The Coding Gain of Multiplexed Wavelet Transforms

Gianpaolo Evangelista, Member, IEEE

Abstract—In a recent paper, we introduced the multiplexed wavelet transform (MWT) and pointed out its potential applications to the analysis, synthesis, processing and coding of pseudo-periodic signals such as voiced speech and music. Coders based on the MWT have been shown to outperform the conventional subband coders when a reliable pitch parameter can be extracted from data. In this paper, we investigate the effects of uniform quantization of the MWT coefficients. We compare the performance of the new coders with that of WT, block-DCT, and KLT coders in terms of the coding gain achieved when optimal bit allocation schemes are adopted.

I. INTRODUCTION

PITCH-SYNCHRONOUS methods in wavelet theory were introduced in [4] and [6], where we defined the multiplexed (MWT) and the pitch-synchronous (PSWT) wavelet transforms. The pitch information that can be extracted from signals was exploited in order to obtain a representation of pseudo-periodic segments of the signal in terms of an average oscillatory behavior plus fluctuations at several scales. Applications of the transforms range from speech and music processing to coding and image texture analysis. In this paper, we will focus on the performance of the MWT as a coder. Transform coding requires quantization of the expansion coefficients. Although—as we showed with the support of experiments in [6]—adaptive and dynamic bit-allocation schemes may improve the performance of the MWT coder, we will confine ourselves to uniform quantization with fixed bit rates. In this case, the analysis of performance is more amenable and the mathematics greatly simplified. In designing the coder, we need to determine the optimal number of bits to be allocated to each quantizer. The performance of the coder may be evaluated by computing the coding gain with respect to PCM. For this purpose, we will extend to the MWT the results presented in [10] and [3] for coders based on the ordinary wavelet transform (WT). We will carry out the evaluation of performance using both real-life and model signals in the form of AR comb processes.

A. Organization of the Paper

In Section II, we review the basic definitions and properties of the MWT and its modified version (the MMWT), and we illustrate the coders that can be implemented with these transforms. In Section III, we provide useful results for computing the variance of the error introduced by coefficient quantization. We derive the optimal bit allocation schemes for both the MWT and MMWT coders and provide expressions for the coding gain. In Section IV, we evaluate the performance of the coders and compare them with the WT, KLT, and DCT coders. Finally, in Section V, we draw our conclusions. The quantizer noise model, on which most of our derivations are based, is discussed in Appendix A. In Appendix B, we provide the proof of a simple orthogonality Lemma stated in Section III. Details about the Lagrange multiplier method for optimum bit allocation may be found in Appendix C.

II. CODERS BASED ON THE MWT

In this section, we briefly review the MWT and its modified version: the MMWT. We describe the coders that can be designed with these transforms and make some preliminary considerations on the quantization noise at the decoder.

A. Multiplexed Wavelet Transforms

Given a sequence of integer estimates \( P(k) \) of the local periods—expressed in number of samples—of a pseudo-periodic signal \( x[k] \), the pitch-synchronous representation of \( x(k) \) is given by the following decomposition:

\[
x(k) = \sum_{i} \sum_{q=0}^{P(i)-1} x_q(i) \delta(k - q - M(i))
\]

where \( \delta(k) = 1 \) if \( k = 0 \) is the Kronecker delta sequence, \( M(k) = \sum_{r=0}^{k-1} P(r) \), and \( x_q(k) = x(q + M(k)) \), \( q = 0, 1, \ldots, P(k) - 1 \) is the \( q \)th section of the signal. The pitch-synchronous wavelet transform (PSWT) is obtained by discrete wavelet transforming each section \( x_q(k) \) individually. In many samples taken from speech or music, the sections are smooth lowpass sequences, except on transients. Therefore, most of the energy is contained in the low-rate wavelet expansion coefficients corresponding to large scales. As we will show, this property may be exploited in coding systems since fewer bits may be allocated to the high-rate coefficients, obtaining a large coding gain. In this paper, we will be mainly concerned with the special case of constant pitch \( P(k) = P \). In this case, (1) simplifies to

\[
x(k) = \sum_{i} \sum_{q=0}^{P-1} x_q(i) \delta(k - q - iP)
\]

where the sections

\[
x_q(k) = x(q + kP)
\]

\[
= \sum_{r} x(r) \delta(r - q - kP), \quad q = 0, 1, \ldots, P - 1
\]

Manuscript received February 23, 1994; revised January 2, 1996. The associate editor coordinating the review of this paper and approving it for publication was Prof. Baru Onaral.
The author is with ACEL, Department of Physical Sciences, Università Federico II Naples, Mostra d’Oltremare, Pad. 20, 80125 Napoli, Italy (e-mail: evangelista@na.infn.it).
Publisher Item Identifier S 1053-587X/96/0521-31. X.
are time-domain versions of the polyphase components of the signal [1]. Over constant-pitch segments of the signal, the PSWT reverts to the MWT, which is defined as the set of coefficients obtained by wavelet transforming each polyphase component individually. This result may be concisely expressed by defining the multiplexed wavelets (MW) and multiplexed scaling (MS) sequences, respectively, as follows:

\[ \zeta_{n,m,q}(r) = \sum_k \delta(r - q - kP)\psi_{n,m}(k) = \zeta_{n,0,0}(r - q - 2^n m P) \]  

(4)

and

\[ \vartheta_{n,m,q}(r) = \sum_k \delta(r - q - kP)\phi_{n,m}(k) = \vartheta_{n,0,0}(r - q - 2^n m P) \]  

(5)

\( n = 1,2,\ldots,m \) integer; \( q = 0,1,\ldots,P - 1 \), where \( \{\psi_{n,m}(k)\} \) is a set of discrete orthogonal wavelets [2], [8], [5], with associated scaling set \( \{\phi_{n,m}(k)\} \). Notice that the multiplexed wavelet \( \zeta_{n,m,q}(k) \) is obtained from the wavelet \( \psi_{n,m}(k) \) by inserting \( P - 1 \) zeros between successive samples and shifting by \( q \) samples. The finite scale \( N \) multiplexed wavelet decomposition of the signal \( x(k) \) is given by the following expansion:

\[ x(k) = \sum_{q=0}^{P-1} \sum_{n=1}^{N} \left( \sum_{r=0}^{n} X_{n,q}(m)\psi_{n,m,q}(r) + X_{N+1,q}(m)\vartheta_{n,m,q}(r) \right) \]  

(6)

where

\[ X_{n,q}(m) = \sum_r x(r)\zeta_{n,m,q}(r), \quad n = 1,2,\ldots,N; \quad q = 0,1,\ldots,P - 1 \]  

(7)

are the multiplexed wavelet coefficients, and

\[ X_{N+1,q}(m) = \sum_r x(r)\vartheta_{n,m,q}(r), \quad q = 0,1,\ldots,P - 1 \]  

(8)

are the scaling coefficients. The set of coefficients (7) and (8) form the MWT, and (6) is the inverse MWT (IMWT). The frequency spectrum of the multiplexed scaling sequence consists of multiple frequency bands that are centered on the harmonics at \( v_p = p/P, \quad p = 0,1,\ldots,[P/2] \). On the other hand, the frequency spectra of multiplexed wavelets have a multiple-band structure consisting of sidebands of the harmonics. As \( n \) grows, these bands narrow and get closer to the harmonics. By means of the MWT, the signal is decomposed into an asymptotically periodic trend—the scaling component—plus fluctuations at several scales. The multiplexed wavelets \( \zeta_{n,m,q}(k) \) are defined over three indices. The first index corresponds to scale, the second to time measured in increments of \( 2^n P \) samples, and the last is the intraperiod index. The last two indices of multiplexed wavelets may be combined into a single progressive time index. One defines the modified multiplexed wavelets (MMW’s) and scaling (MMS’s) sequences, respectively, as follows:

\[ \hat{\zeta}_{n,p}(k) = \sum_{m=0}^{P-1} \sum_{q=0}^{m} \zeta_{n,m,q}(k)\delta(p - q - mP) \]  

(9)

and

\[ \hat{\vartheta}_{n,p}(k) = \sum_{m=0}^{P-1} \sum_{q=0}^{m} \vartheta_{n,m,q}(k)\delta(p - q - mP) \]  

(10)

\( n = 1,2,\ldots,N; \quad p \) integer. The modified multiplexed wavelet decomposition is given by the following expansion:

\[ x(k) = \sum_{p} \left( \sum_{n=1}^{N} \hat{X}_n(p)\hat{\zeta}_{n,p}(k) + \hat{X}_{N+1,p}(k)\hat{\vartheta}_{N,p}(k) \right) \]  

(11)

where

\[ \hat{X}_n(p) = \sum_r x(r)\hat{\zeta}_{n,p}(r), \quad n = 1,2,\ldots,N \]  

(12)

are the MMW coefficients, and

\[ \hat{X}_{N+1}(p) = \sum_r x(r)\hat{\vartheta}_{N,p}(r) \]  

(13)

are the modified scaling coefficients. The ensemble of (12) and (13) form the MMWT, and (11) is the inverse MMWT (IMMWWT). The following relationships hold between MWT and MMWT:

\[ \hat{X}_n(q + mP) = X_n(m), \quad n = 1,2,\ldots,N + 1; \quad q = 0,1,\ldots,P - 1; \quad m \) integer. \]  

(14)

The MWT may be given in a vector notation stemming from the block form of the input defined as the sequence of \( P \times 1 \) vectors:

\[ x(k) = [x(kP) \quad x(kP + 1) \quad \cdots \quad x(kP + P - 1)]^T. \]

By defining the block-MWT coefficients as the \( P \times 1 \) vectors

\[ X_n(m) = [X_{n,0}(m) \quad X_{n,1}(m) \quad \cdots \quad X_{n,P-1}(m)]^T, \]

the block-form of expansion (6) may be written as follows:

\[ x(k) = \sum_m \left( \sum_{n=1}^{N} X_n(m)\psi_{n,m}(k) + X_{N+1,m}(m)\varphi_{N,m}(k) \right) \]  

(15)

where \( \psi_{n,m}(k) \) and \( \varphi_{N,m}(k) \), respectively, are ordinary wavelets and scaling sequences.

B. MWT and MMWT Coders

In coders based on the MWT, the expansion coefficients (7) and (8) are quantized, and the signal is decoded by computing the IMWT of the quantized coefficients. The block diagram of the MWT coder is shown in Fig. 1(a). There, the ordinary WT blocks, each implemented as an \( N + 1 \) channel iterated filterbank, are represented as one input \( N + 1 \) output devices. Outputs labeled \( n = 1,\ldots,N \) correspond to wavelet coefficients at scale level \( n \) and output \( N + 1 \) to scaling coefficients. In order to compute the MWT, the input signal \( x(k) \) is demultiplexed over \( P \) channels that are input
are the multiplexed wavelet and scaling coefficient quantization errors. Finally, the total output error is obtained by multiplexing the components \( e_q(k) \), \( q = 0, 1, \ldots, P - 1 \) of the block output error

\[
e(k) = \sum_{q=0}^{P-1} e_q(r) \delta(k - q - rP).
\]  

(19)

The block diagram of the MMWT coder is shown in Fig. 2(a). Again, the ordinary WT blocks are represented as one input \( N + 1 \) output devices. In order to compute the MMWT, the input signal \( x(k) \) is demultiplexed over \( P \) channels feeding the wavelet transform blocks. The \( P(N+1) \) outputs of the \( P \) WT blocks are multiplexed over \( N + 1 \) channels, using one multiplexer for each scale level wavelet and scaling output. These channels are individually quantized with quantizers \( Q \) obtaining the sequences \( [\hat{X}_n(p)]_Q \). The signal is decoded by taking the IMMWT of the quantized coefficients, as shown in Fig. 2(b). Results similar to those obtained for the MWT coder hold for the MMWT, i.e., the decoded signal

\[
\hat{x}_c(k) = \sum_{p} \left( \sum_{n=1}^{N} [\hat{X}_n(p)]_Q \hat{q}_{n,p}(k) + [\hat{X}_{N+1}(p)]_Q \hat{\varphi}_{N,p}(k) \right)
\]  

(20)

is affected by a total error

\[
\hat{e}(k) = \hat{x}_c(k) - x(k)
\]  

(21)

where

\[
e_{\hat{x}}(p) = [\hat{X}_n(p)]_Q - \hat{X}_n(p)
\]  

(22)

are the coefficient quantization errors.

The main difference between MWT and MMWT coders lies in the number of independent quantizers. A total of \( P(N+1) \) quantizers is needed for the MWT compared with the \( N+1 \) needed for the MMWT. Several types of quantizers may be adopted, depending on the specific signal and on the bit-rate requirement of the coder. If we consider uniform quantizers, each operating with a sufficiently large number of bits, then the coefficient quantization errors may be modeled as additive zero-mean wide-sense stationary noise sources whose variances are proportional to the average variances of the inputs to the quantizers (see Appendix A and [7] and [11]).

The average variance of a zero-mean discrete random process \( y(k) \) is defined as follows:

\[
\sigma_y^2 = \lim_{L \to \infty} \frac{1}{2L} \sum_{k=-L}^{L-1} E\{[y(k)]^2\}.
\]  

(23)

As given in (A-1), in this model, the noise variances of wavelet and scaling coefficients in the MWT coder are the following:

\[
\sigma_{n,q}^2 = E\{e_{X_{n,q}}^2(m)\} = \varepsilon^2 \sigma_{\varphi_{n,q}}^2,
\]  

(24)
where \( R_{n,q} \) denote the number of bits of the \( P(N + 1) \) quantizers. The factor may depend on the indices \( n \) and \( q \) if overflow scaling is individually set. However, in order to simplify our formulae and to make our results compatible with those reported in [10], we use a uniform scaling factor. The general case may be easily derived by means of the following substitutions: \( \sigma_{X_n,q}^2 \rightarrow \epsilon_{n,q}^2 \sigma_{X_n,q}^2 \), \( \epsilon^2 \rightarrow 1 \). Similarly, the noise variances of wavelet and scaling coefficients in the MMWT coder are the following:

\[
\sigma_n^2 = E\{v_n^2, (m)\} = 2^{-2R_n} \sigma_{X_n}^2, \quad n = 1, 2, \ldots, N + 1
\]

(25)

where \( R_n \) are the number of bits of the \( N + 1 \) quantizers. Again, the scaling factor \( \epsilon \) is kept constant over the channels, the general case being retrieved by means of substitutions:

\[
\sigma_{X_n}^2 \rightarrow \epsilon_{n}^2 \sigma_{X_n}^2, \quad \epsilon^2 \rightarrow 1.
\]

The average variances of the MWT and MMWT coefficients are related. Using (14), it is easy to show that \( \sigma_{X_n}^2 (q + mP) = \sigma_{X_{n,q}}^2 (m) \), from which it follows that

\[
\sigma_{X_n}^2 = \frac{1}{P} \sum_{q=0}^{P-1} \sigma_{X_{n,q}}^2.
\]

(26)

### III. Coding Gain

In designing a coder with constant total bit-rate and based on a generic transform \( T \), we have to determine the optimal bit allocation scheme, i.e., how many bits do we have to allocate to each quantizer in order to minimize the output error variance \( \sigma_{e,T}^2 \)? Moreover, in the assumption of optimal bit allocation, we would like to know what is gained by using the transform coder. A relative measure of the system error is given by the coding gain

\[
G_T = \sigma_{e,PCM}^2 / \sigma_{e,T}^2
\]

(27)

of the generic \( T \) coder over straight PCM quantization of the signal. In order to establish the optimal bit allocation scheme and evaluate the coding gain, we need to derive an expression for the total error variance in terms of the error correlation of the coefficients. Using this result, which is shown in Lemma I, it is easy to compute the average total variance of the MWT (Theorem I) and MMWT (Theorem II) coders. In both cases, the expressions we derive are similar, in form, to that of the WT coder [10].

### A. Average Total Error Variance

As in the case of the WT coder, the total output error of the MWT and MMWT coders (19) and (21) is not necessarily a stationary process, even in the assumption that the quantizer noises can be modeled as zero-mean uncorrelated processes. Thus, the total error variance depends on time. However, if we are interested in fixed optimal bit allocation schemes, our objective will be that of minimizing the average variance of the error. The following result will be useful for computing the average variance.

**Lemma I:** Let \( x(k) = \sum_{n} x_n f_n(k) \) be the expansion of a zero-mean random process \( x(k) \) over a complete set of orthogonal sequences \( \{f_n(k)\}_{n \in \mathbb{N}} \). Then,

\[
\sigma_x^2 = \lim_{L \to \infty} \frac{1}{2L} \sum_{k=-L}^{L-1} \sum_{n} \sigma_{X_n}^2 |f_n(k)|^2
\]

(28)

where \( \sigma_x^2 \) is the time-average variance of \( x(n) \), and \( \sigma_{X_n}^2 = E\{|X_n|^2\} \).

A straightforward proof is given in Appendix B. Although coefficient crosscorrelation terms may contribute to the output variance, this Lemma states that they do not affect the average output variance. Clearly, Lemma I continues to hold when the basis sequences are characterized by more than one index, as in the case of wavelets. Next, we will show that the results found in [10] for the total error variance of WT coders extend to both the MWT and the MMWT, provided that the coefficient quantization errors are suitably interpreted. More specifically, we will prove the following theorem.

**Theorem I:** The average variance of the total error of the MWT coder with uniform quantizers modeled as in (24) is the following:

\[
\sigma_e^2 = \sum_{n=1}^{N} 2^{-n} \sigma_n^2 + 2^{-N} \sigma_{N+1}^2
\]

(29)

where

\[
\sigma_n^2 = \frac{1}{P} \sum_{q=0}^{P-1} \sigma_{n,q}^2, \quad n = 1, 2, \ldots, N + 1.
\]

(30)

**Proof:** It is easy to show from (19) that the average variance of the total output error equals the mean of the
average variances of the $P$ block error components
\begin{equation}
\sigma^2_e = \frac{1}{p} \lim_{K \to \infty} \frac{1}{2K} \sum_{k=K}^{K-1} E|e^T(k)\bar{e}(k)|^2.
\end{equation}

Notice that (17) is the expansion of the block error process over a complete set of orthonormal wavelets and scaling sequences. Application of Lemma I and substitution in (31) yields
\begin{equation}
\sigma^2_e = \lim_{K \to \infty} \frac{1}{2KP} \sum_{k=-K}^{K-1} \sum_{m=1}^{N} \left( \sum_{n=1}^{N} v_n^T(m) v_n(m) |\psi_{n,m}(k)|^2 \right)
\end{equation}
\begin{equation}
+ \left( \sum_{n=1}^{N} v_{N+1}^T(m) v_{N+1}(m) |\varphi_{N,m}(k)|^2 \right)
\end{equation}
\begin{equation}
where \quad v_n(m) = [\sigma_{n,0}(m) \sigma_{n,1}(m) \cdots \sigma_{n,P-1}(m)]^T.
\end{equation}
However, as discussed in the previous section and in Appendix A, if the coefficient quantization noises may be modeled as WSS processes whose variances are given in (24), the vectors $v_n(m)$ are constant: $v_n(m) = v_n$. In taking the limit as $K \to \infty$ in (32), it is convenient to let $K = 2^{N-1} L$ and perform the limit over $L$. That is
\begin{equation}
\sigma^2_e = \lim_{L \to \infty} \frac{1}{2NPL} \sum_{k=-2^{N-1} L}^{2^{N-1} L-1} \sum_{n=1}^{N} v_n^T v_n W_n(k)
\end{equation}
where
\begin{equation}
W_n(k) = \sum_{m=1}^{N} |\psi_{n,m}(k)|^2 = \sum_{m=1}^{N} |\varphi_{N,0}(k-2^m m)|^2,
\end{equation}
\begin{equation}
where \quad n = 1, 2, \ldots, N
\end{equation}
and
\begin{equation}
W_{N+1}(k) = \sum_{m=1}^{N} |\varphi_{N,m}(k)|^2 = \sum_{m=1}^{N} |\varphi_{N,0}(k-2^m m)|^2.
\end{equation}
Notice that $W_n(k)$ and $W_{N+1}(k)$ are periodic with period $2^n$ and $2^N$, respectively. Thus
\begin{equation}
2^{N-1} L \sum_{k=-2^{N-1} L}^{2^{N-1} L-1} W_n(k) = 2^{N-n} L \sum_{k=0}^{2^{N-1}-1} W_n(k) = 2^{N-n} L
\end{equation}
and
\begin{equation}
2^{N-1} L \sum_{k=-2^{N-1} L}^{2^{N-1} L-1} W_{N+1}(k) = L \sum_{k=0}^{2^{N-1}-1} W_{N+1}(k) = L
\end{equation}
where the right-most equalities in (36) and (37) hold since wavelets and scaling sequences are normalized. Substitution of the last two results into (33) yields (29).

The following theorem extends the results of Theorem I to the MMWT coder.

**Theorem II:** The average variance of the total error of the MMWT coder with uniform quantizers modeled as in (25) is the following:
\begin{equation}
\bar{\sigma}^2_e = \sum_{n=1}^{N} 2^{-n} \sigma^2_n + 2^{-N} \sigma^2_{N+1}.
\end{equation}

The proof of this theorem is similar to that of Theorem I and will be omitted.

**B. Optimal Bit Allocation**

From Theorems I and II and from [10], we conclude that the average total error variance of the WT, MWT, and MMWT coders can be put in the general form
\begin{equation}
\bar{\sigma}^2_e = \sum_{n=1}^{N} 2^{-n} a_n + 2^{-N} a_{N+1}.
\end{equation}
For the MMWT coder, we have $a_n = \hat{\sigma}^2_n$, where $\hat{\sigma}^2_n$ are the coefficient quantization error variances. According to (29), for the MWT coder, we have $a_n = \sigma^2_n$, with $\sigma^2_n$ given in (30). In both cases, the optimal bit allocation scheme may be found by minimizing (39) subjected to the constant bit-rate constraint. This constraint has the following form:
\begin{equation}
R = \sum_{n=1}^{N} 2^{-n} r_n + 2^{-N} r_{N+1}
\end{equation}
where $R$ is the total bit rate of the coder. In both the WT and MMWT coder, we have $r_n = R_n$, where the $R_n$ are equal to the bit rates of the $N + 1$ quantizers. For the MWT coder, we have $r_n = \frac{1}{P} \sum_{n=0}^{P-1} R_n$, where $R_n$ are equal to the bit rates of the $P(N + 1)$ quantizers. The average total error variance can be expressed in terms of the average variances of the inputs to the quantizers by substituting the suitable model variances (24) or (25). We recall that these models are valid only when a sufficient number of bits per quantizer is allocated. Using the Lagrange multiplier method to minimize (39) subject to (40), we obtain the optimal bit allocation schemes (see Appendix C). For the MMWT case, we have
\begin{equation}
R_n = R + \frac{1}{2} \log_2 \left( \frac{\bar{\sigma}^2_{n+1}}{D_{\text{MMWT}}} \right)
\end{equation}
where
\begin{equation}
D_{\text{MMWT}} = \left( \frac{\bar{\sigma}^2_{N+1}}{\bar{\sigma}^2_n} \right)^{1/2N} \prod_{n=1}^{N} \left( \frac{\bar{\sigma}^2_n}{\bar{\sigma}^2_N} \right)^{1/2N}.
\end{equation}
The value of the Lagrange multiplier is
\begin{equation}
\lambda = 2^{2^{-N}\tau} \bar{\sigma}^2_n = 2^{2\lambda} n = 1, 2, \ldots, N + 1
\end{equation}
showing that optimal bit allocation is achieved when the variances of the errors introduced by each of the quantizers are equal.

The optimal bit allocation scheme for the MWT is the following:
\begin{equation}
R_n = R + \frac{1}{2} \log_2 \left( \frac{\bar{\sigma}^2_{n+1}}{D_{\text{MWT}}} \right)
\end{equation}
where
\begin{equation}
D_{\text{MWT}} = \left( \prod_{n=0}^{P-1} D_q \right)^{1/P}
\end{equation}
with
\begin{equation}
D_q = \left( \frac{\bar{\sigma}^2_{n+1}}{\bar{\sigma}^2_n} \right)^{1/2N} \prod_{n=1}^{N} \left( \frac{\bar{\sigma}^2_n}{\bar{\sigma}^2_{N+1}} \right)^{1/2N}.
\end{equation}
The value of the Lagrange multiplier is
\[ \lambda = 2^{2-2R_{n,\sigma}} \sigma_{n,q}^2 = 2^{2-2R_{n,q}} \]
\[ n = 1, 2, \ldots, N + 1; \quad q = 0, 1, \ldots, P - 1. \]  \hfill (47)

Again, optimal bit allocation requires the variances of the quantization errors to be equalized.

We remark that (41) and (42) are identical in form to the WT case. Conversely, in (45) the geometric means—over the period P—of the variances of the transform coefficients appear.

C. Coding Gain

The optimal allocation of bits per quantizer is given in (41), (42), and (44)–(46), respectively, for the MMWT and the MWT coders. Direct substitution of these quantities in the model variances inserted in the general error variance expression (39) yields, respectively, for the MMWT and MWT coders
\[ \sigma_{MMWT}^2 = \sigma_e^2 = \varepsilon^2 2^{-2R_{e,\sigma}} D_{MMWT} \]  \hfill (48)
\[ \text{and} \]
\[ \sigma_{MWT}^2 = \sigma_e^2 = \varepsilon^2 2^{-2R_{e,\sigma}} D_{MWT}. \]  \hfill (49)

Since the error variance of the PCM coder is \( \sigma_{PCM}^2 = \varepsilon^2 2^{-2R_{e,\sigma}^2} \), the resulting coding gains are
\[ G_{MMWT} = \frac{\sigma_{PCM}^2}{\sigma_{MMWT}^2} \]  \hfill (50)
\[ \text{and} \]
\[ G_{MWT} = \frac{\sigma_{PCM}^2}{\sigma_{MWT}^2}. \]  \hfill (51)

These results hold when the quantizer scaling factors are kept equal and identical to the PCM scaling factor \( \varepsilon_{x}^2 \). By choosing unequal scaling factors, one may design more efficient coders. As already remarked in Section II-B, the variable scaling factor case may be obtained by performing, when appropriate, the following substitutions in (41)–(51)
\[ \varepsilon_{x}^2 \rightarrow \varepsilon_{n}^2, \quad \varepsilon_{n,q}^2 \rightarrow \varepsilon_{n,q}^2 \quad \text{and} \quad \varepsilon^2 \rightarrow \frac{1}{2}. \]

We remark that
\[ G_{MWT} \geq G_{MMWT}, \]  \hfill (52)
when the scaling factors are equal. In fact, from the arithmetic-geometric mean inequality, we have
\[ \frac{1}{P} \sum_{q=0}^{P-1} \sigma_{n,q}^2 \geq \left( \prod_{q=0}^{P-1} \sigma_{n,q}^2 \right)^{1/P}. \]  \hfill (53)

Inserting (26) in (42), we obtain, in view of (53), \( D_{MMWT} \geq D_{MWT} \). The equality sign in (52) holds iff for any fixed \( n = 1, 2, \ldots, N + 1 \), all \( \sigma_{n,q}^2 \) are equal. This happens, for example, when the input is a WSS process. In that case, since decimation by \( P \) does not alter WSS stationarity, the signal sections \( x_n(k) = x(q + kP) \) are (individually) WSS processes with autocorrelations \( R_{x}(k) = R_{x}(kP) \). It is easy to show that the MWT coefficient variances are equal for any \( q \), resulting in
\[ \sigma_{n,q}^2 = \sum_{r} R_{x}(rP) \rho_{\phi_{n,0}}(r), \quad n = 1, 2, \ldots, N, \]  \hfill (54)
and
\[ \sigma_{n,q+1}^2 = \sum_{r} R_{x}(rP) \rho_{\phi_{N,0}}(r) \]  \hfill (55)

where \( \rho_{f}(r) = \sum_k f(k)f(k+r) \) is the deterministic autocorrelation of the generic sequence \( f(k) \).

The MWT coder, requiring more quantizers than the MMWT coder, may be useful if we know \textit{a priori} that the coefficient sequences pertaining to the same scale level and with different interperiod indices are not identically distributed.

The coding gains of the MWT and MMWT coders are greater than unity. The proof is easy if we assume that the input is a wide sense cyclostationary (CWSS) \( \rho \) process with period \( P \), i.e., if the signal autocorrelation has the following property:
\[ R_{x}(k + mP, l + rP) = R_{x}(k, l + (r - m)P). \]  \hfill (56)

In this case, the signal sections \( x_n(k) = x(q + kP) \) are jointly WSS processes with autocorrelation \( R_{x}(k) = R_{x}(q + kP) \). The coefficients of the MWT at each scale level are (individually) WSS processes with variances
\[ \sigma_{n,q}^2 = \sum_{r} R_{x}(r) \rho_{\phi_{n,0}}(r), \quad n = 1, 2, \ldots, N, \]  \hfill (57)
and
\[ \sigma_{n,q+1}^2 = \sum_{r} R_{x}(r) \rho_{\phi_{N,0}}(r). \]  \hfill (58)

The energy preservation property of the ordinary wavelet sets, which stems from their definition in terms of iterated QMF’s, yields
\[ \sum_{n=1}^{N} 2^{-n} \rho_{\phi_{n,0}}(r) + 2^{-N} \rho_{\phi_{N,0}}(r) = \delta_{r,0}. \]  \hfill (59)

From (57)–(59), we conclude that
\[ \sigma_{n,q}^2 = \frac{1}{P} \sum_{q=0}^{P-1} \sigma_{x}^2 = \frac{1}{P} \sum_{q=0}^{P-1} \left( \sum_{n=1}^{N} 2^{-n} \sigma_{x,n,q}^2 + 2^{-N} \sigma_{x,N+1,q}^2 \right). \]  \hfill (60)

Thus, the arithmetic-geometric mean inequality applied to (51) yields \( G_{MMWT} \geq 1 \). Moreover, from (60), (26), (50) and the same inequality, we obtain \( G_{MMWT} \geq 1 \).

If the input is a mean-square periodic process, i.e., if its autocorrelation is doubly periodic, \( R_{x}(n + mP, r + kP) = R_{x}(n, r) \), then the coding gains of the MWT and MMWT coders, as given in (50) and (51), become infinite. In fact, we have
\[ \sigma_{n,q}^2 = \sigma_{x}^2 \left( \sum_{k} \psi_{n,0}(k) \right)^2, \quad n = 1, 2, \ldots, N \]  \hfill (61)

and
\[ \sigma_{n,q+1}^2 = \sigma_{x}^2 \left( \sum_{k} \phi_{N,0}(k) \right)^2. \]  \hfill (62)
The regularity condition [2], expressed in the frequency domain, i.e., \( \psi_{n,0}^{}(0) = 0 \), \( n = 1, 2, \cdots, N \), and \( \hat{\Psi}_{N,0}^{}(0) = 2^{N/2} \), together with (61) and (62) imply that \( \sigma_{X_{n}}^2 = 0 \), and \( \sigma_{X_{N+1,q}}^2 = 2^N \sigma_{X_{1}}^2 \). From (42) and (45), we conclude that \( D_{\text{MMWT}} = D_{\text{MMWT}} = 0 \) and the coding gains become arbitrarily large. This means that at equal data rates, the error variances of the MWT and MMWT coders are infinitely smaller than that of the PCM coder. However, as explained in Appendix A, the error models assumed in (24) and (25) is not valid in the ideal case of zero variance input processes, and our coding gain estimates become unreliable. Nevertheless, it is clear that if all the wavelet coefficient channels have zero variance, one needs only to transmit their average values (side-information) and the scaling coefficients. The latter can be transmitted at an arbitrarily small rate by increasing \( N \) at the cost of arbitrarily increasing the coding delay. Large coding gains may be achieved if the small-scale coefficient variances are small, e.g., if the input process is close to a mean-square periodic process.

IV. PERFORMANCE OF MWT AND MMWT CODERS

In this section, we compare the MWT and MMWT coders with other transform coders such as block-DCT, KLT, and WT. The coding gain is a measure of the energy compaction of the transform and a useful criterion for evaluating the performance of the coders. Other factors such as complexity, coding delay, and perceptual results must be taken into account. The comparison will be carried out using both real-life signals and source models. In order to simplify the analysis, we will consider signals possessing a stable and known pitch such as those recorded from musical instruments. Time-varying pitch could be taken into account by considering the PSWT scheme. A feature of the MWT is that pitch adaptation does not require updating the analysis or synthesis filterbanks since the only parameter to be updated is the multiplexing rate. However, the evaluation of the characteristics of the PSWT coder falls beyond the scope of this paper, and the reader is referred to [6] for experimental results. Synthetic test signals can be generated by means of AR-comb models representing an approximation to voiced speech and music signals. We will compare the coding gains of the coders when the model parameters are varied.

A. Complexity and Coding Delay

The complexities of MWT, MMWT, and WT are the same, and they are equal to \( 2C(1 - 2^{-N}) \), where \( N \) is the number of scale levels [9], [3], [6]. The factor \( C \) is the complexity of a two-band orthogonal QMF bank and is proportional to \( L \) operations per input sample, where \( L \) is the degree of the transfer function. As we can see, the total complexity is roughly independent of \( N \). The complexity of the MWT implemented with short impulse response QMF’s is comparable with that of the block-DCT with large block length. In fact, the number of operations per sample of the block-DCT with block length \( B \) is proportional to \( \log(B) \). For example, the MWT with degree \( L = 5 \) QMF and a DCT with \( B = 1024 \) have about the same complexity. The complexity of the block-KLT with precomputed basis \( B \) operations per sample. Thus, a KLT with block length \( B = 10 \) requires roughly the same operations as an \( L = 5 \) MWT. In Fig. 3, we report the coding gains of the block-KLT (up to \( B = 256 \)) and block-DCT coder as a function of the block length. It is well-known that the KLT yields an optimum energy compaction versus block length characteristic among the class of finite block-length transforms. However, it has a higher complexity. In Fig. 4, we plot the coding gain of the MWT coder as a function of the degree of Daubechies’ flat QMF transfer functions [2]. The test signal was recorded from a violin and sampled at 20 kHz. Similar results were obtained with other single-instrument audio sources. The coding gains of DCT and MWT are approximately the same. For the same complexity, the coding gain of the WT subband coder is much lower (\( G_{\text{WT}} \approx 5.5 \text{ dB} \)). Notice that at short block lengths, large values of \( G_{\text{DCT}} \) may only be obtained if the block length is set to an integer multiple of the signal period. Pitch adaptation is not very efficient in the DCT coder since we need to implement prime factor algorithms. The coding gain of the MWT coder can be increased by optimizing the QMF transfer functions. Under orthogonality and fixed degree constraints, we attempted to maximize the coding gain. The degree \( L = 5 \) QMF frequency response resulting from this optimization conducted on the violin sound is shown in Fig. 5. Notice that the optimal transfer function has a steeper transition band than Daubechies’ QMF and, although we did not apply the lowpass constraint, we obtained it as a result. However, the resulting coding gain was only 0.5 dB larger.

The coding delay of an \( N \)-scale WT with degree \( L \) QMF’s is \( D_{\text{WT}} = (2^N - 1)L \) samples. The coding delays of MWT and MMWT with multiplexing rate \( 1/P \) are \( P \) times larger: \( D_{\text{MMWT}} = D_{\text{MMWT}} = (2^N - 1)LP \) samples. The delay of the DCT is equal to the block-length \( D_{\text{DCT}} = B \). For example, with \( L = 5, N = 4, B = 1024 \), and a sampling rate of 20 kHz, a signal with fundamental frequency of 300 Hz would encounter the following delays: \( D_{\text{DCT}} = 50 \text{ ms}, D_{\text{MMWT}} = D_{\text{MMWT}} = 250 \text{ ms}, \text{and } D_{\text{WT}} = 4 \text{ ms} \). The delay of the MWT coder linearly decreases as the pitch increases.

B. AR Comb Models

In order to evaluate the performance of the MWT and MMWT on classes of signals, we suggest two source models,
namely, the broadband and lowpass AR-comb models. The broadband AR(1)-comb model with period $P$ is defined as
\[ y(k) = \rho_P y(k - P) + w(k) \quad 0 \leq \rho_P < 1 \] (63)
where $w(k)$ is WSS zero-mean white Gaussian noise, and $\rho_P$ is the lag $P$ correlation coefficient. This coefficient represents the correlation between source samples delayed $P$ samples apart. Its value is close to unity for pseudo-periodic signals such as voiced sounds in speech and musical tones. The lowpass AR(1, 1)-comb model is obtained by filtering the AR(1)-comb signal (63) through a single pole lowpass filter
\[ s(k) = \rho_1 s(k - 1) + y(k) \quad 0 \leq \rho_1 < 1. \] (64)
Here, $\rho_1$ is the lag 1 correlation coefficient after removal of the spectrum pole $\rho_P$, and it controls the bandwidth of the process.

In Fig. 6, we report the coding gains of the DCT coder with block-length $B = 1024$ and of the MWT with degree $L = 5$ QMF’s and $N = 1, 2, 3$, all with the same broadband AR(1)-comb input. In the same figure, the coding gains of equally complex KLT and WT could not be reported since they are negligibly small. Since the input signal is a WSS process, the coding gain of the MWT is identical to that of the MMWT, as remarked in the previous section. Notice that the coding gain of the MWT/MMWT is larger than that of the other coders if $N > 1$.

In order to evaluate the effect of the signal bandwidth, consider the AR(1, 1) model with $\rho_P = 0.99$. In Fig. 7, we report a plot of the coding gains of equally complex KLT, WT, MWT, and DCT coders as a function of $\rho_1$. Notice that
for $\rho_1 > 0.8$, the coding gain of the DCT becomes larger than that of the $N = 4$ MWT, whereas the coding gains of both WT and KLT are considerably smaller up to $\rho_1 \approx 0.95$. A feature of the MWT is that the coding gain is independent of the bandwidth of the signal. Thus, the MWT coder performs better than the DCT on broadband signals.

In actual implementations, each quantizer rate must be approximated by an integer number of bits per sample. This may be achieved by rounding the optimum bit-rates. The resulting bit allocation scheme is no longer optimal, and the SNR may be degraded. In Tables I and II, we report on the results of simulations comparing the SNR’s and bit rates of actual MWT, DCT, and PCM coders when the inputs are AR(1, 1) processes with $\rho_1 = 0.7$ and $\rho_1 = 0.3$, respectively, and $\rho_P = 0.99$. The performance of the MWT coder is superior to that of the DCT in terms of both SNR and bit rate when the input is broadband (Table II).

Exact pitch determination is a critical issue to obtain good compression rates or large coding gains with the MWT/MMWT coders. The assumption is that the input is pseudo-periodic with a well-defined fundamental frequency. Multiphonic sounds, like those recorded form an orchestra, cannot be advantageously coded unless we foresee a more complex scheme in which several coders are cascaded, where each pertains to a pseudo-harmonic series. The MWT/MMWT coding gain over aperiodic signals is identical to that of the WT coder.

Independent of its accuracy, the pitch period must be rounded to an integer number of samples in order to compute the multiplexed transforms. In the violin tone example of Fig. 3, the average pitch period of the signal amounted to 64.85 samples, and the coder was tuned to a rounded period of 65 samples. If the pitch of the input signal is constant, it can be accurately estimated so that the largest error is due to rounding. This assumption was comforted by experiments with autocorrelation-based pitch-detection algorithms. In order to test the performance of the coder over pitch mismatch, a set of variable pitch input signals was generated by downsampling by a factor 10 an AR(1, 1) process with $\rho_P = 0.99, \rho_1 = 0.7$, and $P$ ranging from 635 to 645 samples. The effect of pitch rounding on the coding gain of the MWT coder is shown in Fig. 8. A significant drop of gain occurs if the magnitude of the pitch round-off error is larger than 0.25 samples. Pitch misdetection occurring in transients and at low SNR’s may introduce larger gain drops. However, it is reasonable to assume that pitch detection is unreliable only in short time intervals, having little effect on the overall performance. As already remarked, if the pitch is not constant, a pitch-adaptive PSWT-based coding scheme should be used in place of MWT.

### Table I

<table>
<thead>
<tr>
<th>Rate (b.p.s.)</th>
<th>SNR (dB)</th>
<th>Rate (b.p.s.)</th>
<th>SNR (dB)</th>
<th>Rate (b.p.s.)</th>
<th>SNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.93</td>
<td>27.55</td>
<td>3.99</td>
<td>26.54</td>
<td>4</td>
<td>13.61</td>
</tr>
<tr>
<td>4.93</td>
<td>33.15</td>
<td>4.99</td>
<td>32.52</td>
<td>5</td>
<td>19.26</td>
</tr>
<tr>
<td>5.93</td>
<td>39.15</td>
<td>5.99</td>
<td>38.43</td>
<td>6</td>
<td>25.24</td>
</tr>
<tr>
<td>6.93</td>
<td>45.04</td>
<td>6.99</td>
<td>44.41</td>
<td>7</td>
<td>31.02</td>
</tr>
<tr>
<td>7.93</td>
<td>51.12</td>
<td>7.99</td>
<td>50.46</td>
<td>8</td>
<td>36.96</td>
</tr>
<tr>
<td>8.93</td>
<td>57.18</td>
<td>8.99</td>
<td>56.55</td>
<td>9</td>
<td>43.18</td>
</tr>
<tr>
<td>9.93</td>
<td>63.11</td>
<td>9.99</td>
<td>62.44</td>
<td>10</td>
<td>49.18</td>
</tr>
<tr>
<td>10.93</td>
<td>69.11</td>
<td>10.99</td>
<td>68.58</td>
<td>11</td>
<td>55.11</td>
</tr>
<tr>
<td>11.93</td>
<td>75.22</td>
<td>11.99</td>
<td>74.47</td>
<td>12</td>
<td>61.15</td>
</tr>
<tr>
<td>12.93</td>
<td>81.13</td>
<td>12.99</td>
<td>80.51</td>
<td>13</td>
<td>67.23</td>
</tr>
<tr>
<td>13.93</td>
<td>87.24</td>
<td>13.99</td>
<td>86.48</td>
<td>14</td>
<td>73.22</td>
</tr>
<tr>
<td>14.93</td>
<td>93.21</td>
<td>14.99</td>
<td>92.70</td>
<td>15</td>
<td>79.21</td>
</tr>
<tr>
<td>15.93</td>
<td>99.27</td>
<td>15.99</td>
<td>98.55</td>
<td>16</td>
<td>85.24</td>
</tr>
</tbody>
</table>

**V. Conclusion**

In this paper, we have examined the performance of coders based on the multiplexed wavelet transform. We derived...
TABLE II
RATE AND SNR OF MWT, DCT AND PCM CODERS (P = 0.99 AND 1 = 0.3)

<table>
<thead>
<tr>
<th>Rate (b.p.s.)</th>
<th>SNR (dB)</th>
<th>Rate (b.p.s.)</th>
<th>SNR (dB)</th>
<th>Rate (b.p.s.)</th>
<th>SNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.90</td>
<td>27.61</td>
<td>4.01</td>
<td>24.17</td>
<td>4.00</td>
<td>13.57</td>
</tr>
<tr>
<td>4.90</td>
<td>33.34</td>
<td>5.01</td>
<td>29.98</td>
<td>5.00</td>
<td>19.15</td>
</tr>
<tr>
<td>5.90</td>
<td>39.33</td>
<td>6.01</td>
<td>35.99</td>
<td>6.00</td>
<td>24.74</td>
</tr>
<tr>
<td>6.90</td>
<td>45.37</td>
<td>7.01</td>
<td>42.01</td>
<td>7.00</td>
<td>31.17</td>
</tr>
<tr>
<td>7.90</td>
<td>51.34</td>
<td>8.01</td>
<td>48.03</td>
<td>8.00</td>
<td>37.12</td>
</tr>
<tr>
<td>8.90</td>
<td>57.27</td>
<td>9.01</td>
<td>54.13</td>
<td>9.00</td>
<td>43.17</td>
</tr>
<tr>
<td>9.90</td>
<td>63.50</td>
<td>10.01</td>
<td>60.09</td>
<td>10.00</td>
<td>49.10</td>
</tr>
<tr>
<td>10.90</td>
<td>69.42</td>
<td>11.01</td>
<td>66.11</td>
<td>11.00</td>
<td>55.12</td>
</tr>
<tr>
<td>11.90</td>
<td>75.42</td>
<td>12.01</td>
<td>72.12</td>
<td>12.00</td>
<td>61.21</td>
</tr>
<tr>
<td>12.90</td>
<td>81.43</td>
<td>13.01</td>
<td>78.24</td>
<td>13.00</td>
<td>67.18</td>
</tr>
<tr>
<td>13.90</td>
<td>87.54</td>
<td>14.01</td>
<td>84.14</td>
<td>14.00</td>
<td>73.26</td>
</tr>
<tr>
<td>14.90</td>
<td>93.52</td>
<td>15.01</td>
<td>90.20</td>
<td>15.00</td>
<td>79.14</td>
</tr>
<tr>
<td>15.90</td>
<td>99.54</td>
<td>16.01</td>
<td>96.15</td>
<td>16.00</td>
<td>85.17</td>
</tr>
</tbody>
</table>

expressions for the optimal bit allocation schemes and the coding gains of this transform and its modified form—the MMWT—in the assumption that the analysis coefficients are uniformly quantized. We used these results to show, by means of experiments on real-life signals and AR-comb models, that one can achieve larger coding gains than those attained by the WT and DCT coders. This is particularly true when the input is a broadband pseudo-periodic signal. We attempted a design of the QMF's transfer functions by maximizing the coding gain and obtaining only a slight improvement over maximally flat QMF's. The major drawback of the MWT coder is its long coding delay, which is also pitch dependent. Unlike block-DCT, pitch adaptation of the MWT transform can be efficiently performed as in the PSWT. We did not address the issue of extending our results on coding gain to allow for time-varying pitch, which seems to be a formidable task. However, even on constant pitch inputs, exact pitch estimation and period roundoff are critical issues for the correct functioning of the MWT coder. Another open issue is the optimal design of biorhogonal MWT transforms.

CODERS BASED ON THE MULTIPLEXED WAVELET TRANSFORM MAY OUTPERFORM CONVENTIONAL CODERS, PROVIDED THAT THE INPUT SIGNAL IS BROADBAND PSEUDO-PERIODIC—AS IN THE CASE OF THE LPC EXCITATION SIGNAL—and the coding delay is not a critical factor.

APPENDIX A

Throughout this paper, the uniform quantization noise is modeled as a WSS process whose variance is proportional to the average variance of the input to the quantizer. This assumption, which is valid only for a sufficient number of bits and in the hypothesis of no overflow, needs to be justified.

It is shown in [11] that if $x$ is a random process, which we uniformly quantize with quantization step $s$, the resulting output error variance is

$$\sigma^2_x = \sum_m E_m \Phi_x \left( \frac{2\pi m}{s} \right)$$

where $E_0 = \frac{x^2}{12}$, $E_m = (-1)^m \frac{x^2}{2^{2m+1}}$ for $m \neq 0$, and $\Phi_x(\omega) = E(e^{j\omega x})$ is the characteristic function of $x$. Notice that unless $x$ is stationary, the characteristic function is time varying, and so is the output error variance. However, if the quantization step is sufficiently small and if $\Phi_x(\omega)$ decays as $\omega$ grows, which is true for a smooth p.d.f., the approximation $\sigma^2_x = \frac{x^2}{12}$ is fairly good. In the same approximation, the output error mean is zero, and the error process is closely WSS. The above assumption is not valid if, for example, the input is a constant, zero variance process, since its p.d.f. is a Dirac delta and the characteristic function is not decaying.

If $R$ is the number of bits and $\Delta$ the dynamic range of the quantizer, then for a two's complement quantizer, we have $s = \frac{\Delta}{2^R-1}$. In order to limit the overflow probability, we should [7] let the dynamic range be proportional to the standard deviation of the input to the quantizer: $\Delta = a\sigma_x$. In case the input process is not WSS and we desire to keep the quantization step constant, we let $\Delta$ to be proportional to the average standard deviation of the input $\sigma_x$, obtaining

$$\sigma^2_x = \frac{a^2 \sigma^2_x 2^{2(R-1)}}{12} = \varepsilon^2 2^{-R} \sigma^2_x.$$  \hspace{1cm} (A-1)

The constant $\varepsilon^2 = \frac{a^2}{3}$ controls the overflow probability, which, in turn, depends on the input statistics.

APPENDIX B

We need to prove Lemma 1 of Section III, stating that in an orthogonal expansion the coefficient cross-correlation terms do not contribute to the total average variance.

By substituting in the average variance definition (23) into the expansion of the process $x(k) = \sum_n x_n f_n(k)$ over the
orthogonal set \( \{ f_n(k) \}_{n \in \mathbb{N}} \), we obtain
\[
\hat{\sigma}_a^2 = \lim_{L \to \infty} \frac{1}{L} \sum_{k=-L}^{L-1} \sum_{n, m} E \{ X_n X_m^* \} f_n(k) f_m^*(k). \tag{B-1}
\]
In view of the orthogonality of the set \( \{ f_n(k) \}_{n \in \mathbb{N}} \), we have
\[
\lim_{L \to \infty} \frac{1}{L} \sum_{k=-L}^{L-1} f_n(k) f_m^*(k) = 0 \quad \text{when } n \neq m.
\]
Then, a fortiori, we obtain
\[
\lim_{L \to \infty} \frac{1}{L} \sum_{k=-L}^{L-1} \sum_{n, m} E \{ X_n X_m^* \} f_n(k) f_m^*(k) = 0. \tag{B-2}
\]
Finally, from (B-1) and (B-2), we derive (28).

**APPENDIX C**

In order to obtain the optimum bit allocation, one needs to minimize the average total error variance (39) subjected to the constant rate constraint (40). To this purpose, let \( \beta_n = 2^{-n}, n = 1, \cdots, N, \) and \( \beta_{N+1} = 2^{-N} \) so that (39) and (40), respectively, may be put in the following form:
\[
\sigma_a^2 = \sum_{n=1}^{N+1} \beta_n a_n, \tag{C-1}
\]
and
\[
R = \sum_{n=1}^{N+1} \beta_n r_n. \tag{C-2}
\]

First, we will examine the optimum bit allocation of the MWT coder for which \( r_n = \frac{1}{P} \sum_{q=0}^{P-1} R_{n,q} \) and \( a_n = \frac{1}{P} \sum_{q=0}^{P-1} c_{n,q}^2 \). According to the quantization noise model (24), we can let \( \sigma_{a,n}^2 = \gamma_{n,q}^2 2^{-2R_{n,q}} \), where \( \gamma_{n,q} = \varepsilon_2 \sigma_{X_n^q}^2 \), therefore, (C-1) becomes
\[
\sigma_a^2 = \frac{1}{P} \sum_{n=1}^{N+1} \beta_n \sum_{q=0}^{P-1} \gamma_{n,q}^2 2^{-2R_{n,q}}. \tag{C-3}
\]

To solve the constrained minimization problem, we will apply the Lagrange multiplier method. To this purpose, form the auxiliary function
\[
F(R_{1,0}, \cdots, R_{1,P-1}, \cdots, R_{N+1, P-1}, \lambda) = \frac{1}{P} \sum_{n=1}^{N+1} \beta_n \sum_{q=0}^{P-1} \gamma_{n,q}^2 2^{-2R_{n,q}} + \lambda \left( \frac{1}{P} \sum_{n=1}^{N+1} \beta_n \sum_{q=0}^{P-1} R_{n,q} - R \right)
\]
where \( \lambda \) is the Lagrange multiplier. In order to minimize (C-3), it is necessary that all partial derivatives of \( F \) be zero
\[
\begin{cases}
\frac{\partial F}{\partial R_{n,q}} = 0, & n = 1, 2, \cdots, N + 1; q = 0, 1, \cdots, P - 1 \\
\frac{\partial F}{\partial \lambda} = 0
\end{cases} \tag{C-4}
\]
the last equation being equivalent to the constraint. Solution of the first \((N+1)P \) equations of (C-4) yields an expression of the rate parameters in terms of \( \lambda \)
\[
R_{n,q} = -\frac{1}{2} \log_2 \frac{\lambda}{2^{\gamma_{n,q}}}.
\tag{C-5}
\]
In order to determine \( \lambda \), we substitute (C-5) in the constant rate constraint (C-2) and solve for \( \lambda \), observing that \( \sum_{n=1}^{N+1} \beta_n = 1 \). We obtain
\[
\lambda = 2^{-2R+1} D, \tag{C-6}
\]
and
\[
D = \prod_{n=1}^{N+1} \prod_{q=0}^{P-1} c_{n,q}^2 \tag{C-7}
\]
Finally, substitution of (C-6) in (C-5) yields
\[
R_{n,q} = R + \frac{1}{2} \log_2 \frac{\gamma_{n,q}}{D}. \tag{C-8}
\]
Recalling that \( \gamma_{n,q} = \varepsilon_2 \sigma_{X_n^q}^2 \), (C-7) and (C-8) are equivalent to (45) and (46), with \( D = 2^P D_{MMWT} \), and (44), respectively. Furthermore, by inverting (C-5), we obtain (47).

In order to verify that the average variance of the total error actually attains a minimum at the rates indicated in (C-8), observe that the arithmetic-geometric mean inequality applied to (C-1) yields
\[
\sigma_a^2 = \frac{1}{2^N} \sum_{n=1}^{N+1} \gamma_{n,q} 2^{-2R_{n,q}} \geq \gamma_{N+1} \prod_{n=1}^{N} \sigma_{n,q}^2
\]
with equality only if all the \( a_n \) are identical. Furthermore, recalling the definition of \( a_n \) and, again, applying the arithmetic-geometric mean inequality, we obtain
\[
a_n = \frac{1}{P} \sum_{q=0}^{P-1} \sigma_{n,q}^2 \geq \prod_{q=0}^{P-1} \left( \sigma_{n,q}^2 \right)^{1/P} \tag{C-9}
\]
with equality iff, for any fixed \( n \), all \( \sigma_{n,q}^2 \) are identical. Combining the last two inequalities, we have
\[
\sigma_a^2 \geq \prod_{q=0}^{P-1} \left( \frac{1}{P} \sum_{n=1}^{N+1} \left( \sigma_{N+1, q}^2 \right)^{1/2^N P} \right)^{1/2^N P} \prod_{n=1}^{N} \prod_{q=0}^{P-1} \sigma_{n,q}^2 \tag{C-9}
\]
By substituting \( \sigma_{n,q}^2 = \gamma_{n,q} 2^{-2R_{n,q}} \) and applying the constant rate constraint, the last inequality simplifies to
\[
\sigma_a^2 \geq 2^{-2R} D \tag{C-9}
\]
showing that the right-hand side is a constant once the input process and the rate are fixed. The equality sign holds in (C-9) if and only if all \( \sigma_{n,q}^2 \) are identical, which, as shown in (47), is verified by the solution found. Thus, the quantizer rates given in (C-8) actually provide the optimum bit allocation scheme for the MWT coder.

The optimum bit allocation scheme for the MMWT may be found using a similar procedure. However, note that it may be derived from (C-8) as a particular case in which \( R_{n,q} = R_n \) and \( \gamma_{n,q} = \gamma_n = \varepsilon_2 \sigma_{X_n}^2 \), independently of \( q \).
REFERENCES


Gianpaolo Evangelista (M’92) received the laurea in physics, summa cum laude, from the University of Naples in 1984 and the M.Sc. and Ph.D. degrees in electrical engineering from the University of California, Irvine, in 1987 and 1990, respectively. During 1985–1986, he worked at the Centre d’Études de Mathématique et Acoustique Musicales (CEMAMu/CNET), Paris, where he contributed to the development of a DSP-based sound synthesis system, and from 1991 to 1994, he was a research engineer at the Microgravity Advanced Research and Support (MARS) Center, Naples, where he was engaged in research in image processing applied to fluid motion analysis. He is currently a researcher with the Department of Physical Sciences, Federico II University of Naples, Italy. His interests include speech, music, and image processing, coding, and multirate signal processing.

Dr. Evangelista was a recipient of the Fulbright Fellowship.