Book title

Chapter 1

MAPPING TIME SERIES TO NETWORKS: A BRIEF OVERVIEW OF VISIBILITY ALGORITHMS

***Lucas Lacasa**[*], **Bartolo Luque**[**]*
[*] IFISC, Instituto de Física Interdisciplinar y Sistemas Complejos
CSIC-UIB Mallorca, Spain
[**] Dpto. Matemática Aplicada y Estadística,
ETSI Aeronáuticos, Universidad Politécnica de Madrid, Spain

Abstract

In the last years a new approach for making time series analysis has appeared. This new approach considers the mapping of time series to networks, in order to characterize the structure of time series (and therefore the dynamics that generated the series) via characterization of the associated network. It makes use of several metrics recently developed in the so called Complex Network theory, and makes a bridge between this latter discipline and the more general aspects of time series analysis and nonlinear dynamics. While several possibilities have been proposed, here we focus on the so called visibility algorithm, which has received much attention in the last two years. This method has been shown to be well defined as time series correlations are inherited in the associated visibility graphs, opening the possibility of characterizing complex signals from a brand new viewpoint. We will make an overview of the method, addressing two different possible mapping criteria (i.e. the visibility algorithm and the horizontal visibility algorithm) for mapping series into graphs. This method captures the correlations of a time series and encodes it in the topology of the associated graph. After presenting the mapping properties, we will address within the visibility algorithms three fundamental problems in nonlinear time series analysis, namely (i) the network characterization of fractional Brownian motion, (ii) the characterization of uncorrelated processes, and (iii) the discrimination between randomness and chaos through the visibility algorithm.

[*]E-mail address: lucas@ifisc.uib-csic.es, bartolome.luque@upm.es

Keywords: Time series analysis, Complex Networks, Nonlinear Dynamics.

1. Introduction

In the last years a brand new approach in time series analysis has appeared. In a nutshell, time series are mapped into a network representation (where the connections between nodes capture the series structure according to the mapping criteria) and graph theoretical tools are subsequently employed to characterize the properties of the series. Among some possibilities (see for instance [1, 2, 3, 4, 5]) the so called visibility algorithm [3] has received much attention, since it has been shown that it is a well defined method, in the sense that series correlations (including periodicity, fractality or chaoticity) are captured by the algorithm and translated in the associated visibility graph, opening the possibility of building bridges between time series analysis, nonlinear dynamics, and graph theory, with potential applications in a plethora of disciplines. The method, inspired in the concept of visibility [6], proceeds by mapping time series into graphs according to a specific geometric criterion, in order to make use of complex networks techniques [8, 9, 10, 11] for characterize time series. In this chapter we will make an overview of the main results, including the presentation of different possible mapping criteria (that lead to somewhat different -albeit related- visibility graphs) and the subsequent characterization of periodic, random, fractal and chaotic series. This topic is currently evidencing a burst of research activity, and in this sense we assume that the present chapter will not in any case be a final overview but instead, an introduction and a summary of recent results.

2. The Visibility algorithms: two different mapping criteria

Next we will describe the two algorithms for mapping temporal series to graphs, this will be our main concern in this chapter.

2.1. The visibility algorithm

In short, a visibility graph is obtained from the mapping of a time series of n data into a network of n nodes (where each datum is associated to a specific node and where temporal order is preserved in the node labelling) according to the following visibility criterion (see figure 1: Two arbitrary data (t_a, x_a) and (t_b, x_b) in the time series have visibility, and consequently become two nodes in the associated graph, if any other data (t_c, x_c) such that $t_a < t_c < t_b$ fulfills

$$x_c < x_a + (x_b - x_a)\frac{t_c - t_a}{t_b - t_a}. \quad (1)$$

Some basic properties of the mapping include connectedness (the visibility graph is alwas connected by definition), undirectedness (undirected network), and invariance under affine transformations (see figure 2).

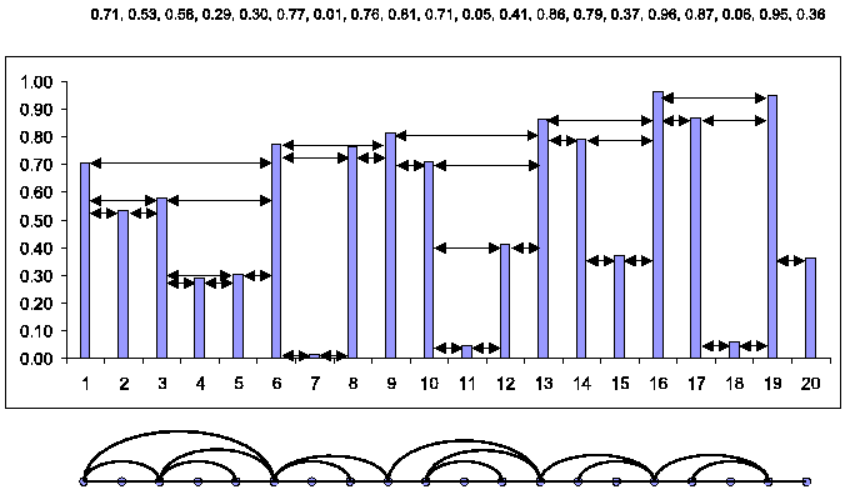

Figure 1. Illustrative example of the visibility algorithm. In the upper part we plot a periodic time series and in the bottom part we represent the graph generated through the visibility algorithm. Each datum in the series corresponds to a node in the graph, such that two nodes are connected if their corresponding data heights fulfill the visibility criterion of equation 1.

2.2. The Horizontal Visibility algorithm

The *horizontal* visibility algorithm maps time series into graphs and it is defined as follows. Let $\{x_i\}_{i=1..N}$ be a time series of N data. The algorithm assigns each datum of the series to a node in the horizontal visibility graph (graph from now on). Two nodes i and j in the graph are connected if one can draw a horizontal line in the time series joining x_i and x_j that does not intersect any intermediate data height (see figure 3 for a graphical illustration). Hence, i and j are two connected nodes if the following geometrical criterion is fulfilled within the time series:

$$x_a, x_b > x_c \text{ for all } c \text{ such that } a < c < b\,. \tag{2}$$

This algorithm is a simplification of the former algorithm. As a matter of fact, notice that given a time series, its horizontal visibility graph is always a subgraph of its associated visibility graph.

Accordingly, as in the former case, the horizontal visibility graph associated to a time series is always:

(*i*) Connected: each node sees at least its nearest neighbors (left-hand side and right-hand side).

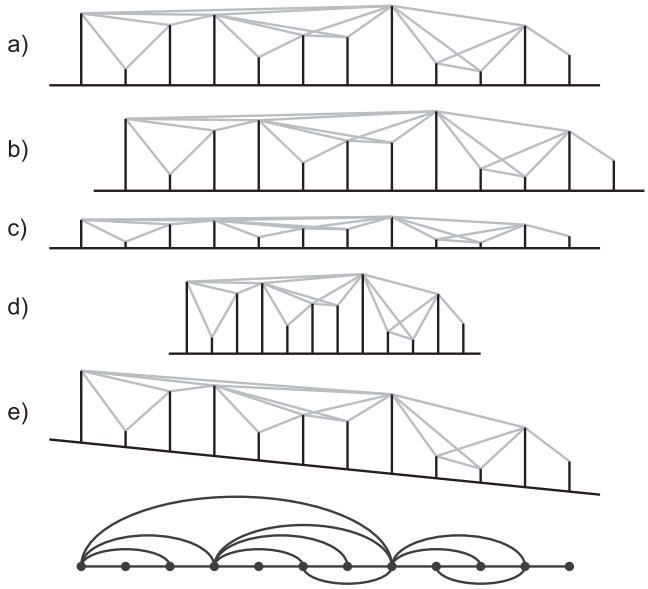

Figure 2. The visibility graph of a time series remains invariant under several transformation of the time series: a) original time series with visibility links b) translation of the data c) vertical rescaling d) horizontal rescaling e) addition of a linear trend to the data. As can be seen in the botom figure, in all these cases the visibility graph remains invariant.

(*ii*) Invariant under affine transformations of the series data: the visibility criterium is invariant under rescaling of both horizontal and vertical axis, as well as under horizontal and vertical translations.

Some other properties can be stated, namely:

(*iii*) Reversible/Irreversible character of the mapping: note that some information regarding the time series is inevitably lost in the mapping from the fact that the network structure is completely determined in the (binary) adjacency matrix. For instance, two periodic series as $..., 3, 1, 3, 1, ...$ and $..., 3, 2, 3, 2, ...$ with the same period $T = 2$ would have the same visibility graph, albeit being quantitatively different. Although the spirit of the visibility graph is to focus on time series structural properties (periodicity, fractality, etc.), the method can be generalized by making use of weighted networks (where the adjacency matrix is not binary and the weights determine the height difference of the associated data for example), if we eventually need to quantitatively distinguish time series like above, for instance. Using weighted networks, the algorithm converts to a reversible one. This feature will be adressed in further work.

(*iv*) Undirected/directed character of the mapping: Although this algorithm generates undirected graphs, note that one could also extract a directed graph (related to the temporal axis direction) in such a way that for a given node one should distinguish two different degrees: an ingoing degree k_{in}, related to how many nodes see a given node i, and an

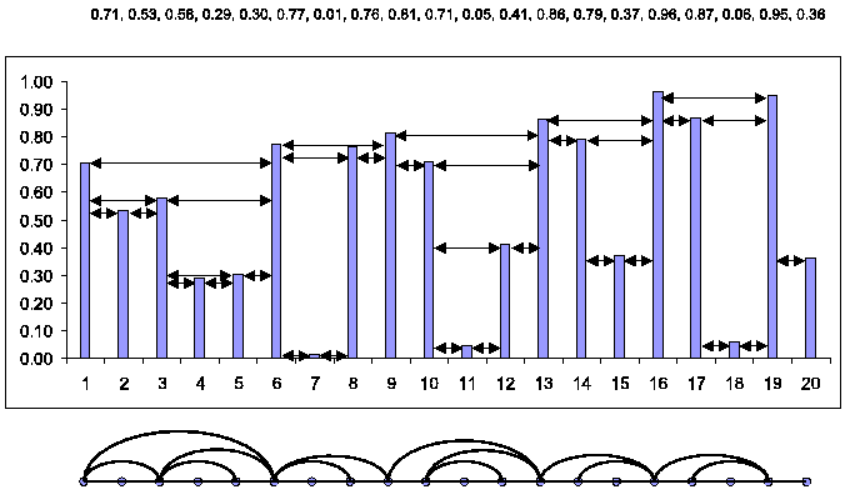

Figure 3. Illustrative example of the horizontal visibility algorithm. In the upper part we plot a time series and in the bottom part we represent the graph generated through the horizontal visibility algorithm. Each datum in the series corresponds to a node in the graph, such that two nodes are connected if their corresponding data heights are larger than all the data heights between them. The data values (heights) are made explicit in the top.

outgoing degree k_{out}, that is the number nodes that node i sees. In that situation, if the direct visibility graph extracted from a given time series is not invariant under time reversion (that is, if $P(k_{in}) \neq P(k_{out})$), one could assert that the process that generated the series is not conservative. In a first approximation we have studied the undirected version, and the directed one will be eventually addressed in further work. While the undirected choice seems to violate causality, note that the same 'causality violation' is likely to take place when performing the DFT of a time series, for instance.

(*vi*) Comparison between geometric criteria: Note that the geometric criterion defined for the horizontal visibility algorithm is more 'visibility restrictive' than its analogous for the general case. That is to say, the nodes within the horizontal visibility graph will have 'less visibility' than their counterparts within the visibility graph. While this fact does not have an impact on the qualitative features of the graphs, quantitatively speaking, horizontal visibility graphs will have typically 'less statistics'. For instance, it has been shown that the degree distribution $P(k)$ of the visibility graph associated to a fractal series is a power law $P(k) \sim k^{-\gamma}$, such that the Hurst exponent H of the series is linearly related to γ [12]. Now, for practical purposes it is more recommendable to make use of the visibility algorithm (in detriment of the horizontal version) when measuring the Hurst exponent of a fractal series, since a good estimation of γ requires at least two decades of statistics in $P(k)$, something which is more likely within the visibility algorithm. In what follows we will show that the simplicity of the horizontal version of the algorithm -which

is computationally faster than the original- allows analytical tractability, and nonetheless, this latter method is well fitted to distinguish different degrees of chaos from a sequence of uncorrelated random variables.

2.3. What is the visibility algorithm actually mapping?

In order to deepen on the geometric interpretation of the visibility graph, let us focus on a periodic series. It is straightforward that its visibility graph is a concatenation of a motif: a repetition of a pattern (see figure 1). Now, which is the degree distribution $P(k)$ of this visibility graph? Since the graph is just a motif's repetition, the degree distribution will be formed by a finite number of non-null values, this number being related to the period of the associated periodic series. This behavior reminds us the Discrete Fourier Transform (DFT), which for periodic series is formed by a finite number of peaks (vibration modes) related to the series period. Using this analogy, we can understand the visibility algorithm as a geometric (rather than integral) transform. Whereas a DFT decomposes a signal in a sum of (eventually infinite) modes, the visibility algorithm decomposes a signal in a concatenation of graph's motifs, and the degree distribution simply makes a histogram of such 'geometric modes'. While the time series is defined in the time domain and the DFT is defined on the frequency domain, the visibility graph is defined on the 'visibility domain'. This is, of course, a hand-waving analogy and further work should study its extent rigorously. For instance, this transform is not, as presented, a reversible one. Reversibility can however be easily obtained weighting the links in the visibility graph with the slope of the visibility line that links the associated data heights. The weighted version of the algorithm and its geometric transform nature are two important open problems. At this point we can comment that whereas a generic DFT fails to capture the presence of nonlinear correlations in time series (such as the presence of chaotic behavior), we will show that the visibility algorithm can clearly distinguish between white noise (i.e. a sequence of identically independent random variables) and chaotic series.

3. Some examples: periodic, random, fractal series

The key question is to know whether the associated graph inherits some structure of the time series, and consequently if the process which generated the time series may be characterized using graph theory. In a first step we will consider periodic series. As a matter of fact, the example plotted in figure 1 is nothing but a periodic series with period 4. The associated visibility graph is ordered, as long as it is constructed by periodic repetition of a pattern. The degree distribution of this graph is formed by a finite number of peaks related to the series period, much in the vein of the Fourier Power Spectrum of a time series. Generically speaking, all periodic time series are mapped into motif-like graphs, the discrete degree distribution being the fingerprint of the time series periods. In the case of periodic time series, its regularity seems therefore to be conserved or inherited structurally in the graph by means of the visibility map.

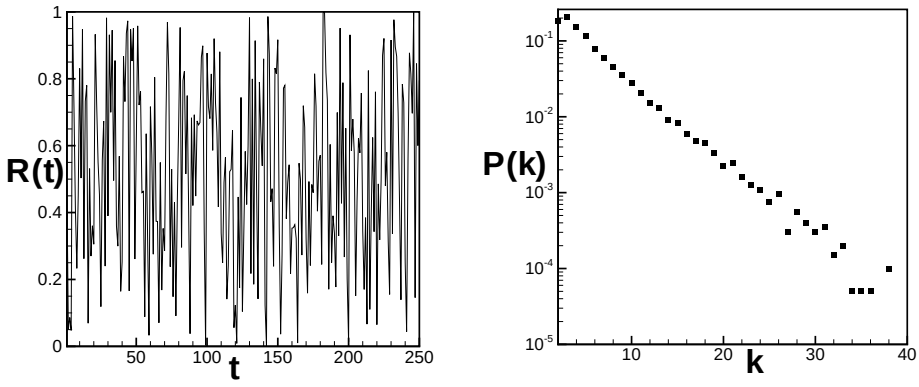

Figure 4. Left figure: First 250 values of $R(t)$, a random series of 10^6 data extracted from $U[0,1]$. Right figure: degree distribution of the visibility graph associated to $R(t)$ (plotted in semi-log). While the beginning of the curve approaches the result of a Poisson process, the tail is clearly exponential. This behavior is due to data with large values (rare events), which are the hubs.

As an opposite to periodic series, in a second step we tackle a random series $R(t)$ of 10^6 data extracted from an uniform distribution in $[0,1]$. Although one would expect in a first moment a Poisson degree distribution in this case (as for uncorrelated random graphs [13]), a random time series has indeed some correlation, since it is an ordered set. In fact, let k_t be the connectivity of the node associated to the data t. If k_t is large (related to the fact that the data has a large value and that consequently it has large visibility), one would expect that k_{t+1} would be relatively small, since the time series is random and two consecutive data with a large value are not likely to occur. It is indeed due to these 'unlikely' large values (the hubs) that the tail of the degree distribution deviates from the Poisson distribution. Two large values in the series data can be understood as two rare events in the random process. The time distribution of these events is indeed exponential [14], therefore we should expect the tail of the degree distribution in this case to be exponential instead of Poissonian, as long as the form of this tail is related to the hub's distribution. In the left side of figure 4 we depict the first 250 values of $R(t)$. In the right side we plot the degree distribution $P(k)$ of its visibility graph. The tail of this distribution fits quite well an exponential distribution, as expected. In a few pages we will tackle the full characterization of random series.

Hitherto, ordered (periodic) series convert into regular graphs, and random series convert into exponential random graphs: order and disorder structure in the time series seem to be inherited in the topology of the visibility graph. Thus, the question arises: What kind of visibility graph is obtained from a fractal time series? This question is in itself interesting at the present time. Recently, the relationship between self-similar and scale-free networks [7, 8] has been intensively discussed [15, 16, 17, 18]. While complex networks usually exhibit the Small-World property [19] and cannot be consequently size-invariant, it has been recently shown [15] that applying fitted box-covering methods and renormalization procedures, some real networks actually exhibit self-similarity.

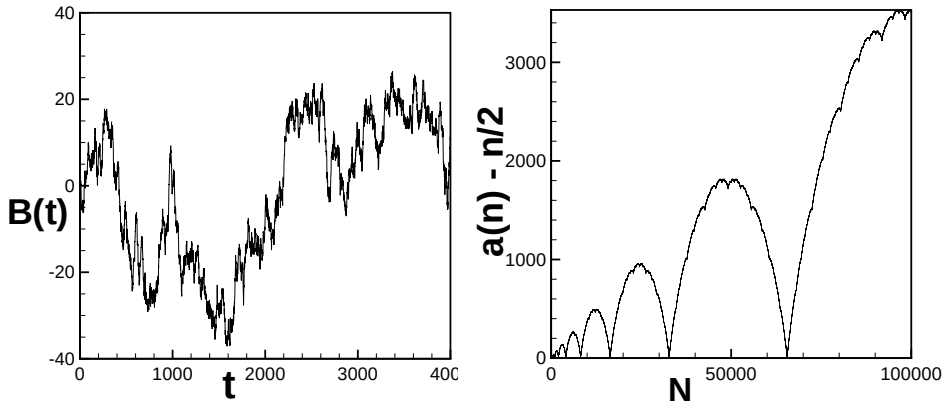

Figure 5. Left: First 4000 data from a Brownian series of 10^6 data. Right: First 10^5 data from a Conway series of $4 \cdot 10^6$ data.

So, whereas self-similarity seems to imply scale-freeness, the opposite is not true in general.

In order to explore these issues in more detail, the following two fractal series will be considered: the well-known Brownian motion $B(t)$ and the Conway series [20]. While the Brownian motion represents a well-known case of self-affinity (indeed, the following relation holds: $B(t) = a^{1/2}B(t/a)$, the Conway series $a(n) - n/2$ is the recursively generated fractal series from:

$$\begin{aligned} &a(1) = a(2) = 1 \\ &a(n) = a(a(n-1)) + a(n - a(n-1));\ n > 2. \end{aligned} \tag{3}$$

In figure 5 we show for illustrative purposes two examples of Brownian and Conway series. In figures 6 and 7 we plot the degree distribution $P(k)$ of their respective visibility graphs and their mean path length $L(N)$ as a function of the series length. First, both series have visibility graphs with degree distributions that correspond to power laws of the shape $P(k) \sim k^{-\alpha}$, where we get different exponents in each case: this result enhances the fact that in the context of the visibility algorithm, power law degree distributions (that is, scale free networks)arise naturally from fractal series. Moreover, this relation seems to be robust as long as the preceding examples show different kinds of fractality: while $B(t)$ stands for a stochastic self-affine fractal, the Conway series is a deterministic series recursively generated. On the other hand, while the Brownian visibility graph seems to evidence the Small-World effect (right figure 6) as $L(N) \sim \log(N)$, the Conway series shows in turn a self-similar relation (right figure 7) of the shape $L(N) \sim N^{\beta}$. This fact can be explained in terms of the so called hub repulsion phenomenon [17]: visibility graphs associated to stochastic fractals such as the Brownian motion or generic noise series do not evidence repulsion between hubs (in these series it is straightforward that the data with

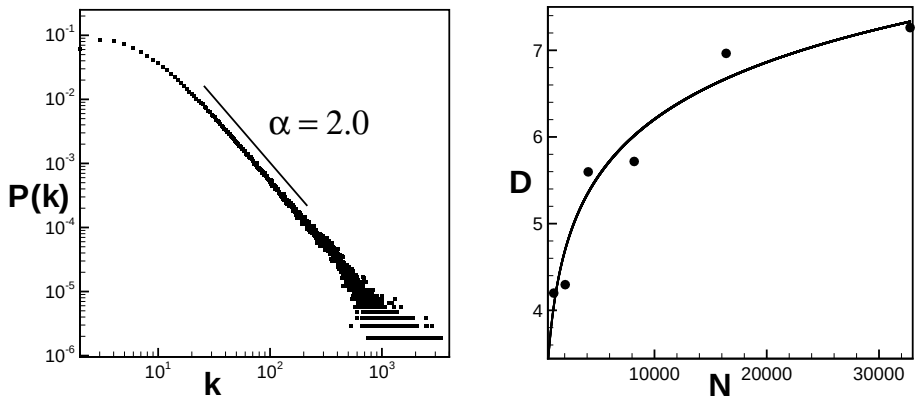

Figure 6. The left figure stands for the degree distribution of the visibility graph associated to the Brownian motion. This one is a power law $P(k) \sim k^{-\alpha}$ with $\alpha = 2.00 \pm 0.01$. In the right part of the figure we plot the mean path length of this network as a function of the network size N. The best fitting provides a logarithmic scaling of the shape $D(N) = -2.54 + 0.95 \ln(N)$. This network shows Small-World effect in addition to be scale-free.

highest values would stand for the hubs, and these data would have visibility between each other), and consequently won't be fractal networks following Song et al. [17]. On the other hand, the Conway series actually evidence hub repulsion: this series is concave-shaped and consequently the highest data won't in any case stand for the hubs; the latter ones would be located much likely in the monotonic regions of the series, which are indeed hidden from each other (effective repulsion) across the series local maxima. The Conway visibility graph is thus fractal.

Since a fractal series is characterized by its Hurst exponent, we may argue that the visibility graph can actually distinguish different types of fractality, something that will be explored in detail in the following section.

4. Characterization of fractal series: visibility graphs of fractional Brownian motion

Self-similar processes such as fractional Brownian motion (fBm) [21] are currently used to model fractal phenomena of different nature, ranging from Physics or Biology to Economics or Engineering. To cite a few, fBm has been used in models of electronic delocalization [22], as a theoretical framework to analyze turbulence data [23], to describe geologic properties [24], to quantify correlations in DNA base sequences [25], to characterize physiological signals such as human heartbeat [26] or gait dynamics [27], to model economic data [28] or to describe network traffic [29, 30, 31]. Fractional Brownian motion $B_H(t)$ is a non-stationary random process with stationary self-similar increments (fractional Gaussian noise) that can be characterized by the so called Hurst exponent,

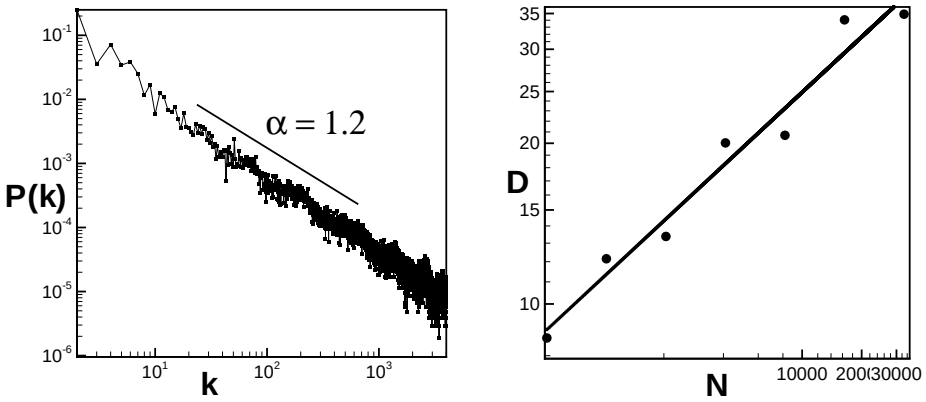

Figure 7. The left figure stands for the degree distribution of the visibility graph associated to the Conway series. This one is a power law $P(k) \sim k^{-\alpha}$ with $\alpha = 1.2 \pm 0.1$. The mean path length as a function of the size N is depicted in the right part of the figure. The best fitting provides a power law scaling of the shape $D(N) = 1.04N^{0.34}$. Then, this network is scale-invariant.

$0 < H < 1$. The one-step memory Brownian motion is obtained for $H = \frac{1}{2}$, whereas time series with $H > \frac{1}{2}$ shows persistence and anti-persistence if $H < \frac{1}{2}$.

While different fBm generators and estimators have been introduced in the last years, the community lacks consensus on which method is best suited for each case. This drawback comes from the fact that fBm formalism is exact in the infinite limit, i.e. when the whole infinite series of data is considered. However, in practice, real time series are finite. Accordingly, long range correlations are partially broken in finite series, and local dynamics corresponding to a particular temporal window are overestimated. The practical simulation and the estimation from real (finite) time series is consequently a major issue that is, hitherto, still open. An overview of different methodologies and comparisons can be found in [32, 33, 34, 35, 36, 37, 38] and references therein.

In this section we show that the visibility graphs derived from generic fBm series are also scale free. This robustness goes further, and we prove that a linear relation between the exponent γ of the power law degree distribution in the visibility graph and the Hurst exponent H of the associated fBm series exists. Therefore, the visibility algorithm provides an alternative method to compute the Hurst exponent and then, to characterize fBm processes.

In fig.8 we have depicted in log-log the degree distribution of the visibility graph associated with three artificial fBm series of 10^5 data, namely an anti-persistent series with $H = 0.3$ (triangles), a memoryless Brownian motion with $H = 0.5$ (squares) and a persistent fBm with $H = 0.8$ (circles). As can be seen, these distributions follow a power law $P(k) \sim k^{-\gamma}$ with decreasing exponents $\gamma_{0.3} > \gamma_{0.5} > \gamma_{0.8}$.
In order to compare γ and H appropriately, we have calculated the exponent of different

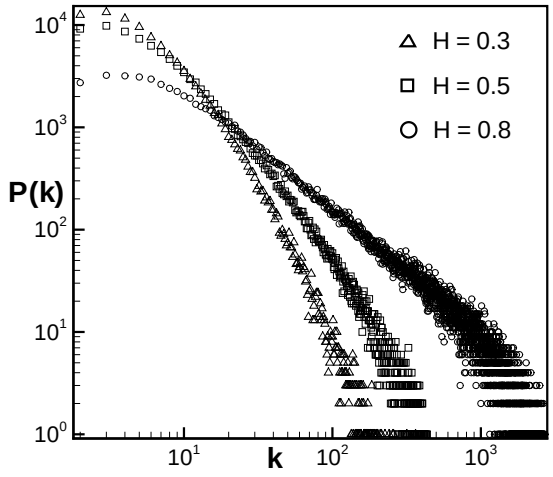

Figure 8. Degree distribution of three visibility graphs, namely (i) triangles: extracted from a fBm series of 10^5 data with $H = 0.3$, (ii) squares: extracted from a fBm series of 10^5 data with $H = 0.5$, (iii) circles: extracted from a fBm series of 10^5 data with $H = 0.8$. Note that distributions are not normalized. The three visibility graphs are scale-free since their degree distributions follow a power law $P(k) \sim k^{-\gamma}$ with decreasing exponents $\gamma_{0.3} > \gamma_{0.5} > \gamma_{0.8}$.

scale free visibility graphs associated with fBm artificial series of 10^4 data with $0 < H < 1$ generated by a wavelet based algorithm [39]. Note at this point that some bias is inevitably present since artificial series generators are obviously not exact, and consequently the nominal Hurst exponents have an associated error [40]. For each value of the Hurst parameter we have thus averaged the results over 10 realizations of the fBm process. We have estimated exponent γ in each case through Maximum Likelihood Estimation (MLE) [41]:

$$\gamma = 1 + n\left[\sum_{i=1}^{n} \log \frac{x_i}{x_{min}}\right]^{-1}, \tag{4}$$

where n is total number of values taken into account, $x_i,\ i = 1, .., n$ are the measured values and x_{min} corresponds to the smallest value of x for which the power law behavior holds. In fig.9 we have represented the relation between γ and H (black circles). As can be seen, a roughly linear relation holds (the best linear fitting is $\gamma = 3.1 - 2H$).

The estimation of H through the visibility method has been applied recently with success to several temporal series: turbulence [42], hurricanes [43] and stock markets [44]. That fBm yields scale free visibility graphs is not that surprising. The most highly connected nodes (hubs) are the responsible for the heavy tailed degree distributions. Within fBm series, hubs are related to extreme values in the series, since a data with a very large value has typically a large connectivity, according to eq. 1. In order to calculate the tail of the distribution we consequently need to focus on the hubs, and thus calculate the probability that an extreme

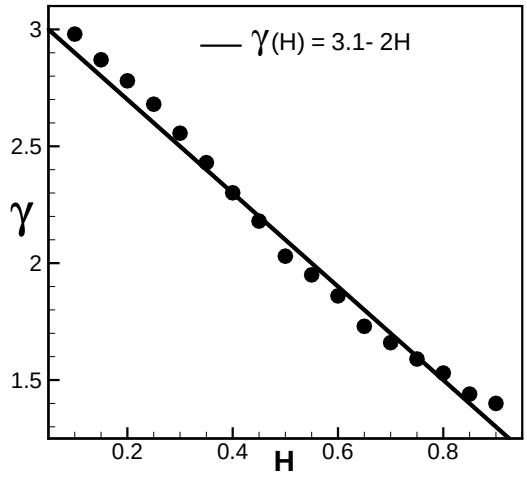

Figure 9. (Black dots) Numerical estimation of exponent γ of the visibility graph associated with a fBm series with exponent H. In each case γ is averaged over 10 realizations of a fBm series of 10^4 data, in order to avoid non-stationary biases (the error bars are included in the dot size). The solid line corresponds to the theoretical prediction $\gamma(H) = 3-2H$. Note that deviations from the theoretical law take place for values of $H > 0.5$ and $H < 0.5$ (strongly correlated or anti-correlated series), where fBm generators evidence finite-size accuracy problems [40], these being more acute the more we move away from the non-correlated case $H = 0.5$.

value has a degree k. Suppose that at time t the series reaches an extreme value (a hub) $B_H(t) = h$. The probability of this hub to have degree T is

$$P(T) \sim P_{fr}(T)r(T), \tag{5}$$

where $P_{fr}(T)$ provides the probability that after T time steps, the series returns to the same extreme value, i.e. $B(t+T) = h$ (and consequently the visibility in t gets truncated in $t+T$), and $r(T)$ is the percentage of nodes between t and $t+T$ that t may see. $P_{fr}(T)$ is nothing but the first return time distribution, which is known to scale as $P_{fr}(T) \sim T^{H-2}$ for fBm series [45]. On the other hand, the percentage of visible nodes between two extreme values is related to the roughness of the series in that basin, that is, to the way that a series of T time steps folds. This roughness is encoded in the series standard deviation [21], such that intuitively, we have $r(T) \sim T^H/T = T^{H-1}$ (this fact has been confirmed numerically). Finally, notice that in this context $T \equiv k$, so eq.5 converts into

$$P(k) \sim k^{H-2}k^{H-1} = k^{2H-3}, \tag{6}$$

what provides a linear relation between the exponent of the visibility graph degree distribution and the Hurst exponent of the associated fBm series:

$$\gamma(H) = 3 - 2H, \tag{7}$$

in good agreement with our previous numerical results. Note in figure 9 that numerical results obtained from artificial series deviate from the theoretical prediction for strongly-correlated ones ($H > 0.5$ or $H < 0.5$). This deviation is related to finite size effects in the generation of finite fBm series [40], and these effects are more acute the more we deviate from the non-correlated case $H = 0.5$. In any case, a scatter plot of the theoretical (eq.7) versus the empirical estimation of $\gamma(H)$ provides statistical conformance with a correlation coefficient $c = 0.99$.

5. Theorems and exact results for the horizontal visibility graphs of random time series

In this section we make a turn and focus on the *horizontal* version of the algorithm. This is a geometrically simpler version that allows full analytical tractability [46] as we will see. As a first step towards a theory of visibility algorithms, we will focus on random uncorrelated series.

5.1. Degree distribution

Consider a bi-infinite time series created from a random variable X with probability distribution $f(x)$ with $x \in [0, 1]$ and let us construct its associated horizontal visibility graph (note that if the distribution's support is a generic interval $x \in [a, b]$, we can rescale to $[0, 1]$ without loss of generality since the associated graph remains invariant, and this also applies to unbounded supports). For convenience, we will label a generic datum x_0 as the 'seed' datum from now on. In order to derive the degree distribution $P(k)$ [7] of the associated graph, we need to calculate the probability that an arbitrary datum with value x_0 has visibility of exactly k other data. If x_0 has visibility of k data, there always will exist two 'bounding data', one on the right-hand side of x_0 and another one on its left-hand side, such that the $k - 2$ remaining visible data will be located inside that window (in fact, $k = 2$ is the minimum possible degree). As these 'inner' data should appear sorted by size, there are exactly $k - 1$ different possible configurations $\{C_i\}_{i=0..k-2}$, where the index i determines the number of inner data on the left-hand side of x_0 (see figure 10 for an illustration of the possible configurations and a labelling recipe of the data in the case $k = 4$). Accordingly, C_i corresponds to the configuration for which i inner data are placed at the left-hand side of x_0, and $k - 2 - i$ inner data are placed at its right-hand side. Each of these possible configurations have an associated probability $p_i \equiv p(C_i)$ that will contribute to $P(k)$ such that

$$P(k) = \sum_{i=0}^{k-2} p_i. \tag{8}$$

Before trying to find a general relation for $P(k)$ and for illustrative purposes, let us study some particular cases. The first and simplest case is $P(k = 2)$, that is, the probability that the seed data has two and only two visible data, the minimum degree. These obviously will be the bounding data, that we will label x_{-1} and x_1 for left-hand side and right-hand side of the seed respectively. The probability that x_0 sees $k \geq 2$ is 1 by construction, since the

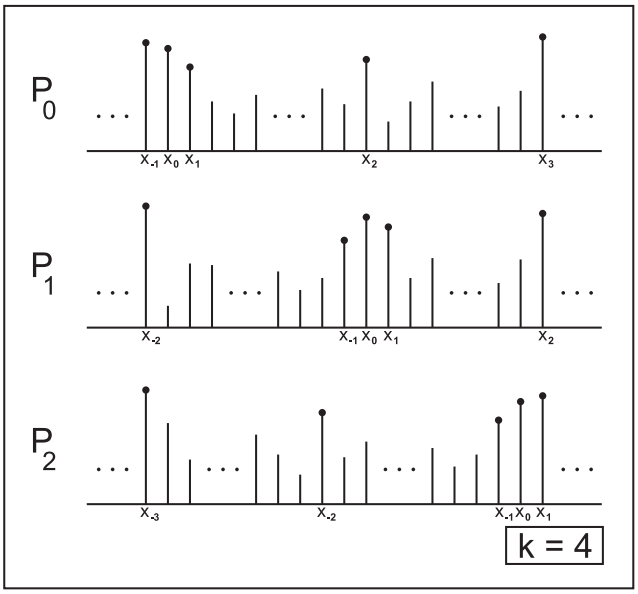

Figure 10. Set of possible configurations for a seed data x_0 with $k = 4$. Observe that the sign of the subindex in x_i indicates if the data is located whether at left-hand side of x_0 (sign minus) or at right-hand side. Accordingly, the bounding's data subindex directly indicates the amount of data located in that side. For instance, C_0 is the configuration where none of the $k - 2 = 2$ inner data are located in the left-hand side of x_0, and therefore the left bounding data is labeled as x_{-1} and the right bounding data is labeled as x_3. C_1 is the configuration for which an inner data is located in the left-hand side of x_0 and another inner data is located in its right-hand side. Finally, C_2 is the configuration for which both inner data are located in the left-hand side of the seed. Notice that an arbitrary number of hidden data can be eventually located among the inner data, what is schematically represented in the figure as a row of vertical lines.

horizontal visibility algorithm assures that any data will always have visibility of its first neighbors. Now, in order to assure that $k = 2$, we have to impose that the bounding data neighbors have a larger height than the seed, that is, $x_{-1} \geq x_0$ and $x_1 \geq x_0$. Then,

$$P(k = 2) = \text{Prob}(x_{-1}, x_1 \geq 0) = \int_0^1 f(x_0)dx_0 \int_{x_0}^1 f(x_1)dx_1 \int_{x_0}^1 f(x_{-1})dx_{-1}. \quad (9)$$

Now, the cumulative probability distribution function $F(x)$ of any probability distribution $f(x)$ is defined as

$$F(x) = \int_0^x f(x')dx', \quad (10)$$

where $dF(x)/dx = f(x)$, $F(0) = 0$ and $F(1) = 1$. In particular, the following relation between f and F holds:

$$f(x)F^{n-1}(x) = \frac{1}{n}\frac{dF^n(x)}{dx}. \quad (11)$$

We can accordingly rewrite and compute equation 9 as

$$P(k=2) = \int_0^1 f(x_0)[1-F(x_0)]^2 dx_0 = \frac{1}{3}, \tag{12}$$

independently of the shape of the probability distribution $f(x)$.

Let us proceed by tackling the case $P(k=3)$, that is, the probability that the seed has three and only three visible data. Two different configurations arise: C_0, in which x_0 has 2 bounding visible data (x_{-1} and x_2 respectively) and a right-hand side inner data (x_1), and the same for C_1 but with the inner data being place at the left-hand side of the seed, so

$$P(k=3) = p(C_0) + p(C_1) \equiv p_0 + p_1.$$

Notice at this point that an arbitrary number r of hidden data $n_1, n_2 ... n_r$ can eventually be located between the inner data and the bounding data, and this fact needs to be taken into account in the probability calculation. The geometrical restrictions for the n_j hidden data are $n_j < x_1,\ j = 1, ..., r$ for C_0 and $m_j < x_{-1},\ j = 1, ..., s$ for C_1. Then,

$$\begin{aligned} p_0 &= \text{Prob}\Big((x_{-1}, x_2 \geq x_0) \cap (x_1 < x_0) \cap (\{n_j < x_1\}_{j=1,...,r})\Big), \\ p_1 &= \text{Prob}\Big((x_{-2}, x_1 \geq x_0) \cap (x_{-1} < x_0) \cap (\{m_j < x_{-1}\}_{j=1,...,s})\Big). \end{aligned} \tag{13}$$

Now, we need to consider every possible hidden data configuration (C_0 without hidden data, C_0 with a single hidden data, C_0 with two hidden data, and so on, and the same for C_1). With a little calculus we come to

$$\begin{aligned} p_0 = &\int_0^1 f(x_0)dx_0 \int_{x_0}^1 f(x_{-1})dx_{-1} \int_{x_0}^1 f(x_2)dx_2 \int_0^{x_0} f(x_1)dx_1 + \\ &\sum_{r=1}^{\infty} \int_0^1 f(x_0)dx_0 \int_{x_0}^1 f(x_{-1})dx_{-1} \int_{x_0}^1 f(x_2)dx_2 \int_0^{x_0} f(x_1)dx_1 \prod_{j=1}^{r} \int_0^{x_1} f(n_j)dn_j \end{aligned}$$

where the first term corresponds the contribution of a configuration with no hidden data and the second sums up the contributions of r hidden data. Making use of the properties of the cumulative distribution $F(x)$ we arrive to

$$p_0 = \int_0^1 f(x_0)dx_0 \int_{x_0}^1 f(x_{-1})dx_{-1} \int_{x_0}^1 f(x_2)dx_2 \int_0^{x_0} \frac{f(x_1)}{1-F(x_1)}dx_1, \tag{14}$$

where we also have made use of the sum of a geometric series. We can find an identical result for p_1, since the last integral on equation 14 only depends on x_0 and consequently the configuration provided by C_1 is symmetrical to the one provided by C_0. We finally have

$$P(k=3) = 2p_0 = -2\int_0^1 f(x_0)(1-F(x_0))^2 \ln(1-F(x_0))dx_0 = \frac{2}{9}, \tag{15}$$

where the last calculation also involves the change of variable $z = 1 - F(x)$. Again, the result is independent of $f(x)$.

Hitherto, we can deduce that a given configuration C_i contributes to $P(k)$ with a product of integrals according to the following rules:

- The seed data provides a contribution of $\int_0^1 f(x_0)dx_0$ (S).
- Each boundary data provides a contribution of $\int_{x_0}^1 f(x)dx$ (B).
- An inner data provides a contribution $\int_{x_j}^{x_0} \frac{f(x)dx}{1-F(x)}$ (I).

These diagrammatic-like rules allow us to schematize in a formal way the probability associated to each configuration. For instance in the case $k = 2$, $P(k)$ has a single contribution p_0 represented by the formal diagram B-S-B, while for $k = 3$, $P(k) = p_0 + p_1$ where p_0's diagram is B-S-I-B and p_1's is B-I-S-B. It seems quite straightforward to derive a general expression for $P(k)$, just by applying the preceding rules for the contribution of each C_i. However, there is still a subtle point to address that will become evident for the case $P(k = 4) = p_0 + p_1 + p_2$ (see figure 10). While in this case C_1 leads to essentially the same expression as for both configurations in $k = 3$ (and in this sense one only needs to apply the preceding rules to derive p_1), C_0 and C_2 are geometrically different configurations. These latter ones are configurations formed by a seed, two bounding and two *concatenated* inner data, and concatenated data lead to concatenated integrals. For instance, applying the same formalism as for $k = 3$, one come to the conclusion that for $k = 4$,

$$p_0 = \int_0^1 f(x_0)dx_0 \int_0^{x_0} \frac{f(x_1)dx_1}{1-F(x_1)} \int_{x_1}^{x_0} \frac{f(x_2)dx_2}{1-F(x_2)} \int_{x_0}^1 f(x_3)dx_3 \int_{x_0}^1 f(x_{-1})dx_{-1}. \quad (16)$$

While for the case $k = 3$ every integral only depended on x_0 (and consequently we could integrate independently every term until reaching the dependence on x_0), having two concatenated inner data on this configuration generates a dependence on the integrals and hence on the probabilities. For this reason, each configuration is not equiprobable in the general case, and thus will not provide the same contribution to the probability $P(k)$ ($k = 3$ was an exception for symmetry reasons). In order to weight appropriately the effect of these concatenated contributions, we can make use of the definition of p_i. Since $P(k)$ is formed by $k-1$ contributions labelled $C_0, C_1 ... C_{k-2}$ where the index denotes the number of inner data present at the left-hand side of the seed, we deduce that in general the $k-2$ inner data have the following effective contribution to $P(k)$:

- p_0 has $k-2$ concatenated integrals (right-hand side of the seed).
- p_1 has $k-3$ concatenated integrals (right-hand side of the seed) and an independent inner data contribution (left-hand side of the seed).
- p_2 has $k-4$ concatenated integrals (right-hand side of the seed) and another 2 concatenated integrals (left-hand side of the seed).
- ...
- p_{k-2} has $k-2$ concatenated integrals (left-hand side of the seed).

Observe that p_i is symmetric with respect to the seed.

Including this modification in the diagrammatic rules, we are now ready to calculate a general expression for $P(k)$. Formally,

$$P(k) = \sum_{j=0}^{k-2} [S][B]^2 [I]_j [I]_{k-2-j}, \tag{17}$$

where the sum extends to each of the $k-1$ configurations, the superindex denotes exponentiation and the subindex denotes concatenation (this latter expression can be straightforwardly proved by induction). In order to solve it, one needs to firstly calculate the concatenation of n inner data integrals $[I]_n \equiv I(n)$, that is

$$I(n) = \int_0^{x_0} \frac{f(x_1)dx_1}{1-F(x_1)} \prod_{j=1}^{n-1} \int_{x_j}^{x_0} \frac{f(x_{j+1})dx_{j+1}}{1-F(x_{j+1})}. \tag{18}$$

The calculation of $I(n)$ is easy but quite tedious. One proceeds to integrate equation 18 step by step (first $n=1$, then $n=2$, and so on), and a recurrence quickly becomes evident. One can easily prove by induction that

$$I(n) = \frac{(-1)^n}{n!} \Big[\ln\left(1-F(x_0)\right) \Big]^n. \tag{19}$$

According to the formal solution 17 and to equation 19, we finally have

$$\begin{aligned} P(k) &= \sum_{j=0}^{k-2} \frac{(-1)^{k-2}}{j!(k-2-j)!} \int_0^1 f(x_0)[1-F(x_0)]^2 [\ln(1-F(x_0))]^{k-2} dx_0 \\ &= 3^{1-k} \sum_{j=0}^{k-2} \frac{(k-2)!}{j!(k-2-j)!} = \frac{1}{3}\left(\frac{2}{3}\right)^{k-2}. \end{aligned} \tag{20}$$

Surprisingly, we can conclude that for every probability distribution $f(x)$, the degree distribution $P(k)$ of the associated horizontal visibility graph has the same exponential form.

In order to check further the accuracy of our analytical results for the case of *finite* time series, we have performed several numerical simulations. We have generated random series of 10^6 data from different distributions $f(x)$ and have generated their associated horizontal visibility graphs. In figure 11 we have plotted the degree distribution of the resulting graphs (triangles correspond to a series extracted from a uniform distribution, while circles and squares correspond to one extracted from a Gaussian and a power law distribution $f(x) \sim x^{-2}$ respectively). The solid line corresponds to the theoretical equation 20, showing a perfect agreement with the numerics.

5.2. Degree versus height

An interesting aspect worth exploring is the relation between data height and the node degree, that is, to study whether a functional relation between the height of a datum and the

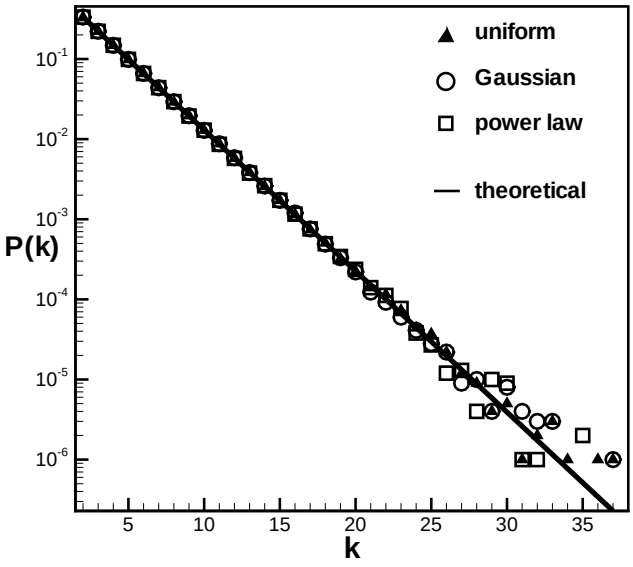

Figure 11. Semi-log plot of the degree distribution of an horizontal visibility graph associated to random series of 10^6 data extracted from a uniform distribution $f(x) = U[0, 1]$ (triangles), a Gaussian distribution (circles), and power law distribution $f(x) \sim x^{-2}$ (squares). Solid line corresponds to equation 20.

degree of its associated node holds. In this sense, let us define $P(k|x)$ as the conditional probability that a given node has degree k provided that it has height x. Observe that $P(k|x)$ can be easily deduced from eq. 20, such that

$$P(k|x) = \sum_{j=0}^{k-2} \frac{(-1)^{k-2}}{j!(k-2-j)!}[1 - F(x)]^2 \cdot [\ln(1 - F(x))]^{k-2}. \tag{21}$$

Notice that probabilities are well normalized and that $\sum_{k=2}^{\infty} P(k|x) = 1$, independently of x. Now, we can define an average value of the degree of a node associated to a datum of height x, $K(x)$, in the following way

$$K(x) = \sum_{k-2}^{\infty} kP(k|x) = 2 - 2\ln(1 - F(x)). \tag{22}$$

Since $F(x) \in [0, 1]$ and $\ln(x)$ are monotonically increasing functions, $K(x)$ will also be monotonically increasing. We can thus conclude that graph hubs (that is, the most connected nodes) are the data with largest values, that is, the extreme events of the series.

In order to check the accuracy of the theoretical prediction within finite series, in fig. 12 we have plotted (circles) the numerical values of $K(x)$ within a random series of 10^6 data extracted from a uniform distribution with $F(x) = x$. The line corresponds to eq. 22, showing a perfect agreement.

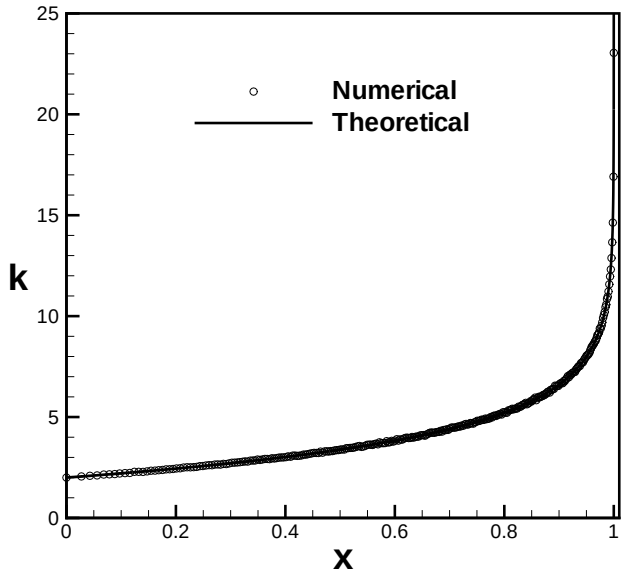

Figure 12. Average degree of a node, as a function of the associated datum's height: (circles) numeric results from a random series of 10^6 data extracted from a uniform distribution $f(x) = U[0, 1]$. The solid line corresponds to the theoretical prediction eq.22, showing a perfect agreement. It comes evident that the hubs stand for the nodes associated to the data with larger values (extreme events).

5.3. Local clustering coefficient distribution

The local clustering coefficient C [8, 9, 10, 11, 7] of an horizontal visibility graph associated to a random series can be easy deduced by means of geometrical arguments. For a given node i, C denotes the rate of nodes connected to i that are connected between each other (observe that in this subsection, C denotes clustering: do not mistake this with the 'C' (configuration) of subsection 5.1). In other words, we have to calculate from a given node i how many nodes from those visible to i have mutual visibility (triangles), normalized with the set of possible triangles $\binom{k}{2}$. In a first step, if a generic node i has degree $k = 2$, these nodes are straightforwardly two bounding data, hence having mutual visibility. Thus, in this situation there exists 1 triangle and $C(k = 2) = 1$. Now if a generic node i has degree $k = 3$, one of its neighbors will be an inner data, which will only have visibility of one of the bounding data (by construction). We conclude that in this situation we can only form 2 triangles out of 6 possible ones, thereby $C(k = 3) = 2/6$. In general, for a degree k we can form $k - 1$ triangles out of $\binom{k}{2}$ possibilities, and then:

$$C(k) = \frac{k-1}{\binom{k}{2}} = \frac{2}{k}, \tag{23}$$

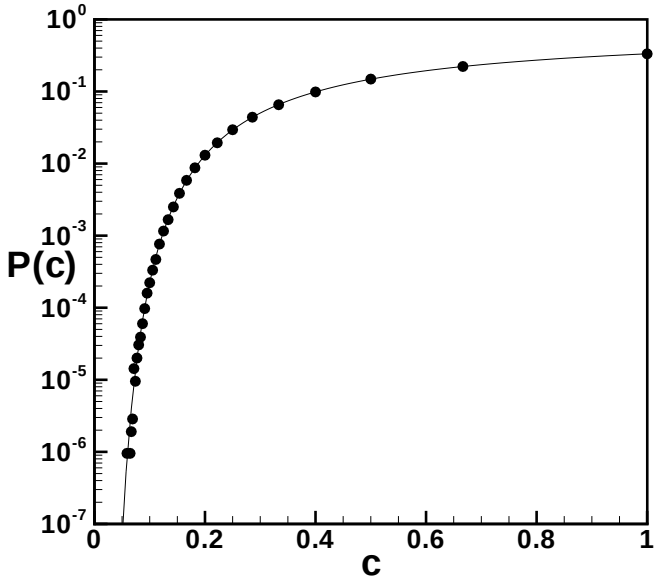

Figure 13. Semi-log plot of the clustering distribution of an horizontal visibility graph associated to random series of 10^6 data extracted from a uniform distribution $f(x) = U[0,1]$ (dots). The solid line corresponds to the theoretical prediction $P(C) = (1/3)(2/3)^{2/C-2}$. In order to avoid border effects we have imposed periodic boundary conditions in the data series.

what indicates a so called hierarchical structure [47]. This relation between k and C allows us to deduce the local clustering coefficient distribution $P(C)$:

$$P(k) = \frac{1}{3}\left(\frac{2}{3}\right)^{k-2} = P(2/C),$$
$$P(C) = \frac{1}{3}\left(\frac{2}{3}\right)^{2/C-2}. \qquad (24)$$

To check the validity of this latter relation within finite series, in figure 13 we depict the clustering distribution of an horizontal visibility graph associated to a random series of 10^6 data (dots) obtained numerically. The solid line corresponds to the theoretical prediction (equation 24), in excellent agreement with the numerics.

5.4. Long distance visibility, mean degree and mean path length

In order to derive the scaling of the mean path length [7], let us first calculate the probability $P(n)$ that two data separated by n intermediate data be two connected nodes in the graph. Consider again a time series extracted from a random variable X with probability distribution $f(x)$ and $x \in [0,1]$, and let us construct its associated horizontal visibility graph.

An arbitrary value x_0 from this series will 'see' x_n (and consequently will be connected to node x_n in the graph) iff $x_i < \min(x_0, x_n)$ for all $x_i,\ i = 1, 2, ..., n-1$. Then $P(n)$ can be expressed as:

$$P(n) = \int_0^1 \int_0^1 f(x_0) f(x_n) dx_0 dx_n \int_0^{\min(x_0,x_n)} \dots \\ \dots \int_0^{\min(x_0,x_n)} f(x_1) \dots f(x_{n-1}) dx_1 \dots dx_{n-1} \tag{25}$$

Since the integration limits are independent, rewriting $x \equiv \min(x_0, x_n)$ we have

$$P(n) = \int_0^1 \int_0^1 f(x_0) f(x_n) F^{n-1}(x) dx_0 dx_n. \tag{26}$$

We can fix x_0 and move x_n without loss of generality, such that the latter equation can be expressed as

$$P(n) = \underbrace{\int_0^1 \int_0^{x_0} f(x_0) f(x_n) F^{n-1}(x_n) dx_0 dx_n}_{\text{the minimum here is } x_n} + \underbrace{\int_0^1 \int_{x_0}^1 f(x_0) f(x_n) F^{n-1}(x_0) dx_0 dx_n}_{\text{the minimum here is } x_0} \tag{27}$$

Applying the definition of $F(x)$ and the relation 11, with a little calculus we get

$$\begin{aligned} P(n) &= \left(\frac{1}{n} - 1\right) \int_0^1 f(x_0) F^n(x_0) dx_0 + \int_0^1 f(x_0) F^{n-1}(x_0) dx_0 \\ &= \frac{2}{n(n+1)}. \end{aligned} \tag{28}$$

Observe that $P(n)$ is again independent of the probability distribution of the random variable X. Notice that the latter result can also be obtained, alternatively, with a purely combinatorial argument that reads as it follows. Take a random series with $n+1$ data and choose its two largest values. This latter pair can be placed with equiprobability in $n(n+1)$ positions, while only two of them are such that the largest values are placed at distance n, so we get $P(n) = \frac{2}{n(n+1)}$ on agreement with the previous development.

At this point, we can calculate the mean degree $< k >$ of the horizontal visibility graph:

$$< k >= \sum k P(k) = \sum_{k=2}^{\infty} \frac{k}{3} \left(\frac{2}{3}\right)^{k-2} = 4, \tag{29}$$

that we can recover from $P(n)$ as:

$$< k >= 2 \sum_{n=1}^{\infty} P(n) = 4. \tag{30}$$

Now, for illustration purposes, in figure 14 we show the adjacency matrix [7] of the horizontal visibility graph associated to a random series of 10^3 data (the entry i, j is filled in black if nodes i and j are connected, and left blank otherwise). Since every data x_i has

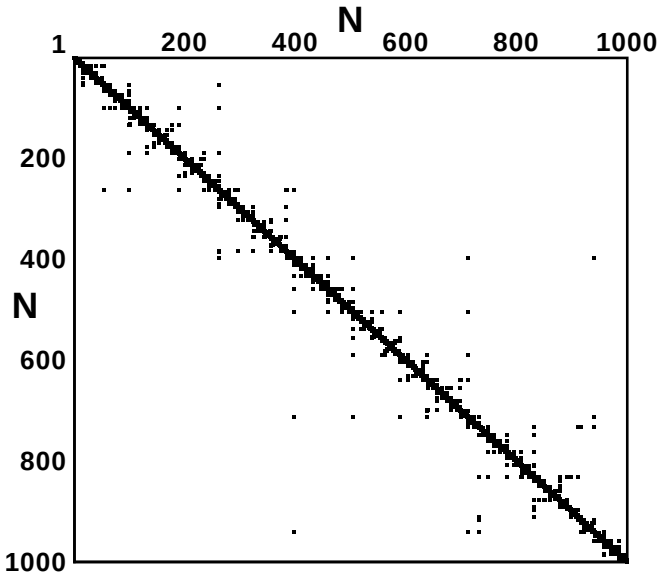

Figure 14. Adjacency matrix of an horizontal visibility graph associated to a random series of 10^3 data.

visibility of its first neighbors x_{i-1}, x_{i+1}, every node i will be connected by construction to nodes $i-1$ and $i+1$: the graph is thus connected. Observe in figure 14 that the graph evidences a typical homogeneous structure: the adjacency matrix is predominantly filled around the main diagonal. Furthermore, the matrix evidences a superposed sparse structure, reminiscent of the visibility probability $P(n) = 2/(n(n+1))$ that introduces some shortcuts in the horizontal visibility graph, much in the vein of the Small-World model [8]. Here the probability of having these shortcuts is given by $P(n)$. Statistically speaking, we can interpret the graph's structure as quasi-homogeneous, where the size of the local neighborhood increases with the graph's size. Accordingly, we can approximate its mean path length $L(N)$ as:

$$L(N) \approx \sum_{n=1}^{N-1} nP(n) = \sum_{n=1}^{N-1} \frac{2}{n+1} = 2\log(N) + 2(\gamma - 1) + O(1/N), \qquad (31)$$

where we have made use of the asymptotic expansion of the harmonic numbers and γ is the Euler-Mascheroni constant. As can be seen, the scaling is logarithmic, denoting that the horizontal visibility graph associated to a generic random series is Small-World [19], according to what figure 14 suggested. In figure 15 we have plotted the numerical results of $L(N)$ (dots) of an horizontal visibility graph associated to several random series of increasing size $N = 2^7, 2^8, \ldots 2^{17}$. The solid line corresponds to the best fit $L(N) = 1.3\log(N) - 1.7$.

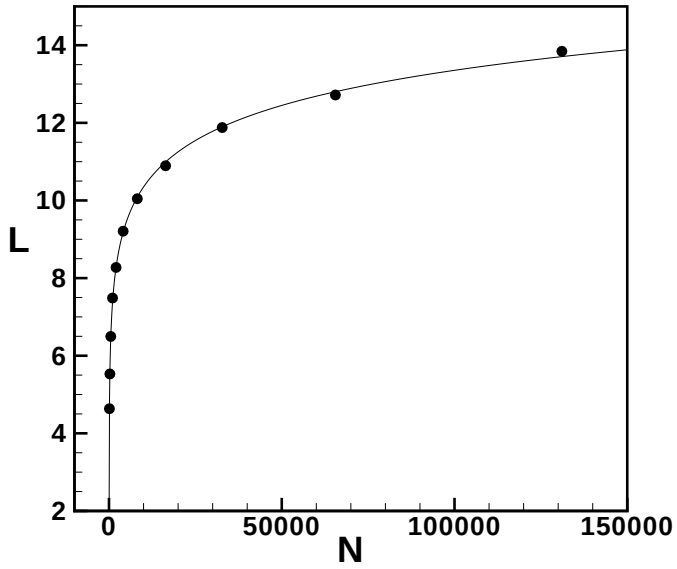

Figure 15. Mean path length $L(N)$ of an horizontal visibility graph associated to random series of $N = 2^7, 2^8, \ldots 2^{17}$ data (dots). The solid line corresponds to the better fit $L(N) = 1.3\log(N) - 1.7$.

6. Application of the theory to discriminate chaotic series

So far we have presented exact results on the topological properties of graphs associated to series of i.i.d. random variables (random series from now on) via the horizontal visibility algorithm. The very first application of this theory can be found in the task of discriminating a random signal from a chaotic one. The task of identifying random processes and more concretely discriminating (low dimensional) deterministic chaotic systems from stochastic processes has been extensively studied in the last decades (see for instance [48, 49, 50, 51, 52, 53, 54]). Essentially, all methods that have been introduced so far rely on two major points: Firstly, chaotic systems have a finite dimensional attractor, whereas stochastic processes arise from an infinitely dimensional one. Being able to reconstruct this latter attractor is thus a clear evidence showing that the time series has been generated by a deterministic system. Secondly, deterministic systems evidence, as opposed to random ones, short-time prediction: the difference between the time evolution of two nearby states will remain rather low for regular systems and increase exponentially fast for chaotic ones, while for stochastic processes this difference should be randomly distributed. Whereas several algorithms relying on the preceding concepts are nowadays available, the great majority of them are purely numerical and/or usually complicated to perform, computationally speaking (these difficulties are eventually more acute for noisy series [55] or high dimensional chaotic ones [56]). Furthermore, even the discrimination between a chaotic series and a series of i.i.d. random variables, something that an autocorrelation function or a power

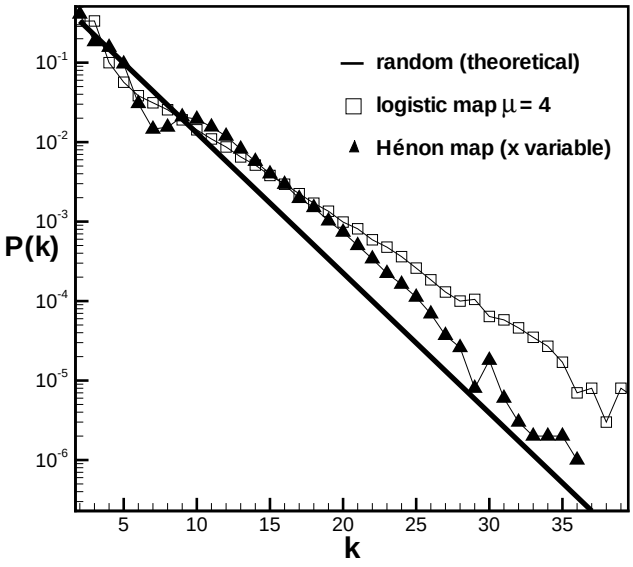

Figure 16. Semi-log plot of the degree distribution of several horizontal visibility graphs associated to: (solid line) theoretical prediction for random series (equation 20), (squares) time series of 10^6 extracted from a logistic map $x_{t+1} = \mu x_t(1 - x_t)$ in the fully chaotic region $\mu = 4$, (black triangles) time series of 10^6 extracted from x variable of the Hénon map $(x_{t+1}, y_{t+1}) = (y_t + 1 - ax_t^2, bx_t)$ in the fully chaotic region ($a = 1.4$, $b = 0.3$).

spectra fails to do but some other methods such as recurrence plots can [57] is nontrivial when the chaotic degree of the series is high, or even when such series is polluted with noise. All these complications provide motivation for a search for new methods that can directly distinguish, in a reliable way, random from chaotic time series, prior to quantifying the dimension [58] and without needs for additional sophisticated techniques such as surrogate data [59] or noise reduction methods [55]. In the preceding sections we have proved that the horizontal visibility graph associated to a random series has well-defined and universal degree distribution, local clustering distribution and $P(n)$, independent of the shape of the random probability distribution $f(x)$. These theorems guarantee that horizontal visibility graphs with other topological properties are not uncorrelated random series. In what follows we explore the reliability of the method to distinguish uncorrelated randomness from chaos in finite series.

6.1. Low-dimensional chaos

In order to test the practical usefulness of this method, we have generated the horizontal visibility graph of several noise-free chaotic series, and have calculated numerically their degree distribution. We have restricted our analysis to discrete systems (maps). In figure 16 we have plotted in semi-log the results of these simulations. In every case and by simple visual inspection we can conclude that $P(k)$ deviates from equation 20: the method is able

to easily distinguish randomness from low-dimensional chaos (similar results are obtained with $P(n)$ and $P(C)$).

Observe at this point that if we shuffle the series data and reproduce the analysis, we would find a degree distribution that now would correspond to equation 20, since shuffling breaks the temporal correlations of the series: such shuffled series would be equivalent to a random series extracted from a probability distribution equal to the system's probability measure (the Beta distribution in the case of the Logistic map). We can deduce that the algorithm captures temporal correlations of time series, and that $P(k)$ plays the role of an autocorrelation function, but with the additional ability of capturing nonlinear correlations. Observe also that this method neither works on the time nor on the frequency domain, since it only makes use of topological features.

6.2. Noisy chaotic series

It is well known that standard methods evidence problems when noise is present in chaotic signals, since even a small amount of noise can destroy the fractal structure of a chaotic attractor and mislead the calculation of chaos indicators such as the correlation dimension or the Lyapunov exponents [55]. In order to check the algorithm's robustness, we have introduced an amount of white noise (measurement noise) in a signal extracted from a fully chaotic Logistic map ($\mu = 4.0$). In figure 17 we plot the degree distribution of its associated visibility graph. Remarkably, the algorithm still discriminates noisy chaotic behavior from randomness even when the noise level reaches the 100% of the signal amplitude.

6.3. Topological properties of chaotic series

Observe in fig. 16 that the series extracted from the Logistic and Hénon maps seem to have an associated visibility graph with a degree distribution which has an exponential tail, yet different to eq. 20. This characteristic can be explained as follows: First, the tail of $P(k)$ is related to the hubs degree. Hubs correspond to the data series that have largest visibility. These are, according to eq. 22, extreme events in the series, whose degree is truncated by other extreme data (statistically speaking). Accordingly, the tail of $P(k)$ essentially reduces to calculate the probability distribution of recurrence times in the series. Within random series, notice that this distribution is straightforwardly exponential (recurrence times in a Poisson process are exponentially distributed [14]), consistent with eq.20. Within chaotic series, recurrence time statistics are related to the concept of Poincaré recurrence time (PRT) [60], which measures the time interval between two consecutive visits of a trajectory to a finite size region of the phase space. As a matter of fact, it has been shown that Poincaré recurrence times are exponentially distributed in several hyperbolic chaotic systems, including the Logistic and Hénon maps (see [61] and references therein). We conjecture that the functional form of $P(k)$ is closely related for chaotic series with their associated Poincaré recurrence time distribution (which deviate from the Poissonian

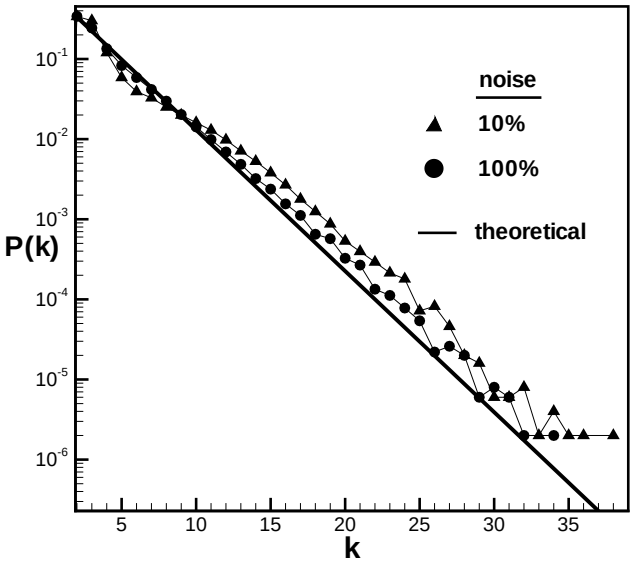

Figure 17. Semi-log plot of the degree distribution of an horizontal visibility graph associated to: (triangles) noisy chaotic series of 10^5 data extracted from the Logistic map ($\mu = 4$) with a measurement noise level of 10% (by amplitude), (circles) idem but for noise level of 100%. The solid line corresponds to the theoretical prediction for random series $P(k) = (1/3)(2/3)^{k-2}$.

statistics, eq. 20, due to deterministic effects), something that will be addressed in future work.

7. Conclusions

The family of visibility algorithms is a method to map information from time series to graphs. This method opens the possibility of building bridges between nonlinear dynamics, time series analysis and graph theory. Hitherto, research on this topic has verified that the method is well defined, in the sense that series correlations are inherited in the network. However this research is on its infancy, and important questions are yet to be studied. Indeed, it is time now to go beyond simple network indicators such as the degree distribution or the mean path length, and study whether if more complex topological features -perhaps mesoscale indicators- relate to different complex dynamics. This is a challenging research program that may have important applications in different branches of science, ranging from finance to biology, to cite some.

On the other hand, a well ground tool merits a mathematically sound background, and therefore a theory of visibility algorithms is also needed. Some of the first steps of this theory have been tackled in this chapter, but the core of such theory is still lacking. This

is another challenging research programme that entangles dynamical systems and graph theory in a single piece.

References

[1] J. Zhang, M. Small, *Phys. Rev. Lett.* **96** 238701 (2006).

[2] X. Xu, J. Zhang, M. Small, *Proc. Natl. Acad. Sci. USA* **105**, 50 (2008).

[3] L. Lacasa, B. Luque, F. Ballesteros, J. Luque and J.C. Nuño, *Proc. Natl. Acad. Sci. USA* **105**, 13 (2008), 4972-4975.

[4] Z. Gao and N. Jin, *Chaos* 19, 033137 (2009).

[5] A.H. Shirazi, G. Reza Jafari, J. Davoudi, J. Peinke, M. Reza Rahimi Tabar, M. Sahimi, *J. Stat. Mech.* (2009) P07046.

[6] M. de Berg, M. van Kreveld, M. Overmans and O. Schwarzkopf, in *Computational Geometry: Algorithms and Applications*, pp. 307-317 (Springer-Verlag, Berlin, 2000).

[7] A.L. Barabási, R. Albert, *Science* **286**, 509 (1999).

[8] R. Albert, A.L. Barabasi, *Rev. Mod. Phys.* 74 (2002).

[9] M.E.J. Newman, *SIAM Review* **45** (2003) 167-256.

[10] S. Dorogovtsev, J.F.F Mendes, *Advances in Physics* **51**, 4 (2002).

[11] S. Bocaletti, V. Latora, Y. Moreno, M. Chávez and D.U. Hwang, *Phys. Reports* **424** (2006) 175-308.

[12] L. Lacasa, B. Luque, J. Luque and J.C. Nuno, *EPL* **86** (2009) 30001.

[13] B. Bollobás, *Modern Graph Theory*, Springer-Verlag, New York Inc. (1998).

[14] W. Feller, *An Introduction to Probability Theory and its Applications*, John Willey and Sons, Inc. (1971).

[15] C. Song, S. Havlin, H.A. Makse, *Nature* **433**,392 (2005).

[16] K.I. Goh, G. Salvi, B. Kahng, D. Kim, *Phys. Rev. Lett.* **96** (2006).

[17] C. Song, S. Havlin, H.A. Makse, *Nat. Phys.* **2**,275 (2006).

[18] J.S.Kim, K.I. Goh, G. Salvi, E. Oh, B. Kahng, D. Kim, *Phys. Rev. E* **75**, 016110 (2007).

[19] D.J. Watts and S.H. Strogatz, *Nature* **393**, 440-442 (1998).

[20] J. Conway, *Some Crazy Sequences*, Lecture at AT&T Bell Labs (1988).

[21] B.B Mandelbrot and J.W Van Ness, *SIAM Review* **10**, 4 (1968) 422-437.

[22] F.A.B.F. de Moura and M.L. Lyra, *Phys. Rev. Lett.* **81**, 17 (1998).

[23] J.F. Muzy, E. Bacry and A. Arneodo, *Phys. Rev. Lett.* **67**, 25 (1991); K. Kiyani *et al.*, *ibid.* **98**, 2111101 (2007).

[24] M.P. Golombek et al., *Nature* **436**, 44-48 (2005).

[25] R.F. Voss, *Phys. Rev. Lett.* **68**, 25 (1992).

[26] P.Ch. Ivanov *et al.*, *Nature* **399**, 461-465 (1999).

[27] J. Haussdorf, *Human Movement Review* **26** (2007) 555-589.

[28] J.A.O. Matos *et al.*, *Physica A* **387**, 15 (2008) 3910-3915.

[29] T. Karagiannis, M. Molle and M. Faloutsos, *IEEE internet computing* **8**, 5 (2004) 57-64.

[30] W.E. Leland *et al.*, *IEEE/ACM Transactions on Networking* **2** (1994) 1-15.

[31] T. Mikosch *et al. The Annals of Applied Probability* , **12**, 1 (2002) 2368.

[32] R. Weron, *Physica A* **312** (2002) 285-299.

[33] B. Pilgram and D.T. Kaplan, *Physica D* **114** (1998) 108-112.

[34] J. W. Kantelhardt, Fractal and multifractal time series, in: *Springer encyclopaedia of complexity and system science* (in press, 2008) preprint arXiv:0804.0747.

[35] B. Podobnik, H.E. Stanley, *Phys. Rev. Lett.* **100**, 084102 (2008).

[36] A. Carbone, *Phys. Rev. E* **76**, 056703 (2007).

[37] J. Mielniczuk and P. Wojdyllo, *Comput. Statist. Data Anal.* **51** (2007) 4510-4525.

[38] I. Simonsen, A. Hansen and O.M. Nes, *Phys. Rev. E* **58**, 3 (1998).

[39] P. Abry and F. Sellan, *Appl. and Comp. Harmonic Anal.*, **3**, 4 (1996) 377-383.

[40] G.A. Horn *et al.Perfomance Evaluation* **64**, 2 (2007). 162-190.

[41] M.E.J. Newmann, *Contemporary Physics* **46**, 5 (2005), 323-351.

[42] C. Liu, W.X. Zhou, and W.K. Yuan, Arxiv preprint arXiv:0905.1831 (2009).

[43] J. B. Elsner, T. H. Jagger, and E. A. Fogarty, *Geophys. Res. Lett.* 36, 16 (2009).

[44] Y. Yang, J. Wang, H. Yang, J. Mang, *Physica A* 388, 20 (2009), Pages 4431-4437.

[45] M. Ding and W. Yang, *Phys. Rev. E*, **52**, 1 (1995).

[46] B. Luque, L. Lacasa, F. Ballesteros and J. Luque, *Phys. Rev. E* **80** (2009)

[47] E. Ravasz, A.L. Somera, D.A. Mongru, Z.N. Oltvai, A.-L. Barabasi, *Science* **297**, 1551 (2002).

[48] P. Grassberger and I. Procaccia, *Phys. Rev. Lett.* **50**, 448 (1983).

[49] J.D. Farmer and J.J. Sidorovich, *Phys. Rev. Lett.* **59** (1987) 845-848.

[50] G. Sugihara and R.M. May, *Nature* **344**, 734 (1990).

[51] A. A. Tsonis and J. B. Elsner, *Nature* **358**, 217 (1992).

[52] D.T. Kaplan and L. Glass, *Phys. Rev. Lett.* **68**, 4 (1992).

[53] O.A. Rosso, H.A. Larrondo, M.T. Martin, A. Plastino, and M.A. Fuentes, *Phys. Rev. Lett.* **99**, 154102 (2007).

[54] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge UNiversity press, Cambridge, 2003).

[55] E.J. Kostelich and T. Schreiber, *Phys. Rev. E* **48**, 1752 (1993).

[56] M. Bauer, H. Heng and W. Martienssen, *Phys. Rev. Lett.* **71**, 4 (1993).

[57] D. Kaplan and L. Glass, *Understanding Nonlinear Dynamics*, Springer-Verlag (1996).

[58] M.B. Kennel and S. Isabelle, *Phys. Rev. A* **46**, 6 (1992).

[59] D. Prichard and J. Theiler, *Phys. Rev. Lett.* **73**, 7 (1994).

[60] G.M. Zaslavsky, *Physics Reports* **371** (2002) pp: 461-580.

[61] M. Benedicks and L. Carleson, *Ann. Math.* **133** (1991) pp: 73169, L.-.S. Young, *Ann. Math.* **147**, 3 (1998) pp: 585-650, M. Hirata, B. Saussol, and S. Vaienti, *Comm. Math. Phys.* **206** (1999) pp: 33-55.