalar case [4], [6].

II. BLOCK REPRESENTATION OF A RECURSIVE PROCESS

Let us assume that the desired signal provided to the adaptive filter can be appropriately described as a recursive process as follows:

$$d(k) = \sum_{j=0}^{N_d} b_j u(k-j) + \sum_{l=1}^{N_d} a_l d(k-l) \quad (1)$$

and let  $f(\cdot)$  be an equally structured, equally dimensioned identifier for the process  $d(\cdot)$ , that is

$$f(k) = \sum_{j=0}^N \hat{b}_j(k) u(k-j) + \sum_{l=1}^N \hat{a}_l(k) f(k-l) \quad (2)$$

For the latter process, form blocks of  $P$  output values, taken starting at  $k=0$ , and index such blocks by  $P \in \mathcal{M}$ . To express each block as a first order recursion on past blocks, we restrict to the case  $P \geq N$ . Then, (2) can be expressed in block form as [5].

$$\mathcal{F}_{(P)} = \hat{\mathcal{A}}_0 \mathcal{F}_{(P)} + \hat{\mathcal{A}}_1 \mathcal{F}_{(P-1)} + \hat{\mathcal{B}} \mathcal{U}_{(P)} \quad (3)$$

where

$$\mathcal{F}_{(P)} \triangleq \begin{bmatrix} f(pP-1) \\ f(pP-2) \\ f(pP-3) \\ \vdots \\ f(pP-1) \end{bmatrix} \quad (4)$$

$$\mathcal{U}_{(P)} \triangleq \begin{bmatrix} u(pP-1) \\ u(pP-2) \\ u(pP-3) \\ \vdots \\ u(pP-1-M) \end{bmatrix} \quad (5)$$

$$\hat{\mathcal{A}} \triangleq \begin{bmatrix} 0 & \hat{a}_1 & \hat{a}_2 & \hat{a}_3 & \dots \\ 0 & 0 & \hat{a}_1 & \hat{a}_2 & \dots \\ 0 & 0 & 0 & \hat{a}_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & 0 & \dots & \hat{a}_1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \quad (6)$$

\*Now with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712.

$$\hat{\mathcal{A}} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \hat{a}_2 & \hat{a}_3 & \hat{a}_4 & \hat{a}_5 & \cdot \\ \hat{a}_1 & \hat{a}_2 & \hat{a}_3 & \hat{a}_4 & \cdot \end{bmatrix} \quad (7)$$

$$\hat{\mathcal{B}} = \begin{bmatrix} \hat{b}_0 & \hat{b}_1 & \hat{b}_2 & \hat{b}_3 & \cdot & \cdot \\ 0 & \hat{b}_0 & \hat{b}_1 & \hat{b}_2 & \cdot & \cdot \\ 0 & 0 & \hat{b}_0 & \hat{b}_1 & \cdot & \cdot \\ \vdots & \vdots & \vdots & \vdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \hat{b}_{M-1} \hat{b}_M \end{bmatrix} \quad (8)$$

The appropriate dimensions for (4) - (8) are  $P \times 1$ ,  $(P+M) \times 1$ ,  $P \times P$ ,  $P \times P$  and  $P \times (P+M)$ , respectively. Manipulating (3), we arrive at the expression

$$\mathcal{F}(p) = \mathcal{A}^* \mathcal{F}(p-1) + \hat{\mathcal{B}}^* \mathcal{U}(p) \quad (9)$$

where for  $I$  an identity matrix of appropriate dimensions,

$$\hat{\mathcal{A}}^* = (I - \mathcal{A}_0)^{-1} \hat{\mathcal{A}} \quad (10)$$

$$\hat{\mathcal{B}}^* = (I - \mathcal{A}_0)^{-1} \hat{\mathcal{B}} \quad (11)$$

Since  $(I - \mathcal{A}_0)$  is upper triangular, its inverse always exists and may be recursively computed [5].

By use of delay operator notation, (9) can be put in a transfer function-like form. Defining

$$q^{-s} \mathcal{F}(p) = \mathcal{F}(p-s) \quad (12)$$

then (9) can be rewritten as

$$\mathcal{F}(p) = (I - q^{-1} \hat{\mathcal{A}})^{-1} \hat{\mathcal{B}}^* \mathcal{U}(p) \quad (13)$$

and it is easily seen that

$$(I - q^{-1} \hat{\mathcal{A}})^{-1} = q(qI - \hat{\mathcal{A}})^{-1}$$

thus this inverse always exists due to the linear independence introduced by the term  $Iq$  to the rows and columns of  $(qI - \hat{\mathcal{A}})$ . Similarly, (1) can be expressed in block form as

$$\mathcal{Z}(p) = (I - q^{-1} \mathcal{A})^{-1} \mathcal{B} \mathcal{U}(p) \quad (14)$$

with  $\mathcal{Z}(p)$ ,  $\mathcal{A}^*$  and  $\mathcal{B}^*$  defined similarly as (4), (10) and (11), respectively.

### III. BLOCK EQUATION ERROR

We can view equations (13) and (14) as a block systems identification scheme. Then the block output error is defined as

$$\mathcal{E}(p) = \mathcal{Z}(p) - \mathcal{F}(p) \quad (15)$$

which can be rewritten, using (13), as

$$\mathcal{E}(p) = (I - q^{-1} \hat{\mathcal{A}})^{-1} \{ (I - q^{-1} \hat{\mathcal{A}}) \mathcal{Z}(p) - \hat{\mathcal{B}}^* \mathcal{U}(p) \} \quad (16)$$

The above equation can be interpreted as the second term passed through the system  $(I - q^{-1} \hat{\mathcal{A}})^{-1}$ . Thus, by an extrapolation of the equation error concept for the scalar case [5], [7] we are led to define in a very natural way the block equation error as

$$\begin{aligned} \Xi(p) &\triangleq (I - q^{-1} \hat{\mathcal{A}})^{-1} \mathcal{E}(p) \\ &= \mathcal{Z}(p) - \hat{\mathcal{A}} \mathcal{Z}(p-1) + \hat{\mathcal{B}}^* \mathcal{U}(p) \end{aligned} \quad (17)$$

Also, defining a two-input/one-output non-recursive block process as

$$\hat{\mathcal{Z}}(p) \triangleq \hat{\mathcal{A}} \mathcal{Z}(p-1) + \hat{\mathcal{B}}^* \mathcal{U}(p) \quad (18)$$

then (17) can be expressed as

$$\Xi(p) = \mathcal{Z}(p) - \hat{\mathcal{Z}}(p) \quad (19)$$

From (17) and (19), we can realize the dual way in which the block equation error can be viewed, where a similar situation occurs for the scalar case [6]. On the one hand, the block equation error can be viewed as a linear transformation (i.e. filtering) of the block output error, as given by (17). On the other hand, it can be viewed as the block error between the process under identification  $\mathcal{Z}(p)$  and the artificially introduced nonrecursive process of (18).

### IV. A BLOCK EXTENDED LMS ALGORITHM

Let us define the following matrices of parameters

$$\hat{\mathcal{Q}} \triangleq \begin{bmatrix} \hat{\mathcal{A}}^* \\ \hat{\mathcal{B}}^* \end{bmatrix} \quad (20)$$

$$\hat{\mathcal{G}} \triangleq \begin{bmatrix} \hat{\mathcal{A}}^* \\ \hat{\mathcal{B}}^* \end{bmatrix} \quad (21)$$

where both have dimensions  $(2P+M) \times P$ , and define a  $(2P+M) \times 1$  information vector as

$$\mathcal{X}(p) \triangleq \begin{bmatrix} \mathcal{Z}(p-1) \\ \mathcal{U}(p) \end{bmatrix} \quad (22)$$

Hence (14) and (18) can be respectively rewritten in a general linear-in-the-parameters form as

$$\mathcal{L}(p) = \hat{\theta}^T \mathcal{Z}(p) \quad (23)$$

$$\hat{\mathcal{L}}(p) = \hat{\theta}^T \mathcal{Z}(p) \quad (24)$$

In an adaptive scheme, the estimated matrix of coefficients in (24) is adjusted once per block. Extrapolating from the scalar case, let us select an index of performance as a quadratic form of the block equation error, computed as the norm squared of (19), that is

$$J(p) = \bar{\mathcal{Z}}^T(p) \bar{\mathcal{Z}}(p) \\ = \mathcal{Z}^T(p) - \mathcal{Z}^T(p) \hat{\theta} | \mathcal{Z}(p) - \hat{\theta}^T \mathcal{Z}(p) | \\ = \mathcal{Z}^T(p) \mathcal{Z}(p) - \mathcal{Z}^T(p) \hat{\theta}^T \mathcal{Z}(p) - \mathcal{Z}^T(p) \hat{\theta} \mathcal{Z}(p) \\ + \hat{\theta}^T(p) \hat{\theta} \mathcal{Z}(p) \quad (25)$$

The equation above describes a time-varying quadratic hypersurface  $J: \mathcal{R}^{P(2P+M)} \rightarrow \mathcal{R}$ . Due to the structural match assumed for  $\mathcal{Z}(p)$  and  $\hat{\mathcal{L}}(p)$ , this hypersurface will have a global minimum at  $\hat{\theta} = \theta$ . Hence, an iterated gradient descent of this surface seems plausible. Computing the gradient of (25) with respect to  $\hat{\theta}$ , we obtain

$$\nabla_{\hat{\theta}} J(p) = -2 \mathcal{Z}(p) \mathcal{Z}^T(p) \\ + 2 \mathcal{Z}(p) \mathcal{Z}^T(p) \hat{\theta} \\ = -2 \mathcal{Z}(p) \bar{\mathcal{Z}}^T(p) \quad (26)$$

Thus, the gradient search for the global minimum of (25) is accomplished through the block extended LMS-type (or BELMS) algorithm

$$\hat{\theta}(p+1) = \hat{\theta}(p) + \mu \mathcal{Z}(p) \bar{\mathcal{Z}}^T(p) \quad (27)$$

where  $\mu$  can be selected as some small, fixed positive step-size for a strict sense steepest-descent, or as an appropriately chosen time-varying  $(2P+M) \times (2P+M)$  diagonal matrix, as an extrapolation of the scalar case [8, chps. 4, 5].

## V. CONVERGENCE ANALYSIS

In a deterministic scalar setting, Mendel [8] has presented comprehensive results for linear-in-the-parameters identification schemes, which is the structure into which equation error-based identifiers can be casted [4], [5]. For this case, a Lyapunov stability analysis can be performed on the resulting recursive algorithm used for parameter adaptation to show asymptotic convergence to the true parameters of the reference system. Fortunately, C.R. Johnson, Jr. [9] has extended Mendel's results to the linear-in-the-parameters multivariable case, which fits the scheme of (23), (24) and (27). Using the results of [9], the following can be shown [5]

**Theorem I:** If the estimated parameter matrix  $\hat{\theta}$  in (24) is updated by means of the algorithm of (27), then for any initial estimate  $\hat{\theta}(0)$  it will happen that

$$\hat{\theta}(p) \rightarrow \theta, \text{ as } p \rightarrow \infty \quad (28)$$

where  $\theta$  is as in (23), provided that

(i) the parameter error vector and the information vector are not orthogonal infinitely many times, of the form

$$[\theta - \hat{\theta}(p)]^T \mathcal{Z}(p) = 0, \forall p \geq p_1 \quad (29)$$

(ii) For  $\mu$  in (27) a scalar, we have

$$0 < \mu < \frac{2}{\mathcal{Z}(p)^T \mathcal{Z}(p)} \quad (30)$$

**Remark I:** (i) gives a nonorthogonality condition similar to those commonly encountered in scalar recursive identification. Such a situation is avoided by the use of a satisfactory (i.e. sufficiently exciting)  $\mathcal{Z}(p)$ . Clearly, a necessary condition is that no  $x_i(p)$ ,  $i \in \{1, 2, \dots, 2P+M\}$  is equal to zero  $\forall p \geq p_1$ .

**Remark II:** just as in the scalar case [8, pag. 208], we can select  $\mu = 2 [\mathcal{Z}(p)^T \mathcal{Z}(p)]^{-1}$  for optimum rate of convergence (in a Lyapunov sense), or as  $\mu = 2 [\mathcal{Z}(p)^T \mathcal{Z}(p)]^{-1}$  to avoid any possible singularity (e.g. see [10, pp. 57-58]).

## VI. COMPUTATIONAL COMPLEXITY ISSUES

By unraveling the matrix equation of (27) into its components and counting the number of arithmetic operations per iteration, it is found [5] that the ratios of multiplies  $\eta_m$  and adds  $\eta_a$  between the block and scalar equation error IIR filters are given by

$$\eta_m = \frac{C_1 E_1}{C_2 E_2} = \left( \frac{2P+M}{N+M+1} \right)^2 \quad (31)$$

$$\eta_a = \frac{C_1 (E_1 - 1)}{C_2 (E_2 - 1)} = \left( \frac{2P+M}{N+M+1} \right) \left( \frac{2P+M-1}{N+M} \right) \quad (32)$$

For the case when  $P=N$ , some numeric values are shown in the following table.

N	M	$\eta_m$	$\eta_a$
1	1	1.000	1.000
2	1	1.563	1.667
3	2	1.778	1.714
4	3	1.890	1.964
5	4	1.960	2.022
6	5	2.007	2.061
7	6	2.041	2.088
8	7	2.066	2.108
9	8	2.086	2.124
10	9	2.103	2.137

Table I: Complexity Ratios

## VII. SUMMARY

In this paper, a block adaptive IIR filter has been presented. The concept of equation error has been extended to the block formulation, as given by (16) and (17). This in turn was used to obtain the block adaptive estimation scheme of (23), (24) and (27). Global asymptotic convergence of this scheme can be shown using available analytical tools, for the deterministic case. For the stochastic case (i.e. noisy measurements of  $d(\cdot)$ ) the necessary tools for a detailed analysis do not seem to be available in the current literature, however bias is to be expected in the estimates, as for the scalar case [5], [8].

An increase in the computational load per iteration-per coefficient is found for the block implementation as compared to the scalar one. Also, preliminary computer simulations have shown a slight decrease in convergence rate for the former as compared to the latter. However, the possibility of parallel implementations of the block formulation make this approach promising. Of course, further studies must be carried out in this respect.

## REFERENCES

- [1] C.S. Burrus, "Block Implementation of Digital Filters", IEEE Trans. on Circuit Theory, Vol. CT-18, No. 6, pp. 697-701, 1971.
- [2] S.K. Mitra and R. Gnanasekaran, "Block Implementation of Recursive Digital Filters — New Structures and Properties", IEEE Trans. on Circuits and Systems, Vol. CAS-25, No. 9, pp. 200-207, 1978.
- [3] G.A. Clark, S.K. Mitra and S.D. Parker, "Block Implementation of Adaptive Digital Filters", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-29, No. 3, pp. 744-752.
- [4] E. Fernández-Gaucherand and J.Y. Cheung, "A Suboptimal Solution to IIR Adaptive Filtering Using an Extended LMS Algorithm", Proc. 22nd Annual Allerton Conf. Comm., Control, and Computing, Univ. Illinois at Urbana-Champaign, pp. 132-191, 1984.
- [5] E. Fernández-Gaucherand, "Adaptive IIR Filtering: Current Results and the Equation Error Approach", MSc. Thesis, School of Elect. Eng. and Comp. Sci., The Univ. of Oklahoma, 1985.
- [6] E. Fernández-Gaucherand and J.Y. Cheung, "On the Similarities Between the Equation Error and the Output Error Approaches in Adaptive IIR Filtering", submitted for publication, 1986.
- [7] K.J. Åström and P. Eykhoff, "Systems Identification - A Survey", Automatica, Vol. 7, pp. 123-162, 1971.
- [8] J.M. Mendel, Discrete Techniques of Parameter Estimation: The Equation Error Formulation, Marcel Dekker, New York, 1973.
- [9] C.R. Johnson, "Adaptive Parameter Matrix and Output Vector Estimation Via an Equation Error Formulation", IEEE Trans. Systems, Man and Cybernetics, Vol. SMC-9, No. 7, pp. 392-397, 1979.
- [10] G.C. Goodwin and K.S. Sin, Adaptive Filtering, Prediction, and Control, Prentice-Hall, Englewood Cliffs, NJ, 1984.