Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2013, Article ID 102834, 14 pages
http://dx.doi.org/10.1155/2013/102834

Research Article

A Study of Wavelet Analysis and Data Extraction from Second-Order Self-Similar Time Series

Leopoldo Estrada Vargas,[1] Deni Torres Roman,[1] and Homero Toral Cruz[2]

[1] *Department of Electrical Engineering, Center of Research and Advanced Studies (CINVESTAV), 45019 Guadalajara, JAL, Mexico*
[2] *Department of Sciences and Engineering, University of Quintana Roo (UQROO), 77019 Chetumal, QROO, Mexico*

Correspondence should be addressed to Leopoldo Estrada Vargas; lestrada@gdl.cinvestav.mx

Received 13 January 2013; Revised 27 May 2013; Accepted 27 May 2013

Academic Editor: Marcelo Moreira Cavalcanti

Statistical analysis and synthesis of self-similar discrete time signals are presented. The analysis equation is formally defined through a special family of basis functions of which the simplest case matches the Haar wavelet. The original discrete time series is synthesized without loss by a linear combination of the basis functions after some scaling, displacement, and phase shift. The decomposition is then used to synthesize a new second-order self-similar signal with a different Hurst index than the original. The components are also used to describe the behavior of the estimated mean and variance of self-similar discrete time series. It is shown that the sample mean, although it is unbiased, provides less information about the process mean as its Hurst index is higher. It is also demonstrated that the classical variance estimator is biased and that the widely accepted aggregated variance-based estimator of the Hurst index results biased not due to its nature (which is being unbiased and has minimal variance) but to flaws in its implementation. Using the proposed decomposition, the correct estimation of the *Variance Plot* is described, as well as its close association with the popular *Logscale Diagram*.

1. Introduction

In the past decades, the self-similar processes and long-range dependence (LRD or long memory) have been applied to the study and modeling of many natural and man-made complex phenomena. These kinds of processes have been particularly attractive in the pursuit of optimal design and configuration of network communications.

The published work of Leland et al. in 1993 and 1994 [1, 2] demonstrated that Ethernet traffic is statistically self-similar and that the commonly used models are unable to capture that fractal behavior, highlighting that a *burstiness* and LRD are present when $H > 0.5$. Since then, researchers have been studying extensively long memory processes and their impact on network performance, for example, Karagiannis et al. stated that the identification of LRD is not trivial and that not all scenarios in modern networks present LRD characteristics, for example, traffic in the Internet backbone is more likely to be Poisson type instead of LRD [3].

Many researchers have also addressed their studies to determine if network traffic is sufficiently modeled by self-similar processes or a more general model is needed, for example, one that considers *multiscaling* or *multifractality* [4–6]. The advantage of the capability to model complex systems with self-similar processes is that the correlation structure is defined by a single parameter: the Hurst index (H).

Unlike other statistics, the Hurst index, although it is mathematically well defined, cannot be estimated unambiguously from real world samples. Several methods have been developed then in order to estimate it. Examples of classical estimators are those based on R/S statistic [7] (and its unbiased version [8]), detrended fluctuation analysis (DFA) [8, 9], maximum likelihood (ML) [10], aggregated variance (VAR) [7], wavelet analysis [11, 12], and so forth. In [13], Clegg developed an empirical comparison of estimators for data in raw form and corrupted. An important observation is that the estimation of the Hurst index may differ from one estimator to another, and the selection of the most adequate estimator is a difficult task. This selection depends greatly on how well the data sample meets the assumptions the estimator is based on. However, through analytical and empirical studies, it has been discovered that the estimators that have the best performance in bias and standard deviation, and, consequently, in mean squared error (MSE) are Whittle ML

and the wavelet-based estimator proposed by Veitch and Abry in [11].

From these two estimators, the wavelet based is computationally simpler and faster [7, 11].

In addition to the Hurst index, other statistical characteristics are needed to describe the phenomenon under study. The most common are the first and second-order statistics, that is, mean, variance, and correlation. The classical estimators of these characteristics have been proposed decades ago, for example, Kenney and Keeping demonstrated in 1939 and 1951 that the classical variance estimator is unbiased for independent and identically distributed Gaussian observations [14, 15]. Confidence interval is also given for these estimations, for example, $P[\overline{X} \in (-3/\sqrt{N}, 3/\sqrt{N})] \approx 99\%$ (where $\overline{X}$ is the sample mean estimated from a sample of size N) for a standardized white noise process. This confidence interval is narrower as the sample size increases.

It has been claimed that most processes satisfy those common assumptions [16]. As it has been expressed, however, other authors (Leland et al. [1], Taqqu et al. [5], Tsybakov and Georganas [17], Veitch and Abry [11], and many others) conclude that traffic characteristics present correlation and that the estimation of these statistics (including confidence interval) with the classical estimators (which do not consider correlation) may lead to estimation errors and, consequently, to wrong decisions or inaccurate models, especially when data presents accentuated LRD.

Several estimators of the Hurst index have been proposed, but many of them do not consider the effect of correlation on the estimation of first- and second-order statistics, thus applying incorrectly the classical formulae. Particularly, it has been claimed that the aggregated variance method can only be used as a heuristic method, and that the *Variance Plot* (also named *Variance-time Plot*; see Section 2.3) can only be used to check whether the time series is self-similar or not and, if so, to obtain a crude guess for the Hurst index [22, page 44]. This work clarifies this point, demonstrating that the *Variance Plot* can be estimated efficiently and that the estimation of the Hurst index from it is actually unbiased and has minimum variance (similarly to the wavelet-based estimator).

This work is motivated by the mentioned importance of the self-similar processes in many areas, especially in the analysis and modeling of Internet traffic, and by the fact that there are still some misunderstandings and bad practices that must be overcome.

This document is organized as follows. Section 2 defines the discrete time self-similarity and some of its statistical properties. Section 3 describes the proposed set of basis functions and the analysis and synthesis equations, which are used to generate self-similar samples whose Hurst index matches the estimator proposed by Veitch et al. and the corrected version of the variance-based estimator. Section 4 defines statistics for the sample mean and variance of self-similar discrete processes; Section 5 explains how the variance-based estimator has been misunderstood and defines the correct estimator, which coincides with the Haar wavelet based in a particular case. Section 6 presents some simulations and measurements and, finally, Section 7 concludes the work.

2. Self-Similarity

Self-similarity describes the phenomenon where certain properties are preserved irrespective of scaling in space or time. Deterministic self-similarity is clearly exemplified by popular figures as Sierpinski's triangle or Koch's snow flake. This form of self-similarity is named scale invariance and makes different scales of the same object undistinguishable. Stochastic self-similarity is not that obvious, it refers to how statistical properties of a stochastic process are preserved under time expansion. Stochastic self-similarity is defined for continuous and discrete time stochastic processes.

Self-similarity (either continuous or discrete time) is tightly related to short- and long-range dependencies (SRD and LRD, resp.). The degree of self-similarity is defined and measured through the so named Hurst index H ($0 < H < 1$). It is known that processes with $H < 0.5$ are SRD, and processes with $H > 0.5$ are LRD. If $H = 0.5$, neither SRD nor LRD are present. For example, the commonly used white Gaussian noise (WGN) has always $H = 0.5$ and does not present any time dependency.

Processes with LRD are also named long memory, as current and future realizations of these are strongly correlated. The dividing line between SRD and LRD processes is not ambiguous; for LRD processes, the autocovariance function is not absolutely convergent (i.e., the sum is not finite), while it is for SRD processes. This work refers only to stochastic discrete time self-similarity.

2.1. Discrete Time Self-Similarity. The definition of discrete time stochastic self-similarity is given in terms of the aggregated processes. Let $\{X_t; t \in \mathbb{N}\}$ be a discrete time series derived from a self-similar process with stationary increments and Hurst index H (H-SSSI). The aggregated time series, derived from X_t is the sequence given by [17]

$$X(m) = \left\{X_k^{(m)};\ k \in \mathbb{N}\right\}, \tag{1}$$

where each term $X_k^{(m)}$ is defined as

$$X_k^{(m)} = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i, \quad k \in \mathbb{N}, \tag{2}$$

where m represents the aggregation level. That is, each new time series is obtained by partitioning the original time series into nonoverlapping blocks of size m and then averaging each block to obtain its respective values.

Let X_t be a covariance stationary discrete time series with mean $\mu_X = 0$, variance σ_X^2 and autocovariance function (ACvF) $\gamma_X(k)$ and $X_k^{(m)}$ its aggregated series. Then it is said that X_t is self-similar (H-SS), if the following holds [18]:

$$X_k^{(m)} \sim m^{H-1} X_t, \tag{3}$$

where $\sim$ means equality in distribution.

The many methods for generating artificial discrete time self-similar sequences are classified in *sequential* and *fixed length*. All these have particular advantages and shortcomings associated to accuracy, generation time, memory or

processing resources, and so forth. Sequential generators are sometimes more appropriate for long duration simulations, but the level of approximation and several parameters are needed, while fixed length only depends on the desired Hurst index but may need to store part or all of the sequence in memory. For the purpose of the simulations described in Section 6, a fixed-length Davies-Harte *fractional Gaussian noise* (FGN) generator is used. FGN is a special type of noise where the autocovariance function has a special shape (expression (8)). Section 3.2 describes a theoretical method to synthesize self-similar discrete sequences using Haar wavelet-based decomposition. Although Daubechies wavelet- (DW-) based approximations are better in mean and variance, the Haar-based method generates a self-similar sequence of specified Hurst index from practically any given sequence, regardless of whether it is self-similar or not if some conditions are met [19]. For other applications, DW are preferred; see [20] as an example.

2.2. Properties of Self-Similar Discrete Time Series. The definition of discrete stochastic self-similarity (3) has important implications about the stochastic process X_t; these implications include the following properties [21].

(i) Zero-mean:

$$E\left(X_t\right) = E\left(X_k^{(m)}\right) = 0. \tag{4}$$

(ii) Power law of the qth order moments:

$$E\left[\left(X_k^{(m)}\right)^q\right] = m^{q(H-1)} E\left[\left(X_t\right)^q\right]. \tag{5}$$

(iii) Power law of the qth order absolute moments:

$$E\left(\left|X_k^{(m)}\right|^q\right) = m^{q(H-1)} E\left(\left|X_t\right|^q\right). \tag{6}$$

2.3. Second-Order Discrete Self-Similarity. The second-order definition of self-similarity is derived from (5) for $q = 2$. The variance of the aggregated time series is defined by the following [17]:

$$\operatorname{var}\left(X_k^{(m)}\right) = m^{2H-2} \operatorname{var}\left(X_t\right). \tag{7}$$

An equivalent definition is

$$\gamma_X^{(m)}(k) = \frac{\sigma_X^2}{2}\left[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}\right], \quad k \geq 0, \tag{8}$$

where $\gamma_X^{(m)}(k)$ is the autocovariance function of $X_k^{(m)}$.

If a discrete time series X_t satisfies these conditions, it is called second-order self-similar with Hurst index H (H-SOSS). Note that the mean of an H-SOSS process is not necessarily zero.

The plot $\log[\operatorname{var}(X_k^{(m)})]$ versus $\log(m)$ is known as *Variance Plot*. It is a straight line of slope $2H - 2$ for self-similar processes. This plot is the basis of the variance-based estimator of the Hurst index. It has been "shown" in the literature that the variance-based estimator underestimates the Hurst index and that the variance-based estimator throws a coarse estimation of the true Hurst index. Section 5 demonstrates that this is a consequence of inadequate implementations of this estimator, that is, the aggregated variance is estimated with the classical formula (36), which is not correct if any correlation exists. An apparent solution is to use the proposed unbiased estimator, but that leads to an ill-conditioned problem: the Hurst index is needed to estimate the variance and vice versa. The solution for this situation is also described (see Section 5.1).

2.4. Wavelet Decomposition and the Logscale Diagram. The wavelet decomposition transforms a signal X_t into a sum of orthogonal components as follows:

$$X_t = \sum_{j=i}^{J} \sum_{k=1}^{2^j} d_X(j,k)\, \psi_{j,k}(t), \tag{9}$$

where each function $\psi_{j,k}(t)$ is derived from a basis function $\psi_0(t)$, namely, the mother wavelet, by scaling and displacement, that is,

$$\psi_{j,k}(t) = 2^{-j/2} \psi_0\left(2^{-j}t - k\right), \tag{10}$$

and coefficients $d_X(j,k)$ is the value at time k of scale j, computed as an inner product between the signal X_t and the wavelet function $\psi_{j,k}(t)$:

$$d_X(j,k) = \left\langle X(t), \psi_{j,k}(t) \right\rangle. \tag{11}$$

The statistic $s_2(j)$ is then defined from these coefficients as

$$S_2(j) = E\left|d_X(j,\cdot)\right|^2, \tag{12}$$

which, for an H-SOSS process, is related to the Hurst index as

$$S_2(j) = c_f C 2^{j(2H-1)}, \tag{13}$$

where the quantity $c_f C$, related to the power of the process, is considered a constant.

The plot $\log_2 S_2(j)$ versus j forms the widely known *Logscale Diagram* described by Veitch and Abry [11]. The *Logscale Diagram* of an H-SOSS process is a straight line of slope $2H - 1$. To obtain an unbiased estimation of the Hurst index based on the *Logscale Diagram*, it is also necessary to subtract the bias that results of averaging the logarithms of the respective variance of real world time series, which is estimated as [11]

$$g_j = \frac{\Psi\left(n_j/2\right)}{\ln(2)} - \log_2\left(\frac{n_j}{2}\right), \tag{14}$$

where n_j is the number of coefficients available at octave, that is

$$E\left\{\log_2\left[\operatorname{var}\left(\widehat{C}_{X,t}^{n,i}\right)\right]\right\} \approx E\left\{\log_2\left[\operatorname{var}\left(\widehat{C}_{X,t}^{n,i}\right)\right]\right\} - g_j. \tag{15}$$

If the *Logscale Diagram* of a time series cannot be adequately modeled with a linear model, possibly the scaling behavior needs to be described with more than one scaling parameter, that is, the Hurst parameter is not adequate (or insufficient) [22, 23]. However, even if the time series under study is not self-similar, the *Logscale Diagram* can show whether or not it presents LRD.

3. An Orthogonal Decomposition Performed by Subtracting Aggregated Series

The time series X_t can be decomposed into a set of time series, each one defined as

$$C_{X,t}^{n,j} = X_t^{(n^{j-1}E)} - X_t^{(n^jE)}, \quad n, j \in \mathbb{N}, \tag{16}$$

where $X_t^{(n^jE)}$ is the time series X_t after two operations, which are as follows.

(1) Aggregation at level n^j, as defined by (1) and (2), that is, $m = n^j$.

(2) Expansion of level n^j, which consists of "repeat" each element of a time series n^j times, that is, $X_t^{(n^jE)} = X_k^{(n^j)}$ for $k = 1 + \lfloor (t-1)/n^j \rfloor$ and $j \in \mathbb{N}$.

These zero-mean components $C_{X,t}^{n,j}$ have three important properties.

(i) They synthesize the original time series without loss (assuming zero-mean), that is,

$$X_t = \sum_j C_{X,t}^{n,j}. \tag{17}$$

(ii) They are pair-wise orthogonal:

$$\left\langle C_{X,t}^{n,j_1}, C_{X,t}^{n,j_2} \right\rangle = 0, \quad j_1 \neq j_2. \tag{18}$$

(iii) If X_t is exactly or at least second-order self-similar, then the variance of its components satisfies

$$\operatorname{var}\left(C_{X,t}^{n,j}\right) = r \cdot \operatorname{var}\left(C_{X,t}^{n,j-1}\right), \tag{19}$$

where

$$r = n^{2H-2}. \tag{20}$$

Another useful property relates the variance of the component to the variances of the aggregated series, that is,

$$\operatorname{var}\left(C_{X,t}^{n,j}\right) = \operatorname{var}\left[X_t^{(n^{j-1}E)}\right] - \operatorname{var}\left[X_t^{(n^jE)}\right]. \tag{21}$$

It is easy to proof (21): from (16), it turns out that $\operatorname{var}[C_{X,t}^{n,j}] = \operatorname{var}[X_t^{(n^{j-1}E)}] + \operatorname{var}[X_t^{(n^jE)}] - 2\operatorname{cov}[X_t^{(n^{j-1}E)}, X_t^{(n^jE)}]$, but as $X_t^{(n^jE)}$ is itself an aggregation of $X_t^{(n^{j-1}E)}$ it turns out that $\operatorname{cov}[X_t^{(n^{j-1}E)}, X_t^{(n^jE)}] = \operatorname{var}[X_t^{(n^jE)}]$, and (21) comes after a substitution.

Properties (i), (ii), and (iii) imply that

$$\begin{aligned} \sigma_X^2 &= \sum_i \operatorname{var}\left(C_{X,t}^{n,j}\right), \\ \sigma_X^2 &= \frac{1}{1-r}\operatorname{var}\left(C_{X,t}^{n,1}\right). \end{aligned} \tag{22}$$

Then, the variance of the jth component is related to the variance of X_t as

$$\operatorname{var}\left(C_{X,t}^{n,j}\right) = (1-r)\, r^{j-1}\sigma_X^2. \tag{23}$$

It is easy to prove the following relation:

$$\operatorname{var}\left(C_{X,t}^{n,j}\right) = n^{-j} S_2(j). \tag{24}$$

An immediate consequence of (24) is that the plot $j + \log_n[\operatorname{var}(C_{X,t}^{n,j})]$ versus j is equivalent to the *Logscale Diagram*. It is a straight line for H-SOSS time series, and the slope is related to the Hurst index so that $s = 2H - 1$.

For example, for $n = 2$ the statistics $\operatorname{var}(C_{X,t}^{2,j})$ and $S_2(j)$ are related as

$$\operatorname{var}\left(C_{X,t}^{2,j}\right) = 2^{-j} S_2(j), \tag{25}$$

and in this case the basis function is the Haar wavelet, which is defined as [24, 25]

$$\psi_0(t) = \begin{cases} +1, & 0 \le t < \dfrac{1}{2}, \\ -1, & \dfrac{1}{2} \le t < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

Figure 1 shows the components obtained from an H-SOSS sample of size 32 and $H = 0.9$. The squared form of the components is due to the expansion of the aggregated series and disappears after the downsampling.

The authors of [26] described the aggregation as an inner product with the signal and the Haar "father" wavelet, and then, the relation between wavelet coefficients and aggregation levels is obvious. At this point there are similarities between this section and that previous work, the most important is that the relation between aggregation levels and Haar wavelet is described by Abry et al. However, two differences must be highlighted: (1) the decomposition presented in this work only coincides with that definition for $n = 2$ in (16); for higher values, the Haar wavelet is not sufficient to describe the components of (16); and (2) authors of [26] discard anyway the estimation of the Hurst index based on the so named "a-aggregation." In this work, it is clarified that the "a-aggregation" is misunderstood leading to incorrect implementations (i.e., the "classical" variance estimator), and a generalization of the "d-aggregation" is proposed.

3.1. General Waveform of the Basis Functions. The orthogonal decomposition defined by (16) can be expressed in terms of an inner product between the signal and a set of orthogonal wavelet-type functions. Let us describe the waveform for the general case.

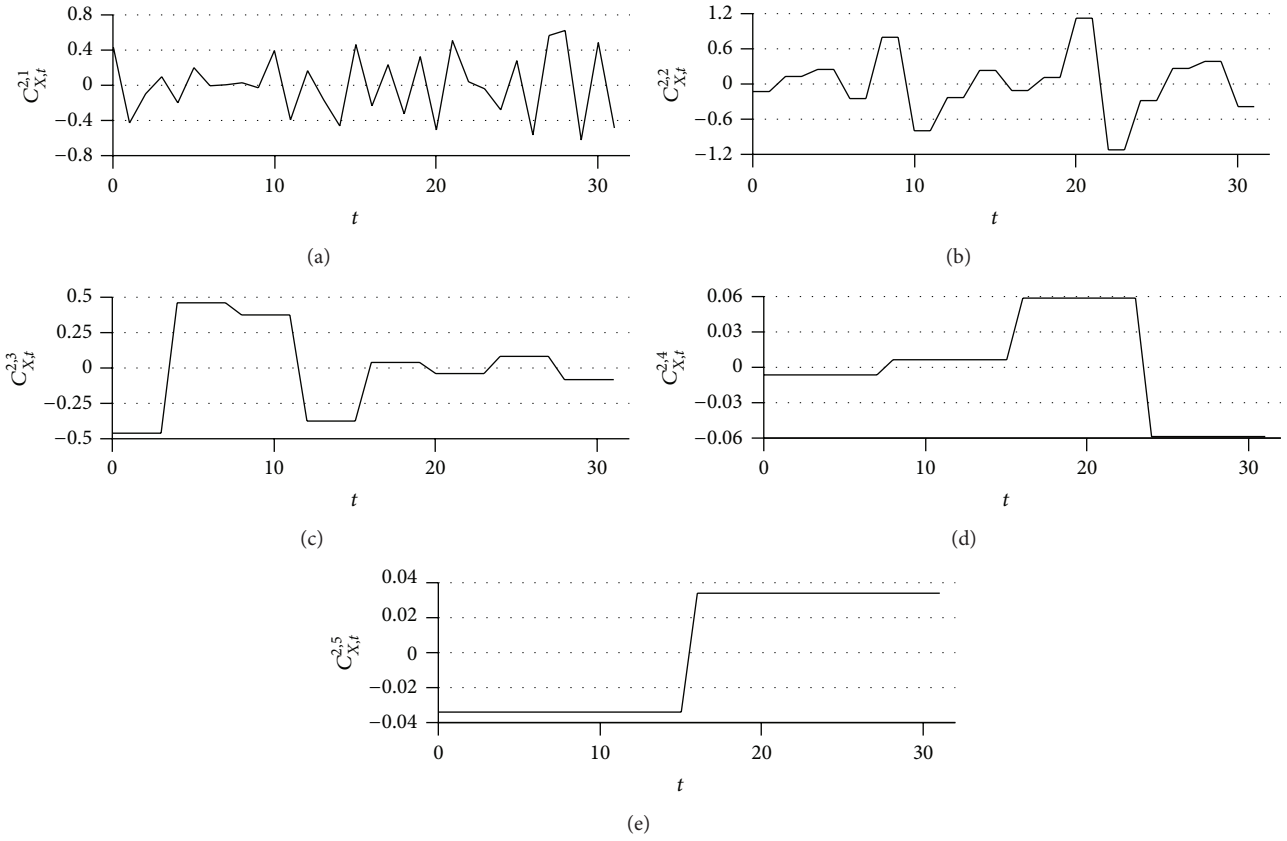

Figure 1: Components of an H-SOSS sample of size 32 and $H = 0.9$.

The wavelet-based function is

$$\psi_n(t) = \begin{cases} 1 - \dfrac{1}{n}, & 0 \le t < \dfrac{1}{n}, \\ -\dfrac{1}{n}, & \dfrac{1}{n} \le t < 1, \\ 0, & \text{otherwise}, \end{cases} \tag{27}$$

and the wavelet functions derived from $\psi_n(t)$ are obtained by three operations: scaling, displacement (similarly to (10)), and a phase shift, that is,

$$\begin{aligned} \psi_{n,j,k,\theta}(t) &= \left\{1 - u\left[n^j(k-1)\right]\right\} \psi_{n,j,k}\left(t - \theta n^{j-1}\right) \\ &\quad + u\left[n^j(k-1)\right] \psi_{n,j,k}\left(t + n^j - \theta n^{j-1}\right), \\ \psi_{n,j,k}(t) &= n^{-j/2} \psi_n\left(n^{-j}t - k\right), \end{aligned} \tag{28}$$

for $j = 1, 2, \ldots, J$, $k = 0, 2, \ldots, n^{J-j} - 1$, and $\theta = 0, 1, 2, \ldots, n-1$. Note that note that $\psi_{n,0,0,0}(t) = \psi_n(t)$.

The function defined by (27) is a generalized form of the Haar wavelet. It is always a rectangle-shaped function, but it is not symmetric about the horizontal axis except for $n = 2$.

Figure 2 shows the basis function for $n = 2$ without phase shift and with a phase shift of 1/2, respectively. Obviously, $\langle X(t), \psi_{2,j,k,0}(t)\rangle = -\langle X(t), \psi_{2,j,k,1}(t)\rangle$, which means that both products give the same information. Redundant information can then be reduced by decimating the sequence of coefficients.

Figure 3 shows the basis function for $n = 3$. Note that the phase shift moves the rectangle of height 2/3 from one-third to another. In this case, there exists also redundant information, as $\psi_{3,0,0,2}(t) = -\psi_{3,0,0,0}(t) - \psi_{3,0,0,1}(t)$ and in the general case $\psi_{n,j,k,\theta}(t) = -\sum_{i \ne \theta} \psi_{n,j,k,i}(t)$, which means that one of the coefficients, for example, the one obtained with the last phase shift, can be discarded. This means that the sequence of coefficients can be downsampled without loss of information. For a sample of length N, it is easy to verify that the number of observations that remains in all components (sequences of coefficients) after the downsampling is $N - 1$, which can be complemented with the sample mean, in the case that this is not zero.

3.2. Wavelet Synthesis of Self-Similar Time Series. A method to synthesize H_1-SOSS from practically any time series, regardless of whether it is or not self-similar or its marginal distribution, is proposed. This method consists of adjusting the sum expressed by (9) with a set of weights, that is,

$$X_{H_1}(t) = \sum_{j=i}^{J} w_j \sum_{k=1}^{2^j} d_X(j,k)\, \psi_{j,k}(t), \tag{29}$$

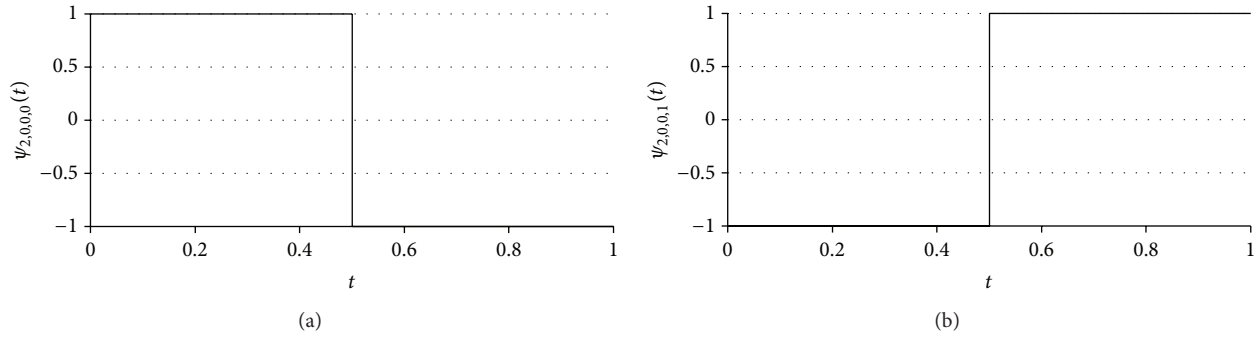

FIGURE 2: Basis functions: (a) $\psi_{2,0,0,0}(t)$ (no phase shift) and (b) $\psi_{2,0,0,1}(t)$ (equal to $\psi_{2,0,0,0}(t)$ with a phase shift of 1/2).

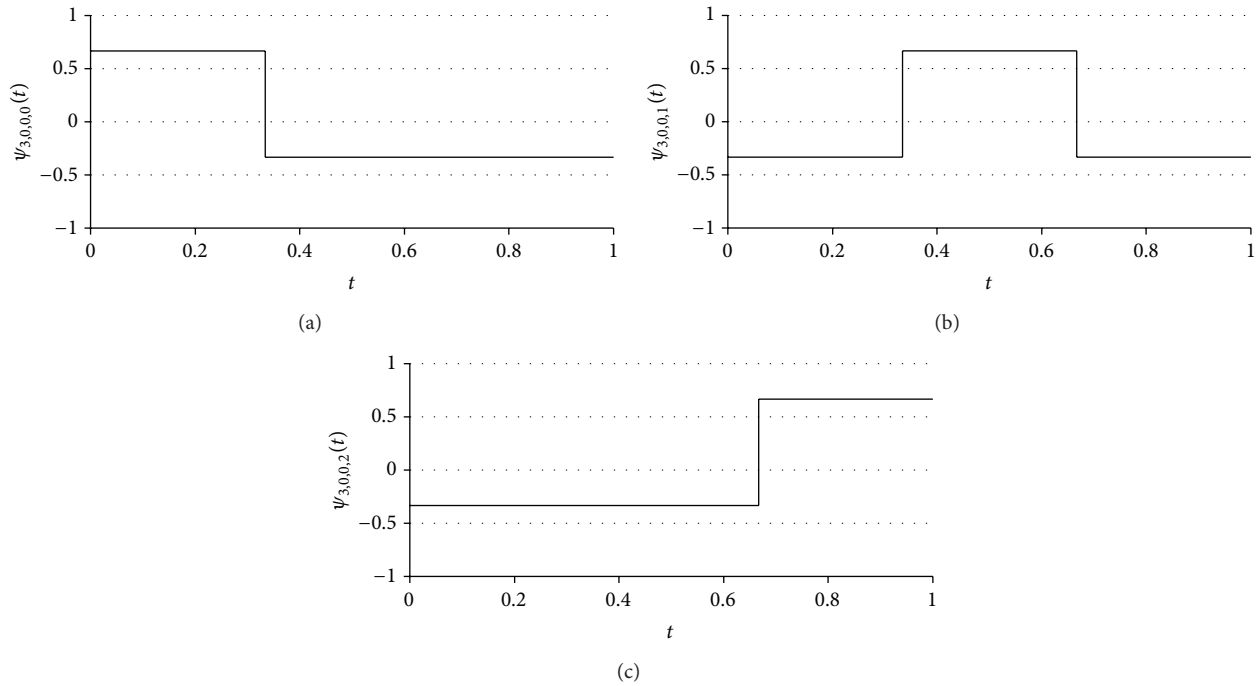

FIGURE 3: Basis functions: (a) $\psi_{3,0,0,0}(t)$ (no phase shift), (b) $\psi_{3,0,0,1}(t)$ ($\psi_{3,0,0,0}(t)$ with a phase shift of 1/3), and (c) $\psi_{3,0,0,2}(t)$ ($\psi_{3,0,0,0}(t)$ with a phase shift of 2/3).

where these weights w_j are defined as

$$w_j = \sqrt{\frac{\widehat{c_f C} \cdot 2^{j(2H_1-1)}}{\widehat{S}_2(j)}}, \tag{30}$$

where $\widehat{S}_2(j)$ and $\widehat{c_f C}$ are the respective estimations of $S_2(j)$ and $c_f C$ (the associated power parameter [4]) from X_t and H_1 is the desired Hurst index of the new synthetic series. It is necessary that $\widehat{S}_2(j) > 0$ for all $j = 1, \ldots, J$.

The weighted sum (29) can also be expressed in terms of the orthogonal components described in Section 3 and defined by (16) as follows:

$$X_t = \sum_{i=1}^{L} w_i C_{X,t}^{n,i}, \tag{31}$$

where the weights w_i are computed as

$$w_i = \sqrt{r^{i-1} \cdot \frac{1-r^L}{1-r} \cdot \frac{\widehat{\sigma}_X^2}{\operatorname{var}\left\{\widehat{C}_{X,t}^{n,i}\right\}}}, \tag{32}$$

where r is defined by (20). Note that the only restriction of (32), similarly to (30), is that the estimated variance $\widehat{C}_{X,t}^{n,i}$ must be nonzero. Even though this synthesis does not depend either on the marginal distribution or the correlation structure of the input signal, it is preferable that this is self-similar (e.g., FGN) and that its Hurst index is close to that of $X_{H_1}(t)$, that is, $H_1 \approx H$.

Pathological behavior can be produced in the output series for some critical conditions, for example, noticeable steps, which may produce a nonstationary signal, result from transforming an SRD, or uncorrelated input signal to an LRD

output with H close to 1. Impulses of very large magnitude (outliers) can be also be produced when, for some j, $\widehat{S}_2(j)$ is close to zero.

A similar methodology was developed by Deléchelle et al. [27], to synthesize fractional Gaussian noise by performing a weighted sum of the intrinsic mode functions (IMF) of a white noise process. The advantages of the proposed synthesis compared to that of Deléchelle et al. are that the components defined by (16) are exactly orthogonal, the relation between the weights for the reconstruction sum are mathematically well defined, and the Hurst index of the synthesized time series matches perfectly the wavelet estimator proposed by Abry et al. [4], that is, the estimated H of the synthesized series is unbiased ($E(\widehat{H}) = H$) and has zero variance ($\text{var}(\widehat{H}) = 0$). The disadvantages are that the components are sequences of squared signals (because of the expansion described in Section 3) and not sinusoids, and noticeable steps arise when synthesizing a time series with high Hurst index, for example, close to 1, from an input that is SRD or weakly correlated. A solution for this problem is to apply interpolation (as in EMD) instead of expanding the series in order to produce softer components (sinusoids or polynomial) instead of square type, with the consequence that the Hurst index is no longer exact, but approximated.

4. Estimation of Mean and Variance Self-Similar Time Series

4.1. Sample Mean. The sample mean of a self-similar process is unbiased: its expected value is the process mean, that is, $E(\overline{X}) = \mu_X$, where $\overline{X} = 1/N \sum_{t=1}^{N} X_t$, regardless of the presence of correlation between observations. However, its variance does not depend only on the sample size (N) but also on the degree of self-similarity (H) of the process as follows:

$$\text{var}\left(\overline{X}\right) = \sigma_X^2 N^{2H-2}, \tag{33}$$

which becomes

$$\text{var}\left(\overline{X}\right) = \frac{\sigma_X^2}{N}, \tag{34}$$

(classical estimator) for $H = 0.5$ (uncorrelated observations). Figure 4 shows the *probability distribution function* (PDF) of the sample mean of standardized FGN.

To derive (33), consider that $\overline{X}$, estimated from a sample of size N, behaves exactly the same as the stationary aggregated process $X_k^{(N)}$, defined by (1), and its variance is determined by the definition of second-order self-similarity (7). Expression (33) can be also derived (for $H > 0.5$) from the autocorrelation coefficient $\rho(k) = 0.5[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}]$ for $k \geq 1$ ($\rho(0) = 1$) and $\text{var}(\overline{X}) = (\sigma_x^2/N^2) \sum_{i,j=1}^{N} \rho(k)$.

Important implications of (33) about the uncertainty of the mean are (1) that it increases with the Hurst index, for example, $\text{var}(\overline{X})$ tends to σ_X^2 as H tends to 1, which makes the sample mean worth a single observation and (2) that it cannot be zero for any case when estimated from a finite-size sample.

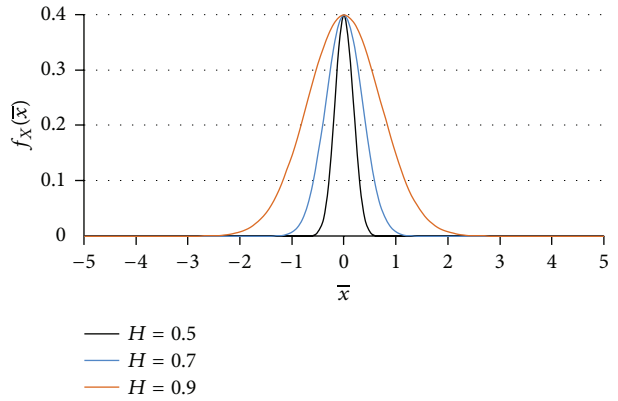

Figure 4: Distribution of the sample mean of standardized fractional Gaussian noise processes with $H = \{0.50, 0.70, 0.90\}$ and $N = 32$.

4.2. Sample Variance. For a high number of observations, sample variance is usually calculated as

$$\widehat{\sigma}_X^2 = \frac{1}{N} \sum_{t=1}^{N} (X_t - \widehat{\mu}_X)^2. \tag{35}$$

It is known that estimator (35) is biased, so for small samples, it is more adequate to use

$$\widehat{\sigma}_X^2 = \frac{1}{N-1} \sum_{t=1}^{N} \left(X_t - \overline{X}\right)^2. \tag{36}$$

Particularly, if the observations X_t are independent and come from a normal distribution, $\widehat{\sigma}_X^2$ is distributed as $\widehat{\sigma}_X^2 \sim (\sigma_X^2/N)\chi_{n-1}^2$, as stated by Cochran's theorem [28, 29]. Formula (36) is the most used estimator of the sample variance [14] but, as Beran indicates in [30], it is needed to know which assumptions this estimator is based on in order to apply it correctly; otherwise, it may be the source of errors that in practice cannot be negligible for all cases.

A self-similar process is uncorrelated only and only if the Hurst index is 0.5. In this particular case, the classical estimator of sample variance, defined by (36), is unbiased.

The expected value of the sample variance defined by (36) is

$$E\left(\widehat{\sigma}_X^2\right) = E\left[\frac{1}{N-1} \sum_{t=1}^{N} (X_t - \widehat{\mu})^2\right], \tag{37}$$

which can be expressed as

$$E\left(\widehat{\sigma}_X^2\right) = \frac{N}{N-1} \left[E\left(X_t^2\right) - E\left(\widehat{\mu}^2\right)\right], \tag{38}$$

then,

$$E\left(\widehat{\sigma}_X^2\right) = \frac{N}{N-1} \left\{\text{var}\left(X_t\right) + \left[E\left(X_t\right)\right]^2 - \left\{\text{var}\left(\widehat{\mu}\right) + \left[E\left(\widehat{\mu}\right)\right]^2\right\}\right], \tag{39}$$

$$E\left(\widehat{\sigma}_X^2\right) = \frac{N}{N-1} \sigma_X^2 \left(1 - N^{2H-2}\right).$$

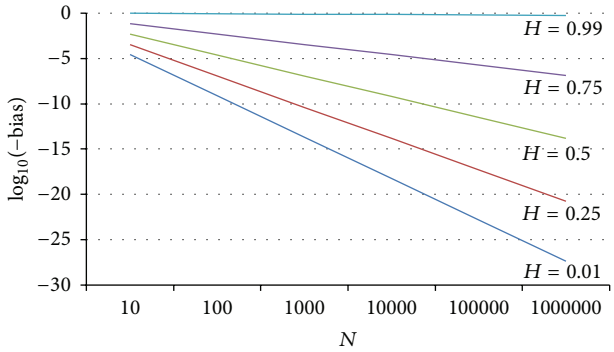

Figure 5: Logarithm of the variance bias for different sample sizes.

Expression (39) proves that the classical estimator (36) is biased (i.e., $E(\widehat{\sigma}_X^2) \neq \sigma_X^2$) for $H \neq 0.5$. It is straightforward that the unbiased variance estimator for self-similar processes is then

$$\widehat{\sigma}_X^2 = \frac{1}{N - N^{2H-1}} \sum_{t=1}^{N} \left(X_t - \overline{X}\right)^2, \tag{40}$$

which obviously becomes (36) for $H = 0.5$.

A plot of $\log_{10}(\widehat{\sigma}_X^2 - \sigma_X^2)$ versus N is shown in Figure 5. Note that, for a fixed sample size, as H increases the estimation of the variance by means of the classical estimator, (36) becomes less significant.

Figure 5 exemplifies that as the sample size is greater, the classical variance estimator has less bias; however, as the Hurst index of the process is greater, the bias is also greater (in magnitude). Note that as H approaches to 1, the variance is considerably underestimated, which makes the classical variance estimator useless.

The variance of the estimated variance of a self-similar time series can be approximated by applying the formula proposed by Yunhua in [31] for $k = 0$, that is,

$$\operatorname{var}\left(\widehat{\sigma}_X^2\right) = \frac{\left(2N^{4H-3} + 8H - 7\right)\left(2H^2\right)(2H-1)^2}{N(4H-3)} + \frac{1}{N}. \tag{41}$$

This approximation is close to the variance of $\widehat{\sigma}_X^2$, with the disadvantage that it has a discontinuity in $H = 0.75$. Further work can be developed in order to verify this approximation and to quantify its error.

Let us mention that, although the proposed estimator of the sample variance is unbiased, its performance relies on the estimation of the Hurst index. This dependence is very noticeable as H approaches to 1, as the statistic $1 - N^{2\widehat{H}-2}$ is especially sensitive to the variation of $\widehat{H}$ under that condition. The derivative $d(1-N^{2\widehat{H}-2})/d\widehat{H} = -2N^{2\widehat{H}-2}\ln(N)$ versus H is shown in Figure 6. Note that as the estimated H increases, the estimation of the mean of the sample variance is more variable.

An immediate implication of this is that processes with Hurst index close to 1 must be carefully treated, as slight deviations of the Hurst index estimation derive in a nonnegligible error in the estimation of the process variance.

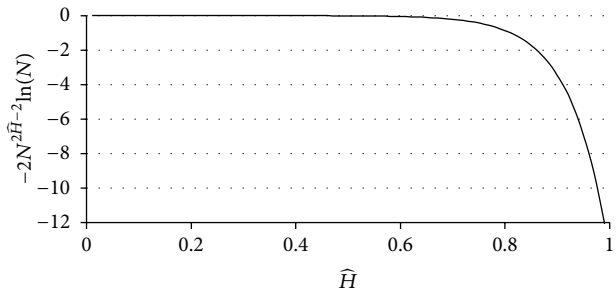

Figure 6: Plot $-2N^{2\widehat{H}-2}\ln(N)$ versus $\widehat{H}$.

4.3. Statistics of the Aggregated Process. The aggregated process $X_k^{(m)}$ (defined by (2)) derived from an H-SS process is also H-SS (self-similar with the same Hurst index). It is also true for the case of H-SOSS processes, and this aggregated process is, by definition, identically distributed to the sample mean obtained from a set of $N = m$ observations (m is the aggregation level, as in (2)), that is, $X_k^{(m)} \sim \widehat{\mu}$. Also, the aggregated sample variance obtained with the classical estimator (36) is biased, as it is well known [32].

The variance of the aggregated series of an H-SOSS process must be then estimated with the unbiased formula (40) adapted to the number of observations in the sample, that is,

$$\operatorname{var}\left(\widehat{X}_k^{(m)}\right) = \frac{1}{N_i - N_i^{2H-1}} \sum_{k=1}^{N_i} \left(X_k^{(m)} - \overline{X_k^{(m)}}\right)^2, \tag{42}$$

where N_i is the size of the series after aggregation (i.e., $N_i = N/m$). Note that the estimation of the sample mean from the aggregated sample is also unbiased, that is, $E[\overline{X_k^{(m)}}] = E(X_t)$, and it is more reliable than the mean estimated from a sample of the same size, that is, $\operatorname{var}[\overline{X_k^{(m)}}] = m^{2H-2}\operatorname{var}(\overline{X})$ if the samples are of equal size.

4.4. Statistics of the Orthogonal Components. Let $\{\widehat{X}_t;\ t = 1, \ldots, N\}$ be a finite-length self-similar time series such that $N = n^J$ ($J < \infty$) and $n \geq 2$ (i.e., N is a power of n); then a set of nonzero J components ($\widehat{C}_{X,t}^{n,j};\ j = 1, \ldots, J$) can be obtained as expressed by the analysis (16). As the components are pair-wise orthogonal, the variance of $\widehat{X}_t$ ($\widehat{\sigma}_X^2$) is the sum of a finite number of variances:

$$\widehat{\sigma}_X^2 = \sum_{j=1}^{J} \operatorname{var}\left(\widehat{C}_{X,t}^{n,j}\right). \tag{43}$$

Expressions (43) and (19) imply that the variance of the ith component ($\widehat{C}_{X,t}^{n,i}$) (computed using formula (35)), which has also finite length, is

$$\operatorname{var}\left(\widehat{C}_{X,t}^{n,i}\right) = \frac{1-r}{1-r^J} r^{i-1} \widehat{\sigma}_X^2, \tag{44}$$

and estimating $\operatorname{var}(C_{X,t}^{n,j})$ as in (23), it follows that

$$\operatorname{var}\left(\widehat{C}_{X,t}^{n,i}\right) = \frac{\operatorname{var}\left(C_{X,t}^{n,j}\right)\left(1 - N^{2H-2}\right)}{1-r^J}. \tag{45}$$

Replacing r^J by N^{2H-2}, it yields

$$\operatorname{var}\left(\widehat{C}_{X,t}^{n,i}\right) = \operatorname{var}\left(C_{X,t}^{n,j}\right). \tag{46}$$

Expression (46) implies that the estimation of the variance of components is unbiased (a desirable property) and, as a consequence, so it is the estimation of the statistic $S_2(j)$. Another implication of (46) is that the estimations of the Hurst index and the power parameter c_fC from the *Logscale Diagram* are unbiased, as have previously been proven by the authors of [11].

4.5. Correlation of the Wavelet Coefficients. Let us describe the autocovariance of the Haar wavelet coefficients, that is, the case for $n = 2$. The structure of this jth component ($C_{X,t}^{2,j}$), as a function of the elements of X_t, is

$$C_{X,\tau}^{2,j} = \frac{x_{2\tau-1}^{(2^{j-1})} - x_{2\tau}^{(2^{j-1})}}{2}. \tag{47}$$

Note that $C_{X,\tau}^{2,j}$ is a downsampled version of $C_{X,t}^{2,j}$, that is, only the first observation of each 2^j of $C_{X,t}^{2,j}$ remains.

Then, two consecutive coefficients are correlated as

$$\begin{aligned}\gamma_{C_j}(k) &= E\left(C_{X,\tau}^{2,j}, C_{X,\tau+k}^{2,j}\right) \\ &= E\left(\frac{x_{2\tau-1}^{(2^{j-1})} - x_{2\tau}^{(2^{j-1})}}{2} \frac{x_{2(\tau+k)-1}^{(2^{j-1})} - x_{2(\tau+k)}^{(2^{j-1})}}{2}\right).\end{aligned} \tag{48}$$

Assuming that X_t represents a zero-mean H-SOSS process, and according to the definition (8), $E(x_{2\tau-1}^{(2^{j-1})} x_{2(\tau+k)-1}^{(2^{j-1})}) = E(x_{2\tau}^{(2^{j-1})} x_{2(\tau+k)}^{(2^{j-1})}) = \gamma_X(2k)$, $E(x_{2\tau-1}^{(2^{j-1})} x_{2(\tau+k)}^{(2^{j-1})}) = \gamma_X(2k+1)$ and $E(x_{2\tau}^{(2^{j-1})} x_{2(\tau+k)-1}^{(2^{j-1})}) = \gamma_X(2k-1)$. Then, (48) is calculated as

$$\gamma_{C_j}(k) = \frac{2^{2H-2}\sigma_X^2}{4}\left[2\rho_X(2k) - \rho_X(2k+1) - \rho_X(2k-1)\right], \tag{49}$$

and the correlation coefficient of the jth component is then

$$\rho_{C_j}(k) = \rho_C(k) = \frac{\left[\rho_X(2k+1) - 2\rho_X(2k) + \rho_X(2k-1)\right]}{2^{2H} - 4}. \tag{50}$$

As (50) shows, the correlation structure is the same for all components, that is, $\rho_{C_j}(k)$ is independent of j. Other implications of (50) are that

$$\begin{aligned}\sum_{k=0}^{\infty} &\rho_{C_j}(k) \\ &= 1 + \frac{-\rho_X(1) - \rho_X(2k+1) + 2\sum_{i=1}^{2K+1}\left[(-1)^{i-1}\rho_X(i)\right]}{2^{2H} - 4},\end{aligned} \tag{51}$$

which can be approximated as

$$\sum_{k=0}^{\infty} \rho_{C_j}(k) \approx -0.052H^2 - 0.311H + 1.168 < \infty, \tag{52}$$

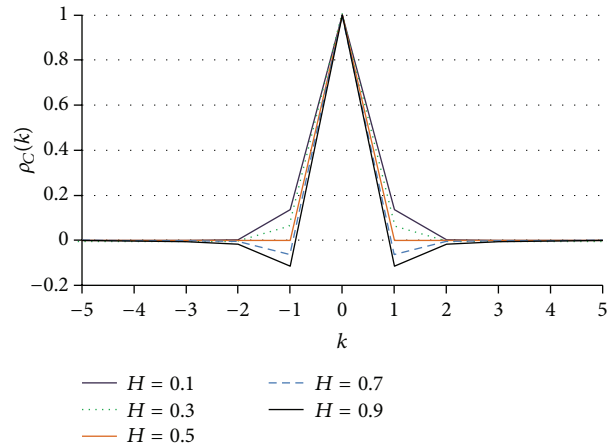

Figure 7: Correlation coefficient of the Haar wavelet components.

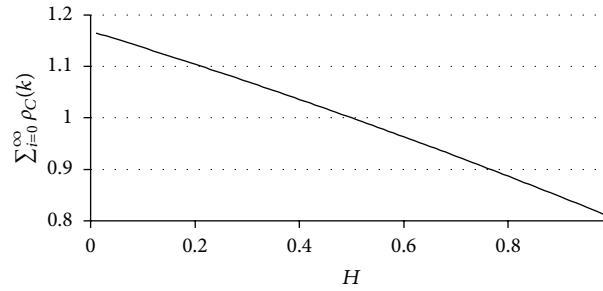

Figure 8: Sum of correlation coefficients of the Haar wavelet components.

that is, the coefficients are weakly correlated and the sum of correlations is finite [33]; furthermore, none of the components (sequences of wavelet coefficient) can be a self-similar time series except for $H = 0.5$, for example, components of a white noise process are also white noise processes. A plot of $\rho_{C_j}(k)$ versus k is shown in Figure 7 for $H = \{0.1, 0.3, 0.5, 0.7, 0.9\}$. The sum $\sum_{k=0}^{\infty} \rho_{C_j}(k)$ versus H is shown in Figure 8.

As the maximum magnitude of $\gamma_{C1}(1)$ is 1/8 (when $H \to 0$), it is said that these coefficients are *quasiuncorrelated*. Note that $\gamma_{C1}(1) = 0$ for $H = 0.5$ and $H \to 1$.

Although the assumption that the estimation of the 1st component variance ($\operatorname{var}(\widehat{C}_{X,t}^{n,1})$) is unbiased is nearly accurate, it may not hold for components of greater order. As the wavelet coefficients $d_X(j,k)$ are almost uncorrelated, the estimation of the component variance (i.e., $\sum_{t=1}^{N/n^i} (\widehat{C}_{X,t}^{n,i})^2$) is approximately $((N - n^i)/N)\operatorname{var}(C_{X,t}^{n,i})$.

5. *Variance Plot*-Based Estimation of the Hurst Index

As described in Section 2.3, the *Variance Plot* is a straight line for self-similar time series but, as many authors have claimed, it underestimates the Hurst index when working with real world data. This is a consequence of the inadequate

estimation of the aggregated variance, caused by the application of the classical formula (36) regardless of whether the original process presents any type of correlation. The solution would be then to apply the unbiased formula (42), but it leads to an ill-conditioned problem: the Hurst index needs to be estimated and known at the same time. Then, it is not that the *Variance Plot* is not adequate to estimate the Hurst index, rather the flaw of those implementations is that the aggregated variance is underestimated.

Nevertheless, it is actually possible to estimate the Hurst index analytically from the *Variance Plot*: the key is to choose the aggregation levels so that they form a geometric series, as explained in Section 5.1. Note that a numerical method can also be applied to estimate simultaneously the *Variance Plot* and the Hurst index, but the proposed solution is computationally simpler and more efficient.

5.1. Analytical Solution to the Ill-Conditioned Problem. Let $\{m_i;\ i = 1, \ldots, M\}$ be the set of aggregation levels such that $m_i = am_{i-1} = a^{i-1}m_1$, $a, m_1 \in \mathbb{N}$ and $a > 1$, that is, the levels of aggregation follow a geometric series, for example, $\{m_i\} = \{2, 4, 8, \ldots, 2^M\}$ or $\{m_i\} = \{10, 100, 1000, \ldots, 10^M\}$, and let $\widehat{\sigma}^2_{X^{(m_i)}}$ be the variance of the aggregated series $X_k^{(m_i)}$ estimated with formula (35). Obviously, $\widehat{\sigma}^2_{X^{(m_i)}}$ is biased, as $\widehat{\sigma}^2_{X^{(m_i)}} = \sigma^2_{X^{(m_i)}}[1 - (N/m_i)^{2H-2}] = \sigma^2_X m_i^{2H-2}(1 - N^{2H-2})$, but then the difference between $\widehat{\sigma}^2_{X^{(m_i)}}$ and $\widehat{\sigma}^2_{X^{(m_{i+1})}}$ is calculated as follows:

$$\widehat{\Delta}_i = \widehat{\sigma}^2_{X^{(m_i)}} - \widehat{\sigma}^2_{X^{(m_{i+1})}} = \sigma^2_X m_1^{2H-2} a^{(i-1)(2H-2)} \left(1 - a^{2H-2}\right), \quad i = 1, \ldots, M-1, \tag{53}$$

and its logarithm

$$\log_a \Delta_i = \log_a \left[\sigma^2_X m_1^{2H-2} \left(1 - a^{2H-2}\right)\right] + (i-1)(2H-2). \tag{54}$$

Finally, the slope (s) of the plot $\log_a \Delta_i$ versus i is obtained (e.g., with a weighted least square regression) and H is estimated as $\widehat{H} = \widehat{s}/2 + 1$. It can be easily proven as $\widehat{\sigma}^2_{X^{(m_i)}} = X_t^{(n^{j-1}E)}$ and substituting it in (16) and (19).

The slope is computed by the following weighted formula:

$$\widehat{s} = \frac{\sum_{i=1}^{M-1} \left(i\widehat{\Delta}_i W_i\right) - \sum_{i=1}^{M-1} (iW_i) \cdot \sum_{i=1}^{M-1} \left(\widehat{\Delta}_i W_i\right)}{\sum_{i=1}^{M-1} (i^2 W_i) - \left[\sum_{i=1}^{M-1} (iW_i)\right]^2}, \tag{55}$$

where the weights are such that $\sum_{i=1}^{M-1}(W_i) = 1$ and they are adequate so that $\widehat{s}$ has minimal variance, for example, $W_i = W_{i-1}/m_1$.

Note that the *Variance Plot* can be estimated without bias, but the Hurt index is not estimated from it. Furthermore, if the aggregation levels are taken as $m_i = 2^i$, the estimator is exactly the same than the one that uses Haar wavelet. The authors of [34] developed an empirical study of estimation of the Hurst index from series with the presence of trends. They conclude that a method named differenced-variance (a variation of the variance-bases estimator) should not be used for estimating the Hurst index. The proposed solution is also a differenced-variance type method, but it can be used to estimate the Hurst index without bias and with optimal variance. Evidently, when working with real world traces, the *Variance Plot* may differ from the straight line, and an additional bias results from the logarithm as $E[\log(\cdot)] \neq \log[E(\cdot)]$. This bias can be subtracted analogously to (15), that is,

$$\log_a \mathrm{var}\left[X_t^{(m_i)}\right] = \log_a \mathrm{var}\left[X_t^{(m_i)}\right] - g_i, \tag{56}$$

where g_i is the bias defined in(14).

6. Simulation and Measurements

6.1. Estimation of the Sample Mean. In order to verify the equations that describe the mean and variance of the sample mean, a set of zero-mean, unitary variance, and FGN time series of size $N_p = 10^6$ observations are generated using an implementation of the generator proposed by Davies and Harte in [35], each for a different Hurst index for $H = \{0.30, 0.50, 0.70, 0.90\}$. Then, the mean is estimated from blocks of size $N = 100$ and the empirical PDF is obtained from the estimations and compared to the classical (34) and proposed (33) estimators. Figure 9(a) shows that the variance of the estimated mean does not fit the classical model when SRD or LRD is present (Figures 9(a), 9(c), and 9(d)), but only for the uncorrelated case (Figure 9(b)). Only proposed estimator (33) represents adequately this phenomenon for the four cases.

6.2. Estimation of the Sample Variance. The followed procedure to verify the proposed estimator of the sample variance ($\widehat{\sigma}^2_X$) consists of the generation of a set of 100 FGN samples of size 1024 for each value of $H = \{0.05, 0.10, 0.15, \ldots, 0.95\}$ and the estimation of the variance using the classical formula (36) and the proposed estimator (40). The respective mean of both estimations for each set was obtained. For the estimation of the Hurst index the wavelet estimator of Veitch and Abry [11] is used.

Figure 10 shows that the classical formula underestimates the variance noticeably for higher values of H. For $H > 0.95$ the estimated variance is less than half of the process variance. The proposed formula (40) does not underestimate the variance, but for high values of H the estimated variance is significantly different from one realization to another. This variation results from the estimation of the Hurst index, as the statistic $1 - N^{2\widehat{H}-2}$ is very sensitive to the variations of $\widehat{H}$, which depends, in turn, on the efficiency of the sample generator. This is an indicator that the generator proposed by Davies and Harte may be less accurate as H is closer to 1.

Figure 11 shows the variance of the estimated variance obtained with the proposed formula (40) compared to the approximation proposed by Yunhua in [31], as expected, the variance of the estimation is close to zero (lower than 0.001) for $H < 0.5$, but as the Hurst index increases, it becomes noticeable when $H > 0.75$. This verifies the observation of [31], which says that beyond $H = 0.75$ the precision of the autocorrelation is about one order lower than when $H < 0.75$.

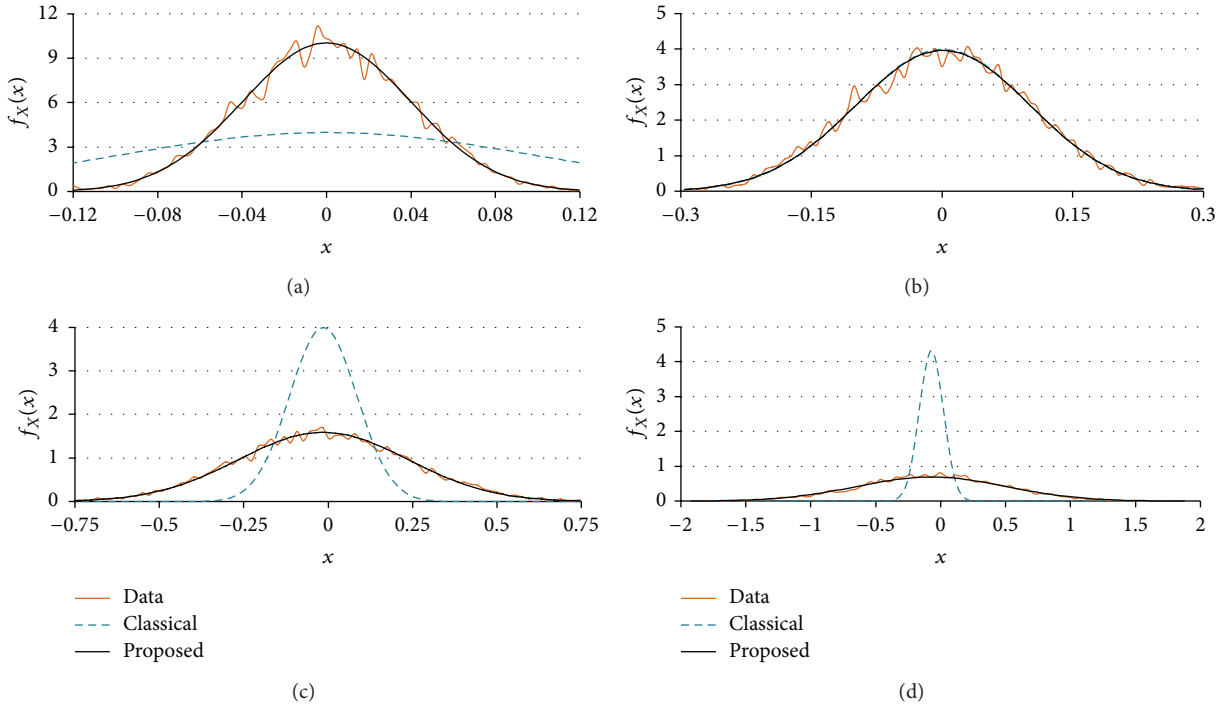

Figure 9: Estimation of the sample mean with $n = 100$ for four cases: (a) $H = 0.30$, (b) $H = 0.50$, (c) $H = 0.70$, and (d) $H = 0.90$.

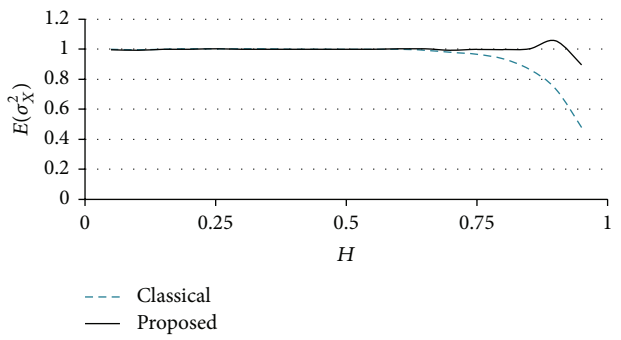

Figure 10: Mean of the estimated variance for $n = 1024$ and $H = \{0.05, 0.10, 0.15, \ldots, 0.95\}$.

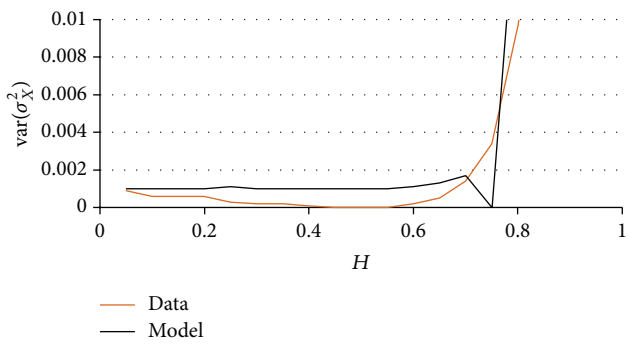

Figure 11: Variance of the estimated variance for $n = 1024$ and $H = \{0.05, 0.10, 0.15, \ldots, 0.95\}$.

6.3. Synthesis of H-SOSS Time Series. To exemplify the proposed wavelet-based synthesis (described in Section 3.2), four time series with respective Hurst index 0.3, 0.5, 0.7, and 0.9 were synthesized from an FGN sample of size 1024. The *Logscale Diagram* of the four new time series were obtained and compared to that of the original sample. The plot X_t versus t for each one; the four synthesized series is shown in Figure 12. One can visually check the presence of positive correlation in Figures 12(c) and 12(d).

The *Logscale Diagram* of these artificial series is shown in Figure 13. Note that the original *Logscale Diagram* of the source sample is not a straight line, but it so is for the synthesized series. Also, the estimated Hurst index of this series is $\widehat{H} = 0.56$ and its estimated *Logscale Diagram* is not a straight line, but the estimated Hurst index of the four generated series is exactly the desired, for example, $\widehat{H} = 0.30$ for the series shown in Figure 12(a) and the same for the others, and their respective *Logscale Diagram* is a straight line.

6.4. Voice over IP (VoIP) Measurements. The jitter behavior of Voice over Internet Protocol (VoIP) traffic by means of networks measurements is analyzed. As result of this analysis, detailed characterization and accurate modeling of this Quality of Service (QoS) parameter is provided. Previous studies have revealed that VoIP jitter can be modeled by self-similar processes with short-range dependence (SRD) or long-range dependence (LRD) [36]. The discovery of LRD (a kind of asymptotic fractal scaling) and weak self-similarity in the VoIP jitter data traces was followed by a further work that shows the evidence for multifractal behavior. The discovery of evidence for multifractal behavior is a richer form of scaling

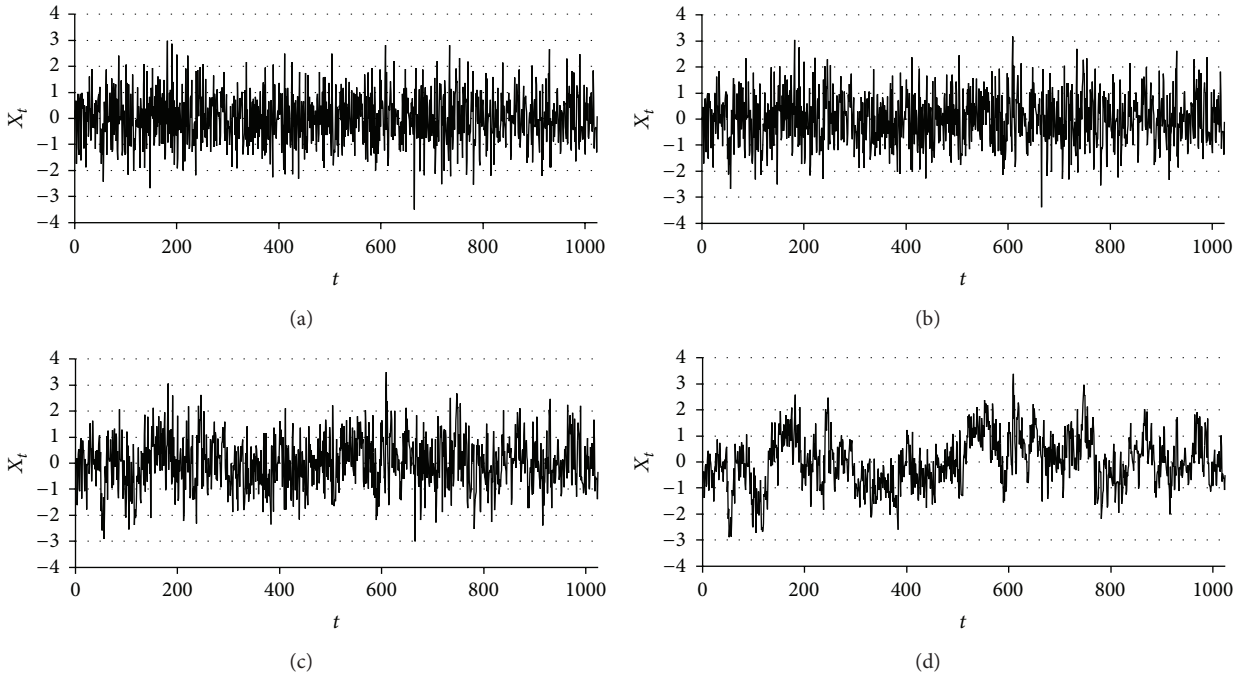

FIGURE 12: Plot versus time of the four synthesized time series. (a) $H = 0.30$, (b) $H = 0.50$, (c) $H = 0.70$, and (d) $H = 0.90$.

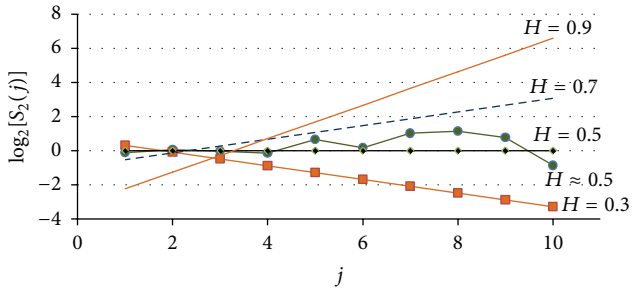

FIGURE 13: *Logscale Diagram* of synthesized time series.

behavior associated with nonuniform local variability, which could lead to a complete and robust model of IP network traffic over all time scales.

Motivated by such concerns, the evidence for multifractal behavior of VoIP jitter data traces is reviewed. In order to accomplish this analysis, the time series of VoIP jitter into a set of time series or components $C_{X,\tau}^{2,j}$ is decomposed as defined by (16). The behavior of these components is used to determine the kind of asymptotic fractal scaling. If the variance of the components of a time series is modeled by a straight line, the time series exhibits monofractal behavior, and a linear regression can be applied in order to estimate the Hurst parameter. On the other hand, if the variance of the components cannot be adequately modeled with a linear model, then the scaling behavior should be described with more than one scaling parameter, that is, the time series exhibits multifractal behavior [23]. In Figures 14 and 15, we show the components behavior of the collected VoIP jitter data traces.

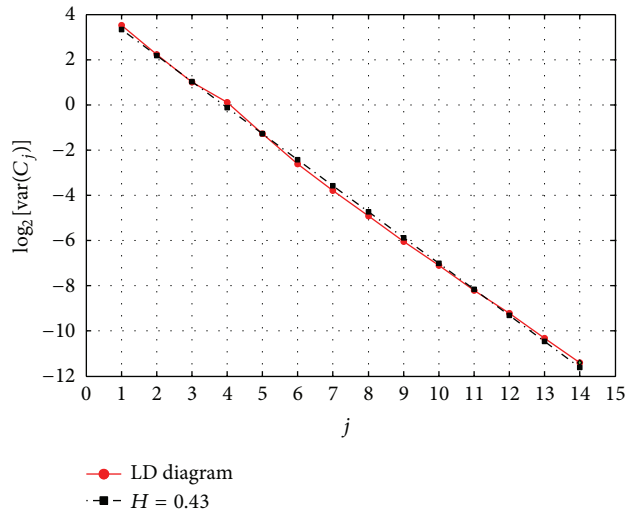

FIGURE 14: Components behavior of VoIP jitter data traces: monofractal behavior.

Figure 14 shows the components behaviors of a VoIP jitter data trace that belongs to the data sets with SRD. It is observed that the variance of the components of this time series is modeled by a straight line; therefore, the time series exhibits monofractal behavior.

Figure 15 shows the components behaviors of a VoIP jitter data trace that belongs to the data sets with LRD. It is observed that the variance of the components of this time series cannot be adequately modeled with a linear model, and the scaling behavior should be described with multiple scaling parameters (biscaling); therefore, this time series exhibits multifractal behavior.

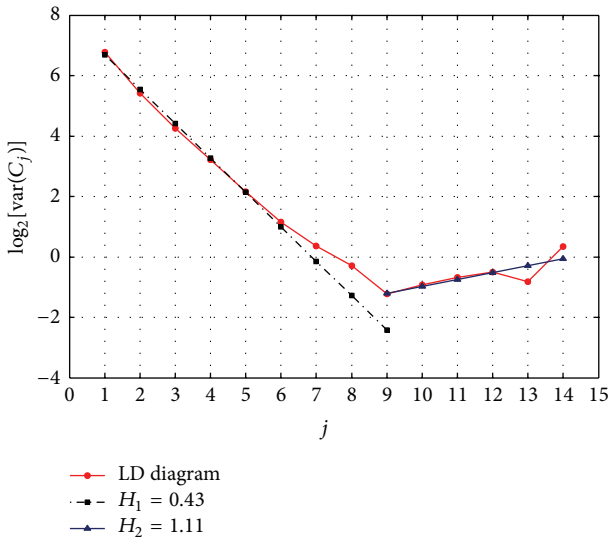

FIGURE 15: Components behavior of VoIP jitter data traces: multifractal behavior.

These results show that VoIP jitter with SRD or LRD exhibit monofractal or multifractal behavior, respectively. This phenomenon explains the behavior of the data traces with SRD and high degree of self-similarity (scale invariance), because the self-similarity is defined for a single scale parameter. On the other hand, the data traces with LRD exhibit weak self-similarity because they have associated nonuniform local variability (multifractal behavior).

The implication of this behavior for VoIP and other interactive multimedia services is that receiver dejitter buffer may not be large enough to mask the jitter with LRD and multifractal characteristics.

7. Conclusion

An orthogonal decomposition, that constitutes a powerful statistic tool to study discrete time series, is presented. The resulting components ($C_{X,t}^{n,j}$) have zero-mean and three desirably properties: (1) they synthesize the source signal without loss, (2) they are pair wise orthogonal, and (3) their variances form a geometric series whose rate is related to the Hurst index.

This decomposition is firstly compared to Haar wavelet based. For a particular case these two coincide, but the proposed one is more general, as other levels of aggregation may occur. In this case, the Haar wavelet is not sufficient and a special class of basis functions are defined, and the wavelet coefficients are obtained by the inner products between the signal under study and a scaled, displaced, and phase shifted versions of the basis function, in contradistinction of other wavelet decomposition, that only apply scaling and displacement. For a fixed scaling and displacement, the last phase shift can be discarded (i.e., the components are downsampled) as it does not provide additional information.

The proposed decomposition can be used to estimate the Hurst index, as the plot $j + \log_n[\mathrm{var}(C_{X,t}^{n,j})]$ versus j is equivalent to the *Logscale Diagram* proposed by Veitch and Abry in [11]. It is a straight line for H-SOSS time series, and the slope is related to the Hurst index so that $s = 2H - 1$.

The components can be also used to synthesize H-SOSS time series by means of the weighted sum defined by (31) and (32), regardless of the distribution of the source signal and whether it is or not self-similar. The synthesis is exact, that is, the Hurst index of the synthesized series is exactly the desired, which is an advantage over the proposed synthesis, that uses IMF from an EMD decomposition, described in [27].

A study of the estimated mean and variance of self-similar time series is also presented. Both statistics, mean and variance, can be estimated without bias, by applying the classic sample mean ($\overline{X} = 1/N \sum_{t=1}^{N} X_t$) and the proposed variance estimator (40). The variance of these estimators depends on the Hurst index of the process. In the case of the sample mean, its variance is an increasing function of H, as expressed by (33), such that $\mathrm{var}(\overline{X}) \in (\sigma_X^2 N^{-2}, \sigma_X^2)$. Note that the sample mean becomes less significant as H approaches to 1. For the variance of $\widehat{\sigma}_X^2$, it can be observed that it has the best performance when $H = 0.5$ (the variance of $\widehat{\sigma}_X^2$ is minimal). For $H < 0.5$, the variance increases as H is lower, but it is still acceptable (e.g., $\mathrm{var}(\widehat{\sigma}_X^2) < 0.001\sigma_X^2$ when $H \to 0$. But when H gets close to 1, the uncertainty of the sample variance increases rapidly, making the estimation of σ_X^2 less significant.

It is demonstrated also (clarifying a popular misunderstanding) that the *Variance Plot* can be used to estimate efficiently the Hurst index. The claims of many researchers about the inefficiency of the *Variance Plot* are the result of an ill-conditioned problem: to estimate the Hurst index the variance of the aggregated series is needed and vice versa, leading to a vicious circle. It is shown how this problem is avoided by estimating the *Variance Plot* using a set of aggregation levels that follow a geometric series and calculating *the differences* between the variances of the corresponding aggregated series; then, a weighted linear regression is applied to estimate the slope ($\widehat{s}$), and the estimated Hurst index is $\widehat{H} = \widehat{s}/2 + 1$. This result gives more significance to the results published by Abry et al. in [26].

Acknowledgment

This work was supported by PROMEP Grant UQROO-PTC-110.

References

[1] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, 1994.

[2] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," in *Proceedings of the ACM SIGCOMM Computer Communication Review*, vol. 23, pp. 183–193, San Francisco, Calif, USA, 1993.

[3] T. Karagiannis, M. L. Molle, and M. Faloutsos, "Long-range dependence: Ten years of internet traffic modeling," *IEEE Internet Computing*, vol. 8, no. 5, pp. 57–64, 2004.

[4] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch, "Multiscale nature of network traffic," *IEEE Signal Processing Magazine*, vol. 19, no. 3, pp. 28–46, 2002.

[5] M. S. Taqqu, V. Teverovsky, and W. Willinger, "Is network traffic self-similar or multifractal?" *Fractals*, vol. 5, no. 1, pp. 63–73, 1997.

[6] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, and D. Veitch, "Multiscale nature of network traffic," *IEEE Signal Processing Magazine*, vol. 19, no. 3, pp. 28–46, 2002.

[7] H. D. J. Jeong, J. S. R. Lee, D. McNickle, and K. Pawlikowski, "Comparison of various estimators in simulated FGN," *Simulation Modelling Practice and Theory*, vol. 15, no. 9, pp. 1173–1191, 2007.

[8] J. Mielniczuk and P. Wojdyłło, "Estimation of Hurst exponent revisited," *Computational Statistics and Data Analysis*, vol. 51, no. 9, pp. 4510–4525, 2007.

[9] P. Shang, Y. Lu, and S. Kamae, "Detecting long-range correlations of traffic time series with multifractal detrended fluctuation analysis," *Chaos, Solitons and Fractals*, vol. 36, no. 1, pp. 82–90, 2008.

[10] G. Horn, A. Kvalbein, J. Blomskøld, and E. Nilsen, "An empirical comparison of generators for self similar simulated traffic," *Performance Evaluation*, vol. 64, no. 2, pp. 162–190, 2007.

[11] D. Veitch and P. Abry, "A wavelet-based joint estimator of the parameters of long-rangé dependence," *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 878–897, 1999.

[12] D. Radev and I. Lokshina, "Advanced models and algorithms for self-similar ip network traffic simulation and performance analysis," *Journal of Electrical Engineering*, vol. 61, no. 6, pp. 341–349, 2010.

[13] R. G. Clegg, "A practical guide to measuring the Hurst parameter," in *Proceedings of the 21st UK Performance Engineering Workshop*, School of Computing Science, Technical Report Series, pp. 43–55, 2006.

[14] J. F. Kenney, *Mathematics of Statistics*, Van Nostrand, New York, NY, USA, 1939.

[15] J. F. Kenney and E. S. Keeping, *Mathematics of Statistics*, part 2, Van Nostrand, Princeton, NJ, USA, 2nd edition, 1951.

[16] T. T. Soong, *Fundamentals of Probability and Statistics for Engineers*, John Wiley & Sons, New York, NY, USA, 2004.

[17] B. Tsybakov and N. D. Georganas, "Self-similar processes in communications networks," *IEEE Transactions on Information Theory*, vol. 44, no. 5, pp. 1713–1725, 1998.

[18] I. W. C. Lee and A. O. Fapojuwo, "Stochastic processes for computer network traffic modeling," *Computer Communications*, vol. 29, no. 1, pp. 1–23, 2005.

[19] D. Radev and I. Lokshina, "Self-similar simulation of IP traffic for wireless networks," *International Journal of Mobile Network Design and Innovation*, vol. 2, no. 3-4, pp. 202–208, 2007.

[20] J. M. Ramírez-Corte, V. Alarcon-Aquino, G. Rosas-Cholula, P. Gomez Gil, and J. Escamilla-Ambrosio, "P-300 rhythm detection using anfis algorithm and wavelet feature extraction in eeg signals," in *Proceedings of the World Congress on Engineering and Computer Science*, San Francisco, Calif, USA, October 2010.

[21] M. S. Taqqu, V. Teverovsky, and W. Willinger, "Is network traffic self-similar or multifractal?" *Fractals*, vol. 5, no. 1, pp. 63–73, 1997.

[22] O. Sheluhin, S. Smolskiy, and A. Osin, *Self-Similar Processes in Telecommunications*, Wiley, 2007.

[23] S. Stoev, M. S. Taqqu, C. Park, and J. S. Marron, "On the wavelet spectrum diagnostic for Hurst parameter estimation in the analysis of Internet traffic," *Computer Networks*, vol. 48, no. 3, pp. 423–445, 2005.

[24] M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, IEEE Press.

[25] N. Rillo, "Introduction to wavelets theory," http://www.mat.ub.es/~soria/TAD-Wavelets.pdf.

[26] P. Abry, D. Veitch, and P. Flandrin, "Long-range dependence: Revisiting aggregation with wavelets," *Journal of Time Series Analysis*, vol. 19, no. 3, pp. 253–266, 1998.

[27] É. Deléchelle, J. C. Nunes, and J. Lemoine, "Empirical mode decomposition synthesis of fractional processes in 1D- and 2D-space," *Image and Vision Computing*, vol. 23, no. 9, pp. 799–806, 2005.

[28] W. G. Cochran, "The distribution of quadratic forms in a normal system, with applications to the analysis of covariance," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 30, no. 2, pp. 178–191, 1934.

[29] Y. Tian and G. P. H. Styan, "Cochran's statistical theorem revisited," *Journal of Statistical Planning and Inference*, vol. 136, no. 8, pp. 2659–2667, 2006.

[30] J. Beran, *Statistics for Long-Memory Processes*, Monographs on Statistics & Applied Probability, Chapman & Hall/CRC, 1994.

[31] R. Yunhua, "Evaluation and estimation of second-order self-similar network traffic," *Computer Communications*, vol. 27, no. 9, pp. 898–904, 2004.

[32] M. Krunz and I. Matta, "Analytical investigation of the bias effect in variance-type estimators for inference of long-range dependence," *Computer Networks*, vol. 40, no. 3, pp. 445–458, 2002.

[33] P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 910–917, 1992.

[34] V. Teverovsky and M. Taqqu, "Testing for long-range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimator," *Journal of Time Series Analysis*, vol. 18, no. 3, pp. 279–304, 1997.

[35] R. B. Davies and D. S. Harte, "Tests for hurst effect," *Biometrika*, vol. 74, no. 1, pp. 95–101, 1987.

[36] H. Toral, *QoS parameters modeling of self-similar VoIP traffic and an improvement to the E model [Ph.D. thesis]*, Electrical Engineering, Telecommunication Section, CINVESTAV, Jalisco, Mexico, 2010.