Visibility Algorithms: A Short Review

Angel M. Nuñez, Lucas Lacasa, Jose Patricio Gomez and Bartolo Luque
Universidad Politécnica de Madrid, Spain
Spain

1. Introduction

1.1 Motivation

Disregarding any underlying process (and therefore any physical, chemical, economical or whichever meaning of its mere numeric values), we can consider a time series just as an ordered set of values and play the naive mathematical game of turning this set into a different mathematical object with the aids of an abstract mapping, and see what happens: which properties of the original set are conserved, which are transformed and how, what can we say about one of the mathematical representations just by looking at the other... This exercise is of mathematical interest by itself. In addition, it turns out that time series or signals is a universal method of extracting information from dynamical systems in any field of science. Therefore, the preceding mathematical game gains some unexpected practical interest as it opens the possibility of analyzing a time series (i.e. the outcome of a dynamical process) from an alternative angle. Of course, the information stored in the original time series should be somehow conserved in the mapping. The motivation is completed when the new representation belongs to a relatively mature mathematical field, where information encoded in such a representation can be effectively disentangled and processed. This is, in a nutshell, a first motivation to map time series into networks.

This motivation is increased by two interconnected factors: first, although a mature field, time series analysis has some limitations, when it refers to study the so called complex signals. Beyond the linear regime, there exist a wide range of phenomena (not exclusive to physics) which are usually embraced in the field of the so called Complex Systems. Under this vague definition lies a common feature: the relevant effect of nonlinearities in their mathematical representation. This feature can be reflected in the temporal evolution of (at least one of) the variables describing the system and necessitates the use of specific tools for nonlinear analysis [1]. Dynamical phenomena such as chaos, long-range correlated stochastic processes, intermittency, multifractality, etc... are examples of complex phenomena where time series analysis is pushed to its own limits. Nonlinear time series analysis develops from techniques such as nonlinear correlation functions, embedding algorithms, multrifractal spectra, projection theorems... tools that increase in complexity parallel to the complexity of the process/series under study. New approaches, new paradigms to deal with complexity are not only welcome, but needed. Approaches that deal with the intrinsic nonlinearity

---

[1] We should note that nonlinearity is not the only feature that characterize a complex system; many interacting parts, randomness and emergence could also be cited but, as we are going to see later, nonlinearity will be sufficient for our purposes in this chapter

by being intrinsically nonlinear, that deal with the possible multiscale character of the underlying process by being designed to naturally incorporate multiple scales. And such is the framework of networks, of graph theory. Second, the technological era brings us the possibility of digitally analyze myriads of data in a glimpse. Massive data sets can nowadays be parsed, and with the aid of well suited algorithms, we can have access and filter data from many processes, let it be of physical, technological or even social garment. It is now time to develop new approaches to filter such plethora of information.

It is in this context that the network approach for time series analysis was born. The family of visibility algorithms constitute one of other possibilities to map a time series into a graph and subsequently analyze the structure of the series through the set of tools developed in the graph /complex network theory. In this chapter we will review some of its basic properties and show some of its first applications.

1.2 Different methods to map time series into graphs

The idea of mapping time series into graphs seems attractive because it lays a bridge between two prolific fields of modern science as Nonlinear Signal Analysis and Complex Networks Theory, so much so that it has attracted the attention of several research groups which have contributed to the topic with different strategies of mapping. While an exhaustive list of such strategies is beyond the scope of this work, we shall briefly outline some of them.

Zhang & Small (2006) developed a method that mapped each cycle of a pseudoperiodic time series into a node in a graph. The connection between nodes was established by a distance threshold in the reconstructed phase space when possible or by the linear correlation coefficient between cycles in the presence of noise. Noisy periodic time series mapped into random graphs while chaotic time series did it into scale-free, small-world networks due to the presence of unstable periodic orbits. This method was subsequently applied to characterize cardiac dynamics.

Xu et al. (2008) concentrated in the relative frequencies of appearance of four-node motifs inside a particular graph in order to classify it into a particular superfamily of networks which corresponded to specific underlying dynamics of the mapped time series. In this case, the method of mapping consisted in embedding the time series in an appropiated phase space where each point corresponded to a node in the network. A threshold was imposed not only in the minimum distance between two neighbours to be eligible (temporal separation should be greater than the mean period of the data) but also to the maximum number of neighbours a node could have. Different superfamilies were found for chaotic, hyperchaotic, random and noisy periodic underlying dynamics, unique fingerprints were also found for specific dynamical systems within a family.

Donner et al. (2010; 2011) presented a technique which was based on the properties of recurrence in the phase space of a dynamical system. More precisely, the recurrence matrix obtained by imposing a threshold in the minimum distance between two points in the phase space (as in Xu et al. (2008)) was interpreted as the adjacency matrix of an undirected, unweighted graph. Properties of such graphs at three different scales (local, intermediated and global) were presented and studied on several paradigmatic systems (Hénon map, Rossler system, Lorenz system, Bernoulli map). The variation of some of the properties of the graphs with the distance threshold was analyzed, the use of specific measures like the local clustering coefficient was proposed as a way for detecting dynamically invariant objects

(saddle points or unstable periodic orbits) and studying the graph properties dependent on the embedding dimension was suggested as a means to distinguish between chaotic and stochastic systems.

Campanharo et al. (2011) contributed with an idea along the lines of Shirazi et al. (2009), Strozzi et al. (2009) and Haraguchi et al. (2009) of a surjective mapping which admits an inverse opperation. This approach opens the reciprocal possibility of benefiting from time series analysis to study the structure and properties of networks. Time series are treated as Markov processes, values are grouped in quantiles which will correspond to nodes in the associated graph. Weighted and directed connections are stablished between nodes as a function of the probability of transition between quantiles. An inverse operation can be defined without any a priori knowledge of the correspondance between nodes and quantiles just by imposing a continuity condition in the time series by means of a cost function defined on the weighted adjacency matrix of the graph. A random walk is performed on the network and a time series with properties equivalent to the original one is recovered. This method was applied to a battery of cases which included a periodic-to-random family of processes parametrized by the probability of transition p, a pair of chaotic systems (Lorentz and Rossler attractors) and two human heart rate time series. Reciprocally, the inverse map was applied to the metabolic network of Arabidopsis Thaliana and to the '97 year Internet Network. Time series obtained were demostrated to exhibit different dynamics.

Among all these methods of mapping, in this chapter we are going to concentrate our attention on the one developed in Lacasa et al. (2008) and subsequent works. To cite some of its most relevant features, we will stress its intrinsic nonlocality, its low computational cost, its straightforward implementation and its quite 'simple' way of inherit the time series properties in the structure of the associated graphs. These features are going to make it easier to find connections between the underlying processes and the networks obtained from them by a direct analysis of the latter. In what follows we will firstly present different versions of the algorithm along with its most notable properties, that in many cases can be derived analytically (theorems are reported when possible). Based on these latter properties, several applications are addressed.

2. Visibility algorithms: Theory

2.1 Natural visibility algorithm: definition

Let $\{x(t_i)\}_{i=1..N}$ be a time series of N data. The natural visibility algorithm (Lacasa et al., 2008) assigns each datum of the series to a node in the natural visibility graph (from now on NVg). Two nodes i and j in the graph are connected if one can draw a straight line in the time series joining $x(t_i)$ and $x(t_j)$ that does not intersect any intermediate data height $x(t_k)$ (see figure 1 for a graphical illustration). Hence, i and j are two connected nodes if the following geometrical criterion is fulfilled within the time series:

$$x(t_k) < x(t_i) + (x(t_j) - x(t_i)) \frac{t_k - t_i}{t_j - t_k}. \tag{1}$$

It can easily checked that by means of the present algorithm, the associated graph extracted from a time series is always:

(*i*) connected: each node sees at least its nearest neighbors (left-hand side and right-hand side).
(*ii*) undirected: the way the algorithm is built up, there is no direction defined in the links.
(*iii*) 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.
(*iv*) "lossy": 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 with the same period as $T1 = ..., 3, 1, 3, 1, ...$ and $T2 = ..., 3, 2, 3, 2, ...$ would have the same visibility graph, albeit being quantitatively different.
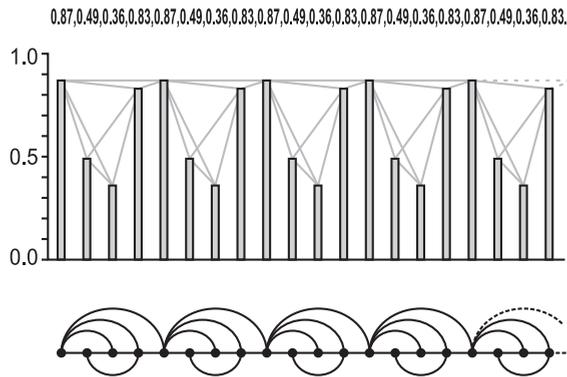

Fig. 1. Illustrative example of the visibility algorithm. In the upper part we plot a periodic time series and in the bottom part we represent the graph generated through the visibility algorithm. Each datum in the series corresponds to a node in the graph, such that two nodes are connected if their corresponding data heights fulfill the visibility criterion of equation 1. Note that the degree distribution of the visibility graph is composed by a finite number of peaks, much in the vein of the Discrete Fourier Transform of a periodic signal. We can thus interpret the visibility algorithm as a geometric transform.

One straightforward question is: what does the visibility algorithm stand for? 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 then defined on the 'visibility domain'. At this point we can mention 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 see that the visibility algorithm can distinguish between stochastic and chaotic series. Of course this analogy is, so far, a simple metaphor to help our intuition (this transform is not a reversible one for instance).

2.2 Horizontal visibility algorithm: definition

An alternative criterion for the construction of the visibility graph is defined as follows: let $\{x_i\}_{i=1..N}$ be a time series of N data. The so called horizontal visibility algorithm (Luque et al., 2009) assigns each datum of the series to a node in the horizontal visibility graph (from now on HVg). 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 2 for a graphical illustration). Hence, i and j are two connected nodes if the following geometrical criterion is fulfilled within the time series:

$$x_i, x_j > x_n \text{ for all } n \text{ such that } i < n < j \tag{2}$$

This algorithm is a simplification of the NVa. In fact, the HVg is always a subgraph of its associated NVg for the same time series (see figure 2). Beside this, the HVg graph will also be (*i*) connected, (*ii*) undirected, (*iii*) invariant under affine transformations of the series and (*iv*) "lossy". Some concrete properties of these graphs can be found in Gutin et al. (2011); Lacasa et al. (2010); Luque et al. (2009; 2011). In the next sections we are going to focus on properties of this particular method as it is a quite more analytically tractable version.

2.3 Topological properties of the HVg associated to periodic series: mean degree

Theorem 2.1. *The mean degree of an horizontal visibility graph associated to an infinite periodic series of period T (with no repeated values within a period) is*

$$\bar{k}(T) = 4\left(1 - \frac{1}{2T}\right) \tag{3}$$

A proof can be found in Núñez et al. (2010).

An interesting consequence of the previous result is that every time series extracted from a dynamical system has an associated HVG with a mean degree $2 \leq \bar{k} \leq 4$, where the lower bound is reached for constant series, whereas the upper bound is reached for aperiodic (random or chaotic) series (Luque et al., 2009).

2.4 Topological properties of the HVg associated to random time series

Let $\{x_i\}$ be a bi-infinite sequence of independent and identically distributed random variables extracted from a continous probability density $f(x)$, and consider its associated HVg. In the following sections we outline some theorems regarding the topological properties of these graphs.

2.4.1 Degree distribution of the visibility graph associated to a random time series

Theorem 2.2. *The degree distribution of its associated horizontal visibility graph is*

$$P(k) = \frac{1}{3}\left(\frac{2}{3}\right)^{k-2}, \, k = 2, 3, 4, ... \tag{4}$$

0.71, 0.53, 0.58, 0.29, 0.30, 0.77, 0.01, 0.76, 0.81, 0.71, 0.05, 0.41, 0.86, 0.79, 0.37, 0.96, 0.87, 0.06, 0.95, 0.36

0.71, 0.53, 0.58, 0.29, 0.30, 0.77, 0.01, 0.76, 0.81, 0.71, 0.05, 0.41, 0.86, 0.79, 0.37, 0.96, 0.87, 0.06, 0.95, 0.36
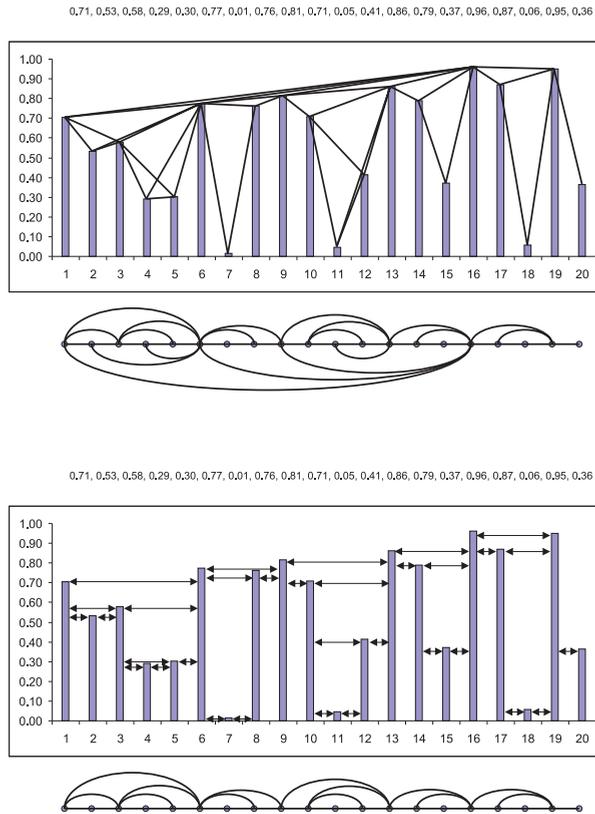

Fig. 2. Illustrative example of the natural and horizontal visibility algorithms. We plot the same time series and we represent the graphs generated through both visibility algorithms below. Each datum in the series corresponds to a node in the graph, such that two nodes are connected if their corresponding data heights fulfill respectively the visibility criteria of equations 1 and 2 respectively.

A lengthy constructive proof can be found in Luque et al. (2009) and alternative, shorter proofs can be found in Núñez et al. (2010).

Observe that the mean degree $\bar{k}$ of the horizontal visibility graph associated to an uncorrelated random process is then:

$$\bar{k} = \sum kP(k) = \sum_{k=2}^{\infty} \frac{k}{3} \left(\frac{2}{3} \right)^{k-2} = 4 \tag{5}$$

in good agreement with the prediction of eq. 3 in the limit $T \to \infty$, i.e. an aperiodic series.

2.4.2 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 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. $P(k|x)$ is easily deduced in Luque et al. (2009), resulting in

$$P(k|x) = \sum_{j=0}^{k-2} \frac{(-1)^{k-2}}{j!(k-2-j)!} [1 - F(x)]^2 \cdot [\ln(1 - F(x))]^{k-2} \tag{6}$$

The average value of the degree of a node associated to a datum of height x, $K(x)$, in then

$$K(x) = \sum_{k-2}^{\infty} kP(k|x) = 2 - 2\ln(1 - F(x)) \tag{7}$$

where $F(x) = \int_{-\infty}^{x} f(x')dx'$.

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.

2.4.3 Local clustering coefficient distribution

The local clustering coefficient C (Boccaletti et al., 2006; Newmann, 2003) of an horizontal visibility graph associated to a random series can be easy deduced by means of geometrical arguments (Luque et al., 2009):

$$C(k) = \frac{k-1}{\binom{k}{2}} = \frac{2}{k} \tag{8}$$

what indicates a so called hierarchical structure (Ravasz et al., 2002). This relation between k and C allows us to deduce the local clustering coefficient distribution $P(C)$:

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

$$P(C) = \frac{1}{3}\left(\frac{2}{3}\right)^{2/C-2} \tag{9}$$

2.4.4 Long distance visibility, mean degree and mean path length

The probability $P(n)$ that two data separated by n intermediate data be two connected nodes in the graph can be demostrated to be (see Luque et al. (2009))

$$P(n) = \left(\frac{1}{n} - 1\right) \int_0^1 f(x_0)F^n(x_0)dx_0 + \int_0^1 f(x_0)F^{n-1}(x_0)dx_0$$

$$= \frac{2}{n(n+1)} \tag{10}$$

where $P(n)$ is independent of the probability distribution f(x) of the random variable. Notice that the latter result can also be obtained, alternatively, with a purely combinatorial argument: 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.

2.4.5 Small World property

If we looked the adjacency matrix (Newmann, 2003) of the horizontal visibility graph associated to a random series (Luque et al., 2009), we would see that every data x_i has 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. 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 (Strogatz, 2001). 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:

$$L(N) \approx \sum_{n=1}^{N-1} nP(n) = \sum_{n=1}^{N-1} \frac{2}{n+1} = 2\log(N) + 2(\gamma - 1) + O(1/N) \tag{11}$$

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 (Newmann, 2003).

2.5 Topological properties of the HVg associated to other stochastic and chaotic processes

It was proved that $P(k) = (1/3)(2/3)^{k-2}$ for uncorrelated random series. To find out a similar closed expression in the case of generic chaotic or stochastic correlated processes is a very difficult task, since variables can be long-range correlated and hence the probabilities cannot be separated (lack of independence). This leads to a very involved calculation which is typically impossible to solve in the general case. However, some analytical developments can be made in order to compare them with our numerical results. Concretely, for Markovian systems global dependence is reduced to a one-step dependence. We will make use of such property to derive exact expressions for $P(2)$ and $P(3)$ in some Markovian systems (both deterministic and stochastic).

2.5.1 Ornstein-Uhlenbeck process: degree distribution

Suppose a short-range correlated series (exponentially decaying correlations) of infinite size generated through an Ornstein-Uhlenbeck process (Van Kampen, 2007), and generate its associated HVg. Let us consider the probability that a node chosen at random has degree $k = 2$. This node is associated to a datum labelled x_0 without lack of generality. Now, this node will have degree $k = 2$ if the datum first neighbors, x_1 and x_{-1} have values larger than x_0:

$$P(k = 2) = P(x_{-1} > x_0 \cap x_1 > x_0) \tag{12}$$

In this case the variables are correlated, so in general we should have

$$P(2) = \int_{-\infty}^{\infty} dx_0 \int_{x_0}^{\infty} dx_{-1} \int_{x_0}^{\infty} dx_1 \, f(x_{-1}, x_0, x_1) \tag{13}$$

We use the Markov property $f(x_{-1}, x_0, x_1) = f(x_{-1})f(x_0|x_{-1})f(x_1|x_0)$, that holds for an Ornstein-Uhlenbeck process with correlation function $C(t) \sim \exp(-t/\tau)$ (Van Kampen, 2007):

$$f(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}} \qquad f(x_2|x_1) = \frac{\exp(-(x_2 - Kx_1)^2/2(1-K^2))}{\sqrt{2\pi(1-K^2)}}, \qquad (14)$$

where $K = \exp(-1/\tau)$.

Numerical integration allows us to calculate $P(2)$ for every given value of the correlation time τ. A procedure to compute $P(3)$ can also be found in Lacasa et al. (2010).

2.5.2 Logistic map: degree distribution

A chaotic map of the form $x_{n+1} = F(x_n)$ does also have the Markov property, and therefore a similar analysis can be applied (even if chaotic maps are deterministic). For chaotic dynamical systems whose trajectories belong to the attractor, there exists a probability measure that characterizes the long-run proportion of time spent by the system in the various regions of the attractor. In the case of the logistic map $F(x_n) = \mu x_n(1 - x_n)$ with parameter $\mu = 4$, the attractor is the whole interval $[0, 1]$ and the probability measure $f(x)$ corresponds to the *beta* distribution with parameters $a = 0.5$ and $b = 0.5$:

$$f(x) = \frac{x^{-0.5}(1-x)^{-0.5}}{\mathrm{B}(0.5, 0.5)} \qquad (15)$$

Now, for a deterministic system, the transition probability is

$$f(x_{n+1}|x_n) = \delta[x_{n+1} - F(x_n)], \qquad (16)$$

where $\delta(x)$ is the Dirac delta distribution. Departing from equation 12, for the logistic map $F(x_n) = 4x_n(1 - x_n)$ and $x_n \in [0, 1]$, we have

$$P(2) = \int_0^1 dx_0 \int_{x_0}^1 f(x_{-1})f(x_0|x_{-1})dx_{-1} \int_{x_0}^1 f(x_1|x_0)dx_1 =$$
$$\int_0^1 dx_0 \int_{x_0}^1 f(x_{-1})\delta(x_0 - F(x_{-1}))dx_{-1} \int_{x_0}^1 \delta(x_1 - F(x_0))dx_1. \qquad (17)$$

Now, notice that, using the properties of the Dirac delta distribution, $\int_{x_0}^1 \delta(x_1 - F(x_0))dx_1$ is equal to one iff $F(x_0) \in [x_0, 1]$, what will happen iff $0 < x_0 < 3/4$, and zero otherwise. Therefore the only effect of this integral is to restrict the integration range of x_0 to be $[0, 3/4]$.

On the other hand,

$$\int_{x_0}^1 f(x_{-1})\delta[x_0 - F(x_{-1})]dx_{-1} = \sum_{x_k^*|F(x_k^*)=x_0} f(x_k^*)/|F'(x_k^*)|,$$

that is, the sum over the roots of the equation $F(x) = x_0$, iff $F(x_{-1}) > x_0$. But since $x_{-1} \in [x_0, 1]$ in the latter integral, it is easy to see that again, this is verified iff $0 < x_0 < 3/4$ (as a matter of fact, if $0 < x_0 < 3/4$ there is always a *single* value of $x_{-1} \in [x_0, 1]$ such that $F(x_{-1}) = x_0$, so the sum restricts to the adequate root). It is easy to see that the particular

value is $x^* = (1 + \sqrt{1 - x_0})/2$. Making use of these piecewise solutions and equation 15, we finally have

$$P(2) = \int_0^{3/4} \frac{f(x^*)}{4\sqrt{1 - x_0}} dx_0 = 1/3, \tag{18}$$

Note that a similar development can be fruitfully applied to other chaotic maps, provided that they have a well defined natural measure. Analytical and numerical developments for $P(3)$ can be found in Lacasa et al. (2010).

2.6 Directed horizontal visibility graph

So far, undirected visibility graphs have been considered, as visibility did not have a predefined temporal arrow. However, such a directionality can be made explicit by making use of directed networks or digraphs (Newmann, 2003).

Let a *directed* horizontal visibility graph (DHVg, Lacasa et al. (2011)) be a horizontal visibility graph, where the degree $k(x_i)$ of the node x_i is now splitted in an *ingoing* degree $k_{\text{in}}(x_i)$, and an *outgoing* degree $k_{\text{out}}(x_i)$, such that $k(x_i) = k_{\text{in}}(x_i) + k_{\text{out}}(x_i)$. The ingoing degree $k_{\text{in}}(x_i)$ is defined as the number of links of node x_i with other *past* nodes associated with data in the series (that is, nodes with $j < i$). Conversely, the outgoing degree $k_{\text{out}}(x_i)$, is defined as the number of links with *future* nodes ($i < j$).
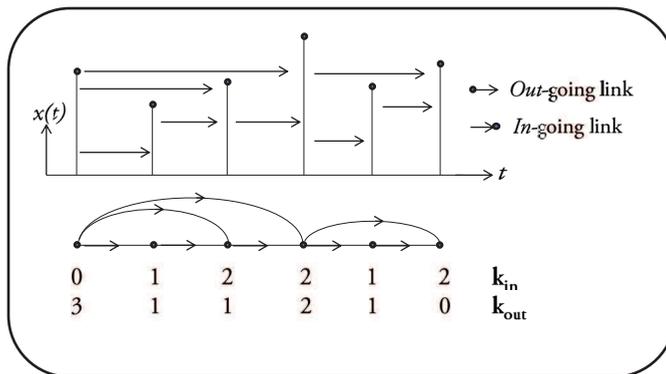

Fig. 3. Graphical illustration of the method. In the top we plot a sample time series $\{x(t)\}$. Each datum in the series is mapped to a node in the graph. Arrows, describing allowed directed visibility, link nodes. The associated directed horizontal visibility graph is plotted below. In this graph, each node has an ingoing degree k_{in}, which accounts for the number of links with *past* nodes, and an outgoing degree k_{out}, which in turn accounts for the number of links with *future* nodes. The asymmetry of the resulting graph can be captured in a first approximation through the invariance of the outgoing (or ingoing) degree series under time reversal.

For a graphical illustration of the method, see figure 3. The degree distribution of a graph describes the probability of an arbitrary node to have degree k (i.e. k links, Newmann (2003)). We define the *in* and *out* (or ingoing and outgoing) degree distributions of a DHVg as the

probability distributions of k_{out} and k_{in} of the graph which we call $P_{\text{out}}(k) \equiv P(k_{\text{out}} = k)$ and $P_{\text{in}}(k) \equiv P(k_{\text{in}} = k)$, respectively.

2.6.1 Uncorrelated stochastic series: degree distribution

Theorem 2.3. *Let $\{x_t\}_{t=-\infty,...,\infty}$ be a bi-infinite sequence of independent and identically distributed random variables extracted from a continuous probability density $f(x)$. Then, both the* in *and* out *degree distributions of its associated directed horizontal visibility graph are*

$$P_{\text{in}}(k) = P_{\text{out}}(k) = \left(\frac{1}{2}\right)^k, \ k = 1, 2, 3, ... \tag{19}$$

Proof. (out-distribution) Let x be an arbitrary datum of the aforementioned series. The probability that the horizontal visibility of x is interrupted by a datum x_r on its right is independent of $f(x)$,

$$\Phi_1 = \int_{-\infty}^{\infty} \int_x^{\infty} f(x)f(x_r)dx_rdx = \int_{-\infty}^{\infty} f(x)[1 - F(x)]dx = \frac{1}{2},$$

The probability $P(k)$ of the datum x being capable of exactly seeing k data may be expressed as

$$P(k) = Q(k)\Phi_1 = \frac{1}{2}Q(k), \tag{20}$$

where $Q(k)$ is the probability of x seeing at least k data. $Q(k)$ may be recurrently calculated via

$$Q(k) = Q(k-1)(1 - \Phi_1) = \frac{1}{2}Q(k-1), \tag{21}$$

from which, with $Q(1) = 1$, the following expression is obtained

$$Q(k) = \left(\frac{1}{2}\right)^{k-1}, \tag{22}$$

which together with equation (20) concludes the proof. $\square$

An analogous derivation holds for the *in* case. This result is independent of the underlying probability density $f(x)$: it holds not only for Gaussian or uniformly distributed random series, but for any series of independent and identically distributed (i.i.d.) random variables extracted from a continuous distribution $f(x)$.

3. Towards a graph theory of time series?

In the preceding section, specific properties of the visibility graphs (either NVg, HVg or the directed version of HVg) associated to different time series have been considered. Relying on the aforementioned dualities between time series structure and network topological features, we proceed here to make the first steps for a graph theoretical analysis of time series and dynamical systems, addressing several nontrivial problems of time series analysis through the visibility algorithm apparatus.

3.1 Estimating the Hurst exponent with NVg

Self-similar processes such as fractional Brownian motion (fBm, Mandelbrot & Van Ness (1968)) are currently used to model fractal phenomena of different nature, ranging from Physics or Biology to Economics or Engineering (see Lacasa et al. (2009) and references therein). A fBm $B_H(t)$ is a non-stationary random process with stationary self-similar increments (fractional Gaussian noise) that can be characterized by the so called Hurst exponent, $0 < H < 1$. 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 Carbone (2007); Kantelhardt (2008); Karagiannis et al. (2004); Mielniczuk & Wojdyllo (2007); Pilgram & Kaplan (1998); Podobnik & Stanley (2008); Simonsen et al. (1998); Weron (2002) and references therein.
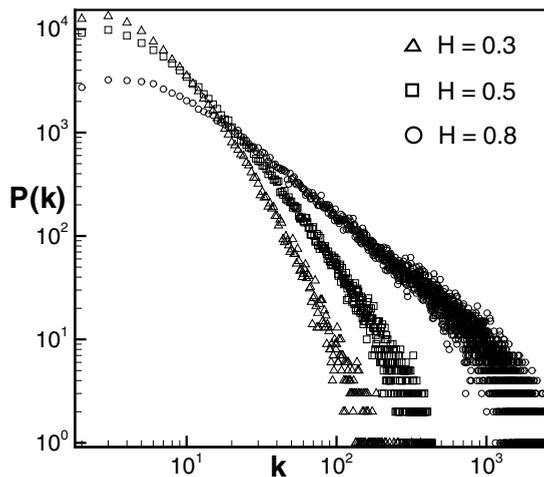

Fig. 4. Degree distribution of three visibility graphs, namely (i) triangles: extracted from a fBm series of 10^5 data with $H = 0.3$, (ii) squares: extracted from a fBm series of 10^5 data with $H = 0.5$, (iii) circles: extracted from a fBm series of 10^5 data with $H = 0.8$. Note that distributions 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 $\gamma_{0.3} > \gamma_{0.5} > \gamma_{0.8}$.

Here we address the problem of estimating the Hurst exponent of a fBm series via the NVg. If we map a fBm time series by means of the NVa, what we get is a scale-free graph (Lacasa et al.,

2008; 2009), see figure 4. As a matter of fact, 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 datum with a very large value has typically a large connectivity (a fact reminiscent of eq. 7). It can be proved (Lacasa et al., 2009) that the degree distribution of a NVg extracted from a fBm with Hurst exponent H shows a power law shape $P(k) \sim k^{-\gamma}$, such that

$$\gamma(H) = 3 - 2H. \tag{23}$$

Numerical analysis corroborated this theoretical relation in Lacasa et al. (2009).

It is well known that fBm has a power spectra that behaves as $1/f^{\beta}$, where the exponent β is related to the Hurst exponent of an fBm process through the well known relation

$$\beta(H) = 1 + 2H. \tag{24}$$

Now according to eqs. 23 and 24, the degree distribution of the visibility graph corresponding to a time series with $f^{-\beta}$ noise should be again power law $P(k) \sim k^{-\gamma}$ where

$$\gamma(\beta) = 4 - \beta. \tag{25}$$

The theoretical prediction eq. 25 was also corroborated numerically in Lacasa et al. (2009). Finally, eq. 24 holds for fBm processes, while for the increments of an fBm process, known as a fractional Gaussian noise (fGn), the relation between β and H turns to be

$$\beta(H) = -1 + 2H, \tag{26}$$

The relation between γ and H for a fGn (where fGn is a series composed by the increments of a fBm) can be deduced to be

$$\gamma(H) = 5 - 2H. \tag{27}$$

In order to illustrate this latter case, we address a realistic and striking dynamics where long range dependence has been recently described. Gait cycle (the stride interval in human walking rhythm) is a physiological signal that has been shown to display fractal dynamics and long range correlations in healthy young adults (Goldenberger et al., 2002; Hausdorff et al., 1996). In the upper part of fig. 5 we have plotted to series describing the fluctuations of walk rhythm of a young healthy person, for slow pace (bottom series of 3304 points) and fast pace (up series of 3595 points) respectively (data available in www.physionet.org/physiobank/database/umwdb/ (Goldberger et al., 2000)). In the bottom part we have represented the degree distribution of their visibility graphs. These ones are again power laws with exponents $\gamma = 3.03 \pm 0.05$ for fast pace and $\gamma = 3.19 \pm 0.05$ for slow pace (derived through MLE). According to eq. 25, the visibility algorithm predicts that gait dynamics evidence $f^{-\beta}$ behavior with $\beta = 1$ for fast pace, and $\beta = 0.8$ for slow pace, in perfect agreement with previous results based on a Detrended Fluctuation Analysis (Goldenberger et al., 2002; Hausdorff et al., 1996). These series record the fluctuations of walk rhythm (that is, the increments), so according to eq. 27, the Hurst exponent is $H = 1$ for fast pace and $H = 0.9$ for slow pace, that is to say, dynamics evidences long range dependence (persistence) (Goldenberger et al., 2002; Hausdorff et al., 1996).
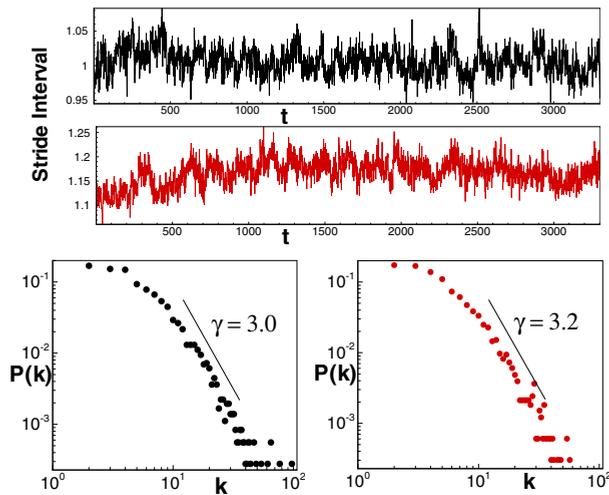
Fig. 5. Black signal: time series of 3595 points from the stride interval of a healthy person in fast pace. Red signal: time series of 3304 points from the stride interval of a healthy person in slow pace. Bottom: Degree distribution of the associated visibility graphs (the plot is in log-log). These are power laws where $\gamma = 3.03 \pm 0.05$ for the fast movement (black dots) and $\gamma = 3.19 \pm 0.05$ for the slow movement (red dots), what provides $\beta = 1$ and $\beta = 0.8$ for fast and slow pace respectively according to eq.25, in agreement with previous results (Goldenberger et al., 2002; Hausdorff et al., 1996).

3.2 Discriminating stochastic vs. chaotic series via HVg

Both stochastic and chaotic processes share many features, and the discrimination between them is indeed very subtle. The relevance of this problem is to determine whether the source of unpredictability (production of entropy) has its origin in a chaotic deterministic or stochastic dynamical system, a fundamental issue for modeling and forecasting purposes. Essentially, the majority of methods (Cecini et al., 2010; Kants H. & Schreiber, 2003) that have been introduced so far rely on two major differences between chaotic and stochastic dynamics. The first difference is that chaotic systems have a finite dimensional attractor, whereas stochastic processes arise from an infinite-dimensional one. Being able to reconstruct the attractor is thus a clear evidence showing that the time series has been generated by a deterministic system. The development of sophisticated embedding techniques (Kants H. & Schreiber, 2003) for attractor reconstruction is the most representative step forward in this direction. The second difference is that deterministic systems evidence, as opposed to random ones, short-time prediction: the time evolution of two nearby states will diverge exponentially fast for chaotic ones (finite and positive Lyapunov exponents) while in the case of a stochastic process such separation is randomly distributed. Whereas some algorithms relying on the preceding concepts are nowadays available, the great majority of them are purely phenomenological and often complicated to perform, computationally speaking. These drawbacks provide the motivation for a search for new methods that can directly distinguish, in a reliable way, stochastic from chaotic time series. We show here that the horizontal visibility algorithm offers a different, conceptually simple and computationally efficient method to distinguish between deterministic and stochastic dynamics, since the

degree distribution of HVGs associated to stochastic and chaotic processes are exponential $P(k) \sim \exp(-\lambda k)$, where for stochastic dynamics $\lambda > \lambda_{un}$ and for chaotic dynamics $\lambda < \lambda_{un}$ (Lacasa et al., 2010), λ_{un} being the uncorrelated case, (theorem 2.2).

3.2.1 Correlated stochastic series

In order to analyze the effect of correlations between the data of the series, we focus on two generic and paradigmatic correlated stochastic processes, namely long-range (power-law decaying correlations) and Ornstein-Uhlenbeck (short-range exponentially decaying correlations) processes. We have computed the degree distribution of the HVg associated to different long-range and short-range correlated stochastic series (the method for generating the associated series is explained in Lacasa et al. (2010)) with correlation function $C(t) = t^{-\gamma}$ for different values of the correlation strength $\gamma \in [10^{-2} - 10^{1}]$ and with an exponentially decaying correlation function $C(t) = \exp(-t/\tau)$. In both cases the degree distribution of the associated HVG can be fitted for large k by an exponential function $\exp(-\lambda k)$. The parameter λ depends on γ or τ and is, in each case, a monotonic function that reaches the asymptotic value $\lambda = \lambda_{un} = \ln(3/2)$ in the uncorrelated limit $\gamma \to \infty$ or $\tau \to 0$, respectively. Detailed results of this phenomenology can be found in (Lacasa et al., 2010). In all cases, the limit is reached from above, i.e. $\lambda > \lambda_{un}$ (see figure 6). Interestingly enough, for the power-law correlations the convergence is slow, and there is still a noticeable deviation from the uncorrelated case even for weak correlations ($\gamma > 4.0$), whereas the convergence with τ is faster in the case of exponential correlations.

3.2.2 Chaotic maps

Poincaré recurrence theorem suggests that the degree distribution of HVgs associated to chaotic series should be asymptotically exponential (Luque et al., 2009). Several deterministic time series generated by chaotic maps have been analyzed:

(1) the α-map $f(x) = 1 - |2x - 1|^{\alpha}$, that reduces to the logistic and tent maps in their fully chaotic region for $\alpha = 2$ and $\alpha = 1$ respectively, for different values of α,
(2) the 2D Hénon map ($x_{t+1} = y_t + 1 - ax_t^2$, $y_{t+1} = bx_t$) in the fully chaotic region ($a = 1.4$, $b = 0.3$);
(3) a time-delayed variant of the Hénon map: $x_{t+1} = bx_{t-d} + 1 - ax_t^2$ in the region ($a = 1.6$, $b = 0.1$), where it shows chaotic behavior with an attractor dimension that increases linearly with the delay d (Sprott, 2006). This model has also been used for chaos control purposes (Buchner & Zebrowski, 2000), although here we set the parameters a and b to values for which we find high-dimensional chaos for almost every initial condition (Sprott, 2006);
(4) the Lozi map, a piecewise-linear variant of the Hénon map given by $x_{t+1} = 1 + y_n - a|x_t|$, $y_{t+1} = bx_t$ in the chaotic regime $a = 1.7$ and $b = 0.5$;
(5) the Kaplan-Yorke map $x_{t+1} = 2x_t \mod (1)$, $y_{t+1} = \lambda y_t + \cos(4\pi x_t) \mod (1)$; and
(6) the Arnold cat map $x_{t+1} = x_t + y_t \mod (1)$, $y_{t+1} = x_t + 2y_t \mod (1)$, a conservative system with integer Kaplan-Yorke dimension. References for these maps can be found in Sprott & Rowlands (2001).

We find that the tails of the degree distribution can be well approximated by an exponential function $P(k) \sim \exp(-\lambda k)$. Remarkably, we find that $\lambda < \lambda_{un}$ in every case, where λ seems

to increase monotonically as a function of the chaos dimensionality [2], with an asymptotic value $\lambda \to \ln(3/2)$ for large values of the attractor dimension (see fig. 6 where we plot the specific values of λ as a function of the correlation dimension of the map (Sprott & Rowlands, 2001)). Again, we deduce that the degree distribution for uncorrelated series is a limiting case of the degree distribution for chaotic series but, as opposed to what we found for stochastic processes, the convergence flow towards λ_{un} is from below, and therefore $\lambda = \ln(3/2)$ plays the role of an effective frontier between correlated stochastic and chaotic processes.
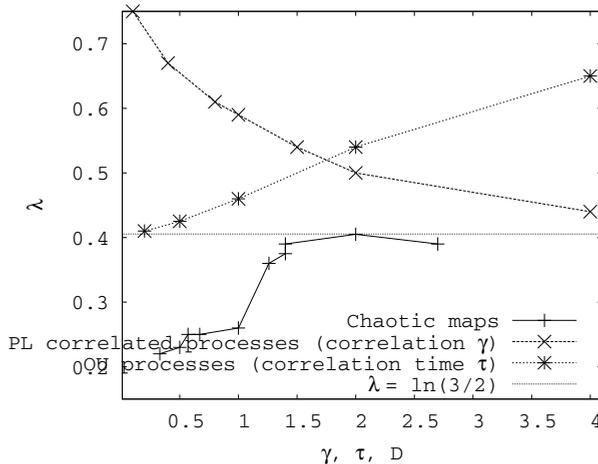

Fig. 6. Plot of the values of λ for several processes, namely: (i) for power-law correlated stochastic series with correlation function $C(t) = t^{-\gamma}$, as a function of the correlation γ, (ii) for Ornstein-Uhlenbeck series with correlation function $C(t) = \exp(-t/\tau)$, as a function of the correlation time τ, and (iii) for different chaotic maps, as a function of their correlation dimension D. Errors in the estimation of λ are incorporated in the size of the dots. Notice that stochastic processes cluster in the region $\lambda > \lambda_{un}$ whereas chaotic series belong to the opposite region $\lambda < \lambda_{un}$, evidencing a convergence towards the uncorrelated value $\lambda_{un} = \ln(3/2)$ (Luque et al., 2009) for decreasing correlations or increasing chaos dimensionality respectively.

In the following section we provide some heuristic arguments supporting our findings, for additional details, numerics and analytical developments we refer the reader to Lacasa et al. (2010).

3.2.3 Heuristics

We argue first that correlated series show lower data variability than uncorrelated ones, so decreasing the possibility of a node to reach far visibility and hence decreasing (statistically speaking) the probability of appearance of a large degree. Hence, the correlation tends to decrease the number of nodes with large degree as compared to the uncorrelated counterpart.

---

[2] This functional relation must nonetheless be taken in a cautious way, indeed, other chaos indicators (such as the Lyapunov spectra) may also play a relevant role in the final shape of $P(k)$ and such issues should be investigated in detail

Indeed, in the limit of infinitely large correlations ($\gamma \to 0$ or $\tau \to \infty$), the variability reduces to zero and the series become constant. The degree distribution in this limit case is, trivially,

$$P(k) = \delta(k-2) = \lim_{\lambda \to \infty} \frac{\lambda}{2} \exp(-\lambda|k-2|),$$

that is to say, infinitely large correlations would be associated to a diverging value of λ. This tendency is on agreement with the numerical simulations (figure 6) where we show that λ monotonically increases with decreasing values of γ or increasing values of τ respectively. Having in mind that in the limit of small correlations the theorem previously stated implies that $\lambda \to \lambda_{un} = \ln(3/2)$, we can therefore conclude that for a correlated stochastic process $\lambda_{stoch} > \lambda_{un}$.

Concerning chaotic series, remember that they are generated through a deterministic process whose orbit is continuous along the attractor. This continuity introduces a smoothing effect in the series that, statistically speaking, increases the probability of a given node to have a larger degree (uncorrelated series are rougher and hence it is more likely to have more nodes with smaller degree). Now, since in every case we have exponential degree distributions (this fact being related with the Poincaré recurrence theorem for chaotic series and with the return distribution in Poisson processes for stochastic series (Luque et al., 2009)), we conclude that the deviations must be encoded in the slope λ of the exponentials, such that $\lambda_{chaos} < \lambda_{un} < \lambda_{stoch}$, in good agreement with our numerical results.

3.3 Noise filtering using HVg: periodic series polluted with noise

In this section we address the task of filtering a noisy signal with a hidden periodic component within the horizontal visibility formalism, that is, we explore the possibility of using the method for noise filtering purposes (see (Núñez et al., 2010) for details). Periodicity detection algorithms (see for instance (Parthasarathy et al., 2006)) can be classified in essentially two categories, namely the time domain (autocorrelation based) and frequency domain (spectral) methods. Here we make use of the horizontal visibility algorithm to propose a third category: graph theoretical methods.

If we superpose a small amount of noise to a periodic series (a so-called *extrinsic* noise), while the degree of the nodes with associated small values will remain rather similar, the nodes associated to higher values will eventually increase their visibility and hence reach larger degrees. Accordingly, the delta-like structure of the degree distribution (associated with the periodic component of the series) will be perturbed, and an exponential tail will arise due to the presence of such noise (Lacasa et al., 2010; Luque et al., 2009). Can the algorithm characterize such kind of series? The answer is positive, since the degree distribution can be analytically calculated resulting in:

$$P(2) = 1/2,$$
$$P(3) = 0,$$
$$P(k+2) = \frac{1}{3}\left(\frac{2}{3}\right)^{k-2}, \; k \geq 2,$$
$$\text{or } P(k) = \frac{1}{4}\left(\frac{2}{3}\right)^{k-3}, \; k \geq 4, \tag{28}$$

that is to say, introducing a small amount of extrinsic uncorrelated noise in a periodic signal introduces an exponential tail in the HVG's degree distribution with the same slope as the one associated to a purely uncorrelated process. The mean degree $\bar{k}$ reads

$$\bar{k} = \sum_{k=2}^{\infty} kP(k) = 4,$$

which, according to equation 3, suggests aperiodicity, as expected.

3.3.1 A graph-theoretical noise filter

Let $S = \{x_i\}_{i=1,...,n}$ be a periodic series of period T (where $n >> T$) polluted by a certain amount of extrinsic noise (without loss of generality, suppose a white noise extracted from a uniform distribution $U[-0.5, 0.5]$), and define the filter f as a real valued scalar such that $f \in [\min x_i, \max x_i]$. The so called filtered Horizontal Visibility Graph (f-HVg) associated to S is constructed as it follows:
(i) each datum x_i in the time series is mapped to a node i in the f-HVg, (ii) two nodes i and j are connected in the f-HVg if the associated data fulfill

$$x_i, x_j > x_n + f, \, \forall \, n \mid i < n < j \,. \tag{29}$$

The procedure of filtering the noise from a noisy periodic signal goes as follows: one generates the f-HVg associated to S for increasing values of f, and in each case proceeds to calculate the mean degree $\bar{k}$. For the proper interval $f_{min} < f < f_{max}$, the f-HVg of the noisy periodic series S will be equivalent to the noise free HVG of the pure (periodic) signal, which has a well defined mean degree as a function of the series period. In this interval, the mean degree will therefore remain constant, and from equation 3 the period can be inferred. As an example, we
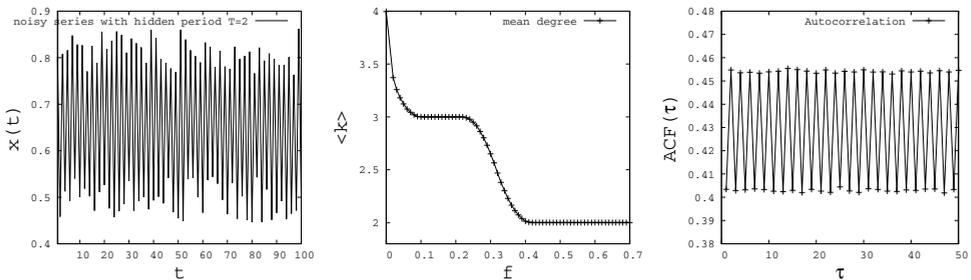

Fig. 7. *Left*: Periodic series of period $T = 2$ polluted with extrinsic noise extracted from a uniform distribution $U[-0.5, 0.5]$ of amplitude 0.1. *Middle*: Values of the HVg's mean degree $\bar{k}$ as a function of the amplitude of the graph theoretical filter. The first plateau is found for $\bar{k} = 3$, which renders a hidden period $T = (2 - \bar{k}/2)^{-1} = 2$. The second plateau corresponding to $\bar{k} = 2$ is found when the filter is large enough to screen each datum with its first neighbors, such that the mean degree reaches its lowest bound. *Right*: Autocorrelation function of the noisy periodic series, which is itself an almost periodic series with period $T = 2$, as it should.

have artificially generated a noisy periodic series of hidden period $T = 2$ (see figure 7). The results of the graph filtering technique yielded a net decreases of the mean degree, which has an initial value of 4 (as expected for the HVg ($f = 0$) of an aperiodic series such as a noisy

periodic signal) and an asymptotic value of 2 (lower bound of the mean degree). The plateau is clearly found at $\bar{k} = 3$, which according to equation 3 yields a period

$$T = \left(2 - \frac{\bar{k}}{2}\right)^{-1} = 2,$$

as expected.

3.3.2 Noisy periodic versus chaotic

Let us now consider a simple case of chaotic map with disconnected attractors. The Logistic map

$$x_{t+1} = \mu x_t (1 - x_t),$$

with $x \in [0, 1]$ and $\mu \in [3.6, 3.67]$ has an attractor that is partitioned in two disconnected chaotic bands, and the chaotic orbit makes an alternating journey between both bands (see fig. f̊igintro). The map is ergodic, but the attractor is not the whole interval, as there is a gap between both chaotic bands. In this situation, the chaotic series is by definition not periodic, however, an autocorrelation function analysis indeed suggests the presence of periodicity, what is reminiscent of the disconnected two-band structure of the attractor. Interestingly enough, applying the aforementioned noise filter technique, at odds with the autocorrelation function, the results suggests that the method does not find any periodic structure, as it should (see Núñez et al. (2010) for details). Furthermore, information of both the phase space structure and the chaotic nature of the map becomes accessible from an analysis of the HVg's degree distribution. First, we find $P(2) = 1/2$, that indicates that half of the data are located in the bottom chaotic band, in agreement with the alternating nature of the chaotic orbit. This is reminiscent of the misleading result obtained from the autocorrelation function. Second, the tail of the degree distribution is exponential, with an asymptotic slope smaller than the one obtained rigorously (Luque et al., 2009) for a purely uncorrelated process. This is, according to Lacasa et al. (2010), characteristic of an underlying chaotic process.

3.4 The period-doubling route to chaos via HVg: Feigenbaum graphs

In low-dimensional dissipative systems chaotic motion develops out of regular motion in a small number of ways or routes, and amongst which the period-doubling bifurcation cascade or Feigenbaum scenario is perhaps the better known and most famous mechanism (Peitgen et al., 1992; Schuster, 1988). This route to chaos appears an infinite number of times amongst the family of attractors spawned by unimodal maps within the so-called periodic windows that interrupt stretches of chaotic attractors. In the opposite direction, a route out of chaos accompanies each period-doubling cascade by a chaotic band-splitting cascade, and their shared bifurcation accumulation points form transitions between order and chaos that are known to possess universal properties (Peitgen et al., 1992; Schuster, 1988; Strogatz, 1994). Low-dimensional maps have been extensively studied from a purely theoretical perspective, but systems with many degrees of freedom used to study diverse problems in physics, biology, chemistry, engineering, and social science, are known to display low-dimensional dynamics (Marvel et al., 2009).

In this section, we offer a distinct view of the Feigenbaum scenario through the specific HVg formalism, and provide a complete set of graphs, which we call Feigenbaum graphs, that encode the dynamics of all stationary trajectories of unimodal maps. We first characterize their

topology via the order-of-visit and self-affinity properties of the maps. We will additionally define a renormalization group (RG) procedure that leads, via its flows, to or from network fixed-points to a comprehensive view of the entire family of attractors. Furthermore, the optimization of the entropy obtained from the degree distribution coincides with the RG fixed points and reproduces the essential features of the map's Lyapunov exponent independently of its sign. A general observation is that the HV algorithm extracts only universal elements of the dynamics, free of the peculiarities of the individual unimodal map, but also of universality classes characterized by the degree of nonlinearity. Therefore all the results presented in this section, while referring to the specific Logistic map for illustrative reasons apply to any unimodal map.

3.4.1 Feigenbaum graphs

According to the HV algorithm, a time series generated by the Logistic map for a specific value of μ (after an initial transient of approach to the attractor) is converted into a Feigenbaum graph (Luque et al., 2011). Notice that this is a well-defined subclass of HV graphs where consecutive nodes of degree $k = 2$, that is, consecutive data with the same value, do not appear, what is actually the case for series extracted from maps (besides the trivial case of a constant series). While for a period T there are in principle several possible periodic orbits, and therefore the set of associated Feigenbaum graphs is degenerate, it can be proved that the mean degree $\bar{k}(T)$ and normalized mean distance $\bar{d}(T)$ of all these Feigenbaum graphs fulfill $\bar{k}(T) = 4(1 - \frac{1}{2T})$ and $\bar{d}(T) = \frac{1}{3T}$ respectively, yielding a linear relation $\bar{d}(\bar{k}) = (4 - \bar{k})/6$ that is corroborated in the inset of figure 8. Aperiodic series ($T \to \infty$) reach the upper bound mean degree $\bar{k} = 4$.

3.4.2 Period-doubling cascade

A deep-seated feature of the period-doubling cascade is that the order in which the positions of a periodic attractor are visited is universal (Schroeder, 1991), the same for all unimodal maps. This ordering turns out to be a decisive property in the derivation of the structure of the Feigenbaum graphs. A plot the graphs for a family of attractors of increasing period $T = 2^n$, that is, for increasing values of $\mu < \mu_\infty$ can be found in (Luque et al., 2011). This basic pattern also leads to the expression for their associated degree distributions,

$$P(n,k) = \left(\frac{1}{2}\right)^{k/2}, \quad k = 2, 4, 6, ..., 2n, \tag{30}$$

$$P(n,k) = \left(\frac{1}{2}\right)^{n}, \quad k = 2(n+1),$$

and zero for k odd or $k > 2(n+1)$. At the accumulation point μ_∞ the period diverges ($n \to \infty$) and the distribution is exponential for all even values of the degree,

$$P(\infty,k) = \left(\frac{1}{2}\right)^{k/2}, \quad k = 2, 4, 6, ..., \tag{31}$$

and zero for k odd. Observe that these relations are independent of the order of the map's nonlinearity: the HV algorithm sifts out every detail of the dynamics except for the basic storyline.
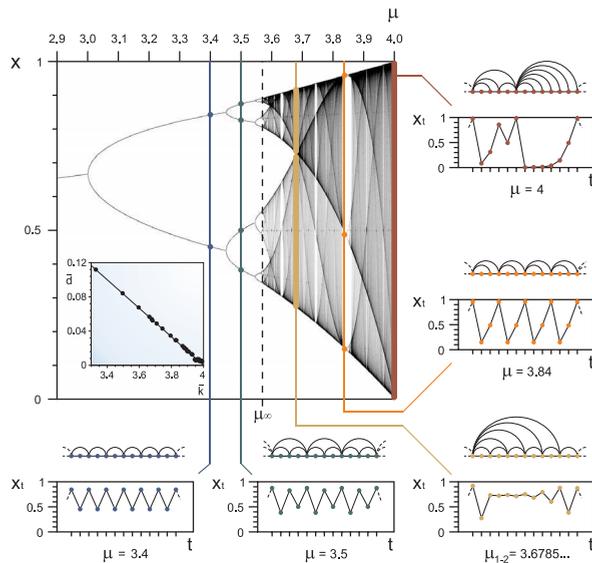
Fig. 8. Feigenbaum graphs from the Logistic map $x_{t+1} = f(x_t) = \mu x_t(1 - x_t)$. The main figure portrays the family of attractors of the Logistic map and indicates a transition from periodic to chaotic behavior at $\mu_\infty = 3.569946...$ through period-doubling bifurcations. For $\mu \geq \mu_\infty$ the figure shows merging of chaotic-band attractors where aperiodic behavior appears interrupted by windows that, when entered from their left-hand side, display periodic motion of period $T = m \cdot 2^0$ with $m > 1$ (for $\mu < \mu_\infty$, $m = 1$) that subsequently develops into m period-doubling cascades with new accumulation points $\mu_\infty(m)$. Each accumulation point $\mu_\infty(m)$ is in turn the limit of a chaotic-band reverse bifurcation cascade with m initial chaotic bands, reminiscent of the self-affine structure of the entire diagram. All unimodal maps exhibit a period-doubling route to chaos with universal asymptotic scaling ratios between successive bifurcations that depend only on the order of the nonlinearity of the map, the Logistic map belongs to the quadratic case. Adjoining the main figure, we show time series and their associated Feigenbaum graphs according to the HV mapping criterion for several values of μ where the map evidences both regular and chaotic behavior (see the text). Inset: numerical values of the mean normalized distance $\bar{d}$ as a function of mean degree $\bar{k}$ of the Feigenbaum graphs for $3 < \mu < 4$ (associated to time series of 1500 data after a transient and a step $\delta\mu = 0.05$), in good agreement with the theoretical linear relation (see the text).

3.4.3 Period-doubling bifurcation cascade of chaotic bands

We turn next to the period-doubling bifurcation cascade of chaotic bands that takes place as μ decreases from $\mu = 4$ towards μ_∞. For the largest value of the control parameter, at $\mu = 4$, the attractor is fully chaotic and occupies the entire interval $[0, 1]$ (see figure 8). This is the first chaotic band $n = 0$ at its maximum amplitude. As μ decreases in value within $\mu_\infty < \mu < 4$ band-narrowing and successive band-splittings (Peitgen et al., 1992; Schroeder, 1991; Schuster, 1988; Strogatz, 1994) occur. In general, after n reverse bifurcations the phase space is partitioned in 2^n disconnected chaotic bands, which are self-affine copies of the first

chaotic band (Crutchfield et al., 1982). The values of μ at which the bands split are called Misiurewicz points (Schroeder, 1991), and their location converges to the accumulation point μ_∞ for $n \to \infty$. Significantly, while in the chaotic zone orbits are aperiodic, for reasons of continuity they visit each of the 2^n chaotic bands in the same order as positions are visited in the attractors of period $T = 2^n$ (Schroeder, 1991). A plot of the Feigenbaum graphs generated through chaotic time series at different values of μ that correspond to an increasing number of reverse bifurcations can be found in(Luque et al., 2011). Since chaotic bands do not overlap, one can derive the following degree distribution for a Feigenbaum graph after n chaotic-band reverse bifurcations by using only the universal order of visits

$$P_\mu(n,k) = \left(\tfrac{1}{2}\right)^{k/2}, \quad k = 2, 4, 6, ..., 2n,$$

$$P_\mu(n, k \geq 2(n+1)) = \left(\tfrac{1}{2}\right)^n, \tag{32}$$

and zero for $k = 3, 5, 7, ..., 2n + 1$. We note that this time the degree distribution retains some dependence on the specific value of μ, concretely, for those nodes with degree $k \geq 2(n+1)$, all of which belong to the top chaotic band. The HV algorithm filters out chaotic motion within all bands except for that taking place in the top band whose contribution decreases as $n \to \infty$ and appears coarse-grained in the cumulative distribution $P_\mu(n, k \geq 2(n+1))$. As would be expected, at the accumulation point μ_∞ we recover the exponential degree distribution (equation 31), *i.e.* $\lim_{n\to\infty} P_\mu(n,k) = P(\infty,k)$.

3.4.4 Renormalization group

Before proceeding to interpret these findings via the consideration of renormalization group (RG) arguments, we recall that the Feigenbaum tree shows a rich self-affine structure: for $\mu > \mu_\infty$ periodic windows of initial period m undergo successive period-doubling bifurcations with new accumulation points $\mu_\infty(m)$ that appear interwoven with chaotic attractors. These cascades are self-affine copies of the fundamental one. The process of reverse bifurcations also evidences this self-affine structure, such that each accumulation point is the limit of a chaotic-band reverse bifurcation cascade. Accordingly, we label $G(m,n)$ the Feigenbaum graph associated with a periodic series of period $T = m \cdot 2^n$, that is, a graph obtained from an attractor within window of initial period m after n period-doubling bifurcations. In the same fashion, $G_\mu(n,m)$ is associated with a chaotic attractor composed by $m \cdot 2^n$ bands (that is, after n chaotic band reverse bifurcations of m initial chaotic bands). Graphs corresponding to $G(1,n)$ and $G_\mu(1,n)$ respectively can be found in (Luque et al., 2011). For the first accumulation point $G(1,\infty) = G_\mu(1,\infty) \equiv G_\infty$. Similarly, in each accumulation point $\mu_\infty(m)$, the identity $G(m,\infty) = G_\mu(m,\infty)$ is fulfilled.

In order to recast previous findings in the context of the renormalization group, let us define an RG operation $\mathcal{R}$ on a graph as the coarse-graining of every couple of adjacent nodes where one of them has degree $k = 2$ into a block node that inherits the links of the previous two nodes. This is a real-space RG transformation on the Feigenbaum graph (Newmann & Watts, 1999), dissimilar from recently suggested box-covering complex network renormalization schemes (Radicchi et al., 2008; Song et al., 2005; 2006). This scheme turns out to be equivalent for $\mu < \mu_\infty$ to the construction of an HV graph from the composed map $f^{(2)}$ instead of the original f, in correspondence to the original Feigenbaum renormalization procedure (Strogatz, 1994). We first note that $\mathcal{R}\{G(1,n)\} = G(1, n-1)$, thus, an iteration of this process yields an RG

flow that converges to the (1st) trivial fixed point $\mathcal{R}^{(n)}\{G(1,n)\} = G(1,0) \equiv G_0 = \mathcal{R}\{G_0\}$. This is the stable fixed point of the RG flow $\forall \mu < \mu_\infty$. We note that there is only one relevant variable in our RG scheme, represented by the reduced control parameter $\Delta\mu = \mu_\infty - \mu$, hence, to identify a nontrivial fixed point we set $\Delta\mu = 0$ or equivalently $n \to \infty$, where the structure of the Feigenbaum graph turns to be completely self-similar under $\mathcal{R}$. Therefore we conclude that $G(1,\infty) \equiv G_\infty$ is the nontrivial fixed point of the RG flow, $\mathcal{R}\{G_\infty\} = G_\infty$. In connection with this, let $P_t(k)$ be the degree distribution of a generic Feigenbaum graph G_t in the period-doubling cascade after t iterations of $\mathcal{R}$, and point out that the RG operation, $\mathcal{R}\{G_t\} = G_{t+1}$, implies a recurrence relation $(1 - P_t(2))P_{t+1}(k) = P_t(k+2)$, whose fixed point coincides with the degree distribution found in equation 31. This confirms that the nontrivial fixed point of the flow is indeed G_∞.

Next, under the same RG transformation, the self-affine structure of the family of attractors yields $\mathcal{R}\{G_\mu(1,n)\} = G_\mu(1,n-1)$, generating a RG flow that converges to the Feigenbaum graph associated to the 1st chaotic band, $\mathcal{R}^{(n)}\{G_\mu(1,n)\} = G_\mu(1,0)$. Repeated application of $\mathcal{R}$ breaks temporal correlations in the series, and the RG flow leads to a 2nd trivial fixed point $\mathcal{R}^{(\infty)}\{G_\mu(1,0)\} = G_{rand} = \mathcal{R}\{G_{rand}\}$, where G_{rand} is the HV graph generated by a purely uncorrelated random process. This graph has a universal degree distribution $P(k) = (1/3)(2/3)^{k-2}$, independent of the random process underlying probability density (see (Lacasa et al., 2010; Luque et al., 2009)).
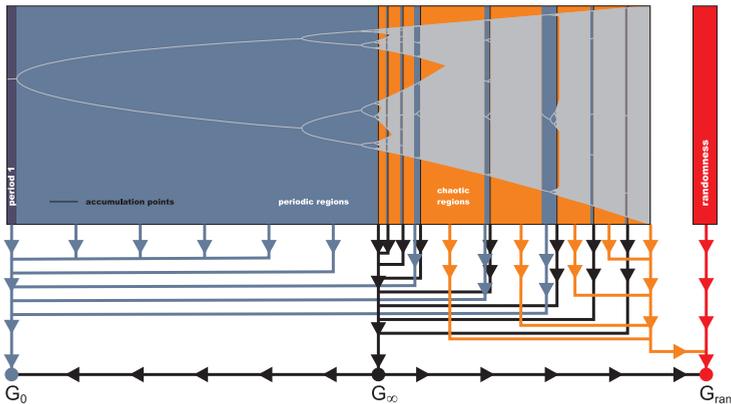

Fig. 9. Illustrative cartoon incorporating the RG flow of Feigenbaum graphs in the whole Feigenbaum diagram: aperiodic (chaotic or random) series generate graphs whose RG flow converge to the trivial fixed point G_{rand}, whereas periodic series (both in the region $\mu < \mu_\infty$ and inside windows of stability) generate graphs whose RG flow converges to the trivial fixed point $G(0,1)$. The nontrivial fixed point of the RG flow $G(\infty,1)$ is only reached through the critical manifold of graphs at the accumulation points $\mu_\infty(m)$.

Finally, let us consider the RG flow inside a given periodic window of initial period m. As the renormalization process addresses nodes with degree $k = 2$, the initial applications of $\mathcal{R}$ only change the core structure of the graph associated with the specific value m. The RG flow will therefore converge to the 1st trivial fixed point via the initial path $\mathcal{R}^{(p)}\{G(m,n)\} = G(1,n)$, with $p \leq m$, whereas it converges to the 2nd trivial fixed point for $G_\mu(m,n)$ via $\mathcal{R}^{(p)}\{G_\mu(m,n)\} = G_\mu(1,n)$. In the limit of $n \to \infty$ the RG flow proceeds towards

the nontrivial fixed point via the path $\mathcal{R}^{(p)}\{G(m, \infty)\} = G(1, \infty)$. Incidentally, extending the definition of the reduced control parameter to $\Delta\mu(m) = \mu_\infty(m) - \mu$, the family of accumulation points is found at $\Delta\mu(m) = 0$. A complete schematic representation of the RG flows can be seen in figure 9.

Interestingly, and at odds with standard RG applications to (asymptotically) scale-invariant systems, we find that invariance at $\Delta\mu = 0$ is associated in this instance to an exponential (rather than power-law) function of the observables, concretely, that for the degree distribution. The reason is straightforward: $\mathcal{R}$ is not a conformal transformation (*i.e.* a scale operation) as in the typical RG, but rather, a translation procedure. The associated invariant functions are therefore non homogeneous (with the property $g(ax) = bg(x)$), but exponential (with the property $g(x + a) = cg(x)$).

3.4.5 Network entropy

Finally, we derive, via optimization of an entropic functional for the Feigenbaum graphs, all the RG flow directions and fixed points directly from the information contained in the degree distribution. Amongst the graph theoretical entropies that have been proposed we employ here the Shannon entropy of the degree distribution $P(k)$, that is $h = -\sum_{k=2}^{\infty} P(k) \log P(k)$. By making use of the Maximum Entropy formalism, it is easy to prove that the degree distribution $P(k)$ that maximizes h is exactly $P(k) = (1/3)(2/3)^{k-2}$, which corresponds to the distribution for the 2nd trivial fixed point of the RG flow G_{rand}. Alternatively, with the incorporation of the additional constraint that allows only even values for the degree (the topological restriction for Feigenbaum graphs $G(1, n)$), entropy maximization yields a degree distribution that coincides with equation 31, which corresponds to the nontrivial fixed point of the RG flow G_∞. Lastly, the degree distribution that minimizes h trivially corresponds to G_0, *i.e.* the 1st trivial fixed point of the RG flow. Remarkably, these results indicate that the fixed-point structure of the RG flow are obtained via optimization of the entropy for the entire family of networks, supporting a suggested connection between RG theory and the principle of Maximum Entropy (Robledo, 1999).

The network entropy h can be calculated exactly for $G(1, n)$ ($\mu < \mu_\infty$ or $T = 2^n$), yielding $h(n) = \log 4 \cdot (1 - 2^{-n})$. Because increments of entropy are only due to the occurrence of bifurcations h increases with μ in a step-wise way, and reaches asymptotically the value $h(\infty) = \log 4$ at the accumulation point μ_∞. For Feigenbaum graphs $G_\mu(1, n)$ (in the chaotic region), in general h cannot be derived exactly since the precise shape of $P(k)$ is unknown (albeit the asymptotic shape is also exponential (Luque et al., 2011)). Yet, the main feature of h can be determined along the chaotic-band splitting process, as each reverse bifurcation generates two self-affine copies of each chaotic band. Accordingly, the decrease of entropy associated with this reverse bifurcation process can be described as $h_\mu(n) = \log 4 + h_\mu(0)/2^n$, where the entropy $h_\mu(n)$ after n reverse bifurcations can be described in terms of the entropy associated with the first chaotic band $h_\mu(0)$. The chaotic-band reverse bifurcation process takes place in the chaotic region in the direction of decreasing μ's, and therefore leads in this case to a decrease of entropy with an asymptotic value of $\log 4$ for $n \to \infty$ at the accumulation point. These results suggest that the graph entropy behaves qualitatively as the map's Lyapunov exponent λ, with the peculiarity of having a shift of $\log 4$, as confirmed in figure 10. This unexpected qualitative agreement is reasonable in the chaotic region in view of the Pesin theorem (Peitgen et al., 1992), that relates the positive Lyapunov exponents of a map with its Kolmogorov-Sinai entropy (akin to a topological entropy) that for unimodal
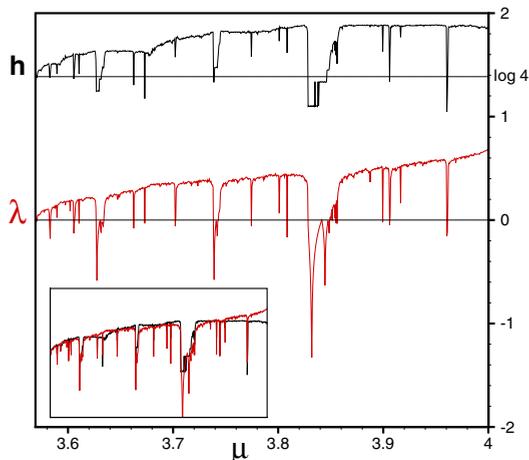
Fig. 10. Horizontal visibility network entropy h and Lyapunov exponent λ for the Logistic map. We plot the numerical values of h and λ for $3.5 < \mu < 4$ (the numerical step is $\delta\mu = 5 \cdot 10^{-4}$ and in each case the processed time series have a size of 2^{12} data). The inset reproduces the same data but with a rescaled entropy $h - \log(4)$. The surprisingly good match between both quantities is reminiscent of the Pesin identity (see text). Unexpectedly, the Lyapunov exponent within the periodic windows ($\lambda < 0$ inside the chaotic region) is also well captured by h.

maps reads $h_{KS} = \lambda, \ \forall \lambda > 0$, since h can be understood as a proxy for h_{KS}. Unexpectedly, this qualitative agreement seems also valid in the periodic windows ($\lambda < 0$), since the graph entropy is positive and varies with the value of the associated (negative) Lyapunov exponent even though $h_{KS} = 0$, hinting at a Pesin-like relation valid also out of chaos which deserves further investigation. The agreement between both quantities lead us to conclude that the Feigenbaum graphs capture not only the period-doubling route to chaos in a universal way, but also inherits the main feature of chaos, *i.e.* sensitivity to initial conditions.

3.5 Measuring irreversibility via HVg

A stationary process x_t is said to be statistically time reversible (hereafter time reversible) if for every n, the series $\{x_1, \cdots, x_n\}$ and $\{x_n, \cdots, x_1\}$ have the same joint probability distributions (Weiss, 1975). Roughly, this means that a reversible time series and its time reversed are, statistically speaking, equally probable. Reversible processes include the family of Gaussian linear processes (as well as Fourier-transform surrogates and nonlinear static transformations of them), and are associated with processes at thermal equilibrium in statistical physics. Conversely, time series irreversibility is indicative of the presence of nonlinearities in the underlying dynamics, including non-Gaussian stochastic processes and dissipative chaos, and are associated with systems driven out-of-equilibrium in the realm of thermodynamics (Kawai et al., 2007; Parrondo et al., 2009). Time series irreversibility is an important topic in basic and applied science. From a physical perspective, and based on the relation between

statistical reversibility and physical dissipation (Kawai et al., 2007; Parrondo et al., 2009), the concept of time series irreversibility has been used to derive information about the entropy production of the physical mechanism generating the series, even if one ignores any detail of such mechanism (Roldan & Parrondo, 2011). In a more applied context, it has been suggested that irreversibility in complex physiological series decreases with aging or pathology, being maximal in young and healthy subjects (Costa et al., 2005; 2008; Yang et al., 2003), rendering this feature important for noninvasive diagnosis. As complex signals pervade natural and social sciences, the topic of time series reversibility is indeed relevant for scientists aiming to understand and model the dynamics behind complex signals.

The definition of time series reversibility is formal and therefore there is not an *a priori* optimal algorithm to quantify it in practice. Several methods to measure time irreversibility have been proposed (Andrieux et al., 2007; Cammarota & Rogora, 2007; Costa et al., 2005; Daw et al., 2000; Diks et al., 1995; Gaspard, 2004; Kennel, 2004; Wang et al., 2005; Yang et al., 2003). The majority of them perform a time series symbolization, typically making an empirical partition of the data range (Daw et al., 2000) (note that such a transformation does not alter the reversible character of the output series (Kennel, 2004)) and subsequently analyze the symbolized series, through statistical comparison of symbol strings occurrence in the forward and backwards series or using compression algorithms (Cover & Thomas, 2006; Kennel, 2004; Roldan & Parrondo, 2011). The first step requires an extra amount of *ad hoc* information (such as range partitioning or size of the symbol alphabet) and therefore the output of these methods eventually depend on these extra parameters. A second issue is that since typical symbolization is local, the presence of multiple scales (a signature of complex signals) could be swept away by this coarse-graining: in this sense multi-scale algorithms have been proposed recently (Costa et al., 2005; 2008). The *time directed* version of the horizontal visibility algorithm is proposed in this section as a simple and well defined tool for measuring time series irreversibility (see Lacasa et al. (2011) for details).

3.5.1 Quantifying irreversibility: DHVg and Kullback-Leibler divergence

The main conjecture of this application is that the information stored in the *in* and *out* distributions take into account the amount of time irreversibility of the associated series. More precisely, we claim that this can be measured, in a first approximation, as the distance (in a distributional sense) between the *in* and *out* degree distributions ($P_{in}(k)$ and $P_{out}(k)$). If needed, higher order measures can be used, such as the corresponding distance between the *in* and *out* degree-degree distributions ($P_{in}(k, k')$ and $P_{out}(k, k')$). These are defined as the *in* and *out* joint degree distributions of a node and its first neighbors (Newmann, 2003), describing the probability of an arbitrary node whose neighbor has degree k' to have degree k.

The Kullback-Leibler divergence (Cover & Thomas, 2006) is used as the distance between the *in* and *out* degree distributions. Relative entropy or Kullback-Leibler divergence (KLD) is introduced in information theory as a measure of distinguishability between two probability distributions. Given a random variable x and two probability distributions $p(x)$ and $q(x)$, KLD between p and q is defined as follows:

$$D(p||q) \equiv \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}, \tag{33}$$

which vanishes if and only if both probability distributions are equal $p = q$ and it is bigger than zero otherwise.

We compare the outgoing degree distribution in the actual (forward) series $P_{k_{\text{out}}}(k|\{x(t)\}_{t=1,\dots,N}) = P_{\text{out}}(k)$ with the corresponding probability in the time-reversed (or backward) time series, which is equal to the probability distribution of the ingoing degree in the actual process $P_{k_{\text{out}}}(k|\{x(t)\}_{t=N,\dots,1}) = P_{\text{in}}(k)$. The KLD between these two distributions is

$$D[P_{\text{out}}(k)||P_{\text{in}}(k)] = \sum_k P_{\text{out}}(k) \log \frac{P_{\text{out}}(k)}{P_{\text{in}}(k)}. \tag{34}$$

This measure vanishes if and only if the outgoing and ingoing degree probability distributions of a time series are identical, $P_{\text{out}}(k) = P_{\text{in}}(k)$, and it is positive otherwise. We will apply it to several examples as a measure of irreversibility.

Notice that previous methods to estimate time series irreversibility generally proceed by first making a (somewhat *ad hoc*) local symbolization of the series, coarse-graining each of the series data into a symbol (typically, an integer) from an ordered set. Then, they subsequently perform a statistical analysis of word occurrences (where a word of length n is simply a concatenation of n symbols) from the forward and backwards symbolized series (Andrieux et al., 2007; Wang et al., 2005). Time series irreversibility is therefore linked to the difference between the word statistics of the forward and backwards symbolized series. The method presented here can also be considered as a symbolization if we restrict ourselves to the information stored in the series $\{k_{\text{out}}(t)\}_{t=1,\dots,N}$ and $\{k_{\text{in}}(t)\}_{t=1,\dots,N}$. However, at odds with other methods, here the symbolization process (i) lacks *ad hoc* parameters (such as number of symbols in the set or partition definition), and (ii) it takes into account *global* information: each coarse-graining $x_t \to (k_{\text{in}}(t), k_{\text{out}}(t))$ is performed using information from the whole series, according to the mapping criterion of fig. 3. Hence, this symbolization naturally takes into account multiple scales, which is desirable if we want to tackle complex signals (Costa et al., 2005; 2008).

3.5.2 Results for correlated stochastic series

The first example of a reversible series with $D[P_{\text{out}}(k)||P_{\text{in}}(k)] = 0$ are uncorrelated stochastic series which were considered in 2.6.1. As a further validation, linearly correlated stochastic processes have also been considered as additional examples of reversible dynamics (Weiss, 1975). An explanation of the method employed to generate the series can be consulted in (Lacasa et al., 2010), and results are summarized in table 1.

3.5.3 Results for a discrete flashing ratchet

A discrete flashing ratchet is an example of thermodynamic system which can be smoothly driven out of equilibrium by modifying the value of a physical parameter (the peak value V of an asymmetric potencial). We make use of a time series generated by a discrete flashing ratchet model introduced in (Roldan & Parrondo, 2010). For $V = 0$ detailed balance condition is satisfied, the system is in equilibrium and trajectories are statistically reversible. In this case both $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ and $D[P_{\text{out}}(k,k')||P_{\text{in}}(k,k')]$ using degree distributions and degree-degree distributions vanish. On the other hand, if V is increased, the system is driven out of equilibrium, what introduces a net statistical irreversibility which increases with V (Roldan & Parrondo, 2010). The amount of irreversibility estimated with KLD increases with

V for both measures, therefore the results produced by the method are qualitatively correct (see (Lacasa et al., 2011) for details). Interestingly enough, the tendency holds even for high values of the potential, where the statistics are poor and the KLD of sequences of symbols usually fail when estimating irreversibility (Roldan & Parrondo, 2010). However the values of the KLD obtained are far below the KLD per step between the forward and backward trajectories, which is equal to the dissipation as reported in (Roldan & Parrondo, 2010).

The degree distributions capture the irreversibility of the original series but it is difficult to establish a quantitative relationship between eq. (34) and the KLD between trajectories. The
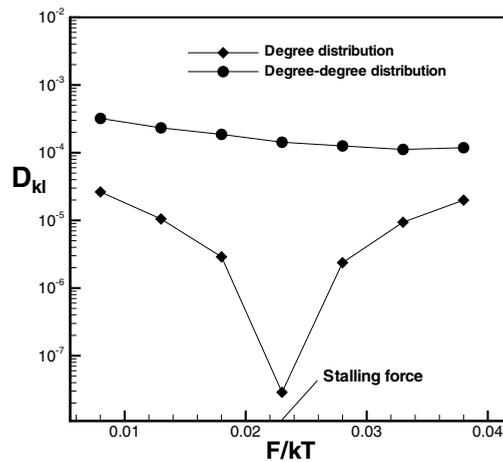
Fig. 11. Irreversibility measures $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ and $D[P_{\text{out}}(k,k')||P_{\text{in}}(k,k')]$ in the flashing ratchet ($r = 2, V = 2kT$) as a function of FL/kT. Here, F is the applied force and L is the spatial period of the ratchet, which in this case is equal to 1. For each value of the force, we make use of a single stationary series of size $N = 10^6$ containing partial information (the state information is removed).

measure based on the degree-degree distribution $D[P_{\text{out}}(k,k')||P_{\text{in}}(k,k')]$ takes into account more information of the visibility graph structure than the KLD using degree distributions, providing a closer bound to the physical dissipation as it is expected by the chain rule (Cover & Thomas, 2006), $D[P_{\text{out}}(k,k')||P_{\text{in}}(k,k')] \geq D[P_{\text{out}}(k)||P_{\text{in}}(k)]$. The improvement is even qualitatively significant in some situations. For instance, when a force opposite to the net current on the system is present (Roldan & Parrondo, 2010), the current vanishes for a given value of the force usually termed as *stalling force*. When the force reaches this value, the system is still out of equilibrium ($V > 0$) and it is therefore time irreversible, but no current of particles is observed if we describe the dynamics of the ratchet with partial information given by the position x. $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ tends to zero when the force approaches to the stalling value (see figure 11). Therefore, our measure of irreversibility (34) fails in this case, as do other KLD estimators based on local flows or currents (Roldan & Parrondo, 2010). However, $D[P_{\text{out}}(k,k')||P_{\text{in}}(k,k')]$ captures the irreversibility of the time series, and yields a positive value at the stalling force(Roldan & Parrondo, 2011).

Series description	$D[P_{\text{out}}(k)\|P_{\text{in}}(k)]$	$D[P_{\text{out}}(k,k')\|P_{\text{in}}(k,k')]$
Reversible Stochastic Processes		
$U[0,1]$ uncorrelated	$3.88 \cdot 10^{-6}$	$2.85 \cdot 10^{-4}$
Ornstein-Uhlenbeck ($\tau = 1.0$)	$7.82 \cdot 10^{-6}$	$1.52 \cdot 10^{-4}$
Long-range correlated		
stationary process ($\gamma = 2.0$)	$1.28 \cdot 10^{-5}$	$2.0 \cdot 10^{-4}$
Dissipative Chaos		
Logistic map ($\mu = 4$)	0.377	2.978
α map ($\alpha = 3$)	0.455	3.005
α map ($\alpha = 4$)	0.522	3.518
Henon map ($a = 1.4, b = 0.3$)	0.178	1.707
Lozi map	0.114	1.265
Kaplan Yorke map	0.164	0.390
Conservative Chaos		
Arnold Cat map	$1.77 \cdot 10^{-5}$	$4.05 \cdot 10^{-4}$

Table 1. Values of the irreversibility measure associated to the degree distribution $D[P_{\text{out}}(k)\|P_{\text{in}}(k)]$ and the degree-degree distribution $D[P_{\text{out}}(k,k')\|P_{\text{in}}(k,k')]$ respectively, for the visibility graphs associated to series of 10^6 data generated from reversible and irreversible processes. In every case chain rule is satisfied, since $D[P_{\text{out}}(k,k')\|P_{\text{in}}(k,k')] \geq D[P_{\text{out}}(k)\|P_{\text{in}}(k)]$. Note that that the method correctly distinguishes between reversible and irreversible processes, as KLD vanishes for the former and it is positive for the latter.

3.5.4 Results for chaotic series

This method was applied to several chaotic series and found that it is able to distinguish between dissipative and conservative chaotic systems. Dissipative chaotic systems are those that do not preserve the volume of the phase space, and they produce irreversible time series. This is the case of chaotic maps in which entropy production via instabilities in the forward time direction is quantitatively different to the amount of past information lost. In other words, those whose positive Lyapunov exponents, which characterize chaos in the forward process, differ in magnitude with negative ones, which characterize chaos in the backward process (Kennel, 2004). Several chaotic maps have been analyzed and the degree of reversibility of their associated time series has been estimated using using KLD, showing that for dissipative chaotic series it is positive while it vanishes for an example of conservative chaos. A summary of results cann be checked in table 1. In every case, we find an asymptotic positive value, in agreement with the conjecture that dissipative chaos is indeed time irreversible.

Finally, we also consider the *Arnold cat map*: $x_{t+1} = x_t + y_t \mod(1), y_{t+1} = x_t + 2y_t \mod(1)$. At odds with previous dissipative maps, this is an example of a *conservative* (measure-preserving) chaotic system with integer Kaplan-Yorke dimension (Sprott & Rowlands, 2001). The map has two Lyapunov exponents which coincide in magnitude $\lambda_1 = \ln(3 + \sqrt{5})/2 = 0.9624$ and $\lambda_2 = \ln(3 - \sqrt{5})/2 = -0.9624$. This implies that the amount of information created in the forward process (λ_1) is equal to the amount of

information created in the backwards process ($-\lambda_2$), therefore the process is time reversible. $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ for a time series of this map asymptotically tends to zero with series size, and the same happens with the degree-degree distributions (see table 1). This correctly suggests that albeit chaotic, the map is statistically time reversible.

3.5.5 Robustness: Irreversible chaotic series polluted with noise

Standard time series analysis methods evidence problems when noise is present in chaotic series. 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 (Kostelich & Schreiber, 1993). In order to check if our method is robust, we add an amount of white noise (measurement noise) to a signal extracted from a fully chaotic Logistic map ($\mu = 4.0$). The results for the KLD of the signal polluted with noise is significantly greater than zero, as it exceeds the one associated to the noise in four orders of magnitude, even when the noise reaches the 100% of the signal amplitude (Lacasa et al., 2011). Therefore our method correctly predicts that the signal is irreversible even when adding a large amount of noise.

4. Summary, perspectives and open problems

In this chapter a review on the state of the art of visibility algorithms as a method to make time series analysis through network theory has been presented. We have reported the properties of natural and horizontal visibility algorithms, and have explored their ability in several problems such as the estimation of Hurst exponent in self-similar (fractal series), the discrimination between uncorrelated, correlated stochastic and chaotic processes, the problem of noise filtering, the problem of determining the amount of irreversibility (i.e. entropy production) of a system, or the generic study of nonlinear systems as they undergo a period-doubling route to chaos.

Before commenting on the plethora of applications and challenging open problems to be faced, a few words on how to be cautious and make good science should be stated. The simplicity and straightforwardness of a method can be tricky, since they could convey the wrong impression to directly produce results when applied to concrete problems. From a physical point of view, the practical interest of this method lies in its ability to reveal properties of the system under study, i.e. to reveal hidden structures in a given series. But this capacity is intimately linked to the strength and extent of the theory behind the method. That is why, before venturing to study complex systems in nature, a method should provide a sufficient theoretical support. In the case under study, it should be clearly stated what information and which properties we are mapping into what and how, before attempting to measure all kind of features in a visibility graph.

According to this, the first general open problem lies just there: to generate a mathematically sound, rigorous theory that explains and shows how time series/dynamical systems properties are mapped into the associated visibility graph. In this review we have outlined the first steps in this direction, but a broad and general theory is still to be completely developed. This theory should deal with questions such as (i) what concrete information are the algorithms mapping? and (ii) how they do so? Once we know this, we can understand what network features are behind multifractality, spatio-temporal chaos, intermittency, quasi-periodicity, and many other complex dynamical processes.

Only when these questions have been rigorously responded, this tool could be ready to be unambiguously used by practitioners, since visibility algorithms will be a new and universal

method to extract information from complex signals. Moreover, the possibility of defining mesoscale measures, which are typically network-based (for instance, modularity, community structure, etc), could be of interest to analyze non-local / multiscale dynamics. The potentials of the method could then apply to study long standing problems in Physics and Society, such as turbulence, stock market dynamics, or physiological signals such as electro-encephalogram, electro-cardiograms, and so on. On this respect, the initial naive approaches in those directions (turbulence (Liu et al., 2009), financial series (Liu et al., 2009; Yang et al., 2009), cardiac series (Shao, 2010)) are nowadays inconclusive because the theory behind the method is not fully developed. Eventually.

5. References

D. Andrieux, P. Gaspard, S. Ciliberto, N. Garnier, S. Joubaud, and A. Petrosyan, Entropy production and time asymmetry in nonequilibrium fluctuations, *Phys. Rev. Lett.* 98, 150601 (2007).

Boccaletti S, Latora V, Moreno Y, Chavez M, & Hwang DU (2006) Complex networks: Structure and dynamics. Phys. Rep. 424, 175.

Buchner T. and Zebrowski J.J., Logistic map with a delayed feedback: Stability of a discrete time-delay control of chaos. *Phys. Rev. E* 63, 016210 (2000).

C. Cammarota and E. Rogora, Time reversal, symbolic series and irreversibility of human heartbeat *Chaos, Solitons and Fractals* 32 (2007) 1649-1654.

Campanharo, Andriana S. L. O., Sirer, M. Irmak, Malmgren, R. Dean, Ramos, Fernando M. & Amaral, Luís A. Nunes (2011). Duality between Time Series and Networks. *PLoS ONE*, Vol. 6, No. 8, e23378.

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

Cencini M., Cecconi F., and Vulpiani A., *Chaos: From Simple Models to Complex Systems*, World Scientific (2010).

M. Costa, A.L. Goldberger, and C.-K. Peng, Broken Asymmetry of the Human Heartbeat: Loss of Time Irreversibility in Aging and Disease, *Phys. Rev. Lett.* 95, 198102 (2005).

M.D. Costa, C.K. Peng and A.L. Goldberger, Multiscale Analysis of Heart Rate Dynamics: Entropy and Time Irreversibility Measures, *Cardiovasc Eng* 8, (2008)

T.M. Cover and J.A. Thomas, *Elements of Information Theory* (Wiley, New Jersey, 2006).

Crutchfield JP, Farmer JD & Huberman BA (1982) Fluctuations and simple chaotic dynamics. Phys. Rep. 92, 2.

C.S. Daw, C.E.A. Finney, and M.B. Kennel, Symbolic approach for measuring temporal 'irreversibility', *Phys. Rev. E* 62, 2 (2000).

C. Diks, J.C. van Houwelingen, F. Takens, and J. DeGoede, Reversibility as a criterion for discriminating time series, *Phys. Lett. A* 201 (1995), 221-228.

Donner, R. V., Zou, Y., Donges, J. F., Marwan, N. & Juergen Kurths (2010). Recurrence networks - A novel paradigm for nonlinear time series analysis. *New Journal of Physics*, 12, 033025.

Donner, R. V., Small, M., Donges, J. F., Marwan, N., Zou, Y., Xiang, R. & Juergen Kurths (2010). Recurrence-based time series analysis by means of complex network methods. *International Journal of Bifurcation and Chaos*, Vol. 21, No. 4, 1019–1046.

P. Gaspard, Time-reversed dynamical entropy and irreversibility in markovian random processes, *J. Stat. Phys.* 117 (2004).

A.L. Goldberger *et al.*, *Circulation* 101, 23 (2000) 215-220.

A.R. Goldenberger *et al.*, *Proc. Natl. Acad. Sci. USA* 99, 1 (2002), 2466-2472.

Gutin, G., Mansour, T. & Severini, S. (2011). A characterization of horizontal visibility graphs and combinatorics on words. *PHYSICA A*, Vol. 390 (12), 2421–2428.

Haraguchi, Y., Shimada, Y., Ikeguchi, T. & Aihara, K. (2009). Transformation from complex networks to time series using classical multidimensional scaling, *Proceedings of the 19th International Conference on Artificial Neural Networks*, Heidelberg, Berlin, ICANN 2009, Springer-Verlag.

J.M. Hausdorff *et al.*, *J. App. Physiol.* 80 (1996) 1448-1457.

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

Kants H. and Schreiber T. *Nonlinear Time Series Analysis*, (Camdrige University Press, 2003).

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

R. Kawai, JMR Parrondo and C Van den Broeck, Dissipation: the phase-space perspective, *Phys. Rev. Lett.* 98, 080602 (2007).

MB Kennel, Testing time symmetry in time series using data compression techniques, *Phys. Rev. E* 69, 056208 (2004).

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

Lacasa L., Luque B., Ballesteros F., Luque J. & Nuño J.C. (2008). From time series to complex networks: the visibility graph. *Proc. Natl. Acad. Sci. USA* 105, 13, 4972-4975.

Lacasa L., Luque B., Nuño J.C. & Luque J. (2009). The Visibility Graph: a new method for estimating the Hurst exponent of fractional Brownian motion. *EPL* 86, 30001 (2009).

Lacasa, L. & Toral, R. (2010). Description of stochastic and chaotic series using visibility graphs. *Phys. Rev. E* 82, 036120 (2010).

Lacasa, L., Núñez, A. M., Roldan, E., Parrondo, J.M.R., & Luque, B. (2011).Time series irreversibility: a visibility graph approach. *arXiv:1108.1691* 2011.

Liu, C., Zhou, W-T. & Yuan, W-K. (2009). Statistical properties of visibility graph of energy dissipation rates in three-dimensional fully developed turbulence. *Physics Letters A*, Vol. 389 (13), 7.

Luque B., Lacasa L., Balleteros F., & Luque J. (2009). Horizontal visibility graphs: exact results for random time series. *Phys Rev E* 80, 046103 (2009).

B. Luque, L. Lacasa, F.J. Ballesteros & A. Robledo (2011). Feigenbaum graphs: a complex network perspective of chaos *PLoS ONE* 6, 9 (2011).

B.B Mandelbrot & J.W Van Ness (1968). Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review* 10, 4 (1968) 422-437.

Marvel, S.A., Mirollo, R.E. & Strogatz, S. H. (2009) Identical phase oscillators with global sinusoidal coupling evolve by Möbius group action. Chaos 19, 043104.

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

Núñez, A. M., Lacasa, L., Valero, E., Gómez, J.P. & Luque, B. (2010). Detecting series periodicity with horizontal visibility graphs. *International Journal of Bifurcation and Chaos*, in press, 2010.

Newmann MEJ & Watts DJ (1999) Renormalization group analysis of the small-world network model. Phys. Lett. A 263:341-346.

M.E.J. Newmann, The structure and function of complex networks, *SIAM Review* 45, 167-256 (2003).

JMR Parrondo, C Van den Broeck and R Kawai, Entropy production and the arrow of time, *New. J. Phys.* 11 (2009) 073008.

Parthasarathy S, Mehta S., and Srinivasan S. (2006). Robust Periodicity Detection Algorithms, Proceedings of the 15th ACM international conference on Information and knowledge management.

Peitgen H.O., Jurgens H., and Saupe D. *Chaos and Fractals: New Frontiers of Science*, Springer-Verlag, New York.(1992).

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

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

Radicchi F, Ramasco JJ, Barrat A, & Fortunato S (2008) Complex Networks Renormalization: Flows and Fixed Points. Phys. Rev. Lett. 101, 148701.

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

Robledo, A. (1999) Renormalization group, entropy optimization, and nonextensivity at criticality. Phys. Rev. Lett. 83, 12.

E. Roldan and JMR Parrondo, Estimating dissipation from single stationary trajectories, *Phys. Rev. Lett.* 105, 15 (2010).

E. Roldan and JMR Parrondo, Dissipation and relative entropy in discrete random stationary states, *in preparation*.

Schroeder M (1991) Fractals, chaos, power laws: minutes from an infinite paradise. Freeman and Co., New York.

Schuster, H.G. (1988) Deterministic Chaos. An Introduction. 2nd revised ed, Weinheim: VCH.

Shao, Z-G. (2010). Network analysis of human heartbeat dynamics. *Applied Physics Letters*, Vol. 96 073703 (2010).

A. H. Shirazi, G. Reza Jafari, J. Davoudi, J. Peinke, M. Reza Rahimi Tabar & Muhammad Sahimi (2011).Mapping stochastic processes onto complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2009, No. 07, P07046.

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

Song C, Havlin S, & Makse HA (2005) Self-similarity of complex networks. Nature 433, 392.

Song C, Havlin S, & Makse HA (2006) Origins of fractality in the growth of complex networks. Nat. Phys. 2.

Sprott J.C., and Rowlands G., Improved correlation dimension calculation, *International Journal of Bifurcation and Chaos* 11, 7 (2001) 1865-1880.

Sprott J.C., High-dimensional dynamics in the delayed Hénon map. *EJTP* 12 (2006) 19-35.

Strogatz, S.H. (1994) Nonlinear dynamics and chaos. Perseus Books Publishing, LLC.

Strogatz, S.H. (2001) Exploring complex networks. Nature 410:268-276.

Strozzi, F., Zaldívar, J. M., Poljansek, K., Bono, F. & Gutiérrez, E. (2009). From complex networks to time series analysis and viceversa: Application to metabolic networks. *JRC Scientific and Technical Reports*, EUR 23947, JRC52892.

Van Kampen N.G., *Stochastic processes in Physics and Chemistry*, Elsevier, The Netherlands (2007).

Q. Wang, S. R. Kulkarni and S. Verdú, Divergence estimation of continuous distributions based on data-dependent partitions, *IEEE Transactions on Information Theory*, 51, 9 (2005).

G. Weiss, Time-reversibility of linear stochastic processes, *J. Appl. Prob.* 12, 831-836 (1975).

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

Xu, X., Zhang, J. & Small, M. (2008). Superfamily phenomena and motifs of networks induced from time series. *PNAS*, Vol. 105, No. 50, 19601–19605.

A.C. Yang, S.S. Hseu, H.W. Yien, A.L. Goldberger, and C.-K. Peng, Linguistic analysis of the humean heartbeat using frequency and rank order statistics, *Phys. Rev. Lett.* 90, 10 (2003).

Yang, Y., Jianbo, W., Yang, H. & Mang, J. (2009). Visibility graph approach to exchange rate series. *PHYSICA A*, Vol. 388 (20), 4431–4437.

Zhang, J. & Small, M. (2006).Complex Network from Pseudoperiodic Time Series: Topology versus Dynamics. *PRL*, Vol. 96, 238701.
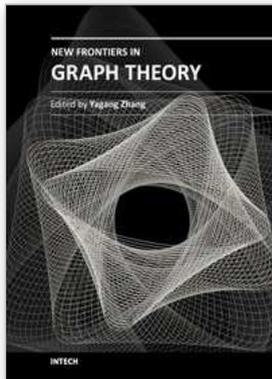
New Frontiers in Graph Theory

Edited by Dr. Yagang Zhang

Nowadays, graph theory is an important analysis tool in mathematics and computer science. Because of the inherent simplicity of graph theory, it can be used to model many different physical and abstract systems such as transportation and communication networks, models for business administration, political science, and psychology and so on. The purpose of this book is not only to present the latest state and development tendencies of graph theory, but to bring the reader far enough along the way to enable him to embark on the research problems of his own. Taking into account the large amount of knowledge about graph theory and practice presented in the book, it has two major parts: theoretical researches and applications. The book is also intended for both graduate and postgraduate students in fields such as mathematics, computer science, system sciences, biology, engineering, cybernetics, and social sciences, and as a reference for software professionals and practitioners.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds