It is easy to verify that

$$\langle N_i(r_i)\rangle_m \cdot M_i(x)$$
$$= N^{-1}r_i^{-(N-1)}(x^{N-1} + r_i x^{N-2} + \cdots + r_i^{N-2}x + r_i^{N-1}). \tag{20}$$

Compare (9) with (20), it follows that $\langle N_i(r_i)\rangle_m \cdot M_i(x) = Q_i(x)$. It is easy to check $\langle Q_i(x)\rangle_{(x-r_i)} = 1$ and $\langle Q_i(x)\rangle_{(x-r_j)} = 0$, for $i \neq j$. By CRTP, $C(x)$ can be obtained as

$$C(x) = \sum_{i=0}^{N-1} c_i^* \langle N_i(r_i)\rangle_m \cdot M_i(x)$$
$$= \sum_{i=0}^{N-1} c_i^* Q_i(x) \tag{21}$$

III. CONCLUSION

From the derivations given in Section II, the polynomial residue number system can be interpreted by the terminology of Chinese remainder theorem for polynomials over a finite ring, which is more familiar for the computer and signal processing societies.

REFERENCES

[1] A. Skavantzos and F. J. Taylor, "On the polynomial residue number system," *IEEE Trans. Signal Processing*, vol 39, no. 2, pp. 376–382, Feb. 1991.
[2] J. H. McClellan and C. M. Rader, *Number Theory in Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

Wavelet Coefficient Computation with Optimal Prefiltering

Xiang-Gen Xia, C.-C. Jay Kuo, and Zhen Zhang

Abstract—Discrete wavelet transform (DWT) is often used to approximate wavelet series transform (WST) and continuous wavelet transform (CWT), since it can be computed numerically. In this research, we first study the accuracy of the computed DWT coefficients obtained from the Shensa algorithm as an approximate of the WST coefficients. Based on the accuracy analysis, we then propose a procedure to design optimal FIR prefilters used in the Shensa algorithm to reduce the approximation error. Finally, numerical examples are presented to demonstrate the performance of the optimal FIR prefilters.

X.-G. Xia is with the Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433-7765 USA.
C.-C. J. Kuo is with the Signal and Image Processing Institute, Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 9089-2564 USA.
Z. Zhang is with the Communication Science Institute, Department of Electrical Engineering-Systems, University of Southern California. Los Angeles, CA 9089-2565 USA.
IEEE Log Number 9401921.

I. INTRODUCTION

Wavelet transforms have recently been recognized as useful tools for various applications such as signal and image processing, numerical analysis and physics. There are three types of wavelet transforms discussed in the literature, namely, continuous wavelet transform (CWT) [4], wavelet series transform (WST) [3], and discrete wavelet transform (DWT) [6], [8]. These transforms using biorthogonal wavelet bases are briefly summarized below. We use the notation

$$f_{jk}(t) \triangleq 2^{j/2} f(2^j t - k), \quad j, k \in \mathbf{Z},$$

and

$$f_{a,b}(t) = |a|^{-1/2} f\left(\frac{t-b}{a}\right), \quad a \neq 0, b \in \mathbf{R}.$$

Let $\psi(t)$ and $\tilde{\psi}(t)$ be, respectively, a real wavelet function and its dual such that $\{\psi_{jk}(t)\}_{j,k}$ and $\{\tilde{\psi}_{jk}(t)\}_{j,k}$ form a biorthogonal wavelet basis in $L^2(\mathbf{R})$. Then, for $f(t) \in L^2(\mathbf{R})$, its CWT with respect to the wavelet $\psi(t)$ is defined as

$$\mathrm{CWT}\{f(t); a, b\} \triangleq \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt$$

where a and b are called the scale and time parameters, respectively. The WST of $f(t)$ is obtained by sampling its CWT in the scale-time plane (a, b) with the so-called "dyadic" grid, i.e.

$$\mathrm{WST}\{f(t); j, k\}$$
$$= \mathrm{CWT}\{f(t); a = 2^{-j}, b = k2^{-j}\}, j, k \in \mathbf{Z}.$$

Thus, the WST coefficients, also denoted by $b_{j,k}$, can be determined by

$$b_{j,k} \triangleq \mathrm{WST}\{f(t); j, k\} = \int_{-\infty}^{\infty} f(t)\psi_{jk}(t)dt, \quad j, k \in \mathbf{Z}. \tag{1}$$

Moreover, $f(t)$ can be reconstructed via

$$f(t) = \sum_j \sum_k b_{j,k} \tilde{\psi}_{jk}(t).$$

The orthogonal wavelet is a special case of the biorthogonal one by requiring $\psi(t) = \tilde{\psi}(t)$. If the t as well as parameters (a, b) all take discrete values, which are recognized as a natural wavelet transform for the discrete-time signal $f(m\triangle t)$ with $m \in \mathbf{Z}$, the resulting transform is called the DWT of $f(t)$. It is clear that only the DWT coefficients can be computed numerically, and the CWT and WST coefficients have to be approximated by the DWT coefficients in practice.

Several numerical algorithms have been proposed to compute the DWT coefficients such as the Mallat algorithm [6], the "à trous" algorithm of Holschneider *et al.* [5], and the Shensa algorithm [8] as a unified approach for the former two. Efficient implementations and detailed computational complexity analysis for these algorithms were discussed by Rioul and Duhamel [7]. However, an important issue which has not yet been addressed is the numerical accuracy of the computed DWT coefficients $b'_{j,k}$ with respect to the true WST coefficients $b_{j,k}$ as defined in (1). This was considered as an open problem in the work by Rioul and Duhamel [7] and Shensa [8]. In this research, after a brief review of some results from wavelet theory in Section II, we derive formulas to characterize the error between the computed and true wavelet coeficients in Section III. With such an error analysis, we develop a procedure to design the optimal FIR prefilter $q[n]$ to reduce the error as much as possible in Section

IV. Numerical examples are given in Section V to demonstrate the performance of the optimal prefilters for both orthogonal and biorthogonal wavelets.

II. REVIEW OF BIORTHOGONAL WAVELETS AND DISCRETE WAVELET TRANSFORM

We review basic properties of biorthogonal wavelets, and refer to [1] for more details. Consider a real mother wavelet function $\psi(t)$, the associated scaling function $\phi(t)$, and their dual functions $\tilde{\psi}(t)$ and $\tilde{\phi}(t)$ such that $\{\psi_{jk}(t)\}_{j,k \in \mathbf{Z}}$ and $\{\tilde{\psi}_{jk}(t)\}_{j,k \in \mathbf{Z}}$ form a biorthogonal wavelet basis in $L^2(\mathbf{R})$. Let the Fourier transform of $f(t) \in L^2(\mathbf{R})$ be denoted by

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$$

and

$$H(\omega) = \sum_n h_n e^{-in\omega} \quad \text{and} \quad G(\omega) = \sum_n g_n e^{-in\omega}$$

be the associated filters of $\phi(t)$ and $\psi(t)$, respectively. Similarly, we associate the following dual filters

$$\tilde{H}(\omega) = \sum_n \tilde{h}_n e^{-in\omega} \quad \text{and} \quad \tilde{G}(\omega) = \sum_n \tilde{g}_n e^{-in\omega}$$

with the dual wavelet and scaling functions $\tilde{\psi}(t)$ and $\tilde{\phi}(t)$, respectively. One can derive the following well known properties for biorthogonal wavelets (see Chapter 5 in [1]):

$$\hat{\phi}(\omega) = \prod_{k=1}^{\infty} H(2^{-k}\omega), \quad \text{and} \quad \hat{\tilde{\phi}}(\omega) = \prod_{k=1}^{\infty} \tilde{H}(2^{-k}\omega),$$

$$\begin{cases} H(\omega)\overline{\tilde{H}(\omega)} + G(\omega)\overline{\tilde{G}(\omega)} &= 1, \\ H(\omega)\tilde{H}(\omega + \pi) + G(\omega)\tilde{G}(\omega + \pi) &= 0, \end{cases} \quad (2)$$

$$\sum_k \hat{\phi}(\omega + 2k\pi)\hat{\tilde{\phi}}(-\omega - 2k\pi) = 1.$$

Note that one possible solution for the second equation of the system (2) is

$$G(\omega) = e^{-i\omega}\overline{\tilde{H}(\omega + \pi)} \quad \text{and} \quad \tilde{G}(\omega) = e^{-i\omega}\overline{H(\omega + \pi)}, \quad (3)$$

which is often imposed to simplify the filter design procedure.

Let $(\tilde{V}_j)_{j \in \mathbf{Z}}$ denote the multiresolution wavelet subspaces in $L^2(\mathbf{R})$ corresponding to the dual scaling function $\tilde{\phi}(t)$, and $f_J(t)$ be the projection of $f \in L^2(\mathbf{R})$ in $\tilde{V}_J$ for an arbitrarily fixed integer J. In mathematical terms, we can write

$$f_J(t) = \sum_k c_{J,k} \tilde{\phi}_{Jk}(t) = \sum_{j<J}\sum_k b_{j,k}\tilde{\psi}_{jk}(t)$$

where

$$c_{J,k} = \int_{-\infty}^{\infty} f(t)\phi_{Jk}(t)dt \quad (4)$$

and $b_{j,k}$ is defined in (1). The coefficients $b_{j,k}$ with $j < J$ can be obtained from $c_{J,k}$ via

$$c_{j-1,k} = \sqrt{2}\sum_n h_{n-2k}c_{j,n} \quad (5)$$

and

$$b_{j-1,k} = \sqrt{2}\sum_n g_{n-2k}c_{j,n} \quad (6)$$

for $j = J, J-1, \ldots, J_c + 1$. Besides, one can reconstruct $c_{J,k}$ from $c_{J_c,k}$ and $b_{j,k}$, $J_c \le j \le J$, via

$$c_{j,n} = \sqrt{2}\left(\sum_k \tilde{h}_{n-2k}c_{j-1,k} + \sum_k \tilde{g}_{n-2k}b_{j-1,k}\right).$$

To compute wavelet coefficients with the Shensa algorithm, we first perform a prefilering on the sampled signal $x[n] = f(n/2^J)$ to obtain a new sequence $x'[n]$, i.e.

$$x'[n] = \sum_m x[m]q[n-m] \quad (7)$$

and then apply the recursion (5) and (6) to $x'[n]$. The DWT coefficients of $x'[n]$, denoted by

$$\text{DWT}\{x'[n]; j, k\} \triangleq b_{j,k}^{(S)}, \quad j < J, k \in \mathbf{Z}$$

are called the wavelet coefficients obtained from the Shensa algorithm. The well known Mallat algorithm is in fact a special case of the Shensa algorithm by choosing $q[n] = \delta[n]$.

III. ERROR ESTIMATION OF COMPUTED WAVELET COEFFICIENTS

We see from the previous discussion that the difference between the WST coefficients $\{b_{j,k}\}$ and the DWT coefficients $\{2^{-J/2}b_{j,k}^{(S)}\}$ results from the difference between the input sequences $c_{J,k}$ in (4) and $2^{-J/2}x'[k]$ given by (7). This can be written as

$$b_{j,k} - 2^{-J/2}b_{j,k}^{(S)} = \text{DWT}\{c_{J,n} - 2^{-J/2}x'[n]; j, k\}.$$

This error difference can be analyzed with two steps. That is, we first find an expression for the difference between $c_{J,k}$ and $2^{-J/2}x'[k]$, and then analyze the recursion (5) and (6) of the DWT.

For the first step, we have

$$\begin{aligned} c_{J,k} &= 2^{J/2}\int_{-\infty}^{\infty} f(t)\phi(2^J t - k)dt \\ &= 2^{J/2}\int_{-\infty}^{\infty} f(t)\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{\phi}(\omega)e^{i\omega(2^J t - k)}d\omega dt \\ &= \frac{2^{J/2}}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(t)e^{i\omega 2^J t}dt\right)\hat{\phi}(\omega)e^{-i\omega k}d\omega \\ &= \frac{2^{J/2}}{2\pi}\int_{-\infty}^{\infty}\hat{f}(-2^J\omega)\hat{\phi}(\omega)e^{-i\omega k}d\omega. \end{aligned}$$

In addition, we know

$$\begin{aligned} 2^{-J/2}x'[k] &= 2^{-J/2}\sum_m x[m]q[k-m] \\ &= 2^{-J/2}\sum_m \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(\omega)e^{i\omega m/2^J}d\omega q[k-m] \\ &= \frac{2^{J/2}}{2\pi}\int_{-\infty}^{\infty}\hat{f}(2^J\omega)\sum_m e^{i\omega m}d\omega q[k-m] \\ &= \frac{2^{J/2}}{2\pi}\int_{-\infty}^{\infty}\hat{f}(2^J\omega)\sum_m e^{i\omega(m-k)}q[k-m]e^{i\omega k}d\omega \\ &= \frac{2^{J/2}}{2\pi}\int_{-\infty}^{\infty}\hat{f}(-2^J\omega)Q(\omega)e^{-i\omega k}d\omega \end{aligned}$$

where $Q(\omega) = \sum_n q[n]e^{in\omega}$. Therefore, it is concluded that for $k \in \mathbf{Z}$,

$$\begin{aligned} & c_{J,k} - 2^{-J/2}x'[k] \\ & = \frac{2^{J/2}}{2\pi}\int_{-\infty}^{\infty}\hat{f}(-2^J\omega)(\hat{\phi}(\omega) - Q(\omega))e^{-ik\omega}d\omega. \end{aligned}$$

For the second step, let $s[n]$ be an arbitrary discrete sequence used as the input to the filters $H(\omega)$ and $G(\omega)$ so that it is decomposed into an approximation sequence $a[n]$ and a detail sequence $d[n]$:

$$a[k] = \sqrt{2} \sum_n h_{n-2k} s[n] \qquad (8)$$

and

$$d[k] = \sqrt{2} \sum_n g_{n-2k} s[n]. \qquad (9)$$

Let $S(\omega) = \sum_n s[n] e^{-in\omega}$. Take Fourier transforms of (8) and (9),

$$A(2\omega) = \frac{1}{\sqrt{2}} \left[\overline{H(\omega)} S(\omega) + \overline{H(\omega + \pi)} S(\omega + \pi) \right]$$

and

$$D(2\omega) = \frac{1}{\sqrt{2}} \left[\overline{G(\omega)} S(\omega) + \overline{G(\omega + \pi)} S(\omega + \pi) \right].$$

Therefore,

$$\sum_n |a[n]|^2 + \sum_n |d[n]|^2$$

$$= \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} |A(\omega)|^2 d\omega + \int_{-\pi}^{\pi} |D(\omega)|^2 d\omega \right)$$

$$= \frac{1}{4\pi} \left[\int_{-\pi}^{\pi} \left(\left(\left| H\left(\frac{\omega}{2}\right) \right|^2 + \left| G\left(\frac{\omega}{2}\right) \right|^2 \right) \left| S\left(\frac{\omega}{2}\right) \right|^2 \right. \right.$$

$$+ \left(\left| H\left(\frac{\omega}{2} + \pi\right) \right|^2 + \left| G\left(\frac{\omega}{2} + \pi\right) \right|^2 \right) \left| S\left(\frac{\omega}{2} + \pi\right) \right|^2 \right) d\omega$$

$$+ 2 \int_{-\pi}^{\pi} \mathrm{Re}\left(\overline{H\left(\frac{\omega}{2}\right)} H\left(\frac{\omega}{2} + \pi\right) \right.$$

$$\left. \left. + \overline{G\left(\frac{\omega}{2}\right)} G\left(\frac{\omega}{2} + \pi\right) \right) S\left(\frac{\omega}{2}\right) \overline{S\left(\frac{\omega}{2} + \pi\right)} \right) d\omega \right]$$

$$\leq \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} \left(|H(\omega)|^2 + |G(\omega)|^2 \right) |S(\omega)|^2 d\omega$$

$$+ \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} \left(|H(\omega + \pi)|^2 + |G(\omega + \pi)|^2 \right) |S(\omega + \pi)|^2 d\omega$$

$$+ \frac{1}{4\pi} \int_{-\pi}^{\pi} \left| \mathrm{Re}\left(\overline{H\left(\frac{\omega}{2}\right)} H\left(\frac{\omega}{2} + \pi\right) + \overline{G\left(\frac{\omega}{2}\right)} G\left(\frac{\omega}{2} + \pi\right) \right) \right|$$

$$\left(\left| S\left(\frac{\omega}{2}\right) \right|^2 + \left| \overline{S\left(\frac{\omega}{2} + \pi\right)} \right|^2 \right) d\omega$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(|H(\omega)|^2 + |G(\omega)|^2 \right) |S(\omega)|^2 d\omega$$

$$+ \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} \left| \mathrm{Re}\left((\overline{H(\omega)} H(\omega + \pi) + \overline{G(\omega)} G(\omega + \pi) \right) \right|$$

$$\left(|S(\omega)|^2 + \left| \overline{S(\omega + \pi)} \right|^2 \right) d\omega$$

$$\leq C_{\max} \frac{1}{2\pi} \int_{-\pi}^{\pi} |S(\omega)|^2 d\omega$$

$$= C_{\max} \sum_n |s(n)|^2.$$

Thus,

$$\sum_n |a[n]|^2 + \sum_n |d[n]|^2 \leq C_{\max} \sum_n |s(n)|^2$$

where

$$C_{\max} = \max_{\omega \in [-\pi, \pi]} \left(|H(\omega)|^2 + |G(\omega)|^2 \right)$$

$$+ \max_{\omega \in [-\pi/2, \pi/2]} |\mathrm{Re}(\overline{H(\omega)} H(\omega + \pi)$$

$$+ \overline{G(\omega)} G(\omega + \pi))|.$$

In particular, if

$$\begin{cases} |H(\omega)|^2 + |G(\omega)|^2 = 1, \\ \mathrm{Re}[\overline{H(\frac{\omega}{2})} H(\frac{\omega}{2} + \pi) + \overline{G(\frac{\omega}{2})} G(\frac{\omega}{2} + \pi)] = 0, \end{cases} \forall \omega \in [-\pi, \pi]$$

$$(10)$$

then

$$\sum_n |a[n]|^2 + \sum_n |d[n]|^2 = \sum_n |s[n]|^2. \qquad (11)$$

Thus, the decomposition (8)–(9) preserves energy.

With the results obtained in Steps 1 and 2, we can easily estimate the error between $b_{j,k}$ and $2^{-J/2} b_{j,k}^{(S)}$. For orthogonal wavelets, since $H(\omega) = \tilde{H}(\omega)$ and $G(\omega) = \tilde{G}(\omega)$, condition (10) holds as a direct consequence of (2). Thus, by recursively applying (11), we have the following error estimate:

$$\sum_{j \leq J-1} \sum_k |b_{j,k} - b_{j,k}^{(S)}|^2 = \sum_k |c_{J,k} - 2^{-J/2} x'[k]|^2.$$

In particular, if $f(t)$ is $2^J \pi$ band-limited, we have

$$\sum_{j \leq J-1} \sum_k |b_{j,k} - b_{j,k}^{(S)}|^2$$

$$= \frac{2^{J-1}}{\pi} \int_{-\pi}^{\pi} |\hat{f}(-2^J \omega)|^2 |Q(\omega) - \hat{\phi}(\omega)|^2 d\omega$$

$$= \frac{2^{J-1}}{\pi} \int_{-\pi}^{\pi} |\hat{f}(-2^J \omega)|^2 |A(\omega) - \hat{\phi}(\omega)|^2 d\omega. \qquad (12)$$

However, equations given by (10) in general fail for the biorthogonal wavelet case so that it is difficult to derive a precise error formula as done in the orthogonal case. However, since

$$c_{j-1,k} - 2^{-J/2} c_{j-1,k}^{(S)} = \sqrt{2} \sum_n h_{n-2k} \left(c_{j,n} - 2^{-J/2} c_{j,n}^{(S)} \right),$$

$$b_{j-1,k} - 2^{-J/2} b_{j-1,k}^{(S)} = \sqrt{2} \sum_n g_{n-2k} \left(c_{j,n} - 2^{-J/2} c_{j,n}^{(S)} \right),$$

for $j = J, J-1, \ldots$, we have

$$\sum_k |c_{j-1,k} - 2^{-J/2} c_{j-1,k}^{(S)}|^2$$

$$+ \sum_k |b_{j-1,k} - 2^{-J/2} b_{j-1,k}^{(S)}|^2$$

$$\leq C_{\max} |c_{j,n} - 2^{-J/2} c_{j,n}^{(S)}|^2, \quad j \leq J.$$

By recursively applying this error bound, we obtain

$$\sum_k |b_{j,k} - 2^{-J/2} b_{j,k}^{(S)}|^2 \leq C_{\max}^{J-j} |c_{J,n} - 2^{-J/2} c_{J,n}^{(S)}|^2, \quad j \leq J-1.$$

With respect to a $2^J \pi$ band-limited signal $f(t)$, we have the following upper bound on the error estimate:

$$\sum_k | b_{j,k} - 2^{-J/2} b_{j,k}^{(S)} |^2$$

$$\leq C_{\max}^{J-j} \frac{2^{J-1}}{\pi} \int_{-\pi}^{\pi} | \hat{f}(-2^J \omega) |^2 \left| Q(\omega) - \hat{\phi}(\omega) \right|^2 d\omega.$$

$$(13)$$

where $j \leq J - 1$.

IV. OPTIMAL FIR PREFILTER DESIGN

By examining (12) and (13), we see that for a signal with energy concentrated in frequency band $[-2^J \pi, 2^J \pi]$, the integral term

$$C_{f,\phi}(Q) = \int_{-\pi}^{\pi} |\hat{f}(-2^J \omega)|^2 |Q(\omega) - \hat{\phi}(\omega)|^2 d\omega \qquad (14)$$

is the dominant term of the error. Thus, it is natural to minimize $C_{f,\phi}(Q)$ for the design of prefilter $Q(\omega)$ to reduce the error between $b_{j,k}$ and $2^{-J/2} b_{j,k}^{(S)}$.

(a)

(b)
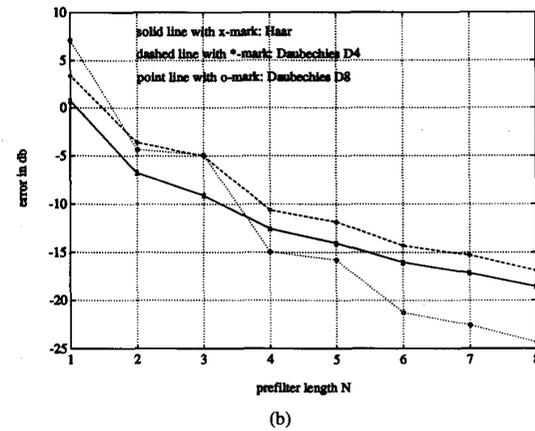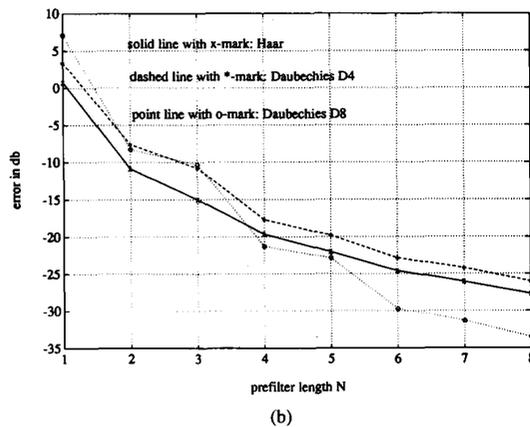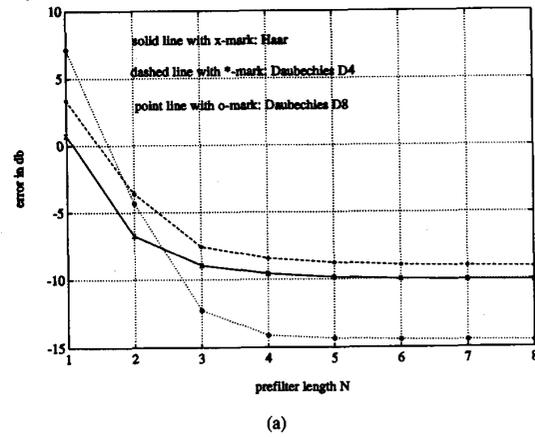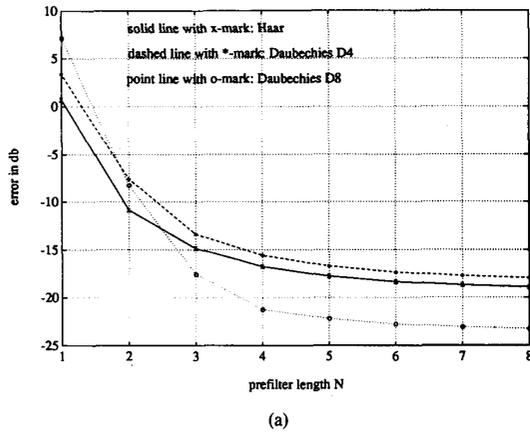
Fig. 1. Errors for: (a) causal; (b) noncausal signal dependent optimal prefilters in Example 1.

(a)

(b)

Fig. 2. Errors for: (a) causal; (b) noncausal signal independent optimal prefilters in Example 1.

TABLE I
OPTIMAL PREFILTERS FOR THE D_4 BASIS WITH $a = 0$

length N	$q_o[n]$, $0 \leq n \leq N-1$				
1	1.0000				
2	0.6466	0.5846			
3	0.6466	0.5846	-0.1975		
4	0.6466	0.5846	-0.1975	0.0980	
5	0.6466	0.5846	-0.1975	0.0980	-0.0675
6	0.6466	0.5846	-0.1975	0.0980	-0.0675
	0.0517				
7	0.6466	0.5846	-0.1975	0.0980	-0.0675
	0.0517	-0.0420			
8	0.6466	0.5846	-0.1975	0.0980	-0.8975
	0.0517	-0.0420	0.0353		
9	0.6466	0.5846	-0.1975	0.0980	-0.0675
	0.0517	-0.0420	0.0353	-0.0305	
10	0.6466	0.5846	-0.1975	0.0980	-0.0675
	0.0517	-0.0420	0.0353	-0.0305	0.0269

TABLE II
OPTIMAL PREFILTERS FOR THE D_4 BASIS WITH $a = 0.1$

length N	$q_o[n]$, $0 \leq n \leq N-1$				
1	1.0000				
2	0.6292	0.5527			
3	0.6179	0.5884	-0.1805		
4	0.6154	0.5942	-0.1975	0.0854	
5	0.6145	0.5961	-0.2015	0.0968	-0.0573
6	0.6140	0.5969	-0.2029	0.0998	-0.0659
	0.0431				
7	0.6138	0.5973	-0.2036	0.1010	-0.0683
	0.0500	-0.0345			
8	0.6136	0.5975	-0.2039	0.1015	-0.0693
	0.0520	-0.0402	0.0287		
9	0.6135	0.5976	-0.2042	0.1018	-0.0698
	0.0528	-0.0419	0.0336	-0.0246	
10	0.6135	0.5977	-0.2043	0.1020	-0.0700
	0.0532	-0.0427	0.0351	-0.0288	0.0214

TABLE III
SIGNAL DEPENDENT OPTIMAL PREFILTERS FOR THE BIORTHOGONAL WAVELET

length N_1	$q_o[n]$, $-N_1 \leq n \leq N_1$					
0	1.0000					
1	0.6766	0.4347	-0.0976			
2	-0.0783	0.6766	0.4347	-0.0976	0.0573	
3	0.0518	-0.0783	0.6766	0.4347	-0.0976	0.0573
	-0.0405					
4	-0.0379	0.0518	-0.0783	0.6766	0.4347	-0.0976
	0.0573	-0.0405	0.0314			
5	0.0298	-0.0379	0.0518	-0.0783	0.6766.	0.4347
	-0.0976	0.0573	-0.0405	0.0314	-0.0256	

TABLE IV
SIGNAL INDEPENDENT OPTIMAL PREFILTERS FOR THE BIORTHOGONAL WAVELET

length N_1	$q_o[n]$, $-N_1 \leq n \leq N_1$					
0	1.0000					
1	0.6749	4349	-0.0965			
2	-0.0770	0.6763	0.4348	-0.0974	0.0563	
3	0.0508	-0.0780	0.6765	0.4347	-0.0976	0.0571
	-0.0397					
4	-0.0371	0.0516	-0.0782	0.6766	0.4347	-0.0976
	0.0572	-0.0403	0.0307			
5	0.0291	-0.0377	0.0517	-0.0782	0.6766	0.4347
	-0.0976	0.0572	-0.0404	0.0312	-0.0250	

We consider the determination of the optimal filter $q_o[n]$ (or $Q_o(\omega)$) which has a finite length and minimizes the cost function $C_{f,\phi}(Q)$. To do so, let us expand $C_{f,\phi}(Q)$ as a function of $q[n]$

$$C_{f,\phi}(Q) = \int_{-\pi}^{\pi} |\hat{f}(-2^J\omega)|^2 \left| \sum_n q[n]e^{in\omega} - \hat{\phi}(\omega) \right|^2 d\omega$$

$$= A \sum_n q^2[n] - \sum_{n<m} B_{nm} q[n]q[m]$$

$$+ \sum_n C_n q[n] + D$$

where

$$A = \int_{-\pi}^{\pi} |\hat{f}(-2^J\omega)|^2 d\omega \qquad (15)$$

$$B_{nm} = 2\int_{-\pi}^{\pi} |\hat{f}(-2^J\omega)|^2 \cos((n-m)\omega)d\omega, \quad \text{for } n \neq m \qquad (16)$$

$$C_n = \int_{-\pi}^{\pi} |\hat{f}(-2^J\omega)|^2 \mathrm{Re}(\hat{\phi}(\omega)e^{-in\omega})d\omega \qquad (17)$$

$$D = \int_{-\pi}^{\pi} |\hat{f}(-2^J\omega)|^2 |\hat{\phi}(\omega)|^2 d\omega.$$

The solution of the above problem can be solved from the following equation

$$\frac{\partial C_{f,\phi}(Q)}{\partial q[n]} = 0, \quad \text{for } n \in \mathbf{Z}. \qquad (18)$$

Let the prefilter be an FIR filter of length N with filter coefficients $q[n] = 0$ for $n < N_1$ and $n \geq N_2$ and $N_2 - N_1 = N > 0$. Then, with (18), one can show that the optimal $q_o[n]$, $N_1 \leq n \leq N_2 - 1$, can be solved from the following linear equations:

$$\mathbf{T}_1 \mathbf{q}_o = \mathbf{C}_1,$$

where $\mathbf{T}_1 = (t_{nm})_{N_1 \leq n,m < N_2}$ with

$$t_{nm} = \begin{cases} 2A, & n = m, \\ -B_{nm}, & n \neq m, \end{cases}$$

$$\mathbf{C}_1 = (-C_{N_1}, -C_{N_1+1}, \ldots, -C_{N_2-2}, -C_{N_2-1})^t,$$

$$\mathbf{q}_o = (q_o[N_1], q_o[N_1+1], \ldots, q_o[N_2-2], q_o[N_2-1])^t$$

and A, B_{nm} and C_n are defined by (15), (16) and (17), respectively.

For some applications, it may be preferable that the prefilter does not depend on a signal while some partial knowledge of signals is available, say, the rough shape of signal spectra. To compute the WST coefficients for this class of signals, we may adopt a procedure using a nonnegative weighting function $F(\omega)$ to replace the term $|\hat{f}(-2^J\omega)|^2$ in (14). For example, consider

$$F(\omega) = e^{-a\omega^2}, \, a \geq 0 \qquad (19)$$

where the parameter a can be adjusted according to different applications. If the input signal is known in advance, it is natural that the error $b_{j,k}^{(S)}$ resulting from the signal-independent optimal prefilter is larger than that resulting from a signal-dependent optimal prefilter.

V. NUMERICAL EXAMPLES

We consider two numerical examples of optimal FIR prefiltering for the Shensa algorithm with orthogonal and biorthogonal wavelets, respectively.

Example 1—Orthogonal Wavelets: The test function $f(t)$ is a $2^6\pi$ band-limited signal whose Fourier spectrum is of the form

$$\hat{f}(\omega) = \begin{cases} e^{-(\omega/100)^2}, & |\omega| < 2^6\pi, \\ 0, & \text{otherwise.} \end{cases}$$

The wavelet bases studied include the Haar wavelets, the Daubechies D_4 and D_8. Both causal and noncausal prefilters are considered. For

a non-causal prefilter with length N, we always choose $-N_1 + 1 \geq N_2 - 1 \geq -N_1$ where $N = N_2 - N_1$.

First, we examine the error when the signal-dependent optimal prefilter is used in WST coefficient computation by the Shensa algorithm and plot the result in Fig. 1. We can see the significant improvement of the Shensa algorithm with a prefilter of length $N > 1$ from the Mallat algorithm corresponding to the case $N = 1$. Also, the error resulted from the Shensa algorithm by using a non-causal optimal prefilter is smaller than the one by using a causal optimal prefilter when the prefilter length $N \geq 4$. Second, we examine the error when the signal-independent optimal prefilter is used. The weighting function $F(\omega)$ used is $F(\omega) = e^{-a w^2}$ with $a = 0$ and 0.1. For $a = 0$, the optimal prefilter $q_o[n]$ can be solved from

$$\min_{q[n]} \int_{-\pi}^{\pi} \left| \sum_{N_1 \leq n < N_2} q[n] e^{in\omega} - \hat{\phi}(\omega) \right|^2 d\omega.$$

Thus, the optimal $q_o[n]$ is exactly the coefficients d_n, $N_1 \leq n < N_2$, of the Fourier series expansion of $\hat{\phi}(\omega)$ in $[-\pi, \pi]$. This implies that the optimal $q_o[n]$ with shorter length is the truncation of the one with longer length. We show in Tables I and II two designs of signal-independent optimal causal prefilters with lengths from 1 to 10 where the wavelet basis is the Daubechies D_4. Fig. 2 shows the error between $b_{j,k}$ and $b_{j,k}^{(S)}$ with the same parameters as in Fig. 1 except the prefilters used here are signal-independent with $F(\omega) = F_2(\omega)$. We see that the error resulting from the Shensa algorithm with $N > 1$ is also much smaller than the one resulting from the Mallat algorithm with $N = 1$. The properties in Fig. 2 are similar to the ones in Fig. 1 except the errors in Fig. 2 are generally larger than the ones in Fig. 1.

Example 2—Biorthogonal Wavelets: The biorthogonal wavelet basis adopted here is given in [2]. The filters are

$$H(\omega) = \left(\frac{1 + e^{-i\omega}}{2} \right)^2$$

and

$$\tilde{H}(\omega) = \left(\frac{1 + e^{i\omega}}{2} \right)^2 e^{-2i\omega} \left(1 + 2 \sin^2 \left(\frac{\omega}{2} \right) \right)$$

and filters $G(\omega)$ and $\tilde{G}(\omega)$ are determined according to (3). The test function considered is

$$f(t) = \frac{\sin(2^6 \pi t)}{\pi t}$$

which is $2^J \pi$ band-limited with $J = 6$. We focus on noncausal filters $q(-N_1), q(-N_1 + 1), \ldots, q(N_1 - 1), q(N_1)$ of length $N = 2N_1 + 1$, and both signal dependent and signal independent optimal prefilters are designed. For the signal independent case, we choose the weighting function

$$F(\omega) = e^{-a\omega^2}, \quad a = 0.01.$$

Tables III and IV show the filter coefficients for N_1 ranging from 0 to 5.

The errors e_j, $j = 5, 4, 3$, between the true WST coefficients and the computed ones are defined as

$$e_5 = \sum_{k=-120}^{120} \left| b_{5,k} - \frac{1}{8} b_{5,k}^{(S)} \right|^2,$$

$$e_4 = \sum_{k=-60}^{60} \left| b_{4,k} - \frac{1}{8} b_{4,k}^{(S)} \right|^2,$$

$$e_3 = \sum_{k=-30}^{30} \left| b_{3,k} - \frac{1}{8} b_{3,k}^{(S)} \right|^2.$$

(a)

prefilter length N

(b)
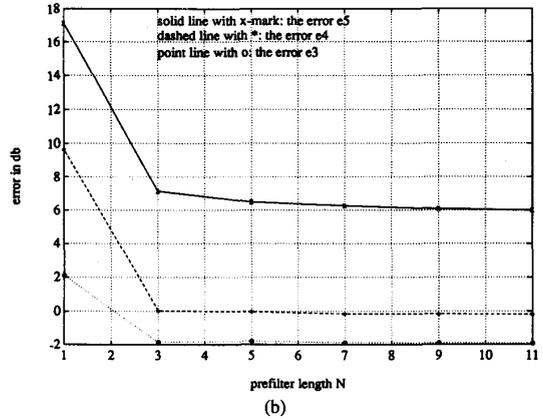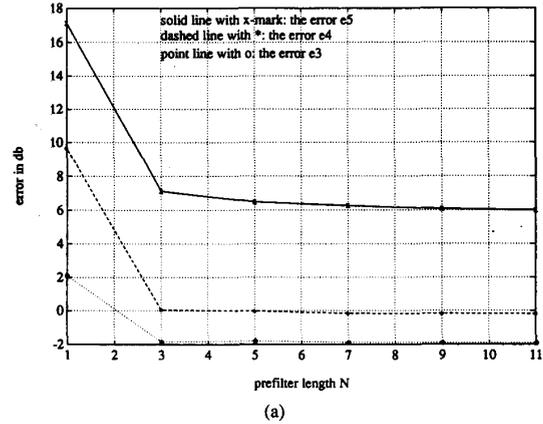
Fig. 3. The errors e_5, e_4 and e_3 between desired $b_{j,k}$ and the computed $b_{j,k}^{(S)}$ with (a) signal dependent and (b) signal independent optimal prefilters as given by Tables III and IV.

and plotted in Fig. 3. Again, we see a clear advantage of the Shensa algorithm with optimal prefilters of even very short length. There is no substantial difference for signal dependent and signal independent cases if an appropriate parameter a of the weighting function in (19) is used for this particular test problem.

VI. CONCLUSION

In this work, we studied the error estimate between the true WST coefficients and the computed ones from samples of a continuous time signal by using the Shensa algorithm with biorthogonal wavelet bases. We discussed the design of optimal prefilters used in the Shensa algorithm, and showed that they provide significant improvement on the accuracy of computed wavelet coefficients over the Mallat algorithm.

REFERENCES

[1] C. K. Chui, *An Introduction to Wavelets*. New York: Academic, 1992.
[2] A. Cohen, "Biorthogonal wavelets," in *Wavelets: A Tutorial in Theory and Applications*, C. K. Chui, Ed. New York: Academic, 1992, pp. 123–152.
[3] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, no. 5, pp. 961–1005, Sept. 1990.

[4] P. Goupillaud, A. Grossmann and J. Morlet, "Cycle-octave and related transforms in seismic signal analysis," *Geoexploration*, vol. 23, 1984/85, pp. 85–102.

[5] M. Holschneider, R. Kronland-Martinet, J. Morlet and P. Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets, Time-Frequency Methods and Phase Space*, A. G. J. M. Combes and P. Tchamitchian, Eds. Berlin: Springer, IPTI, 1989, pp. 286–297.

[6] S. Mallat, "Multifrequency channel decompositions of images and wavelet methods," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, 1989, pp. 2091–2110.

[7] O. Rioul and P. Duhamel, "Fast algorithms for discrete and continuous wavelet transform," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, Mar. 1992, pp. 569–586.

[8] M. J. Shensa, "Affine wavelets: wedding the Atrous and Mallat algorithms," *IEEE Trans. Signal Processing*, vol. 40, no. 10, Oct. 1992, pp. 2464–2482.

Efficient Closed-Form Estimation of Multivariate Moving-Average Processes Using Higher Order Statistics

Hong Chen, Tao Chen, and Tianping Chen

Abstract—In this paper, an improved algorithm is proposed for estimation of nonGaussian, nonminimum phase, multivariate moving average (MA) processes using higher order cumulants. This algorithm improves upon earlier results [3] [11] and contains development beyond existing algorithms. It provides a closed-form solution to estimating the MA parameter matrices (up to a post-multiplication by a permutation matrix), and (under certain assumptions) eliminates the indeterminacy associated with scaling. The algorithm is theoretically derived and tested via computer simulations. In addition, it will be shown that this algorithm is computationally more efficient than the one in [11]. Finally, the effect of imperfect input data on our algorithm is tested via simulations.

I. Introduction

Estimation of multivariate moving-average processes is one of the fundamental problems in time series analysis, with applications in multivariate linear prediction, spectral estimation [9], image coding and multivariate control. Conventional methods based only on second-order statistics generally require the assumption that the MA models are of minimum phase. The use of higher-order statistics makes it possible to estimate MA models, which may be nonminimum phase, noncausal and/or non-Gaussian [4]. Moreover, a cumulant-based estimator is insensitive to additive Gaussian observation noises, white or colored.

Because of these features, cumulant-based methods have received considerable interest [1]-[13]. Although much of the work in this area is concentrated on the single channel case, third-order cumulant-

based methods were recently reported for estimating parameters of multivariate MA models in [3] and [4]. It was shown in [3] and [11] that the information contained in the third-order output cumulants theoretically is sufficient for the identification of the multivariate MA parameter matrices, up to a post-multiplication by a permutation matrix. In particular, when driven by i.i.d. (independent and identically distributed) input signals, these parameters satisfy a set of nonlinear equations involving third-order output cumulants, now referred to as *cumulant-based identification (CBI)* equations [3] [11]. Similar CBI equations of any (higher) order were derived using the Kronecker product in [10].

Although the issue of theoretic identifiability has been well addressed by these papers, there is a need for practical algorithms. However, solving these simultaneous nonlinear CBI equations is nontrivial. A closed form solution for the bivariate case was given in [1]; adaptive algorithms have been proposed in [1] and [6]. A solution was proposed in [3] when the initial model matrix is in a triangular form. Recently, an interesting eigenstructure-based *iterative* algorithm for solving the third-order CBI equations was proposed in [11], which identifies the MA parameter matrices up to a post-multiplication by a permutation matrix *and* a diagonal (scaling) matrix.

In this correspondence, we improve upon the results in [3] and [11] and independently develop a closed-form identification algorithm to solve this problem. Instead of relying on some intermediate matrix variables whose positive-definiteness is required, and must be guaranteed by the clever use of *iterative* procedures such as a perceptron algorithm [11], this improved algorithm adopts an alternative approach which circumvents this difficulty. In addition, this algorithm removes the indeterminacy of scaling in the final results. Finally, the computational efficiency is considerably improved.

The authors wish to point out that recently, an improvement [12] was independently made to solve the same problem in closed form, while retaining the scaling and permutation indeterminacy.

II. Problem Formulation

Let $\{\boldsymbol{y}(t) \in \boldsymbol{R}^n\}$ be a stationary process. The third-order cumulants of $\{\boldsymbol{y}(t)\}$ are defined as the set of $n \times n$ matrices [3] [11]

$$C_i(m_1, m_2) = \mathrm{E}\{\boldsymbol{y}(t + m_1)\boldsymbol{y}^T(t)y_i(t + m_2)\}, \quad i = 1, 2, \cdots, n \tag{1}$$

where $y_i(t)$ denotes the ith component of vector $\boldsymbol{y}(t)$ and $\mathrm{E}\{\cdot\}$ denotes the mathematical expectation operator. Consider a stationary n-variate, qth order MA vector process given by

$$\boldsymbol{y}(t) = \sum_{k=0}^{q} H(k)\boldsymbol{x}(t - k) + \boldsymbol{n}(t) \tag{2}$$

satisfying the following assumptions [3] [11]:

1) $H(0)$ and $H(q) \in \boldsymbol{R}^{n \times m}$ are of full column rank.

2) $\{\boldsymbol{x}(t)\}$ is an m-variate, zero-mean, stationary and nonGaussian process with components $\{x_i(t)\}$. Spatial order m is not necessarily known.

3) $\{\boldsymbol{n}(t)\}$ with components $\{n_i(t)\}$ is an n-variate, Gaussian, perhaps colored, zero-mean process independent of $\{\boldsymbol{x}(t)\}$.

4) $\mathrm{E}\{x_i(t_1)x_j(t_2)x_k(t_3)\} = \begin{cases} 1, & \text{for } i = j = k \\ & \text{and } t_1 = t_2 = t_3 \\ 0, & \text{otherwise.} \end{cases}$