such an approach is local. We can choose the interpolation points nonuniformly by assigning more interpolation points around irregular locations. The disadvantage of ISS is that the generated curves are not polynomial splines.

V. CONCLUSION

In this correspondence, we propose the ISS approach for the initialization of a wavelet transform, which is very important for further applications. Moreover, the proposed method also provides an efficient evaluation of inner products of wavelets and their derivatives, which is essential in the wavelet–Galerkin method. We formulate a general procedure for the computation of exact derivative values of the interpolatory fundamental function at dyadic points. The error analysis shows that the method is exact for certain polynomials. The numerical experiments have shown the high accuracy and efficiency of the method.

REFERENCES


Abstract—We propose a bootstrap version of a bispectrum-based test for departure from Gaussianity that achieves high power while maintaining the level of significance, even for small sample sizes. The proposed procedure can be also used to set confidence bands for a measure of the bicoherence of stationary random signals.

Index Terms—Bispectrum, bootstrap, Gaussianity test.

I. INTRODUCTION

Tests for departure from Gaussianity of stationary signals are of practical importance in many disciplines such as biomedicine, radar, sonar, and structural vibration analysis [1].

Let \( X_t \) be a real-valued stationary stochastic process with an unspecified distribution function \( F_X \). We wish to test whether \( F_X \) is Gaussian based on observations \( X_0, \ldots, X_T \). In the case where the data constitutes a set of independent and identically distributed (i.i.d.) random variables, we could use the classical goodness-of-fit tests based on the \( \chi^2 \) or the Kolmogorov–Smirnov statistic. More powerful tests include the Shapiro–Wilk test, the D’Agostino test, and tests based on the characteristic function [2]. Problems arise, however, when the data is correlated. Most tests proposed for testing departure from Gaussianity for correlated data are based on higher order cumulants or polyspectra.

Gasser [3] suggested a test for departure from Gaussianity based on normalized sample measures of skewness and kurtosis. This is a simple test, but its performance suffers as a result. An improvement of Gasser’s test has been suggested by Giannakis and Tsatsanis [4]. The authors propose an approach based on testing for zero the third- and fourth-order cumulants. The tests differ from Gasser’s in that the authors utilize more than the zeroth lag of the cumulants. A limitation of the method is that the derived distribution of the sample cumulants does not hold unless the number of data points is large.

Hinich has described a procedure based on the bispectrum to test for departure from Gaussianity of a stationary stochastic process [5]. This test follows from the work of Subba Rao and Gabr [6]. Hinich’s test has been used in many applications to decide whether a stationary process is Gaussian, although the test is not able to detect symmetric alternatives. A more appropriate procedure to test for departure from Gaussianity of a stationary stochastic process is based on testing the trispectrum as well as the bispectrum for zero [7]. Accurate estimation of the trispectrum generally requires large sample sizes and is not feasible in many practical situations.

In this correspondence, we design a new test by incorporating the bootstrap [8], [9] into a bispectrum-based test and demonstrate its superior performance compared with Subba Rao and Gabr’s test.

Manuscript received April 15, 1997; revised July 13, 1998. The associate editor coordinating the review of this paper and approving it for publication was Dr. Phillip A. Regalia.

A. M. Zoubir is with the Australia Telecommunications Research Institute and the School of Electrical and Computer Engineering, Curtin University of Technology, Perth, Australia.

D. R. Iskander is with the Center for Eye Research, Queensland University of Technology, Brisbane, Australia.

Publisher Item Identifier S 1053-587X(99)01355-0.
This correspondence is organized as follows. In the next section, we propose our bispectrum based test for departure from Gaussianity using a bootstrap procedure. In Section III, simulation results are presented, and Section IV follows with a conclusion.

II. A Bootstrap Based Approach for Testing for Departure from Gaussianity

A. Testing for Departure from Gaussianity Using the Bispectrum

Let \( X_0, \cdots, X_{T-1} \) be observations from a strictly stationary real-valued time-series \( X_t, t \in \mathbb{Z} \) with mean zero, finite variance, and bispectrum \( C_{X_XX}(\omega_j, \omega_k) = \pi, \omega_j, \omega_k \leq \pi \). For a Gaussian process (as well as any other symmetric process), the bispectrum is identically zero. Thus, testing whether the bispectrum is zero may be considered to be testing for departure from Gaussianity. Rejection of the hypothesis implies that the stationary process is non-Gaussian; otherwise, the process may be non-Gaussian, but the data is consistent with a zero bispectrum. Often, such tests are referred to as Gaussianity tests, although they cannot detect symmetric alternatives \([5], [6]\). Nevertheless, bispectrum-based tests have found extensive use among signal processing practitioners \([10]\).

To test departure from Gaussianity, we would test \( C_{X_XX}(\omega_j, \omega_k) = 0 \), when \( \omega_j \) and \( \omega_k \) are restricted to \( 0 \leq \omega_k \leq \omega_j, \omega_k + 2\omega_j \leq 2\pi \). Taking points outside this region is unnecessary due to the symmetry of the bispectrum.

B. Estimation of the Bispectrum

Let

\[
I_{X_XX}(\omega_j, \omega_k) = \frac{1}{T} d_X(\omega_j)d_X(\omega_k)d_X(\omega_j + \omega_k) - \pi \leq \omega_j, \omega_k \leq \pi
\]

(1)

denote the biperiodogram of the sample, where

\[
d_X(\omega_i) = \sum_{t=0}^{T-1} X_t \exp \{-j\omega_i t\}, \quad -\pi \leq \omega_i \leq \pi
\]

(2)

and \( d_X \) is the complex conjugate of \( d_X \). An estimate of \( C_{X_XX}(\omega_j, \omega_k) \) can be obtained by smoothing the biperiodogram with a two-dimensional (2-D) window \( W(\omega_j, \omega_k) \) as

\[
\hat{C}_{X_XX}(\omega_j, \omega_k) = \frac{1}{T} \sum_{k_1=1}^{[T/2]} \sum_{k_2=1}^{[T/2]} \left[ \frac{1}{T} \right] W(\omega_j - 2\pi k_1, \omega_k - 2\pi k_2) \times I_{X_XX}(\omega_j - 2\pi k_1, \omega_k - 2\pi k_2)
\]

(3)

where \([::] \) is the floor operator. Alternatively, we can estimate \( C_{X_XX}(\omega_j, \omega_k) \) by dividing the \( T \) observations into \( P \) nonoverlapping segments of \( N \) consecutive measurements, calculating the biperiodogram \( I_{X_XX}^{(i)}(\omega_j, \omega_k) \) as in (1) for each segment \( i = 1, \cdots, P \) and averaging to obtain

\[
\hat{C}_{X_XX}(\omega_j, \omega_k) = \frac{1}{P} \sum_{i=1}^{P} I_{X_XX}^{(i)}(\omega_j, \omega_k).
\]

(4)

It is known that under mild regularity conditions and large \( T \), the estimated bispectral density \( \hat{C}_{X_XX}(\omega_j, \omega_k) \), when \( (\omega_j, \omega_k) \) is restricted to the region \( 0 < \omega_k \leq \omega_j, \omega_k + 2\omega_j \leq 2\pi \), is approximately distributed as a complex normal variable. If \( \omega_j = 0 \), the estimate is approximately distributed as a real normal variable. Similar results hold for the estimator in (4) (see, for example, \([5]\)).

Validity of these asymptotic results is, however, questionable when the sample size is small. In the next section, we develop a test that outperforms existing bispectrum-based tests even for small sample sizes.

C. A Bootstrap-Based Procedure for Testing Departure from Gaussianity

Let \( X_0, \cdots, X_{T-1} \) be divided into \( P \) nonoverlapping segments of \( N \) consecutive measurements. We will consider estimating the bispectrum at discrete frequencies \( \omega_j = 2\pi j/N \) and \( \omega_k = 2\pi k/N \) for each segment and will consider the triangular grid (let \( N \) be even)

\[
\mathcal{D} = \{0 < k < j, 2j + k \leq N\}
\]

This region does not include the boundary \( k = 0 \) to reduce the computational burden of the bootstrap procedure by avoiding separate resampling schemes for complex and real data.

Our procedure is based on the interpretation of the bispectral estimation problem as the approximate regression

\[
I_{X_XX}^{(i)}(\omega_j, \omega_k) = C_{X_XX}(\omega_j, \omega_k) + \varepsilon_{j,k}\sqrt{V(\omega_j, \omega_k)}
\]

(5)

where

\[
V(\omega_j, \omega_k) = N\sigma_{X_XX}(\omega_j)C_{X_XX}(\omega_j, \omega_k)C_{X_XX}(\omega_j + \omega_k)
\]

\[
\times \left[ 1 + 2\delta(j - k) + \delta(N - 2j - k) + 4\delta(N - 3j)\delta(N - 3k) \right]
\]

(6)

\( C_{X_XX}(\omega) \) is the spectrum of the time-series \( X_t \), and \( \delta(k) \) is Kronecker’s delta function. This regression is consistent with asymptotic results \([11]\) that claim that for large \( N \), \( E I_{X_XX}^{(i)}(\omega_j, \omega_k) = C_{X_XX}(\omega_j, \omega_k) + O(N^{-1}) \) and \( \text{Var} I_{X_XX}^{(i)}(\omega_j, \omega_k) = V(\omega_j, \omega_k) + O(1) \) hold. We shall assume that the residuals \( \varepsilon_{j,k} \) are independent and identically distributed random variates for a reasonably large \( N \). Consequently, the mean and variance of \( \varepsilon_{j,k} \) are zero and one, respectively.

The steps for testing departure from Gaussianity of \( X_t \) are highlighted in Table I. Several steps in Table I and their motivation are explained in \([9]\). For example, the threshold of the bootstrap test is obtained as follows. From the collection \( \hat{C}_1, \cdots, \hat{C}_H \), we form the empirical distribution of the test statistic \( \hat{C} \) under the null hypothesis. Thus, by ranking the elements of this collection into increasing order, we select a threshold \( \hat{C}_{(i)} = [(B+1)/(1-\alpha)] \), where \( \alpha \) determines the level of the test.

The variance estimates \( \hat{\sigma}(\omega_j, \omega_k)^2 \) of \( \hat{C}_{X_XX}(\omega_j, \omega_k) \) and \( \hat{\sigma}^2(\omega_j, \omega_k)^2 \) of \( \hat{C}_{X_XX}(\omega_j, \omega_k) \) are constructed as follows. For each bootstrap biperiodogram \( \hat{I}_{X_XX}^{(i)}(\omega_j, \omega_k) \), we repeat steps 3–6 of the algorithm (nested bootstrap) a small number of times (e.g., \( B = 25 \)), replacing \( \hat{C}_{X_XX}(\omega_j, \omega_k) \) and \( \hat{I}_{X_XX}^{(i)}(\omega_j, \omega_k) \) with \( \hat{C}_{X_XX}(\omega_j, \omega_k) \) and \( \hat{I}_{X_XX}^{(i)}(\omega_j, \omega_k) \), respectively. Then, we estimate the variance of \( \hat{C}_{X_XX}(\omega_j, \omega_k) \) by

\[
\hat{\sigma}^2(\omega_j, \omega_k)^2 = \frac{1}{B_1} - \frac{1}{B_1} \sum_{b=1}^{B_1} \left( \hat{C}_{X_XX}(\omega_j, \omega_k) \right)^2 - \frac{1}{B_1} \sum_{b=1}^{B_1} \left( \hat{C}_{X_XX}^{(b)}(\omega_j, \omega_k) \right)^2
\]

(7)

where \( \hat{C}_{X_XX}^{(b)}(\omega_j, \omega_k), b = 1, \cdots, B_1 \) is a bootstrap version of \( \hat{C}_{X_XX}(\omega_j, \omega_k) \).

\( ^{1}\)We note that the magnitude of the correlation between biperiodograms \( I_{X_XX}^{(i)}(\omega_j, \omega_k) \) and \( I_{X_XX}^{(i)}(\omega_j, \omega_k) \) is approximately of order \( O(N^{-2}) \) for \( j \neq j' \) and \( k \neq k' \) (or “exclusive or” cases \([5]\)).
A variance estimate of \( \hat{\sigma}^2(\omega_j, \omega_k) \) is computed by running steps 4–6 a small number of times (say) to obtain \( \hat{\sigma}^2(\omega_j, \omega_k) \), and applying (7), where \( \hat{\sigma}^2(\omega_j, \omega_k) \) is replaced by \( \hat{\sigma}^2(\omega_j, \omega_k) \). For more details on the nested bootstrap, see [8].

The size of the data for which we can still assume i.i.d. (but at the same time, we cannot apply asymptotic results for testing Gaussianity) is of great importance. This issue could have been solved in a satisfactory manner only if there had existed a closed-form expression of the distribution of the bispectrum estimator for a finite sample. We have investigated this issue numerically and found, using simple procedures, that for 512 data points, the i.i.d. assumption of the residuals is reasonable. This is also reflected in the results of the test that maintains the level of significance (see Section III). Increasing the data size to larger values, say 1024 data points, would not warrant the use of the bootstrap due to computational complexity, but this may be the case where we can resort to asymptotic tests. We also note that the use of an independent data bootstrap method is by no means a limitation of the proposed test. If we knew the correlation structure of the residuals, we would be able to incorporate this into the test. Otherwise, we could use a general model for the errors such as a linear regression and resample the innovations.

### III. Simulation Results

This section demonstrates the performance results of the proposed test for departure from Gaussianity discussed in Section II-C. In simulations, three processes were considered: a white process, an autoregressive process of order five, and a moving-average process of order two. Specifically, we assumed

\[
X_t = Y_t, 
\]

and

\[
X_t = 0.5X_{t-1} - 0.6X_{t-2} + 0.3X_{t-3} - 0.4X_{t-4} + 0.2X_{t-5} + Y_t, 
\]

where \( Y_t \) is an independent random process with distribution \( \mathcal{N}(0, 1) \), which varies according to Tables II–V.

The estimated power of Subba Rao and Gabr’s test (with a square grid) are presented in Tables II and III for \( T = 256 \) and \( T = 512 \), respectively. The parameters of the test (see [6] for details) were set to \( M = 255, K = 7, d = 6, \) and \( r = 2 \) for \( T = 256 \); and \( M = 511, K = 9, d = 6, \) and \( r = 3 \) for \( T = 512 \). The estimated
TABLE IV
PERCENTAGE OF REJECTED HYPOTHESES FOR ZERO BISPECTRUM
AT 5% LEVEL OF SIGNIFICANCE USING BOOTSTRAP
BASED PROCEDURE FOR T = 256 AND N = 16

<table>
<thead>
<tr>
<th>Distribution</th>
<th>i.i.d.</th>
<th>AR(5)</th>
<th>MA(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N(0,1)</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>U(0,1)</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$\chi^2_2$</td>
<td>96</td>
<td>72</td>
<td>63</td>
</tr>
<tr>
<td>$\chi^2_4$</td>
<td>69</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>Laplace</td>
<td>16</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>K(1,1)</td>
<td>78</td>
<td>42</td>
<td>20</td>
</tr>
<tr>
<td>LogN</td>
<td>100</td>
<td>91</td>
<td>77</td>
</tr>
</tbody>
</table>

TABLE V
PERCENTAGE OF REJECTED HYPOTHESES FOR ZERO BISPECTRUM
AT 5% LEVEL OF SIGNIFICANCE USING BOOTSTRAP
BASED PROCEDURE FOR T = 512 AND N = 16

<table>
<thead>
<tr>
<th>Distribution</th>
<th>i.i.d.</th>
<th>AR(5)</th>
<th>MA(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N(0,1)</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>U(0,1)</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$\chi^2_2$</td>
<td>100</td>
<td>100</td>
<td>91</td>
</tr>
<tr>
<td>$\chi^2_4$</td>
<td>98</td>
<td>70</td>
<td>36</td>
</tr>
<tr>
<td>Laplace</td>
<td>13</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>K(1,1)</td>
<td>100</td>
<td>92</td>
<td>64</td>
</tr>
<tr>
<td>LogN</td>
<td>100</td>
<td>100</td>
<td>94</td>
</tr>
</tbody>
</table>

power of the bispectrum based test that utilizes the bootstrap are given in Tables IV and V for $T = 256$ ($N = 16$, $P = 16$) and $T = 512$ ($N = 16$, $P = 32$), respectively. In each case, the number of resamplings was set to $B = 199$ and $B_1 = 25$ for the main and nested bootstrap, respectively. The results were obtained using 100 independent replications for each distribution. The nominal level of significance was set to 5%.

1) Discussion: Tables II and III show that Subba Rao and Gabr’s test does not maintain the preset 5% level of significance for symmetric processes, e.g., for colored Gaussian or independent/colored Laplace process. Only in a few cases, the test maintains the nominal 5% level of significance, such as for independent and correlated Gaussian processes, independent uniform processes for $T = 256$, and independent and correlated $[AR(5)]$ processes for $T = 512$. The parameter setting we chose for Subba Rao and Gabr’s test corresponds to the case where the level of significance is maintained for the independent Gaussian process. This may not be the best approach to maximize the power of the test. However, the problem of choosing optimal parameters in the sense of maximizing the power for all alternatives (omnibus test) is beyond the scope of this correspondence.

On the other hand, the test based on the bootstrap maintains the nominal level of significance at below the preset 5%, except in the case of independent uniform and Laplace processes for $T = 512$. Tables II–V also show that the power of the test based on the bootstrap is comparable with that achieved using Subba Rao and Gabr’s test, but often higher, especially in the case of an independent process (second column in Tables II–V). This may be due to the fact that the assumption made about the independence of $e_j,k$ is more accurate in the case of an independent process. In the case of correlated data, we may choose to use another resampling procedure based on the smoothed biperiodogram of (3). In that case, we would resample the residuals centered around the coarse grid rather than all the residuals in the principle domain. Such a technique has been tested and found to achieve better results for correlated processes. However, it has two limitations. First, the computational burden of such a procedure is much greater than the one based on the averaged biperiodogram (4). Second, the choice of the parameters (e.g., grid size) is difficult to optimize.

In summary, the results show that the bootstrap procedure is able to maintain the nominal level of significance, whereas Subba Rao and Gabr’s test does not, at least for the parameters setting we chose. Even though a comparison of power is appropriate only if the tests maintain the nominal level of significance, we found that the test achieves power comparable with or better than Subba Rao and Gabr’s test.

2) Extension of Bootstrap Tests: The bootstrap procedure can also be applied to cumulant-based tests such as the ones proposed in [3] and [4]. This research is currently under study. Preliminary results showed that the tests are sensitive to error modeling and parameter choice of the so-called dependent data bootstrap procedures. Results will be presented elsewhere. Another interesting application of the bootstrap is with more recently developed tests based on the characteristic function [12], [13] (see also [2], [14]).

3) Confidence Bands: The bootstrap-based procedure presented may also be used to set confidence bands for higher order spectra or cumulants [15]. Specifically, repeating steps 4–6 a large number of times, we can obtain a total of $B$ bootstrap estimates of a measure of the bicoherence, as $\hat{C}_{XX}\{1\}(\omega_j, \omega_k) = \hat{C}_{XX}\{1\}(\omega_j)\hat{C}_{XX}\{1\}(\omega_k)$.

After sorting these estimates at each frequency pair $(\omega_j, \omega_k)$, we can determine the confidence bands using estimated percentiles obtained as in Step 9 of Table I.

IV. CONCLUSIONS

We have presented a bispectrum based test for departure from Gaussianity that can be used in situations where only a relatively small amount of data points is available. Simulation results have shown that the test achieves high power when compared with the test proposed by Subba Rao and Gabr while maintaining the level of significance. The choice of Subba Rao and Gabr’s test for comparison was purely arbitrary, and similar comparisons can be made with other tests, such as Hinich’s test. The proposed test can be used in situations where neither multiple realizations of the process nor information about the data correlation are available. The proposed procedure can also be used for setting confidence bands for a measure of the bicoherence.

ACKNOWLEDGMENT

The authors are grateful to H.-T. Ong for his helpful comments and for his assistance in the implementation of the tests.

REFERENCES

[6] T. Gabr's test, but often higher, especially in the case of an independent process. In the case of correlated data, we may choose to use another resampling procedure based on the smoothed biperiodogram of (3). In that case, we would resample the residuals centered around the coarse grid rather than all the residuals in the principle domain. Such a technique has been tested and found to achieve better results for correlated processes. However, it has two limitations. First, the computational burden of such a procedure is much greater than the one based on the averaged biperiodogram (4). Second, the choice of the parameters (e.g., grid size) is difficult to optimize.
Performance Analysis of Wavelets in Embedded Zerotree-Based Lossless Image Coding Schemes

V. N. Ramaswamy, N. Ranganathan, and K. R. Namuduri

Abstract—In this correspondence, we present a modification to the scanning approach in the set partitioning algorithm proposed by Said and Pearlman to exploit the correlation in a local neighborhood. The wavelet filters are characterized based on the wavelet coefficients obtained after the wavelet transform. Two new criteria are proposed for evaluating the performance of wavelets in lossless image coding applications: cumulative zerotree count and monotone spectral ordering of subbands produced after wavelet transform in a multiresolution scheme. Several wavelet filters are evaluated to test the evaluation criteria. The experimental results are presented to justify the proposed performance criteria.

Index Terms—Embedded zerotree wavelet, lossless image coding, performance analysis, subband coding.

I. INTRODUCTION

The general framework for embedded zerotree wavelet (EZW) image compression algorithm [5] consists of three stages, as shown in Fig. 1. In this framework, the selection of the wavelet filter is extremely critical. The choice of wavelet has an impact on the characteristics of the resulting wavelet coefficients. The characteristics of the wavelet coefficients affect the sizes of the significant map and the residue, which in turn has an impact on the compression efficiency at the final stage. We define compression efficiency as

\[ CE = \frac{\text{Orig. Image Size} - \text{Compressed Image Size}}{\text{Orig. Image Size}} \times 100. \]

What constitutes an appropriate wavelet filter for image coding? What are the properties that form the criteria to evaluate a wavelet filter for image coding? What effect does the evaluating procedure have on image coding? These are open issues that are currently being studied by researchers, and the topics still need further investigation. Gestenson et al. [6] evaluated wavelet filters for image compression applications by characterizing a filter bank according to its impulse response and step response in addition to regularity. Recently, Calderbank et al. [7] evaluated several filters for lossless image coding application and observed that there are other factors besides the number of analyzing vanishing moments that affect the compression efficiency [4], [9]. In this correspondence, we attempt to address these questions and present two new criteria for evaluating the performance of wavelets in lossless image compression application. The compression scheme is based on the set partitioning in hierarchical trees (SPIHT) algorithm proposed in [8]. We propose a modification to the scanning approach in the SPIHT and other EZW algorithms to exploit correlation in a local neighborhood. An analogy is drawn between our proposed criterion for performance evaluation of wavelets and the concept of majorization of subbands proposed by Vaidyanathan [3] for optimal orthonormal filter banks. We characterize wavelets based on the characteristics (ordering) of the wavelet coefficients obtained after the wavelet transform. In fact, it is not the wavelet filter alone that determines the compression efficiency of the system. It needs to be emphasized that in the three-stage framework described above, the second and third stages also contribute significantly to the compression efficiency apart from the wavelet transform in the first stage.

The organization of this paper is as follows. In Section II, modifications of the SPIHT algorithm that exploit correlation in a local neighborhood are described. In Section III, we derive two new criteria for performance evaluation of wavelets in lossless image coding. The experimental results that justify our evaluation criteria are presented in Section IV. Section V provides the summary and conclusions.

II. MODIFICATIONS TO THE SPIHT ALGORITHM

In theory, it is well known that any pair of biorthogonal filters exhibit perfect reconstruction of the data [1]. In practice, biorthogonal filters with rational filter coefficients are preferable for image coding applications, although Calderbank et al. [9] have also used the lifting scheme to convert the real coefficients into integer values. In our scheme, the transformed coefficients, which are obtained using the biorthogonal filter pair, are encoded as \( 2 \times 2 \) blocks using a top-down splitting and bottom-up scanning, which we call the \( z \)-scan encoding algorithm. The addresses of the \( 2 \times 2 \) blocks of transformed coefficients are obtained from the size of the original image using a top-down hierarchical splitting. The coefficients are then encoded in a bottom-up fashion starting from the \( 2 \times 2 \) block at the top-left corner of the subband. The four coefficients in the \( 2 \times 2 \) blocks (i.e., the leaves of the tree) are coded in a \( z \)-scan format (top-left, top-right, bottom-left, and bottom-right). Again, for a \( 4 \times 4 \) block in any subband, \( 2 \times 2 \) blocks are scanned in a left to right and top to bottom \( (z \text{-scan}) \) format. This process is carried out until the size