term in (A.12) be negligible compared with (35), we must further require that $(T/B^2)e^{-\beta_G/B}$ approaches zero. This rather unusual requirement simply assures that the modeling errors (bias squared) be negligible compared with the error variance.

The derivation of the bound in (30) follows similarly by slightly modifying the definition of $\boldsymbol{D}$.

References

[1] S. Bellini and F. Rocca, "Asymptotically efficient blind deconvolution," *Signal Process.*, vol 20, no. 1, pp. 193–209, 1990.

[2] D. R. Brillinger, *Time Series, Data Analysis and Theory*. San Francisco, CA: Holden Day, 1981.

[3] J. Cardoso, "On the performance of orthogonal source separation algorithms," in *Proc. EUSIPCO*, Edinburgh, U.K., Sept. 1994.

[4] P. Comon, "Independent component analysis, a new concept?," *Signal Process.*, vol. 36, no. 3, Apr. 1994.

[5] ——, "Contrasts for multichannel blind deconvolution," *IEEE Signal Processing Lett.*, vol. 3, pp. 209–211, July 1996.

[6] P. Comon, C. Jutten, and J. Hérault, "Blind separation of sources, Part II: Problems statement," *Signal Process.*, vol 24, no. 1, pp. 11–20, July 1991.

[7] A. Gorokhov and J. Cardoso, "Equivariant blind deconvolution of MIMO channels," in *Proc. SPAWC*, Paris, France, Apr. 1997.

[8] C. Jutten and J. Hérault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal Process.*, vol 24, no. 1, pp. 1–10, July 1991.

[9] J. L. Lacoume and P. Ruiz, "Separation of independent sources from correlated inputs," *IEEE Trans. Signal Processing*, vol. 40, pp. 3074–3078, Dec. 1992.

[10] L. Ljung, *System identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[11] E. Moreaum and O. Macchi, "High order contrasts for self-adaptive source separation," *Int. J. Adaptive Contr. Signal Process.*, to be published.

[12] D. T. Pham, "Blind separation of instantaneous mixture of sources via an independent component analysis," *IEEE Trans. Signal Processing*, vol. 44, pp. 2768–2770, Nov. 1996.

[13] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[14] D. Sengupta and S. Kay, "Efficient estimation of parameters for non-Gaussian autoregressive processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 785–794, June 1989.

[15] O. Shalvi and E. Weinstein, "Maximum likelihood and lower bounds in system identification with non-Gaussian inputs," *IEEE Trans. Inform. Theory*, vol. 40, pp. 328–339, Mar. 1994.

[16] H. L. Van-Trees, *Detection, estimation, and Modulation Theory*, Part I. New York: Wiley, 1968.

[17] E. Weinstein, "Estimation of trajectory parameters from passive array measurements," Ph.D. dissertation, Yale Univ., New Haven, CT, 1978.

[18] E. Weinstein, M. Feder, and A. V. Oppenheim, "Multi-channel signal separation by decorrelation," *IEEE Trans. Speech Audio Processing*, vol. 1, pp. 405–413, Oct. 1993.

[19] D. Yellin and B. Friedlander, "Blind multi-channel identification and deconvolution: performance bounds," in *Proc. 8th SSAP Workshop*, Greece, 1996.

[20] D. Yellin and E. Weinstein, "Multi-channel signal separation: Methods and analysis," *IEEE Trans. Signal Processing*, vol. 44, pp. 106–118, Jan. 1996.

[21] D. Yellin and B. Friedlander, "Multi-channel system identification and deconvolution: Performance bounds," Tech. Rep., Dept. Elect. Comput. Eng., Univ. California, Davis, Oct. 1996.

A Modified QRD for Smoothing and a QRD-LSL Smoothing Algorithm

Jenq-Tay Yuan

***Abstract*—This paper introduces a *modified QR-decomposition* (QRD) that extends the method of QRD to a more general case to solve the least-squares lattice smoothing problems. We show that the conventional QRD is a special form of the modified QRD that occurs when no *future* data values are used. Within the framework of the modified QRD procedure, an *order-recursive* QRD-based least-squares lattice (QRD-LSL) smoothing algorithm is formulated. The algorithm combines all the desirable features of the standard QRD-LSL filtering algorithm with a more accurate smoothing process. The results of some computer simulations of a channel equalizer are also presented.**

***Index Terms*— Least-squares lattice, QR-decomposition, QRD-LSL, smoothing.**

I. Introduction

One of the most important families of fast algorithms in *order-recursive* adaptive filtering are those of the conventional recursive *least-squares lattice* (LSL) algorithms [1]–[5], [16]. Many of these fast algorithms tend to suffer from some form of numerical instability due to finite-precision effects [4], [10]. The *QR-decomposition* (QRD) technique is, in general, well-conditioned and numerically stable [4], [6], [8]. Furthermore, a useful property of the QRD technique is that, upon solving an Nth-order filtering problem, the solutions to all lower order problems are obtained as a byproduct [9], [10]. This property is also found in the least-squares (LS) *lattice filters.* Accordingly, it is possible to solve the LSL problems using the QRD technique [4], [10], [12]. Indeed, the LSL algorithm for adaptive filtering based on the QRD is endowed with a highly desirable set of features that include robustness to roundoff error, superior numerical properties, modularity, and a high level of computational efficiency [4]. Extensive computer simulations have shown that the QRD-LSL algorithm has excellent numerical properties [7], [12], [13].

Smoothing differs from filtering in that not only the "subspace of past and present observations" but the "subspace of *future* observations" are taken into account in estimating the present desired signal. The smoothing process is known to be more accurate than the filtering process since the former is more "complete" than the latter in terms of the available information at a certain time step [14]. A smoother can be realized by introducing a suitable delay into any filter that makes the filter "noncausal" in the sense that a linear combination of the present, past, and future observations can be used to estimate the present desired signal. However, once delay is introduced into a QRD-LSL filter, the order-recursive property no longer holds. Higher order noncausal filters, or smoothers, cannot be built from lower order ones simply by adding more lattice stages as more "*future*" observations are used to estimate the present desired signal. An orthogonal basis theorem was thus developed in [5] for the design of *order-recursive LSL smoothers.* However, the recursive LSL smoothing algorithm derived in [5] requiring the inversion of

Manuscript received June 5, 1997; revised September 29, 1998. This work was supported by the National Science Council, R.O.C., under Contract NSC 87-2213-E-030-004. The associate editor coordinating the review of this paper and approving it for publication was Dr. Sergios Theodoridis.

The author is with the Department of Electronic Engineering, Fu Jen Catholic University, Taipei, Taiwan, R.O.C. (e-mail: yuan@mars.ee.fju.edu.tw).

Publisher Item Identifier S 1053-587X(99)03253-5.

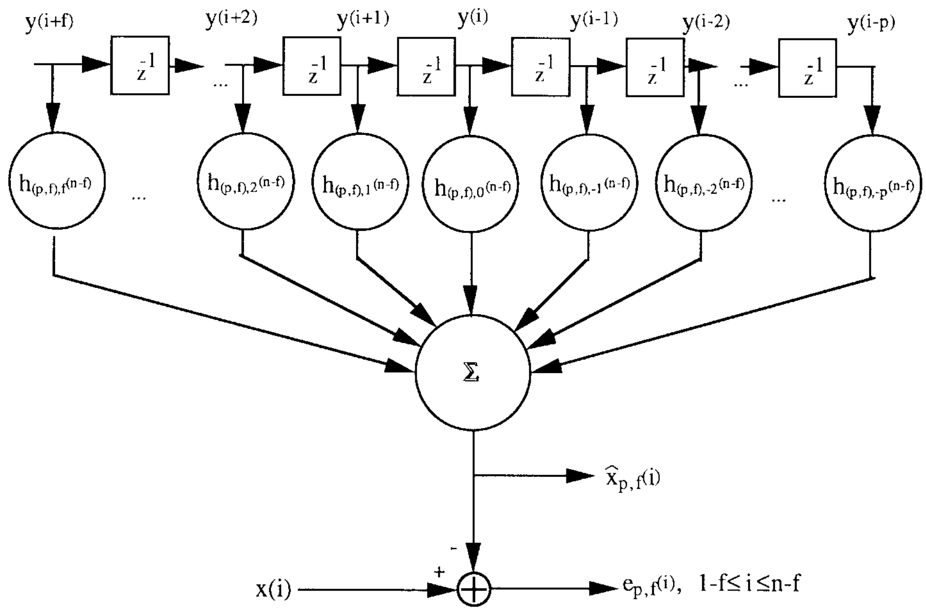


Fig. 1. Direct-form realization of the Nth or (p, f)th order FIR least-squares smoother.

the correlation matrix of the input data may suffer from some form of numerical instability and inaccuracy due to its numerical sensitivity to limited arithmetic precision. This paper extends the well-conditioned and numerically stable QRD-LSL filtering algorithm [4], [12] to the QRD-LSL smoothing algorithm while retaining all the desirable features of the former. The QRD-LSL smoothing algorithm is order recursive in the sense that we can always increase the smoother order by adding more "past" stages as well as more "future" stages, whereas the "old" part of the smoother still remains optimal. This property is particularly useful when there is no prior knowledge as to what the final value of the smoother order should be. The total computational complexity of the proposed algorithm scales linearly with the smoothing order N.

II. FIR Smoothers and Orthogonal Basis Sets

Consider the direct-form realization of an Nth-order FIR least-squares smoother shown in Fig. 1. The desired sequence $x(i)$ is estimated from its current, p past, and f "future" observations $y(i)$ (the data sequence) for $i = 1, 2, \cdots, n$. The length of the observations n is variable. The order $N = p + f$. We will refer to any Nth-order smoother that uses p past and f future data values as a (p, f)th-order smoother, where $N = p + f$ is assumed implicitly. The estimation error is

$$\begin{aligned} e_{\boldsymbol{p},\boldsymbol{f}}(i) &= x(i) - \hat{x}_{\boldsymbol{p},\boldsymbol{f}}(i) \\ &= x(i) - \mathbf{h}^{\boldsymbol{T}}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f})\mathbf{y}_{N+1}(\mathbf{i}+\mathbf{f}) \\ &\qquad 1 - f \le i \le n - f \end{aligned} \tag{1}$$

where

$$\mathbf{y}^{\boldsymbol{T}}_{N+1}(\mathbf{i}+\mathbf{f}) = [y(i+f), y(i+f-1), \cdots, y(i-p)] \tag{2}$$

and

$$\begin{aligned} \mathbf{h}^{\boldsymbol{T}}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f}) = [&h_{(p,f),f}(n-f), \cdots, h_{(p,f),1}(n-f) \\ &h_{(p,f),0}(n-f), h_{(p,f),-1}(n-f), \cdots \\ &h_{(p,f),-p}(n-f)]. \end{aligned} \tag{3}$$

The vector $\mathbf{h}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f})$ contains the fixed coefficients of the (p, f)th-order FIR smoother and will be chosen, at time $n-f$, for least-squares estimation error over the time interval $1 - f \le i \le n - f$ with the prewindowing condition on the data, that is, $y(i) = 0$ for $i \le 0$. Equation (1) can be written in matrix form as

$$\mathbf{e}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f}) = \mathbf{x}(\mathbf{n}-\mathbf{f}) - \mathbf{Y}_{N+1}(\mathbf{n})\mathbf{h}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f}) \tag{4}$$

where

$$\mathbf{x}^{\boldsymbol{T}}(\mathbf{n}-\mathbf{f}) = [x(1-f), x(2-f), \cdots, x(n-f)] \tag{5}$$

$$\mathbf{e}^{\boldsymbol{T}}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f}) = [e_{\boldsymbol{p},\boldsymbol{f}}(1-f), e_{\boldsymbol{p},\boldsymbol{f}}(2-f), \cdots, e_{p,f}(n-f)] \tag{6}$$

$$\mathbf{Y}_{N+1}(\mathbf{n}) = \begin{bmatrix} y(1) & 0 & & 0 \\ y(2) & y(1) & \cdots & 0 \\ y(3) & y(2) & & 0 \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ y(n) & y(n-1) & & -y(n-N) \end{bmatrix}. \tag{7}$$

We use $\{\ \}$ to represent each row of elements of matrix $\mathbf{Y}_{\boldsymbol{N}+\mathbf{1}}(\mathbf{n})$ as the linear subspace spanned by the columns of matrix $\mathbf{Y}_{\boldsymbol{N}+\mathbf{1}}(\mathbf{n})$. For example, the bottom row of elements of matrix $\mathbf{Y}_{\boldsymbol{N}+\mathbf{1}}(\mathbf{n})$ constitutes the linear subspace $\mathbf{Y}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f}) = \{y(n-N), \cdots, y(n-1), y(n)\}$, which may be viewed as the subspace of the current observation [i.e., $y(n-f)$], p past observations [i.e., $y(n-f-1), \cdots, y(n-N)$], and f future observations [i.e., $y(n-f+1), \cdots, y(n)$]. The optimum coefficients in (3) can be chosen by minimizing the squared Euclidean norm of the error vector $\|\Lambda^{1/2}(\mathbf{n})\mathbf{e}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f})\|^2$, where $\Lambda(\mathbf{n}) = \text{diag}[\lambda^{n-1}, \lambda^{n-2}, \cdots, 1]$ is the $n \times n$ exponential weighting matrix in which $0 \ll \lambda \le 1$.

Since a smoother uses both past and future observations to estimate the desired sequence, we need to consider the problem of increasing order $N \to N+1$ by increasing either f or p by one when developing an order-update recursion for the estimation error. This order-update recursion can be accomplished by embedding an $(N+1)$st-order prediction lattice into a LSL smoother of order $(p+1, f)$ or $(p, f+1)$. It was first shown in [5] that when a sequence of p past and f future observations are considered, appropriate combinations of f delayed forward prediction errors and p delayed backward prediction errors form $C^f_N = N!/[(N-f)!f!]$ sets of orthogonal

TABLE I
SUMMARY OF THE QRD-LSL SMOOTHING ALGORITHM

1. Computations

(I) Predictions: For time n = 1, 2,..., and prediction order m = 1, 2,..., M, where M is the final prediction order [4].

$B_{m-1}(n-1) = \lambda B_{m-1}(n-2) + (e^B_{m-1}(n-1))^2$, where $e^B_i(n)$ is the backward prediction error of order i.

$$c_{b,m-1}(n-1) = \frac{\lambda^{1/2}(B_{m-1}(n-2))^{1/2}}{(B_{m-1}(n-1))^{1/2}}$$

$$s_{b,m-1}(n-1) = \frac{e^B_{m-1}(n-1)}{(B_{m-1}(n-1))^{1/2}}$$

$$e^F_m(n) = c_{b,m-1}(n-1)e^F_{m-1}(n) - \lambda^{1/2}s_{b,m-1}(n-1)\pi^F_{m-1}(n-1)$$

$$\pi^F_{m-1}(n) = \lambda^{1/2}c_{b,m-1}(n-1)\pi^F_{m-1}(n-1) + s_{b,m-1}(n-1)e^F_{m-1}(n)$$

$F_{m-1}(n) = \lambda F_{m-1}(n-1) + (e^F_{m-1}(n))^2$, where $e^F_i(n)$ is the forward prediction error of order i.

$$c_{f,m-1}(n) = \frac{\lambda^{1/2}(F_{m-1}(n-1))^{1/2}}{(F_{m-1}(n))^{1/2}}$$

$$s_{f,m-1}(n) = \frac{e^F_{m-1}(n)}{(F_{m-1}(n))^{1/2}}$$

$$e^B_m(n) = c_{f,m-1}(n)e^B_{m-1}(n-1) - \lambda^{1/2}s_{f,m-1}(n)\pi^B_{m-1}(n-1)$$

$$\pi^B_{m-1}(n) = \lambda^{1/2}c_{f,m-1}(n)\pi^B_{m-1}(n-1) + s_{f,m-1}(n)e^B_{m-1}(n-1)$$

(II) Smoothing: For time n = 1, 2,..., start from f = 0 and p = −1. N = p + f. Additional p past and f future stages can be increased by computing any of C^f_N combinations of part (a) and part (b).

(a) One additional future observation is taken into account:

$$F_{N+1}(n) = \lambda F_{N+1}(n-1) + (e^F_{N+1}(n))^2 \quad (21)$$

$$c_{f,N+1}(n) = \frac{\lambda^{1/2}F^{1/2}_{N+1}(n-1)}{F^{1/2}_{N+1}(n)} \quad (19)$$

$$s_{f,N+1}(n) = \frac{e^F_{N+1}(n)}{F^{1/2}_{N+1}(n)} \quad (20)$$

$$\varepsilon_{p,f+1}(n-f-1) = c_{f,N+1}(n)\varepsilon_{p,f}(n-f-1) - \lambda^{1/2}s_{f,N+1}(n)\rho^F_{p,f+1}(n-1) \quad (22)$$

bases. The orthogonality among all the elements within each of these orthogonal bases has been referred to as the orthogonal basis theorem (see [5]).

III. A MODIFIED QR-DECOMPOSITION

In this section, we modify the well-known QR-decomposition based on the orthogonal basis theorem so that it will be suited to implement the order-recursive QRD-LSL smoothers. It can be shown that an $n \times n$ orthogonal matrix $\overline{\mathbf{Q}(\mathbf{n})}$ can always be constructed from one of the C^f_N orthonormal basis sets, each of which provides an orthonormal basis for the linear subspace $\mathbf{Y}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n}-\mathbf{f})$ such that [15] (due to space limitation, details are omitted here)

$$\overline{\mathbf{Q}(\mathbf{n})}\mathbf{Y}_{N+1}(\mathbf{n}) = \begin{bmatrix}\mathbf{Q}^{\boldsymbol{T}}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n})\\ \mathbf{S}^{\boldsymbol{T}}(\mathbf{n})\end{bmatrix}\mathbf{Y}_{N+1}(\mathbf{n}) = \begin{bmatrix}\mathbf{R}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n})\\ \mathbf{O}\end{bmatrix} \quad (8)$$

where $\mathbf{Q}^{\boldsymbol{T}}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n})$ contains the first $(N+1)$ rows of $\overline{\mathbf{Q}(\mathbf{n})}$, whereas $\mathbf{S}^{\boldsymbol{T}}(\mathbf{n})$ contains the remaining rows. Since C^f_N possible sequences can be used, the matrix $\mathbf{R}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n})$ in (8) can display C^f_N different forms, all of which, however, contain one $(f+1)\times(f+1)$ left-hand lower triangular matrix and one $(p+1)\times(p+1)$ right-hand upper triangular matrix. We indicate, in particular, that $\mathbf{R}_{\boldsymbol{p},\boldsymbol{f}_{\mathbf{BFBFBF}\cdots}}(\mathbf{n})$ using the BFBFBF... sequence with $p=f$, can be as in (9), shown on the next page [15], where $F^{1/2}_M(n-j)$ and $B^{1/2}_M(n-j)$ are the square roots of the minimum sum of the mth-order forward and backward prediction error squares, respectively, and the symbol x in (9) denotes either a zero or a nonzero element whose value is not of direct interest. In this correspondence, we consider only the BFBFBF... sequence [5], among all C^f_N permissible sequences, for the implementation of a QRD-LSL smoother of an arbitrary-order (p,f). For example, for a (7, 4)th-order smoother, we only consider the BFBFBFBFBBB sequence. Other sequences can be similarly obtained. Notice that unlike the upper triangular form

TABLE I
SUMMARY OF THE QRD-LSL SMOOTHING ALGORITHM (*Continued.*)

$$\rho^{F}_{p,f+1}(n) = \lambda^{1/2} c_{f,N+1}(n)\rho^{F}_{p,f+1}(n-1) + s_{f,N+1}(n)\varepsilon_{p,f}(n-f-1) \tag{23}$$

$$\gamma^{1/2}_{p,f+1}(n) = c_{f,N+1}(n)\gamma^{1/2}_{p,f}(n), \tag{25}$$

where $\gamma^{1/2}_{p,f}(n)$ is the square root of the likelihood variable.

$$e_{p,f+1}(n-f-1) = \gamma^{1/2}_{p,f+1}(n)\varepsilon_{p,f+1}(n-f-1), \tag{24}$$

where $e_{p,f+1}(n-f-1)$ is the a posteriori smoothing error.

(b) One additional past observation is taken into account:

$$B_{N+1}(n) = \lambda B_{N+1}(n-1) + \left(e^{B}_{N+1}(n)\right)^2 \tag{26}$$

$$c_{b,N+1}(n) = \frac{\lambda^{1/2} B^{1/2}_{N+1}(n-1)}{B^{1/2}_{N+1}(n)} \tag{27}$$

$$s_{b,N+1}(n) = \frac{e^{B}_{N+1}(n)}{B^{1/2}_{N+1}(n)} \tag{28}$$

$$\varepsilon_{p+1,f}(n-f) = c_{b,N+1}(n)\varepsilon_{p,f}(n-f) - \lambda^{1/2} s_{b,N+1}(n)\rho^{B}_{p+1,f}(n-1) \tag{29}$$

$$\rho^{B}_{p+1,f}(n) = \lambda^{1/2} c_{b,N+1}(n)\rho^{B}_{p+1,f}(n-1) + s_{b,N+1}(n)\varepsilon_{p,f}(n-f) \tag{30}$$

$$\gamma^{1/2}_{p+1,f}(n) = c_{b,N+1}(n)\gamma^{1/2}_{p,f}(n) \tag{31}$$

$$e_{p+1,f}(n-f) = \gamma^{1/2}_{p+1,f}(n)\varepsilon_{p+1,f}(n-f) \tag{32}$$

2. Initialization

(c) Auxiliary parameter initialization: For order m = 1, 2,...,M, set

$\pi^{F}_{m-1}(0) = \pi^{B}_{m-1}(0) = 0$

$\rho^{B}_{p,f}(0) = \rho^{F}_{p,f}(0) = 0$, for all p and f

(d) Soft constraint initialization: For order m = 0, 1,...,M, set

$F_m(0) = B_m(-1) = \delta$, where δ is a small positive constant.

(e) Data initialization: For n = 1, 2,..., set

$e^{F}_0(n) = e^{B}_0(n) = y(n)$

$\varepsilon_{-1,0}(n) = x(n)$, where y(n) is the input and x(n) is the desired response at time n.

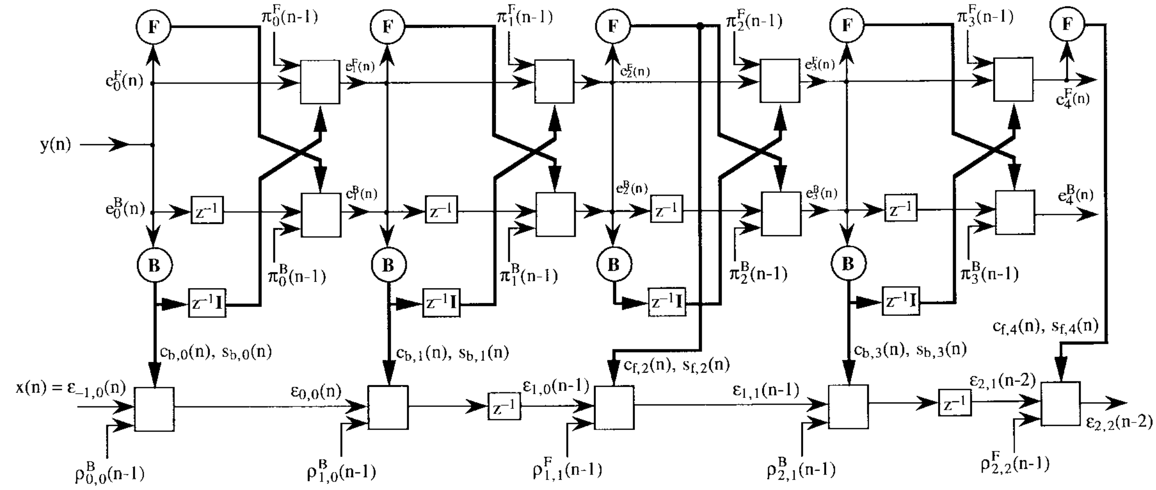
$\varepsilon_{p,f}(n) = 0$ for $n \le 0$ and all p and f

$\gamma_{-1,0}(n) = 1$

we have in the conventional QRD in which the diagonal elements give the square roots of backward prediction error energies of all orders owing to the use of current and past observations only to estimate the present desired signal [4], [10]. In $\mathbf{R}_{\boldsymbol{p},\boldsymbol{f}_{\mathrm{BFBFBF}\cdots}}(\mathbf{n})$, however, the first f diagonal elements give the square roots of forward prediction error energies corresponding to the f future observations $y(n), \cdots, y(n-f+1)$, and the remaining $(p+1)$ diagonal elements give the square roots of backward prediction error energies corresponding to the current observation $y(n-f)$ and the p past observations $y(n-f-1), \cdots, y(n-N)$. We will refer to the result in (8) as the *modified QRD for smoothing* and refer to the form of the lower and upper triangle matrices in $\mathbf{R}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n})$ as the *LU triangular form.* The modified QR-decomposition provides a theoretical basis for the development of the QRD-LSL smoothers (see Section IV and [11]). For the special case in which $(p, f) = (N, 0)$ (i.e., none of the future observations is considered), each column of $\mathbf{Q}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n})$ can be shown to be connected with backward prediction errors. Therefore, the LU triangular form in the matrix $\mathbf{R}_{\boldsymbol{p},\boldsymbol{f}}(\mathbf{n})$ in

$$\mathbf{R}_{\mathbf{p},\mathbf{f}_{\mathrm{BFBFBF}\ldots}}(\mathbf{n}) = \begin{bmatrix} F_N^{1/2}(n) & 0 & \cdots & 0 & \times & \cdots & \times \\ \times & F_4^{1/2}(n-f+2) & \ddots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & F_2^{1/2}(n-f+1) & 0 & \times & \cdots & \times \\ \times & \cdots & \times & B_0^{1/2}(n-f) & \times & \cdots & \times \\ \times & \cdots & \times & 0 & B_1^{1/2}(n-f) & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \ddots & B_3^{1/2}(n-f+1) & \times \\ \times & \cdots & \times & 0 & \cdots & 0 & B_{N-1}^{1/2}(n-1) \end{bmatrix} \tag{9}$$

(8) is reduced to matrix $\mathbf{R}_{\boldsymbol{N},0}(\mathbf{n})$, which contains the $C_N^0 = 1$ unique $(N+1) \times (N+1)$ upper triangular matrix whose diagonal elements give the square roots of backward prediction error energies of all orders. This is exactly the case for the conventional QRD [4], [10].



(F): The computation of the **forward** cosine-sine pair of rotation parameters $c_{f,m}(n)$ and $s_{f,m}(n)$

(B): The computation of the **backward** cosine-sine pair of rotation parameters $c_{b,m}(n)$ and $s_{b,m}(n)$

Fig. 2. Signal-flow graph of the (2, 2)th-order QRD-LSL smoother using the sequence BFBF.

IV. Order-Recursive QRD-LSL Smoothers

To develop the order-recursive QRD-LSL smoothers, we suppose that at time $n-1$, the $(n-1)\times(N+1)$ weighted data matrix $\Lambda^{1/2}(\mathbf{n}-1)\mathbf{Y}_{N+1}(\mathbf{n}-1)$ [see (7)] has already been reduced to the LU triangular form by the $(n-1)\times(n-1)$ orthogonal matrix $\overline{\mathbf{Q}(\mathbf{n}-1)}$. We thus have

$$\overline{\mathbf{Q}(\mathbf{n}-1)}\Lambda^{1/2}(\mathbf{n}-1)\mathbf{Y}_{N+1}(\mathbf{n}-1)=\begin{bmatrix}\mathbf{R}_{p,f}(\mathbf{n}-1)\\ \mathbf{O}\end{bmatrix} \tag{10}$$

where $\mathbf{R}_{p,f}(\mathbf{n}-1)$ is an LU triangular matrix, and $\mathbf{O}$ is the $(n-N-2)\times(N+1)$ null matrix. With the solution at time $(n-1)$ and the new observations for time n, we have (11), shown at the bottom of the page. To complete the LU triangularization, we define the orthogonal matrix $\hat{\mathbf{Q}}_{p,f}(\mathbf{n})$ that is used to annihilate the new observations at the bottom row of $\mathbf{R}'_{p,f}(\mathbf{n})$ such that

$$\hat{\mathbf{Q}}_{p,f}(\mathbf{n})\mathbf{R}'_{p,f}(\mathbf{n})=\begin{bmatrix}\mathbf{R}_{p,f}(\mathbf{n})\\ \mathbf{O}\\ \mathbf{o}^T\end{bmatrix} \tag{12}$$

where the $(N+1)\times(N+1)$ matrix $\mathbf{R}_{p,f}(n)$ corresponds to a complete LU triangular portion of the $n\times(N+1)$ weighted data matrix $\Lambda^{1/2}(\mathbf{n})\mathbf{Y}_{N+1}(\mathbf{n})$. The diagonal matrix $\hat{\mathbf{Q}}_{p,f}(\mathbf{n})$ can be formed as the product of $(N+1)$. Givens rotations as $\hat{\mathbf{Q}}_{p,f}(\mathbf{n})=\mathbf{Q}_{0,0}(\mathbf{n})\mathbf{Q}_{1,0}(\mathbf{n})\mathbf{Q}_{1,1}(\mathbf{n})\mathbf{Q}_{2,1}(\mathbf{n})\mathbf{Q}_{2,2}(\mathbf{n})\cdots\mathbf{Q}_{p,f}(\mathbf{n})$ for the BFBFBF... sequence. Note that unlike the conventional QRD, the sequences of Givens rotations are applied by first annihilating $y(n)$ in $\mathbf{R}'_{p,f}(\mathbf{n})$ in (11) and then successively annihilating the resulting bottom row by moving to the right in N steps. In the modified QRD, however, the current observation $y(n-f)$ is first annihilated by $\mathbf{Q}_{0,0}(\mathbf{n})$, and then, the annihilations proceed bidirectionally leftwards (for an "F") and rightwards (for a "B") in accordance with a particular sequence chosen (e.g., BFBFBF... in this case) until the leftmost observation $y(n)$ and the rightmost observation $y(n-N)$ are both annihilated to transform $\mathbf{R}'_{p,f}(\mathbf{n})$ into the LU triangular matrix, as shown in (12). We thus can obtain a recursion that relates the updated value of the $n\times n$ orthogonal matrix $\overline{\mathbf{Q}(\mathbf{n})}$ to the old value of the orthogonal matrix $\overline{\mathbf{Q}(\mathbf{n}-1)}$

$$\overline{\mathbf{Q}(\mathbf{n})}=\hat{\mathbf{Q}}_{p,f}(\mathbf{n})\begin{bmatrix}\overline{\mathbf{Q}(\mathbf{n}-1)} & \mathbf{o}\\ \mathbf{o}^T & 1\end{bmatrix}. \tag{13}$$

To obtain the optimum value of $\mathbf{h}_{p,f}(\mathbf{n}-\mathbf{f})$, we use $\overline{\mathbf{Q}(\mathbf{n})}$ to rotate (4) into [11]

$$\begin{bmatrix}\mathbf{Q}^T_{p,f}(\mathbf{n})\\ \mathbf{S}^T(\mathbf{n})\end{bmatrix}\Lambda^{1/2}(\mathbf{n})\mathbf{e}_{p,f}(\mathbf{n}-\mathbf{f}) =\begin{bmatrix}\mathbf{p}_{p,f}(\mathbf{n}-\mathbf{f})-\mathbf{R}_{p,f}(\mathbf{n})\mathbf{h}_{p,f}(\mathbf{n}-\mathbf{f})\\ \boldsymbol{v}_{p,f}(\mathbf{n}-\mathbf{f})\end{bmatrix}. \tag{14}$$

The optimal weight vector containing the optimum coefficients of the (p,f)th-order FIR smoother can be obtained from

$$\mathbf{p}_{p,f}(\mathbf{n}-\mathbf{f})=\mathbf{R}_{p,f}(\mathbf{n})\mathbf{h}_{p,f}(\mathbf{n}-\mathbf{f}). \tag{15}$$

Note that since $\mathbf{R}_{p,f}(\mathbf{n})$ is LU triangular, the optimum coefficients may readily be solved by back substitution. This choice of $\mathbf{h}_{p,f}(\mathbf{n}-\mathbf{f})$ yields a minimum value of the sum of weighted error squares

$$E_{p,f_{\min}}(n)=\min_h \|\Lambda^{1/2}(\mathbf{n})\mathbf{e}_{p,f}(\mathbf{n}-\mathbf{f})\|^2=\|\mathbf{v}_{p,f}(\mathbf{n}-\mathbf{f})\|^2. \tag{16}$$

$$\begin{bmatrix}\overline{\mathbf{Q}(\mathbf{n}-1)} & \mathbf{o}\\ \mathbf{o}^T & 1\end{bmatrix}\Lambda^{1/2}(\mathbf{n})\mathbf{Y}_{N+1}(\mathbf{n})=\mathbf{R}'_{p,f}(\mathbf{n})=\begin{bmatrix}\lambda^{1/2}\mathbf{R}_{p,f}(\mathbf{n}-1)\\ \mathbf{O}\\ y(n),\, y(n-1),\,\cdots,\, y(n-f+1),\, y(n-f),\, y(n-f-1),\,\cdots,\, y(n-N)\end{bmatrix} \tag{11}$$

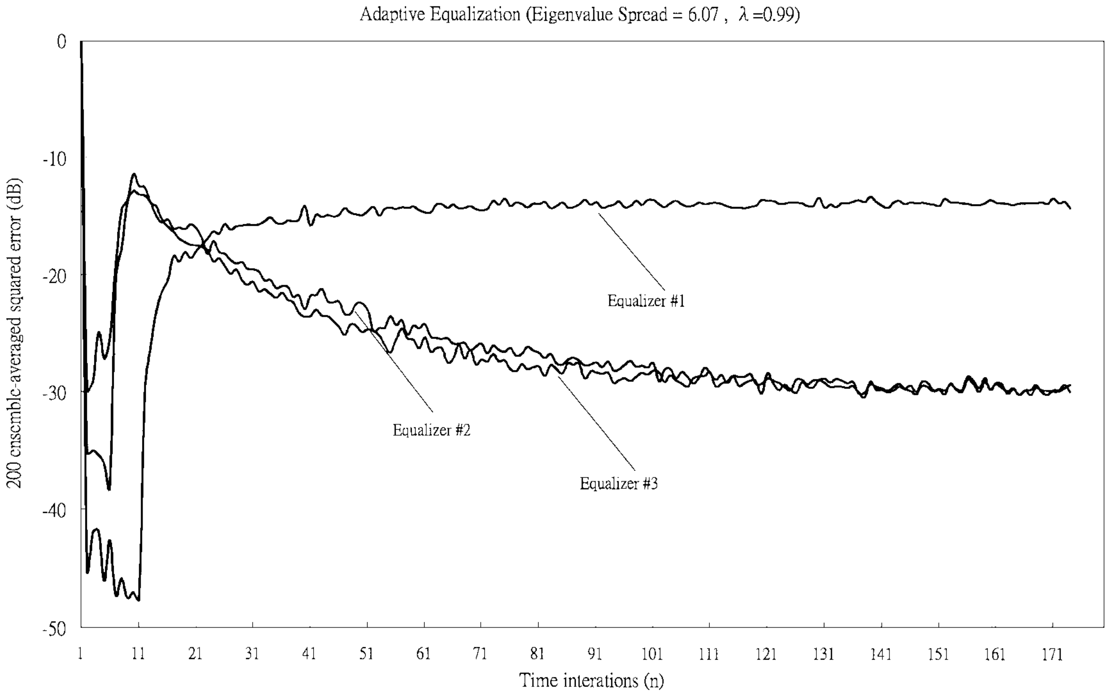


Fig. 3. Learning curves for the three equalizers (eigenvalue spread = 6.07).

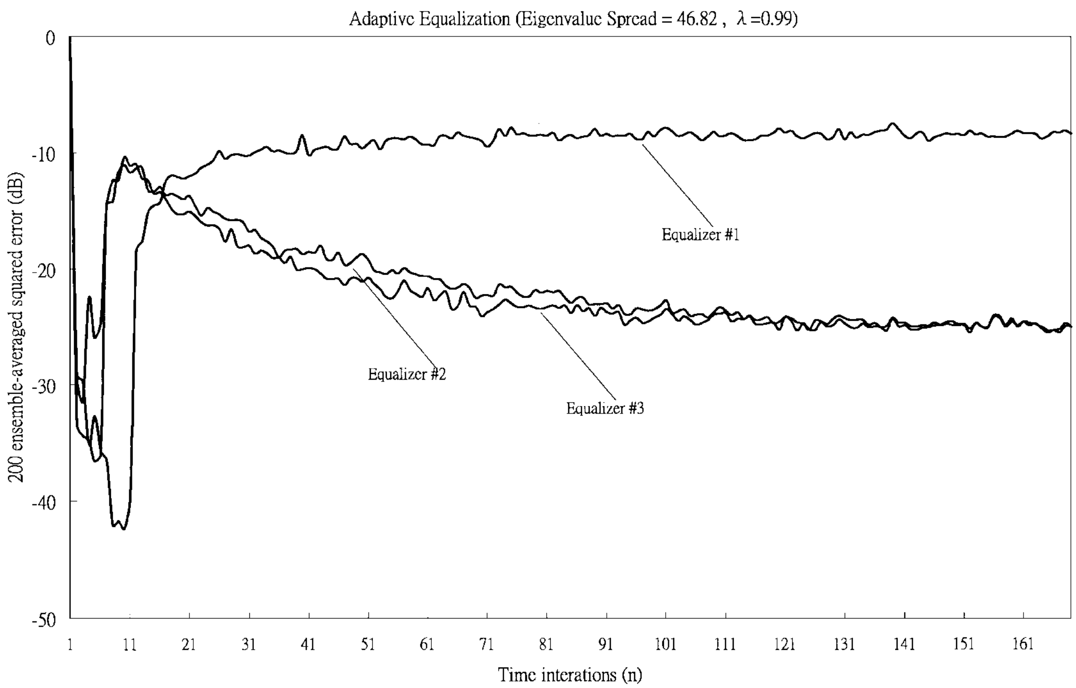


Fig. 4. Learning curves for the three equalizers (eigenvalue spread = 46.82).

Both vectors $\mathbf{p}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f})$ and $\mathbf{v}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f})$ in (14) can be computed recursively by using the same diagonal matrix $\hat{\mathbf{Q}}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n})$

$$\begin{bmatrix}\mathbf{p}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f})\\ \mathbf{v}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f})\end{bmatrix} = \hat{\mathbf{Q}}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n})\begin{bmatrix}\lambda^{1/2}\mathbf{p}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f}-1)\\ \lambda^{1/2}\mathbf{v}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f}-1)\\ x(n-f)\end{bmatrix} = \begin{bmatrix}\mathbf{p}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f})\\ \lambda^{1/2}\mathbf{v}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f}-1)\\ \varepsilon_{p,\,f}(n-f)\end{bmatrix} \tag{17}$$

where $\varepsilon_{\boldsymbol{p},\,\boldsymbol{f}}(n-f)$ can be referred to as the *angle-normalized joint-process estimation error* of $x(n-f)$ by using the current observation, f future observations, and p past observations. The elements of vector $\mathbf{p}^{\boldsymbol{T}}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n}-\mathbf{f}) = [\rho^F_{\boldsymbol{p},\,\boldsymbol{f}}(n),\,\cdots,\,\rho^F_{1,\,1}(n-f+1),\,\rho^B_{0,\,0}(n-f),\,\rho^F_{1,\,0}(n-f),\,\cdots,\,\rho^F_{p,\,f-1}(n-1)]$ are referred to as the joint-process auxiliary parameters and play a key role in implementing a $(p,\,f)$th-order-recursive QRD-LSL smoother. From (14), the auxiliary parameters can be shown to be the coefficients of the projection of the desired signal vector $\mathbf{x}(\mathbf{n}-\mathbf{f})$ onto the orthonormal basis set provided by the rows of matrix $\mathbf{Q}^{\boldsymbol{T}}_{\boldsymbol{p},\,\boldsymbol{f}}(\mathbf{n})$.

The development of the recursions for computing the auxiliary parameters and angle-normalized joint-process estimation error for a QRD-LSL smoother can be seen in [11] and is omitted here due to space limitation. Equations (19)–(32) in Table I are final results of the development and constitute the *QRD-LSL smoothing algorithm* [11]. A signal-flow graph of the QRD-LSL smoothing algorithm showing the (2, 2)th-order QRD-LSL smoother using the sequence BFBF is shown in Fig. 2. Note that a QRD-LSL smoother may also be implemented by using the multichannel filtering algorithms of unequal channel lengths based on orthogonal Givens rotations developed by Rontogiannis and Theodoridis [17] for the special case of two input channels. However, the algorithm developed in this paper is different from [17] in the sense that one of C_N^f orthogonal basis vectors of the subspace of both past and future observations is constructed first followed by the projection of the present desired signal vector onto the orthogonal basis vector successively. As a result, successive stages of the QRD-LSL smoothers developed in this correspondence are decoupled. This decoupling property permits dynamic assignment and rapid automatic determination of the most effective smoother length.

V. Computer Simulations

In this section, we present results of computer simulations of adaptive equalization of a linear channel having unknown distortion. The simulations closely follow that of [16]. A polar form pseudo-random signal $x(n)$ is applied to a channel having unit pulse response

$$h_n = \begin{cases} \frac{1}{2}\left[1+\cos\left(\frac{2\pi}{W}(n-2)\right)\right], & n = 1, 2, 3 \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

The observation $y(n)$ is the sum of the channel output and an independent white Gaussian noise with variance 0.001. The adaptive equalizer attempts to correct the distortion produced by the channel and the additive noise. We compared the performances of three equalizers, each having order $N = 10$ (11 taps). Equalizer #1 was a tenth-order QRD-LSL filter. Equalizer #2 was a tenth-order QRD-LSL filter with five units time delay (i.e., five "future" observations were used) of the type described in [16]. Equalizer #2 would have possessed the order-recursive property were it not for the five units of delay. As noted earlier, once delay is introduced into a QRD-LSL filter, the order-recursive property is lost. Equalizer #3 was a (5, 5)th-order QRD-LSL smoother possessing the order-recursive property of the type described in this correspondence. Of the $C_{10}^5 = 252$ possible realizations of a (5, 5)th-order QRD-LSL smoother, we used the sequencing BFBFBFBFBF. The parameter W in (18) was set equal to 2.9 and 3.5 to provide for eigenvalue spreads 6.078 and 46.82, respectively.

The learning curves for the three equalizers are shown in Figs. 3 and 4. Each learning curve was obtained by ensemble-averaging the squared value of the *a posteriori* error over 200 independent trials of the experiment. It can be seen from the plots that the steady-state mean squared error of noncausal filters including the smoother and the filter with delay is about 15 dB less than that of a causal filter. It can also be seen that the rate of convergence of the (5, 5)th-order QRD-LSL smoother is a little faster than that of the tenth-order filter with delay. Additional realizations including the sequencing BBBBBFFFFF and the sequencing FFFFFBBBBB were tried. The simulation results revealed that the alternating sequence BFBFBF-BFBF (or FBFBFBFBFB) and equalizer #2 displayed the fastest and slowest initial transient performance respectively compared with other sequences of the (5, 5)th-order smoother, although their differences were not large. We conjecture that this is because signal autocorrelation functions are typically monotonically decreasing. In addition, the two alternating sequences that use the present observation first immediately followed by the use of its nearest-neighboring observations to estimate the present desired signal make earlier use of the correlation between the desired signal and the corresponding observations than other sequences. On the other hand, the realization of equalizer #2 that actually corresponds to the BBBBBBBBBB sequence with five units time delay introduced to the desired signal $x(n)$ makes the latest use of the correlation among all possible sequences.

VI. Conclusions

This work extends the QRD-LSL algorithm from filtering to smoothing with identical computational cost. All the desirable features found in the QRD-LSL filtering algorithm are also shared by the QRD-LSL smoothing algorithm with a finite delay but with a significant reduction in minimum mean square error. Due to the construction of the *orthogonal basis set* of the subspace of both *past* and *future* observations, successive stages of the QRD-LSL smoother developed in this paper are *decoupled*. This decoupling (or order-recursive) property gives the QRD-LSL smoother a computationally efficient, modular, latticelike structure.

References

[1] M. Morf and D. T. Lee, "Recursive least squares ladder forms for fast parameter tracking," in *Proc. 1978 IEEE Conf. Decision Contr.,* San Diego, CA, pp. 1362–1367.

[2] D. T. Lee, M. Morf, and B. Friedlander, "Recursive least-squares ladder estimation algorithms," *IEEE Trans. Circuits Syst.,* vol. CAS-28, pp. 467–481, 1981.

[3] H. Lev-Ari, T. Kailath, and J. Cioffi, "Least-squares adaptive lattice and transeveral filters: A unified geometric theory," *IEEE Trans. Inform. Theory,* vol. IT-30, pp. 222–236, 1984.

[4] S. Haykin, *Adaptive Filter Theory,* 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[5] J. T. Yuan and J. A. Stuller, "Least squares order-recursive lattice smoothers," *IEEE Trans. Signal Processing,* vol 43, pp. 1058–1067, May 1995.

[6] C. R. Ward, P. J. Hargrave, and J. G. McWhirter, "A novel algorithm and architecture for adaptive digital beamforming," *IEEE Trans. Antennas Propagat.,* vol. AP-34, pp. 338–346, Mar. 1986.

[7] M. D. Levin and C. F. N. Cowan, "The performance of eight recursive least squares adaptive filtering algorithms in a limited precision environment," in *Proc. Euro. Signal Process. Conf.,* 1994, Edinburgh, U.K., pp. 1261–1264.

[8] I. K. Proudler, J. G. McWhirter, and T. J. Shepherd, "QRD-based lattice filter algorithms," in *Proc. SPIE Adv. Algorithms Architecture Signal Process. IV,* 1989, vol. 1152.

[9] I. K. Proudler, J. G. McWhirter, and T. J. Shepherd, "The QRD-based least squares lattice algorithm for wide-band beamforming and adaptive filtering," in *Proc. ISSPA,* Brisbane, Australia, Aug. 27–31, 1990.

[10] P. A. Regalia and M. G. Bellanger, "On the duality between fast QR methods and lattice methods in least squares adaptive filtering," *IEEE Trans. Signal Processing,* vol. 39, pp. 879–891, Apr. 1991.

[11] J. T. Yuan, "QR-decomposition for noncausal filtering—Part II: A QR-decomposition-based LSL smoothing algorithm," in *IASTED Int. Conf. Modeling Simulation,* Pittsburgh, PA, May 14–17, 1997, pp. 354–358.

[12] I. K. Proudler, J. G. McWhirter, and T. J. Shepherd, "Computationally efficient QR decomposition approach to least squares adaptive filtering," in *Proc. Inst. Elect. Eng. F.,* vol. 138, no. 4, Aug. 1991.

[13] B. Yang, and J. F. Bohme, "Rotation-based RLS algorithms: Unified derivations, numerical properties and parallel implementations," *IEEE Trans. Signal Processing,* vol. 40, pp. 1151–1167, 1992.

[14] P. Strobach, *Linear Prediction Theory—A Mathematical Basis for Adaptive Systems.* Berlin, Germany: Springer-Verlag, 1990.

[15] J. T. Yuan, "QR-decomposition for noncausal filtering—Part I: A modified QR-decomposition," in *Proc IASTED Int. Conf. Modeling Simulation,* Pittsburgh, PA, May 14–17, 1997, pp. 349–353.

[16] E. H. Satorius and J. D. Pack, "Applications of least squares lattice algorithms to adaptive equalization," *IEEE Trans. Commun.,* vol. COMM-29, pp. 136–142, Feb. 1981.

[17] A. Rontogiannis and S. Theodoridis, "A new highly parallel multichannel fast QRD-LS adaptive algorithm," in *Proc. EUSIPCO,* 1996, Trieste, Italy.