Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments, and Application to Acoustic Echo Cancellation

André Gilloire and Martin Vetterli, *Senior Member, IEEE*

Abstract—**Adaptive filtering in subbands is a new technique for the real-time identification of large impulse responses like the ones encountered in acoustic echo cancellation. This technique generally allows computational savings as well as better convergence behavior.**

We give first an exact analysis of the critically subsampled two band modelization scheme. We demonstrate that adaptive cross-filters between the subbands are necessary for modelization with small output errors; moreover, we show that perfect reconstruction filter banks can yield exact modelization. We extend those results to the critically subsampled multiband schemes, and we show that important computational savings can be achieved by using good quality filter banks. Then we consider the problem of adaptive identification in critically subsampled subbands, and we derive an appropriate adaptation algorithm. We give a detailed analysis of the computational complexity of all the discussed schemes, and we verify experimentally the theoretical results that we have obtained. Finally, we discuss the adaptive behavior of the subband schemes that we have tested. We generally observe some degradation of the convergence performance in comparison with conventional schemes; however, the overall performance could be acceptable in practical use.

I. Introduction

ADAPTIVE identification of linear systems modelized by their impulse responses has been studied extensively, and a number of efficient algorithms has been proposed for that purpose [1]–[3]. The adaptive transversal filtering algorithms yield an estimate of the impulse response in a direct form, and the output of the model is obtained as the convolution of the system input signal by the estimated impulse response. The adaptation algorithms used for real-time identification are generally of a gradient type (LMS), and are known to achieve good performances in applications to communications systems like adaptive equalization and adaptive echo cancellation for data transmission. However, in some new applications like acoustic echo cancellation, several problems arise that limit the efficiency of those "classical" algorithms.

Those problems have motivated a new approach: adaptive filtering in subbands (e.g., [6]–[12]), with the double purpose of reducing the computational complexity and of improving the convergence speed of the algorithm. In adaptive filtering in subbands, the input signal and the system output signal are split into adjacent frequency subbands by analysis filter banks; then each subband signal is subsampled, and the adaptive filtering algorithm is applied to those signals. The identified impulse response is the system impulse response filtered by the corresponding subband filter. The subsampling leads to a greater computational efficiency than the usual full-band scheme, and the processing in separate subbands makes better convergence speed possible in the case of the LMS algorithm, since in each subband the adaptation step size can be matched to the energy of the input signal in that band.

Because of the subsampling, aliased versions of the input may appear in the output, especially if the filter bank is critically sampled [8], [13]. Solutions to this problem have been proposed, including spectral gaps between subbands [14], oversampling of the filter banks [11], and adaptive cross-terms between the subbands [9]. Spectral gaps may impair the subjective quality (especially when the number of subbands is large); oversampling can be costly if the oversampling ratio is large, and the speed of adaptation of the cross-terms can be problematic. Some frequency domain adaptive filtering schemes [15]–[17] which are based on transmultiplexers or short-time Fourier transforms, have a close connection with the subband adaptive filtering scheme discussed in the present paper. Note that these schemes are often highly oversampled (by a factor of 2 to 4), thus aliasing components are small enough to be practically neglected. The main reason for considering these schemes was their high computational efficiency.

In our presentation, we will consider essentially critically subsampled systems; we will emphasize the multirate nature of these systems, thus focusing on issues like aliasing suppression. In order to do so, we will use tools

Manuscript received March 11, 1989; revised June 1, 1991. This work was supported in part by the U.S. National Science Foundation under Grants CDR-84-21402 and MIP-88-08-277.

A. Gilloire is with the Centre National d'Etudes des Télécommunications, CNET LAA/TSS/CMC, 22301 Lannion Cedex, France.

M. Vetterli is with the Department of Electrical Engineering, Center for Telecommunications Research, Columbia University, New York, NY 10027.

IEEE Log Number 9201055.

which have been introduced in the context of multirate filter banks [18], [19], [20]. Note also that such multirate systems with time-invariant input/output behavior have a close connection with block processing schemes [21], [22]. We will also derive an adaptive identification algorithm consistent with the critical subsampling framework. In that sense, the results we present are extensions to the adaptive case of some recent results on alias-free block-processing schemes.

As is well known [4], the acoustic echo appears when one or both ends in a communication use audio terminals with acoustic feedback from the loudspeaker to the microphone, like hands-free telephone sets and teleconference systems (Fig. 1). The speech $x(t)$ from the remote terminal is diffused into the room by a loudspeaker, the local user's speech $n(t)$ is picked up by a microphone and sent to the remote user. Because of the acoustic transmission path from the loudspeaker to the microphone (including the reflections on walls, furniture, and people inside the room), some remote speech $y(t)$ is fed back to the remote user. Basically, an echo canceller is an adaptive filter $h(t)$ fed by the speech $x(t)$; $h(t)$ modelizes the acoustic echo path and yields at its output a replica $\hat{y}(t)$ of the "echo" signal $y(t)$ which is subtracted from it.

The length of the acoustic echo path impulse responses (often several hundred milliseconds) leads to adaptive filters with very large numbers of taps (several thousands). Another difficulty arises from the nature of speech $x(t)$, which is a badly suited signal for identification because of its nonflat spectrum. The corresponding spread of the eigenvalues in the autocorrelation matrix of the signal slows down the adaptation process in the LMS case [5]. Note also that since the acoustic echo path is time varying, $h(t)$ has to be adjusted continuously in time during the communication. The tracking capability is therefore an important characteristic of adaptive identification algorithms for acoustic echo cancellation.

The experimental results that we will present come mainly from two subband schemes. Although the computational efficiency generally increases with the number of subbands, two subband schemes are practically useful since they are structurally well suited for transmission of wide-band speech as well as telephone speech. For example, they could be directly embedded within subband speech coding schemes [34]. Another reason to consider two subband schemes is that practical systems should have short transmission delays, since the overall delay is presently a critical point in network planning considerations [35].

The main points addressed in the paper are as follows.

Section II presents a detailed analysis of the two-band scheme with critical subsampling by two for modelization of linear systems. It is shown how aliasing can be perfectly cancelled by using cross-filters.

In Section III, we generalize the analysis to the multiband case (critically sampled). We show that, under weak conditions on the filters, nearest neighbor cross-filters will suppress aliasing as well.

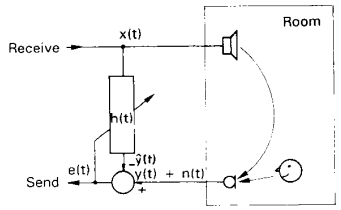

Fig. 1. Origin of the acoustic echo. A basic echo canceller $h(t)$ is shown.

We consider the problem of adaptive identification in Section IV. As a first solution, we apply directly the conventional LMS adaptive algorithm to the static structure developed in Sections II and III. Then we derive a specific adaptation algorithm in the purpose to avoid the overdetermination problem inherent in the first solution, and we compute a convergence bound for the adaptation step size of this algorithm.

Section V is concerned with computational complexity of the proposed adaptive identification schemes.

Then, Section VI applies the proposed schemes to the problem of acoustic echo cancellation, both with synthetic and real signals and systems. Experimental results on convergence rates and asymptotic errors as well as on tracking capability illustrate the different schemes. Besides, some comparisons are made with an oversampled two subband scheme. These experimental results are discussed, and some qualitative remarks are made that explain peculiarities of the adaptive processes in the proposed schemes.

II. Two Subbands Scheme

A two channel subband identification scheme is depicted in Fig. 2. We assume in the following that z transforms of all signals and filters exist on the unit circle. Thus, from Fig. 2, we see that the signal $X(z)$ is filtered by the system $S(z)$; the system output has a "noise signal" $N(z)$ added to it (this noise signal can actually be the useful signal, as in acoustic echo cancellation where it corresponds to the near end speech), and it is then split by an analysis filter bank $(H_0(z), H_1(z))$ and subsampled to yield the two "system subband signals" $Y_0(z)$ and $Y_1(z)$. The identification path first splits the signal $X(z)$ by an identical analysis filter bank, and then models the system in the subband domain by a matrix $\boldsymbol{C}_m(z)$ to yield the two estimated system subband signals $\hat{Y}_0(z)$ and $\hat{Y}_1(z)$. The subband error signals are obtained as $E_i(z) = Y_i(z) - \hat{Y}_i(z)$, $i = 0, 1$. The total error $E(z)$ is obtained after passing the subband error signals through a synthesis filter bank $(G_0(z), G_1(z))$. The model matrix $\boldsymbol{C}_m(z)$ has to be adjusted adaptively so as to minimize the total error $E(z)$; in practice, it is adjusted so as to minimize the subband error signals. This point will be discussed further in this section.

Assume first that we have perfect half-band filters, both in the analysis and in the synthesis filter banks. In that case, the subsampling is perfectly valid (no aliased component is folded back) and the two channels can be treated

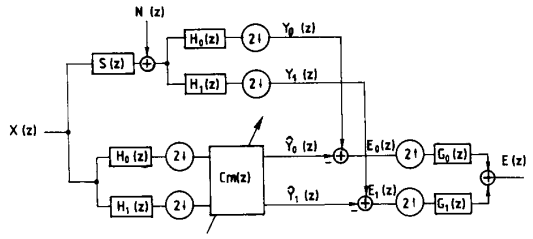

Fig. 2. Two-channel subband identification scheme. The system $S(z)$ is modeled by $\boldsymbol{C}_m(z)$ in the subband domain. We will refer to the upper half as the "system path," while the lower left part is often called the "identification path."

completely independently for the identification process. After the separate identification of the channels, upsampling and interpolation by the perfect half-band filters yield an output which is identical with the one that would have been produced by a full-band identification scheme. In that ideal case, the matrix $\boldsymbol{C}_m(z)$ is diagonal, with the lower and higher frequency halves of the system as diagonal elements.

In the nonideal case, that is when the filters are only approximations of ideal half-band filters, it will be shown that $\boldsymbol{C}_m(z)$ cannot be diagonal, and that the cross-terms are necessary for a correct system identification, at least if one insists on critical sampling.

Note that this section presents a deterministic analysis of the system, where we assume that the system is known; that will allow us to derive exactly the structure of the solution.

In the analysis below, we will use the modulation domain approach [19], because it is well suited for interpretation. This presentation will be succinct, and we refer the reader to the literature for details on multirate filter banks [18], [19], [23].

A. Modulation Domain Analysis

Let us use the modulation expansion of the signals and filters [19]. Defining

$$\boldsymbol{x}_m(z) = [X(z) \quad X(-z)]^T \tag{1}$$

we define similarly

$$\boldsymbol{H}_m(z) = \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} \tag{2}$$

as well as

$$\boldsymbol{S}_m(z) = \begin{bmatrix} S(z) & 0 \\ 0 & S(-z) \end{bmatrix}. \tag{3}$$

Then, the output of the system path is equal to [9]

$$\begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} H_0(z^{1/2}) & H_0(-z^{1/2}) \\ H_1(z^{1/2}) & H_1(-z^{1/2}) \end{bmatrix} \cdot \begin{bmatrix} S(z^{1/2})X(z^{1/2}) \\ S(-z^{1/2})X(-z^{1/2}) \end{bmatrix} = \frac{1}{2} \boldsymbol{H}_m(z^{1/2}) \boldsymbol{S}_m(z^{1/2}) \boldsymbol{x}_m(z^{1/2}). \tag{4}$$

The output of the identification path is

$$\begin{bmatrix} \hat{Y}_0(z) \\ \hat{Y}_1(z) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} C_{0,0}(z) & C_{0,1}(z) \\ C_{1,0}(z) & C_{1,1}(z) \end{bmatrix} \begin{bmatrix} H_0(z^{1/2}) & H_0(-z^{1/2}) \\ H_1(z^{1/2}) & H_1(-z^{1/2}) \end{bmatrix} \cdot \begin{bmatrix} X(z^{1/2}) \\ X(-z^{1/2}) \end{bmatrix} = \frac{1}{2} \boldsymbol{C}_m(z) \boldsymbol{H}_m(z^{1/2}) \boldsymbol{x}_m(z^{1/2}). \tag{5}$$

The subband error signals are zero if

$$\boldsymbol{C}_m(z^2) \boldsymbol{H}_m(z) = \boldsymbol{H}_m(z) \boldsymbol{S}_m(z). \tag{6}$$

To interpret this result, we make the following choice for the filters in the banks:

$$H_0(z) = H(z) \tag{7a}$$

$$H_1(z) = H(-z) \tag{7b}$$

where $H(z)$ is a "good" half-band low-pass filter (therefore, $H(-z)$ is a "good" half-band high-pass filter). This choice is usual in subband coding [18], even if it does only lead approximately to a perfect reconstruction system. It is called the classical QMF solution. This choice is made for clarity of explanation and because it is usually good in practice. The inverse of $\boldsymbol{H}_m(z)$ is

$$[\boldsymbol{H}_m(z)]^{-1} = \frac{1}{\det \boldsymbol{H}_m(z)} \begin{bmatrix} H(z) & -H(-z) \\ -H(-z) & H(z) \end{bmatrix}. \tag{8}$$

It can be shown that for reasonably long filters [32] $\det \boldsymbol{H}_m(z)$ is approximately equal to

$$\det \boldsymbol{H}_m(z) \approx \alpha z^{-L_f+1} \tag{9}$$

where L_f is the length of the filter $H(z)$ (which is assumed even). Thus, the inverse of $\boldsymbol{H}_m(z)$ (within a delay) can be approximated by

$$\boldsymbol{G}_m(z) = z^{-L_f+1} [\boldsymbol{H}_m(z)]^{-1} \approx \frac{1}{\alpha} \begin{bmatrix} H(z) & -H(-z) \\ -H(-z) & H(z) \end{bmatrix}. \tag{10}$$

A causal $\boldsymbol{C}_m(z^2)$ can be obtained as (from (6)) [9]

$$z^{-L_f+1} \boldsymbol{C}_m(z^2) \approx \boldsymbol{H}_m(z) \boldsymbol{S}_m(z) \boldsymbol{G}_m(z) \approx \frac{1}{\alpha} \begin{bmatrix} H^2(z)S(z) - H^2(-z)S(-z) & H(z)H(-z)[S(-z) - S(z)] \\ H(z)H(-z)[S(z) - S(-z)] & H^2(z)S(-z) - H^2(-z)S(z) \end{bmatrix}. \tag{11}$$

Since z^{-L_f+1} as well as the right-hand side of (11) are odd functions of z, $\boldsymbol{C}_m(z^2)$ is indeed an even function of z. From (11), it is immediate that $\boldsymbol{C}_m(z)$ is diagonal only if either of the conditions below are satisfied [9], [11]:

i) $H(z)H(-z) = 0$: that is, $H(z)$ has infinite attenuation above a quarter and below three quarters of the sampling frequency.

ii) $S(z) - S(-z) = 0$: the system $S(z)$ has only even indexed samples.

Obviously, ii) does not correspond to a general physical system.

The condition i), together with the requirement for perfect transmission of the useful signal, shows that $H(z)$ has to be a perfect half-band filter in order to get a diagonal $\boldsymbol{C}_m(z)$. That is, $\boldsymbol{C}_m(z)$ cannot be the correct model matrix of a general physical system if it is diagonal, and there have to be cross-terms between the channels in order to identify the system correctly in the subbands.

Recall that $\boldsymbol{C}_m(z)$ in (11) is only an approximation, because $\boldsymbol{G}_m(z)$ was approximated in (10). The effect of such an approximate solution will be discussed below. Note that the off-diagonal terms in (11) have a factorized form which can be used in an implementation, and that if $H(z)$ is a "good" half-band filter, the energy present in the off-diagonal part becomes generally small compared to the diagonal terms (because $H(z)H(-z)$ is small except on a small frequency range around a quarter of the sampling frequency).

B. Remarks on the Two Band Scheme

Still within the deterministic framework used so far, we will discuss some effects and tradeoffs.

1) Absence of Cross-Terms: Assume that the off-diagonal terms in (11) are zero, and the diagonal terms are unchanged. Then, with $G_0(z) = H(z)$ and $G_1(z) = -H(-z)$, it can be verified [9] that the error signal at the output contains mainly aliased input $X(z)$. This is exactly what was noted experimentally in [8].

2) Perfect Aliasing Cancellation: One can verify that using an aliasing cancellation filter bank (with $G_0(z) = H(z)$ and $G_1(z) = -H(z)$) indeed cancels aliasing, i.e., the output error does not contain aliased input. It is easily verifiable that with $\boldsymbol{C}_m(z)$ as in (11), the output error equals

$$E(z) = [H^2(z) - H^2(-z)] \cdot [z^{-L_f+1} - (H^2(z) - H^2(-z))]S(z)X(z). \quad (12)$$

There is no aliasing term $X(-z)$ in the product on the right side. The second term of this product, which should be very small for well-designed filters like QMF's [32], shows that $S(z)X(z)$ cannot be perfectly cancelled anymore.

3) Perfect Reconstruction Filter Banks: Instead of using an approximate inverse of the analysis bank in (10), one can write an exact inverse (within a delay) given that the filter bank is a perfect reconstruction one. Various perfect reconstruction filter banks are possible (having linear phase [29] or being lossless [23]).

In the lossless two channel case, the analysis filters obey [33]: $H_0(z) = H(z)$ and $H_1(z) = z^{-L_f+1}H(-z^{-1})$, while the synthesis filters are given by: $G_0(z) = z^{-L_f+1}H(z^{-1})$ and $G_1(z) = H(-z)$. $H(z)$ is a prototype filter (of length L_f even) such that the filter bank is lossless, that is,

$$H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = 1. \quad (13)$$

The model matrix $\boldsymbol{C}_m(z)$ can be derived straightforwardly. It is similar to the QMF case (11). The main difference is that $\boldsymbol{C}_m(z)$ is a perfect model of the system, and, thus, the error can be made theoretically exactly zero. This holds also as far as the reconstruction of $N(z)$ is concerned.

When the filter bank is lossless, the L_2 norm of the total output error $E(z)$ is equal to the L_2 norm of the error vector $[E_0(z) \quad E_1(z)]^T$, because the filter bank can be seen as a generalized rotation that preserves L_2 norms. Therefore it will be convenient and advantageous to use the subband errors in the adaptive algorithm, rather than the total error. Note that a QMF bank with sufficiently selective filters is not far from being lossless, and thus the same argument applies in that case also.

III. Multiband Scheme

Assume now that we generalize the system so as to divide signals into N bands, each subsampled by N. Each filter in the analysis bank is chosen as a bandpass filter of bandwidth f_s/N (if the filters are real, they will have two conjugate parts of bandwidth $f_s/2N$ each). Furthermore, we assume that the filters are selective enough so that they overlap only with adjacent filters:

$$H_i(z) \cdot H_j(z) \approx 0, \qquad |i - j| > 1. \quad (14)$$

Since sufficiently selective filters can always be obtained if the complexity can be afforded, the product in (14) can be made reasonably close to zero.

The amplitude response for such a filter bank is schematically shown in Fig. 3. A convenient class of such filters which has been studied for subband coding of speech is the class of pseudo-QMF filters [24]–[27]. The ith filter of such a bank is obtained by cosine modulation of a low-pass prototype filter with cutoff frequency $f_s/4N$.

The matrix $\boldsymbol{H}_m(z)$ (see (2)) is now of size $N \times N$ and contains the filters and their modulated versions (by the Nth root of unity $W = e^{-j2\pi/N}$):

$$\boldsymbol{H}_m(z) = \begin{bmatrix} H_0(z) & H_0(Wz) & \cdots & H_0(W^{N-1}z) \\ H_1(z) & H_1(Wz) & \cdots & H_1(W^{N-1}z) \\ \vdots & \vdots & \cdots & \vdots \\ H_{N-1}(z) & H_{N-1}(Wz) & \cdots & H_{N-1}(W^{N-1}z) \end{bmatrix}. \quad (15)$$

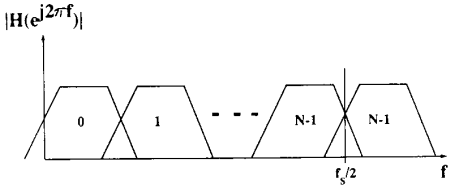

Fig. 3. Amplitude response of a real filter bank of size N where only adjacent filters overlap.

Similarly to (3), $S_m(z)$ is a diagonal matrix of the form

$$S_m(z) = \begin{bmatrix} S(z) & \vdots & \cdots & \vdots \\ & S(Wz) & \cdots & \\ \vdots & \vdots & \cdots & S(W^{N-1}z) \end{bmatrix}. \quad (16)$$

From (6), the subband error signals are zero if

$$C_m(z^N) = H_m(z) S_m(z) [H_m(z)]^{-1}. \quad (17)$$

For the sake of discussion, assume for now that the analysis filter bank is obtained from a prototype low-pass filter $H(z)$ by complex modulation with the Nth root of unity. Similarly, a synthesis bank is obtained from a prototype $G(z)$ and produces (approximately) a perfect reconstruction system (this can only be approximately true [20], [28]).

$G_m(z)$ can then be considered as an (approximate) inverse of $H_m(z)$ (within a delay):

$$H_m(z) \cdot G_m(z) \approx z^{-K} I. \quad (18)$$

Furthermore, both $H(z)$ and $G(z)$ are sufficiently selective so that (14) holds also between $H_i(z)$ and $G_j(z)$ (the difference $|i - j|$ is taken *modulo* N). Then, it is easy to verify that the product (from (17))

$$C_m(z^N) \approx z^K H_m(z) S_m(z) G_m(z) \quad (19)$$

is actually tridiagonal:

$$C_m(z^N) \approx z^K \begin{bmatrix} C_{0,0}(z) & C_{0,1}(z) & 0 & \cdots & 0 & C_{0,N-1}(z) \\ C_{1,0}(z) & C_{1,1}(z) & C_{1,2}(z) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ C_{N-1,0}(z) & 0 & 0 & \cdots & C_{N-1,N-2}(z) & C_{N-1,N-1}(z) \end{bmatrix} \quad (20a)$$

with

$$C_{l,l}(z) \approx z^K \sum_{i=0}^{N-1} S(W^{i-l}z) H(W^i z) G(W^i z) \quad (20b)$$

$$C_{l,l+1}(z) \approx z^K \sum_{i=0}^{N-1} S(W^{i-l}z) H(W^i z) G(W^{i+1}z) \quad (20c)$$

$$C_{l,l-1}(z) \approx z^K \sum_{i=0}^{N-1} S(W^{i-l}z) H(W^i z) G(W^{i-1}z) \quad (20d)$$

and similar definitions of $C_{0,N-1}(z)$ and $C_{N-1,0}(z)$ (these terms appear because the lowest and highest frequency filters are neighbors in the complex filter case).

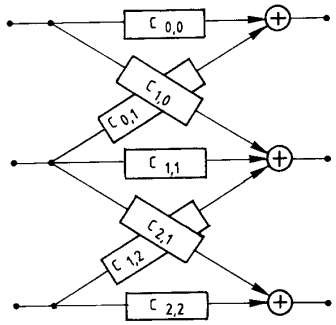

Fig. 4. Identified system in the critically sampled multiband real case ($N = 3$).

The important point here is that the matrix $C_m(z)$ is not a full matrix, but has only cross-terms from adjacent bands.

Note that $C_{i,j}(z)$ are functions of z^N only (one can verify that $C_{i,j}(W^k z) = C_{i,j}(z)$). Thus, the length of the corresponding filters is of the order of

$$L_C = (L_H + L_G + L_S)/N. \quad (21)$$

Note that because the cross-terms $C_{l,l+1}(z)$ and $C_{l,l-1}(z)$ are obtained from the product of adjacent filters, they can usually be approximated by shorter filters than the main terms.

In the case of real filter banks, the convenience of pseudo-QMF's is that a good approximation of the inverse filter bank (therefore of the inverse $H_m(z)$) is known. The approximation can even be perfect in certain cases (see [29] for details). Going through a similar development as in the complex filter bank case ((19)–(21) above) one can verify that, given that only adjacent filters overlap, the matrix $C_m(z)$ is again tridiagonal ($C_{0,N-1}(z)$ and $C_{N-1,0}(z)$ are no longer present in the real filter case).

Thus, the model system $C(z)$ is a network as depicted in Fig. 4 where there are only nearest neighbor cross-terms. Failure to include these cross-terms will lead to the apparition of aliased versions of the input in the output.

IV. Adaptive Identification Algorithms

In Sections II and III we have analyzed the problem of linear system modeling with critically subsampled subbands schemes. Now, the model matrix $C_m(z)$ (11), (19) has to be made adaptive in order to implement the identification of the unknown system $S(z)$. In this section, we discuss two adaptive algorithms which can be used in that purpose.

A. Direct Adaptation of the Filters in the Model Matrix

The most direct way to adapt the matrix $\boldsymbol{C}_m(z)$ is to use for each of its terms an adaptive filter of proper length (theoretically $L_c = L_f + L_s/2 - 2$). One could then take advantage of the subband splitting of the input signal $x(t)$, which should lead to a better convergence of the LMS algorithm because the eigenvalue spread is smaller in the subbands. This assumption is valid only for the diagonal terms of the model matrix (see experimental results in Section VI).

The corresponding implementation of adaptive filtering in subbands is shown schematically in Fig. 5 for the two subband case. The main filter and the cross-filter which contribute to each subband i, $i = 0, 1$ are adapted by a classical algorithm, for example, the LMS (the adaptation path is symbolized by the dotted lines) which uses at each "block" of two samples the error $e_i(m)$ at the output of the corresponding subband (m is the block number). This algorithm can be generalized straightforwardly to the multiband case with the limitation to adjacent cross-filters, according to the simplified scheme described in Section III.

Considering the two subband case, we note that the direct adaptation of the filters in the model matrix leads to the separate identification of four filters of length $L_c = L_f + L_s/2 - 2$, although the length of the unknown system response S is L_s. This corresponds indeed to an overdetermined problem, for which the optimum solution (i.e., $\hat{\boldsymbol{C}}(z) = \boldsymbol{C}(z)$ as given by (11)) may be hard to find. This overdetermination explains probably the convergence difficulties observed in the experiments (see Section VI).

B. Adaptation of the System-Dependent Components

To overcome this problem, another adaptation scheme is proposed which does not exhibit overdetermination and thus yields better performances. This scheme tries to identify directly the system impulse response, in the frame of the subband structure.

Denoting by $\hat{S}$ the estimated system impulse response, one tries to minimize the power of the output error $e(t)$ relatively to the samples of $\hat{S}$. As noted in Section II, if the synthesis filter matrix $\boldsymbol{G}_m(z)$ is paraunitary, the power of the output error is equal to the sum of the powers of the subband errors at its inputs; then the minimization of the output error power is equivalent to the minimization of this sum. We define an optimality criterion as follows. Setting:

$$J(\hat{S}) = \sum_{i=0}^{N-1} \langle e_i^2(m) \rangle \tag{22}$$

where $\langle \cdot \rangle$ denotes some averaging operation (e.g., mathematical expectation), we have the optimal solution $\hat{S} = S^*$ iff $J(\hat{S})$ minimum. This corresponds indeed to

$$\frac{\partial J(\hat{S})}{\partial \hat{s}_k} = 0, \qquad k = 0, 1, \cdots, L_s - 1. \tag{23}$$

S^* is equal to the unknown system impulse response S,

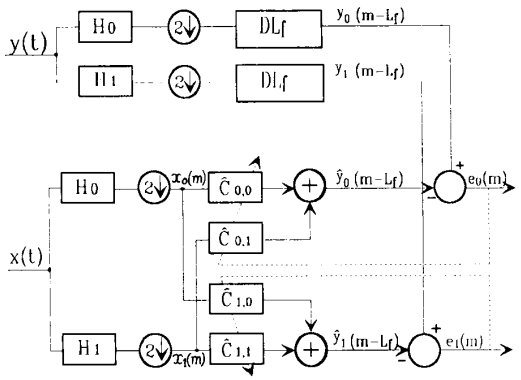

Fig. 5. Two subband adaptive identification with direct adaptation of the filters in the model matrix (DL_f means L_f samples delay).

since the subband errors are linear functions of the impulse response samples $\hat{s}_k$ and because they are equal to zero when the model matrix satisfies (11) or (19).

In the sequel, we will develop the computations for the two-band case. We discuss more briefly the multiband case in the Appendix I.

1) LMS Adaptation Algorithm: We specialize here the computation to the LMS case. We use then a modified criterion:

$$\tilde{J}(\hat{S}) = e_0^2(m) + e_1^2(m). \tag{24}$$

Taking the partial derivatives $\tilde{J}(\hat{S})$ relatively to the samples of $\hat{S}$, we get the components of the "instantaneous gradient" vector. Then, the LMS adaptation algorithm will have the general form:

$$\hat{s}_k(m) = \hat{s}_k(m-1) - \mu(m)\left[2e_0(m)\frac{\partial e_0(m)}{\partial \hat{s}_k} + 2e_1(m)\frac{\partial e_1(m)}{\partial \hat{s}_k}\right], \quad k = 0, 1, \cdots L_s - 1. \tag{25}$$

2) Components of the Gradient: Using (5)–(8) and (9)–(11) (where we assume that exact equalities hold), we can get expressions for the partial derivatives of the subband errors relatively to the even and odd coefficients of the estimated impulse response $\hat{S}$. To obtain closed-form expressions, we use the polyphase decomposition of the filters responses, i.e.,

$$H(z) = H_e(z^2) + z^{-1}H_o(z^2) \tag{26}$$

and we define "filtered input signals" as follows:

$$X_{i,c}(z) = 2H_e(z)H_o(z)X_i(z), \qquad i = 0, 1 \tag{27}$$

$$X_{i,s}(z) = 2[H_e^2(z) + z^{-1}H_o^2(z)]X_i(z), \qquad i = 0, 1 \tag{28}$$

$$X_{i,d}(z) = 2[H_e^2(z) - z^{-1}H_o^2(z)]X_i(z), \qquad i = 0, 1 \tag{29}$$

where $X_0(z)$ and $X_1(z)$ result from the subband splitting of $X(z)$. Then we get for $k = 2p$:

$$\frac{\partial e_0(z)}{\partial \hat{s}_{2p}} = -2z^{-p-1}X_{0,c}(z) \tag{30}$$

$$\frac{\partial e_1(z)}{\partial \hat{s}_{2p}} = -2z^{-p-1}X_{1,c}(z) \tag{31}$$

and for $k = 2p + 1$:

$$\frac{\partial e_0(z)}{\partial \hat{s}_{2p+1}} = -z^{-p-1}[X_{0,s}(z) - X_{1,d}(z)] \tag{32}$$

$$\frac{\partial e_1(z)}{\partial \hat{s}_{2p+1}} = -z^{-p-1}[X_{0,d}(z) - X_{1,s}(z)]. \tag{33}$$

3) Expressions of the Convolutions: We have to implement the convolutions in the model matrix according to (11). Using the polyphase decomposition of the estimated system response, i.e.,

$$\hat{S}(z) = \hat{S}_e(z^2) + z^{-1}\hat{S}_o(z^2) \tag{34}$$

we get for the estimated system outputs in the subbands:

$$\hat{y}_0(z) = z^{-1}\{2\hat{S}_e(z)X_{0,c}(z) + \hat{S}_o(z)[X_{0,s}(z) - X_{1,d}(z)]\} \tag{35}$$

$$\hat{y}_1(z) = z^{-1}\{2\hat{S}_e(z)X_{1,c}(z) + \hat{S}_o(z)[X_{0,d}(z) - X_{1,s}(z)]\}. \tag{36}$$

For each "block" of two samples, we have thus to implement four convolutions on $L_s/2$ samples, and we have to adapt L_s coefficients. Since the adaptation of each coefficient requires 2 multiplications and 2 additions, the total computational load (convolution + adaptation) is equivalent to the one of the full-band LMS transversal filter.

C. Convergence Condition for the Nonoverdetermined Algorithm

In this subsection, we will limit our analysis to the two subband case for more tractability. The convergence of the algorithm (25) towards the optimal solution $S^* = S$ is controlled by the adaptation step size μ. Using a well-known technique [2, ch. 7], we study the convergence to zero of the coefficient error vector:

$$V(m) = S - \hat{S}(m). \tag{37}$$

Assuming $\mu(m) = \mu$ for the analysis, we get from (25)

$$V(m) = V(m-1) + \mu\left[2e_0(m)\frac{\partial e_0(m)}{\partial \hat{S}} + 2e_1(m)\frac{\partial e_1(m)}{\partial \hat{S}}\right]. \tag{38}$$

It is convenient there to split $V(m)$ in its even and odd components $V_e(m)$ and $V_o(m)$, as we did before to derive the practical form of the adaptation algorithm. Using (30)–(33) and (35), (36) and assuming for simplicity that the output additive noise is zero, we get from (38):

$$\begin{aligned} V_e(m) = V_e(m-1) &- 4\mu[2X_{0,c}X_{0,c}^T V_e(m-1) \\ &+ X_{0,c}(X_{0,s}^T - X_{1,d}^T)V_o(m-1)] \\ &-4\mu[2X_{1,c}X_{1,c}^T V_e(m-1) \\ &+ X_{1,c}(X_{0,d}^T - X_{1,s}^T)V_o(m-1)] \end{aligned} \tag{39}$$

$$\begin{aligned} V_o(m) = V_o(m-1) &- 2\mu[2(X_{0,s} - X_{1,d})X_{0,c}^T V_e(m-1) \\ &+ (X_{0,s} - X_{1,d})(X_{0,s}^T - X_{1,d}^T)V_o(m-1)] \\ &-2\mu[2(X_{0,d} - X_{1,s})X_{1,c}^T V_e(m-1) \\ &+ (X_{0,d} - X_{1,s})(X_{0,d}^T - X_{1,s}^T)V_o(m-1)] \end{aligned} \tag{40}$$

where $X_{0,c}$, $X_{1,c}$, $X_{0,s}$, $X_{1,d}$, $X_{0,d}$, and $X_{1,s}$ denote the vectors of size $L_s/2$ containing the samples of the filtered input signals defined by (27)–(29) up to the current "block" m, and ^{T} denotes transposition of vectors and matrices. Note that the unit delays in (30)–(33) and (35), (36) have been dropped out, and the delays on the system path have been changed into $z^{-L_f/2+1}$ accordingly.

Equations (39) and (40) can be grouped in a first-order recurrent system as follows:

$$V_e(m) = [\boldsymbol{I}_{L_s/2} - \mu\boldsymbol{A}(m)]V_e(m-1) - \mu\boldsymbol{B}(m)V_o(m-1) \tag{41i}$$

$$V_o(m) = -\mu\boldsymbol{B}^T(m)V_e(m-1) + [\boldsymbol{I}_{L_s/2} - \mu\boldsymbol{D}(m)]V_o(m-1) \tag{41ii}$$

where $\boldsymbol{I}_{L_s/2}$ is the $L_s/2 \times L_s/2$ identity matrix, and

$$\boldsymbol{A}(m) = 8[X_{0,c}X_{0,c}^T + X_{1,c}X_{1,c}^T] \tag{42}$$

$$\boldsymbol{B}(m) = 4[X_{0,c}(X_{0,s}^T - X_{1,d}^T) + X_{1,c}(X_{0,d}^T - X_{1,s}^T)] \tag{43}$$

$$\boldsymbol{D}(m) = 2[(X_{0,s} - X_{1,d})(X_{0,s}^T - X_{1,d}^T) + (X_{0,d} - X_{1,s})(X_{0,d}^T - X_{1,s}^T)] \tag{44}$$

are $L_s/2 \times L_s/2$ matrices which depend on the filtered input signals up to the current block m. Let

$$\boldsymbol{\Phi} = \begin{bmatrix} E[\boldsymbol{A}(m)] & E[\boldsymbol{B}(m)] \\ E[\boldsymbol{B}^T(m)] & E[\boldsymbol{D}(m)] \end{bmatrix} \tag{45}$$

where $E[\cdot]$ denotes the mathematical expectation (we assume that $x(t)$, and consequently the filtered input signals, are stationary processes). Assuming that μ is "sufficiently small," we can use a classical approach [2]. We write for the system (41i–ii):

$$\begin{bmatrix} E[V_e(m)] \\ E[V_o(m)] \end{bmatrix} = [\boldsymbol{I}_{L_s} - \mu\boldsymbol{\Phi}]\begin{bmatrix} E[V_e(m-1)] \\ E[V_o(m-1)] \end{bmatrix} \tag{46}$$

where $\boldsymbol{I}_{L_s}$ is the $L_s \times L_s$ identity matrix.

It is shown in Appendix 2 that the symmetric block-matrix $\boldsymbol{\Phi}$ is positive definite. It can be shown by a classical reasoning [2] that the behavior of the mean coefficient error vector is governed by the eigenvalues of $\boldsymbol{\Phi}$, which are all strictly greater than zero. In particular, this vector converges exponentially to zero provided that $\mu < 1/\lambda_{\max}$, where $\lambda_{\max}$ is the largest eigenvalue of $\boldsymbol{\Phi}$.

As it is well known, this condition is not sufficient to insure the convergence of the MSE to its minimum. Using again the classical approach developed in [2], we can derive a convergence condition for the MSE. Using (35), (36), and the instantaneous squared error (24), we get

straightforwardly:

$$e_0^2(m) + e_1^2(m) = \frac{1}{2}\,[V_e^T(m) \quad V_o^T(m)] \begin{bmatrix} \boldsymbol{A}(m) & \boldsymbol{B}(m) \\ \boldsymbol{B}^T(m) & \boldsymbol{D}(m) \end{bmatrix} \begin{bmatrix} V_e(m) \\ V_o(m) \end{bmatrix}. \tag{47}$$

The rest of the analysis is then similar to the one developed in [2]. In particular, the MSE for a known coefficient error vector $[V_e^T(m) \quad V_o^T(m)]$ can be written as

$$E[e_0^2(m) + e_1^2(m)] = \frac{1}{2}\,[V_e^T(m) \quad V_o^T(m)]\,\Phi \begin{bmatrix} V_e(m) \\ V_o(m) \end{bmatrix} \tag{48}$$

and the mean coefficient error vector propagates according to (46). Then we get the bound for μ:

$$\mu < \frac{2}{\operatorname{tr} \Phi} = \mu_{\max}. \tag{49}$$

Let us now gain some further insight in the form of Φ for a particular case. With general QMF filter banks, the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{D}$ have complicated terms which depend on the polyphase components of the QMF filter H. Using the "trivial" QMF filte bank derived from $H(z) = 0.5(1 + z^{-1})$, we get much more tractable expressions:

$$\boldsymbol{A}(m) = X_{e,2m}X_{e,2m}^T + X_{o,2m}X_{o,2m}^T \tag{50}$$

$$\boldsymbol{B}(m) = X_{e,2m}X_{o,2m}^T + X_{o,2m}X_{e,2m-2}^T \tag{51}$$

$$\boldsymbol{D}(m) = X_{o,2m}X_{o,2m}^T + X_{e,2m-2}X_{e,2m-2}^T \tag{52}$$

where $X_{e,2m}$ and $X_{o,2m}$ denote, respectively, the even and odd polyphase components of the vector of L_s input samples of $x(t)$ at time $t = 2m$ (m is the block number).

Computing the expectations of (50)–(52) to get the matrix Φ may be a difficult task, since in general no simple relations exist between the autocorrelations and cross correlations of the polyphase components of a signal and its autocorrelation function. In the white noise case, however, the polyphase components $X_{e,2m}$ and $X_{o,2m}$ are uncorrelated, and (45) simplifies to

$$\Phi = \begin{bmatrix} 2\sigma_x^2 \boldsymbol{I}_{L_s/2} & 0 \\ 0 & 2\sigma_x^2 \boldsymbol{I}_{L_s/2} \end{bmatrix} = 2\sigma_x^2 \boldsymbol{I}_{L_s} \tag{53}$$

where σ_x^2 is the variance of the input signal $x(t)$. The upper bound $\mu_{\max}$ is then equal to $1/L_s\sigma_x^2$, i.e., half the upper bound for the adaptation step size of the standard LMS algorithm with a filter of length L_s.

V. Computational Complexity

Besides convergence behavior, reduced computational complexity is another attractive feature of adaptive filtering in subbands. We will demonstrate this point by comparing the computational complexity of a full band scheme with that of N band schemes, both in the critical and oversampled cases.

In this section, we are going to count multiplications per input sample in the various schemes, and it is understood that the number of additions has a very similar behavior as well.

The full-band filter with the LMS adaptation algorithm uses $2L_s + 1$ multiplications per input sample. This number is increased typically by 2 multiplications in the case of the normalized LMS algorithm.

The complexity of the two subband scheme described in Section II with the overdetermined adaptation algorithm is given by three filter bank evaluations and one adaptive sytem matrix with four elements of length $(L_f + L_s)/2 - 2$. The complexity of one filter bank evaluation per input sample varies from $(L_f + 1)/4$ multiplications (in the case of linear phase perfect reconstruction filter banks) to $L_f/2$ multiplications (in the classical QMF case) [29]. This leads to a total of approximately

$$2L_s + \frac{7}{2}L_f \quad \text{multiplications/sample.} \tag{54}$$

In (54), we took the usual QMF filter bank, and we neglected any possible savings from the factorization of the cross-filters. Assuming that $L_s >> L_f$, we see that the computational complexity is comparable with that of the full-band adaptive filter.

The two subband scheme with the nonoverdetermined adaptation algorithm requires three filter bank evaluations, the computation of six "filtered input signals" and four convolutions and adaptations with the polyphase components of the estimated system impulse response. Per block of two samples, we have for the filtered input signals $6L_f - 2$ multiplications, for the convolutions $4L_s/2 + 2 = 2L_s + 2$ multiplications, and for the adaptation $2(L_s + 1)$ multiplications. In the usual QMF case, this leads to a total of approximately

$$2L_s + \frac{9}{2}L_f \quad \text{multiplications/sample.} \tag{55}$$

We see that the computational complexity is again comparable with that of the full-band adaptive filter. Note that if the "trivial" QMF filter bank (for which $L_f = 2$) is used, the contributions of the filter banks and of the computation of the "filtered input signals" to the total number of multiplications are negligible.

In this two band case, savings can be obtained by using an adaptive fast running convolution scheme [22], [30] instead of the direct convolutions involving the system-dependent components. It will lead to about 25% savings over the time-domain implementation, but can be generalized to more bands yielding higher savings.

We will now consider N band schemes, with $N > 2$, and verify if these schemes can lead to substantial computational savings. A N band system, with subsampling by M, requires the evaluation of 3 filter banks (2 analysis and 1 synthesis bank) for every set of M input samples. Assuming that N is a power of 2, this leads to $3(N \log_2 N + L_f)/M$ multiplications/sample. It is assumed that the

filter banks are modulated and thus, three size N fast transforms (like DCT's or real DFT's) together with polyphase input and output filterings are only required per M input samples.

Further computational gains can be obtained when computing the polyphase filters in the frequency domain as well [31], but such schemes are quite involved and will not be considered further.

When using the overdetermined adaptation algorithm with critical subsampling ($M = N$), the adaptive part, between the analysis and the synthesis bank, consists of N main filters and $2N - 2$ cross-filters. The main filters are of length $(2L_f + L_s)/N$, and we will assume that, thanks to the factorization of cross-filters, a third of this length should be sufficient for those ones (this value comes from simulation results). Since all those filters are adaptive and require 2 multiplications per tap in the LMS case, the system identification part requires approximately

$$\frac{10}{3N} \cdot (L_s + 2L_f) \quad \text{multiplications/sample.} \tag{56}$$

The total number of multiplications for the filter banks and the system identification is therefore of the order of

$$\frac{1}{N}\left[\frac{10}{3} L_s + \frac{29}{3} L_f + 3N \log_2 N\right]$$

$$\text{multiplications/sample} \tag{57}$$

which shows that the dominant term (the factor of L_s) is about $5/3N$ times smaller than the full-band one, that is, the improvement in computational complexity is proportional to the number of subbands. Note, however, that L_f will increase with the number of channels, but is usually smaller than L_s.

Looking at the nonoverdetermined adaptation algorithm, we note that the direct use of (A1.1) for the adaptation of the coefficients of the estimated system impulse response leads to N multiplications and N additions for each coefficient. The computation of the estimated output in each subband (A1.5) requires N convolutions with the polyphase components of the estimated system impulse response, which are of size L_s/N. The adaptation formulas can be arranged in the form of convolution matrices of size $N \times N$; therefore significant computational gains (for large values of N) can be achieved by using fast convolution schemes. We can also save some computations by further decomposition of the filters responses and input signals in their polyphase components. If we use FFT's to compute the convolutions, the number of multiplications for the adaptive part can be reduced to approximately $3L_s \cdot \log_2 N$ for each block of N samples. We have then in total

$$\frac{1}{N}\left[3(L_s + N) \log_2 N + O(L_f)\right]$$

$$\text{multiplications/sample.} \tag{58}$$

Therefore, the computational gain over the full-band LMS algorithm is only of the order of $3 \log_2 N/N$.

Let us now make some comparisons with oversampled multiband schemes. In those schemes, the cross-filters are no longer present; then the adaptive part requires

$$\frac{N}{M} \cdot 2(L_s + 2L_f) \quad \text{multiplications/sample} \tag{59}$$

and the total system uses

$$\frac{M}{N}\left[2\frac{(L_s + 2L_f)}{M} + 3 \log_2 N\right] + 3\frac{L_f}{M}$$

$$\text{multiplications/sample.} \tag{60}$$

Assuming that N/M is close to 1 and that L_f is still reasonably small compared to L_s, we see an improvement by a factor of M (that is, close to N) with respect to the full-band system. Using (57), (58), and (60), it is possible to find the tradeoffs possible between critically sampled systems and oversampled ones, and this by adjusting the oversampling ratio N/M as well as the filter lengths in the banks.

In conclusion, we have verified that adaptive filtering in subbands can be computationally attractive, with gains of the order of the number of bands (the gain seems to be smaller in the case of the nonoverdetermined algorithm). Note that this fact has been used in transmultiplexer based adaptive systems, but with relatively large oversampling ratios ($2 \leq N/M \leq 4$ typically [15]) as well as in schemes for echo cancellation [16], [17].

VI. Experimental Results with Application to Acoustic Echo Cancellation

In this section we use acoustic echo cancellation as a framework for the experimental study of adaptive filtering in subbands with the overdetermined and nonoverdetermined adaptation algorithms. The corresponding subbands' adaptive schemes are named in the sequel overdetermined subbands (OSB) and nonoverdetermined subbands (NOSB's), respectively. We consider also the full-band adaptive filter in the purpose of comparisons.

We verify some important characteristics that have been derived theoretically or indicated in Sections II–IV:

1) it is shown that aliasing cancellation can be achieved with adaptive cross-filters;
2) it is shown that the use of perfect reconstruction filter banks leads to the perfect identification of the system;
3) it is verified that the nonoverdetermined adaptation algorithm generally yields better performances than the overdetermined one.

The convergence behavior is discussed for both adaptation algorithms. Besides, the tracking capability of a continuously time-varying system is examined.

A. Experimental Setup

The arrangement depicted in Fig. 1 was simulated in floating-point arithmetic. Three input signals were used: white noise, noise with a speech-like spectrum (USASI noise, ANSI standard S1.4-1961), and real speech; the sampling frequency was 16 kHz.

For the convergence experiments, the unknown system was generally simulated by a truncated impulse response measured in a real room, and the system output signals $y(t)$ were obtained by convolving the input signals $x(t)$ with those impulse responses. The adaptive filters had sufficient lengths to avoid system impulse response truncation effects. Thus the asymptotic behavior of the investigated schemes can be observed, because it is not masked by the residual echo which would persist with untruncated impulse responses.

In the purpose to simulate a time-varying system for the tracking experiments, a wood plate was moved between a loudspeaker and a microphone in a small room. The signals at the input (loudspeaker) and output (microphone) of the system (room) were recorded synchronously. The input signal was the USASI noise.

The experimental results are displayed in the form of total error decay curves (generally normalized by the short-time power of the output signal $y(t)$). In the sequel we call "asymptotic error" the error which is observed after the initial decay.

B. Schemes Using the Overdetermined Adaptation Algorithm

We discuss OSB schemes with two and eight subbands. In the two subband schemes QMF filter banks were used with filters of length $L_f = 32$ (type D [32]) as well as perfect reconstruction filter banks with short filters ($L_f = 4$). In the eight subband scheme, pseudo-QMF filter banks were used with filters of length 128 obtained by interpolation of a QMF prototype of length 16 (type B [32]). When QMF or pseudo-QMF filter banks were used, the cross-filters were factorized in a fixed part and in an adaptive part (system dependent) according to (11) and (20c–d). Adjacent cross-filters only were implemented in the eight subband scheme, according to the tridiagonal form derived in Section III.

In all the experiments discussed below we used the normalized least means squares algorithm (NLMS) [2]. The adaptation step size μ was computed according to $\mu = \nu/X^TX$ where ν is the adaptation gain constant and X is the vector of input signal samples stored in the adaptive filter. ν was generally set to 0.5 for the main filters; this value was always used for the full-band adaptive filter (with $\nu = 1$ the NLMS algorithm becomes unstable). For the cross-filters, ν was generally set to 0.0625 for optimal performances. The adaptation of the two subband scheme was done according to Fig. 5; the same adaptation structure was extended straightforwardly to the eight subband scheme.

1) OSB Schemes with QMF Filter Banks: The error decay curves for the two and eight subband OSB schemes and for the full-band adaptive filter are shown on Fig. 6 (input signal: USASI noise). The three curves exhibit similar initial parts; however, after the error has decayed by about 35 dB (2 subbands) and 27 dB (8 subbands), the convergence of the OSB schemes slows down and the errors go down to asymptotic values of about −50 dB (2 subbands) and −43 dB (8 subbands).

TABLE I
ERROR DECAY RATES (DECIBELS/SECOND) FOR THE FULL-BAND SCHEME AND FOR THE TWO SUBBAND OSB SCHEME WITH FACTORIZED CROSS-FILTERS (GLOBAL VALUES AND VALUES IN EIGHT ADJACENT FREQUENCY BANDS ARE GIVEN)

Freq. kHz	0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	Global
Full-band	−74.	−76.	−76.	−66.	−62.	−52.	−48.	−41.	−73.
Subband	−86.	−102.	−113.	−48.	−30.	−193.	−192.	−82.	−83.

These asymptotic errors are well above the one of the full-band adaptive filter because nonperfect QMF filter banks were used (see Section II). Some improvement of the transient behavior was observed in the higher part of the spectrum (see Table I); this improvement does not appear in Fig. 6 because the energy of the error is dominant in the lower frequencies.

In the absence of cross-filters the aliased components in the output error were not cancelled, as expected from the results of Section II (see experimental results in [8], [9]). Other experiments showed that when there was no factorization of the cross-filters the convergence speed at all frequencies was severely degraded; this will be discussed at the end of this section.

With speech signals at the input the convergence was significantly degraded in comparison with the full-band adaptive filter. This degradation probably comes from interactions between the main filters' and the cross-filters' adaptations.

2) Comparison with a Two Subband Oversampled Scheme: A two subband adaptive filter with an oversampling ratio of 5/4 was implemented. Its transient and asymptotic behavior with USASI noise at the input was found similar to the one of the two subband OSB scheme.

3) Two Subbands OSB Scheme with Perfect Reconstruction Filter Banks: As demonstrated in Section II, the use of perfect reconstruction filter banks should lead to zero asymptotic error at the output of the identification scheme; this property has been experimentally verified. Short filters with length $L_f = 4$ were used for demonstration purposes. The coefficients of H_0 and H_1 were, respectively, 1, 3, 3, 1 and 1, 3, −3, −1; the coefficients of G_0 and G_1 were, respectively, −1, 3, 3, −1, and 1, −3, 3, −1 [29].

The test was made with a white noise at the system input. The error curve is shown on Fig. 7, with the error curves for the full-band adaptive filter and for the two subband scheme with QMF filters discussed above. The asymptotic error of the scheme with perfect reconstruc-

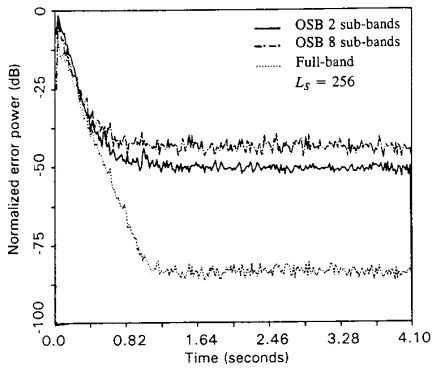

Fig. 6. Error decay curves for the two and eight subbands OSB schemes compared with the error curve for the full-band scheme (input: USASI noise).

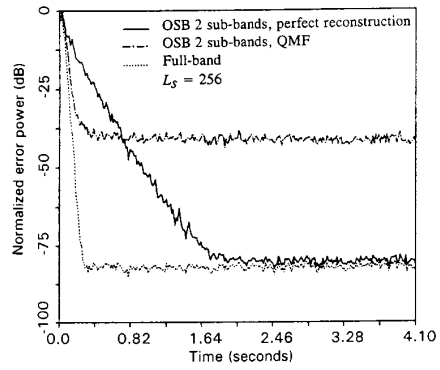

Fig. 7. Error decay curves for the two subband OSB schemes with perfect reconstruction and QMF filter banks compared with the error curve for the full-band scheme (input: white noise).

tion filter banks is very close to the error of the full-band filter; this demonstrates the perfect identification property (the observed residual error is due to the quantization of the system input and output signals).

Those last results show that the adaptive filtering process in subbands based on the feedback of the subbands' errors is able to identify perfectly a system. This was also verified by identifying a short synthetic impulse response: the asymptotic values of the elements in the identified model matrix were exactly the same as when computed from (11).

C. Two Subband Scheme Using the Nonoverdetermined Adaptation Algorithm

Since the derivation of the nonoverdetermined algorithm as well as the convergence analysis in Section IV focused on the two subband structures, we will discuss experimental results for the scheme with two subbands only.

A QMF filter bank with filters of length $L_f = 32$ was used (the same as for the OSB scheme), as well as a perfect reconstruction "trivial" QMF filter bank derived from $H(z) = 0.5(1 + z^{-1})$. The estimated system impulse response was adapted according to (25) (even and odd components were adapted separately).

1) Bound to the Adaptation Step Size: We have observed experimentally that with the USASI noise as the input signal $x(t)$, the instability appeared when μ was slightly greater than the bound (49). With a white noise at the input, the instability appeared when μ was equal to twice the bound. With input signals having lower correlation than the USASI noise (for example, white noise filtered by a first-order low-pass or high-pass filter), the instability appeared for values of μ between the bound (49) and twice this bound. Note that those bounds were essentially independent of the filter bank choice. For practical computations, the expectations were replaced by time averages over $L_s/2$ samples.

These results show that the simplified analysis derived from [2] is only approximatively adequate, although it gives a correct order of magnitude for the upper bound of the adaptation step size.

2) Convergence Behavior: The fastest convergence with the USASI noise was obtained with μ approximately equal to half the bound (49); in this case, the convergence was equivalent to the LMS one during an initial period, after which it became slower. The same behavior was observed when using both filter banks (Fig. 8). The fastest convergence with the white noise was obtained with μ equal to the bound (49), and in this case the convergence curve was identical to the full-band LMS one.

Note that the asymptotic error obtained when using the QMF filter bank is significantly higher than the one obtained with the "trivial" perfect reconstruction filter bank, which is very close to the error of the full-band adaptive filter. This corresponds indeed to the perfect modeling condition derived in Section II.

With speech at the input, the optimum value of μ was close to the optimum value found for the USASI noise; the convergence curve was then approximately the same as the one of the full-band scheme (Fig. 9). Other experiments have shown that with speech at the input the behavior of the two subband NOSB scheme was similar to the one of the subband oversampled scheme considered previously.

Overall, the performances of the NOSB scheme are significantly better than those of the OSB scheme: in particular, the asymptotic error is much smaller. However, the convergence speed is not improved in comparison with the full-band scheme.

D. Tracking Capability

The adaptation gain constants yielding the highest convergence speeds were used in the experiments. The length of the identified part of the impulse response was 50 ms. It is shown in Fig. 10 that the tracking capabilities of the OSB and NOSB two subband schemes considered are neither improved nor degraded in comparison with the one of the full-band adaptive filter. Experiments with the oversampled two subband scheme led to the same conclusion.

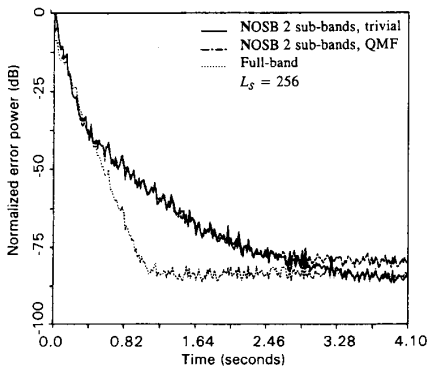

Fig. 8. Error decay curves for the two subband NOSB schemes with trivial and QMF filter banks compared with the error curve for the full-band scheme (input: USASI noise).

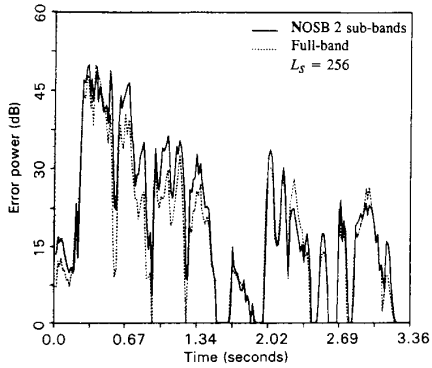

Fig. 9. Error decay curve for the two subband NOSB scheme with trivial filter banks compared with the error curve for the full-band scheme (input: speech).

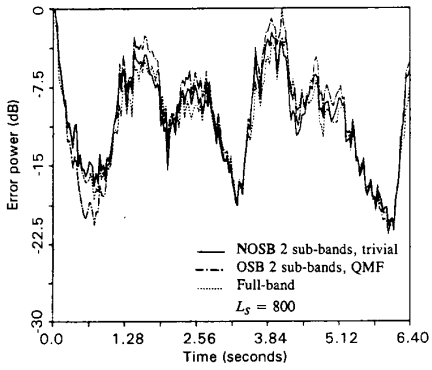

Fig. 10. Error curves for the two subband OSB and NOSB schemes and for the full-band scheme (time-varying system; input: USASI noise).

Similar experiments showed that the tracking capability of the OSB eight subband scheme was slightly degraded in comparison with the one of the full-band scheme (this is probably linked to the problem of choosing the adaptation step size of the cross-filters).

E. Further Comments on the Experimental Results

The OSB and NOSB schemes with perfect reconstruction filter banks are asymptotically optimal, i.e., they are able to identify the system without bias. This is an interesting result from a theoretical point of view, because it extends the general framework of the identification techniques.

Note also that the nonzero asymptotic errors that can be observed in the nonperfect QMF case (and also in the oversampled case) are not detrimental in acoustic echo cancellation, because in this application the identification is systematically biased by the truncation of the impulse response, and the resulting output error is generally larger than the error due to the nonperfect character of the aforementioned subband schemes.

The convergence performances of the OSB schemes are strongly dependent on the cross-filters: it has been observed that the adaptation of the main filters is slowed down as soon as the cross-filters are adapting. This can be intuitively explained by interactions between the adaptive processes in the main filters and in the cross-filters, which are made less critical by the factorization of the cross-filters; nevertheless, the adaptation step sizes of the cross-filters had to be chosen small for satisfactory performances.

The overall convergence speeds of the OSB and NOSB schemes were not found significantly better than the one of the full-band adaptive filter. Nevertheless, such schemes would have the practical advantage of reduced computational complexity in comparison with the full-band adaptive filter, as shown in Section V.

Double-talk aspects were not investigated in details. A theoretical study of those aspects would be certainly very complex. Note that experiments (not discussed in this paper) with the OSB schemes, where a noise was added to the echo signal, did not show significant performance degradations (in terms of residual echo minus noise) in comparison with the full-band scheme.

Overall, the theoretical results from the previous sections have been verified (more or less approximately); thus, the multirate framework that we have used seems to be adequate for the analysis of subband adaptive filtering schemes. Note that the theoretical analysis of the adaptive behavior is still to be completed.

VII. Conclusion

In this paper we have shown how to solve theoretically and experimentally the adaptive identification problem with critically sampled multirate systems. First, we have shown that the (approximately) perfect modelization of linear systems with critically sampled subband schemes can be achieved only if the model matrix incorporates nondiagonal terms, i.e., cross-filters between the subbands. Then we have made these schemes adaptive; for that purpose, we have proposed two adaptation algorithms; the first one (OSB) leads to overdetermination

problems which are avoided by the second one (NOSB). The additional complexity corresponding to the cross-filters can be reduced if the filter banks are chosen with good selectivity. The adaptive behavior of OSB schemes is dependent on this selectivity, and the factorization of the cross-filters is necessary for satisfactory performances. The NOSB schemes are much less dependent on the filter banks characteristics, but they are less computationally efficient.

Overall, the results did not meet the expectations originally put into critically sampled subbands schemes. Although generally superior to the OSB schemes, the NOSB ones did not give better convergence performances than the full-band adaptive filter. Since both convergence gains and computational efficiency can be best achieved with oversampled schemes, oversampling is still the way to go. For better computational efficiency, one can use NOSB fast running filter based filtering [36]; but no convergence gain can be obtained directly over the full-band scheme.

Appendix 1
Multiband Nonoverdetermined Algorithm

The LMS adaptation algorithm for more than two subbands is obtained straightforwardly from (25):

$$\hat{s}_k(m) = \hat{s}_k(m-1) - \mu(m) \sum_{l=0}^{N-1} 2e_l(m) \frac{\partial e_l(m)}{\partial \hat{s}_k}, \quad k = 0, 1, \cdots, L_s - 1 \tag{A1.1}$$

where m is the block number (the block size is now N). Assuming that the filter banks are sufficiently selective, we can restrict the system matrix to the tridiagonal form (20a), and we get

$$\frac{\partial e_l(z^N)}{\partial \hat{s}_k} = -\frac{\partial C_{l,l}(z)}{\partial \hat{s}_k} X_l(z^N) - \frac{\partial C_{l,l+1}(z)}{\partial \hat{s}_k} X_{l+1}(z^N) - \frac{\partial C_{l,l-1}(z)}{\partial \hat{s}_k} X_{l-1}(z^N), \quad l = 0, 1, \cdots, N-1 \tag{A1.2}$$

(the second indexes in the terms $C_{l,l+1}$ and $C_{l,l-1}$ are to be taken *modulo* N in the complex filter case).

As in the two subband case, it is convenient to consider the partial derivatives in (A1.1) relatively to the polyphase components of the estimated system impulse response. We set

$$\hat{S}(z) = \sum_{p=0}^{N-1} z^{-p} \hat{S}_p(z^N). \tag{A1.3}$$

Using (20b–d), we can obtain the expression for those partial derivatives. We get, for example,

$$\frac{\partial C_{l,l}(z)}{\partial \hat{s}_{p,q}} = z^K \sum_{i=0}^{N-1} H(W^i z) G(W^i z) (W^{i-l} z)^{-p} z^{-Nq} \tag{A1.4}$$

where $\hat{S}_{p,q}$ is the qth coefficient of the pth polyphase component of $\hat{S}$. Note that we have replaced the approximate equality in (20b) by an exact one.

Using again (20b–d) we obtain for the estimated system outputs in the subbands

$$\hat{y}_l(z^N) = z^K \sum_{p=0}^{N-1} \hat{S}_p(z^N) \left\{ \sum_{i=0}^{N-1} H(W^i z) \cdot G(W^i z)(W^{i-l} z)^{-p} X_l(z^N) + \sum_{i=0}^{N-1} H(W^i z) G(W^{i+l} z)(W^{i-l} z)^{-p} X_{l+1}(z^N) + \sum_{i=0}^{N-1} H(W^i z) G(W^{i-1} z)(W^{i-l} z)^{-p} X_{l-1}(z^N) \right\}. \tag{A1.5}$$

Note that since (A1.4) and (A1.5) are functions of z^N, they can be decimated.

As in the two subband case, the components of the gradient are equal to linear combinations of "filtered subbands input signals," and the estimated system outputs are obtained from convolutions of the same combinations of these signals with the polyphase components of $\hat{S}$.

Appendix 2
Proof of the Positive Definiteness of Φ

The matrix Φ (45) is positive definite iff for any nonzero vector V of size L_s the following inequality is satisfied:

$$V^T \Phi V > 0. \tag{A2.1}$$

Let $V^T = [V_1^T \; V_2^T]$ be a nonzero vector; V_1 and V_2 are vectors of size $L_s/2$. If V_1 (respectively, V_2) is zero, then V_2 (respectively, V_1) is nonzero. Since $E[\boldsymbol{A}(m)]$ and $E[\boldsymbol{D}(m)]$ are sums of covariance matrices, and therefore are positive definite, it can be shown immediately that (A2.1) is verified.

Let us turn now to the case when both V_1 and V_2 are nonzero vectors. Since they are deterministic quantities, the product $V^T \Phi V$ can be written

$$V^T \Phi V = E[V_1^T \boldsymbol{A}(m) V_1 + V_1^T \boldsymbol{B}(m) V_2 + V_2^T \boldsymbol{B}^T(m) V_1 + V_2^T \boldsymbol{D}(m) V_2]. \tag{A2.2}$$

Using the expressions (42)–(44) for the matrices $\boldsymbol{A}(m)$, $\boldsymbol{B}(m)$, and $\boldsymbol{D}(m)$, we define the scalar quantities (which are nonzero for general input signals):

$$C_1 = 2V_1^T X_{0,c} + V_2^T (X_{0,s} - X_{1,d}) \tag{A2.3}$$

$$C_2 = 2V_1^T X_{1,c} + V_2^T (X_{0,d} - X_{1,s}). \tag{A2.4}$$

Making use of the associativity of the product of matrices and vectors, we get straightforwardly

$$C_1^2 + C_2^2 = 0.5\{V_1^T \boldsymbol{A}(m) V_1 + V_1^T \boldsymbol{B}(m) V_2 + V_2^T \boldsymbol{B}^T(m) V_1 + V_2^T \boldsymbol{D}(m) V_2\}. \tag{A2.5}$$

Taking the expectations on both sides of (A2.5), we obtain the required result (A2.1).

ACKNOWLEDGMENT

The authors would like to thank in particular one of the reviewers for his excellent and very thorough review which improved the manuscript substantially.

REFERENCES

[1] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. Cambridge: M.I.T. Press, 1983.
[2] M. L. Honig and D. G. Messerschmitt, *Adaptive Filters: Structures, Algorithms, and Applications*. Boston: Kluwer, 1984.
[3] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[4] O. Horna, "Cancellation of acoustic feedback," *COMSAT Tech. Rev.*, vol. 12, pp. 319-333, 1982.
[5] M. Bellanger, *Adaptive Digital Filtering and Signal Analysis*. New York: Marcel Dekker, 1987.
[6] W. Kellermann, "Kompensation akustischer Echos in Frequenzteilbändern," *Frequenz*, vol. 39, no. 7/8, pp. 209-215, 1985.
[7] I. Itoh *et al.*, "An acoustic echo canceller for teleconference," in *Proc. IEEE ICC'85* (Chicago, IL), pp. 1498-1502.
[8] A. Gilloire, "Experiments with subband acoustic echo cancellers for teleconferencing," in *Proc. IEEE ICASSP'87* (Dallas, TX), pp. 2141-2144.
[9] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands," in *Proc. IEEE ICASSP'88* (New York, NY), pp. 1572-1575.
[10] A. Gilloire and J. F. Zurcher, "Achieving the control of the acoustic echo in audio terminals," in *Proc. EUSIPCO'88* (Grenoble, France), pp. 491-495.
[11] W. Kellermann, "Analysis and design of multirate systems for cancellation of acoustical echoes," in *Proc. IEEE ICASSP'88* (New York, NY), pp. 2570-2573.
[12] J. Chen, H. Bes, J. Vandewalle, and P. Janssens, "A new structure for subband acoustic echo canceller," in *Proc. IEEE ICASSP'88* (New York, NY), pp. 2574-2577.
[13] C. P. J. Tzeng, "An analysis of a subband echo canceller," in *Proc. IEEE GLOBECOM'87*, pp. 49.1.1-49.1.4.
[14] H. Yasukawa, S. Shimada, and I. Furukawa, "Acoustic echo canceller with high speech quality," in *Proc. IEEE ICASSP'87* (Dallas, TX), pp. 2125-2128.
[15] E. R. Ferrara, "Frequency-domain adaptive filtering," in *Adaptive Filters*, C. F. N. Cowan and P. M. Grant, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[16] P. T. Sommen, P. J. Van Gerwen, H. J. Kotmans, and A. J. E. M. Janssen, "Convergence analysis of a frequency-domain adaptive filter with exponential power averaging and generalized window function," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 788-798, July 1987.
[17] M. R. Asharif, F. Amano, S. Unagami, and K. Murano, "Acoustic echo canceller based on frequency bin adaptive filtering (FBAF)," in *Proc. IEEE GLOBECOM'87*, pp. 49.2.1-49.2.5.
[18] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
[19] M. Vetterli, "A theory of multirate filter banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 356-372, Mar. 1987.
[20] P. P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having perfect reconstruction property," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 476-492, Apr. 1987.
[21] P. P. Vaidyanathan and S. K. Mitra, "Polyphase networks, block digital filtering, LPTV systems, and alias-free QMF banks: A unified approach based on pseudocirculants," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 381-391, Mar. 1988.
[22] M. Vetterli, "Running FIR and IIR filtering using multirate filter banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 730-738, May 1988.
[23] P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions, and perfect-reconstruction technique," *IEEE ASSP Mag.*, vol. 4, no. 3, pp. 4-20, July 1987.
[24] H. J. Nussbaumer, "Pseudo QMF filter bank," *IBM Tech. Disclosure Bull.*, 1981.
[25] J. Masson and Z. Picel, "Flexible design of computationally efficient nearly perfect QMF filter banks," in *Proc. IEEE ICASSP'85* (Tampa, FL), pp. 541-544.
[26] P. L. Chu, "Quadrature mirror filter design for an arbitrary number of equal bandwidth channels," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 203-218, Feb. 1985.
[27] R. V. Cox, "The design of uniformly and nonuniformly spaced pseudoquadrature mirror filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1090-1096, Oct. 1986.
[28] M. Vetterli, "Filter banks allowing perfect reconstruction," *Signal Processing*, vol. 10, pp. 219-244, Apr. 1986.
[29] M. Vetterli and D. Le Gall, "Perfect reconstruction FIR filter banks: Some properties and factorizations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1057-1071, July 1989.
[30] Z. J. Mou and P. Duhamel, "Fast FIR filtering: Algorithms and implementations," *Signal Processing*, vol. 13, pp. 377-384, Dec. 1987.
[31] H. J. Nussbaumer, "Polynomial transform implementation of digital filter banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 616-622, June 1983.
[32] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. IEEE ICASSP'80* (Denver, CO), pp. 291-294.
[33] P. P. Vaidyanathan and P. Q. Hoang, "Lattice structures for optimal design and robust implementation of two-channel perfect reconstruction QMF banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 81-94, Jan. 88.
[34] CCITT Recommendation G722 (Wideband Speech Coding), Blue Book, 1990.
[35] CCITT Recommendation G131 (Echo and Delay), Blue Book, 1989.
[36] J. Benesty and P. Duhamel, "A fast exact least mean square adaptive algorithm," *IEEE Trans. Signal Processing*, to be published, Dec. 1992.

André Gilloire was born in France in 1948. He received the Doctorat de 3ème Cycle in 1977 from the University of Paris-Sud.

Since 1976 he has been with the National Research Center for Telecommunications (CNET), Lannion, France, where he first worked on speech coding. Since 1981 his research interests have been digital signal processing and adaptive filtering techniques, and their applications to audio terminals like hands-free telephony and teleconference systems.

Martin Vetterli (S'86-M'86-SM'90) was born in Switzerland in 1957. He received the Dipl. El.-Ing. degree from the Eidgenössische Technische Hochschule Zürich, Switzerland, in 1981, the master of science degree from Stanford University, Stanford, CA, in 1982, and the Doctorat ès Science degree from the Ecole Polytechnique Fédérale de Lausanne, Switzerland, in 1986.

In 1982, he was a Research Assistant at Stanford University, and from 1983 to 1986 he was a Researcher at the Ecole Polytechnique. He has worked for Siemens and AT&T Bell Laboratories. In 1986, he joined Columbia University in New York where he is currently Associate Professor of Electrical Engineering, a member of the Center for Telecommunications Research, and Codirector of the Image and Advanced Television Laboratory. His research interests include multirate signal processing, wavelets, computational complexity, signal processing for telecommunications, and digital video processing.

Dr. Vetterli is a member of SIAM and ACM, a member of the MDSP committee of the IEEE Signal Processing Society, and of the editorial boards of *Signal Processing*, *Image Communication*, and *Annals of Telecommunications*. He received the Best Paper Award of EURASIP in 1984 for his paper on multidimensional subband coding, and the Research Prize of the Brown Bovery Corporation (Switzerland) in 1986 for his thesis, and the IEEE Signal Processing Society's 1991 Senior Award (DSP Technical Area) for a 1989 transactions paper with D. LeGall on FIR filter banks.