structed from all the AWT coefficients and from reduced coefficients are investigated, and the data reduction criteria are obtained from uniform resampling. The proposed reconstruction algorithm is found to allow increased rate of reduction.

The auditory wavelet transform simulates the human auditory periphery as a first-order approximation because the wavelet theory requires the use of time invariant filters that all have the same shape on a logarithmic scale. The filtering function at each point along the length of the human cochlea is dynamically adjusted according to the input sound pressure and other factors. The variable filters should be realized in future analysis/synthesis auditory models. The auditory wavelet transform and the reconstruction algorithm may nevertheless improve signal production for auditory psychological experiments and other applications.

ACKNOWLEDGMENTS

We thank Masaaki Hondo, Kazuhiko Kakehi, and Tatsuya Hirahara, as well as the members of our research group, for their valuable advice and comments.

REFERENCES

[1] L. R. Rabiner and R. W. Schafer, *Digital Signal Processing of Speech Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1978, pp. 250–354.
[2] J. B. Allen, J. L. Hall, A. Hubbard, T. S. Neely, and A. Tubis, Eds., *Peripheral Auditory Mechanisms*. Berlin: Springer-Verlag, 1985.
[3] J. M. Combes, A. Grossmann, and P. Tchamitchian, Eds., *Wavelets*. Berlin: Springer-Verlag, 1989.
[4] R. Kronland-Martinet, J. Morlet, and A. Grossmann, "Analysis of sound patterns through wavelet transforms," *Int. J. Pattern Recognition, Artificial Intell.*, vol. 1, no. 2, pp. 97–126, 1987.
[5] J. S. Lienard, "Speech analysis and reconstruction using short-time elementary waveforms," *IEEE ICASSP'87*, pp. 948–951.
[6] R. W. Gerchberg and W. R. Saxton, "A practical algorithm for the determination of phase," *Optik*, vol. 35, no. 2, pp. 237–246, 1972.
[7] M. H. Hayes, J. S. Lim, and A. V. Openheim, "Signal reconstruction from phase or magnitude," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp.672–680, Dec. 1980.
[8] S. H. Nawab, T. F. Quatieri, and J. S. Lim, "Signal reconstruction from short-time Fourier transform magnitude," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 986–998, Aug. 1983.
[9] D. W. Griffin and J. S. Lim, "Signal estimation from modified short-time Fourier transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 236–243, Apr. 1984.
[10] X. Yang, K. Wang, and S. A. Shamma, "Auditory representations of acoustic signals," *IEEE Trans. Informat. Theory*, vol. 38, pp.824–839, Mar. 1992.
[11] H. Fletcher, *Speech and Hearing in Communication*. New York: Krieger, 1972.
[12] T. Kirihata and H. Sakaguchi, "Digital simulation of Basilar membrane by It(z) transformation," *Proc. Autumn Meeting Acoust. Soc. Jpn.*, 1-5-3, Oct. 1986, pp. 253–254.
[13] T. Irino and K. Kawahara, "Simulation of ear using fluid dynamics model of Cochlea," NTT Basic Res. Labs. Tech. Rep. ISRL-89-1, July 1989.
[14] J. O. Pickles, *An Introduction to the Physiology of Hearing*, 2nd ed. San Diego, CA: Academic, 1988.
[15] T. Irino and K. Kawahara, Wavelet analysis and synthesis using the impulse response of an auditory peripheral model," *Proc. Spring Meeting Acoust. Soc. Jpn.*, 1-8-1, Mar. 1991, pp. 351–352.
[16] ——, "A comparative study on the subjective assessment measure and the distance measure of the signal reconstructed by auditory wavelet transform," *Proc. Spring Meeting Acoust. Soc. Jpn.*, 1-6-15, Mar. 1992, pp. 391–392.
[17] ——, "Signal reconstruction from modified wavelet transform—An application to auditory signal processing," *IEEE ICASSP'92*, vol. 1, Mar. 1992, pp. 85–88.

Fractal Estimation from Noisy Data via Discrete Fractional Gaussian Noise (DFGN) and the Haar Basis

Lance M. Kaplan and C.-C. Jay Kuo

***Abstract*—We show that the application of the discrete wavelet transform (DWT) using the Haar basis to the increments of fractional Brownian motion (fBm), also known as discrete fractional Gaussian noise (DFGN), yields coefficients which are weakly correlated and have a variance that is exponentially related to scale. Similar results were derived by Flandrin, Tewfik, and Kim for a continuous-time fBm going through a continuous wavelet transform (CWT). The new theoretical results justify an improvement to a fractal estimation algorithm recently proposed by Wornell and Oppenheim. The performance of the new algorithm is compared with that of Wornell and Oppenheim's algorithm in numerical simulation.**

I. INTRODUCTION

To model stochastic processes that exhibit significant correlation for large lags, Mandelbrot and Van Ness [11] introduced fractional Brownian motion (fBm) which is a generalization of normal Brownian motion. The fBm $B_H(t)$ is a zero mean nonstationary Gaussian random process with the covariance function

$$r_{B_H}(t, s) = \frac{\sigma^2}{2}\left[|t|^{2H} + |s|^{2H} - |t - s|^{2H}\right] \tag{1.1}$$

where the parameters σ^2 and $0 < H < 1$ characterize the process. The parameter H controls the "roughness" of the fBm such that an individual realization of the process has a fractal dimension [10]

$$D = 2 - H.$$

The H parameter also controls the shape of the average spectral density defined as [3]

$$S(f) = \frac{c}{|f|^{\gamma_b}}$$

where

$$\gamma_b = 2H + 1. \tag{1.2}$$

As a result, the fBm serves as a good model for $1/f$ processes where $1 < \gamma_b < 3$, which represents the infrared (IR) catastrophe case [18]. The IR case is the most common case of $1/f$ processes, and many examples of these $1/f$ processes can be found in nature and even economics [7].

The continuous-time fBm is of only theoretical interest. For practical computation, we have to sample the continuous-time fBm via

$$B[k] = B_H(k\Delta x), \qquad k \in \mathbf{Z} \tag{1.3}$$

Manuscript received July 28, 1992; revised June 10, 1993. The Guest Editor coordinating the review of this paper and approving it for publication was Dr. Ahmed Tewfik. This work was supported by a National Science Foundation Graduate Fellowship and the National Science Foundation Young Investigator Award under Grant ASC-9258396.

The authors are with the Signal and Image Processing Institute and the Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 90089-2564.

IEEE Log Number 9212186.

where Δx is the sampling period. The increment

$$X[k] = B[k+1] - B[k] \tag{1.4}$$

defines a sequence known as the discrete fractional Gaussian noise (DFGN). The variance of the fBm has a self-similar property which provides a variance relation of [8]

$$\text{var}\,(B[k+p] - B[k]) = |p|^{2H} \text{var}\,(B[k+1] - B[k]). \tag{1.5}$$

It can also be shown [8] that the DFGN is a zero mean stationary Gaussian process that is characterized by its autocorrelation

$$r[k] = \frac{\sigma^2}{2} |\Delta x|^{2H} [\,|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}]. \tag{1.6}$$

Without loss of generality, we set $\Delta x = 1$ for the following discussion.

The estimation of the fractal dimension of the continuous-time fBm from its real discrete data is a very important problem, which can be applied to linear prediction problems and texture classification [8]. Many fractal estimation techniques exist [5], [8], [13], but they cannot handle noisy measurements. Most of the fractal estimators use regression analysis to take advantage of the exponential progressions, which are characteristics of fractals, in the frequency or space domain. In fact, the exponential progression for the coefficients of the continuous wavelet transform (CWT) of the continuous-time fBm was studied in [3], [4], [14], [16], which leads to estimators using wavelets.

Wornell and Oppenheim [18] created a fractal estimator algorithm using the wavelet transform to derive an approximate maximum likelihood estimator when a $1/f$ process is embedded in white Gaussian noise. Wornell and Oppenheim's algorithm first passes sampled $1/f$ noise measurements through the DWT based on an algorithm of Mallat [9]. That is, given a sampled signal $a_0[k]$ of the finest scale $m = 0$, we calculate the approximation coefficients $a_m[k]$, and the detail coefficients $d_m[k]$ of coarser scale $m > 0$ recursively via

$$a_{m+1}[k] = \sum_{k'=-\infty}^{\infty} h[2k - k']\, a_m[k'], \tag{1.7}$$

$$d_{m+1}[k] = \sum_{k'=-\infty}^{\infty} g[2k - k']\, a_m[k'] \tag{1.8}$$

where $h[k]$ and $g[k]$ must satisfy the quadrature mirror filter (QMF) constraint, i.e., $g[k] = (-1)^k h[1-k]$. Then, the average power, or variance, of the wavelet coefficients, $d_m[k]$, for each scale is calculated. Finally, the variance estimates are put through an estimate-maximize (EM) algorithm to find the maximum likelihood estimate of H. The EM algorithm is based on the fact that the wavelet coefficients are uncorrelated and that for a given scale, the variance of the coefficients is

$$\text{var}\,[d_m[k]] = \sigma^2 2^{\gamma_b m} + \sigma_w^2 \tag{1.9}$$

where γ_b is related to H via (1.2) and σ_w^2 is the variance of the initial white noise process. The algorithm also estimates σ^2 and σ_w^2. The results of these estimates, however, are of no concern in this paper. The EM algorithm can be tuned for three modes: σ_w^2 is unknown, $\sigma_w^2 > 0$ is known, and $\sigma_w^2 = 0$ is known.

Wornell justified the variance progression in (1.9) by arguing that the wavelet transform acts as a Karhunen–Loève-like expansion for $1/f$ noise [17]. That is, if $d_m[k]$ is indeed uncorrelated over scale and time and if the variance progression of the coefficients does follow (1.9) where $\sigma_w^2 = 0$, then the actual process is nearly $1/f$ in the sense that the power spectrum is bounded by

$$\frac{c_1}{|f|^{\gamma_b}} \le S(f) \le \frac{c_2}{|f|^{\gamma_b}}.$$

The variance progression was also validated in [16], where it is proved that if fBm is passed through a continuous wavelet transform then (1.9) is correct and that the $d_m[k]$ terms are weakly correlated in scale and time. To be precise, the cross correlation of $d_m[k]$ and $d_n[l]$ decays asymptotically as

$$E[d_m[k]\, d_n[l]] \sim O(|2^m k - 2^n l|^{2(H-R)}) \tag{1.10}$$

where R is the number of vanishing moments of the selected wavelet basis.

Consequently, the wavelet basis must be selected such that $R \ge 2$ to whiten the fBm, which means the support of the basis must be at least of length 4 [1]. Errors in the computation of Wornell's algorithm appear because not all wavelet coefficients can be computed exactly for finite supported fBm by using a wavelet basis of length greater than 2. For this case, it is often that the DWT computation assumes periodic extension of available data. However, the extrapolation method may cause some harmful effects. Besides, Wornell and Oppenheim's algorithm must perform the DWT on the discretely sampled fBm. Then, the variance progression of (1.9) for the wavelet coefficients is biased and this error leads to an under estimate of the parameter H when only a small amount of measurements are available [4].

To fix the problem, we first show in Section II that when the increments of the sampled fBm (or the DFGN), is set equal to the finest scale wavelet approximation coefficients and when the Haar basis is selected, the DWT coefficients are weakly correlated and have a variance that is exponentially related to scale. Then, we discuss in Section III a modified fractal estimation algorithm by using the DFGN and the Haar wavelet transform. The new algorithm will be compared with the algorithm of Wornell and Oppenheim in Section IV, and Section V concludes the paper.

II. Properties of DWT Coefficients of DFGN Using Haar Basis

Wornell and Oppenheim's algorithm uses the sampled fBm which is a nonstationary process and can be generated by the DFGN. Since the sampled fBm and the DFGN are causally invertible, they share the same amount of information. By applying the CWT to the fBm, the nonstationarity falls into the CWT approximation coefficients. For the DWT case, however, even though the wavelet coefficients are stationary for the Haar basis, the recursive computation of the coefficients causes the nonstationarity to propagate through scale and bias the variance progression. As a fix to the problem, we consider the application of the DWT to the DFGN.

A. Theory

In this subsection, we will examine the statistical properties of the coefficients of the discrete Haar transform applied to DFGN. Analogous discussion for the case of the CWT applied to fBm appears in [4], [16]. A more general treatment of the statistical properties of CWT coefficients of stochastic processes can be found in [2]. We will show that when the finest approximation coefficients are set to be equal to the DFGN, the DWT coefficients have many desirable properties. Roughly speaking, they follow a nice variance progression per scale, and as lag increases, the correlation of the coefficients decay much faster than that of the correlation of the DFGN. To make the correlation comparison, it is instructive to

perform a Taylor series expansion on (1.6) to understand the asymptotic decay of $r[k]$. The following Taylor series expression is very useful for $|\alpha| < |t|$

$$\left|1 + \frac{\alpha}{t}\right|^{2H} = \sum_{r=0}^{\infty} \frac{\alpha^r}{r!} V(r) |t|^{-r} \tag{2.1}$$

where

$$V(r) = \prod_{s=0}^{r-1} (2H - s).$$

Then, by substituting (2.1) into (1.6), we can express $r[k]$ for $k \neq 0$ as

$$r[k] = \frac{\sigma^2}{2} |k|^{2H} \sum_{r=1}^{\infty} \frac{2}{(2r)!} V(2r) |k|^{-2r}. \tag{2.2}$$

The rate of decay of $r[k]$ is controlled by the first coefficient of the summation. The Taylor series expansion in (2.2) helps to demonstrate the whitening effect of the DWT. The properties of the wavelet coefficients are stated and proved below.

Theorem 1: Let $B[k]$ be sampled fBm with a parameter $0 < H < 1$, and $h[n]$ and $g[n]$ be the scaling and wavelet filter realizations of the Haar basis as given in (1.7) and (1.8). Define the stochastic process $d_m[k]$ as the detail wavelet coefficients of the DWT where the finest scale approximation $a_0[k]$ is the increment $X[k] = B[k+1] - B[k]$ of the sampled fBm. Then,

a) for fixed scale m the variance of $d_m[k]$ is

$$\text{var}\,[d_m[k]] = 2^{\gamma(m-1)} \sigma^2 (2 - 2^{\gamma}) \tag{2.3}$$

where

$$\gamma = 2H - 1 \tag{2.4}$$

b) for fixed scale m the autocorrelation $r_{d_m}[k-l]$ of $d_m[k]$ decays as $O(|k-l|^{2(H-2)})$ for all k, l such that $|k-l| > 1$.

Proof: Since $a_0[k] = X[k]$, the autocorrelation of $a_0[k]$ is given by (1.6). First, we want to show that the autocorrelation $r_{a_m}[k]$ of the approximation coefficients $a_m[k]$ for scale $m \geq 0$ as calculated via (1.7) is related to the autocorrelation $r[k]$ of the increments $X[k]$ by

$$r_{a_m}[k] = 2^{(2H-1)m} r[k]. \tag{2.5}$$

It is obvious that (2.5) holds for $m = 0$. Let us assume

$$r_{a_{m-1}}[k] = 2^{(2H-1)(m-1)} r[k]. \tag{2.6}$$

Due to (1.7) and the Haar basis, we have

$$a_m[k] = \frac{1}{\sqrt{2}} (a_{m-1}[2k] + a_{m-1}[2k+1]). \tag{2.7}$$

By combining (2.6) and (2.7), it is easy to show that the autocorrelation of $a_m[k]$ is

$$r_{a_m}[k] = \frac{2^{(2H-1)(m-1)}}{2} (2r[2k] + r[2k+1] + r[2k-1]). \tag{2.8}$$

By substituting (1.6) into (2.8), one obtains

$$\begin{aligned} r_{a_m}[k] &= 2^{(2H-1)m} \frac{\sigma^2}{2} (|k+1|^{2H} + |k-1|^{2H} - 2|k|^{2H}) \\ &= 2^{(2H-1)m} r[k]. \end{aligned}$$

Thus, by induction, (2.5) holds for all $m \geq 0$. Next, we examine the autocorrelation $r_{d_m}[k]$ of the detail coefficients $d_m[k]$. From (1.8) and the Haar basis,

$$d_m[k] = \frac{1}{\sqrt{2}} (a_{m-1}[2k] - a_{m-1}[2k+1]). \tag{2.9}$$

By combining (2.5) and (2.9), one gets

$$r_{d_m}[k] = \frac{2^{(2H-1)(m-1)}}{2} (2r[2k] - r[2k+1] - r[2k-1]). \tag{2.10}$$

Thus, (2.3) is obtained by setting $k = 0$ in (2.10) and using (1.6), and Part (a) is proved.

Substitution of (1.6) in (2.10) yields

$$\begin{aligned} r_{d_m}[k] = \frac{2^{(2H-1)(m-1)}}{2} \frac{\sigma^2}{2} &(4|2k+1|^{2H} + 4|2k-1|^{2H} \\ &- |2k+2|^{2H} - |2k-2|^{2H} - 6|2k|^{2H}). \end{aligned} \tag{2.11}$$

By factoring out the $2k$ term in (2.11) and using the Taylor series expansion (2.1), we have for $|k| > 1$

$$r_{d_m}[k] = 2^{(2H-1)m} \frac{\sigma^2}{2} |k|^{2H} \sum_{r=0}^{\infty} \frac{P(r)}{r!} V(r) |k|^{-r}, \tag{2.12}$$

where

$$P(r) = 4(\tfrac{1}{2})^r + 4(-\tfrac{1}{2})^r - 1 - (-1)^r - 6\delta(r).$$

Since,

$$P(r) = 0 \qquad \text{for } r = 0, 1, 2, 3$$

the autocorrelation of $d_m[k]$ decays as expressed in Part (b). □

It is worthwhile to point out several features for the above result. First, the scaling filter (1.7) of the Haar basis takes full advantage of the self-similarity of the fBm as characterized by (1.5). In particular, for approximation coefficients $a_m[k]$, each increase in scale m is like subsampling the fBm by a factor of two and, due to (1.6), a 2^{2H} term pops up in the expression for the correlation of $a_m[k]$ by doubling the value of k. Note also that the fBm increment $X[k]$ approximates the first-order differentiation so that the asymptotic behavior of the DFGN correlation from (2.2) is $O(|k|^{2H-2})$ and the variance progression is governed by (2.4) rather than (1.2). Finally, the regularity of the Haar wavelet filter causes the first term in the Taylor series expansion in (2.2) to disappear. Theorem 1 can be generalized below for the correlation $r_d[m, n; k, l]$ of $d_m[k]$ and $d_n[l]$ between scales m and n.

Theorem 2: Let $d_m[k]$ be defined as in Theorem 1. Then,

a) The correlation $r_d[m, n; k, l]$ of $d_m[k]$ and $d_n[l]$ is

$$\begin{aligned} r_d[m, n; k, l] &= E[d_m[k] d_n[l]] \\ &= \frac{\sigma^2}{2} 2^{-(m+n)/2} (2|2^m k - 2^n l \\ &\quad + 2^m - 2^{n-1}|^{2H} \\ &\quad + 2|2^m k - 2^n l + 2^{m-1} - 2^n|^{2H} + 2|2^m k \\ &\quad - 2^n l + 2^{m-1}|^{2H} \\ &\quad + 2|2^m k - 2^n l - 2^{n-1}|^{2H} - 4|2^m k - 2^n l \\ &\quad + 2^{m-1} - 2^{n-1}|^{2H} \\ &\quad - |2^m k - 2^n l + 2^m - 2^n|^{2H} - |2^m k - 2^n l|^{2H} \\ &\quad - |2^m k - 2^n l + 2^m|^{2H} - |2^m k - 2^n l - 2^n|^{2H}). \end{aligned} \tag{2.13}$$

b) For given m and n the correlation $r_d[m, n; k, l]$ of $d_m[k]$ decays as $O(|2^m k - 2^n l|^{2(H-2)})$ for all k and l such that $|2^m k - 2^n l| > \max(2^n, 2^m)$.

Proof: See [6]. □

Theorem 2 proves that the detail DWT coefficients are weakly correlated over both time and scale. The whitening effect is much stronger for the special case of normal Brownian motion ($H = 1/2$). By setting $H = 1/2$ in (2.13), then the correlation is zero as long as $2^m k - 2^n l \geq 2^n$ or $2^m k - 2^n l \leq -2^m$. Thus, we have the following corollary.

Corollary 1: Given the increments of $B[k]$ with $H = 1/2$, the detail wavelet coefficients $d_m[k]$ from the Haar transform are independent, i.e.

$$r_d[m, n; k, l] = 0$$

provided that m and n are fixed and that k and l are outside the region defined as

$$2^{n-m} l - 1 < k < 2^{n-m}(l + 1).$$

This result is similar to one found in [4] for the case of the CWT and continuous fBm, which in turn is a special case of results on the CWT of stochastic processes with orthogonal increments as presented in [2]. Since the increments of Brownian motion are white, Corollary 1 just states that the Haar transform of white noise produces white noise. This fact is true for any orthogonal wavelet basis and is used in [18].

The whitening effect for general DFGN will be verified experimentally in Section II-C. One advantage of using the DFGN and the Haar basis is that because the DWT coefficients are virtually uncorrelated for $0 < H < 1$ via Haar transform, the periodic extension of data problem is avoided. By comparing Part (b) of Theorem 2 to (1.10), we may say that the application of the discrete Haar transform to the DFGN, in a sense, behaves like the application of the continuous wavelet transform with the Daubechies D4 basis to the corresponding continuous-time fBm.

B. Extension to Higher Order Wavelet Filters

It seems that some results presented in Section II-A can be extended to wavelet bases of higher order. It is our conjecture that when the DFGN is passed through the DWT using a Daubechies filter [1] of length $2R$ (or any orthogonal filter with R vanishing moments), the R vanishing moments of the filter should cancel out the first R terms in the correlation expansion of (2.2) so that the detail wavelet coefficients decay at a rate of

$$E[d_m[k]\, d_n[l]] = O(|2^m k - 2^n l|^{2(H-1-R)}). \tag{2.14}$$

In fact, a technique originated in [12] was used by Tewfik and Kim in [15] to verify the decay of (2.14). Their work is based on multiscale signal processing of sampled data, and it covers the special case of orthonormal wavelet basis. When the correlation structure of (1.6) is treated as a continuous function, the function will be continuously differentiable for $k > 1$. Based on [15], it can be argued that if a length $2R$ Daubechies filter is used for DWT implementation, the detail wavelet coefficients are bounded by

$$E[d_m[k]\, d_n[l]] \leq \sup_{0 \leq k', l' < 2R} \left| \frac{d^{2R} r[t]}{dt^{2R}} \right|_{t = 2^m k - 2^n l - k' + l'}. \tag{2.15}$$

By taking the $2R$ derivatives of (2.2), the asymptotic decay (2.14) is verified and experimental verification is provided in Section III-C. Thus, the DWT should be able to whiten the DFGN for any choice of wavelet basis with $R > 0$ vanishing moments.

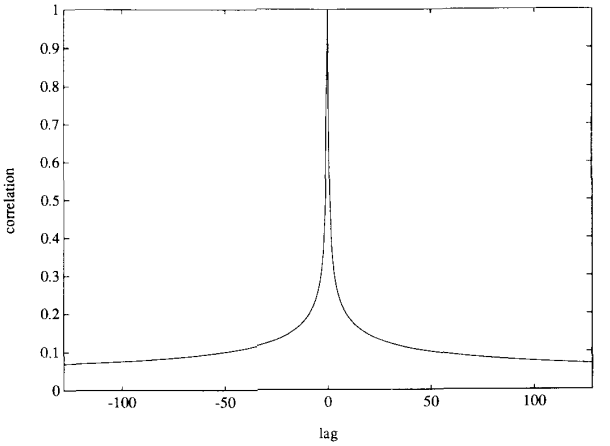

Fig. 1. Autocorrelation of DFGN with $H = 0.8$.

It is believed, however, that only the Haar basis ($R = 1$) takes advantage of the self-similarity (1.5) of the fBm so that the variance progression is not biased. For the case of the CWT, [16, Theorem 1] states that the detail coefficients have the same statistics within a scaling factor for each scale, and this scaling factor controls the "clean" variance progression. For the DWT case, however, the similar correlation functions of the detail coefficients translates into the fact that the approximation correlation functions must be the same between scale within a scaling factor. In the proof of Theorem 1, the variance progression (2.3) is a direct result of the fact that as the approximation coefficients become coarser, the correlation structure remains the same with the exception of a scaling factor. This nice feature does not occur for higher order Daubechies or B-spline filters (the Haar filter is the first order member of both families of filters). The bias in the variance progression is verified experimentally in Section IV. Even if the higher order filters did not suffer from variance bias, they would be of limited value because of the windowing problem that results from the DWT implementation. Longer filters might create coefficients with correlation that decays faster than the Haar filters, but it is shown in Section II-C that the faster decay is not significant for real data.

C. Experimental Verification

To demonstrate the above discussion, the correlation of DFGN for fBm with $H = 0.8$ is examined. Fig. 1 shows the theoretical autocovariance function of the DFGN process for lags -128 to 128 as given by (1.6). We clearly see that there is a rapid drop from lag 0 to other nonzero lags, but the function is slowly converging to zero. When the increments of a fBm realization of length 128 is treated as a vector and is put through a Haar transform, the output can also be treated as a vector where element 1 represents the coarsest approximation coefficient and elements 2 through 128 represent the detail coefficients from coarsest to finest. Because the approximation coefficient is of no interest, it is discarded and the resulting output vector is of length 127. By using (2.13), a theoretical covariance matrix is computed and displayed in Fig. 2. Cross-correlation between overlapping time segments of different scales and the hyperbolic decay of the peaks are both evident in the figure. Finally, 256 realizations of fBm increments of length 128 were generated using the Cholesky approach as in [8]. These samples were put through the Haar transform and an average covariance matrix was computed. Fig. 3 shows the matrix, and the figure

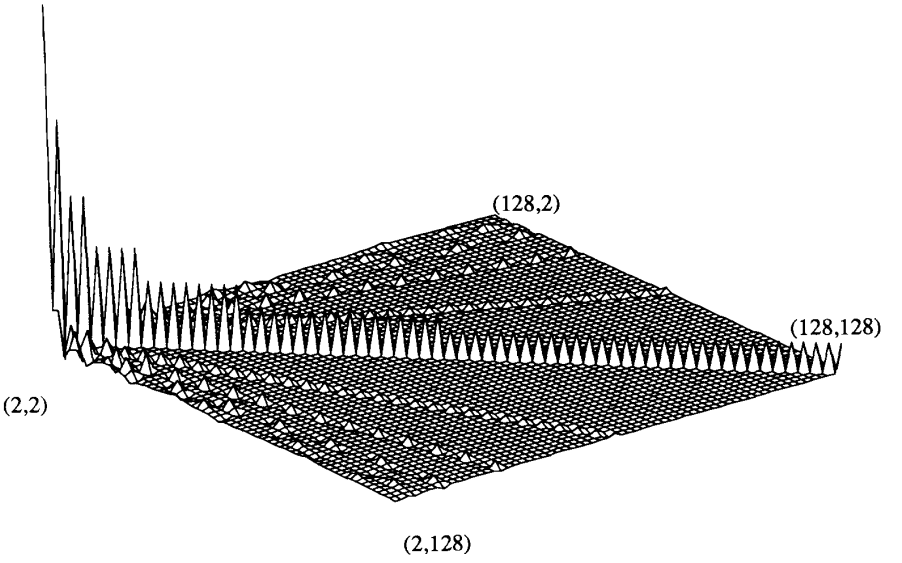

Fig. 2. Magnitude of entries of the theoretical covariance matrix of the wavelet coefficients with the Haar basis and $H = 0.8$.

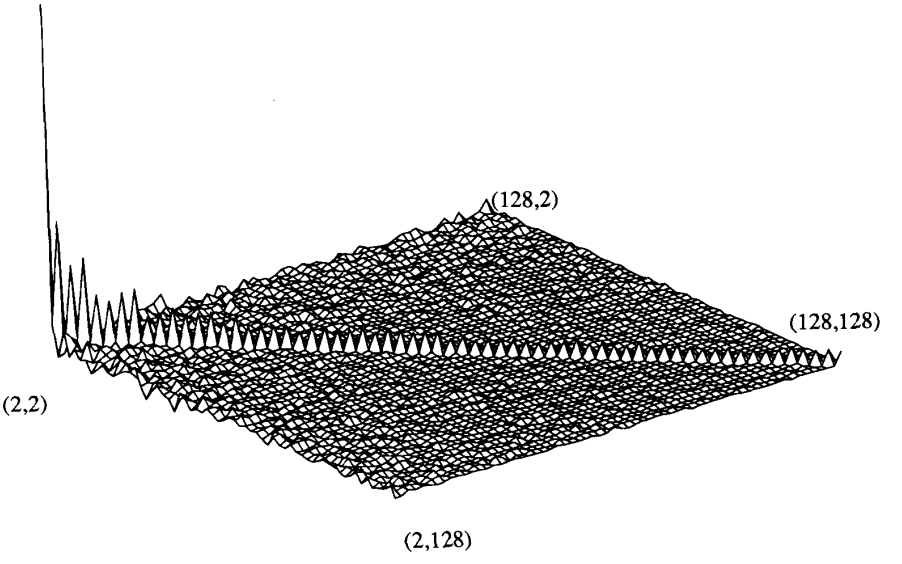

Fig. 3. Magnitude of entries of the experimental covariance matrix of the wavelet coefficients with the Haar basis and $H = 0.8$.

displays a large diagonal peak with small off-diagonal elements. The variance progression is also evident. These experimental results validate the theory in Section II-A.

To verify that higher order Daubechies filters do whiten the DFGN, we use the D4 and D16 bases to perform the same numerical experiment. In these experiments when the length of the filter became too long for the DWT implementation, a smaller filter is chosen just as in [16]. Figs. 4 and 5 show the covariance matrices for the wavelet coefficients using the D4 and D16 bases, respectively. The whitening effect is verified. The faster decay of the peaks for the D16 filter, however, is not evident. The fact that a decay should occur seems more crucial than the actual rate of decay for real data. Besides, the higher order bases suffer from variance bias as will be tested in Section IV.

III. The Modified Algorithm

The theoretical results suggest that the Wornell and Oppenheim's algorithm can be improved by first computing the increments so that the problems of variance bias and periodic data extensions can be avoided. Then, we apply the Haar transform to the DFGN, compute the average variance of detail coefficients $d_m[k]$ for each scale and use the EM algorithm to find a maximum like-

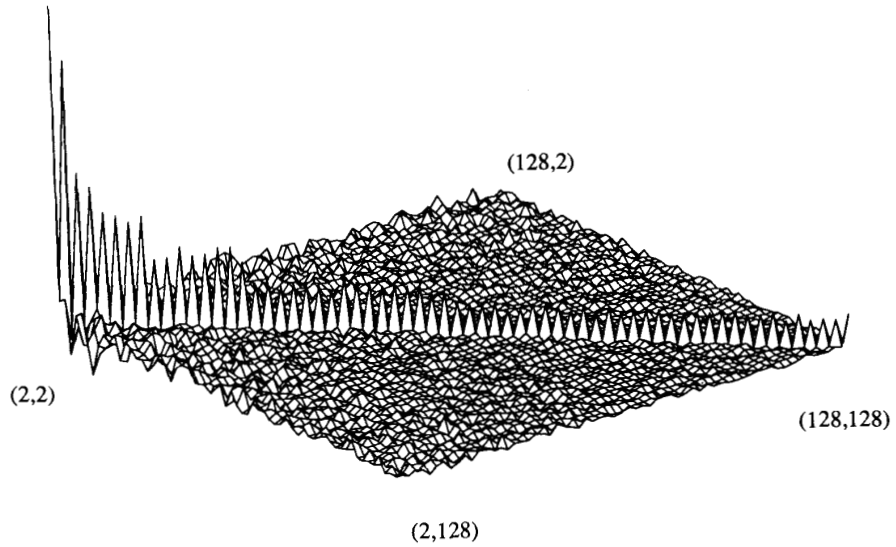

Fig. 4. Magnitude of entries of the experimental covariance matrix of the wavelet coefficients with the Daubechies D4 basis and $H = 0.8$.

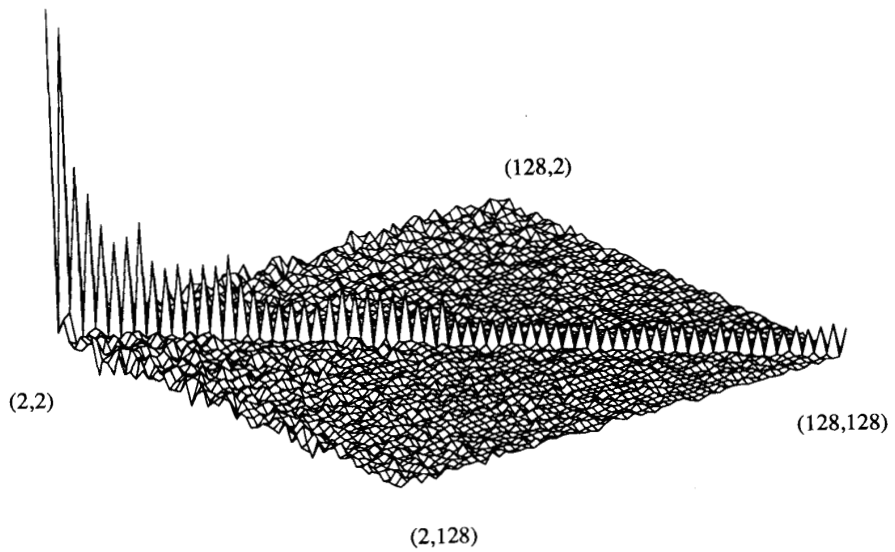

Fig. 5. Magnitude of entries of the experimental covariance matrix of the wavelet coefficients with the Daubechies D16 basis and $H = 0.8$.

lihood estimate of H. It is important to note that the EM algorithm must be altered to reflect the variance progression of the noisy data.

To create a fractal estimator for noisy measurements, one must understand what happens to white noise when its increments are passed through the DWT. Let us assume that the additive white Gaussian noise $w[k]$ is independent of the $1/f$ process and has variance σ_w^2. Then, the white noise increment $w_i[k] = w[k+1] - w[k]$ has a autocorrelation structure of

$$r_{w_i}[k] = \begin{cases} 2\sigma_w^2 & \text{if } k = 0 \\ -\sigma_w^2 & \text{if } |k| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

One can verify that the process $w_i[k]$ is in fact a zero mean stationary Gaussian process characterized by the autocorrelation (1.6) with $H = 0$ and $\sigma^2 = 2\sigma_w^2$. It is important to note that even though the increment is of the form (1.6), white noise is not the same as fBm of $H = 0$ because the way DFGN is added up to generate fBm, the first element of the discrete-time fBm with $H = 0$ is correlated with all future samples. Since the increments, however, are equivalent, the same theory for the DFGN hold for $w_i[k]$. The variance progression of $w_i[k]$ can be determined from (2.3) by setting $H = 0$. Since the DWT is linear and white noise is independent of the fBm process, the wavelet coefficients from the white noise increment is independent of the wavelet coefficients of the DFGN,

and the variance progression of noisy DFGN is

$$\sigma_m^2 = \text{var}\,[d_m[k]] = 2^{\gamma(m-1)}\sigma^2(2 - 2^{\gamma}) + 2^{-(m-1)}3\sigma_w^2 \quad (3.1)$$

where γ is defined in (2.4). In a sense, the modified EM algorithm, formulated according to (3.1), will separate two fBm signals, i.e., one with $H = 0$ and the other one with unknown H.

Using the variance progression given in 3.1 and the "whiteness" assumption, it is easy to derive that the likelihood function for the given data is

$$L(\Theta) = -\frac{1}{2}\sum_{m=1}^{M} N(m)\left[\frac{\hat{\sigma}_m^2}{\sigma_m^2} + \ln(2\pi\sigma_m^2)\right]$$

where $\hat{\sigma}_m^2$ is the sample variance, $N(m)$ is the number of available samples at scale m, and Θ is the parameter space. The derivation of formulas for the estimate maximize (EM) algorithm can be found in [18, Appendix A] with trivial changes, and thus, will not be repeated here, and the new EM formulas are listed in [6, Appendix A]. The new EM algorithm can be tuned for the same three modes as the original EM algorithm that was discussed in Section I.

The modified algorithm can be directly compared to a very common fractal estimation technique known as the variance estimator that was used as a control algorithm in [8]. It is basically a space domain version of the algorithm presented in [13]. The variance estimator uses the self-similar property (1.5). The variance of the increments are computed for different lags, i, and then regression analysis on a log–log scale is used to determine H. By looking at the proof of Theorem 1, one can see that in effect, the variance estimator performs the regression analysis on the approximation coefficients. In other words, the modified algorithm adds the wavelet filtering step of (1.8) to whiten the coefficients so that a maximum likelihood estimate of H is easy to formulate. Even without the maximum likelihood formulations, a regression analysis after the wavelet filtering step should provide for more accurate estimates because the coefficients are virtually independent.

IV. Simulations

Wornell and Oppenheim's (WO) algorithm and our modified algorithm were compared using simulated fBm data. The simulated data was generated using the Cholesky decomposition approach as in [8]. This method is chosen because it provides accurate fBm realizations. In this section, the signal to noise ratio is computed as

$$\text{SNR} = 10\log_{10}\frac{\sigma^2}{\sigma_w^2} \quad (4.1)$$

where σ^2 is the fBm increment power and σ_w^2 is the noise power. Motivation to use 4.1 for SNR instead of considering an approach used in [8] is provided by [6]. To implement the WO algorithm, the 16-tap Daubechies wavelet basis [1] was chosen, and any wavelet coefficients that could not be computed accurately were discarded. For our simulations, 64 samples of noise-free fBm of various lengths where $H = 0.25$, $H = 0.5$, $H = 0.75$, and $H = 0.9$ were created.

The WO and modified algorithms were first tested on the noise-free fBm data of length 2048. The resulting mean, standard deviation, and root mean square (RMS) error of the value of H as estimated by the two algorithms under the noise-free assumption are given in Table I. The problem of variance bias is very much evident in the WO algorithm. The standard deviation of the WO algorithm is also slightly larger. Table II shows the resulting statistics of the estimated parameter H when the two algorithms are tuned to search for a noise level. It appears that the WO algorithm, when searching for a noise floor, treats the variance bias as noise so that a less biased estimate of H can be found, but the standard deviation becomes some what larger than that of the modified algorithm. Table II also shows that the RMS error of the two algorithms are comparable for data of length 2048. The performance of the modified algorithm in the noise search mode, however, beats the WO algorithm for shorter data as demonstrated by Table III.

TABLE I
Means and Standard Deviations for the H Estimate with 2048 Noise-Free Measurements in the Noise-Free Mode

True	WO			Modified		
H	mean	std	rms	mean	std	rms
0.90	0.846	0.021	0.058	0.899	0.017	0.017
0.75	0.683	0.026	0.072	0.748	0.017	0.017
0.50	0.398	0.021	0.109	0.499	0.016	0.016
0.25	0.082	0.022	0.169	0.252	0.016	0.016

TABLE II
Means and Standard Deviations for the H Estimate with 2048 Noise-Free Measurements in the Noise Search Mode

True	WO			Modified		
H	mean	std	rms	mean	std	rms
0.90	0.894	0.032	0.032	0.917	0.019	0.025
0.75	0.760	0.040	0.041	0.770	0.029	0.035
0.50	0.510	0.043	0.044	0.520	0.021	0.030
0.25	0.272	0.050	0.055	0.290	0.038	0.055

TABLE III
Means and Standard Deviations for the H Estimate with 512 Noise-Free Measurements in the Noise Search Mode

True	WO			Modified		
H	mean	std	rms	mean	std	rms
0.90	0.908	0.088	0.088	0.926	0.039	0.047
0.75	0.745	0.103	0.103	0.768	0.041	0.045
0.50	0.536	0.106	0.112	0.542	0.050	0.065
0.25	0.275	0.152	0.154	0.312	0.069	0.093

Next, white noise was added to the fBm data so that the SNR is 10 dB. The two algorithms were tested in the noise search mode. The RMS errors of the H estimate are plotted versus data length in Figs. 6 through 9 for various values of H. It is clear that the modified algorithm is superior for moderate data lengths while the two algorithms perform almost equally well at length 2048.

Flandrin [4] showed that the bias in the variance progression, that leads to an underestimate of H, decreases as the scale becomes coarser. The EM algorithm, however, gives more credence to the variance estimates of finer scales because they contain more samples of data. As a result, the WO algorithm operating in the noise-free mode has difficulty in decreasing the error in the estimate of H as the number of samples increase. A simple fix to this problem is to just throw away the data form the finer scales before the EM algorithm is used. Fig. 10 shows the mean value of the H estimate when the WO algorithm is implemented in noise free mode via the D4, D8, and D16 basis on noise free samples of fBm data with $H = 0.25$ and where the finest k scale detail coefficients are discarded. We see that excluding the finest scales do improve the estimate of H, even though the WO algorithm still underestimates H.

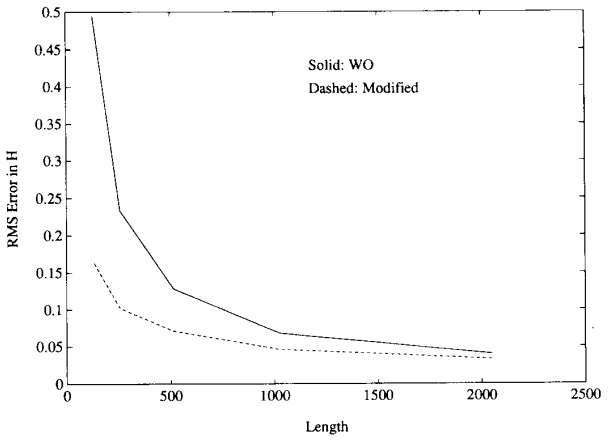

Fig. 6. RMS error of the estimate for $H = 0.9$.

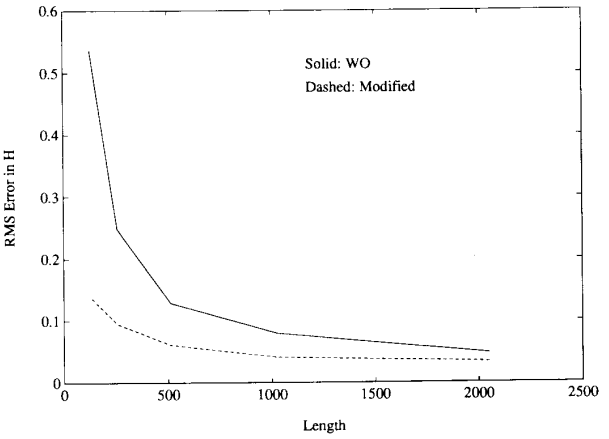

Fig. 8. RMS error of the estimate for $H = 0.5$.

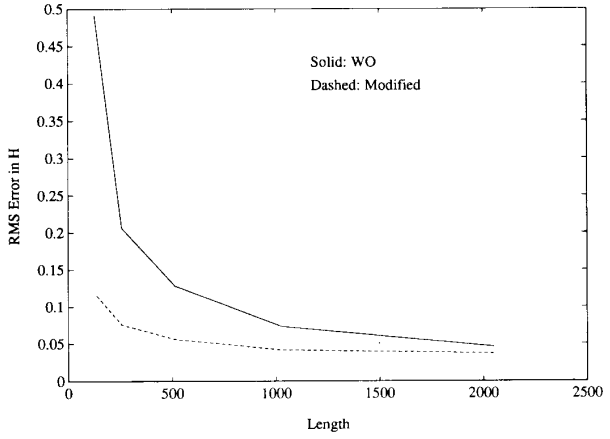

Fig. 7. RMS error of the estimate for $H = 0.75$.

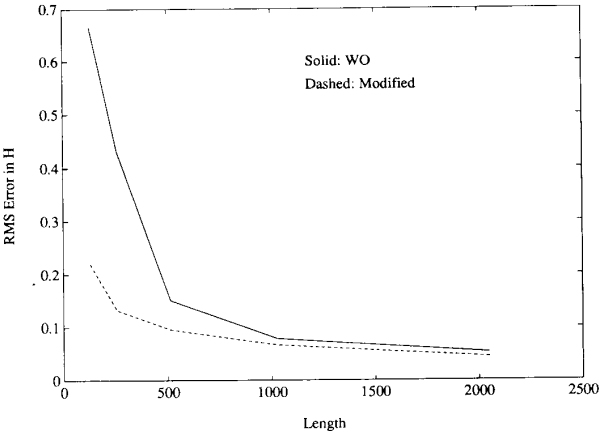

Fig. 9. RMS error of the estimate for $H = 0.25$.

Note also, that the performance of the WO algorithm is not really sensitive to the choice of basis. Next, the same test was applied with the two algorithms in the noise search mode, and the results are shown in Fig. 11. It appears that the noise searching process will help the WO algorithm to provide virtually bias-free estimates. Tables II and III, however, show that the price of the lower bias appears as increased standard deviation.

Since the method of finer scale exclusion offers a way to see if variance bias does not exist in an algorithm, we also use this method to test the variance progression of the modified algorithm extended to higher order Daubechies filters. Fig. 12 shows the result of finer scale exclusion using the same data as in the previous paragraph for the modified algorithm (the Haar basis) and it extension using the D4, D8, and D16 filter. The variance progression bias of (2.3) for higher order filters is evident in the figure.

We want to comment that throwing away the finer scales is not a practical fix to the bias problem, because the finer scales contain most of the data. For example, if data in the 3 finest scales are discarded, only 12.5% of the original data is used for the EM algorithm. As seen in Fig. 10, the quality of the estimator is affected by the absence of the 3 finest scales for the D4 basis. A better way to handle the bias problem is to use an algorithm that does not suffer from the variance bias. The modified algorithm, in noise free mode, is such an algorithm, and its lack of variance bias is clearly seen in Fig. 10.

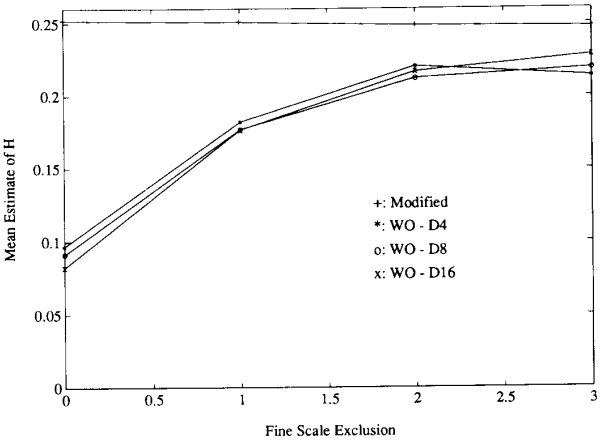

Fig. 10. Mean value of H of various estimators with fine scale exclusion for $H = 0.25$ (noise free mode).

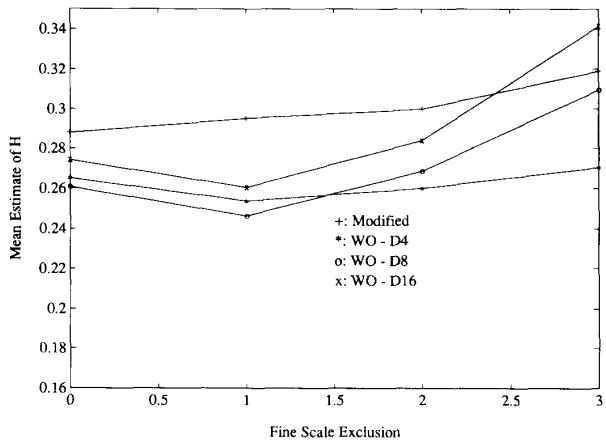

Fig. 11. Mean value of H of various estimators with fine scale exclusion for $H = 0.25$ (noise search mode).

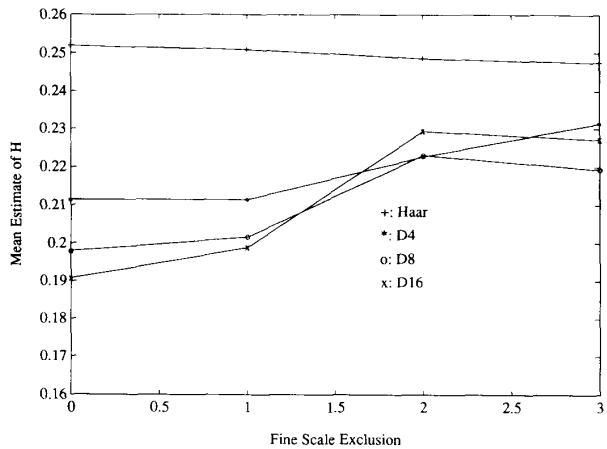

Fig. 12. Mean value of H of various DFGN estimators with fine scale exclusion for $H = 0.25$ (noise free mode).

V. Conclusions and Extensions

A theory about the DWT coefficients of the increments of the fBm using the Haar basis has been presented. It has also been shown that Wornell and Oppenheim's fractal estimation algorithm can be modified as a result of our new theory. The modified algorithm improves the accuracy of the H estimate for moderate data lengths of the fBm. For longer lengths, both our algorithm and Wornell and Oppenheim's algorithm can find very good estimates. For short data lengths of the fBm in additive noise, the modified algorithm is still unreliable, and it is so far not clear how to improve a wavelet based fractal estimator for short data lengths. Although highly regular filters may zero out terms in the Taylor series expansion of the autocorrelation of DFGN, they cause a bias in variance progression. How to avoid the bias of high-order filters is also of interest.

References

[1] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909-996, Nov. 1988.

[2] R. Dijkerman, V. Badrinath, and R. R. Mazumdar, "Multiscale representation of stochastic processes using compactly supported wavelets," in *Proc. IEEE-SP Int. Symp. Time-Frequency Anal.*, Oct. 1992, pp. 185-188.

[3] P. Flandrin, "On the spectrum of fractional Brownian motions," *IEEE Trans. Informat. Theory*, vol. 35, pp. 197-199, Jan. 1989.

[4] —, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. Informat. Theory*, vol. 38, pp. 910-917, Mar. 1992.

[5] N. Gache, P. Flandrin, and D. Garreau, "Fractal dimension estimators for fractal Brownian motions," in *IEEE ICASSP-91*, Mar. 1991, pp. 3557-3560.

[6] L. M. Kaplan and C.-C. J. Kuo, "Fractal estimation from noisy measurements via discrete fractional Gaussian noise (dfGn) and the Haar basis," Univ. of Southern California, Los Angeles, July 1992, Tech. Rep. SIPI 212.

[7] M. S. Keshner, "$1/f$ noise," *Proc. IEEE*, vol. 70, pp. 212-218, Mar. 1982.

[8] T. Lundahl, W. J. Ohley, S. M. Kay, and R. Siffert, "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Trans. Med. Imaging*, vol. MI-5, pp. 152-161, Sept. 1986.

[9] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 11, pp. 674-693, July 1989.

[10] B. Mandelbrot, *The Fractal Geometry of Nature*. San Francisco, CA: Freeman, 1982.

[11] B. Mandelbrot and J. W. V. Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Rev.*, vol. 10, pp. 422-437, Oct. 1968.

[12] Y. Meyer, *Ondelettes et Operateurs*. Paris, France: Herman, 1990.

[13] A. P. Pentland, "Fractal-based description of natural scenes," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. PAMI-6, pp. 661-674, Nov. 1984.

[14] J. Ramanathan and O. Zeitouni, "On the wavelet transform of fractional Brownian motion," *IEEE Trans. Informat. Theory*, vol. 37, pp. 1156-1158, July 1991.

[15] A. H. Tewfik and M. Kim, "Fast multiscale statistical signal processing algorithms," to appear in *IEEE Trans. Signal Processing*.

[16] —, "Correlation Structure of discrete wavelet coefficients of fractional Brownian motion," *IEEE Trans. Informat. Theory*, vol. 38, pp. 904-909, Mar. 1992.

[17] G. W. Wornell, "A Karhunen-Loève-like expansion for $1/f$ processes via wavelets," *IEEE Trans. Informat. Theory*, vol. 36, pp. 859-861, July 1990.

[18] G. W. Wornell and A. V. Oppenheim, "Estimation of fractal signals from noisy measurements using wavelets," *IEEE Trans. Signal Processing*, vol. 40, pp. 611-623, Mar. 1992.

The Strömberg Wavelet and the Franklin System

Djalil Kateb and Karim Drouiche

***Abstract*—The Franklin system is an orthonormal basis of $L^2[0, 1]$. We give a multiresolution analysis of $L^2[0, 1]$ and study the reconstruction error made when the Strömberg analysis (on [0, 1]) is performed instead of the Franklin one. We derive a fast algorithm which enables one to analyze and synthesize a given signal on an interval. Special considerations are held near the edges where the extension of the interval is no longer needed (zero-padding method, etc.). Numerical simulations are provided to emphasize the performances of our results.**

Manuscript received August 20, 1992; revised June 8, 1993. The guest editor coordinating the review of this paper and approving it for publication was Dr. Pierre Duhamel.

D. Kateb is with the Department Statistiques, Universite Paris-Sud, Orsay, France.

K. Drouiche is with Department Signal, Telecom-Paris, 75634, Paris Cedex 13, France.

IEEE Log Number 9212204.