Hurst Exponent Estimation using Natural Visibility Graph Embedding in Fisher-Shannon Plane

T. Averty, A. O. Boudraa, D. Daré-Emzivat

Arts et Métiers Institute of Technology, École navale, IRENav, BCRM Brest, CC 600, Brest, 29240, France

Abstract

In this article, two important stochastic processes, namely fractional Brownian motions (fBm) and fractional Gaussian noises (fGn) are analyzed, within a Fisher-Shannon framework. These processes are well suited for the realistic modeling of phenomena occurring across various domains in science and engineering. An unique feature that characterizes both fBm and fGn, is the Hurst parameter H, that measures the long/short range dependence of such stochastic processes. In this paper, we show that these processes, from which we extract the degree distribution of the associated natural visibility graph (NVG), can be located in an informational plane, defined by normalized Shannon entropy S and Fisher information measure F, in order to estimate their Hurst exponents. The aim of this work is to map signals onto this informational plane, in which a reference backbone is built using generated fBm and fGn processes with known Hurst exponents. To show the effectiveness of the developed graphical estimator, some real-world data are analyzed, and it found that the H estimated by our method are quite comparable to those obtained from four well-known estimators of the literature. Besides, estimation of H parameter is very fast and requires a reduced number of samples of the input signal. Using the constructed reference backbone in the Fisher-Shannon plane, the associated H exponent can be easily estimated by a simple orthogonal projection of the point (S, F) extracted from the truncated degree distribution of the considered signal NVG representation.

Keywords: Shannon entropy, Fisher information measure, Visibility graph, Hurst exponent, fBm, fGn

1. Introduction

Time series measurements provide valuable information for investigating and describing natural phenomena and dynamic processes of complex systems. Information-based measures of randomness or regularity of time series have received considerable attention in many fields such as physics, biology or finance [1]. In the system identification domain, parameter estimation is crucial and is carried out in the presence of perturbations. The noisy observations constitute the time series involved in the parameter identification process [2, 3]. Various strategies have been developed to understand and analyze the complex behavior of non-linear time series [4], including information theory, machine learning and time-frequency methods. Recently, a bridge has been built between signals and complex networks [5, 6, 7]. This has led to new perspectives on the analysis of these signals using graph theory, which proposes its own set of tools and methods. More precisely, time series or signals analysis has greatly benefited from graph representations as they provide a mapping able to deal with issues of non-linearity and multi-scale. To this end, some methods have been developed to construct a graph from an univariate time series, where this graph inherits several structural properties of the time series [8]. The graph can produce insights that are not visible by classical time series approach. Consequently, the dynamic characteristics of the time series are studied by analyzing the topological structure of the graph. One pertinent attribute that can be extracted from these graphs is the degree distribution, which, although extremely simple, can be used for example to identify radar emitters [9] or to classify EEG signals [10, 11]. Among the most popular algorithms for mapping a graph from an ordered set of time series samples is the visibility graph (VG) introduced by Lacasa *et al.* [7]. This graph is computationally efficient, provides a deterministic non-parametric representation of the time series, and is able to deal with non-stationarity and non-linearity issues of the time series [7, 12].

Several information theory quantifiers, such as Shannon entropy, Fisher information measure (FIM) or Fisher-Shannon plane, have also been introduced for analyzing time series and understanding the behavior of associated systems [4], where FIM describes the local changes of the underlying distribution whereas Shannon entropy is a global quantifier of this distribution informational content. Bercher and Vignat highlight the joint use of these two quantifiers and defined the Fisher-Shannon informational plane for characterizing the non-stationary behavior of complex time series [4]. Then, this graphical tool makes it possible to locate, on a two-dimensional plane, a time series or dynamic system based upon its Shannon entropy and its FIM.

Few information-based tools are dedicated to the study of graphs and time series seen as graphs. Recent works of the literature have shown the interest of this Fisher-Shannon informational plane, combined with horizontal visibility graph (HVG) to analyze stochastic time series [13, 14, 15, 16, 17, 18].

However, those studies have been limited to fractional Brownian motion (fBm) processes [18] and the crucial problem of the estimation of the Hurst parameter has not been tackled. In the present work we follow the same strategy and, in addition to fBm processes, extend the study to fractional Gaussian noise (fGn), and propose a graphical estimator of the Hurst exponent H. More precisely, we extend the work of Gonçalves *et al.* [18] by using this informational plane in which we locate the degree distributions of natural visibility graph (NVG) constructed from fBm and fGn processes to estimate their Hurst parameters. To the best of our knowledge, this is the first time that NVG algorithm associated with this informational plane is used to estimate quantitatively the H parameter. We show the effectiveness of the proposed tool by estimating such a parameter of some real-world time series. These estimates of H are compared to those of four well known estimators of the domain using a qualitative analysis in that, to the best of our knowledge, there is no "golden standard" method for estimating this parameter.

2. Fractional Brownian motion & fractional Gaussian noise

The fBm is a non-stationary and self-similar process that exhibits short/long-term dependence, self-similarity and power-law spectra. With fractional character depending on the parameter $H \in (0, 1)$, the fBm is continuous, has stationary increments and is defined [19] as a zero-mean Gaussian process $B_H(t)$, $t \in [0, T]$, with a covariance equals to

$$\mathbb{E}[B_H(t)B_H(s)] = 0.5\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right). \tag{1}$$

For $H \in [0, 1/2]$ the process exhibits short dependence and its increments are negatively correlated. When $H = 1/2$, we find the standard Brownian motion. Finally, when $H \in [1/2, 1]$, the process exhibits long-range dependence and its increments are positively correlated [20]. Due to its non-stationary character, the fBm is not suited for modeling stationary processes. That is why increments of fBm, called fractional Gaussian noise (fGn) defined as $G_H(t) = B_H(t) - B_H(t-1)$, are used and verify the stationary property. Both fGn and fBm, are employed to model a large class of natural phenomena, and most of their statistical properties depend on the parameter H. To construct a fGn of parameter H, we can then construct a fBm of parameter H and differentiate it and to study a fGn of parameter H, it may be easier to integrate it to see it as a fBm of parameter H (after removing the mean so that the cumulative sum in the discrete case is fBm-like). Indeed, many signal processing methods or Hurst estimators are essentially based on fBm [21].

3. Visibility graph representation in a Fisher-Shannon informational plane

3.1. Visibility graph (VG)

Several methods for encoding time series in graphs have been proposed such as the VG which is a widely used approach [7, 22]. The resulting network offers new insights, often revealing non-trivial properties about the data they represent, while preserving the information of the original time series. Topological analysis of this network provides new information about the original data, while preserving its original properties. The measurements are based on calculations of networks characteristics such as centrality, degree or connectivity. Lacasa *et al.* [7] have shown that the NVG associated with a fractal time series is a scale-free graph and that it is possible to extract information about the long-term dependence and fractality of the time series from this graph. The NVG transforms a signal $\mathbf{x} = (x_i)_{1\leq i\leq n}$ of n samples into a graph with n vertices constructed thanks to a simple geometric criterion: an edge exists between vertices i and j if and only if any other sample x_k placed between x_i and x_j satisfies

$$x_k < x_j + (x_i - x_j)\frac{j-k}{j-i}, \quad \forall k \in [\![i, j]\!]. \tag{2}$$

What makes NVG so interesting is its ability to analyze a wide class of time series, in particular those from very complex physical or biological systems, with a low algorithmic construction complexity of $O(n \log n)$.

3.2. Degree distribution

One motivation of this work is the ability to estimate the H parameter of fBm and fGn processes from the slope of the degree distribution of the associated horizontal visibility graph (HVG) [23, 24], a variant of the natural one [25]. Indeed, degree distribution is a simple attribute to capture information about a graph, and still gives clues about the structure of the graph. It has been shown the interest to classify non-stationary signals such as EEG signals using degree distribution [10, 11]. Recalling that the degree $\deg(k)$ of a vertex k is the number of edges incident on it, the degree distribution of the graph is defined by

$$\mathbf{p} = (p_i)_{1\leq i\leq n}, \quad p_i = \frac{\mathrm{Card}\left(\{k : \deg(k) = i\}\right)}{n} \tag{3}$$

where p_i denotes the ratio of vertices having i neighbors. We assume here that the simple degree distribution gives enough

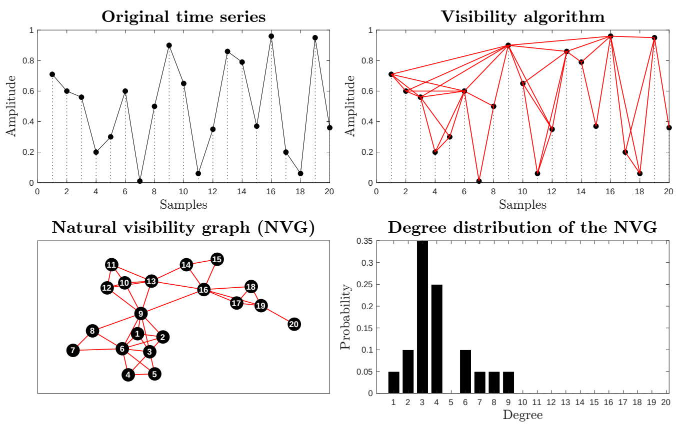

Figure 1: Using the NVG algorithm (top right) on a time series (top left) to get its NVG (bottom left) and extract its degree distribution (bottom right).

information about the VG, even if others can be used: weight distribution [18] if the VG is weighted, the distance distribution [26] or the centrality distribution [27]. Figure 1 illustrates the different steps to map an input signal into NVG and the extraction of its degree distribution. Figure 2 shows fBm draws of 50,000 points for different H values, and the degree distributions (zoomed for degrees between 1 and 50) of the NVG constructed from these signals. It is clear that these distributions are quite different from one another: for a small H value, the distributions are very tight around degree 2-3 because, due to the temporal structure of the process, there are few samples that see a significant number of neighbors. For large values of H, on the other hand, it is noticeable that some samples see numerous neighbors, making the tail of the distribution thicker. All these observations then give hope that distributions constitute a suitable tool to discriminate between H parameters. It is then necessary to characterize these distributions. A way to do it is to combine two informational quantifiers: a local and a global one detailed in the following subsection.

3.3. Fisher information and Shannon entropy

In this work we exploit the potential of Fisher-Shannon informational plane, originally introduced to analyze time series after extracting and estimating their amplitude distribution [4]. The idea is to combine a global informational quantifier like the normalized Shannon entropy (NSE) [28] and the Fisher information measure (FIM) employed to characterize local changes in the distribution [29]. In their work, Vignat and Bercher motivate the use of this plane because FIM makes it possible to characterize the non-stationary behavior of signals, whereas NSE has limitations for this task. Among the properties about this informational plane, the authors proved in the continuous case the existence of a boundary reached when the studied random variable is Gaussian [4]. Let us recall that the Shannon entropy of a random variable X whose probability density function is $f_X(x)$ is given by [28]

$$S(X) = -\int f_X(x) \log_2 f_X(x)\,\mathrm{d}x. \tag{4}$$

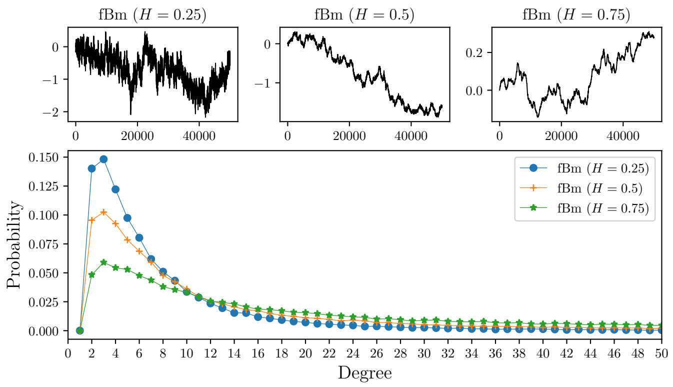

Figure 2: Draws containing 50,000 points of fBm for a $H \in \{0.25, 0.5, 0.75\}$ and the degree distribution of their associated natural visibility graphs (zoomed for degrees between 1 and 50).

In our case, *i.e.* when the distribution $\mathbf{p} = (p_i)_{1\leq i\leq n}$ is discrete, we consider the NSE defined by

$$S(\mathbf{p}) = -\frac{1}{\log_2(n)} \sum_{i=1}^{n} p_i \log_2 p_i. \tag{5}$$

As for the FIM, whose continuous version was introduced as follows [29]

$$F(X) = \int \left(\frac{\partial}{\partial x} f_X(x)\right)^2 \frac{\mathrm{d}x}{f_X(x)} \tag{6}$$

we consider the discretization given by [30]

$$F(\mathbf{p}) = \frac{1}{2} \sum_{i=1}^{n-1} \left(\sqrt{p_{i+1}} - \sqrt{p_i}\right)^2. \tag{7}$$

3.4. Construction of an informational backbone

To illustrate the representation of simulated fBm in the Fisher-Shannon plane, we generated them for H varying from 0.1 to 0.9 in steps of 0.1, each having 50,000 points. As it can be seen on Figure 2, the salient information is contained in the first degree values then it is not necessary to build the degree distribution $\mathbf{p} = (p_i)_{1\leq i\leq n}$ defined by Eq. (3) for n equals to the number of points (50,000 in this case). That is why we only calculate the degree distribution up to an empirically determined degree threshold ε set to 100. More importantly, this makes the method less dependent upon the number of points in the time series being studied. Figure 3 illustrates the constructed informational backbone, which shows an almost linear relationship that makes the estimation of H easier. This figure depicts in gray the embedding of 10 draws of these simulated processes in the Fisher-Shannon plane and the black points correspond to the centroids of the draws for each H value. These black points form the reference backbone onto which we will project the real time series to estimate their Hurst parameters. A pseudo-code algorithm for the backbone construction is proposed in Algorithm 1 to ensure reproducibility.

Algorithm 1 Construction of a backbone $\mathbf{B}$ for fBm in Fisher-Shannon information plane

Input: Number of samples n, number of draws N, vector of L Hurst exponents $\mathbf{h} = (h_i)_{1\leq i\leq L}$ and the degree threshold ε.
Output: Backbone $\mathbf{B}$.

```
for i = 1 : L do
    s ← 0
    f ← 0
    for t = 1 : N do
        x ← FBM(n, h_i)
        G ← NVG(x)                                  ▷ Equation (2)
        p ← degree_distribution(G, 1 : 1 : ε)       ▷ Equation (3)
        s_t ← S(p)                                  ▷ Equation (5)
        f_t ← F(p)                                  ▷ Equation (7)
    B_{i,1} ← mean(s)
    B_{i,2} ← mean(f)
```

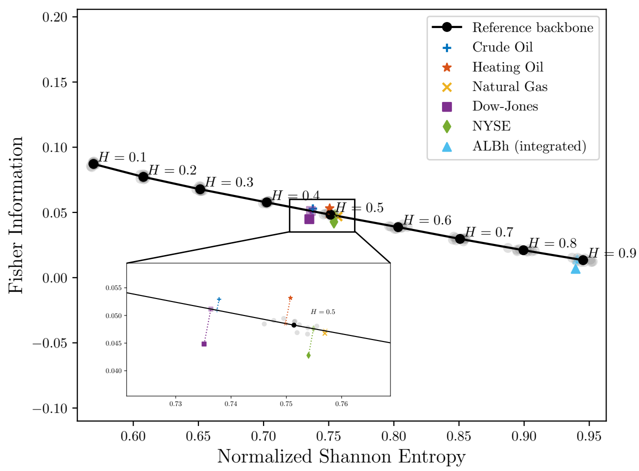

Figure 3: Reference backbone consisting of centroids (black points) of 10 draws embeddings of 50,000 points (gray points) of fBm with H ranging from 0.1 to 0.9. Real time series embeddings are represented by colored markers whose orthogonal projections on the reference backbone are also depicted. A zoom view is provided to show these orthogonal projections.

For a complete overview, the choice of the NVG rather than HVG is justified by the fact that it is not necessary to have many samples in the generated signals to obtain an accurate backbone (50,000 are sufficient in this case), unlike the HVG, which requires more than 1,000,000 samples [18]. This problem had already been mentioned by Gonçalves *et al.* in their work [18], which was solved here by truncating degree distributions of NVG.

4. Hurst exponent estimation of real time series

Once the reference backbone is constructed with the above-mentioned centroids in the Fisher-Shannon informational plane (black points linked in the Figure 3), we can first save it for later use and it is possible to apply the same method to any real-world signal (assuming it behaves as fBm or fGn). For a given real time series, we construct its NVG from which we extract the degree distribution up to a degree equal to $\varepsilon = 100$. We calculate the Fisher information measure F and the normalized Shannon entropy S, and then we orthogonally project the point (S, F) onto the lines joining the centroids in order to estimate the Hurst exponent. This methodology is summarized in Figure 4 and a pseudo-code for this estimation method, which we will denote **VG/FS**, is proposed in Algorithm 2. It is important to keep in mind that if the backbone has already been built, the main complexity of this estimation method comes from the construction of the NVG, which can reach $O(n \log n)$. So it is not an unfeasible method from a computational point of view.

To show the effectiveness of the proposed graphical estimator, six real-world data are quantitatively analyzed. This set of real data is detailed, with associated H parameters estimated by conventional estimators, in reference [21]. Among these time series, daily observations from NYMEX spot-month fu-

Algorithm 2 Method **VG/FS** to estimate the Hurst exponent H of a time series $\mathbf{x}$

Input: Backbone $\mathbf{B} = \{(S_i^c, F_i^c)\}_{1 \leq i \leq L}$ containing L centroids in Fisher-Shannon plane corresponding to Hurst exponents $(h_i)_{1 \leq i \leq L}$ and degree threshold ε.

Output: Estimated Hurst exponent H.

if $\mathbf{x}$ is fGn-like **then**
  $\mathbf{x} \leftarrow \text{cumsum}(\mathbf{x})$
$G \leftarrow \text{NVG}(\mathbf{x})$ ▹ Equation (2)
$\mathbf{p} \leftarrow \text{degree\_distribution}(G, 1\!:\!1\!:\!\varepsilon)$ ▹ Equation (3)
$S \leftarrow S(\mathbf{p})$ ▹ Equation (5)
$F \leftarrow F(\mathbf{p})$ ▹ Equation (7)
Find the two centroids (S_j^c, F_j^c) and (S_k^c, F_k^c) (whose Hurst exponents are respectively h_j and h_k) closest to (S, F).
$X \leftarrow \dfrac{(S_k^c - S_j^c)(S - S_j^c) + (F_k^c - F_j^c)(F - F_j^c)}{(S_k^c - S_j^c)^2 + (F_k^c - F_j^c)^2}$
$S_{\text{proj}} \leftarrow S_j^c + X(S_k^c - S_j^c)$
$F_{\text{proj}} \leftarrow F_j^c + X(F_k^c - F_j^c)$
$d \leftarrow \sqrt{(S_j^c - S_{\text{proj}})^2 + (F_j^c - F_{\text{proj}})^2}$
$d_{\text{tot}} \leftarrow \sqrt{(S_j^c - S_k^c)^2 + (F_j^c - F_k^c)^2}$
$H \leftarrow h_j + (h_k - h_j)\dfrac{d}{d_{\text{tot}}}$

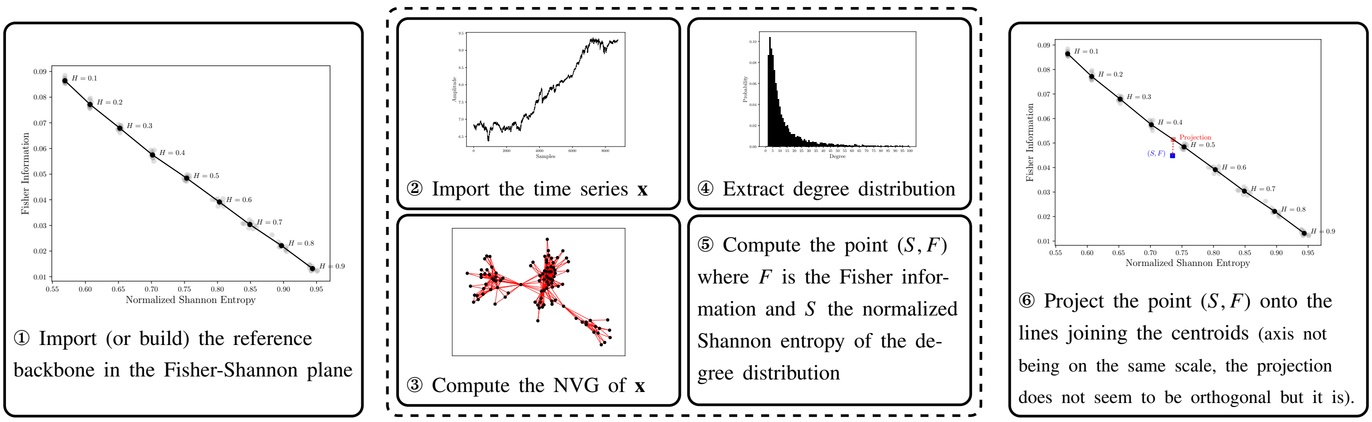

Figure 4: Flowchart of our proposed strategy to estimate the Hurst exponent H.

ture prices for three energy commodities are used: crude oil, heating oil and natural gas. Daily observations of two US stock market indexes are considered: Dow Jones Industrial Average and New-York Stock Exchange. All these financial time series are fBm-like and their Hurst parameters are around $H \approx 0.5$ [31, 21] but in order to show that our strategy works even on signals modeled by a fGn process, hourly prices for Alberta electricity market are considered [32]. This time series is indeed behaving as fGn with Hurst parameter around $H \approx 0.9$ [21]. To estimate the Hurst parameter of this time series with our method, we center and integrate it in order to have a fBm-like signal. All these time series are embedded into the Fisher-Shannon plane using the approach shown in Figure 4. These embeddings and their orthogonal projections are represented by the colored points in Figure 3. The first thing to be noticed is that time series with a Hurst parameter close to 0.5 are located in the neighborhood of the corresponding reference point, as is the signal with a Hurst parameter close to 0.9.

Results of our estimator and those of four methods of the literature are reported in Table 1. The first estimator is the power spectral density (**PSD**) method [21, 33]. By recalling that fBm and fGn are often modeled using colored noises, *i.e.* signals with PSD following a power law with parameter β, a fGn of Hurst parameter H can be modeled with $\beta \in [-1, 1]$ ($\beta = 2H-1$) whereas a fBm can be modeled with $\beta \in [1, 3]$ ($\beta = 2H + 1$). The second estimator is wavelet approach (**WAV**) based on the fact that the variance of the wavelet $W_{a,b}$ follows a power-law a^{β} (the same above-mentioned relations between H and β can then be used) [34]. The third estimator is rescaled range analysis (**R/S**) [35]. The time series of length N is divided into shorter time series of length n. The rescaled range $R(n)/S(n)$ is then calculated for each n value, and it can be proved that $\mathbb{E}[R(n)/S(n)] \propto n^H$ when $n \to \infty$. The fourth estimator is the detrented fluctuation analysis (**DFA**)[36]. This method divides the cumulative sum of the centered time series into consecutive segments of length n on which linear fits are performed. Then, the root-mean-square $\sigma(n)$ of the root-mean-square deviations of the local fluctuations (difference between the segments and the associated linear fit) is computed and $\sigma(n)$ follows a power-law n^{α} where α is an equivalent of the H parameter. Results presented in the Table 1 show that the estimated values H obtained by our estimator **VG/FS** are very consistent with those given by **PSD**, **WAV**, **R/S** and **DFA**.

As already mentioned, there is no "golden standard" method against which to compare our estimates. To provide a qualitative analysis, the boxplot-like in Figure 5 is proposed and shows, among all conventional estimators, the minimal, the maximal and the average one. Our **VG/FS** estimates are depicted with the same marker as those of Figure 3. For the Heating Oil time series, our method overestimates the H parameter compared with classical ones while for the Dow-Jones financial time series, it is underestimated. For all other signals, our estimates are close to the average one, being almost identical in the case of the NYSE and Natural Gas time series. These findings highlight the fact that the NVG and the informational plane are able to capture the persistence of the time series *via* the H estimates.

	VG/FS	PSD	WAV	R/S	DFA
Crude Oil (fBm)	0.471	0.423	0.452	0.498	0.466
Heating Oil (fBm)	0.496	0.441	0.443	0.423	0.458
Natural Gas (fBm)	0.510	0.433	0.465	0.654	0.481
Dow-Jones (fBm)	0.469	0.491	0.480	0.554	0.481
NYSE (fBm)	0.506	0.491	0.487	0.545	0.505
ALBh (fGn)	0.895	0.923	0.829	0.838	0.855

Table 1: Hurst parameter estimates by our **VG/FS** strategy and conventional estimators: **PSD** / **WAV** / **R/S** / **DFA**.

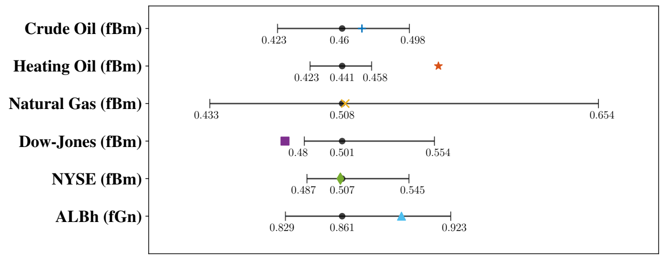

Figure 5: Min-Average-Max Hurst estimators of real-world time series. Conventional methods and our **VG/FS** estimator are depicted with black dot and colored marker (those in Figure 3) respectively.

5. Estimation errors on synthetic data

Although the **VG/FS** method is based on the use of a backbone in the Fisher-Shannon information plane constructed from synthetic fBm signals, it is important to quantify *a posteriori*, *i.e.* after construction of a backbone, say the one in Figure 3, the error that our estimator commits on other fBm synthetic signals. A quantitative analysis is performed with a number of draws equals to 20, with an increasing number of samples constituting fBm signals governed by Hurst exponents H varying from 0.1 to 0.9 in steps of 0.05. Each time series is fed into our estimator and, for each target H, mean and standard deviation of the 20 estimates are calculated. We do this for fBm consisting of 1 000, 2 500, 5 000, 10 000, 25 000 and 50 000 samples. Results are reported in Table 2 (for clarity, the H values are displayed in steps of 0.1). One can note that the number of samples has an obvious influence on the quality of the estimate. For a number of points equal to 50 000, the average estimate is almost always equal to the target H with a very low standard deviation. But the method also performs well with short time series. If we consider the fBm drawn with 2 500 samples, the averages are not far from the target values. A global analysis is done and the results presented in Figure 6 show the effect of the number of samples in signals on mean absolute error (MAE) of H estimates. It is clear that the curve of MAE against the number of samples decreases drastically as the time series becomes longer. In fact, the MAE falls below 0.01 when the signals exceed 20 000 samples. It can even be argued that the MAE of 0.025 for signals of 2 500 samples is already relatively correct, compared with the differences that may exist between the minimum and maximum estimator among the conventional ones proposed in Table 1 and illustrated in Figure 5, which are in all cases greater than 0.03.

		Number of samples in synthetic time series					
		1 000	**2 500**	**5 000**	**10 000**	**25 000**	**50 000**
Hurst exponent	$H = 0.1$	0.13 ± 0.02	0.12 ± 0.01	0.11 ± 0.01	0.10 ± 0.00	0.10 ± 0.00	0.10 ± 0.00
	$H = 0.2$	0.16 ± 0.02	0.17 ± 0.01	0.19 ± 0.01	0.19 ± 0.01	0.20 ± 0.00	0.20 ± 0.00
	$H = 0.3$	0.25 ± 0.02	0.28 ± 0.02	0.29 ± 0.01	0.29 ± 0.01	0.30 ± 0.01	0.30 ± 0.00
	$H = 0.4$	0.35 ± 0.03	0.39 ± 0.02	0.39 ± 0.02	0.40 ± 0.01	0.40 ± 0.01	0.40 ± 0.01
	$H = 0.5$	0.45 ± 0.04	0.49 ± 0.03	0.50 ± 0.01	0.51 ± 0.01	0.50 ± 0.01	0.50 ± 0.01
	$H = 0.6$	0.56 ± 0.04	0.60 ± 0.03	0.60 ± 0.02	0.60 ± 0.01	0.60 ± 0.01	0.60 ± 0.01
	$H = 0.7$	0.67 ± 0.05	0.71 ± 0.03	0.72 ± 0.02	0.71 ± 0.02	0.71 ± 0.01	0.70 ± 0.01
	$H = 0.8$	0.77 ± 0.04	0.81 ± 0.03	0.82 ± 0.02	0.81 ± 0.02	0.81 ± 0.02	0.79 ± 0.01
	$H = 0.9$	0.85 ± 0.06	0.87 ± 0.02	0.87 ± 0.02	0.89 ± 0.01	0.89 ± 0.01	0.89 ± 0.01

Table 2: Averages and standard deviations of the estimates given by our **VG/FS** method, according to different target values of the Hurst exponent H used to generate synthetic fBm. The number of samples in these synthetic time series varies from 1 000 to 50 000.

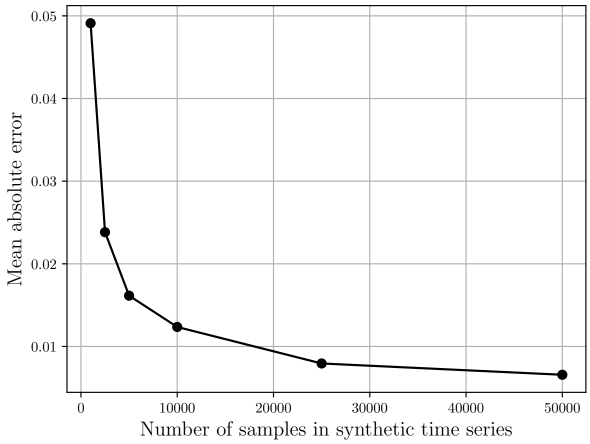

Figure 6: Mean absolute error of our **VG/FS** estimator on synthetic fBm (20 draws for Hurst exponent varying from 0.1 to 0.9 in steps of 0.05, *i.e.* 340 time series) as a function of the number of samples in the draws.

6. Conclusion and perspectives

In this article the potential of the Fisher-Shannon informational plane associated with natural visibility graph (NVG) representation of the input time series, is investigated as a tool to characterize both fBm and fGn stochastic processes by estimating their H parameters. A graphical estimator of this parameter, based upon this informational plane, is constructed and its effectiveness illustrated on real data. Degree distribution is extracted from the constructed graph to characterize topological structure and to capture the dynamics of the transformed input time series, using information theory quantifiers, namely normalized Shannon entropy and Fisher information measure. To the best of our knowledge, this is the first time that NVG associated with such an informational plane is used to estimate quantitatively the H parameter. The obtained results highlight that the estimated H are very consistent with the ones drawn from four estimators of the literature such as the DFA and wavelet approaches. This allows to efficiently estimate the H parameter. It is obvious that a broad class of real time series must be analyzed to confirm the obtained findings, especially for different values of H. Thanks to the reference backbone, built in the informational plane with different H values, the estimation of the Hurst exponent of an input time series is very fast. Moreover, the backbone construction with NVG requires a reduced number of samples in the initial synthetic fBm compared to HVG. Although it is the fastest variant of the visibility graphs, HVG involves at least a few million samples in these synthetic fBm in order to get a backbone as accurate as the one obtained with NVG [18], which makes it less efficient. The results also show the added value of characterizing the stochastic processes using their topology as graphs in order to better exploit the embedded information instead of their original temporal structures as it can be the case with other methods [34]. Also, these results shed new light on the Fisher-Shannon information plane to analyze complex and non-stationary signals. The proposed graphical estimator relies on the assumption that the input time series is correctly classified beforehand as fBm-like process. As future work we plan to develop a strategy associated with this graphical estimator to relax this assumption in order to consider native fGn-like time series. In addition, it will be interesting to investigate other information plane exploiting the Rényi, Tsallis or Havrda-Charvát entropies to further improve the estimation of the H parameter. Moreover, an optimal value for the threshold ε can be sought. Finally, we would like to investigate others methods of converting time-series to graphs that can produce more insights than NVG algorithm.

References

[1] Y. Wang, P. Shang, Analysis of Shannon-Fisher information plane in time series based on information entropy, Chaos: An Interdisciplinary J. Nonlinear Sc. 28 (10) (2018).

[2] V. Stojanovic, N. Nedic, D. Prsic, L. Dubonjic, Optimal experiment design for identification of arx models with constrained output in non-gaussian noise, Applied Mathematical Modelling 40 (13-14) (2016) 6676–6689.

[3] V. Stojanovic, N. Nedic, Identification of time-varying oe models in presence of non-gaussian noise: Application to pneumatic servo drives, International Journal of Robust and Nonlinear Control 26 (18) (2016) 3974–3995.

[4] C. Vignat, J.-F. Bercher, Analysis of signals in the Fisher–Shannon information plane, Physics Lett. A 312 (2003) 27–33.

[5] J. Zhang, M. Small, Complex network from pseudoperiodic time series: Topology versus dynamics, Phys. Rev. Lett. 96 (23) (2006) 238701.

[6] X. Xu, J. Zhang, M. Small, Superfamily phenomena and motifs of networks induced from time series, PNAS 105 (50) (2008) 19601–19605.

[7] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J. C. Nuno, From time series to complex networks: The visibility graph, PNAS 105 (13) (2008) 4972–4975.

[8] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. Xiang, J. Kurths, Recurrence-based time series analysis by means of complex network methods, International Journal of Bifurcation and Chaos 21 (04) (2011) 1019–1046.

[9] K. Chi, J. Shen, Y. Li, Y. Li, S. Wang, Multi-function radar signal sorting based on complex network, IEEE Signal Processing Letters 28 (2020) 91–95.

[10] T. Averty, D. Daré-Emzivat, A.-O. Boudraa, Détection d'épilepsie dans les signaux EEG par graphe de visibilité et un noyau de SVM adapté, in: Colloque GRETSI, 2022, pp. 1–4.

[11] G. Zhu, Y. Li, P. Wen, Analysing epileptic EEGs with a visibility graph algorithm, in: Int. Conf. Biomed. Eng. Inf., 2012, pp. 432–436.

[12] L. Lacasa, B. Luque, J. Luque, J. C. Nuno, The visibility graph: A new method for estimating the Hurst exponent of fractional Brownian motion, Europhysics Lett. 86 (3) (2009) 30001.

[13] L. C. Carpi, O. A. Rosso, P. M. Saco, M. G. Ravetti, Analyzing complex networks evolution through Information Theory quantifiers, Physics Lett. A 375 (4) (2011) 801–804.

[14] B. Zhang, P. Shang, J. Liu, Transition-based complexity-entropy causality diagram: A novel method to characterize complex systems, Commun. Nonlinear Sci. Numer. Simul. 95 (2021) 1–17.

[15] B. Shang, P. Shang, Generalized Shannon–Fisher index: An effective method to quantify the instability of multivariate time series, Commun. Nonlinear Sci. Numer. Simul. 117 (2023) 1–15.

[16] B. Shang, P. Shang, Effective instability quantification for multivariate complex time series using reverse Shannon-Fisher index, Chaos, Solitons & Fractals 160 (2022) 1–9.

[17] B. A. Gonçalves, L. Carpi, O. A. Rosso, M. G. Ravetti, A. P. F. Atman, Quantifying instabilities in financial markets, Physica A Stat. Mech. Appl. 525 (2019) 606–615.

[18] B. A. Gonçalves, L. Carpi, O. A. Rosso, M. G. Ravetti, Time series characterization via horizontal visibility graph and Information Theory, Physica A Stat. Mech. Appl. 464 (2016) 93–102.

[19] B. B. Mandelbrot, J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM review 10 (4) (1968) 422–437.

[20] R. Jennane, R. Harba, G. Jacquet, Méthodes d'analyse du mouvement brownien fractionnaire: théorie et résultats comparatifs, Traitement du Signal 18 (5-6) (2001) 419–436.

[21] F. Serinaldi, Use and misuse of some hurst parameter estimators applied to stationary and non-stationary financial time series, Physica A: Statistical Mechanics and its Applications 389 (14) (2010) 2770–2781.

[22] Y. Zou, R. V. Donner, N. Marwan, J. F. Donges, J. Kurths, Complex network approaches to nonlinear time series analysis, Physics Reports 787 (2019) 1–97.

[23] M. G. Ravetti, L. C. Carpi, B. A. Gonçalves, A. C. Frery, O. A. Rosso, Distinguishing noise from chaos: objective versus subjective criteria using horizontal visibility graph, PloS one 9 (9) (2014) e108004.

[24] W.-J. Xie, W.-X. Zhou, Horizontal visibility graphs transformed from fractional Brownian motions: Topological properties versus the Hurst index, Physica A Stat. Mech. Appl. 390 (20) (2011) 3592–3601.

[25] B. Luque, L. Lacasa, F. Ballesteros, J. Luque, Horizontal visibility graphs: Exact results for random time series, Phys. Rev. E 80 (4) (2009) 046103.

[26] T. A. Schieber, et al., Information theory perspective on network robustness, Physics Lett. A 380 (3) (2016) 359–364.

[27] T. Nie, Z. Guo, K. Zhao, Z.-M. Lu, Using mapping entropy to identify node centrality in complex networks, Physica A Stat. Mech. Appl. 453 (2016) 290–297.

[28] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 623–656.
URL https://api.semanticscholar.org/CorpusID:55379485

[29] R. A. Fisher, On the mathematical foundations of theoretical statistics, Philosophical Trans. Roy. Soc. London. Series A 222 (594-604) (1922) 309–368.

[30] P. Sánchez-Moreno, R. Yánez, J. Dehesa, Discrete densities and Fisher information, in: Proc. Int. Conf. Difference Equations and Applications. Istanbul, Turkey, 2009, pp. 291–298.

[31] A. Serletis, A. A. Rosenberg, The hurst exponent in energy futures prices, Physica A: Statistical Mechanics and its Applications 380 (2007) 325–332.

[32] J. Alvarez-Ramirez, R. Escarela-Perez, Time-dependent correlations in electricity markets, Energy Economics 32 (2) (2010) 269–277.

[33] P. Gilfriche, V. Deschodt-Arsac, E. Blons, L. M. Arsac, Frequency-specific fractal analysis of postural control accounts for control strategies, Frontiers in Physiology 9 (2018) 293.

[34] P. Flandrin, Wavelet analysis and synthesis of fractional brownian motion, IEEE Transactions on information theory 38 (2) (1992) 910–917.

[35] C.-F. Li, Rescaled-range and power spectrum analyses on well-logging data, Geophysical Journal International 153 (1) (2003) 201–212.

[36] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, A. L. Goldberger, Mosaic organization of dna nucleotides, Physical review e 49 (2) (1994) 1685.