



This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.


C O M M O N S D E E D

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:



Attribution. You must attribute the work in the manner specified by the author or licensor.



Noncommercial. You may not use this work for commercial purposes.



No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

Residual Echo Signal in Critically Sampled Subband Acoustic Echo Cancellers Based on IIR and FIR Filter Banks

Oğuz Tanrikulu, *Member, IEEE*, Buyurman Baykal, *Member, IEEE*,
 Anthony G. Constantinides, *Senior Member, IEEE*,
 and Jonathon A. Chambers, *Member, IEEE*

Abstract—The residual echo signal characteristics of critically sampled subband acoustic echo cancellers are analyzed. For finite impulse response (FIR) filter banks, the residual echo signal usually has a relatively broad spectral nature around the subband edges. The residual echo signal of power symmetric infinite impulse response (PS-IIR) filter banks, on the other hand, has very narrowband spectral components around the subband edges. These components can be efficiently removed with PS-IIR notch filters that integrate neatly into the filter banks without introducing perceptually noticeable degradation to the near-end speech. This solution has very low computational complexity and does not impinge on the system performance. Simulation studies with recordings from the cockpit of a car, based on a fast QR least-squares adaptive algorithm, demonstrate the potential of this approach for a practical AEC system.

I. INTRODUCTION

ACOUSTIC echo cancellation (AEC) is a critical problem in teleconferencing systems and hands-free mobile telephone terminals [1], [2]. A variety of techniques have been proposed that strike a balance between performance and computational complexity [2]–[12]. An extensive bibliography can be found in [13].

Acoustic echoes are generated due to the coupling between the loudspeaker and microphone, as illustrated in Fig. 1. The far-end speaker is using a conventional handset that blocks the acoustic path between the loudspeaker and microphone. At the near end, the speaker is not in physical contact with the terminal. The microphone picks up the far-end speech after it has travelled through the acoustic medium, and consequently, echo is heard at the far end. The AEC unit in Fig. 1, therefore, has the duty of electrically synthesizing and subtracting the acoustic echo from the microphone signal.

The length of the acoustic echo path is dependent on the environment. For a car at 8 KHz sampling frequency, it is of the order of hundreds of samples, and for a room, it can be a few thousand [5]. Therefore, the computational complexity of the AEC unit may be very high and critically dependent

Manuscript received October 15, 1995; revised October 14, 1996. This work was supported in part by the EC ESPRIT Project 6166 "FREE TEL" and the U.K., EPSRC research project GR/K48617 "Multirate Echo Cancellation." The associate editor coordinating the review of this paper and approving it for publication was Dr. Jelena Kovacevic.

The authors are with the Department of Electrical and Electronic Engineering, Imperial College of Science, Technology, and Medicine, London, U.K.
 Publisher Item Identifier S 1053-587X(97)02585-3.

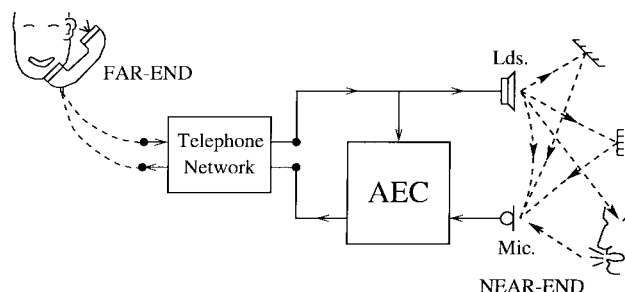


Fig. 1. Acoustic echo cancellation.

on the echo cancellation algorithm, which is an adaptive filter implemented in the time [14], [7] or frequency domain [15] or in subbands [5], [16], [10], [17], [8]. Gradient-descent type adaptive algorithms such as the normalized least mean-square (NLMS) algorithm [18] have slow convergence speed. Recent developments highlight the importance of employing stabilized fast recursive least squares (FRLS) [2], [7] or affine projection [14], [19] algorithms for high performance, but the computational complexity problem still remains. Adaptive filtering in subbands is an elegant way of reducing the computational complexity, which generally increases the convergence speed due to the reduced ill conditioning of each subband signal [5].

There are two drawbacks associated with *critically sampled* subband AEC systems. First, the subband signals are aliased, which limits the echo cancellation performance [20] and second, the multirate analysis banks (MRAB's) and the multirate synthesis banks (MRSB's) introduce extra delay and computational complexity. Cross-adaptive filtering was proposed by Gilloire and Vetterli [21] in order to solve the *interband aliasing* problem. They performed a study of the generic critically sampled two-band AEC unit with finite impulse response (FIR) MRAB and MRSB and demonstrated the necessity of performing cross-channel identification (CRCI) [5]. In their deterministic framework, it was shown that if the MRAB and MRSB possess the perfect reconstruction (PR) property, the residual echo can be made identically zero, provided that all CRCI's are also performed without error. For CRCI, cross-adaptive filters were proposed, but this solution does not deliver the required performance despite the computational complexity being increased significantly. Oversampled MRAB and MRSB have also been used [4], [12]

to eliminate the aliasing at the expense of greatly increased computational complexity, delay, and memory requirements.

In this paper, the drawbacks of critically sampled subband AEC systems are addressed. Power symmetric infinite impulse response (PS-IIR) multirate filterbanks [22]–[24], [16], [8] are proposed for a lower level of aliasing and computational complexity. FIR multirate filterbanks with aliasing cancellation (AC) and PR properties are also investigated. In Section II, a brief summary of the properties of PS-IIR filters are presented along with the AC-FIR and PR-FIR filters used in the paper. Section III introduces the generic two-band AEC unit and the associated variables used in the subsequent sections [5]. In Section IV, approximate expressions are obtained for the residual echo signal when PS-IIR or AC-FIR multirate filterbanks are used without CRCI. It is shown that if the prototype lowpass filter (FIR or IIR) is highly selective, the residual echo signal will contain coherent spectral components (tonal components). CRCI is not suitable under such circumstances since it is necessary to use very long cross-adaptive filters [24], [25]. In Section V, notch filtering is proposed in order to remove the tonal components. In Section VI, the phase distortion imposed by the PS-IIR multirate filterbanks on the near-end speech is investigated. Analytical results for the amplitude reconstruction and the group delay of the cascaded MRAB and MRSB are presented. Informal subjective tests show that the phase distortion is perceptually reduced after introducing notch filters. In particular, without the notch filters, the number of subbands cannot be increased beyond 4, whereas with notch filtering, 16 subbands can be used. Compact implementations of the notch filters and the MRAB and MRSB of the PS-IIR multirate filterbanks are discussed in Section VII. Computational complexity issues related to implementation are also addressed. In Section VIII, the use of notch filtering is demonstrated, and several AEC results for real recordings are presented, where a fast QR least-squares adaptive algorithm [26], [18] is used in subbands. Finally, Section IX draws conclusions on the proposed AEC system.

II. PROPERTIES OF THE PS-IIR FILTERS

PS-IIR filters have the transfer function [23]

$$H_0(z) = \frac{A_0(z^2) + z^{-1}A_1(z^2)}{2} \quad (1)$$

where $A_i(z^2)$, $i = 0, 1$ are cascaded allpass networks [27], [22], [28] defined as

$$A_i(z^2) = \prod_{j=1}^{P_i} \frac{\alpha_{i,j} + z^{-2}}{1 + \alpha_{i,j}z^{-2}}, \quad i = 0, 1 \quad (2)$$

where P_i is the corresponding order. In order to obtain a lowpass filter with stopband frequency θ_s , the stopband energy can be minimized over a dense set of grids [23], [29]. Near optimal initialization of the allpass coefficients $\alpha_{i,j}$ and their discrete reoptimization are described in [16] and [30]. The connections between PS-IIR and Elliptic, Butterworth, and Chebyshev filters are extensively discussed in [22], and an alternative design method is given.

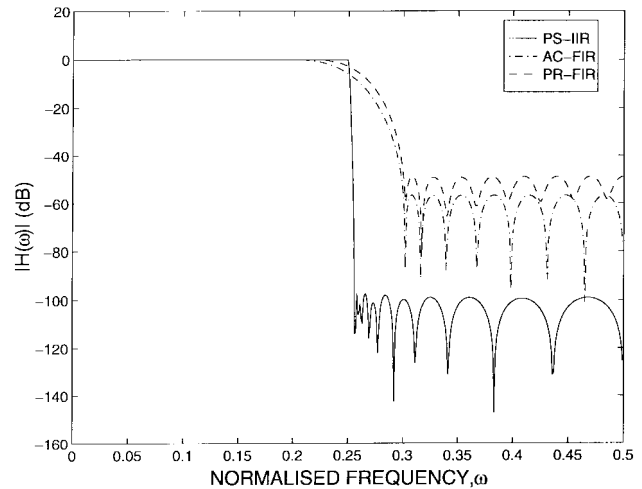


Fig. 2. Amplitude spectra of the PS-IIR in Table I and typical AC-FIR and PR-FIR with 32 taps, $\omega = \theta/2\pi$.

TABLE I
ALLPASS COEFFICIENTS OF THE PS-IIR FILTER IN FIG. 2

$\alpha_{0,1}$	$\alpha_{0,2}$	$\alpha_{0,3}$	$\alpha_{0,4}$	$\alpha_{0,5}$	$\alpha_{0,6}$
0.040407	0.297311	0.593341	0.798278	0.914901	0.984964
$\alpha_{1,1}$	$\alpha_{1,2}$	$\alpha_{1,3}$	$\alpha_{1,4}$	$\alpha_{1,5}$	
0.149350	0.452729	0.708912	0.865132	0.953132	

Once a prototype lowpass filter is designed, the corresponding mirror image highpass filter is given by

$$H_1(z) = H_0(-z) = \frac{A_0(z^2) - z^{-1}A_1(z^2)}{2}. \quad (3)$$

Let us note that in PS-IIR filters [23]

$$\theta_s > \frac{\pi}{2} \quad (4)$$

and therefore, they are only suitable for a two-band decomposition. If more than two subbands are required, a *binary-tree subband decomposition* [22] must be used [8], [29]. During the analysis, two subbands are assumed, but the conclusions still apply if there are more subbands.

To facilitate our discussion, prototype lowpass PS-IIR, AC-FIR, and PR-FIR filters are designed whose amplitude spectra are shown in Fig. 2. For the PS-IIR filter, $P_0 = 6$, $P_1 = 5$, and $\theta_s = 1.6085$ rad. The corresponding allpass coefficients are presented in Table I. The number of taps in the FIR filters is chosen as 32 due to stringent specifications on the AEC computational complexity and delay in hands-free mobile terminals [25]. For the AC-FIR and the PR-FIR filters, the stopband frequency is $\theta_s = 1.8908$ rad. The AC-FIR is equiripple and linear phase. The PR-FIR is lossless and is designed by using the spectral factorization technique in [22].

III. GENERIC TWO BAND AEC UNIT

The generic two-band AEC unit is shown in Fig. 3. $X(z)$ is the loudspeaker signal (in the Z -domain) that represents the far-end speech. $S(z)$ represents the acoustic echo path that is assumed to be linear and time invariant [5]. $M(z)$ is the acoustic echo signal captured by the microphone. $W(z)$ is the

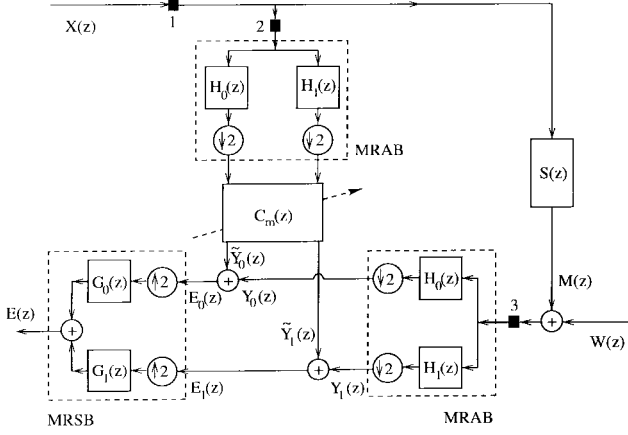


Fig. 3. Generic two-band AEC unit.

near-end speech. For the sake of clarity, $W(z)$ is assumed to be zero when there is speech activity at the far end (single-talk) [24], [25]. $W(z)$ will be explicitly used in Section VI when analyzing the phase distortion introduced by the PS-IIR multirate filter banks. $H_i(z)$, $i = 0, 1$ are the MRAB filters, and $G_i(z)$, $i = 0, 1$ are the corresponding MRSB filters. $E_i(z)$, $i = 0, 1$, and $E(z)$ are, respectively, the subband and the fullband residual echo signals. $\hat{Y}_i(z)$ and $Y_i(z)$, $i = 0, 1$ are, respectively, the predicted and the actual subband echo signals after decomposition, and $\mathbf{C}_m(z)$ is the predictor. In the analysis, $\mathbf{C}_m(z)$ is a 2×2 matrix whose off-diagonal terms represent the CRCI.

According to the modulation domain notation of Vetterli [31], [5], let us define

$$\mathbf{x}_m(z) \triangleq [X(z)X(-z)]^T \quad (5)$$

$$\mathbf{S}_m(z) \triangleq \begin{bmatrix} S(z) & 0 \\ 0 & S(-z) \end{bmatrix} \quad (6)$$

$$M(z) \triangleq S(z)X(z) \quad (7)$$

$$\mathbf{H}_m(z) \triangleq \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} \quad (8)$$

$$\mathbf{g}_m(z) \triangleq [G_0(z)G_1(z)]^T \quad (9)$$

$$\mathbf{C}_m(z) \triangleq \begin{bmatrix} C_{0,0}(z) & C_{0,1}(z) \\ C_{1,0}(z) & C_{1,1}(z) \end{bmatrix}. \quad (10)$$

We then have

$$\mathbf{y}_m(z) \triangleq \begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix} = \frac{1}{2}\mathbf{H}_m(z^{1/2})\mathbf{S}_m(z^{1/2})\mathbf{x}_m(z^{1/2}) \quad (11)$$

$$\hat{\mathbf{y}}_m(z) \triangleq \begin{bmatrix} \hat{Y}_0(z) \\ \hat{Y}_1(z) \end{bmatrix} = \frac{1}{2}\mathbf{C}_m(z)\mathbf{H}_m(z^{1/2})\mathbf{x}_m(z^{1/2}) \quad (12)$$

and

$$\mathbf{e}_m(z) = \begin{bmatrix} E_0(z) \\ E_1(z) \end{bmatrix} = \mathbf{y}_m(z) - \hat{\mathbf{y}}_m(z) \quad (13)$$

$$E(z) = \mathbf{g}_m^T(z)\mathbf{e}_m(z^2) = \frac{1}{2}\mathbf{g}_m^T(z) \cdot [\mathbf{H}_m(z)\mathbf{S}_m(z) - \mathbf{C}_m(z^2)\mathbf{H}_m(z)]\mathbf{x}_m(z). \quad (14)$$

Therefore, the residual echo $E(z)$ is zero when

$$\mathbf{C}_m(z^2) = \mathbf{H}_m(z)\mathbf{S}_m(z)[\mathbf{H}_m(z)]^{-1} \quad (15)$$

where

$$[\mathbf{H}_m(z)]^{-1} = \frac{1}{\det[\mathbf{H}_m(z)]} \begin{bmatrix} H_1(-z) & -H_0(-z) \\ -H_1(z) & H_0(z) \end{bmatrix}. \quad (16)$$

IV. TWO-BAND PROCESSING WITHOUT CROSS-CHANNEL IDENTIFICATION

Operating a critically sampled subband AEC without CRCI is equivalent to substituting $C_{0,1}(z) = 0$ and $C_{1,0}(z) = 0$ in (10). Consequently, as long as the prototype filter $H(z)$ is nonideal, the residual echo signal will contain components due to the aliasing between the subband signals.

A. PS-IIR Multirate Filter Banks

For aliasing cancellation in cascaded MRAB and MRSB [22], we use

$$\begin{aligned} H_0(z) &= H(z) & H_1(z) &= H(-z) \\ G_0(z) &= 2H(z) & G_1(z) &= -2H(-z). \end{aligned} \quad (17)$$

Using (1) and (2) in (8), we have

$$\det[\mathbf{H}_m(z)] = H^2(z) - H^2(-z) = z^{-1}A_0(z^2)A_1(z^2) \quad (18)$$

and, thus, (19), shown at the bottom of the page, where $\Psi(z) \triangleq H(z)H(-z)$. If we ignore the off-diagonal terms in (19) (no CRCI) and use (18), we have

$$\begin{aligned} z^{-1}A_0(z^2)A_1(z^2)E(z) &= -\Psi(z)[S(z) - S(-z)](2\Psi(z)X(z) \\ &\quad + [H^2(z) + H^2(-z)]X(-z)). \end{aligned} \quad (20)$$

In order to gain more insight into (20), we evaluate it on the unit circle and write

$$\begin{aligned} e^{-j\theta}A_0(2\theta)A_1(2\theta)E(\theta) &= -\Psi(\theta)[S(\theta) - S(\theta - \pi)](2\Psi(\theta)X(\theta) \\ &\quad + [H^2(\theta) + H^2(\theta - \pi)]X(\theta - \pi)). \end{aligned} \quad (21)$$

When $H(\theta)$ is a highly selective lowpass prototype PS-IIR filter such as shown in Fig. 2, $\Psi(\theta)$ is nonzero over a very narrow bandwidth around $\theta = \pi/2$. Furthermore, $H(\theta)$ has a real impulse response: $H(-\theta) = H^*(\theta)$. Therefore, we make the following approximation:

$$\Psi(\theta) \approx \Psi_a(\theta) = |H(\pi/2)|^2(\delta(\theta - \pi/2) + \delta(\theta + \pi/2)) \quad (22)$$

where $\delta(\theta)$ is the impulse function. In (22), a nonzero spectral component is allowed at $\theta = \pi/2$, which corresponds to having an almost ideal brickwall amplitude spectrum. It is easy to show that at $\theta = \pm\pi/2$, $H^2(\theta) + H^2(\theta - \pi) = 0$. Furthermore, we have from (1) $|H(\pi/2)|^2 = 1/2$ and $A_i(\pm\pi/2) = (-1)^{P_i}$, $i = 0, 1$. Therefore, when (22) is valid,

$$z^{-1}A_0(z^2)A_1(z^2)\mathbf{C}_m(z^2) = \begin{bmatrix} H^2(z)S(z) - H^2(-z)S(-z) & \Psi(z)[S(-z) - S(z)] \\ \Psi(z)[S(z) - S(-z)] & H^2(z)S(-z) - H^2(-z)S(z) \end{bmatrix} \quad (19)$$

the right-hand side of (21) is nonzero only at $\theta = \pm\pi/2$, and (21) becomes

$$E(\theta) \approx (-1)^{P_0+P_1} \text{Im}[S(\pi/2)](X(\pi/2)\delta(\theta - \pi/2) + X(-\pi/2)\delta(\theta + \pi/2)). \quad (23)$$

In the time domain, we have

$$e_n \approx \frac{1}{\pi} (-1)^{P_0+P_1} \text{Im}[S(\pi/2)] [\text{Re}[X(\pi/2)] \cdot \cos(\pi/2n) - \text{Im}[X(\pi/2)] \sin(\pi/2n)]. \quad (24)$$

Unless $\text{Im}[S(\pi/2)] = 0$, which is not realistic considering the arbitrary nature of the acoustic echo path, (24) is nonzero. When $x_n = 0$, then $e_n = 0$, which is consistent with our practical experience in the ‘‘single-talk’’ condition. In other words, when there is silence in the loudspeaker, the residual echo is also silent. On the other hand, when there is speech activity at the far end, a *tonal component* appears in the residual echo signal [24], [25]. In general, if there are more than two subbands, tonal components will become visible around each subband division frequency. Although these components might be of low energy, they are perceptually quite disturbing to the far-end speaker.

B. AC-FIR Multi-Rate Filter Banks

Let $H(z) = \sum_{k=0}^L h_k z^{-k}$ (L : odd) be a *linear-phase* low-pass filter. We choose the analysis and synthesis bank filters as in (17) for aliasing cancellation. If L is sufficiently high and $H(z)$ has unity passband gain, we have $\det[\mathbf{H}_m(z)] \approx z^{-L}$ [31], [5], and therefore, from (15), we have (25), shown at the bottom of the page. When L is large [31],

$$H^2(z) + H^2(-z) \approx z^{-L} \quad (26)$$

and after ignoring the off-diagonal terms in (25) (no CRCI), we have approximately

$$z^{-L} E(z) \approx -\Psi(z)[S(z) - S(-z)] \cdot (2\Psi(z)X(z) + z^{-L}X(-z)). \quad (27)$$

If the stopband attenuation is sufficiently high, an accurate approximation for $\Psi(\theta)$ would then be a modulated raised cosine of the form

$$\Psi_\alpha(\theta) = \begin{cases} \frac{1}{8} \left\{ 1 - \cos\left(\frac{2\pi}{\pi - 2\theta_s}[\theta - \theta_s]\right) \right\} e^{-jL\theta} & \pi - \theta_s < \theta < \theta_s \\ \frac{1}{8} \left\{ 1 - \cos\left(\frac{2\pi}{2\theta_s - \pi}[\theta + \theta_s]\right) \right\} e^{-jL\theta} & -\theta_s < \theta < \theta_s - \pi \\ 0 & \text{otherwise.} \end{cases} \quad (28)$$

$|\Psi(\theta)|$ and $|\Psi_\alpha(\theta)|$ are shown in Fig. 4. On the unit circle, from (27) and (28)

$$E(\theta) \approx -e^{jL\theta} \Psi_\alpha(\theta) [S(\theta) - S(\theta - \pi)] \cdot (2\Psi_\alpha(\theta)X(\theta) + e^{-jL\theta}X(\theta - \pi)). \quad (29)$$

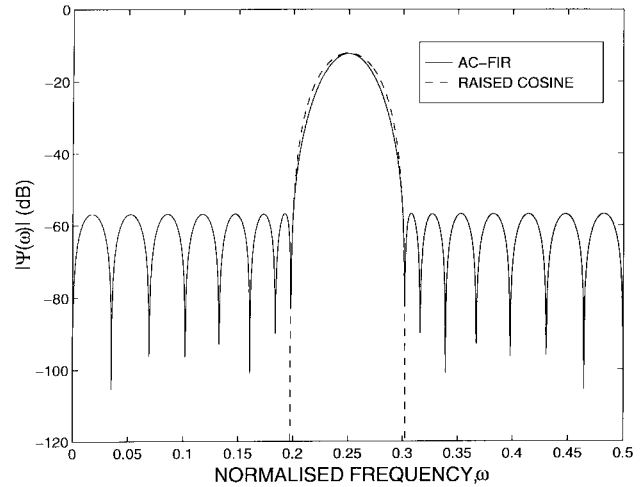


Fig. 4. $|\Psi(\omega)|$ and $|\Psi_\alpha(\omega)|$ for the AC-FIR filter in Fig. 2 $\omega = \theta/2\pi$.

The inverse discrete-time Fourier transform $\mathcal{F}^{-1}\{\cdot\}$ of (29) is

$$e_n \approx -2f_n * \{\sin^2(\pi/2n)s_n\} * \{2f_n * x_{n-L} + (-1)^{n-L}x_{n-L}\}. \quad (30)$$

In (30), $*$ denotes the linear convolution operation, and

$$f_n = \frac{1}{4\pi} \cos(\pi/2n) \sin([\theta_s - \pi/2]n) \cdot \left(\frac{1}{n} - \frac{n}{n^2 - \left(\frac{2\pi}{\pi - 2\theta_s}\right)^2} \right) \quad (31)$$

is a rapidly decreasing function in time. We can now comment on two cases, depending on the value of θ_s .

Case 4.1—(θ_s close to π): f_n becomes shorter in the time domain, but the aliasing bandwidth increases, which is clearly undesirable for high performance AEC.

Case 4.2—(θ_s close to $\pi/2$): $H(z)$ becomes closer to an ideal brickwall filter, similar to the PS-IIR filter described in the previous sections. In the time domain, f_n is longer and e_n in (30) will be similar to a cosine-type waveform. In other words, a tonal component will appear at $\theta = \pi/2$. However, reasonably short FIR filters are used in practice. These filters usually have a wider transition bandwidth than the PS-IIR filters. Therefore, we are almost never able to hear any audible tonal components in the residual echo e_n .

Similar results can be obtained for PR-FIR multirate filter-banks when there is no CRCI.

V. NOTCH FILTERING AND TONAL COMPONENTS

Clearly, $\Psi(\theta)$ is a key element in the analysis. If $\Psi(\theta) = 0$, in which case, $H(\theta)$ has an ideal brickwall amplitude

$$z^{-L} \mathbf{C}_m(z^2) \approx \begin{bmatrix} H^2(z)S(z) - H^2(-z)S(-z) & \Psi(z)[S(-z) - S(z)] \\ \Psi(z)[S(z) - S(-z)] & H^2(z)S(-z) - H^2(-z)S(z) \end{bmatrix} \quad (25)$$

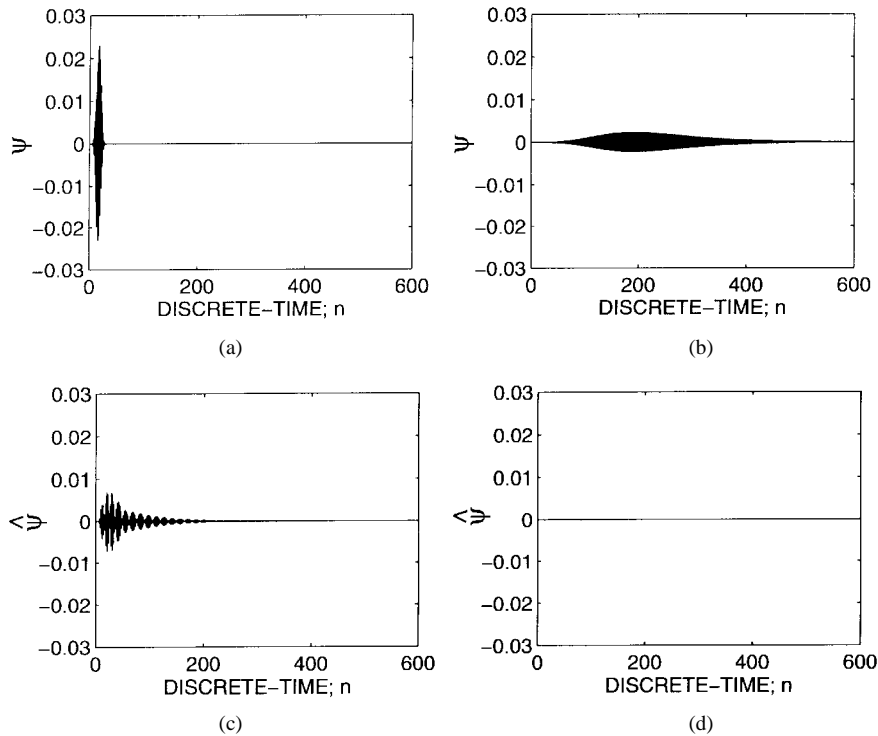


Fig. 5. ψ_n and $\hat{\psi}_n$ for (a), (c) the AC-FIR and (b), (d) the PS-IIR filters in Fig. 2.

spectrum, then (25) and (19) would be diagonal matrices, and consequently, CRCI would not be necessary [5]. The outputs of the MRAB are uncorrelated in this case. The time-domain characteristics of $\Psi(\theta)$ and its level indicate the lengths of $C_{0,1}(z)$ and $C_{1,0}(z)$ that must be identified. If $H(\theta)$ is an FIR filter with θ_s not too close to $\pi/2$, then the level of $\Psi(\theta)$ is high, but it is spread over a short range in the time domain. Therefore, reasonably short length cross-adaptive filters can be used [5]. This is clear in Fig. 5(a), where $\psi_n \triangleq \mathcal{F}^{-1}\{\Psi(\theta)\}$ is shown for the AC-FIR filter in Section II. However, using cross-adaptive filters would decrease the convergence speed of the overall AEC system and add a considerable amount of computational complexity, especially when the number of subbands is increased above two. The cause of the decrease in convergence speed is due to the adaptive nature of this solution. In the two-band case, the adaptation of the elements of $C_m(z)$ corresponds to the operation of two multichannel adaptive filters. It is well known that the convergence of such adaptive filters is more problematic [32].

On the other hand, if a highly selective PS-IIR filter such as in Fig. 2 is used, then the level of $\Psi(\theta)$ is spread over a much wider range in the time domain, as indicated by Fig. 5(b), and although its level is low, it is perceptually annoying to the far-end speaker. In this case, a tonal component occurs around the subband division frequency, as shown in Section IV-A. Therefore, rather than using cross-adaptive filters, notch filters with narrow bandwidths can be used to attenuate the input of the MRAB such that the cross-correlation between subband signals is almost completely removed. Unlike cross-adaptive filtering, the solution of notch filters is a deterministic one with significantly lower computational complexity and, therefore, does not impinge on the AEC performance.

In order to minimize the computational complexity, it is convenient to use an IIR notch structure based on allpass networks [22], [27], [29]. Consider a prototype lowpass filter of the form

$$\tilde{N}(z) = \frac{1}{2} \left(\frac{\alpha_{0,1} + z^{-2}}{1 + \alpha_{0,1}z^{-2}} + z^{-1} \right) \quad (32)$$

which is a special case of (1) for $(P_0, P_1) = (1, 0)$. By using the allpass lowpass to bandstop frequency transformation [33]

$$z^{-1} \leftarrow \frac{\zeta + z^{-2}}{1 + \zeta z^{-2}} \quad (33)$$

in (32), the desired notch filter can be obtained. The frequency transformation (33) makes the second-order allpass stage in (32) fourth-order allpass, which can be written as a cascade of two second-order allpass stages with real coefficients. This is useful during finite-precision implementation in order to minimize the round-off noise [34]. Thus, the notch filter $N(z)$ has the general form

$$N(z) = \frac{1}{2} \left(\hat{N}(z) + \frac{\zeta + z^{-2}}{1 + \zeta z^{-2}} \right) \quad (34)$$

where

$$\hat{N}(z) = \frac{\gamma_2 + \gamma_1 z^{-1} + z^{-2}}{1 + \gamma_1 z^{-1} + \gamma_2 z^{-2}} \frac{\gamma_2 - \gamma_1 z^{-1} + z^{-2}}{1 - \gamma_1 z^{-1} + \gamma_2 z^{-2}}. \quad (35)$$

A prototype lowpass filter of the form (32) with $\alpha_{0,1} = 0.3$ is spectrally transformed as described above. The resulting notch filter coefficients are $\zeta = 0.806325$, $\gamma_1 = 0.171050$ and $\gamma_2 = 0.891673$. Obviously, $N(\theta)$ is nonlinear phase. We leave the discussion on this to the next section, where the phase distortion is discussed.

Define the notch filter matrix

$$\mathbf{N}_m(z) \triangleq \begin{bmatrix} N(z) & 0 \\ 0 & N(-z) \end{bmatrix}. \quad (36)$$

Note that from (34) and (35), $N(z) = N(-z)$, which yields

$$\det[\mathbf{N}_m(z)] = N^2(z). \quad (37)$$

In Fig. 3, there are several candidate positions in which notch filters can be inserted. These are shown as 1, 2, and 3. We evaluate the feasibility of these positions below.

Case 5.1 (Notch Filter at 1): In this case, (19) is valid, and we have

$$E(z) = \frac{1}{2} \mathbf{g}_m^T(z) [\mathbf{H}_m(z) \mathbf{S}_m(z) - \mathbf{C}_m(z^2) \mathbf{H}_m(z) \mathbf{N}_m(z) \mathbf{x}_m(z)]. \quad (38)$$

If the cross-channel terms $C_{0,1}(z)$ and $C_{1,0}(z)$ are ignored (no CRCI)

$$\begin{aligned} & e^{-j\theta} A_0(2\theta) A_1(2\theta) E(\theta) \\ &= -\Psi(\theta) N(\theta) [S(\theta) - S(\theta - \pi)] \\ & \cdot (2\Psi(\theta) X(\theta) + [H^2(\theta) + H^2(\theta - \pi)] X(\theta - \pi)). \end{aligned} \quad (39)$$

Therefore, the coherent part of the spectrum defined by $\Psi(\theta)$ is filtered by $N(\theta)$, and clearly, the tonal component around $\theta = \pi/2$ is suppressed, provided that $N(\theta)$ is suitably designed to ensure $\Psi(\theta)N(\theta) \approx 0$, $\theta \in [0, \pi]$.

Case 5.2 (Notch Filter at 2): We have

$$E(z) = \frac{1}{2} \mathbf{g}_m^T(z) [\mathbf{H}_m(z) \mathbf{S}_m(z) - \mathbf{C}_m(z^2) \mathbf{H}_m(z) \mathbf{N}_m(z) \mathbf{x}_m(z)] \quad (40)$$

and (41), shown at the bottom of the page. If $C_{0,1}(z)$ and $C_{1,0}(z)$ are ignored, we get (21). Therefore, the residual echo signal properties remain unchanged, and a tonal component appears around $\theta = \pi/2$. Obviously, the spectrum of the loudspeaker signal is rejected around the notch, and the echo signal around this frequency cannot be predicted and, hence, cancelled.

Case 5.3 (Notch Filter at 3): We have

$$E(z) = \frac{1}{2} \mathbf{g}_m^T(z) [\mathbf{H}_m(z) \mathbf{N}_m(z) \mathbf{S}_m(z) - \mathbf{C}_m(z^2) \mathbf{H}_m(z) \mathbf{x}_m(z)] \quad (42)$$

and (43), shown at the bottom of the page. If $C_{0,1}(z)$ and $C_{1,0}(z)$ are ignored, we get (39) again, and thus, the tonal component can be removed. The difference here is in the

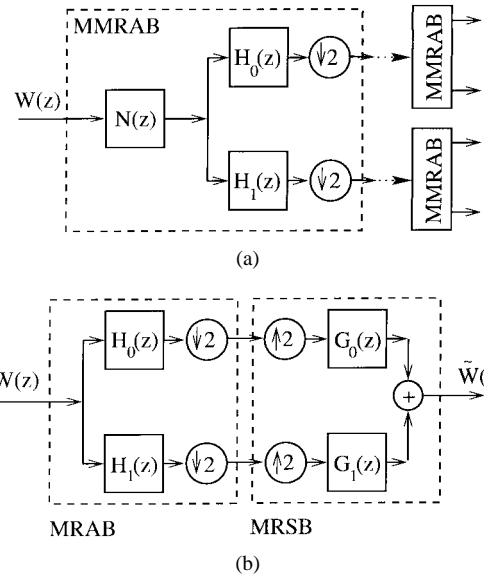


Fig. 6. (a) Modified multirate analysis banks (MMRAB). (b) Cascaded MRAB and MRSB.

identification matrix $\mathbf{C}_m(z)$. Since the diagonal terms are also filtered by the notch $N(z)$, which has a long impulse response, longer $C_{0,0}(z)$ and $C_{1,1}(z)$ are required. This may not be possible in practice due to the restrictions on the computational complexity. In practice, it is observed that although the tonal components are removed, the AEC performance generally decreases compared with Case 5.1 or Case 5.4.

Case 5.4 (Notch Filters at 2 and 3): When there is no near-end speech $W(z) = 0$, this is similar to Case 5.1 as far as the expressions for residual echo $E(\theta)$ and $\mathbf{C}_m(z^2)$ are concerned. However, in Case 5.1, the far-end speech is notch filtered and, therefore, distorted. Furthermore, if more than two subbands are required, more notches are needed, and the order of the notch filter will increase substantially. Such a notch filter is more difficult to implement than (34) in finite precision. Moreover, the notch filter in Case 5.1 operates in full rate. Thus, placing notch filters at 2 and 3 is the best configuration. We modify the MRAB in Fig. 3 by adding a notch filter as shown in Fig. 6(a). This is a modular structure in that if more than two subbands are required, the MMRAB can be used repeatedly as illustrated. Therefore, the notch filters after the first subband decomposition operate at reduced rates and, hence, require fewer computations than the notch filter required in Case 5.1.

$$z^{-1} A_0(z^2) A_1(z^2) N(z) \mathbf{C}_m(z^2) = \begin{bmatrix} H^2(z) S(z) - H^2(-z) S(-z) & \Psi(z) [S(-z) - S(z)] \\ \Psi(z) [S(z) - S(-z)] & H^2(z) S(-z) - H^2(-z) S(z) \end{bmatrix} \quad (41)$$

$$z^{-1} A_0(z^2) A_1(z^2) \mathbf{C}_m(z^2) = N(z) \begin{bmatrix} H^2(z) S(z) - H^2(-z) S(-z) & \Psi(z) [S(-z) - S(z)] \\ \Psi(z) [S(z) - S(-z)] & H^2(z) S(-z) - H^2(-z) S(z) \end{bmatrix} \quad (43)$$

From (39), define

$$\hat{\Psi}(\theta) \triangleq \Psi(\theta)N(\theta) \quad (44)$$

which represents the reduced cross-correlation between the subband signals for properly designed $N(\theta)$.

For the prototype AC-FIR and PS-IIR filters, $\psi_n = \mathcal{F}^{-1}\{\Psi(\theta)\}$ is shown in Fig. 5(a) and (b). After introducing the notch filters as in Case 5.4, we have $\hat{\psi}_n = \mathcal{F}^{-1}\{\hat{\Psi}(\theta)\}$, which are, respectively, shown for the AC-FIR and PS-IIR filters in Fig. 5(c) and (d). The energies of ψ_n and $\hat{\psi}_n$ compare as follows: For the PS-IIR filter, the energy decreased from 7.57×10^{-4} to 1.07×10^{-7} . There is almost no correlation between the subband signals (and hence aliasing). For the AC-FIR filter, the energy reduced from 4.12×10^{-3} to 8.68×10^{-4} , the narrowband notch filter is, therefore, ineffective, and there remains significant correlation between subband signals. If notch filters with wide bandwidths are used, the near-end speech will be distorted significantly [35]. In Fig. 5(c), $\hat{\psi}_n$ is spread over a wider range. Consequently, long $C_{0,1}(z)$ and $C_{1,0}(z)$ would be necessary if CRCI was carried out.

VI. NEAR-END SPEECH RECONSTRUCTION, PHASE DISTORTION, AND DELAY

Thus far, the issues concerning the phase distortion introduced by the IIR subband decomposition have not been addressed. In order to concentrate on this point, consider the cascaded MRAB and MRSB in Fig. 6(b), where $W(z)$ is the near-end speech, and $\tilde{W}(z)$ is its reconstruction.

Case 6.1 (PR-FIR): In lossless perfect reconstruction filter banks [22] with real coefficients and length $L+1$ (L :odd), the analysis bank filters are given by

$$H_0(z) = H(z), \quad H_1(z) = -z^{-L}H(-z^{-1}) \quad (45)$$

and the corresponding synthesis bank filters are

$$G_0(z) = 2z^{-L}H(z^{-1}), \quad G_1(z) = -2H(-z). \quad (46)$$

With these choices, the input-output transfer function in Fig. 6(b) is

$$T(z) \triangleq \frac{\tilde{W}(z)}{W(z)} = z^{-L}. \quad (47)$$

In other words, $W(z)$ is identical to $\tilde{W}(z)$ except for some delay.

Case 6.2 (AC-FIR): The transfer function is

$$T(z) = H^2(z) - H^2(-z). \quad (48)$$

In this case, if L is odd, there is an implicit notch at $\theta = \pi/2$ [26]. This is not a desirable notch since it generally has a wide bandwidth. If L is even, there is no notch, but some amplitude distortion occurs around $\theta = \pi/2$. There is no phase distortion for linear phase $H(z)$ since

$$T(\theta) = (|H^2(\theta)| - |H^2(\theta - \pi)|)e^{-jL\theta}, \quad (49)$$

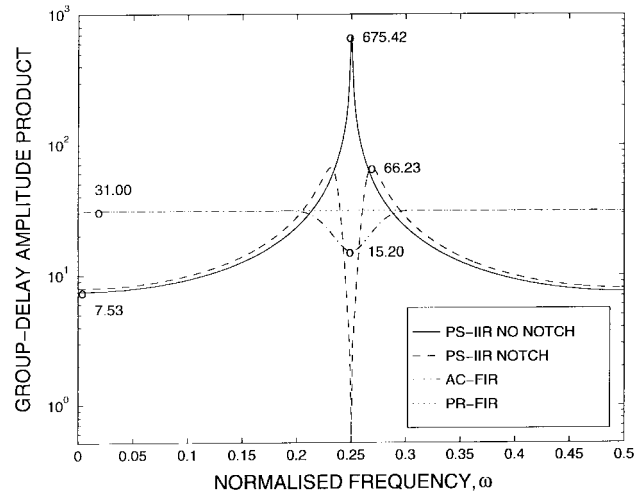


Fig. 7. Group-delay amplitude product for IIR and FIR multirate filterbanks for two-band decomposition.

Case 6.3 (PS-IIR): We have

$$T(z) = z^{-1}A_0(z^2)A_1(z^2). \quad (50)$$

Since $|A_0(\theta)| = |A_1(\theta)| = 1$, the amplitude reconstruction is perfect. However, the phase reconstruction is not. The phase response $\angle(\cdot)$ of $T(z)$ is

$$\angle T(\theta) = -\theta + \angle A_0(2\theta) + \angle A_1(2\theta) \quad (51)$$

$$\angle A_i(2\theta) = -2 \sum_{j=0}^{P_i-1} \arctan \left(\frac{1 - \alpha_{i,j}}{1 + \alpha_{i,j}} \tan(\theta) \right) \quad (52)$$

where $i = 0, 1$. When the notch filter is introduced

$$\tilde{T}(z) = z^{-1}A_0(z^2)A_1(z^2)N(z) \quad (53)$$

$$\angle \tilde{T}(\theta) = -\theta + \angle A_0(2\theta) + \angle A_1(2\theta) + \angle N(\theta). \quad (54)$$

Therefore, the amplitude reconstruction is no longer perfect. However, this is not a serious problem since $N(z)$ has a narrow bandwidth. Besides, the notch attenuates the part of the spectrum where the phase distortion is most significant.

To clarify this point, we need a measure *specific* to the structure in Fig. 6(b). A useful measure to jointly investigate the amplitude and phase reconstruction is the group-delay amplitude product (GAP) $\mathcal{G}(\theta)$, which is defined by

$$\mathcal{G}(\theta) \triangleq -|T(\theta)| \frac{\partial \angle T(\theta)}{\partial \theta}. \quad (55)$$

This is a sensible measure since at least one of the reconstruction properties (amplitude or phase) is perfect for the filter banks under investigation. Knowing this, if $\mathcal{G}(\theta)$ is a constant in $\theta \in [0, \pi]$, then the amplitude and the phase reconstructions are perfect. The gains of the cascaded MRAB and MRSB's are almost unity except the AC-FIR case around $\theta = \pi/2$. Therefore, if $\mathcal{G}(\theta)$ is close to unity, then the AEC system has a low delay. For the filter banks under investigation, $\mathcal{G}(\theta)$ for a two-band decomposition is shown in Fig. 7.

For the PR-FIR multirate filterbanks, we have the desirable situation $\mathcal{G}(\theta) = L$. However, L can generally not be made very small since selective prototype lowpass filters are

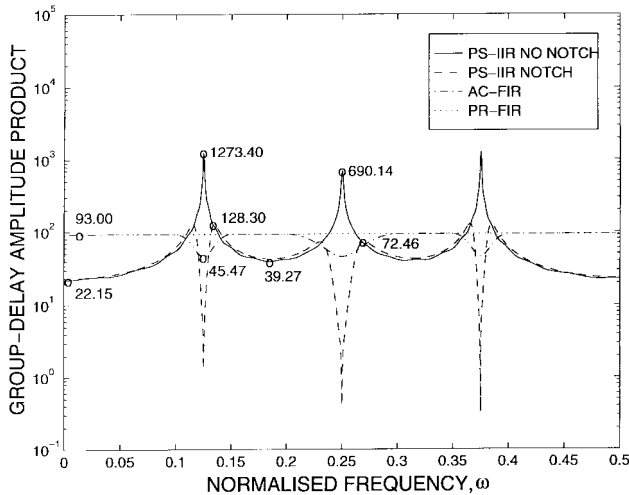


Fig. 8. Group-delay amplitude product for IIR and FIR multirate filterbanks for four-band decomposition.

more desirable for lower aliasing. For the AC-FIR multirate filterbanks, we have

$$\mathcal{G}(\theta) = L(|H^2(\theta)| - |H^2(\theta - \pi)|) \quad (56)$$

which is almost constant, except around $\theta = \pi/2$.

We investigate the PS-IIR multirate filterbanks in two steps: with and without notch filtering. If there is no notch filter, we have

$$\mathcal{G}(\theta) = -\frac{\partial \angle T(\theta)}{\partial \theta}. \quad (57)$$

This is not a constant either, and moreover, its value is very high around $\theta = \pi/2$, which is solely due to the significant phase distortion since the amplitude reconstruction is perfect. However, when the notch filter is introduced, we have

$$\mathcal{G}(\theta) = -|N(\theta)| \frac{\partial \angle \tilde{T}(\theta)}{\partial \theta}. \quad (58)$$

This indicates that although the tonal component around $\theta = \pi/2$ has a high delay, its amplitude is suppressed by the notch filter as shown in Fig. 7.

Before we proceed to the next section, it is worth looking into a four-band binary-tree subband decomposition [22]. Due to the binary-tree subband decomposition, the aliasing bandwidth around $\theta = \pi/4, 3\pi/4$ will be half the aliasing bandwidth around $\theta = \pi/2$. The input-output transfer function can be easily obtained as

$$T(z) = z^{-3} A_0(z^2) A_0(z^4) A_1(z^2) A_1(z^4) N(z) N(z^2). \quad (59)$$

In this case, the notch bandwidths are in agreement with the aliasing bandwidths around $\theta = \pi/4, \pi/2$, and $3\pi/4$. The GAP for the four-band binary-tree subband decomposition is shown in Fig. 8. The effect of the notch filtering is evident.

An informal listening test is performed to evaluate subjectively the phase distortion on the near-end speech. Female and male speech recordings are used. The audience is asked to compare the quality of the speech reconstructed at the output of the cascaded PS-IIR analysis and synthesis banks with the

TABLE II
MEAN OPINION SCORE TEST RESULTS ON NEAR-END SPEECH RECONSTRUCTION

No. of subbands (B)	MOS	
	PS-IIR with notches in MMRABs	PS-IIR without notches
2	4.75	4.63
4	4.62	4.04
8	4.54	2.71
16	4.46	2.08
32	2.67	1.38

original recordings. Mean opinion scores (MOS's) are obtained according to the following impairment scale:

- 5) imperceptible,
- 4) perceptible but not annoying,
- 3) slightly annoying,
- 2) annoying,
- 1) very annoying.

The MOS results in Table II show that without the notch filters, subband decomposition based on PS-IIR filters produces perceptually disturbing artifacts when the number of subbands is increased beyond 4. On the other hand, after introducing the notch filters, the number of subbands can be increased to 16. This is due to the attenuation of the amplitudes of the highly phase distorted frequencies by the notch filters.

Based on Figs. 7 and 8, a comparison between the delays introduced by AC-FIR, PR-FIR, and PS-IIR subband processing schemes can also be performed. In Fig. 7, for FIR based two-band decomposition at 8-KHz sampling rate with 32-tap prototype lowpass filters, the delay on the near-end speech is 3.9 ms (31/8000). Note that this figure would increase if a more selective prototype lowpass filter was used. For the PS-IIR filter, the group delay is nonuniform across the spectrum, and its value is lower than what is obtained for FIR based decomposition except around $\theta = \pi/2$. At $\theta = \pi/2$, the delay is 84.4 ms. However, after introducing the notch filter, a spectral null is introduced at this frequency, and furthermore, the neighboring frequencies with high phase distortion are also attenuated. Based on the GAP measure, the *maximum* value of the delay with the notch filters is around 8.3 ms. In Fig. 8, for a four-band binary-tree decomposition with 32-tap FIR filters, the group-delay is 11.6 ms. The PS-IIR decomposition again yields lower delay except around $\theta = \pi/4, \pi/2$, and $3\pi/4$. In particular, without the notch filter the group delay at $\theta = \pi/4$ is 159.2 ms, but after introducing the notch filter, the highly phase distorted components are suppressed, and the *maximum* delay indicated by the GAP measure is around 16 ms.

VII. IMPLEMENTATION AND COMPUTATIONAL COMPLEXITY

The MMRAB and MRSB of the PS-IIR filter-bank can be efficiently implemented in polyphase form as shown in Fig. 9. In each phase, $A_0(z)$ and $A_1(z)$ are cascaded first-order allpass sections. To implement these sections, compact implementations in Fig. 10(b) can be used. Compared with the canonical implementation in Fig. 10(a), the compact implementations require only one multiplier. Furthermore, compact implementations are more robust to multiplication roundoff noise when implemented in finite-precision arithmetic [16], [22], [34],

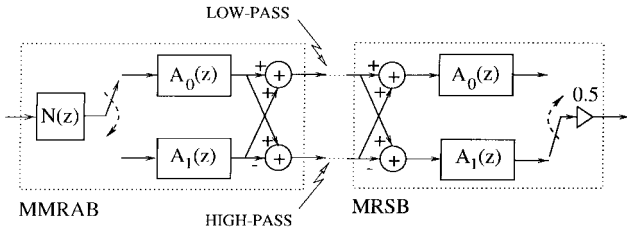


Fig. 9. MMRAB and MRSB of the PS-IIR subband decomposition in polyphase form.

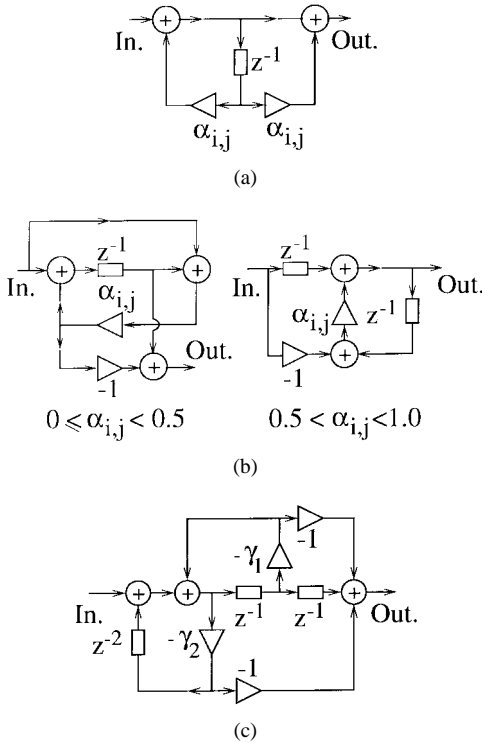


Fig. 10. Allpass network implementations. (a) First-order canonical. (b) Compact. (c) Second-order compact implementation.

[23]. Similarly, the notch filter in (34) can be implemented by using the compact implementations of the second-order allpass networks [34]. A suitable compact implementation is shown in Fig. 10(c).

The number of multiplications per sampling interval for the AC-FIR, PR-FIR, and PS-IIR MRAB and the notch filter are presented in Table III. Note that the number of subbands is denoted by B , which is an integer power of two, and each subband is equispaced in $\theta \in [0, \pi]$. Using the formulae in Table III, the total number of multiplications per sampling interval for the MRAB's, MRSB's, and the notch filters in a B -band AEC are computed and tabulated in Table IV. Note that 32-tap AC-FIR, PR-FIR, and the PS-IIR filter in Fig. 2 are used. The computational efficiency of the PS-IIR filterbanks is evident and would be more visible if longer prototype FIR filters were used for better selectivity. It is also important that the computational complexity of the proposed PS-IIR-based subband AEC system is clearly much lower than subband decomposition with oversampling or cross-adaptive filtering.

In terms of the memory requirement, the PS-IIR subband decomposition with notch filters is also more efficient than

TABLE III
NUMBER OF MULTIPLICATION OPERATIONS PER SAMPLING INTERVAL FOR THE NOTCH FILTER AND THE BINARY-TREE MRABS OF THE PS-IIR (POLYPHASE), AC-FIR (POLYPHASE), AND PR-FIR (PARAUNITARY LATTICE [22]) FILTERBANKS

Notch Filter	PS-IIR	AC-FIR	PR-FIR
$5 \log_2 B$	$0.5(P_0 + P_1) \log_2 B$	$0.5(L + 1) \log_2 B$	$0.5(L + 3) \log_2 B$

TABLE IV
NUMBER OF MULTIPLICATIONS PER SAMPLING INTERVAL IN THE BINARY-TREE ANALYSIS AND SYNTHESIS BANKS OF AEC WITH THE PS-IIR (POLYPHASE), AC-FIR (POLYPHASE) AND PR-FIR (PARAUNITARY LATTICE [22]) BANKS IN FIG. 2

No. of subbands (B)	PS-IIR +			
	PS-IIR	notches in MMRABs	AC-FIR	PR-FIR
2	16.5	26.5	48	51
4	33	53	96	102
8	49.5	79.5	144	153
16	66	106	192	204
32	82.5	132.5	240	255

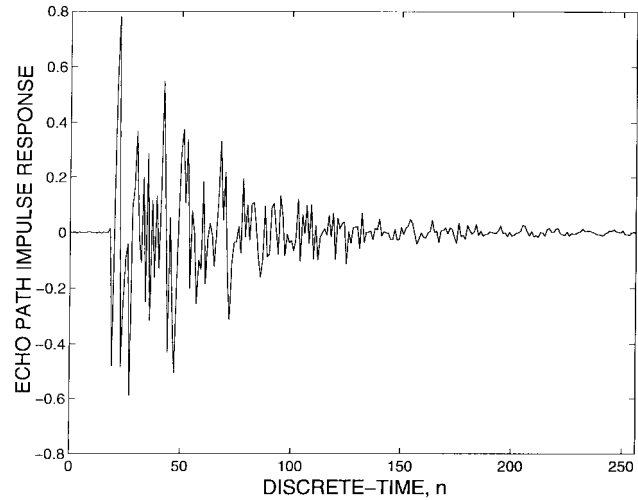


Fig. 11. Echo path impulse response measured in a car cockpit.

AC-FIR or PR-FIR subband decomposition schemes. For instance, for two subbands with PS-IIR decomposition implemented in polyphase form and the notch filter, both realized by using the compact implementations in Fig. 10, 44 memory locations are required to store the filter coefficients and the internal variables. On the other hand, for a two-band AC-FIR subband decomposition with a 32-tap filter, 62 memory locations are required.

VIII. SIMULATIONS

In the first part, no CRCI is performed, and the residual echo signal is investigated for highly selective PS-IIR multirate filterbanks, where the prototype lowpass filter in Fig. 2 is used. The number of subbands B is two. The loudspeaker signal x_n is chosen as zero-mean and white for persistent excitation. An echo path impulse response (Fig. 11) that has 256 samples, measured in a car cockpit, is used.

A fast QR least-squares adaptive algorithm [26, 18] is used in the identification of only the diagonal elements of $C_m(z)$. After the steady state is reached, the residual echo signal is averaged over 20 nonoverlapping blocks, each block being

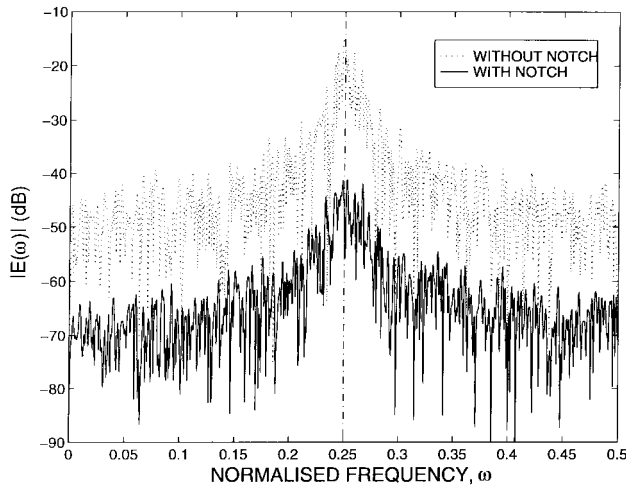


Fig. 12. Amplitude spectrum of the residual echo with and without notch filtering, $\omega = \theta/2\pi$.

2×10^3 samples. The above is repeated after inserting notch filters at 2 and 3 in Fig. 3. The amplitude spectra of the residual echo signals are shown in Fig. 12. The tonal component is clearly visible around $\theta = \pi/2$. After introducing the notch filters, this component is attenuated, and the power of the residual echo signal is decreased around 24 dB's.

In the second part, a set of AEC experiments is conducted with five speech recordings obtained in a car cockpit under a single-talk condition. The sampling frequency was 8 KHz. A binary-tree subband decomposition is used to obtain four equispaced subbands in 0-4 KHz. The AEC performance is assessed by using two objective measures. The first one is the average segmental echo return loss enhancement (A-SERLE), which is defined as

$$\text{A-SERLE} \triangleq \frac{1}{N_S} \sum_{i=1}^{N_S} \left\{ 10 \log_{10} \left(\frac{\sum_{j=(i-1)S}^{iS-1} m_j^2}{\sum_{j=(i-1)S}^{iS-1} e_j^2} \right) \right\} \quad (60)$$

where

- S segment size (with 32 ms duration in our case),
- N_S number of segments,
- m_n echo signal,
- e_n residual echo signal.

The second objective measure is the time of initial convergence (TIC) to a specified level of A-SERLE. The same fast QR adaptive algorithm is used in all subbands. The forgetting factor of this algorithm is chosen as 0.99 in each subband. The number of adaptive coefficients for each subband was 64. The results are in Table V. During this experiment, no CRCI was used. The PR-FIR configuration yields the lowest A-SERLE performance. The highest performance is obtained with the PS-IIR filter-banks when used with the notch filters. The convergence times to 10 dB A-SERLE is the same for all cases. Perhaps the last point to clarify is how much of the ERLE improvement in notched PS-IIR filterbanks is due to the removal of energy from the microphone signal $M(z)$. For

TABLE V
AVERAGE SEGMENTAL ERLE AND TIC FOR REAL
SPEECH RECORDINGS. NUMBER OF SUBBANDS IS 4

FILTER-BANK	A-SERLE (dB)	TIC-10dB (msec.)
PR-FIR	21.81	32.0
AC-FIR	23.83	32.0
PS-IIR	22.40	32.0
PS-IIR+notch	25.32	32.0

four-band processing, the notch filtered microphone signal is $\hat{M}(z) = N(z)N(z^2)M(z)$. The A-SERLE based on \hat{m}_n is found to be 24.38 dB, which is still higher than all the other A-SERLE values in Table V. Let us note that there is also amplitude distortion in the QMF-FIR filterbanks due to the nonperfect amplitude reconstruction.

IX. CONCLUSIONS

Adaptive filtering in subbands when used with powerful adaptive algorithms is an elegant solution for acoustic echo cancellation. The subband nature of the echo canceller reduces the computational complexity of adaptive filtering. The leakage across the subbands (aliasing) is generally a limiting factor. For FIR MRAB and MRSB, cross-adaptive filtering [5] and oversampling [4], [12] have been proposed to mitigate aliasing. However, these solutions increase significantly the computational complexity and/or the delay.

In this paper, an alternative solution is proposed where the filterbanks are forced to be highly selective by using prototype lowpass PS-IIR filters that have several advantages. First, their computational complexity is very low due to the small number of allpass coefficients required, compact implementation of allpass stages, and the polyphase implementation of the resulting multirate filterbanks. Furthermore, PS-IIR filters have good finite-precision characteristics [16]. Second, the aliasing is confined to very narrow bandwidths around the subband division frequencies. When there is no CRCI, the residual echo signal contains tonal components. This result is also shown to be true for FIR subband systems if the filterbanks are highly selective. In this case, an elegant solution is to insert properly designed notch filters into the MRAB's. The proposed notch filters are also PS-IIR filters that are easy to design and implement. Notch filtering attenuates the aliasing between subband signals and suppresses the tonal components that are disturbing to the far-end speaker. The phase distortion on the near-end speech signal is also addressed. The results show that notch filters also mitigate the phase distortion in PS-IIR multirate filterbanks. During listening tests, no perceptually audible distortions occurred with notch filters when the number of subbands is increased to 16. Without notch filters, the phase distortion was annoying if more than four subbands were used.

The suppression of the tonal component is demonstrated experimentally. The echo cancellation performances for the AC-FIR, PR-FIR, PS-IIR, and PS-IIR+notches are compared by using real recordings from a car cockpit. A fast QR least-squares adaptive algorithm is used in subbands. The results show that the PS-IIR multirate filterbanks, when combined with notch filtering, produce the most cost-effective solution and the highest echo cancellation performance.

ACKNOWLEDGMENT

The authors would like to thank the associate editor and the anonymous reviewers for their valuable comments.

REFERENCES

- [1] D. A. Berkley and O. M. M. Mitchell, "Seeking the ideal in "Hands-free" telephony," *Bell Labs. Rec.*, pp. 318–324, Nov. 1974.
- [2] A. Gilloire, "Recent advances in adaptive filtering algorithms for acoustic echo cancellation," in *Proc. Int. Workshop Acoust. Echo Noise Control*, Roros, Norway, June 1995, pp. 115–134.
- [3] W. Kellermann, "Analysis and design of multirate systems for cancellation of acoustical echoes," in *ICASSP-88*, New York, NY, Apr. 1988, pp. 2570–2573.
- [4] D. T. M. Slock, "Fractionally-spaced subband and multiresolution adaptive filters," in *ICASSP-91*, Toronto, Ont., Canada, May 1991, pp. 3693–3696.
- [5] A. Gilloire and M. Vetterli, "Adaptive filtering in sub-bands with critical sampling: Analysis, experiments and application to acoustic echo cancellation," *IEEE Trans. Signal Processing*, vol. 40, pp. 1862–1875, Aug. 1992.
- [6] G. V. Moustakidis and S. Theodoridis, "Fast Newton transversal filters—A new class of adaptive estimation algorithms," *IEEE Trans. Signal Processing*, vol. 39, pp. 2184–2193, Oct. 1991.
- [7] T. Petillon, A. Gilloire, and S. Theodoridis, "The fast Newton transversal filter: An efficient scheme for acoustic echo cancellation in mobile radio," *IEEE Trans. Signal Processing*, vol. 42, pp. 509–518, Mar. 1994.
- [8] J. E. Hart, P. A. Naylor, and O. Tanrikulu, "Polyphase all-pass IIR structures for sub-band acoustic echo cancellation," in *Proc. EUROSPEECH-93*, Berlin, Germany, Sept. 1993, pp. 1813–1816.
- [9] F. Capman, J. Boudy, and P. Lockwood, "Acoustic echo cancellation using a fast QR-RLS algorithm and multirate schemes," in *Proc. ICASSP-95*, Detroit, MI, May 1995, pp. 969–972.
- [10] M. Courville and P. Duhamel, "Adaptive filtering in sub-bands using a weighted criterion," in *Proc. ICASSP-95*, Detroit, MI, pp. 985–988, May 1995.
- [11] J. Benesty, F. Amand, A. Gilloire, and Y. Grenier, "Adaptive filtering algorithms for stereophonic acoustic echo cancellation," in *Proc. ICASSP-95*, Detroit, MI, May 1995, pp. 3099–3102.
- [12] P. A. Naylor and J. E. Hart, "Subband acoustic echo control using noncritical frequency sampling," in *Proc. EUSIPCO-96*, Trieste, Italy, Sept. 1996, pp. 37–40.
- [13] E. Hansler, "The hands-free telephone problem: A second annotated bibliography update," in *Int. Workshop Acoustic Echo Noise Control*, Roros, Norway, June 1995, pp. 107–114.
- [14] S. Gay and S. Tavathia, "The fast affine projection algorithm," in *Proc. ICASSP-95*, Detroit, MI, May 1995, pp. 3023–3026.
- [15] E. Moulines, O. A. Amrane, and Y. Grenier, "The generalized multideelay adaptive filter—Structure and convergence analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 14–28, Jan. 1995.
- [16] O. Tanrikulu, B. Baykal, A. G. Constantinides, J. A. Chambers, and P. A. Naylor, "Finite-precision design and implementation of all-pass polyphase networks for echo cancellation in sub-bands," in *Proc. ICASSP-95*, Detroit, MI, May 1995, pp. 3039–3042.
- [17] U. Iyer, M. Nayeri, and H. Ochi, "IIR sub-band adaptive filtering," in *36th Midwest Symp. Circuits Sys.*, Detroit, MI, Aug. 1993, pp. 125–128.
- [18] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 2nd ed., 1991.
- [19] B. Baykal, "Undetermined recursive least squares adaptive filtering," Ph.D. dissertation, Imperial Coll. Sci., Technol., Medicine, London, U.K., 1995.
- [20] A. Gilloire, "Experiments with sub-band acoustic echo cancellers for teleconferencing," in *Proc. ICASSP-87*, Dallas, TX, Apr. 1987, pp. 2141–2144.
- [21] A. Gilloire and M. Vetterli, "Adaptive filtering in sub-bands," in *Proc. ICASSP-88*, New York, Apr. 1988, pp. 1572–1575.
- [22] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [23] R. A. Valenzuela and A. G. Constantinides, "Digital signal processing schemes for efficient interpolation and decimation," *Proc. Inst. Elec. Eng.*, vol. 130, pp. 225–235, Dec. 1983.
- [24] O. Tanrikulu, B. Baykal, A. G. Constantinides, and J. A. Chambers, "Residual signal in sub-band acoustic echo cancellers," in *Proc. EUSIPCO-96*, Trieste, Italy, Sept. 1996, pp. 21–24.
- [25] O. Tanrikulu, B. Baykal, and P. A. Naylor, Eds., "Advanced solutions for echo cancellation." EC FREETEL (no. 6166), Tech. Rep. D.3.123.2, 1994.
- [26] J. G. Proakis, C. M. Rader, F. Ling, and C. L. Nikias, *Advanced Digital Signal Processing*. New York: Academic, 1992.
- [27] P. P. Vaidyanathan, S. K. Mitra, and Y. Neuvo, "A new approach to the realization of low-sensitivity IIR digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 350–361, Apr. 1986.
- [28] F. Harris and A. G. Constantinides, "Design and implementation of efficient resampling filters using polyphase recursive allpass filters," in *Proc. 25th Asilomar Conf. Signals Syst., Comput.*, Pacific Grove, CA, Nov. 1991, pp. 1031–1036.
- [29] O. Tanrikulu, "Adaptive signal processing algorithms with accelerated convergence and noise immunity," Ph.D. dissertation, Imperial Coll. Sci., Technol., Medicine, London, U.K., 1995.
- [30] O. Tanrikulu and M. Kalkan, "Design and discrete re-optimization of all-pass based power-symmetric IIR filters," *Electron. Lett.*, vol. 32, pp. 1458–1460, Aug. 1996.
- [31] M. Vetterli, "A theory of multirate filter-banks," *IEEE Trans. Signal Processing*, vol. SP-35, pp. 356–372, Mar. 1987.
- [32] N. Kalouptsidis and S. Theodoridis, *Adaptive System Identification and Signal Processing Algorithms*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [33] A. G. Constantinides, "Frequency transformations for digital filters," *Proc. Inst. Elec. Eng.*, vol. 117, pp. 1585–1590, Aug. 1970.
- [34] S. K. Mitra and K. Hirano, "Digital all-pass networks," *IEEE Trans. Circuits Syst.*, vol. CAS-21, pp. 688–700, Sept. 1974.
- [35] A. Zitzewitz, "Considerations on acoustic echo cancelling based on realtime experiments," in *Proc. EUSIPCO-90*, Barcelona, Spain, Sept. 1990, pp. 1987–1990.



Oğuz Tanrikulu (M'96) was born in Ankara, Turkey, in 1969. He received the B.Sc. degree (High Hons.) in electrical engineering from the Middle East Technical University (METU), Ankara, Turkey, in 1990 and the M.Sc. and Ph.D. degrees in electrical engineering from the Imperial College of Science, Technology, and Medicine, London, U.K., in 1991 and 1995, respectively.

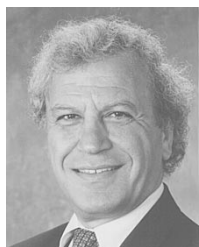
In 1992, he was appointed as Research Assistant at Imperial College, and until 1995, he was involved in the EC ESPRIT project "FREETEL." Currently, he is a Research Associate at Imperial College, working on the EPSRC research project "Multirate Echo Cancellation." His research interests are in adaptive algorithm design and analysis, filter banks, echo cancellation, and blind equalization for hands-free and wireless systems, optimization theory, and higher order statistics.

In 1990, Dr. Tanrikulu was awarded the British Council scholarship funded by the Foreign and Commonwealth Office to pursue the M.Sc. program at Imperial College.



Buyurman Baykal (M'93) was born in Ankara, Turkey, in 1968. He received the B.Sc. degree (High Hons.) from the Middle East Technical University (METU), Ankara, Turkey, in 1990, and the M.Sc. (Distinction) and Ph.D. degrees from the Imperial College of Science, Technology, and Medicine, London, U.K. in 1992 and 1995, respectively, all in electrical and electronic engineering.

He held the British Council fellowship from 1991 to 1992 for postgraduate studies in the U.K. He has been with Imperial College since 1993 as a research assistant and research scientist, where he has been involved with various research projects on echo cancellation, noise reduction, and channel equalization. His research interests include blind and nonblind adaptive algorithms and their applications in communications, wireless systems, and hands-free communications.



Anthony G. Constantinides (SM'78) is Professor of Signal Processing and the Head of the Signal Processing and Digital Systems Section of the Department of Electrical and Electronic Engineering at Imperial College, London, U.K. He has been actively involved with research in various aspects of digital filter design, digital signal processing, and communications for a number of years. His research spans a wide range of digital signal processing, both from the theoretical as well as the practical points of view. His recent work has been directed

toward the demanding signal processing problems arising from the area of telecommunications. He has published a range of books and papers in learned journals in the area of digital signal processing and its applications.

Dr. Constantinides served as the first president of the European Association for Signal Processing (EURASIP) and has contributed in this capacity to the establishment of the *European Journal for Signal Processing*. He has been on, and is currently serving as, a member of many technical program committees of IEEE international conferences. He organized the first ever international series of meetings on digital signal processing in London, initially in 1967, and in Florence (with V. Cappellini) since 1972. In 1985, he was awarded the Honor of Chevalier, Palmes Academiques, by the French government, and in 1996, he was promoted to Officer of the Palmes Academiques. He holds several visiting professorships and other fellowships and honors around the world. He was recently elected to the Board of Governors of the IEEE Signal Processing Society. He is also serving as a member of the IEEE Signal Processing Society Technical Committee on Neural Networks for Signal Processing.



Jonathon A. Chambers (M'93) was born in Peterborough, U.K., in 1960. After an electronics artificer apprenticeship in the Royal Navy, he received the first class B.Sc.(Hons.) degree in electrical and electronic engineering from the Polytechnic of Central London, U.K., receiving the Robert Mitchell Medal as the top graduate in 1985. He was then appointed to a lectureship at the Polytechnic of Central London, teaching courses in digital signal processing. He received the Ph.D. degree in adaptive signal processing in 1990 after studying at Cambridge

University, U.K., and Imperial College, London, U.K.

He spent three years as a research scientist at Schlumberger Cambridge Research, applying adaptive signal processing techniques to oilfield-related applications. He returned to a lectureship in Signal Processing in the Department of Electrical and Electronic Engineering at Imperial College in 1994. He has authored or co-authored many technical publications on adaptive signal processing and is currently running research projects in adaptive signal processing and its application in mobile communication systems.

Dr. Chambers is a member of the IEE professional group committee E5 on signal processing, a guest editor for the *International Journal of Adaptive Control and Signal Processing*, and an associate editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING.