A Technique for Realizing Linear Phase IIR Filters

Scott R. Powell and Paul M. Chau, *Member, IEEE*

Abstract—**A new real-time IIR filter structure is presented that possesses exact phase linearity with 10 ~ 1000 times fewer general multiplies than conventional FIR filters of similar performance and better magnitude characteristics than equiripple or maximally flat group delay IIR filters. This structure is based on a novel technique using local time reversal and single pass sectioned convolution methods to realize a real-time recursive implementation of the noncausal transfer function $H(z^{-1})$. When combined with a causal recursive filter, the precisely linear phase transfer function $H(z) \cdot H(z^{-1})$ results with fewer multiplies and significantly reduced storage requirements than previously reported noncausal recursive approaches. No constraints are placed on the magnitude or phase response characteristics of $H(z)$. The new linear phase IIR filter can be designed using standard magnitude-only IIR filter design programs and can be made unconditionally free from parasitic zero-input limit cycle or overflow oscillations. The resulting group delay is typically constant to within 10^{-5} of the average delay. For a given filter specification, the number of multiplies per sample is typically within a factor of two and the roundoff noise within one-half bit of a causal magnitude-only elliptic IIR implementation. Computer simulations are used to verify the expected phase linearity and magnitude response. A design approach is presented including a technique for establishing the number of samples per convolution section necessary to insure that a time invariant system results with linear phase and magnitude response equal to $|H(z)|^2$. The relationship between the number of samples per convolution section and the total harmonic distortion of the system is also verified through simulation. A comparison to direct FIR implementations for nine example filter specifications is presented to demonstrate the computational advantage of the new linear phase IIR structure.**

I. Introduction

A HIGHLY desirable characteristic of almost any filter, digital or analog, is linear phase response. Precise linear phase requires that the filter transfer function satisfy the relationship $H(z) = H(z^{-1})$ implying that poles and zeros exist in mirror-image pairs. Although this condition can be easily satisfied by all-zero FIR implementations, it would appear that IIR filters cannot be designed to exhibit exact linear phase since poles with radius of less than one require mirror-image poles with radius greater than one rendering the filter unstable. Unfortunately, linear phase FIR filters generally require many times more multiplies per output sample to satisfy stringent magnitude response specifications than IIR filters. In linear phase applications where the computational burden of an FIR approach is excessively high, schemes for deriving acceptable phase responses from IIR filters may provide an appropriate alternative.

Although precisely linear phase is not generally attainable from conventional IIR structures, several techniques have been suggested to approximate passband phase linearity. For example, IIR filtering can be followed by phase equalization via an allpass filter [8]. However, the increased computational burden often results in more multiplies per sample than an optimal FIR implementation [9]. More commonly, IIR filters are designed with equiripple or maximally flat group delay [10]. Although these filters often result in poor magnitude characteristics, transmission zeros can be introduced to offer some improvement in the magnitude response [11]. Alternatively, optimization can be used to simultaneously approximate both magnitude and phase response specifications [12], [13]. Since high amplitude selectivity and phase linearity are generally conflicting criteria, [14] and [15] suggest IIR structures based on a parallel connection of all-pass subfilters which possess the desirable property that decreasing phase error results in decreasing magnitude error. This structure can result in fewer multiplications than an FIR filter but the order increases much faster than a direct IIR filter as the magnitude specifications become increasingly difficult. Although the above techniques substantially improve the phase response of recursive filters, only approximately linear phase results which may not be suitable in phase sensitive applications.

For the limited class of problems where filtering is permissible off-line, a noncausal filter with zero phase can be realized using a combination of multipass filtering and time reversal [1], [2]. By using different filters for each pass, this scheme can be used to realize general noncausal filter functions where poles outside the unit circle are not necessarily mirror images of poles inside the unit circle. Unlike the causal IIR filter design problem, there is no conflict between stability and phase error in designing general noncausal IIR filters. Optimization-based design strategies have been applied to the design of general noncausal IIR filters resulting in significant performance improvements over causal IIR filters in terms of magnitude and phase error [7]. The computational advantage of using noncausal recursive filters versus linear phase FIR filters is clearly indicated in [2]–[7], and the results of this paper. However, an IIR filter based implementation of this

Manuscript received October 23, 1989; revised September 27, 1990. This work was supported in part by the Hughes Aircraft Company fellowship program and the NSF Initiation Grant MIP-99-10437.

The authors are with the University of California, San Diego, La Jolla, CA 92037.

IEEE Log Number 9102461.

1053-587X/91/1100-2425$01.00 © 1991 IEEE

scheme is only possible for finite length inputs since this is the only condition under which true time reversal can be realized directly.

Noncausal filters can be extended to process infinite length input sequences using block processing techniques [3]–[6]. Using this technique, a continuous input is divided into finite sections and each section is filtered separately by a two-pass noncausal filter. The resulting output sections are truncated after an "appropriate" number of samples and are combined using overlap-add or overlap-save methods to form a continuous output. However, the section truncation introduces an unavoidable systematic error which degrades the linearity of the overall system. The magnitude of this error depends on the section length in comparison with the significant length of the impulse response of the overall system's transfer function. To keep this error acceptably small, sections must be very long to implement even modest filtering operations which requires substantial amounts of data storage. Additionally, each output section is created from a two-pass filtering operation, forward and reverse. Each final output sample is created from samples from two adjacent output sections. Thus, four filtering operations must be performed for each final output sample. The large storage and computational requirements of these approaches make VLSI implementations difficult and limit their use in real-time processing applications.

In this paper, a technique is developed for noncausal IIR processing of infinite length input sequences where block processing is performed after the first pass of a conventional two-pass noncausal IIR filter. The noncausal portion of the overall transfer function is created in the first pass and the causal portion in the second pass. This new structure offers two distinct advantages over previous noncausal approaches:

1) Reduced Data Storage: For the linear phase case, the transfer function of the first pass is the square root of the overall transfer function with a much shorter significant impulse response length. Since section length is proportional to the impulse response length, much shorter section lengths are required than if block processing is performed after both passes (see Fig. 6).

2) Reduced Computation: Block processing is used to create the noncausal portion only. Overlap-add operations are performed on adjacent output sections from a single pass requiring only two filtering operations to create a single output sample. A single filtering operation is then used to implement the causal portion. Thus, a total of three filtering operations are necessary for each final output sample as opposed to the previously required four.

By performing the first and second passes with the same filter $H(z)$, the phase of any IIR filter can be made exactly linear without degrading or constraining the magnitude response or altering the scaling or dynamic range properties of $H(z)$. Using this time reversed section scheme, precisely linear phase filters can be implemented with efficient IIR structures designed solely by magnitude specifications.

The new time reversed section technique used to realize exactly linear phase IIR filters is described in Section II along with block diagram implementations and analytical derivations of system equations. Section III presents an analytical investigation of the effects of finite section length on the sectional convolution. A simulation methodology is developed in Section IV to address the special requirements of simulating a time reversed section filter. In Section V, a design example is presented, together with computer simulations to illustrate performance, in terms of overall magnitude response and phase linearity, as a function of finite section length. Nine example filter specifications are used in Section VI to compare the performance and complexity of the time reversed section technique to a direct FIR implementation. Finally, in Sections VII and VIII, future work and variations are given and conclusions discussed.

II. Time Reversed Section Scheme

A necessary condition for the causality of linear phase filters is that the impulse response be finite in duration. It is possible to construct a family of infinitely long continuous time functions with appropriate zero crossings which result in finite length sampled versions [16]. However, stable IIR implementations of this restricted class of functions are noncausal, and causal IIR implementations are unstable. The well-known LaGrange structure [9] uses recursive structures to realize a finite length impulse response but relies on pole-zero cancellation which is difficult to carry out in practice because of dynamic range limitations. A recursive structure can only be made to have linear phase if 1) its impulse response is of finite duration, or 2) schemes can be devised which result in an apparent truncation of the filter's impulse response. In this section, we develop a scheme to approximate the truncation of the impulse response from any stable recursive filter and show how this scheme can be used to perform the equivalent of noncausal linear phase filtering. No additional constraints are placed on the recursive filter's architecture, pole locations, or filter order. Standard IIR filter design programs can generally be used. The new linear phase structure results in filters with fewer distinct multiplications per sample than direct FIR implementations because the impulse response values are implemented recursively. Also, the resulting structure is free from zero-input limit cycles regardless of word length, arithmetic, or implementation.

Neglecting parasitic oscillations, the impulse response from the physical realization of an IIR function is always of finite duration. This is true because at some point, the impulse response decays to a value which is less than the resolution afforded by a finite word length. If the input sequence to a recursive filter is divided into sections of length L, where $L \geq$ impulse response length of the physical realization, the corresponding output sections will be of length $\leq 2L - 1$. Using the well-known overlap-add method for sectioned convolution [9], the output can be

constructed from a superposition of finite length responses from adjacent input sections of length L:

$$y(n) = \sum_{k=0}^{\infty} y_k(n), \qquad \text{where}$$

$$y_k(n) \equiv \begin{cases} y(n) & kL \le n \le (k+1)L - 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$\begin{aligned} y_k(n) &= \sum_{m=n-L}^{n} x(m)h(n-m) \\ &= \sum_{m=kL}^{n} x(m)h(n-m) + \sum_{m=n-L}^{kL-1} x(m)h(n-m) \\ &= X_k^L + X_{k-1}^T \end{aligned} \tag{1}$$

where $X_k^L \equiv$ first L outputs from $h(n) * x_k(n)$, and $X_{k-1}^T \equiv$ last L outputs from $h(n) * x_{k-1}(n)$.

The implications of the effective truncation to $2L$ samples of each section is discussed in the following section. Fig. 1 presents an implementation of (1) where the input is split into sections which are filtered separately and combined to form the output. Duplicate filters $H(z)$ are recursive structures with physical impulse response lengths $\le L$. The output is composed of the leading response (first L output samples) from one section plus the trailing response (L output samples after input is removed) from the preceding section. Note that since the state variables of each recursive filter are alternatively reset to zero, the overall structure cannot sustain zero input limit cycle or overflow oscillations. An analogous structure corresponding to the overlap-save method is also possible. This configuration also possesses zero input limit cycle immunity and does not require the addition of the filter outputs as in the overlap-add method.

The overlap-add and overlap-save schemes can also be used to implement noncausal time reversed convolution, $h(-n) * x(n)$. Fig. 2 illustrates the section sequences resulting from a true time reversed convolution between $x(n)$ and $h(-n)$. Note that output sequences are computed by "sliding" the impulse response, $h(-n)$, past the input in the reverse time direction. With the assumption that $h(-n) = 0$ for $n > L$, each section of the output sequence only depends on two adjacent sections of the input convolved with $h(-n)$:

$$\begin{aligned} y_k(n) &= h(-n) * x(n) \qquad kL \le n \le (k+1)L - 1 \\ &= \sum_{m=n}^{n+L} x(m)h(n-m) \\ &= \sum_{m=n}^{(k+1)L-1} x(m)h(n-m) \\ &\quad + \sum_{m=(k+1)L}^{n+L} x(m)h(n-m) \\ &= h(-n) * x_k(n) + h(-n) * x_{k+1}(n). \end{aligned} \tag{2}$$

Similar to forward time convolution, each output section $kL \le n \le (k+1)L - 1$ consists of the leading response

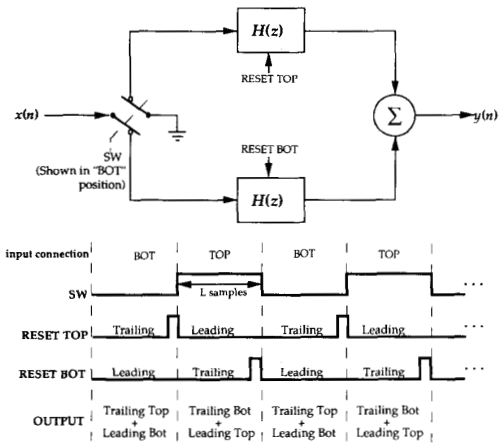


Fig. 1. Overlap-add IIR truncator block diagram and timing.

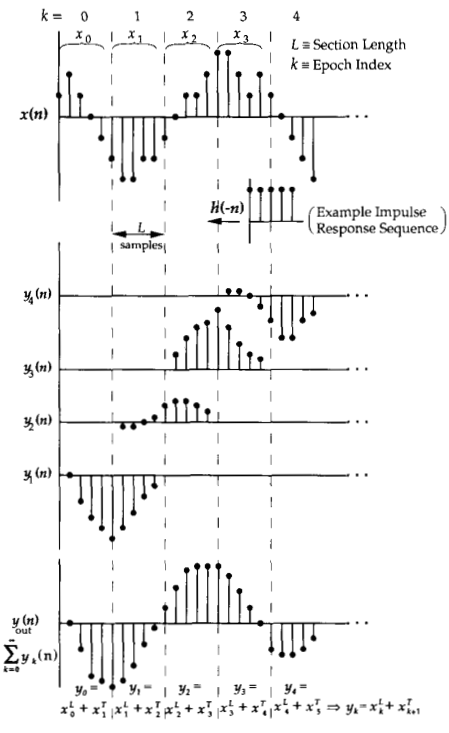


Fig. 2. Overlap-add sequences resulting from $y(n) = h(-n) * x(n)$ with L sample long sections.

(first L outputs) from $h(-n) * x_k(n) \equiv X_k^L$ plus the trailing response (last L outputs) from $h(-n) * x_{k+1}(n) \equiv X_{k+1}^T$, i.e., $y_k = X_k^L + X_{k+1}^T$. This suggests the following scheme:

1) Time reverse each input section $x_k(n)$ to create $x_k(-n)$ and save the responses of $h(n) * x_k(-n)$.

2) Create output sections $y_k(-n)$ consisting of the trailing response from the current input section plus the leading response from the previous input section.

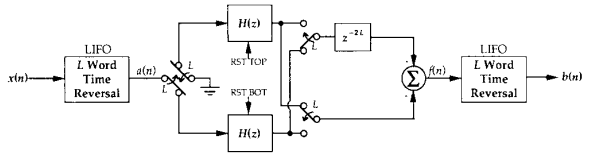


Fig. 3. Block diagram implementation of $b(n) = h(-n) * x(n)$ using overlap-add sectioned convolution.

3) Time reverse output sections $y_k(-n)$ and combine according to $X_k^L + X_{k+1}^T$ to create $y(n) = h(-n) * x(n)$.

Fig. 3 is a block diagram of this scheme requiring two identical copies of the original recursive filter $H(z)$, a total of $4L$ words of LIFO and shift register memory, and an adder. The timing of this scheme is the same as shown in Fig. 1. Fig. 3 represents a standalone realization of a noncausal filter $H(z^{-1})$ where all poles and zeros are outside the unit circle. The time reversed section technique permits the filter to operate on continuous, arbitrary input sequences of infinite length.

A key aspect of this scheme is a real-time implementation of the L word time reversal operation. This function is responsible for splitting continuous data streams into sections L samples long and reversing the sequence of each section. For example, a forward time data stream of $x(n) = \{x(0), x(1), \cdots\}$ results in the following locally time reversed data stream:

$$a(n) = \{\underbrace{x(L-1), \cdots, x(0)}_{\text{Section 0}}, \underbrace{x(2L-1), \cdots, x(L)}_{\text{Section 1}},$$

$$\underbrace{x(3L-1), \cdots, x(2L)}_{\text{Section 2}}, x(4L-1), \cdots\}.$$

A single L word LIFO memory stack can be used to implement the required L word time reversal operation. To illustrate, assume block processing is to be performed on sections containing 5 samples ($L = 5$). In this case, the first section, $k = 1$, consists of input samples $x(0)$–$x(4)$, the second section, $k = 2$, consists of input samples $x(5)$–$x(9)$, and so on. Fig. 4 illustrates snapshots of the contents of a 5 word LIFO implementing a continuous 5 word time reversal operation. Initially, each sample is loaded sequentially from the top until the first input section fills the LIFO ($n = 0$–4). Next, the LIFO input and output ports are swapped so that the LIFO is in the state indicated by the top left snapshot. The first section ($k = 1$) is read out sequentially from the LIFO's top port in reverse order, as required, while the next input section is sequentially loaded into the LIFO from the bottom ($n = 5$–9). When the first section has been completely read out (in reverse order) the LIFO contains the second input section. The LIFO ports are again swapped leaving the LIFO in the state shown in the $n = 10$ snapshot. The operation continues by sequentially unloading the second section (in reverse order) and loading in the third section ($n = 10$–14). This process is carried out indefinitely providing the necessary continuous L word time reversal operation.

The output sequence from the reversed time convolu-

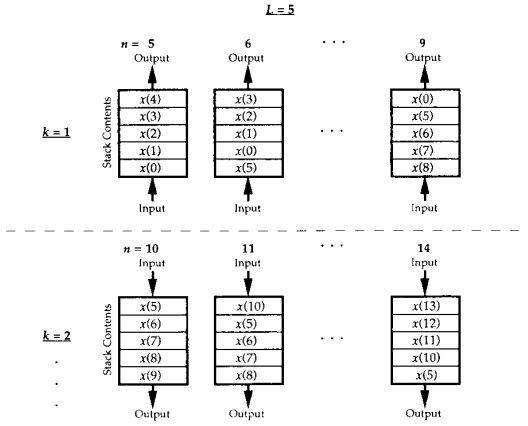


Fig. 4. Single LIFO stack realization of local time reversal; memory contents for successive sample periods. k = section (epoch) index, n = sample index.

tion implementation of Fig. 3 is analytically derived below. The impulse response $h(n)$ is assumed to be zero for $n > L$. The implications of this assumption are considered in the following sections:

$$\alpha(n) \equiv \sum_{k=0}^{\infty} \alpha_k(n)$$

$$\alpha_k(n) \equiv \begin{cases} \alpha(n) & kL \le n \le (k+1)L - 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha = x, a, f, b$

$$a_k(n) = x_k[(2k+1)L - 1 - n] = x_k(n')$$

(in-place reversal)

$$\left\{\begin{array}{l}\text{note: } a_{k-p}(n-q) = x_{k-p}(n' - 2pL + q) \\ p, q = \text{integers}\end{array}\right\} \tag{3}$$

$$f(n) = \sum_{k=0}^{\infty} [h(n) * a_{k-1}(n) + h(n) * a_{k-2}(n - 2L)]$$

$$= \sum_{k=0}^{\infty} [h(n) * x_{k-1}(n' - 2L) + h(n) * x_{k-2}(n' - 2L)]$$

where $n' = (2k+1)L - 1 - n$ (4)

$$b(n) = \sum_{k=0}^{\infty} [h(n') * x_{k-1}(n + 2L) + h(n') * x_{k-2}(n + 2L)] \quad \text{(in-place)}$$

$$= \sum_{k=0}^{\infty} h(L-1-n) * x_{k-1}(n) + h(L-1-n) * x_{k-2}(n)] \quad \text{(non in-place)}$$

since $h_k(n') = 0 \ \forall k \neq 0$

$$= h(L-1-n) * x(n) = h(-n) * x(n-L-1)$$

from (2). (5)

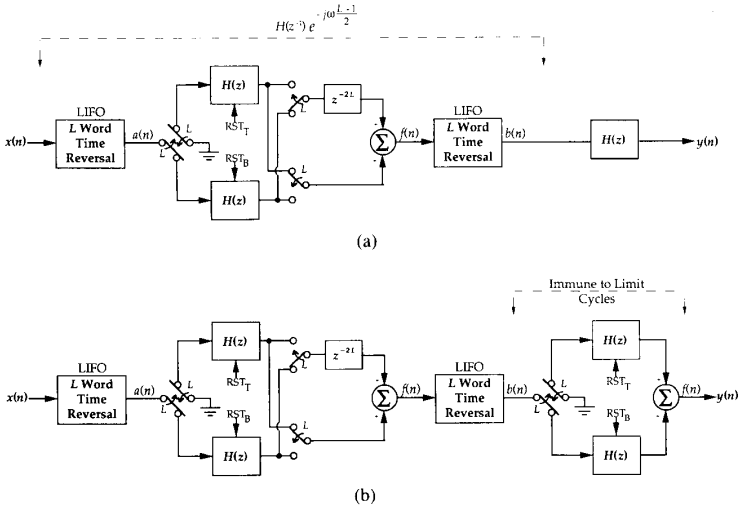


Fig. 5. Time reversed section linear phase filter implementing $H(z)H(z^{-1})e^{-j\omega(L-1)/2}$. (a) Straight implementation. (b) Limit cycle free implementation.

Note that $k - 2$ is substituted for k in (2) to arrive at (5). From (5), the equivalent transfer function for the structure of Fig. 3 is the noncausal function $H(z^{-1})e^{-j\omega(L-1)/2}$. Because the state variables are reset with each new section, the structure of Fig. 3 is unconditionally free from zero input limit cycle oscillations. Note that for clarity, an in-place (delay less) time reversal was assumed in (3). In practice, an L word time reversal operation implemented as in Fig. 4 results in an L sample delay. The final results of (5) incorporate the additional $2L$ sample delay.

Typically, recursive implementations of a function $H(z)$ require scaling to avoid internal overflow. The values of the internal scale factors are usually determined with the assumption that the input is scaled in the L_2 norm sense, i.e., $\|X(e^{j\omega})\|_2 = 1$ [17]. It follows from (3) that $A_k(e^{j\omega})$ differs from $X_k(e^{j\omega})$ only in phase. If $X_k(e^{j\omega})$ is scaled in the L_2 sense

$$\|A_k(e^{j\omega})\|_2 = \left[\frac{1}{2\pi}\int_{-\pi}^{\pi}|A_k(e^{j\omega})|^2\right]^{1/2} = \left[\frac{1}{2\pi}\int_{-\pi}^{\pi}|X_k(e^{j\omega})|^2\right] = \|X_k(e^{j\omega})\|_2 = 1 \tag{6}$$

implying $A_k(e^{j\omega})$ is also scaled in the L_2 sense. Thus, the LIFO time reversal operation preceding the IIR subfilter $H(z)$ in Fig. 3 does not affect $H(z)$'s internal scaling methodology. This also implies that if $H(z)$ is stable with input $x_k(n)$, it is also stable with input $a_k(n)$.

A linear phase filter can now be formed by cascading the time reversed section implementation of the noncausal function $H(z^{-1})$ shown in Fig. 3 with the original causal IIR filter $H(z)$. By passing the sequence $(b)n = h(-n) * x(n)$ back through $H(z)$, the equivalent transfer function becomes

$$H_{eq}(z) = H(z)H(z^{-1})e^{-j\omega(L-1)/2} \tag{7}$$

which is linear phase with magnitude equal to the square of the magnitude of the original IIR filter $H(z)$. A block diagram of this linear phase IIR filter is shown in Fig. 5(a) and represents the fundamental result of this paper. Note that if the second filter is replaced with a different function, $H_c(z)$, a general noncausal filter, $H(z^{-1})H_c(z)$, results with arbitrary pole and zero locations.

Computer simulations presented in the following sections verify that the time reversed section filter of Fig. 5(a) results in a faithful implementation of the linear phase transfer function (7) using IIR filters for $H(z)$. To guarantee zero input limit cycle immunity, the final filter stage $H(z)$ can be implemented using the IIR truncation scheme of Fig. 1 at the expense of an additional filter operation. This limit cycle free, linear phase IIR filter is shown in Fig. 5(b). Note that a linear phase IIR filter also results from a parallel combination of $H(z^{-1})$ and $H(z)$ but with less desirable overall magnitude characteristics [3], [9].

The IIR subfilter $H(z)$ in Fig. 5 can be any type of filter function (Butterworth, Chebyshev, elliptic, etc.) and can be of arbitrary order implemented with any desired structure (direct form, lattice, GIC, etc.) [18]. If the overall linear phase filter, $H_{eq}(z)$, is to satisfy filter specifications for ω_p, ω_s, δ_p, and δ_s; a standard magnitude-only filter design program is used to design a subfilter $H(z)$ satisfying ω_p, ω_s, $\delta_p/2$, and $\delta_s/2$ (δ_p and δ_s specified in decibels) since the overall magnitude response is equal to the square of the magnitude of $H(z)$. Thus, the stopband attenuation of $H(z)$ is improved by a factor of 2 and the passband ripple of $H(z)$ is degraded by a factor of 2.

Since the equivalent transfer function of the time reversed section filter is $|H_{eq}(z)| = |H(z)|^2$, the order of the subfilter $H(z)$ will always be less than the order of a direct nonlinear phase causal IIR filter, $H_c(z)$, whose perfor-

mance is equal to $H_{eq}(z)$. This means that the number of multiplies and the noise gain of $H_c(z)$ will always be greater than the subfilter $H(z)$. The time reversed section implementation, $H_{eq}(z)$, of Fig. 5(a) has three times the number of multiplies and twice the noise gain of the subfilter $H(z)$. In other words, the number of multiplies and noise gain of the time reversed section implementation is directly related to the subfilter, $H(z)$, but is only related in relative terms to an equivalent causal implementation $H_c(z)$.

III. Finite Section Length Effects

The principal assumption for the operation of Figs. 1, 3, and 5, is that an output created from a continuous convolution with an infinite impulse response is equivalent to an output created from sectioned convolutions with a truncated impulse response for sufficiently long section lengths L. The effects of this assumption are most easily understood from the IIR truncator of Fig. 1. The analysis for this case can be extended to the structures of Figs. 3 and 5. The exact derivation for the output of Fig. 5 is given below:

$$x_k(n),\ y_k(n) \equiv \begin{cases} x(n),\ y(n) & kL \le n \le (k+1)L - 1 \\ 0 & \text{otherwise} \end{cases}$$

$$u_k(n) \equiv \begin{cases} 1 & kL \le n \le (k+2)L - 1 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

$$y(n) = \sum_{k=0}^{\infty} \{[h(n) * x_k(n)] \cdot u_k(n)\} \tag{9}$$

$$U_k(e^{j\omega}) = \frac{\sin(\omega k)}{\sin(\omega/2)}\, e^{-j(L-1)\omega} \tag{10}$$

$$Y_k(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_k(e^{j\theta}) H(e^{j\theta}) U_k(e^{j(\omega-\theta)})\, d\theta. \tag{11}$$

If the input is less than $2L$ samples long, the system of Fig. 1 is linear and a relationship can be established between the section length L and the error associated with the truncation [1]. One effect of the truncation under this condition is the addition of a ripple in the passband response. To help counteract this effect, [5] suggests placing a ripple correction filter in parallel with the actual filter. In the two-pass filtering case for implementing $H(z)H(z^{-1})$, [3] develops a scheme to calculate a nonzero initial state for the second filtering pass, based on the final state of the first pass, which helps counteract the effects of the truncation. Although these techniques significantly reduce the error associated with truncation when processing finite length input sequences and offer some reduction in error for infinite length sequences, they both require additional multipliers and an equivalent reduction in error can be achieved by using a slightly longer section length.

For the most useful case where the input is much longer than L, (11) shows that the system of Fig. 1 is slightly time variant since $Y_k(e^{j\omega})/X_k(e^{j\omega})$ is not constant for all sections. Within each L sample section, an analogy can be drawn to the window method used to design FIR filters based on truncated IIR impulse responses. The more samples per section, the smaller the width of the main lobe of $U_k(e^{j\omega})$ and the better each section's transfer function approximates $H(e^{j\omega})$. An alternate viewpoint is to consider that for L *sufficiently large*, the response of each section decays to a small enough value such that $h(n) \cong 0$ before $U_k(e^{j\omega}) = 0$ so that $[h(n) * x_k(n)] \cdot u_k(n) \cong h(n) * x_k(n)$ and each section's impulse response $Y_k(e^{j\omega})/X_k(e^{j\omega})$ approximates $H(e^{j\omega})$.

Kormylo and Jain [2] suggest expanding the overall ideal filter transfer function into its partial fraction expansion consisting of a parallel combination of first-order filters and selecting the section length L to be equal to four times the "equivalent time constant" of the first-order section with the longest transient response. Czarnach [3] develops a bound on the time domain error as a function of the overall ideal filter impulse response but does not indicate an acceptable range for this error which could be used to select an appropriate section length L. Since the goal is to choose a sufficiently long section length, L, such that the magnitude and phase response of the overall filter is not noticeably different from the ideal response $H(z)H(z^{-1})$ we develop an analytic bound on the frequency domain error below and present the effects of finite section length on the overall transfer function via simulation in Section V.

To determine an acceptable value for the convolution section length, L, consider the first epoch ($k = 0$) of the IIR truncator shown in Fig. 1. As before, the analysis for this case can be extended to the time reversed section structures of Figs. 3 and 5. Assuming L to be infinite, the ideal output of Fig. 1 is

$$Y(e^{j\omega}) = \sum_{n=0}^{\infty} y(n) e^{-j\omega n} \tag{12}$$

for finite values of L

$$Y_L(e^{j\omega}) = \sum_{n=0}^{L-1} y(n) e^{-j\omega n}$$

$$= Y(e^{j\omega}) - \sum_{n=L}^{\infty} y(n) e^{-j\omega n} \tag{13}$$

but

$$y(n) = \sum_{m=L}^{n} h(m) x(m-n) \qquad \text{for } m \ge L \tag{14}$$

and

$$y(n) \le \sum_{m=L}^{\infty} |h(m)| \equiv \epsilon(L) \qquad \text{assuming } |x(n)| \le 1\ \forall n. \tag{15}$$

Thus, the difference between the nontruncated and truncated output spectrums at any frequency ω is always less than $\epsilon(L)$. $\epsilon(L)$ can be made arbitrarily small but must be less than or equal to the maximum allowable stopband or passband ripple specifications so as not to affect the overall magnitude response. If L' is defined as the minimum section length where $\epsilon(L)$ is less than the filter's ripple specifications, L' is determined by the following relationship:

$$\epsilon(L') = \sum_{n=L'}^{\infty} h(n) = \min(\delta_p, \delta_s) \tag{16}$$

where $h(n)$ is the impulse response of the IIR filter satisfying δ_p, and δ_s. It is known from IIR scaling theory that (15) is a pessimistic bound on $y(n)$ [9, p. 318]. Thus (16) will generally yield values for L' where $\epsilon(L')$ is actually less than δ_p, and δ_s and represents a worst case approach. Note that $h(n)$ in (16) represents the impulse response of the subfilter $H(z)$ of Fig. 5(a). In previous approaches [2]–[7], $h(n)$ in (16) would represent the impulse response of the overall filter $H(z)H(z^{-1})$ which is significantly longer than the impulse response of $H(z)$ and results in much larger values for L'. Fig. 6 plots $\epsilon(L)$ versus L for the new time reversed section approach and the previous approach. The larger section lengths required of the previous approach result in larger storage requirements.

IV. Simulation Methodology

To determine how well the schemes of Figs. 1, 3, and 5 perform functions equivalent to $H(z)$, $H(z^{-1})$, and $|H(z)|^2 e^{j\omega(L-1)/2}$, respectively, Fortran models of these systems have been created. As mentioned above, these systems perform differently for input sequences with less than L samples than for inputs with more than L samples. Thus, the equivalent transfer function of these systems cannot be determined by applying an impulse (sequence of length $1 \ll L$) to the input and performing a Fourier transform of the output. The effects of finite section lengths to the overall transfer function are only seen for inputs with length much greater than L. We used the impulse response from a high order IIR all-pass function as an input sequence. This sequence, $x_{ap}(n)$, is 4096 samples long and contains more than 2000 samples with significantly large values. Since L is typically less than 300, $x_{ap}(n)$ permits simulations with at least 6–7 convolution sections (epochs). The frequency spectrum of $x_{ap}(n)$ is flat to within 0.00012 dB over the entire frequency range. Thus, when $x_{ap}(n)$ is applied as the input to a simulation model of Figs. 1, 3, or 5, the magnitude of the equivalent transfer function, including finite section length effects, can be determined directly by performing an FFT of the output response.

One frequency domain measure of the nonlinearity of a system is the total harmonic distortion (THD) of the output in response to a single frequency sinusoidal input. In practice, finite length input and output sequences must be used to measure THD when using a simulation model. The actual input is $x(n) = W_r(n) \sin(\omega_0 n)$ (W_r = rectangular window) with a frequency spectrum of

$$X(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\sin(\theta N/2)}{\sin(\theta/2)} \delta_{w_0}(\omega - \theta)\, d\theta$$

$$= \frac{\sin[(\omega - \omega_0)N/2)]}{\sin[(\omega - \omega_0)/2]}. \tag{17}$$

If $\omega_0 = 2\pi l/N$, where l is an integer, (17) is zero at intervals of $\omega = 2\pi(k + l)/N$ for all $k \neq N - l$. Thus, an N point FFT of $x_{\text{in}} = X(e^{j2\pi k/N})$ will also be zero for all $k \neq N - l$ which is equivalent to the spectrum of an infinitely long sinusoid at the frequency intervals $\omega = 2\pi k/N$. Assuming the response to a single frequency sinusoid contains harmonic distortion of the form $\Sigma A_k \sin(k\omega_0)$ because of system nonlinearities, the same arguments can be used to show that an FFT of a finite length output is the same as the spectrum of an infinitely long output which resulted from applying an infinitely long sinusoidal input. Thus, the THD is determined by applying a sinusoidal input sequence containing an integer number of cycles every N points to the simulation models and performing an N point FFT of the output. The input used is many times longer than the section length L and the output is not considered valid until all start-up transients have died out. The THD of the sinusoidal input used is approximately -130 dB.

V. Time Reversed Section Performance

A linear phase filter having the specifications $F_p = 0.3$, $F_s = 0.325$, $\delta_p = 0.01$ dB, and $\delta_s = 70$ dB will be used as an example where $F_{p(s)}$ is the normalized pass(stop)band frequencies and $\delta_{p(s)}$ is the pass(stop)band ripple. The time-reversed section structure of Fig. 5(a) implements the desired linear phase filtering operation with 7 distinct multipliers (21 total multiplies) and 46 additions per output sample. A direct form linear phase FIR filter satisfying this specification requires 73 distinct multiplies and 145 additions per output sample.

Using the time-reversed section architecture, each IIR subfilter $h(z)$ must satisfy the specifications

$$F_p = 0.3, \quad F_s = 0.325$$

$$\delta_p = -0.005 \text{ dB}, \quad \delta_s = -35 \text{ dB}. \tag{18}$$

The minimum order filter which will satisfy these specifications is an elliptic type. The formulas presented in [19] were used to design a seventh-order elliptic function meeting the desired specifications. Each seventh-order IIR subfilter is implemented as a parallel interconnection of two IIR all-pass sections as described in [20]. Since the subfilter can only have half the passband ripple of the overall filter specifications, the low passband sensitivity of the parallel all-pass structure makes it well suited for

this application. The subfilter transfer function is given by (19):

$$H(z) = \frac{a_0 + z^{-1}}{1 + a_0 z^{-1}} \cdot \frac{b_0 + b_1 z^{-1} + z^{-2}}{1 + b_1 z^{-1} + b_0} + \frac{c_0 + c_1 z^{-1} + z^{-2}}{1 + c_1 z^{-1} + c_0 z^{-2}} \cdot \frac{d_0 + d_1 z^{-1} + z^{-2}}{1 + d_1 z^{-1} + d_0 z^{-2}} \tag{19}$$

$$a_0 = 0.1295576 \qquad c_0 = 0.6733035$$

$$b_0 = 0.2783396 \qquad c_1 = 0.5954800$$

$$b_1 = 0.3937937 \qquad d_0 = 0.9135559$$

$$d_1 = 0.7115493.$$

L' is defined as the number of samples where $\epsilon(L)$, the difference between the truncated transfer function and the ideal transfer function, is less than the allowable passband and stopband ripple specifications. Fig. 6 plots $\epsilon(L)$ as a function of section length L. As shown in Fig. 6, $L' = 200$ samples for the subfilter of (19). Note that $L' > 400$ samples must be used to obtain equal performance from previous approaches. The subfilter magnitude response and group delay variation is shown in Fig. 7. The uncorrected group delay varies by more than 20 dB from the average value demonstrating the highly nonlinear phase characteristic typical of elliptic filters.

The simulation techniques described in the preceding section were used to measure performance as a function of section length L of the time reversed section architecture (Fig. 5(a)) with IIR elliptic subfilter (19). A very long section length is first used to determine the performance of the time reversed section filter in the absence of finite section length effects. Fig. 7 shows the measured response of the architecture of Fig. 5(a) without finite section length effects ($L = 500$). As predicted by (7), the time reversed section technique removed the 20-dB group delay variation of the IIR elliptic subfilter (Fig. 7(b)) resulting in exact linear phase. As also predicted by (7), the stopband attenuation and passband ripple of the composite filter is twice that of the subfilter and satisfies the desired specifications of (18). The THD measured at the output is -130 dB, the resolution of the input sinusoid, indicating a linear and time-invariant system.

Figs. 8(a), (b) show the measured response of the time reverse section filter when $L = L' = 200$ samples. The magnitude response is indistinguishable from the ideal magnitude response and the group delay is nearly constant with a very small amount of variation (<0.0005 dB). The THD measured at the output is -110 dB indicating the amount of nonlinearity present due to finite section length effects is extremely small. Several other filters with differing specifications have been investigated and all have been found to possess ideal magnitude characteristics with constant group delay (± 0.0005 dB) when the section

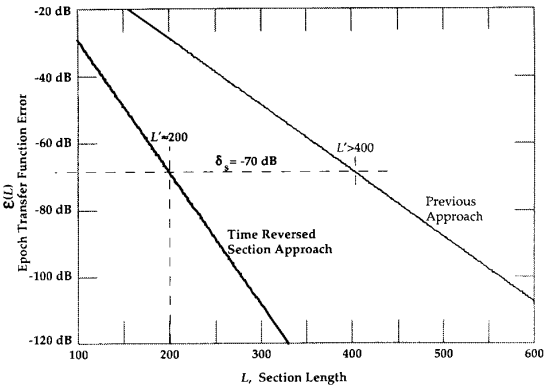


Fig. 6. Maximum transfer function error as a function of section length. Subfilter specifications: $F_s = 0.3$, $F_p = 0.325$, $\delta_p = -0.005$ db; $\delta_s = -35$ dB.

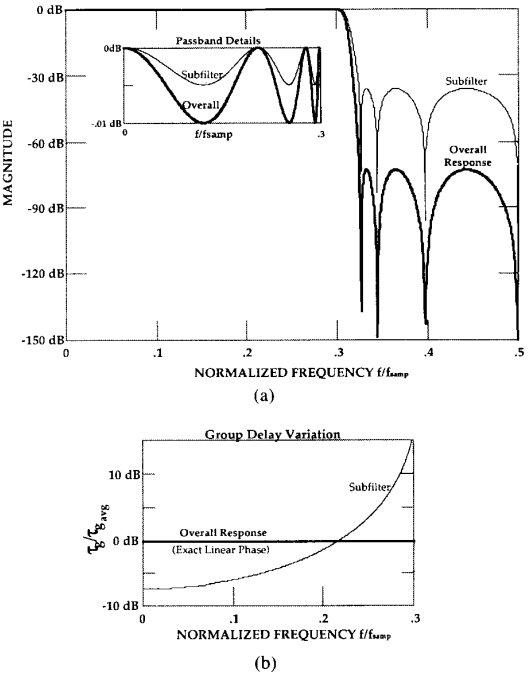


Fig. 7. Response of time reversed section filter (Fig. 4(a)) with section length $L \gg L'$. (a) Magnitude response and passband details. $F_s = 0.3$, $F_p = 0.325$, $\delta_p = -0.01$ dB, $\delta_s = 70$ dB, $L \gg L'$. (b) Group delay variation from average.

length L is set equal to the subfilter's value of L'. Thus, it appears that with $L \geq L'$, the nonideal behavior of (11) due to a finite section length is negligible and the linear phase function of (7) is valid.

As the section length L is decreased below L', the total harmonic distortion appearing at the output and the group delay variation of the measured transfer function increase. Figs. 9 and 10 plot total harmonic distortion and group delay variation measured as a function of section length L for the time reverse section filter using (19). Both of these curves begin to increase rapidly after the section

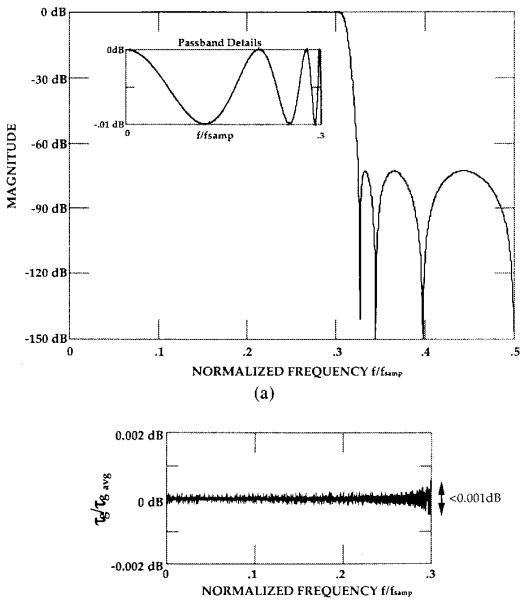


Fig. 8. Response of time reversed section filter (Fig. 4(a)) for $L = L' = 200$. (a) Magnitude response and passband details. $F_s = 0.3$, $F_p = 0.325$, $\delta_p = -0.01$ dB, $\delta_s = -70$ dB, $L = L' = 200$. (b) Group delay variation from average.

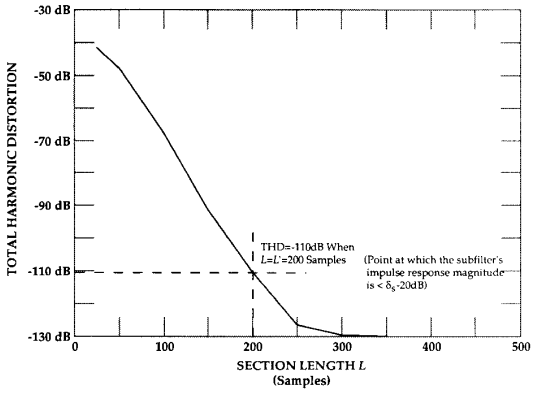


Fig. 9. THD versus section length for time reversed section filter example (Fig. 4(a) implementation). Fundamental samples. frequency = 3/1024. $F_s = 0.3$, $F_p = 0.325$, $\delta_p = -0.01$ dB, $\delta_s = -70$ dB.

length is decreased below the subfilter's value of $L' = 200$. Fig. 11 shows the output spectrum resulting from a single sinusoidal input for section length $L = 150$, less than $L' = 200$. The clear presence of harmonics in this output spectrum result from the nonlinearity of the actual transfer function with short section lengths as predicted by (11). As the section length is further decreased, the harmonics present in Fig. 11 increase, the group delay variation increases, the equivalent transfer function continues to follow the desired response but contains sharp variations (''noise'') of increasingly larger magnitude.

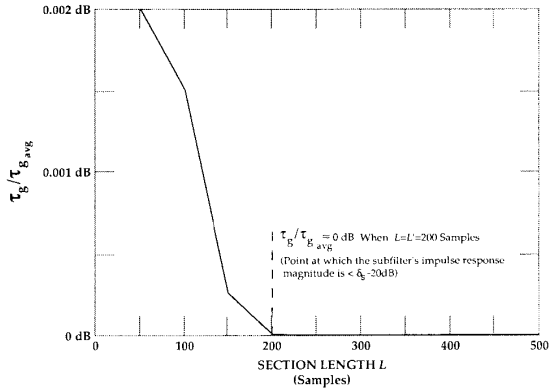


Fig. 10. Group delay variation versus section length for time reversed section filter example (Fig. 4(a) implementation). $F_s = 0.3$, $F_p = 0.325$, $\delta_p = -0.01$ dB, $\delta_s = -70$ dB.

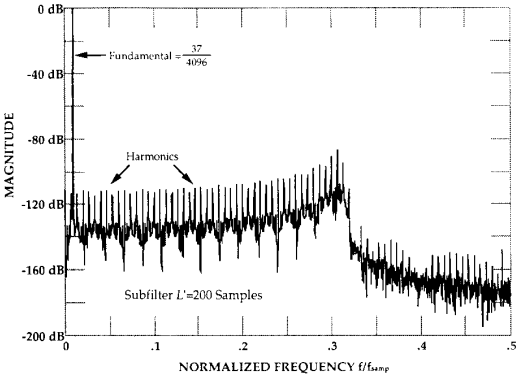


Fig. 11. Spectrum of single frequency response (Fig. 4(a) implementation) demonstrating the onset of harmonic distortion due to finite section length $L = 150$ ($< L'$). $F_s = 0.3$, $F_p = 0.325$, $\delta_p = -0.01$ dB, $\delta_s = -70$ dB, $L = 150$ samples.

VI. Performance Comparison

The number of multiplies per sample and the noise gain of the linear phase IIR filter of Fig. 5(a), $H_{eq}(z)$, are compared to a direct form linear phase FIR implementation, $H_N(z)$, and a nonlinear phase causal IIR implementation $H_c(z)$. Comparisons are made for the nine different filter specifications given in Table I. The FIR filter length N was estimated using the empirical equations in [21] and verified using a standard Parks and McClellan FIR design program [22]. Elliptic filters are used for the IIR subfilter implementation $H(z)$ since they possess the sharpest transition bandwidth for a given order. The time reverse section scheme forces the overall transfer function $H_{eq}(z)$ to have linear phase regardless of the phase of the IIR subfilter $H(z)$ so that the extremely nonlinear phase response of an elliptic filter is not important. The magnitude only causal IIR implementations $H_c(z)$ are also elliptic filters and provide a measure of the minimum possible number of multiplies per sample for comparison.

Elliptic filter coefficients are determined using the methods presented in [19]. The elliptic filters, $H(z)$ and

TABLE I
NINE EXAMPLE FILTER SPECIFICATIONS FOR COMPARING THE TIME REVERSED SECTION FILTER TO OTHER APPROACHES

Example	F_p	F_s	δ_p (dB)	δ_s (dB)
1	0.05	0.0650	−0.100	−62
2	0.10	0.1250	−0.005	−82
3	0.15	0.2000	−0.005	−82
4	0.20	0.2025	−0.100	−62
5	0.25	0.2550	−0.005	−80
6	0.30	0.3250	−0.010	−70
7	0.35	0.3900	−0.100	−72
8	0.40	0.4005	−0.005	−62
9	0.45	0.4700	−0.006	−70

TABLE II
COMPLEXITY AND NOISE PERFORMANCE OF THE TIME REVERSED SECTION FILTER COMPARED TO DIRECT FORM FIR AND NONLINEAR PHASE IIR IMPLEMENTATIONS

	IIR $H_c(z)$		Time Rev. Section $H_{eq}(z)$		Direct Form FIR $H_N(z)$	
Example	# Mult	Noise Gain (dB)	# Mult	Noise Gain (dB)	# Mult	Noise Gain (dB)
1	9	24.9	15	25.7	78	22.0
2	11	21.1	21	23.8	79	22.0
3	11	18.5	21	21.5	47	19.7
4	13	27.3	28	29.9	467	29.7
5	17	25.9	33	28.8	386	28.9
6	11	19.5	21	22.5	73	21.7
7	9	18.4	15	20.6	33	18.2
8	19	34.2	39	36.7	3278	38.2
9	9	21.9	15	27.7	87	22.4

$H_c(z)$, are implemented as parallel all-pass structures. As described by [20], specific all-pass networks are chosen for each filter example so as to minimize overall quantization noise variance. The noise gain for each filter, $H_{eq}(z)$ and $H_c(z)$, is determined by the methods of [20] and [23]. The noise gain of the FIR filter $H_N(z)$ is determined assuming a direct form implementation with rounding after each multiplication [9].

As discussed in Section II, although the number of multiplies and the noise gain of $H_{eq}(z)$ can be determined from the subfilter $H(z)$, only a relative relationship exists between $H_{eq}(z)$ and $H_c(z)$. Table II presents the results of the comparisons. In all cases, the number of multiplies per sample for the time reversed section filter $H_{eq}(z)$ is within a factor of about 2 of the minimum achievable with a nonlinear phase elliptic filter $H_c(z)$. The noise gain of all three implementations are within 1 b (1/2 b typical) of each other with $H_{eq}(z)$ typically the highest. As can be seen in Table II, the advantage of using the time reversed section architecture instead of a linear phase FIR filter increases for sharp transition band filters. In these cases (examples 4, 5, 8), the time reversed section architecture requires 10–100 times fewer multiplies than an FIR implementation, yet this approach provides linear phase, whereas the standard causal elliptic IIR does not.

VII. FUTURE WORK

As shown in (9), each output section is multiplied by a rectangular window of length $2L$. Ideally, $U_k(e^{j\omega})$ would be an impulse at $\omega = 0$. Thus, windows having narrow-width main lobes and low-height sidelobes are the most desirable. For a given length, the spectrum of a rectangular window has the narrowest main lobe but the greatest first sidelobe magnitude when compared to other commonly used window functions. Nonrectangular windows may permit the time reversed section architecture to use shorter section lengths than L'. However, windows with coefficients other than 2^n require additional multipliers [24]. The tradeoffs associated with introducing nonrectangular windows into the time reversed section scheme are currently under investigation.

Using the overlap-save method with the time reversed section technique is also under investigation. As mentioned earlier, this architecture requires one less adder than the overlap-add implementation of Fig. 5. One possible overlap-save implementation is shown in Fig. 12. Initial investigations of this architecture have shown promising results.

VIII. CONCLUSIONS

A time reversed section scheme has been presented for implementing continuous input noncausal filters with reduced storage and computation requirements over previous approaches. A primary application of this approach is the implementation of exactly linear phase filters using computationally efficient recursive structures designed with magnitude-only criteria. The time reversed section scheme reduces the number of multiplies and additions per output sample for linear phase filters to within a factor of about two of a magnitude-only causal IIR implementation. The scaling properties of a linear phase time reversed section implementation are equivalent to a magnitude-only IIR implementation and the noise performance is approximately equivalent. This new linear phase IIR architecture can be made immune to zero-input limit cycle and overflow oscillations even if the recursive subfilters exhibit both. A relationship was derived for determining the section length, L', where the truncated and ideal transfer functions are equal. It was demonstrated that for section lengths greater than or equal to L', the time reversed section scheme results in a linear, time-invariant system with linear phase and magnitude response equal to the square of the magnitude of the IIR subfilter. Simulation techniques have been developed to determine a time reverse section filter's equivalent transfer function and harmonic distortion as a function of section length. For applications where linear phase is important and a sharp transition band is desired, the time reversed section technique can reduce the number of multiplies and additions per output sample by hundreds or even thousands over direct form FIR implementations.

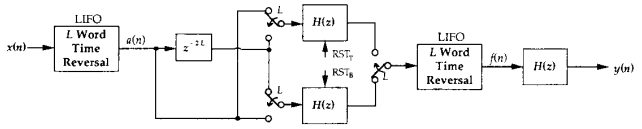


Fig. 12. Possible implementation of overlap-save based time reversed section linear phase filter.

ACKNOWLEDGMENT

The authors wish to thank G. Yacoub, W. McKnight, F. Harris, and H. Whitehouse for their helpful critiques and suggestions.

REFERENCES

[1] A. J. Gibbs, "On the frequency domain responses of causal digital filters," Ph.D. dissertation, Univ. of Wisconsin, Madison, WI, 1969.

[2] J. J. Kormylo and V. K. Jain, "Two-pass recursive digital filter with zero phase shift," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 384–387, Oct. 1974.

[3] R. Czarnach, "Recursive processing by noncausal digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, no. 3, pp. 363–370, June 1982.

[4] K.-P. Estola, "Efficient linear phase FIR filter structures exploiting recursive subfilters," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 1988, pp. 305–308.

[5] P. Jarvilehto and K.-P. Estola, "A new modular VLSI filter architecture using computationally efficient recursive digital filter topology," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 1988, pp. 1301–1304.

[6] R. Czarnach, H. W. Schussler, and G. Rohrlein, "Linear phase recursive digital filters for special applications," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, May 1982, pp. 1825–1828.

[7] G. Cortelazzo and M. R. Lightner, "The use of multiple criterion optimization for frequency domain design of noncausal IIR filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 1, pp. 126–135, Feb. 1985.

[8] Z. Jing, "A new method for digital all-pass filter design," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 11, pp. 1557–1564, Nov. 1987.

[9] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[10] J. P. Thiran, "Recursive digital filters with maximally flat group delay," *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 659–664, Nov. 1971.

[11] S. N. Hazra, "Linear phase IIR filter with equiripple stopband," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, no. 3, June 1983.

[12] G. Cortelazzo and M. R. Lightner, "Simultaneous design in both magnitude and group delay of IIR and FIR filters based on multiple criterion optimization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 5, pp. 949–967, Dec. 1984.

[13] H. Baher, "Digital filters with finite-band approximation to constant amplitude and delay, and arbitrary selectivity," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 1986, pp. 657–659.

[14] C. W. Kim and R. Ansari, "Approximately linear phase IIR filters using allpass sections," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 1986, pp. 661–664.

[15] M. Renfors and T. Saramaki, "A class of approximately linear phase digital filters composed of all-pass subfilters," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 1986, pp. 678–681.

[16] M. A. Clements, "On causal linear phase IIR digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 4, Apr. 1989.

[17] L. B. Jackson, "On the interaction of roundoff noise and dynamic range in digital filters," *Bell Syst. Tech. J.*, vol. 49, pp. 159–184, 1970.

[18] H. Samueli, T.-J. Lin, and W. S. Lao, "A comparison of recursive digital filter structures suitable for high-speed custom VLSI implementation," in *Proc. IEEE 21st Asilomar Conf. Signals, Syst., Comput.*, Nov. 1987, pp. 23–27.

[19] L. Gazsi, "Explicit formulas for lattice wave digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-32, no. 1, pp. 68–88, Jan. 1985.

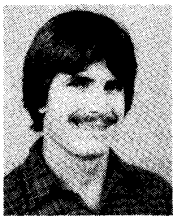
[20] P. P. Vaidyanathan, S. K. Mitra, and Y. Neuvo, "A new approach to the realization of low sensitivity IIR digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, Apr. 1986.

[21] L. R. Rabiner, J. K. Kaiser, O. Herrmann, and M. T. Dolan, "Some comparisons between FIR and IIR digital filters," *Bell Syst. Tech. J.*, vol. 53, pp. 305–331, Feb. 1974.

[22] T. W. Parks and J. H. McClellan, "A program for the design of linear phase finite impulse response digital filters," *IEEE Trans. Circuit Theory*, vol. CT-19, pp. 189–194, Mar. 1972.

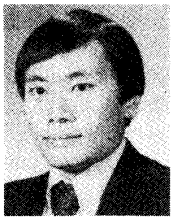
[23] S. K. Mitra and K. Hirano, "Digital all-pass networks," *IEEE Trans. Circuits Syst.*, vol. CAS-21, no. 5, Sept. 1974.

[24] K. M. M. Prabhu and H. Renganthan, "Optimum binary windows for discrete Fourier transforms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 1, pp. 216–220, Feb. 1986.



Scott R. Powell received the B.S. and M.S. degrees in electrical engineering from Oregon State University, Corvallis, OR, in 1982 and 1983, respectively. He is currently working toward the Ph.D. degree in electrical engineering as a Hughes Aircraft fellow at the University of California, San Diego.

With over eight years experience in VLSI circuit design and signal processing, he is currently Head of the Analog and Digital Design Section at Hughes Aircraft Company's Technology Center in Carlsbad, CA, responsible for research and development of high performance signal processing algorithms, architectures, and integrated circuits. His research interests include reduced complexity signal processing algorithms, sigma-delta data conversion, and VLSI circuit design.



Paul M. Chau (S'75–M'82) received the B.S. degree in electrical engineering from the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, in 1977, the M.S. degree from Syracuse University, Syracuse, NY, in 1981, and the Ph.D. degree from Cornell University, Ithaca, NY, in 1987.

From 1977 to 1982, he was with the General Electric Company, Syracuse, NY, and worked on radar and sonar systems, computer-aided design, test engineering, and advanced development engineering. He also worked at GE Corporate Research and Development Center on a VLSI floating-point processor project. His research interests include VLSI architectures and design for digital signal processing, computer architectures, and floating-point processor systems. He has been a member of the faculty in the Electrical and Computer Engineering Department at the University of California, San Diego, since 1987.

Dr. Chau is a member of Tau Beta Pi and Eta Kappa Nu.