Simplified Adaptive IIR Filters Based on Optimized Orthogonal Prefiltering

August N. Kaelin, Allen G. Lindgren, *Senior Member, IEEE,* and George S. Moschytz, *Fellow, IEEE*

Abstract—In order to reduce the circuit complexity associated with the estimation of echoes coming from systems with a long impulse response, we consider an estimator which is based on prefiltered input data. We propose a design of this prefilter which is optimal for a given *system environment.* In doing so, we represent the unknown discrete-time system by a set of *characteristic* impulse responses, which adequately describe the *variety* of the system. For such an environment we determine the optimum poles of a recursive prefilter. These poles are assumed to be *fixed* during the on-line LMS estimation process, which estimates the unknown echo by linearly weighting the prefilter states. An echo canceler for a typical European telephone subscriber-loop environment is used as a practical example. For this example the prefilter is optimized and realized with an orthogonal-state (lattice) filter. This not only reduces the computational costs—if compared to a conventional FIR filter design—but also permits a substantial speed-up of the on-line LMS adaptation process.

I. INTRODUCTION

IN many applications it is desired to estimate the response of a given but unknown system. Typical applications are in adaptive echo cancellation and system identification. While the results presented here are of general interest, our motivation for obtaining them was the problem of echo cancellation for systems with a very long impulse response such as echo paths in digital subscriber loop transmission systems. The conventional finite impulse response (FIR) filters used in such systems require filter lengths in the order of hundreds of taps. Adaptive infinite impulse response (IIR) filters introduce other problems, foremost among them being the requirement for stability. However, it has been shown, e.g., [1] and [2], that fixed (nonadaptive) recursive prefiltering can reduce the complexity of such echo cancelers considerably. This concept is expanded on in this paper by designing a simplified adaptive IIR filter which makes use of a prefilter which is optimized for a given system environment. We show that for classes of systems for which certain characteristics are known in a statistical sense, such a simplified adaptive IIR filter can obtain high performance.
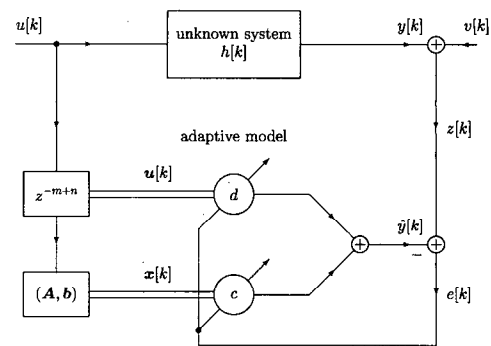
Fig. 1. Linear echo cancelation with a simplified adaptive IIR filter.

In Section II we introduce the proposed simplified adaptive IIR filter and state the associated prefilter design problem. In Section III we identify some fundamental relations relevant to the optimization problem. Section IV describes a method of representing the possible variations of the unknown system by (e.g., measured) impulse responses. An appropriate system space is then defined in terms of a minimal set of orthogonal impulse responses. We continue with the development of the optimization scheme (Section V), which is then illustrated by two examples (Section VI); a simple first-order example and the more practical design of an echo canceler for a European telephone subscriber-loop environment. In Section VII the optimized prefilter of the latter example is used to discuss so-called *orthogonal* filter realizations by means of the inverse lattice filter. Finally, Section VIII compares such a filter realization with nonorthogonal ones and verifies its superiority with the help of numerical simulations. We show, with the help of our echo canceler example, that an adaptive IIR filter which is based on such a prefilter not only reduces the computational costs—compared to the conventional FIR filter—but also speeds up the on-line adaptation process of the filter.

II. STATEMENT OF THE PROBLEM

Referring to the linear echo cancellation scheme depicted in Fig. 1, we assume the independent noise processes $u[k]$ and $v[k]$ to be stationary and with zero mean. For any discrete time k, the available $z[k]$ represents the data $v[k]$ disturbed by the unknown echo $y[k]$. The echo signal itself is assumed to be generated by a system with unknown impulse response $h[k]$ driven by the input $u[k]$. To recover $v[k]$, the echo must be reconstructed and subtracted from $z[k]$.

In reconstructing the echo, the objective of the adaptive model is to match the impulse response $h[k]$ of the unknown system, subject to the following constraints. The model transfer function has a fixed denominator of degree n and an adjustable numerator of degree $m \geq n$. In a general state-space model this means, that in

$$\hat{y}[k] = c^T x[k] + d^T u[k] \tag{1}$$

with input data vector $u[k] \triangleq (u[k], u[k-1], \cdots, u[k-m+n])^T$ and nth-order state vector $x[k] \triangleq (x_0[k], x_1[k], \cdots, x_{n-1}[k])^T$ propagated as

$$x[k+1] = Ax[k] + bu[k-m+n], \tag{2}$$

only the output vectors $c \triangleq (c_0, c_1, \cdots, c_{n-1})^T$ and $d \triangleq (d_0, d_1, \cdots, d_{m-n})^T$ are adjustable. In combining these vectors into the weight vector $w^T \triangleq (c^T, d^T)$, the estimation of the echo is written as $\hat{y}[k] = w^T \varphi[k]$, where the augmented state vector $\varphi^T[k] \triangleq (u[k]^T, x[k]^T)$ denotes the state of the complete model. A fixed nth-order *prefilter* (2) is used to generate the state of the complete mth-order model. It should be noted that d allows $h[0]$, $h[1]$, $\cdots$, $h[m-n]$ to be matched exactly, whereas the remaining part of $h[k]$ may only be matched *approximately* with the help of c.

The *system environment*, i.e., the space of possible systems which the adaptive model must cope with, is represented by a set of M *characteristic* responses $h_i[k]$. For each system i, the model response $\hat{h}[k]$ can be optimized over the adjustable coefficient vectors c and d. In measuring the matching quality, we introduce the vector notation

$$h_i \triangleq (h_i[m-n+1], h_i[m-n+2], \cdots, h_i[m-n+N])^T, \tag{3}$$

for $i = 1, 2, \cdots, M$ and with N sufficiently large in order to describe the relevant part of the system response. The appropriate relevant part of the model response will be denoted by $\hat{h}$.

In matching a measured impulse response, Steiglitz [3] considered the minimization of a sum-square error to be the "ideal problem." In generalizing this error criterion for our environment, we introduce the error measure

$$J \triangleq \sum_{i=1}^{M} \rho_i^2 \min_{c} \|h_i - \hat{h}\|^2 \tag{4}$$

where for every system i the possibility of assigning a weight ρ_i has been considered. As usual, $\| \cdot \|$ denotes the 2-norm of a vector. It is noted that if ρ_i^2 is equal to the probability p_i that system i is identified, the error measure J corresponds—assuming $u[k]$ to be a white noise process with variance σ_u^2—to the echo return enhancement (ERE) in the mean, i.e.,

$$J = \frac{1}{\sigma_u^2} \sum_{i=1}^{M} p_i \min_{c,d} E[(y_i[k] - \hat{y}[k])^2]. $$

Again, the index i refers to the appropriate system. Now, the optimum prefilter design problem is stated as follows: *Given the system environment (3)—with appropriate weights ρ_i—find*

for the mth-order model (1) the optimum nth-order prefilter (2) which minimizes the error criterion (4).

In Section VIII, we will apply the well-known Least Mean Square (LMS) algorithm [4], a stochastic gradient algorithm, for the real-time adjustment of w. We will use the adjustment rule

$$w[k+1] = w[k] + \mu\varphi[k]e[k] \tag{5}$$

where $\varphi[k]e[k]$ is the stochastic gradient of the Mean Squared Error (MSE) $E[e^2[k]]$ with respect to the weight vector w and $e[k] = z[k] - \varphi^T[k]w[k]$, the error signal according to Fig. 1. The diagonal step-size matrix μ allows a tradeoff between the speed of the adaptation process and its misadjustment. For details see, e.g., [4].

Referring to our adaptive model (see Fig. 1), it is noted that, as an alternative, the input vector b could be adjusted with A and c fixed. However, our system configuration has the advantage that with the stochastic gradient vector being equal to the complete state vector $\varphi[k]$ times the error $e[k]$, all the quantities required for the above adjustment rule are readily available.

It is well known that for the form of the LMS algorithm defined by (5), the best performance results if the covariance matrix $E[\varphi[k]\varphi^T[k]]$ is diagonal, i.e., if the states are orthogonal. Such an *orthogonal* realization of the prefilter will be discussed in more detail in Sections VII and VIII. In what follows we identify some relations which are relevant to our optimization problem.

III. SOME FUNDAMENTAL RELATIONS

We start by denoting the error for the ith system as

$$J_i \triangleq \min_{c} \|h_i - \hat{h}\|^2. \tag{6}$$

In reformulating this expression, we write the model response as

$$\hat{h} = C^T c = Gc \tag{7}$$

where the $n \times N$ matrix

$$C \triangleq (b, Ab, A^2 b, \cdots, A^{N-1} b)$$

is the controllability matrix and

$$G \triangleq C^T$$

its transpose. Equation (6) can then be rewritten as

$$J_i = \min_{c} \|h_i - Gc\|^2 = \|h_i - \hat{h}_i\|^2 \tag{8}$$

where $\hat{h}_i$ is the optimized model response for system i. The solution to this linear least-square (LS) problem, in which the coefficient vector c is optimized for system i, is given as

$$c_i = (G^T G)^{-1} G^T h_i.$$

Note that this requires $G^T G$ to be invertible, which can be guaranteed if G or, equivalently, C has its full rank n, i.e.,

if the prefilter (2) is controllable [5]. Substituting c_i into (7) results in the optimized model response with

$$\hat{h}_i = P_G h_i \qquad (9)$$

where the orthogonal projection

$$P_G \doteq G(G^T G)^{-1} G^T \qquad (10)$$

projects $\hat{h}_i$ into the space spanned by the columns of our model space, G.

To proceed, we note that our objective is to select the prefilter (2) such that G optimally matches the given system environment. This depends only on the eigenvalues of A or, equivalently, on the coefficients of the characteristic equation. This is shown next.

The characteristic equation, i.e., the nth-order denominator of our model, is denoted by $D(z) = 1 + \alpha_1 z^{-1}, \cdots, + \alpha_n z^{-n}$. With the Cayley–Hamilton theorem (see, e.g., [5, p. 658]), A satisfies its own characteristic equation, i.e.,

$$A^n + \alpha_1 A^{n-1} + \alpha_2 A^{n-2} + \cdots + \alpha_n I = O. \qquad (11)$$

Hence, we have

$$A^T G = O \qquad (12)$$

with the $(N - n) \times N$ matrix

$$A^T \doteq \begin{pmatrix} \alpha_n & \alpha_{n-1} & \cdots & \alpha_1 & 1 & & 0 \\ & \alpha_n & \alpha_{n-1} & \cdots & \alpha_1 & 1 & \\ & & \ddots & \ddots & & \ddots & \ddots \\ 0 & & & \alpha_n & \alpha_{n-1} & \cdots & \alpha_1 & 1 \end{pmatrix}.$$

Equation (12) implies

$$P_{\mathcal{A}} P_G = O,$$

where $P_{\mathcal{A}}$ is the orthogonal projection based on $\mathcal{A}$. For a controllable prefilter (2), rank (P_G) = co-rank $(P_{\mathcal{A}})$ = n and we conclude that:

$$P_{\mathcal{A}} \doteq \mathcal{A}(\mathcal{A}^T \mathcal{A})^{-1} \mathcal{A}^T = I - P_G. \qquad (13)$$

Now, (8), together with (9), becomes

$$J_i = \|P_{\mathcal{A}} h_i\|^2 = h_i^T P_{\mathcal{A}} h_i \qquad (14)$$

since $P_{\mathcal{A}}$ is a projection (i.e., $P_{\mathcal{A}} = P_{\mathcal{A}} P_{\mathcal{A}}$).

Note that the coefficients of the characteristic equation (11), combined in the parameter vector

$$\alpha \doteq (1, \alpha_1, \alpha_2, \cdots, \alpha_n)^T, \qquad (15)$$

define $\mathcal{A}$ and its associated orthogonal projection $P_{\mathcal{A}}$ and, with (14), the error in matching system i. Furthermore, since $J = \sum_{i=1}^{M} \rho_i^2 J_i$, it also defines the measure of matching the whole environment. Hence, any prefilter (2) having the same α (i.e., the same pole locations) is equivalent in the sense that it yields the same error. A prefilter optimization will therefore not involve the input vector b. As far as the prefilter remains controllable, it can be selected arbitrarily. We will use this flexibility later when looking for an orthogonal realization that has better convergence properties for the on-line LMS adaptation algorithm than nonorthogonal ones (see Sections VII and VIII).

IV. THE SYSTEM ENVIRONMENT

In what follows, we discuss a description of an appropriate system environment. Such a system environment has to be sufficiently large in order to "train" the prefilter. The goal is a performance which is sufficient for any system the adaptive filter will encounter (not only the supplied "training" systems.) To start with, we assume that a "large" number (M) of impulse responses h_i, together with appropriate weights ρ_i, or equivalent information about the variety of the system, is available to the designer. In reducing this environment to a tractable number of systems, we will show that linearly dependent h_i's can be eliminated, since they do not contribute additional information with regard to the variety of the system.

In doing so, we condense the weighted environment (3) into the matrix

$$E \doteq (\rho_1 h_1, \rho_2 h_2, \cdots, \rho_M h_M) \qquad (16)$$

and combine the J_i's according to (14) into the error measure for the total environment. This gives[1]

$$J(E, \mathcal{A}) = \text{tr}\{E^T P_{\mathcal{A}} E\} \qquad (17)$$

with $P_{\mathcal{A}}$ the orthogonal projection as defined in (13).

We proceed by carrying out a singular-value decomposition (SVD) of E. With the unitary matrices (U_1, U_2, U_3) and (V_1, V_2), E can be written as

$$E = (U_1, U_2, U_3) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix},$$

where the diagonal matrices Σ_1 and Σ_2 contain the singular values in decreasing order, weighting U_1 and U_2, respectively. With $P_{\mathcal{A}}$ a projection, such an SVD allows the error measure (17) to be partitioned into

$$J(E, \mathcal{A}) = J(E_1, \mathcal{A}) + J(E_2, \mathcal{A}), \qquad (18)$$

where $J(E_i, \mathcal{A}) = \text{tr}\{E_i^T P_{\mathcal{A}} E_i\}$ is based on the transformed (orthogonal) sub-environment $E_i \doteq E V_i = U_i \Sigma_i$.

Since the projection $P_{\mathcal{A}}$ is orthogonal, the measures of the sub-environments are bounded by

$$0 \leq J(E_i, \mathcal{A}) \leq \text{tr}\{\Sigma_i^2\}.$$

With this bounds, the following relation can be obtained from (18):

$$J(E_1, \mathcal{A}) \leq J(E, \mathcal{A}) \leq J(E_1, \mathcal{A}) + \text{tr}\{\Sigma_2^2\}. \qquad (19)$$

Hence, any prefilter (2) yielding the error $J(E_1, \mathcal{A})$ in the sub-environment E_1 achieves in the original environment E an error which is upper bounded by $J(E_1, \mathcal{A}) + \text{tr}\{\Sigma_2^2\}$.

If $\text{tr}\{\Sigma_2^2\}$ is zero, the neglected E_2 contains only linearly dependent impulse responses, which do not contain additional information with respect to the variety of the unknown system. For an environment which is based on real measurements, we collect in $\text{tr}\{\Sigma_2^2\}$ the smallest singular values of E until the design accuracy of the filter is reached. Within this design accuracy, the reduced environment E_1 is equivalent to

---

[1] The trace operator is denoted by $\text{tr}\{.\}$.

the original environment E. Based on such an environment description, we now develop our optimization scheme for the prefilter.

V. DEVELOPMENT OF THE OPTIMIZATION SCHEME

We start by defining the response matrix for the ith impulse response h_i as

$$H_i \triangleq \begin{pmatrix} h_{i,n+1} & h_{i,n} & \cdots & h_{i,1} \\ h_{i,n+2} & h_{i,n+1} & \cdots & h_{i,2} \\ \vdots & \vdots & & \vdots \\ h_{i,N} & h_{i,N-1} & \cdots & h_{i,N-n} \end{pmatrix}$$

with $h_{i,j}$ denoting the jth component of h_i and by introducing with

$$H^T \triangleq (\rho_1 H_1^T, \rho_2 H_2^T, \cdots, \rho_M H_M^T)$$

an appropriately weighted collection for the whole environment.

We now substitute in (14) the identity

$$A^T h_i = H_i \alpha$$

and obtain for the error measure $J = \sum_{i=1}^{M} \rho_i^2 J_i$ the weighted 2-norm

$$J = \|H\alpha\|_Q^2, \tag{20}$$

where $\| \cdot \|_Q^2 \triangleq (\cdot)^T Q(\cdot)$. Here the weight matrix is block-diagonal and given as

$$Q = \begin{pmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_M \end{pmatrix} \tag{21}$$

with $(N - n) \times (N - n)$ block entries $Q_i \triangleq (A^T A)^{-1}$. Minimizing the error measure J with respect to the prefilter poles or, equivalently, with respect to α, we obtain a weighted LS problem, stated by

$$J_{\min} = \min_{\alpha} \|H\alpha\|_Q^2. \tag{22}$$

Since Q is positive definite, the factorization $Q = L^T L$ is guaranteed, which in turn allows the optimization problem (22) to be restated as

$$J_{\min} = \min_{\alpha} \|LH\alpha\|^2. \tag{23}$$

Since the filter matrix L is dependent on α, the minimization problem (23) is nonlinear. However, in applying the idea of the Steiglitz–McBride algorithm [6] to our "multi-system" environment, we may first estimate the filter matrix L. An estimate of the parameter vector α is then obtained by solving the remaining linear LS problem. Iterating this procedure results in a refinement of α. This iterative procedure is summarized as following:

1) Compute $L^{(l)}$ defined by $Q^{(l)} = L^{(l)T} L^{(l)}$, then
2) find $\alpha^{(l+1)}$ such that $\min_{\alpha} \|L^{(l)} H\alpha\|^2$

where (l) refers to the previous estimate of α. It is noted that the minimum of the Steiglitz–McBride algorithm is always very close to the global minimum, but not identical [7]. This was confirmed as discussed in the next section. This "global" convergence allows one to initialize the algorithm with $\alpha^{(0)} = (1, 0, \cdots, 0)^T$. The accuracy of the iterative refinement can be measured by

$$J^{(l)} = \|L^{(l)} H\alpha^{(l)}\|^2, \tag{24}$$

which is the error measure for $\alpha^{(l)}$.

It should be noted that the weighted LS problem (22) considered here is very sensitive to roundoff errors when Q weights high-order prefilters with poles close to the unit circle. For high-order prefilters, we found an iterative algorithm which is based on an explicit computation of $Q^{(l)}$ to become numerically unstable. However, in making use of the special structure of A [cp. (13)], a robust solution can be derived which computes $L^{(l)}$ directly from $A^{(l)}$. This will be discussed next. For convenience, we will thereby drop the iteration index l.

Consider the following partitioning of A:

$$A^T = N - n$$

$$\begin{pmatrix} \alpha_n & \cdots & \alpha_1 & 1 & & & & 0 \\ & \ddots & \vdots & \alpha_1 & 1 & & & \\ & & \alpha_n & \vdots & \ddots & \ddots & & \\ & & & \alpha_n & & & 1 & \\ & & & & \ddots & & & \\ 0 & & & & \alpha_n & \cdots & \alpha_1 & 1 \end{pmatrix}$$

$$\underbrace{}_{A_1^T} \underbrace{}_{A_2^T}$$

Note that the $(N - n) \times (N - n)$ Matrix A_2 has the inverse

$$A_2^{-1} = \begin{pmatrix} 1 & \beta_1 & \cdots & \beta_{N-n-1} \\ & \ddots & \ddots & \vdots \\ & & & \beta_1 \\ 0 & & & 1 \end{pmatrix} \tag{25}$$

with coefficients β_i related to α_j by the filter process

$$\beta_i = -\sum_{j=0}^{n} \beta_{i-j} \alpha_j, \quad \text{for } i > 0$$

and initialized with $\beta_0 = 1, \beta_{-1} = 0, \beta_{-2} = 0, \cdots$. With this partitioning we express the block entries of Q as

$$(A^T A)^{-1} = (A_1^T A_1 + A_2^T A_2)^{-1}.$$

Noting[2] that $(A_2^T A_2)^{-1} = A_2^{-1} A_2^{-T}$ is already a good approximation for $(A^T A)^{-1}$, we use the Sherman–Morrison–Woodbury formula (e.g., [8, p. 51]) to express the exact inverse as

$$(A^T A)^{-1} = A_2^{-1} [I - A_2^{-T} A_1^T (I + A_1 A_2^{-1} \\ \cdot A_2^{-T} A_1^T)^{-1} A_1 A_2^{-1}] A_2^{-T}.$$

[2]For brevity we use the notation $A_2^{-T} \triangleq (A_2^{-1})^T$

To continue we substitute $\mathcal{A}_1\mathcal{A}_2^{-1}$, a $n \times (N - n)$ matrix, by its singular-value decomposition

$$\mathcal{A}_1\mathcal{A}_2^{-1} = USV^T \qquad (26)$$

and finally obtain

$$(\mathcal{A}^T\mathcal{A})^{-1} = \mathcal{A}_2^{-1}V(I + S^T S)^{-1}V^T\mathcal{A}_2^{-T}.$$

Hence, the singular-value decomposition of $\mathcal{A}_1\mathcal{A}_2^{-1}$, a $n \times (N - n)$ matrix, allows Q, a $M(N - n) \times M(N - n)$ matrix, to be factorized as $L^T L$ with

$$L = \begin{pmatrix} L_1 & & 0 \\ & \ddots & \\ 0 & & L_M \end{pmatrix} \qquad (27)$$

and block entries $L_i = (I + S^T S)^{-1/2}V^T\mathcal{A}_2^{-T}$.

To summarize, the SVD (26) together with (27) allows to compute the filter matrix L directly from $\mathcal{A}$ (cp. (13)). This is step 1 of the proposed iterative procedure. Step 2 is nothing more than a conventional linear LS problem. The reader is referred to [8] for a suitable LS solver. In using the standard built-in LS-solver of the numeric computational program MATLAB,[3] we found accurate prefilters for n up to approximately 12 (with a floating point relative accuracy of 10^{-16}.) A MATLAB program with the complete design procedure is available from the authors.

VI. EXAMPLES

In order to provide some insight into the proposed prefilter design, we apply it, in the following, i) to a simple first-order environment and ii) to the more practical design of an echo canceler for the European telephone subscriber-loop environment.

Example I. A First-Order Environment: The system environment is given by

$$h_i = (1, \lambda_i, \lambda_i^2, \cdots, \lambda_i^N)^T, \qquad i = 1, 2, 3, \qquad (28)$$

where $\lambda_1 = 0.9, \lambda_2 = 0.8$ and $\lambda_3 = 0.7$. We assumed $N = 40$, which allows the corresponding exponentials to be modeled with a coefficient error of less than 1%. In order to equally weight the three cancelation levels, we use $\rho_i^2 = 1/(3\|h_i\|^2)$ resulting in

$$J = \sum_{i=1}^{3} \frac{1}{3} \frac{\|h_i - \hat{h}_i\|^2}{\|h_i\|^2}.$$

Note that with three prefilter poles located at 0.9, 0.8, and 0.7, the three exponentials would be exactly matched. Here we assume a *reduced-order* adaptive filter with two fixed prefilter poles and a total of 3 adjustable coefficients, i.e., $n = 2$ and $m = 2$. Starting with $\alpha^{(0)} = (1, 0, 0)^T$, the proposed iterative algorithm, according to Section V, yields the optimum prefilter with $J^{(\infty)} = -35.39$ dB in about 4 iterations, as shown in Table I. This table contains the coefficients of the characteristic equation and the obtainable error measures $J^{(l)}$ for every iteration l. The optimum prefilter poles p_1 and p_2,
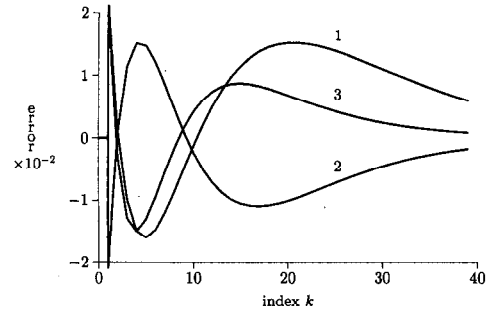
Fig. 2. Example I. Obtained impulse-response errors if $h_i[k], i = 1, 2, 3$, are estimated with the prefilter poles at 0.8845 and 0.7446 (worst-case optimization).

TABLE I
EXAMPLE I. ITERATIVE PREFILTER DESIGN

iteration l	1	2	3	4
α_1	-1.6085	-1.6379	-1.6390	-1.6391
α_2	0.6402	0.6648	0.6658	0.6659
$J^{(l)}$ [dB]	-34.60	-35.41	-35.39	-35.39

i.e., the roots of the characteristic equation $z^2 + \alpha_1 z + \alpha_2 = 0$, are given as 0.8955 and 0.7436, respectively, whereas the error measures for the individual exponentials are obtained as -40.4 dB, -28.53 dB, and -29.56 dB. Furthermore, it is noted that $\rho_1 = 0.9285, \rho_2 = 2.5191$, and $\rho_3 = 2.8126$ results (with the prefilter poles at 0.8845 and 0.7446) in a worst case optimization with all errors at -30.0 dB. With optimally adjusted d_0, c_0 and c_1 (using (8) and (9)), the obtained errors in the impulse responses, i.e., $h_i[k] - \hat{h}_i[k]$, are shown in Fig. 2. The figure shows an exact match of the first value $h_i[0]$ and a close approximation of the remaining values $h_i[1], h_i[2], \cdots h_i[40]$.

Referring to Table I, it is finally noted that $J^{(2)} = -35.41$ dB is slightly smaller then the iteratively found "minimum" $(J^{(\infty)} = -35.39$ dB). However, in all our examples we found the obtained "minimum" always very close to the global minimum. This coincides with what has been reported in [7] (for the comparable single-system case.)

Example II. A Telephone Subscriber-Loop Environment: As a more practical example, we consider the design of an echo canceler for digital data transmission systems in a typical telephone subscriber-loop environment. The considered characteristic impulse responses are taken from measurements on a test hybrid circuit [9] terminated with a HDSL-loop simulator (ILS-2) from Wandel & Goltermann realizing the ETSI (European Telecommunications Standards Institute) test loops DRT/TM-3002 sampled at 58 4kHz. The first 80 of the 200 samples considered are depicted in Fig. 3.

Setting $\rho_i^2 = 1/M$, the error measure J is equal to the echo return enhancement (ERE) in the mean, if a white input $u[k]$ is assumed. Assuming the acceptable level at -60 dB, the adaptive IIR-filter models given in Table II are comparable. In order to meet these requirements, Table II gives the required model degree m for every optimized nth-order prefilter, along with a complexity measure C. This is defined by the number of multiplications per sample interval required for both the
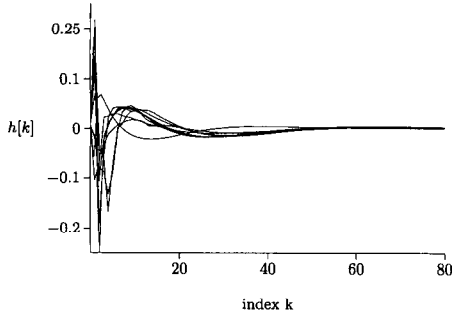
Fig. 3. Example II. European subscriber-loop environment; echo impulse responses for 11 ETSI test loops.

TABLE II
EXAMPLE II. COMPLEXITY OF VARIOUS ADAPTIVE FILTER
MODELS WITH ERE OF AT LEAST -60 dB IN THE MEAN

m	96	72	48	32	21	17	17
n	0	1	2	3	4	5	6
C	192	145	98	67	46	39	40

TABLE III
EXAMPLE III. ITERATIVE DESIGN OF THE OPTIMUM
5TH-ORDER PREFILTER FOR A 17TH-ORDER MODEL

iteration l	1	$\cdots 4$	$\cdots 6$	7
α_1	-1.8053	-4.5854	-4.6317	-4.6318
α_2	0.5306	8.4152	8.5883	8.5888
α_3	0.3669	-7.7261	-7.9694	-7.9701
α_4	0.0233	3.5486	3.7009	3.7014
α_5	-0.1056	-0.6523	-0.6881	-0.6882
$J^{(l)}$ [dB]	-41.22	-60.15	-60.05	-60.05

filtering process and the corresponding LMS adaptation of the coefficient vectors c and d. The prefilter itself is assumed to be realized orthogonally, with an inverse lattice structure based on one-multiplier sections as discussed in Section VII.

As shown in Table II, a remarkable reduction in complexity can be obtained, compared to the conventional FIR-filter canceler. As an example, using an optimized prefilter with $n = 5$ poles; the degree of our adaptive model is reduced to $m = 17$, and the complexity to $C = 39$, compared to $m = 96$ and $C = 192$ for the conventional FIR-filter model ($n = 0$). Comparing the corresponding complexity measures shows that a prefilter with 5 poles is optimum for this environment. It is noted that an increase to $n = 6$ prefilter poles does not further decrease the overall complexity of the model. It seems that 5 poles can completely describe the "recursive" part of the corresponding impulse responses.

The iterative design procedure is demonstrated with the help of this optimum model ($n = 5$, $m = 17$). As is shown in Table III, the algorithm converges in 6 iterations. Again, a slight derivation from the global minimum can be observed.

VII. AN ORTHOGONAL PREFILTER REALIZATION

As mentioned earlier, an orthogonal prefilter realization can increase the performance of the on-line LMS-adaptation process. A filter structure is said to be *orthogonal* if, at any time, all internal variables are uncorrelated and have unit
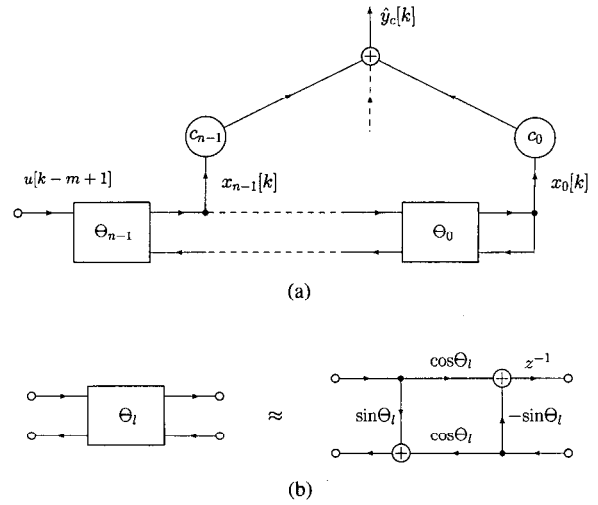

Fig. 4. (a) Structure of the inverse lattice filter. (b) Lattice section.

variance assuming a stationary white noise input [10, p. 450]. Referring to our system (1), this means

$$E\left[\begin{pmatrix} x[k] \\ u[k] \end{pmatrix} (x^T[k], u^T[k])\right] = I. \qquad (29)$$

For our prefilter (2), with $E[u[k]u^T[k]] = I$, (29) is reduced to

$$E\left[\begin{pmatrix} x[k] \\ u[k - m + n] \end{pmatrix} (x^T[k], u[k - m + n])\right] = I. \qquad (30)$$

As an example of an orthogonal structure, which fulfills (30), we consider the inverse lattice structure [10, p. 476] shown in Fig. 4.

Assume

$$\frac{P(z)}{D(z)} = c^T(zI - A)^{-1}b$$

to be the proper part of the model transfer function with the nth-order polynomials in z^{-1}

$$P_l(z) = \sum_{i=0}^{l} p_{l,i} z^{-i}$$

$$D_l(z) = \sum_{i=0}^{l} d_{l,i} z^{-i},$$

where $l = n$. The parameters of the lattice sections, $k_l \hat{=} \sin \Theta_l$, can than be obtained—according to [11]—from $P_n(z)$ and $D_n(z)$ as

$$zB_l(z) = D_l(z^{-1})z^{-l}$$

$$k_{l-1} = d_{l,l}$$

$$D_{l-1}(z) = [D_l(z) - k_{l-1}zB_l(z)]/(1 - k_{l-1}^2)$$

$$c_l = p_{l,0}$$

$$P_{l-1}(z) = P_l(z) - c_l z^{-(n-l)} D_l(z),$$

for $l = n, n - 1, \cdots, 1$ with $c_0 = P_{0,0}$. Note that only the denominator $D(z)$, or its equivalent lattice parameters k_l ($\hat{=} \sin \Theta_l$), are precomputed, whereas the numerator $P(z)$ is adapted on-line with the adjustable coefficient vector c.

(a)

$$\lambda_l = \sqrt{\frac{1 + k_l}{1 - k_l}}$$
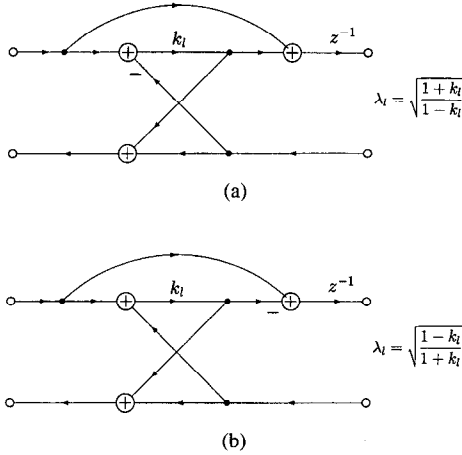

(b)

$$\lambda_l = \sqrt{\frac{1 - k_l}{1 + k_l}}$$

Fig. 5. One-multiplier lattice sections.

Each lattice section represents a rotation which can be implemented very efficiently using look-up techniques (with the so-called CORDIC algorithm, see e.g., [12, pp. 348]). However, if the normalization requirement (to unity power) is eased, simple one-multiplier lattice sections can be used instead of the normalizing rotators. Fig. 5 gives two one-multiplier lattice sections [11], which—replacing the rotators—scale the (normalized) state variables to

$$x_l[k] \leftarrow x_l[k] \prod_{i=l}^{n} \lambda_i,$$

with λ_i given as in Fig. 5.

VIII. A COMPARISON OF DIFFERENT PREFILTER REALIZATIONS

In investigating the influence of various prefilter realizations on the on-line LMS adaptation of the adaptive model according to Fig. 1, we again consider Example II. We compare the conventional FIR filter ($m = 96$) with the optimized IIR-filter model ($n = 5, m = 17$) where two different prefilter realizations have been considered. These are an orthogonal filter realization utilizing the inverse lattice filter of Fig. 4, as discussed in Section VII, and a parallel filter realization (residue-only adjustment scheme) with first-order sections for the real poles and second-order sections for the complex-conjugate poles.

In investigating the performance of the different adaptive filter structures, we use them as echo cancelers according to Fig. 1 in a 2B1Q HDSL transceiver [9] for a typical ETSI test loop according to Fig. 6. We set $E[v^2[k]]$ to 10^{-4}, which is a worst-case level for the "far-end" signal $v[k]$ to be estimated. The misadjustment, defined by

$$\mathcal{M} \doteq \lim_{k \to \infty} \frac{E[(y[k] - \hat{y}[k])^2]}{E[v^2[k]]}$$

is set to $\mathcal{M} = 10^{-2}$ guaranteeing a reasonable accuracy of this estimation. These values allow for an excessive MSE, $E[(y[k] - \hat{y}[k])^2]$, of 10^{-6} which is within the design accuracy of the filters. The misadjustment $\mathcal{M}$ of the filter is related to
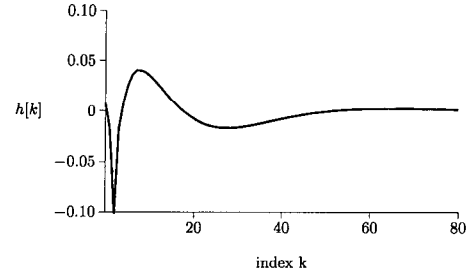

index k

Fig. 6. Example II. The echo impulse response for a typical ETSI test loop.
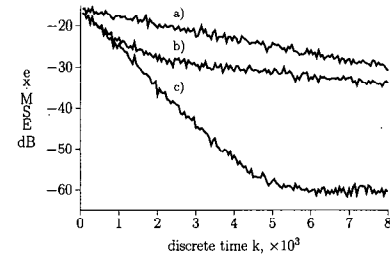

discrete time k, $\times 10^3$

Fig. 7. Example II. Learning curve (excessive MSE, averaged over 100 runs) for different echo-canceler structures: (a) a conventional FIR filter ($m = 96, n = 0$) and an optimum IIR filter ($m = 17, n = 5$) with the prefilter; (b) realized as a parallel structure; and (c) realized as an orthogonal structure.

the μ matrix (see, e.g., [4]) as follows:

$$\mu = \frac{1}{m + 1} \frac{2\mathcal{M}}{1 + \mathcal{M}} \, \text{diag}^{-1} \left(E[\varphi[k]\varphi^T[k]] \right).$$

The performed simulations are based on white input noise $u[k]$ with variance normalized to $\sigma_u^2 = 1$, and the excessive MSE obtained by averaging over 100 runs. The resulting learning curves are depicted in Fig. 7 for all three designs. It is emphasized that all three designs converge to the same final excessive MSE of -60 dB. Note, however, that the optimum IIR filter requires 18 coefficients to be adapted, which is far less than the 97 coefficients required for the comparable FIR filter. As shown in Fig. 7, this results in a correspondingly fast decay of the learning curves for the IIR-filter realizations, at least up to approximately -30 dB. To continue with this fast decay, either a more sophisticated adaptation algorithm than (5), or, alternatively, an orthogonal filter realization such as design c), must be used.

IX. CONCLUSION

A method has been presented of finding linearly adjustable IIR filters resulting in reduced computational costs and improved performance for a given system environment. A priori knowledge about the variety of the unknown echo path is used to find a set of fixed, but optimally placed poles for the IIR filter. For the pole optimization an iterative procedure is derived, which is based on the well-known Steiglitz–McBride algorithm. The sensitivity problem for high-order filter designs is discussed and a robust solution is derived. Furthermore, using the flexibility of the inverse lattice filters, an orthogonal filter realization is discussed which allows any desired stable pole placement. The use of such an orthogonal filter not only decreases the computational costs—when compared to a

conventional FIR filter solution—but also increases the speed of convergence of the on-line LMS adaptation process.

ACKNOWLEDGMENT

## REFERENCES

[1] K. Hayashi, S. Sugimoto, and F. Mano, "Design of 2B1Q transceiver for ISDN subscriber loops," in *Proc. Int. Conf. Commun.*, pp. 228–232, 1989.

[2] G. W. Davidson and D. D. Falconer, "Reduced complexity echo cancellation using orthonormal functions," *IEEE Trans. Circuits Syst.*, vol. 38, no. 1, pp. 20–28, 1991.

[3] K. Steiglitz and L. E. McBride, "On the simultaneous estimation of poles and zeros in speech analysis," *IEEE Trans. Circuits Syst.*, vol. 3, pp. 229–234, 1977.

[4] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[5] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[6] K. Steiglitz and L. E. McBride, "A technique for identification of linear systems," *IEEE Trans. Automat. Contr.*, vol. 10, pp. 461–464, 1965.

[7] H. Fan and M. Doroslovacki, "On 'global convergence' of Steiglitz–McBride adaptive algorithm," *IEEE Trans. Circuits Syst.*, vol. 40, no. 2, pp. 73–87, 1993.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1991.

[9] S. Ruegsegger and N. Imlig, "HDSL—Echokompensator," Institute for Signal- and Information-Processing, ETH-Zurich, Switzerland, Tech. Rep. 6505, 1993.

[10] R. A. Roberts and C. T. Mullis, *Digital Signal Processing*. MA: Addison-Wesley, 1987.

[11] A. H. Gray and J. D. Markel, "Digital lattice and ladder filter synthesis," *IEEE Trans. Audio and Electroacoust.*, vol. 21, no. 6, pp. 491–500, 1973.

[12] R. E. Blahut, *Fast Algorithms for Digital Signal Processing*. MA: Addison-Wesley, 1987.

August N. Kaelin received the diploma and the Dr. Sc. Techn. degrees in electrical engineering from the Swiss Federal Institute of Technology (ETH) in 1983 and 1990, respectively.

In 1983, he joined the Institute for Signal and Information Processing at ETH where he is currently an Assistant Professor. His general interests are analog and digital filters, estimation and detection theory, and system identification. His specific research areas include various aspects of adaptive filters and its application in communication systems, hearing aids, and active noise control.
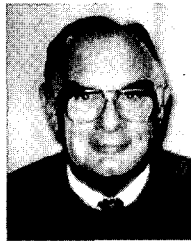
Allen G. Lindgren (M'57–SM'70) received the B.S.E.E. degree from Clarkson University and the M.S. and Ph.D. degrees from the University of Connecticut.

He is currently Professor of Electrical Engineering at the University of Rhode Island, Kingston, where he served as Department Chairman from 1984 to 1990. His research interests are signal and system theory. He has worked on problems in nonlinear systems, signal processing, control theory, system identification and computer aided imaging. He has more than 50 publications in these areas and has directed more than 25 research dissertations. He is a consultant to industry and government laboratories and spent a sabbatical leave with the US Navy Underwater Systems Center at Newport, RI. He has been a visiting professor at Brown University (1970) and the Swiss Federal Institute of Technology at Zurich (1979, 1984, and 1990) with other research visits to ETH in the summers of 1991–1993.

Dr. Lindgren has received awards for teaching at the University of Connecticut and University of Rhode Island. He is a member of Eta Kappa Nu, Sigma Xi, Past Student Branch Counselor, and Past President, and Executive Officer of the IEEE Providence Section. He also serves on IEEE technical committees and is a reviewer.

George S. Moschytz (F'78) was born on April 18, 1934.

He has been with RCA Labs Zurich from 1960 to 1962 and Bell Telephone Labs from 1962 to 1972. Since 1973, he has been Professor and Director of the Institute for Signal & Information Processing, SFIT. His current interests include analog, digital, and adaptive filters (switched-capacitor, switched-current, and current-mode); neural networks; analysis of speech, acoustical (e.g., alarm) signals, electromyograms, and other biosignals.

Dr. Moschytz received the Outstanding Paper Award at the Electronic Components Conference in 1969. His books include *Linear Integrated Networks: Fundamentals* (1974) and *Design* (1975) (Van Nostrand Reinhold); *Active Filter Design Handbook* (Wiley, 1981), *MOS Switched-Capacitor Filters: Analysis and Design*, (IEEE Press Book, 1984). He is the President of the Swiss Chapter on Digital Communications, a member of the Swiss Academy of Engineering Sciences, and Eta Kappa Nu Honor Society. Has held several terms in the Adcom of the IEEE Circuits and Systems Society, the Editorial Board of the PROCEEDINGS OF THE IEEE, and was an Associate Editor of the *IEEE Circuits and Systems Magazine* and the associate editor of several other technical journals.