Sequential visibility-graph motifs

Jacopo Iacovacci[*] and Lucas Lacasa[†]

School of Mathematical Sciences, Queen Mary University of London, Mile End Road, E14NS London, United Kingdom

Visibility algorithms transform time series into graphs and encode dynamical information in their topology, paving the way for graph-theoretical time series analysis as well as building a bridge between nonlinear dynamics and network science. In this work we introduce and study the concept of sequential visibility-graph motifs, smaller substructures of n consecutive nodes that appear with characteristic frequencies. We develop a theory to compute in an exact way the motif profiles associated with general classes of deterministic and stochastic dynamics. We find that this simple property is indeed a highly informative and computationally efficient feature capable of distinguishing among different dynamics and robust against noise contamination. We finally confirm that it can be used in practice to perform unsupervised learning, by extracting motif profiles from experimental heart-rate series and being able, accordingly, to disentangle meditative from other relaxation states. Applications of this general theory include the automatic classification and description of physical, biological, and financial time series.

I. INTRODUCTION

The interdisciplinary field of network science [1–4] has integrated in the last 15 years under a single paradigm of tools and techniques coming from the mathematics (combinatorics and graph theory), physics (statistical physics), and computer science (machine learning and data mining) communities, in the task of exploring, characterizing, and modeling the structure and function of large and complex networks arising in nature, technology, and society. Perhaps one of the most interesting concepts that has emerged within this synergy is that of network motifs, small subgraphs appearing with statistically significant frequencies that are suggested to represent building blocks of network architecture [5]. This local topological feature has proved to be very useful for classifying large graphs in areas such as biochemistry, neuroscience, or ecology, to cite a few (for instance networks that process information of any garment seem to share similar motif statistics [5]), or for understanding the interplay between a network's local structure and function [6]. One can even use the local information gathered by motif statistics to compare networks of different sizes, enabling a classification of networks in terms of superfamilies [7]. Both the role played by network motifs as well as the computational problem of efficiently extracting network motifs [8] are areas of currently active research.

Of course, this useful structural descriptor—and in general any topological measure—is narrowed down to those data sets and systems that have a natural representation in terms of graphs. As a matter of fact, in some of the most challenging and complex systems that scientists face nowadays (let it be spatiotemporal chaotic, or turbulent systems, the financial system, brain activity, etc), information is available in the form of temporal streams of data: series describing the time evolution of certain observables. Interestingly enough, in recent years a novel branch in data analysis has started to

transform time series into graph-theoretical representations. Among other interesting possibilities [9–13], the family of visibility algorithms [14–17] stand out as computationally simple methods to transform time series into networks which are capable of mapping seemingly hidden structure of the series and the underlying dynamics into graph space, with the peculiarity of often being analytically tractable [14]. Here we extend, via visibility algorithms, a tailored notion of network motifs to the realm of time series analysis and classification [11].

The rest of the paper goes as follows. After recalling the basics of visibility (VG) and horizontal visibility graphs (HVG), we define sequential VG or HVG motifs (Sec. II) and develop a mathematical theory for the HVG case (Sec. III) that allows us to easily derive analytical expressions for the motif profiles of several classes of stochastic and deterministic dynamical systems. We prove, accordingly, that sequential HVG motifs are informative features that can easily distinguish among different types of complex dynamics. In Sec. IV we further show that such discrimination is robust, even when the signals under study are polluted by large amounts of measurement noise, enabling its use in empirical (experimental) time series, i.e., in practical problems. We summarize our results on synthetic time series in Sec. V and finally make use of this methodology in a real scenario in Sec. VI, where we are able classify different physiological time series and efficiently disentangle meditative from general relaxation states by using the motif profiles (only five numbers per subject) extracted from heartbeat time series. In Sec. VII we conclude.

II. VISIBILITY GRAPHS AND MOTIFS

Visibility algorithms [14–17] are a family of methods to map time series into graphs, in order to exploit the tools of graph theory and network science to describe and characterize both the structure of time series and their underlying dynamics. Let $\mathcal{S} = \{x(t)\}_{t=1}^{T}$ be a real-valued time series of T data. A so-called natural visibility graph (VG) is a planar graph of T nodes in association with $\mathcal{S}$, such that (i) every datum $x(i)$ in the series is related to a node i in the graph (hence the graph
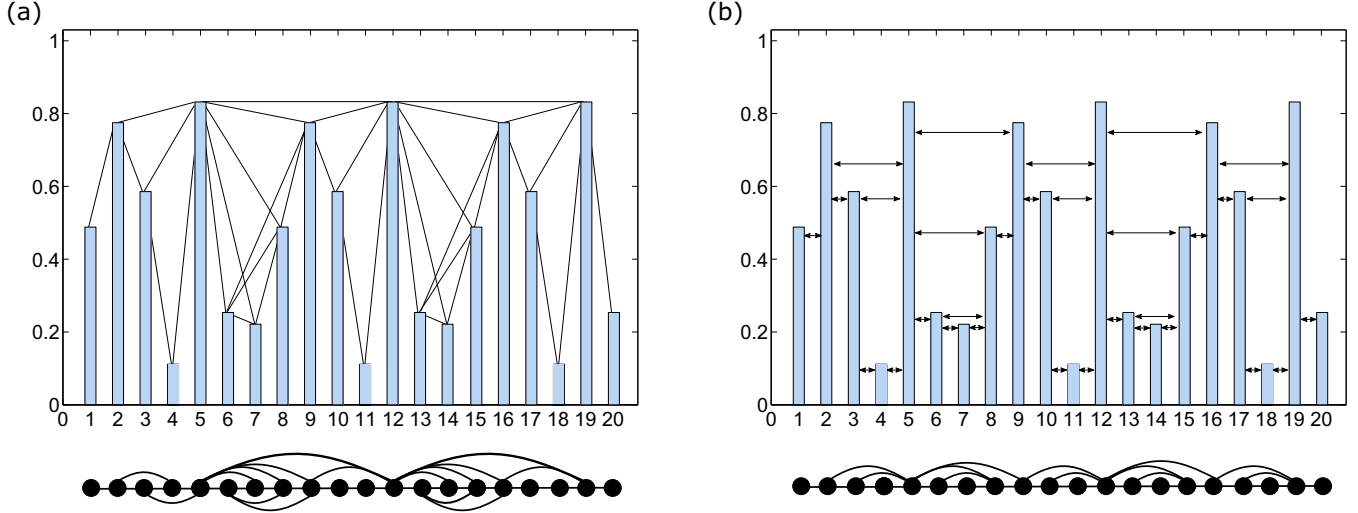
———————
[*]j.iacovacci@qmul.ac.uk
[†]l.lacasa@qmul.ac.uk

FIG. 1. Schematic of two families of visibility algorithms. (a) Natural visibility algorithm applied to 20 data points of a periodic time series (top) and the corresponding visibility graph (VG) (bottom); each datum in the series corresponds to a node in the graph and two nodes are connected if their corresponding data heights show mutual visibility (see the text). (b) Horizontal visibility algorithm applied to the same series (top) and the corresponding horizontal visibility graph (HVG) (bottom); each datum in the series corresponds to a node in the graph and two nodes are connected if their corresponding data heights show horizontal visibility (see the text).

nodes inherit a natural ordering), and (ii) two nodes i and j are connected by an edge if any other datum $x(k)$ where $i < k < j$ fulfils the following *convexity* criterion:

$$x_k < x_i + \frac{k-i}{j-i}[x_j - x_i], \quad \forall k : i < k < j.$$

By construction, VGs are connected graphs with a natural Hamiltonian path given by the sequence of nodes $(1,2,\ldots,T)$, whose topology is invariant under a set of basic transformations in the series, including horizontal and vertical translations. An illustration of this method is shown in panel (a) of Fig. 1, where we plot a time series and its associated VG. VGs inherit in their topology the structure of the time series, in such a way that periodic, random, and fractal series map into motif-like, random exponential, and scale-free networks, respectively. It has been shown that VGs are well suited to handle nonstationary data [18–20].

A so-called horizontal visibility graph (HVG) is defined as a subgraph of the VG, obtained by restricting the visibility criterion and imposing horizontal visibility instead. In this case, two nodes i and j are connected by an edge in the HVG if any other datum $x(k)$ where $i < k < j$ fulfills the following *ordering* criterion:

$$x_k < \inf(x_i, x_j), \quad \forall k : i < k < j.$$

Such subgraph is indeed an outerplanar graph [21] [see Fig. 1, panel (b), for an illustration]. Interestingly, HVGs inherits some of the properties of VGs and, on top of that, are computationally more efficient [42] and analytically tractable. Accordingly, several analytical properties of these families of graphs [15,22], associated with different classes of dynamics including canonical routes to chaos [23–26], have been investigated in recent years. For instance, for the class of Markovian processes with an integrable invariant measure the values of the degree distribution $P(k)$ can be calculated analytically using a formal diagrammatic theory [22].

We are now ready to introduce a new topological property of VGs and HVGs.

Definition (sequential VG and HVG n-node motifs). Consider a VG or HVG of N nodes, associated with a time series of N data, and label the nodes according to the natural ordering induced by the arrow of time (i.e., the trivial Hamiltonian path). Set $n < N$ and consider, sequentially, all the subgraphs formed by the sequence of nodes $\{s, s+1, \ldots, s+n-1\}$ (where s is an integer that takes values in $[1, N-n+1]$) and the edges from the VG or HVG only connecting these nodes: these are defined as the sequential n-node motifs of the VG or HVG. This is akin to defining a sliding window of size n in graph space that initially covers the first n nodes and sequentially slides, in such a way that for each window, one can associate a motif by (only) considering the edges between the n nodes belonging to that window.

Note that, importantly, this definition differs from the one of a standard network motif (which looks at the frequencies of appearance of all subgraphs of a given size, without imposing any restriction on the nodes forming a given subgraph), as here it is required that the labels of the nodes appearing in a motif are in strict sequential order—this is consistent with the vertex ordering of the natural Hamiltonian path induced by construction in the VGs and HVGs. That is, in order to preserve in graph space the dynamical information of the series, the n nodes of an n-size motif are taken in sequential order, and only those edges that connect nodes from the motif are considered. For readability, from now on we will call these simply VG and HVG motifs but the reader should not get confused and be reminded that these are not directly the standard notion of network motifs computed on a VG or HVG. Some basic properties of these motifs are as follows:

(1) Trivially, there are a total of $N - n$ motifs (which can be the same motifs or not) within each VG or HVG.

(2) Each motif is a subgraph of the original VG or HVG. Moreover, HVG motifs are outerplanar and have a trivial

TABLE I. Enumeration of all 3- and 4-node motifs. Each motif can be characterized according to a hierarchy of inequalities in the associated time series. Note that for real-valued aperiodic dynamics the type-II 4-node motif has a null probability of occurrence as the probability that two data in the time series repeat vanishes almost surely (if, on the other hand, the series only take values from a finite set then this motif has a finite probability). For the rest, the probability of each motif reduces to the measure of the set of inequalities (see the text).

Motif label	Motif type	Inequality set
1	•–•–•	$\{\forall(x_0,x_2), x_1 > x_0\} \cup \{\forall x_0, x_1 < x_0, x_2 < x_1\}$
2	•⌒•–•	$\{\forall x_0, x_1 < x_0, x_2 > x_1\}$
1	•–•–•–•	$\{\forall(x_0,x_1), x_2 < x_1, x_3 < x_2\} \cup \{\forall(x_0,x_3), x_1 > x_0, x_2 > x_1\}$
2	•–•–•–•	$\{\forall x_0, x_1 < x_0, x_2 = x_1, x_3 > x_2\}$
3	•–•–•–•	$\{\forall x_0, x_1 < x_0, x_1 < x_2 < x_0, x_3 < x_2\} \cup \{\forall(x_0,x_3), x_1 < x_0, x_2 > x_0\}$
4	•–•–•–•	$\{\forall x_0, x_1 > x_0, x_2 < x_1, x_3 > x_2\} \cup \{\forall x_0, x_1 < x_0, x_2 < x_1, x_2 < x_3 < x_1\}$
5	•–•–•–•	$\{\forall x_0, x_1 < x_0, x_1 < x_2 < x_0, x_3 > x_2\}$
6	•⌒•⌒•–•	$\{\forall x_0, x_1 < x_0, x_2 < x_1, x_3 > x_1\}$

Hamiltonian path; thus HVG motifs are also HVGs [21]. As a result, there are only 6 admissible motifs of size 4, and 2 admissible motifs of size 3 (see Table I for an enumeration).

(3) Computational complexity: computing motifs in both VGs and HVGs is extremely efficient. If instead of exploring the motif occurrence in the structure of the adjacency matrix, one directly examines the set of inequalities reported in Table I, one directly has an algorithm that runs in *linear* time $O(N)$ for HVG motifs. A similar complexity is found for VG motifs [31].

As is done traditionally with network motifs [7], we can compare VGs or HVGs associated with different time series and dynamics by comparing the relative occurrence of each motif inside a VG or HVG. In order to do that, we introduce the extension to the VG or HVG realm of a significance profile.

Definition (VG or HVG motif profile $\mathbf{Z}^n$). Let p be the total number of admissible VG or HVG motifs with n nodes. Assign to each of these p motifs a label from 1 to p (that is, choose an ordering for the motifs). The motif assigned with the label i will be called a type-i motif. Then, we define the n-node VG or HVG motif significance profile $\mathbf{Z}^n$ (or simply HVG motif profile) of a certain time series of size N as the vector function $\mathbf{Z}^n : n \in \mathbb{N} \to [\mathbb{P}_1^n, \ldots, \mathbb{P}_p^n] \in [0,1]^p$ whose output is a vector of p components, where the ith component, $\mathbb{P}_i^n$, is the relative frequency of the type-i motif.

Several technical comments are in order.

(1) First, since $\mathbf{Z}^n$ are n-dimensional real vectors, any L_p norm induces a natural similarity measure (distance) between two graphs.

(2) Second, $\mathbf{Z}^n$ has, by construction, unit L_1 norm, as $\sum_{i=1}^{p} |\mathbb{P}_i^n| = \sum_{i=1}^{p} \mathbb{P}_i^n = 1$.

(3) Third, note that if one considers dynamical processes instead of individual time series, then the estimated relative frequencies $\mathbb{P}_i^n$ for an individual realization of the dynamical process converge for infinitely long series to the probabilities of type-i motif associated with the process. For the motif profile to be a well-defined feature of a certain dynamical process, it needs to be self-averaging. We check this property by estimating $\mathbf{Z}^n$ for an ensemble of realizations of the process, computing the mean $\langle \mathbb{P}_i^n \rangle$ and standard deviation $\sqrt{\langle [\mathbb{P}_i^n]^2 \rangle - \langle \mathbb{P}_i^n \rangle^2}$ over this ensemble, and checking that the standard deviation is small (meaning that a single realization

provides a good description of the average behavior). As we will show below, both VG and HVG motif profiles have very good self-averaging properties. In any case, for every dynamical process considered in this work, instead of $\mathbb{P}_i^n$ we compute $\langle \mathbb{P}_i^n \rangle$ and $\sqrt{\langle [\mathbb{P}_i^n]^2 \rangle - \langle \mathbb{P}_i^n \rangle^2}$, but for readability, from now on we will drop the $\langle \cdot \rangle$ for the elements of the motif profile, as we found that for the size of the series used in the numerical analysis, $\sqrt{\langle [\mathbb{P}_i^n]^2 \rangle - \langle \mathbb{P}_i^n \rangle^2}$ was very small and hence $\mathbb{P}_i^n \approx \langle \mathbb{P}_i^n \rangle$.

(4) Fourth, note at this point that the definition of the VG or HVG motif profile is different from standard profiles (significance profile, subgraph ratio profile) defined in the literature [7], as in the latter case, they make use of a null model (ensemble of randomized networks) to appropriately normalize each frequency. The rationale for this normalization is that one wants to compare motif statistics across very different networks (with different sizes and degree sequences), so variations in the motif relative frequencies only due to size effects need to be removed to be able to correctly compare across different networks. The reader will quickly come to the conclusion that, in the context of VGs and HVGs, the null model is not a randomized ensemble of the graph under study (which would not yield a VG or HVG with high probability), but on the contrary, it should be the VG or HVG of a randomization of the *time series* under study. In other words, normalization in the case of VG or HVG profiles should deal with the motif statistics of uncorrelated random series (i.i.d. white noise or surrogate series that preserve certain structures) with similar probability densities to those of the series under study. In the next section we will prove that, in the case of HVGs (which will be the family of visibility graphs under study), such null model has a universal motif profile, independent of the probability density of the i.i.d. process. Therefore, it is not necessary in this case to normalize each profile accordingly as this would only yield a trivial, constant rescaling.

For illustration purposes, let $n = 4$, and consider two different dynamical processes: (i) white Gaussian noise described by the map $x_t = \xi$, where ξ are independent and identically distributed (i.i.d.) Gaussian random variables $\xi \sim \mathcal{N}[0,1]$, and (ii) chaotic dynamics given by the fully chaotic logistic map $x_{t+1} = 4x_t(1 - x_t)$. In order to estimate the probability of appearance of each of the motifs, we have generated a
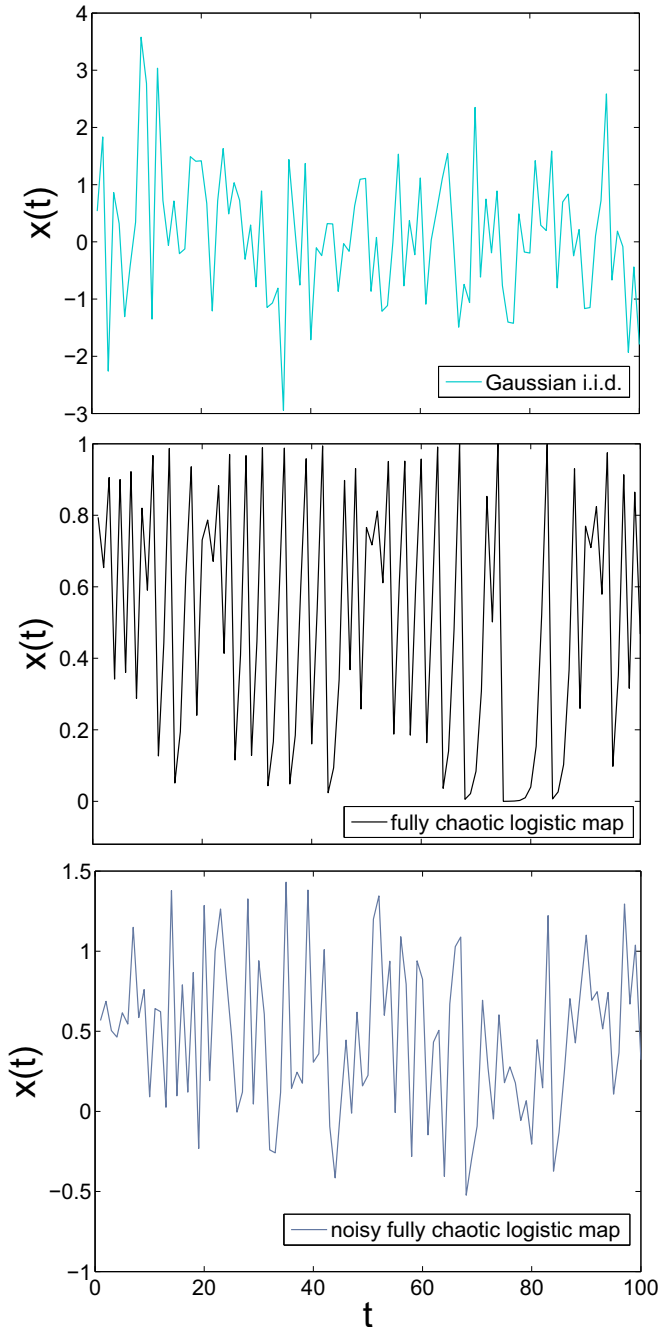
FIG. 2. Sample time series from (a) i.i.d. Gaussian white noise, (b) fully chaotic logistic map, and (c) fully chaotic logistic map polluted with a certain amount of extrinsic white noise are shown for illustrative purposes. Visibility graph motifs can be extracted from these series to reveal differences in their intrinsic structure.

FIG. 3. Four-node motif profiles $\mathbf{Z}^4$ associated with Gaussian white noise (red squares) and with a fully chaotic logistic map (black stars) extracted respectively from HVG (a) and VG (b). Each dot represents the relative frequency of a given motif, averaged over an ensemble of 100 realizations of each process (time series of $N = 10^4$ data per realization). Standard deviations of each motif's relative frequency over the ensemble are plotted as error bars, which are not visible as error bars fall inside the symbols. We conclude that these motifs can be used to distinguish between deterministic and stochastic dynamics.

time series of size $N = 10^4$ data for both processes (sample time series can be seen in the top panels of Fig. 2), and we have computed the relative frequencies of each motif. Results, averaged over an ensemble of 100 realizations, are shown in Fig. 3 (error bars describing the ensemble standard deviation are contained inside the symbols); in panel (a) we plot the HVG motif profile, whereas in panel (b) we plot the VG profile. As we can see, in every case the type-II motif is absent. The simple
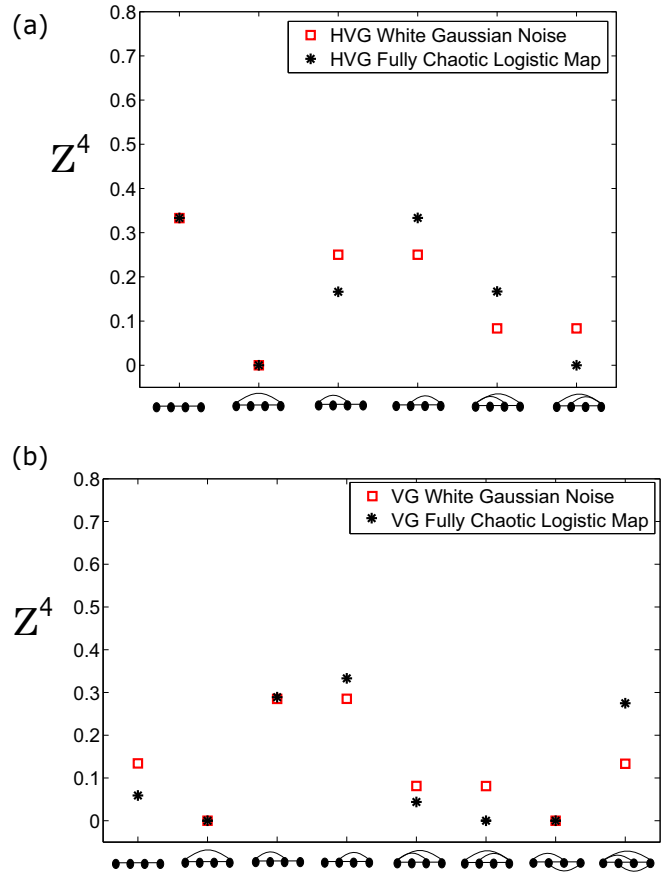
reason is that this profile is absent for irregular (aperiodic) real-valued time series, by construction (see Table I).

For the chaotic process, some other motifs are absent: this is related to forbidden patterns arising in chaotic dynamics. More importantly, in both panels, the average relative frequency of some motifs seems to be different for both dynamical processes, enabling the possibility of using both HVG and VG motif profiles to distinguish among different dynamical origins. From now on we will focus our motif analysis on the horizontal visibility graphs (HVGs) alone, and comparison with the VG case is left for future work [31]. In the next section we advance a theory to compute the motif profile $\mathbf{Z}^n$ in an exact way for different classes of dynamical systems. We will confirm that HVG motifs can indeed distinguish several kinds of dynamics, and we will explore how to build on this peculiar property for feature-based classification.

III. THEORY

In order to numerically explore and compute the frequency of each HVG motif, one can generate the HVG associated with a given time series and count the presence of each motif directly from the adjacency matrix. However, in this section we will show that it is not necessary to do that as, via the zero-order terms of a diagrammatic expansion recently advanced [22], we can also work out the motif occurrence directly from the exploration of the time series, that enables motif computation in linear time. This will allow us to build a theory by which the motif profiles can be computed exactly for a large set of classes of dynamics that fulfill certain properties. Let us consider a dynamical process $\mathcal{H} : \mathbb{R} \to \mathbb{R}$ with a smooth invariant measure $f(x)$ that fulfils the Markov property. That is, from a probabilistic point of view, conditional probabilities fulfill $f(x_n|x_{n-1}, x_{n-2}, \dots) = f(x_n|x_{n-1})$, where $f(x_n|x_{n-1})$ is the transition probability distribution [note that this concept has a clear meaning in random dynamical systems, whereas for deterministic systems, say maps $x_{t+1} = \mathcal{H}(x_t)$, the Markov property is also trivially fulfilled with $f(x_2|x_1) = \delta(x_2 - \mathcal{H}(x_1))$, where $\delta(x)$ is the Dirac-delta distribution]. The key element is that for these processes, each HVG motif has a probability of appearance as a subgraph that can *directly* be computed as the measure of a set of ordering inequalities that take place in the time series. For instance, for $n = 3$ and $n = 4$, probabilities associated with the appearance of a certain motif are based on integrals of the form

$$\int f(x_0)dx_0 \int f(x_1|x_0)dx_1 \int f(x_2|x_1)dx_2 \quad (1)$$

for $n = 3$, and

$$\int f(x_0)dx_0 \int f(x_1|x_0)dx_1 \int f(x_2|x_1)dx_2 \int f(x_3|x_2)dx_3 \quad (2)$$

for $n = 4$. The range of integration and the shape of the conditional probabilities are particular for each motif and each process, respectively. First, the range of integration fully determines the motif. In Table I we depict the conditions in the time series that have to be fulfilled among n consecutive data $x_0, x_1, \dots, x_{n-1}$ to yield a certain motif of size n in the HVG, for $n = 3, 4$ (extension to arbitrary n is easy but gets cumbersome as n increases). It can be proved quite easily that a given motif appears in an HVG if and only if these ordering restrictions are fulfilled in the time series. These restrictions directly translate in the integration range of the probabilities; we illustrate this principle in an example. The first motif, $\mathbf{Z}_1^4$, according to Table I is guaranteed when 4 consecutive values x_0, x_1, x_2, and x_3 are such that $\{\forall(x_0, x_1), x_2 < x_1, x_3 < x_2\} \cup \{\forall(x_0, x_3), x_1 > x_0, x_2 > x_1\}$. Accordingly, if $x \in [a,b] \subset \mathbb{R}$, the probability of this event is

$$\mathbf{Z}_1^4 \equiv \mathbb{P}_1^4 = \int_a^b f(x_0)dx_0 \int_a^b f(x_1|x_0)dx_1 \int_a^{x_1} f(x_2|x_1)dx_2$$
$$\times \int_a^{x_2} f(x_3|x_2)dx_3 + \int_a^b f(x_0)dx_0$$

$$\times \int_{x_0}^b f(x_1|x_0)dx_1 \int_{x_1}^b f(x_2|x_1)dx_2$$

$$\times \int_a^b f(x_3|x_2)dx_3. \quad (3)$$

Analogous expressions can be found for the rest of the probabilities that form the motif profile $\mathbf{Z}$. These terms are nothing but the contributions to the degree distribution at zero order from a diagrammatic expansion in the number of hidden nodes [22]. From a geometric point of view, the first motif will not appear in fast-fluctuating signals and hence deals with the degree of smoothness of a time series at short (order n) scales, whereas the other motifs deal with certain fluctuation shapes. Accordingly, in those processes where the degree of smoothness can vary—such as in fractional Brownian motion, where the smoothness of the signal increases with the Hurst exponent—we would expect that the first motif is particularly informative, whereas for fast-fluctuating series we expect this motif to be less informative. Integrals accounting for the probabilities are easy to deal with; in several cases these are exactly solvable, and in general one can solve them up to arbitrary precision with any symbolic programming software. In what follows we determine the motif profiles for i.i.d. (white noise), colored noise with exponentially decaying correlations, and deterministic chaos (fully chaotic logistic map). We show that $\mathbf{Z}^4$ capture enough information to easily distinguish different processes and thus represent excellent features for series classification.

Relation with ordinal patterns. At this point it is important to highlight the relation between the probability of occurrence of given HVG motifs and the probability of occurrence of so-called ordinal patterns [27,30]. In the theory proposed by Bandt and Pompe in [27] for the case of the embedding dimension equal to 4 one proceeds to map each local time series segment of size 4 into an ordering symbol of 4 letters from the alphabet {0,1,2,3} (where the largest value maps to the letter 0, the second largest to 1, the third largest to 2, and the smallest to 3). There are $4! = 24$ permutations, defining 24 symbols (ordinal patterns) whose frequencies are then counted to measure the so-called permutation entropy that acts as a complexity measure of the series [27]. Interestingly, the probability of occurrence of each HVG motif indeed reduces to the probability of occurrence of a set of possible ordinal patterns (this is no longer the case for VG motifs [31]). For instance, $\mathbf{Z}_1^4$ is the probability of finding any of the ordinal patterns 0123,1023,1203,1230,2103,2130,2310, or 3210, and similarly the rest of the motif probabilities can be linked to the probability of appearance of different sets of ordinal patterns. Accordingly, HVG motifs indeed induce a particular partition of the set of ordinal patterns. The HVG motif profile is thus intimately linked with the so-called permutation spectrum [29] that accounts for the histogram of ordinal patterns.

A. i.i.d.

Let us start by considering time series generate by i.i.d. uniform random variables $\xi \sim U[0,1]$. In this case we have $a = 0, b = 1, f(x) = 1$, and $f(x|y) = f(x) \; \forall y$, and simply enough, probabilities defined by Eqs. (1) and (2) easily

factorize. According to Table I, after a little bit of calculus we find

$$\mathbf{Z}^3 = \left[\frac{2}{3},\frac{1}{3}\right]; \quad \mathbf{Z}^4 = \left[\frac{8}{24},0,\frac{6}{24},\frac{6}{24},\frac{2}{24},\frac{2}{24}\right]. \quad (4)$$

Note that these results are in perfect quantitative agreement with numerics performed for finite-size series (left panel of Fig. 3); we will show in the next subsection that results for finite series converge quite fast to the (asymptotic) theory as the series size increases. Interestingly, results indeed coincide despite the fact that the theoretical values were computed for *uniform* white noise ($f(x) = 1$), while the numerics in Fig. 3 were performed on *Gaussian* white noise [where $f(\cdot)$ is the Gaussian function]. This suggests that i.i.d. may have a universal HVG motif profile, indeed independent of $f(\cdot)$. We now state and prove a theorem that actually guarantees this result.

Theorem 1. Consider a bi-infinite series of i.i.d. random variables extracted from a continuous distribution $f(x)$ with support (a,b), where $a,b \in \mathbb{R}$. Then the probability of finding n-node HVG motifs (with $n = 3,4$) follows Eq. (4), independently of the shape of $f(x)$.

Proof. The proof is a constructive one. We only give here the explicit proof for $\mathbb{P}_1^4$, as the proofs for the rest of the probabilities follow analogously. We rely on the cumulative distribution function $F(x)$, defined as $\int_a^x f(x')dx' = F(x)$, with properties $F(a) = 0, F(b) = 1$, and

$$f(x)F^{n-1}(x) = \frac{dF^n(x)}{ndx}. \quad (5)$$

We have

$$\mathbb{P}_1^4 = \int_a^b f(x_0)dx_0 \int_a^b f(x_1)dx_1 \int_a^{x_1} f(x_2)dx_2$$
$$\times \int_a^{x_2} f(x_3)dx_3 + \int_a^b f(x_0)dx_0 \int_{x_0}^b f(x_1)dx_1$$
$$\times \int_{x_1}^b f(x_2)dx_2 \int_a^b f(x_3)dx_3.$$

Using the properties of $F(x)$, the first term above is then

$$\int_a^b f(x_0)dx_0 \int_a^b f(x_1)dx_1 \int_a^{x_1} f(x_2)dx_2 \int_a^{x_2} f(x_3)dx_3$$
$$= \int_a^b f(x_0)dx_0 \int_a^b f(x_1)dx_1 \int_a^{x_1} f(x_2)F(x_2)dx_2$$
$$= \int_a^b f(x_0)dx_0 \int_a^b f(x_1)\frac{F^2(x_1)}{2}dx_1$$
$$= \int_a^b \frac{f(x_0)}{6}dx_0 = \frac{1}{6},$$

and analogously for the second term,

$$\int_a^b f(x_0)dx_0 \int_{x_0}^b f(x_1)dx_1 \int_{x_1}^b f(x_2)dx_2 \int_a^b f(x_3)dx_3$$
$$= \int_a^b f(x_0)dx_0 \int_{x_0}^b f(x_1)[1-F(x_1)]dx_1$$

$$= \int_a^b f(x_0)\left[\frac{1}{2} - F(x_0) + \frac{F^2(x_0)}{2}\right]dx_0$$
$$= \frac{F(x_0)}{2} - \frac{F^2(x_0)}{2} + \frac{F^3(x_0)}{6}\Big|_a^b = \frac{1}{6}, \quad (6)$$

hence $\mathbb{P}_1^4 = 2/6 = 8/24$, coinciding with the result for uniform and Gaussian series, and being independent of $f(x)$. The rest of the elements in $\mathbf{Z}^4$ are computed analogously. $\blacksquare$

As a matter of fact, the independency from $f(x)$ can be trivially extended for an arbitrary size of the motif n. This is intuitive so we only give here the strategy of a proof. The main ingredient which is required for this independency to hold $\forall n$ is that the limits of the nth integral are either the extremes of the distribution support a,b [where the cumulative distribution $F(x)$ takes the constant values 0 and 1, respectively, and independently of $f(x)$], or other variables $x_0 \ldots x_{n-1}$. In this latter case, one can use iteratively the property in Eq. (5) to solve these integrals up to the last one (in x_0), whose range is always (a,b) and where $F(a) = 0, F(b) = 1$ can be finally applied, to give a result which will not depend on the precise shape of $f(x)$.

According to Theorem 1, Gaussian, uniform, power law, etc., uncorrelated random series all have the same HVG motif profiles. As a by-product, for any kind of sufficiently long time series $\{x_t\}_{t=1}^N$ where $x_t \in f(x)$ and $f(x)$ is continuous, if we randomize (shuffle) the time series, the motif profile of the randomized series is equal to Eq. (4). This is the reason why, at odds with the standard definition of a network's motif profile, for HVGs we do not need to rescale $\mathbf{Z}$ in any way to be able to compare across different time series and dynamical process.

Another notable consequence of Theorem 1 is that it guarantees that series for which $\mathbf{Z}^4$ differ (even in the case of sufficiently long time series) from Eq. (4) are not uncorrelated random series. This suggests a simple test for randomness [15]. For instance, one can use a Pearson's χ^2 hypothesis test, where the null hypothesis is that the observed time series of N data is random and uncorrelated (white noise). The test statistic is then

$$\chi^2 = (N - n)\sum_{i=1}^p \frac{\left[\mathbb{P}_i^4(\text{observed}) - \mathbb{P}_i^4(\text{i.i.d.})\right]^2}{\mathbb{P}_i^4(\text{observed})}. \quad (7)$$

χ^2 upper-critical values with $p - 1$ degrees of freedom, for $p = 6$ ($n = 4$), are 11.07 and 15.086 at the 95% and 99% significance level (meaning that values of the χ^2 larger than 11.07 suggest that the observed series is not random at the 95% significance level). More rigorously, as the type-II motif is forbidden for aperiodic dynamics, we have only $p = 5$ different motifs of size $n = 4$, so the χ^2 upper-critical values should be considered for 4 degrees of freedom: 9.49 (95%) and 13.28 (99%).

B. Deterministic chaos: Fully chaotic logistic map

As previously stated, deterministic maps $x_{t+1} = \mathcal{H}(x)$ are indeed Markovian, and for these situations the conditional probability is simply $f(x_2|x_1) = \delta(x_2 - \mathcal{H}(x_1))$, where $\delta(x)$ is the Dirac-delta distribution. Therefore Eqs. (1) and (2), combined with inequality sets given in Table I, can be
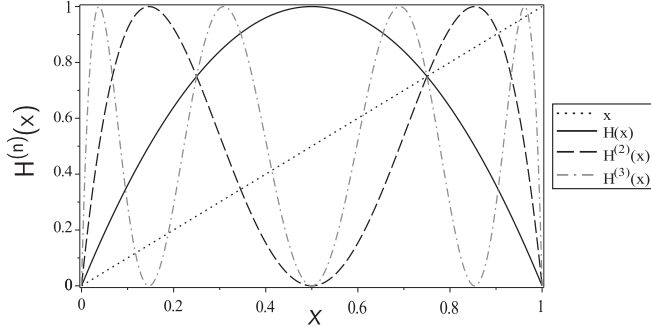
FIG. 4. Cobweb plot of the iterates of the fully chaotic logistic map $\mathcal{H}(x) = 4x(1-x)$.

used to compute the motif profiles for different deterministic processes. In these cases, one has to deal with simple integrals of the form

$$\int_p^q \delta(x-y)dx = \begin{cases} 1, & y \in [p,q], \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

While in principle any deterministic process can be studied, we are interested in complex signals, so we focus on irregular, aperiodic dynamics. As a paradigmatic case, we tackle the fully chaotic logistic map

$$\mathcal{H}(x) = 4x(1-x), \quad x \in [0,1], \quad f(x) = \frac{1}{\pi\sqrt{x(1-x)}}.$$

In this case, $f(x)$ is the invariant measure that describes in a probabilistic way the average time spent by a chaotic trajectory in each region of the attractor. Let us start by considering $\mathbf{Z}^3 := (\mathbb{P}_1^3, \mathbb{P}_2^3)$, for which

$$\mathbb{P}_1^3 = \int_0^1 f(x_0)dx_0 \int_{x_0}^1 \delta(x_1 - \mathcal{H}(x_0))dx_1$$
$$\times \int_0^1 \delta(x_2 - \mathcal{H}^2(x_0))dx_2,$$

$$\mathbb{P}_2^3 = \int_0^1 f(x_0)dx_0 \int_0^{x_0} \delta(x_1 - \mathcal{H}(x_0))dx_1$$
$$\times \int_{x_1}^1 \delta(x_2 - \mathcal{H}^2(x_0))dx_2.$$

According to property in Eq. (8), the Dirac-delta integrals only have the effect of shrinking the range of integration of x_0. For instance, for $\mathbb{P}_1^3$, the integral in x_1 requires $\mathcal{H}(x_0) > x_0$, whereas the integral in x_2 simply requires $\mathcal{H}^2(x_0) \in [0,1]$. While the latter inequality is fulfilled for all $x_0 \in [0,1]$ (and thus has no effect), the former one requires $x_0 \in [0,3/4]$. This can be easily seen from the cobweb plot of $\mathcal{H}(x)$ and its iterates (see Fig. 4): $\mathcal{H}(x) > x$ for $x \in [0,3/4]$. Altogether,

$$\mathbb{P}_1^3 = \int_0^{3/4} f(x_0)dx_0 = 2/3.$$

On the other hand, motif normalization imposes $\mathbb{P}_2^3 = 1/3$. The same result is obviously found if we compute $\mathbb{P}_2^3$ explicitly: in this case the integral in x_1 requires $\mathcal{H}(x_0) < x_0$, which holds when $x_0 \in [3/4,1]$, and the integral in x_2 requires $\mathcal{H}^2(x_0) > x_1 \leftrightarrow \mathcal{H}^2(x_0) > \mathcal{H}(x_0)$. Looking at the cobweb plots, this final

condition is met in two subintervals, so the intersection with the first condition yields a final interval $x_0 \in [3/4,1]$, for which

$$\mathbb{P}_2^3 = \int_{3/4}^1 f(x_0)dx_0 = 1/3,$$

as expected. These results coincide with those found for i.i.d. series, meaning that $\mathbf{Z}^3$ does not capture enough structure to distinguish both processes. Let us proceed in an equivalent way to compute $\mathbf{Z}^4 = (\mathbb{P}_1^4, \dots, \mathbb{P}_6^4)$. It becomes evident that integrals associated with x_n deal with the cobweb plots of $\mathcal{H}(x), \mathcal{H}^2(x), \dots, \mathcal{H}^n(x)$. Accordingly, these integrals are ultimately related to the structure of fixed points of $\mathcal{H}^n(x)$, and to the solutions of equations of the form $\mathcal{H}^r(x) = \mathcal{H}^s(x)$ for some r and s. We only have algebraic closed expressions for the fixed points of $\mathcal{H}(x) \to \{0,3/4\}$ and $\mathcal{H}^2(x) \to \{0, \frac{5-\sqrt{5}}{8}, 3/4, \frac{5+\sqrt{5}}{8}\}$ [for $n \geq 3$, $\mathcal{H}^n(x)$ is a polynomial of order larger or equal to 6 and according to Abel-Ruffini's theorem, the set of fixed points does not have in general an algebraic expression; however we can compute them up to arbitrary precision]. Other values of interest include the roots of $\mathcal{H}^3(x) = \mathcal{H}^2(x)$, and especially the largest one $x = 1/2 + \sqrt{3}/4$.

Let us show how to compute one of these motif probabilities. For instance,

$$\mathbb{P}_5^4 = \int_0^1 f(x_0)dx_0 \int_0^{x_0} \delta(x_1 - \mathcal{H}(x_0))dx_1$$
$$\times \int_{x_1}^{x_0} \delta(x_2 - \mathcal{H}^2(x_0))dx_2 \int_{x_2}^1 \delta(x_3 - \mathcal{H}^3(x_0))dx_3, \quad (9)$$

which reduces to

$$\mathbb{P}_5^4 = \int_p^q f(x_0)dx_0,$$

where $[p,q]$ can be hierarchically obtained as
$\mathcal{H}(x_0) < x_0 \cap [0,1] \Rightarrow x_0 \in [3/4,1]$;
$\mathcal{H}^2(x_0) < x_0 \cap \mathcal{H}^2(x_0) > \mathcal{H}(x_0) \cap [3/4,1] \Rightarrow x_0 \in [\frac{5+\sqrt{5}}{8},1]$;
$\mathcal{H}^3(x_0) > \mathcal{H}^2(x_0) \cap [\frac{5+\sqrt{5}}{8},1] \Rightarrow x_0 \in [x_p,1]$, where x_p is the second largest root fulfilling $\mathcal{H}^3(x_p) = \mathcal{H}^2(x_p)$, i.e., $x_p = 1/2 + \sqrt{3}/4$. Altogether,

$$\mathbb{P}_5^4 = \int_{1/2+\sqrt{3}/4}^1 \frac{1}{\pi\sqrt{x_0(1-x_0)}}dx_0$$
$$= \frac{1}{\pi}B_{[\frac{1}{2}+\frac{\sqrt{3}}{4},1]}\left(\frac{1}{2},\frac{1}{2}\right) = \frac{1}{6}(=4/24)$$

(where B is the incomplete Beta function), which is indeed quite different from the result found for i.i.d., $\mathbb{P}_5^4$(i.i.d.) = 2/24.

Similar arguments can be used to obtain analytically the rest of the probabilities (explicit computations are put in Appendix A), finding

$$\mathbf{Z}^3 = \left[\frac{2}{3},\frac{1}{3}\right]; \quad \mathbf{Z}^4 = \left[\frac{8}{24},0,\frac{4}{24},\frac{8}{24},\frac{4}{24},0\right]. \quad (10)$$

Comparing this set of motif probabilities with the result for i.i.d. [Eq. (4)], we can conclude that $\mathbf{Z}^4$ distinguishes the fully chaotic logistic map from a purely uncorrelated stochastic process. Note, of course, that a similar derivation

can be performed in other deterministic maps; in this sense the methodology is general (however one encounters problems when the attractor has a fractal dimension, and one needs to carefully choose a proper integration theory). These exact results are also in excellent quantitative agreement with numerics performed in finite series (left panel of Fig. 3), so convergence to the theory with series size is quite fast, enabling its use in empirical cases. To be more precise, in the next subsection we make a study of how fast results for short time series converge to the asymptotic theory as series size increases.

C. Convergence of finite series

In order to be more precise about the convergence speed of finite-size numerics to the theory (which in rigor only holds for bi-infinite time series), we have computed for series of size N the numeral estimate $\mathbf{Z}^4(N)$ for both i.i.d. and the fully chaotic logistic map, and compare it with the asymptotic values $\mathbf{Z}^4$. Results are plotted in Fig. 5, where we plot $\Phi(N) = \langle \mathbf{Z}^4(N) \rangle / \mathbf{Z}^4$ as a function of the series size N (the average is with respect to realizations). Results indicate that convergence to the asymptotic theory is already reached for $N \ll 10^4$ (which is the conservative size that is used throughout this work).

D. Stochastic processes with correlations

To round off the theory section, and to explore how results deviate from i.i.d. for correlated stochastic processes, we consider colored noise with exponentially decaying correlations as described by the AR(1) process:

$$
\begin{aligned}
x_0 &= \xi_0, \\
x_t &= r x_{t-1} + \sqrt{(1-r^2)}\xi_t, \quad t \geqslant 1,
\end{aligned}
\tag{11}
$$

where $\xi_t \sim \mathcal{N}(0,1)$ is Gaussian white, and r, $0 < r < 1$ is a parameter that tunes the correlation. The autocorrelation function $C(t)$, which describes the correlation of the position at x_{t_0} and x_{t_0+t}, decays exponentially $C(t) = e^{-t/\tau}$, where the characteristic time $\tau = 1/\ln(r)$. In the limit $r \to 0$, the correlations vanish and the process reduces to a white noise signal. The limit $r \to 1$ is more delicate, but intuitively in this limit the process gets completely correlated and tends to be constant $x_{t+1} = x_t \ \forall t$.

This is a family of models parametrized by the coefficient r. For $0 < r < 1$, these models are indeed Gaussian, Markovian, and stationary, with a probability density $f(x)$ and transition probability $f(x_2|x_1)$

$$
f(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}},
$$

$$
f(x_2|x_1) = \frac{\exp\{-(x_2 - rx_1)^2/[2(1-r^2)]\}}{\sqrt{2\pi(1-r^2)}},
$$

respectively. Since x are Gaussian variables they can vary in $(-\infty, \infty)$. We focus on $\mathbf{Z}^4$ that we know gave good discriminatory results between i.i.d. and chaos. For illustration,
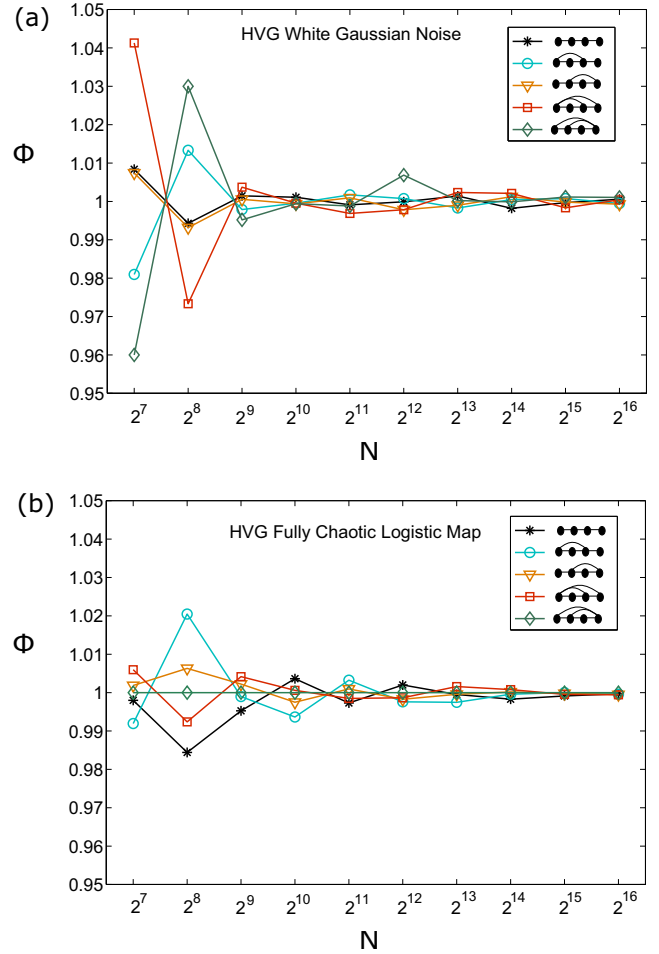

FIG. 5. The measured frequency of appearance rescaled by its theoretical value Φ is plotted for each motif associated with Gaussian white noise (a) and with a fully chaotic logistic map (b) as a function of the time series size N; results are averaged over 100 realizations. The curves oscillate with fast decreasing amplitude around the value 1 (for 2^9 the amplitude is less than 10^{-2}) indicating fast asymptotic convergence of the measured motif profile to the theoretical profile in both cases.

the first element reads

$$
\begin{aligned}
\mathbb{P}_1^4 = &\int_{-\infty}^{\infty} \frac{e^{\frac{-x_0^2}{2}}}{\sqrt{2\pi}} dx_0 \int_{-\infty}^{\infty} \frac{e^{\frac{-(x_1-rx_0)^2}{2(1-r^2)}}}{\sqrt{2\pi(1-r^2)}} dx_1 \\
&\times \int_{-\infty}^{x_1} \frac{e^{\frac{-(x_2-rx_1)^2}{2(1-r^2)}}}{\sqrt{2\pi(1-r^2)}} dx_2 \int_{-\infty}^{x_2} \frac{e^{\frac{-(x_3-rx_2)^2}{2(1-r^2)}}}{\sqrt{2\pi(1-r^2)}} dx_3 \\
&+ \int_{-\infty}^{\infty} \frac{e^{\frac{-x_0^2}{2}}}{\sqrt{2\pi}} dx_0 \int_{x_0}^{\infty} \frac{e^{\frac{-(x_1-rx_0)^2}{2(1-r^2)}}}{\sqrt{2\pi(1-r^2)}} dx_1 \\
&\times \int_{x_1}^{b} \frac{e^{\frac{-(x_2-rx_1)^2}{2(1-r^2)}}}{\sqrt{2\pi(1-r^2)}} dx_2 \int_{-\infty}^{\infty} \frac{e^{\frac{-(x_3-rx_2)^2}{2(1-r^2)}}}{\sqrt{2\pi(1-r^2)}} dx_3. \quad (12)
\end{aligned}
$$

For any particular value of r, these integrals can be evaluated up to arbitrary precision using Mathematica [32]. In Table II we report the theoretical values of $\mathbf{Z}^4(r)$ for $r \in [0.02-0.99]$.

TABLE II. Theoretical values of $\mathbf{Z}^4(r)$ for the AR(1) process evaluated at different values of the coefficient r.

r	$\mathbb{P}_1^4$	$\mathbb{P}_2^4$	$\mathbb{P}_3^4, \mathbb{P}_4^4$	$\mathbb{P}_5^4, \mathbb{P}_6^4$
0.02	0.3370	0	0.2482	0.0833
0.04	0.3406	0	0.2464	0.0833
0.06	0.3443	0	0.2446	0.0832
0.08	0.3478	0	0.2429	0.0831
0.1	0.3514	0	0.2412	0.0830
0.2	0.3690	0	0.2333	0.0822
0.3	0.3862	0	0.2260	0.0809
0.4	0.4030	0	0.2192	0.0793
0.5	0.4196	0	0.2130	0.0772
0.6	0.4359	0	0.2072	0.0748
0.7	0.4521	0	0.2018	0.0722
0.8	0.4681	0	0.1967	0.0692
0.9	0.4841	0	0.1920	0.0660
0.95	0.4919	0	0.1897	0.0643
0.97	0.4945	0	0.1888	0.0636
0.99	0.4973	0	0.1879	0.1879

These are in perfect agreement with numerical simulations performed on finite series of size $N = 10^4$ (ensemble averaged over 100 realizations) for $r = \{0, 0.1, 0.3, 0.5, 0.7, 0.9, 0.99\}$, as shown in Fig. 6. As $r > 0$ the profiles deviate from i.i.d. and thus, again, these features can easily distinguish between exponentially colored and white noise.

IV. ROBUSTNESS

In the preceding section we have developed a general theory to compute explicitly the motif profile of HVGs associated with a given type of dynamics. We have applied this theory to find theoretical expressions in the case of white and colored noise as well as chaotic dynamics, and have shown that these predictions perfectly match the results found in numerical
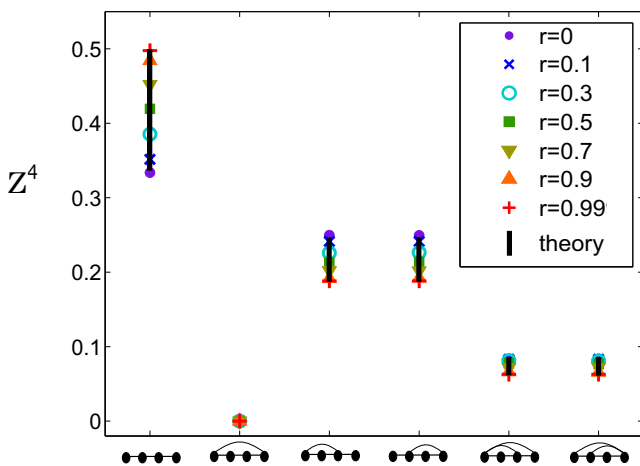

FIG. 6. HVG significance profile $\mathbf{Z}^4$ for AR(1) processes described by Eq. (11), for different values of the correlation coefficient r. When r increases the appearance probability of a motif of type I increases while the rest of the probabilities decrease. This is simply due to the fact that finding constant sequences $x_{t+3} = x_{t+2} = x_{t+1} = x_t$ becomes more probable as r increases.

simulations for reasonably short time series. The theory (which is exact in the limit of infinite-size series) is thus correct also in the case of short time series. These are nonetheless only idealized models; empirical time series, however, even if they comply to a particular dynamical system, are usually polluted with measurement noise. Therefore, before being able to apply this technique to real world phenomena, we need to assess its robustness and reliability against noise contamination. To do that, we consider a situation where a chaotic time series is contaminated with different amounts of white noise, and explore the ability of $\mathbf{Z}^4$ to detect the chaotic signal. Formally, we pollute a chaotic signal $x(t)$ with uniform white noise $\xi(a)$ and thus construct a noisy chaotic signal $Y(t)$ such that

$$
\begin{aligned}
Y_t &= x_t + \xi, \\
x_t &= 4x_{t-1}(1 - x_{t-1}), \\
\xi &\sim U[0,a], \quad 0 \leqslant a \leqslant 1,
\end{aligned}
\tag{13}
$$

where a tunes the noise power. The noise-to-signal ratio of the signal Y_t is defined as $\text{NSR} = \sigma_\xi^2 / \sigma_Y^2$ (where $\sigma_\bullet^2$ denotes the variance of signal $\bullet$); thus NSR will increase monotonically with a. For $\text{NSR} \ll 1$, the noise contamination is small. Any technique that is able to distinguish $Y(t)$ and $\xi(t)$ for increasing values of NSR is said to be robust to noise. For $\text{NSR} = 1$ the levels of the signal and the noise contamination are comparable and for $\text{NSR} > 1$ the underlying chaotic signal is effectively hidden. Of course, when a reaches a certain value it will not be possible any more to distinguish the underlying chaotic nature of the time series by looking at the motif profile. To estimate this threshold we can use two different tests:

(1) The first test makes use of the (L_1) distance in motif space between the signal and the noise $d(a) = |\mathbf{Z}^4(Y) - \mathbf{Z}^4(\text{i.i.d.})|$. This is just a simple, motif-based similarity metric between two graphs, that we use here to measure the similarity between two series. Ideally, the threshold of distinguishability is the smallest value of a for which $d(a) = 0$. However, in practice, as we are dealing with finite-size series, there will always be a small uncertainty associated with small finite-size deviations from the theory. That is, if one estimates the $\mathbf{Z}^4(\text{i.i.d.})$ with an ensemble average of m realizations of a finite random time series of N data, then for each element in the profile, the standard deviation of the estimate $\mathbb{P}_i^4$ will be a finite value (that converges to zero as N and m increases). We define $\sigma(\mathbf{Z}^4(\text{i.i.d.}))$ as the vector where the ith term is such standard deviation, for the same values of N and m used in the analysis of $Y(t)$. Then, we define the *uncertainty threshold* a^* as the smallest value of a such that $d(a) \leqslant |\sigma(\mathbf{Z}^4(\text{i.i.d.}))|$ (intuitively, a^* is the smallest value for which we do not know whether the difference in the motif profile between the empirical results and the theory is due to the fact that there is a chaotic signal underlying the process, or just due to finite-size effects).

(2) The second possibility is to use a Pearson's χ^2 hypothesis test such as Eq. (7) with 4 degrees of freedom, where the null hypothesis is that $Y(t)$ (the observed series) is just white noise (no hidden signal). In this latter case, we are not taking into account the deviations associated with finite-size effects in the profile of i.i.d., though. If $\chi^2 < 9.49$, then we cannot reject the null hypothesis at the 95% significance level: this is the limit of what we could call certain distinguishability.
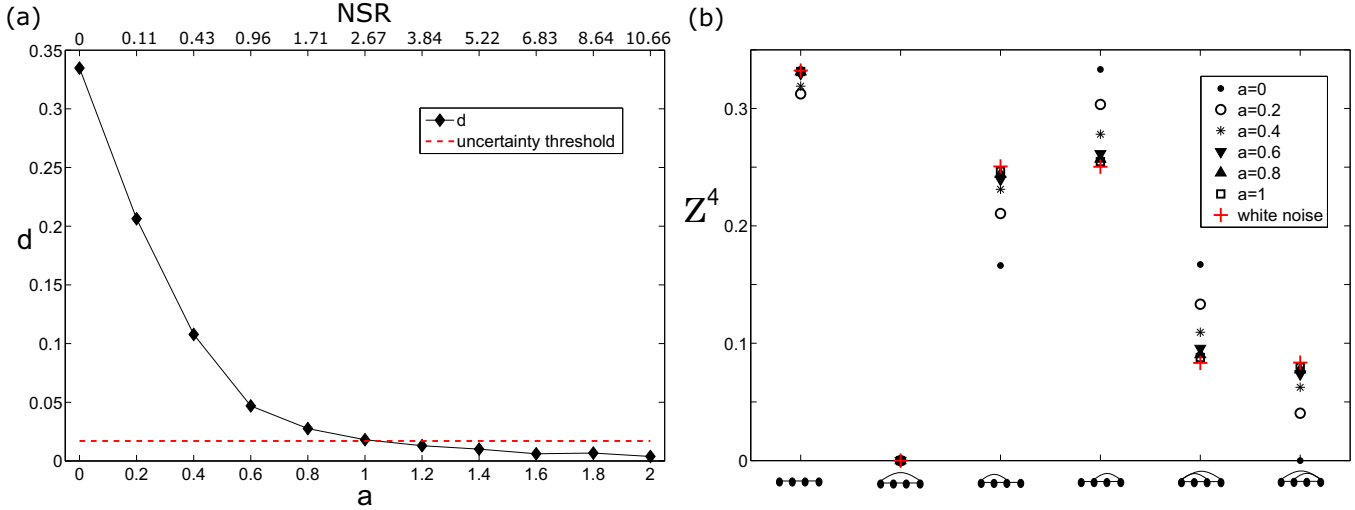
FIG. 7. Robustness of motif profiles for chaotic series (fully chaotic logistic map) polluted with white noise. (a) By increasing the amount of extrinsic noise (parametrized by a) the distance in motif space between the noisy chaotic signal and white noise decreases (see the text). The method is extremely robust as one can distinguish the noisy chaotic signal from pure white noise up to a noise-to-signal ratio NSR ≈ 2.67. (b) The 4-node motif profile of the noisy chaotic signal Y_t for different degrees of noise contamination (a). Motifs III, IV, V, and VI are the most informative as they concentrate most of the profile variability.

For each value of the parameter a, we have simulated a time series of $N = 10^4$ steps from the process $Y(t)$, and results were ensemble averaged over $m = 100$ realizations. In panel (b) of Fig. 7, we plot the motif profile as a function of a. It is interesting to observe that the probabilities which vary most with a are related to types III, IV, V, and VI, while type I seems to maintain approximately the same rate of appearance (we will show later that this is not always the case). In the panel (a) of the same figure we plot $d(a)$. As expected, $d(a)$ is a monotonically decreasing function of a, and we find $a^* \approx 1$. Remarkably, this corresponds to a value of the noise-to-signal ratio NSR ≈ 2.67. This is indeed confirmed by Pearson's χ^2 test, where we found that the limit for confidently rejecting the null hypothesis— certain distinguishability—is $a \approx 1$ (i.e., NSR ≈ 2.67). These results prove that $\mathbf{Z}^4$ is indeed an extremely robust feature with respect to measurement noise contamination, hence useful for applications.

V. PRINCIPAL COMPONENT ANALYSIS

According to the last sections, we can conclude that the HVG $\mathbf{Z}^4$ is an informative feature of complex dynamics. Here we summarize and gather the findings on i.i.d., fully chaotic logistic maps (with and without noise contamination), and colored noise, and we complement those with additional chaotic maps (Ricker's map, cubic map, sine map). Each process is described by the six-dimensional vector $\mathbf{Z}^4$ (although in practice this space is 5-dimensional as $\mathbb{P}_2^4 = 0$). As this representation is obviously not very convenient for readability, we have projected each point into a 2-dimensional space spanned by the principal components of the data. We recall that principal component analysis (PCA) [33] is a common statistical procedure to perform dimensionality reduction on data. It uses an orthogonal transformation to project our set of observations, originally described in $\mathbb{R}^6$—where each direction describes the probability of occurrence of a given

motif, this being possibly correlated among observations— into a lower dimensional subspace spanned by the so-called principal components, obtained from the eigenvectors of the data set covariance matrix. These particular directions are such that (i) they are orthogonal, (ii) the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (i.e., uncorrelated with) the preceding components. If the data can be efficiently projected in a lower dimensional space, then the eigenvalues associated with each of the principal components sum up a large percentage of the data variability. In that case, the projection is said to be faithful, and constitutes an accurate description of the data.

To summarize, the following processes have been considered (for all of them, we have estimated $\mathbf{Z}^4$ from a time series of $N = 10^4$ points, and have averaged this over 100 realizations):

(1) White noise (i.i.d.) with Gaussian, exponential, uniform, and power-law probability densities.

(2) Chaotic maps—in particular, the fully chaotic logistic map $x_{t+1} = 4x_t(1 - x_t)$, Ricker's map $x_{t+1} = 20x_t e^{-x_t}$, the cubic map $x_{t+1} = 3x_t(1 - x_t^2)$, and the sine map $x_{t+1} = \sin(\pi x_t)$.

(3) The noisy logistic map with $a = \{0.2, 0.4, 0.6, 0.8, 1.0\}$.

(4) Colored noise for $r = \{0.1, 0.3, 0.5, 0.7, 0.9, 0.99\}$.

The projection into the space spanned by the first two principal components is shown in Fig. 8. Interestingly, these first two components capture about 98.3% of the variability of the set of variables $\{\mathbf{Z}^4\}$. This means that motif probabilities are indeed highly correlated, and as few as two real numbers per time series seem already enough to describe them. The patterns related to the different processes in this 2-dimensional component space help visualize some of the results previously found and make interesting considerations:
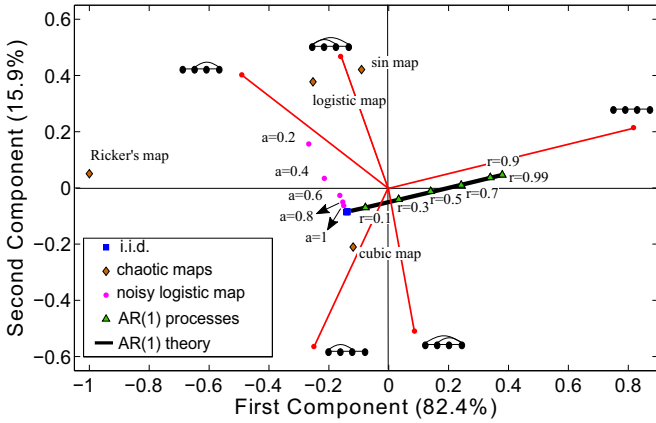
FIG. 8. Two-dimensional projection obtained via principal component analysis on $\mathbf{Z}^4$ for time series generated from different deterministic and stochastic processes: different white noise series respectively with Gaussian, exponential, uniform, and power law (blue squares), chaotic maps (brown diamonds), noisy logistic map for different levels of contamination (purple dots), and different stochastic correlated AR(1) processes (green triangles). The relative weight of each motif in this projection principal component analysis is also plotted using red solid axes.

(1) All the i.i.d. processes have the same coordinates in the 2-dimensional space which do not correspond to the coordinates of any other class of processes considered. Indeed according to the theory, i.i.d. processes share the same $\mathbf{Z}^4$.

(2) Red solid axes describe the projection of each motif in this new basis (see also Table III) and give an idea of which motif types are more related to different processes, thus helping to interpret a particular trajectory in this space, as a given process changes. For instance colored noise which interpolates between white noise ($r \to 0$) and a constant series ($r \to 1$) projects into a straight linelike trajectory, departing from the i.i.d. coordinates and following the direction where type-I motif increases as r increases. Analogously, as the noise level a increases the noisy logistic map interpolates between the fully chaotic logistic coordinate and i.i.d. following a specific path.

(3) The distance in this space between i.i.d. and the ($a = 1$)-noisy logistic map gives us a rough idea of the distinguishability or coarse-graining distance, a lower bound below which any two processes cannot be distinguished.

TABLE III. Weights of each motif in the 2-dimensional projection of the set of all dynamical processes analyzed (i.i.d., white noise, colored noise with exponentially decaying correlations, chaotic maps, noisy chaotic logistic map).

	First component	Second component
	0.814	0.2114
	-8.6×10^{-18}	1.5×10^{-16}
	-0.2506	-0.5670
	-0.4926	0.4030
	-0.1569	0.4608
	0.0852	-0.5090

We conclude that $\mathbf{Z}^4$ is a highly informative and robust feature, which in principle could be used to assess similarities and differences across empirical complex signals. To test this hypothesis, in the final section we will explore this idea and will show that clustering of complex physiological processes is possible with this simple feature.

VI. UNSUPERVISED LEARNING: DISENTANGLING MEDITATIVE FROM OTHER RELAXATION STATES USING HVG MOTIF PROFILES FROM HEART RATE TIME SERIES

It is well known that meditation has a measurable effect on well-being. In particular, neuroscience has shown that meditation promotes EEG high-amplitude gamma synchronization [34], or increases sustained attention [35] among others effects on the brain [36]. In this final section we explore, via an HVG motif profile analysis, whether one can distinguish purely meditative states from general states of relaxation by only looking at a single physiological indicator: the heart rate series [37,38]. This analysis is based on experiments performed in a former publication [39]. Data are freely available online [40].

A. Data

Data are collected for five different groups of healthy subjects [39].

(1) The first group of 4 subjects (two women and two men in the age range 20–52) were expert Kundalini yoga meditators. Their heart rate was recorded for approximately fifteen minutes before the yoga practice (premeditative state) and for approximately one hour during the breathing and chanting exercises (meditative state) (a total of 8 time series).

(2) The second group comprised 8 Chinese chi meditation practitioners (five women and three men in the age range 26–35) relatively novice in the practice. The heart rate of the subjects was recorded for approximately five hours during the premeditation (premeditative state) and for approximately one hour during the meditation session (meditative state) (a total of 16 time series).

To better compare the premeditation and meditation states, three healthy, nonmeditating control groups were considered from a database of retrospective electrocardiogram (ECG) signals:

(3) a spontaneous breathing group of 13 subjects (eight women and five men in the age range 25–35) during sleeping hours (general relaxation state) (a total of 13 time series);

(4) a group of 9 elite triathlon athletes (six women and three men, age range 21–55) in the prerace period during sleeping hours (general relaxation state) (a total of 9 time series);

(5) a group of 14 subjects (nine women and five men, age range 20–35) during supine metronomic breathing at 0.25 Hz (a total of 14 time series).

Sample time series from each group are plotted in Fig. 9. In the original study the authors addressed the frequency spectra and observed prominent heart rate oscillations in the time series recorded during the two meditation practices with a peak in the range 0.025–0.35 Hz, and an overall variability of these series with respect to those from nonmeditative states.
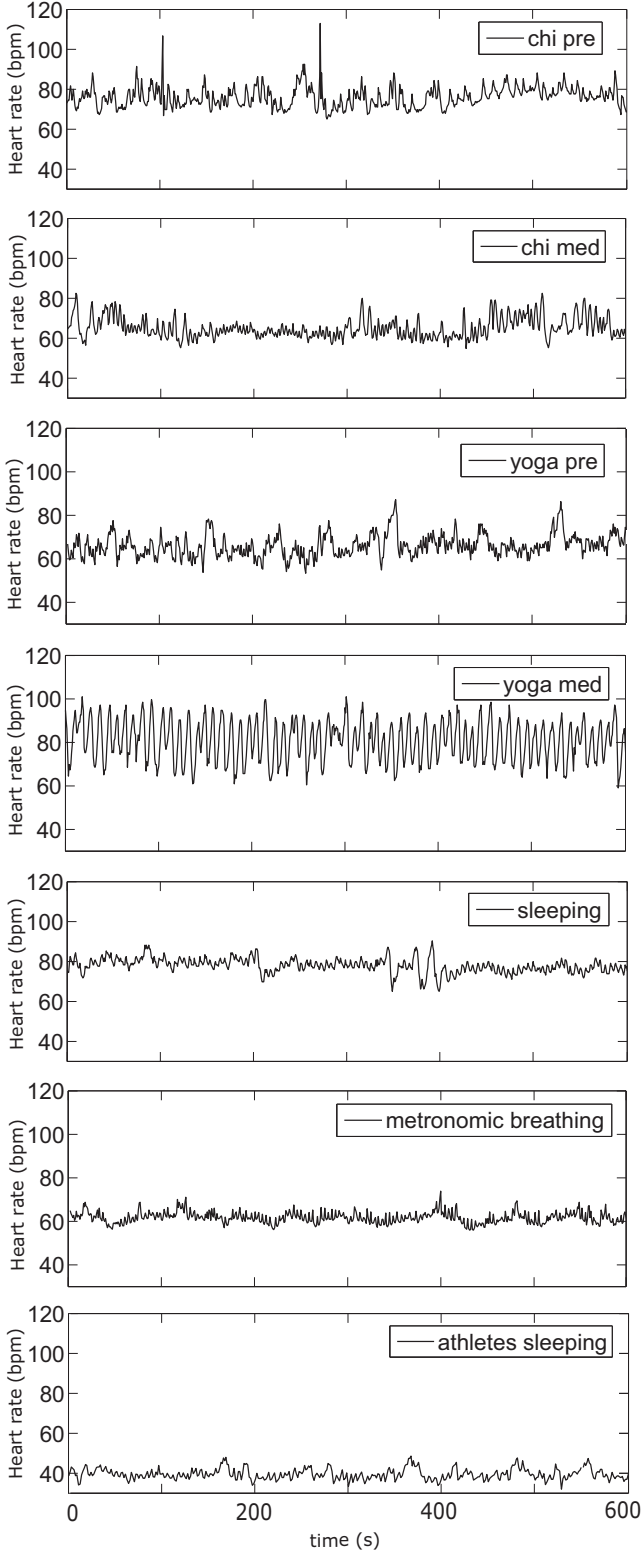
FIG. 9. Sample heart rate time series from patients in meditative and nonmeditative states.

B. Unsupervised clustering based on HVG motif profiles

The total data set is made of a total of 60 time series (60 observations). *A priori*, we assume that each series is a different process. For each subject and state, we extract from the heart

beat series the corresponding $\mathbf{Z}^4$ (detailed results are put in Appendix B).

As a first analysis, we only consider the expert meditators (first group) performing two different tasks and we explore whether $\mathbf{Z}^4$ can disentangle the two tasks. Results are shown in panel (a) of Fig. 10. In PCA space, we have 8 points scattered over the subspace spanned by the first two principal components. These aggregate more than 99% of the data variance and are thus a faithful projection. Interestingly, already a visual inspection clusters the 4 subjects in the meditative state (red circles, right-hand side of the plane) from those in the premeditative state (green squares). A simple k-means algorithm [33] with $k = 2$ correctly distinguishes the two states by assigning different clusters to both states (a black dotted oval is depicted with the purpose of visualizing the result of the k-means clustering).

In a second step, we consider the second group, formed now by novice chi meditators before and during the practice. We repeat the analysis in the panel (b) of Fig. 10. Again the first two principal components capture more than 99% of the variability of the motifs considered. The scores related to the first principal component are very close to the ones found for the yoga data subset (see Appendix B). For this meditation technique however it is not that easy to perfectly distinguish premeditative from meditative state clusters: the partition obtained with the k-means algorithm with input $k = 2$ (visualized by the black dotted line) contains "false meditators" and "false nonmeditators." In order to quantify the performance of the clustering we use the so-called *purity coefficient* [41] defined by

$$\text{purity} = \frac{n_m^1 + n_n^0}{n_m^1 + n_m^0 + n_n^0 + n_n^1}, \qquad (14)$$

where n_m^1 is the number of meditators in cluster 1, which is defined as the cluster where most of the meditators are found; n_m^0 is the number of meditators in cluster 0, which is defined as the cluster where most of the nonmeditators are found ("false nonmeditators"); n_n^1 is the number of nonmeditators in cluster 1 ("false meditators"); n_n^0 is the number of nonmeditators in cluster 0. Purity takes value in [0,1] and was measured for the different partitions reported in Fig. 10 (see Table IV); in this case we found purity $\simeq 0.83$. Now, as in this experiment the subjects were inexperienced chi meditators, it is plausible that some of them were not able to concentrate or perform the task adequately, which would put their motif profile mixed among the premeditative state subjects. As we can see in the figure, there is some evidence of finding the "false nonmeditators" intertwined among nonmeditators, but not the "false meditators" intertwined among meditators.

We then perform the same analysis by considering data from the first two groups (yoga group and the chi group) altogether. Here we also aim at distinguishing meditative from premeditative states; however this is in principle much more delicate and problematic as we have different subjects performing different tasks. The results are reported in panel (c) of Fig. 10, and are consistent with the first two analyses conducted before. In PCA space, the first two principal components still capture more than 99% of the data variability (scores are reported in Appendix B). The k-means algorithm
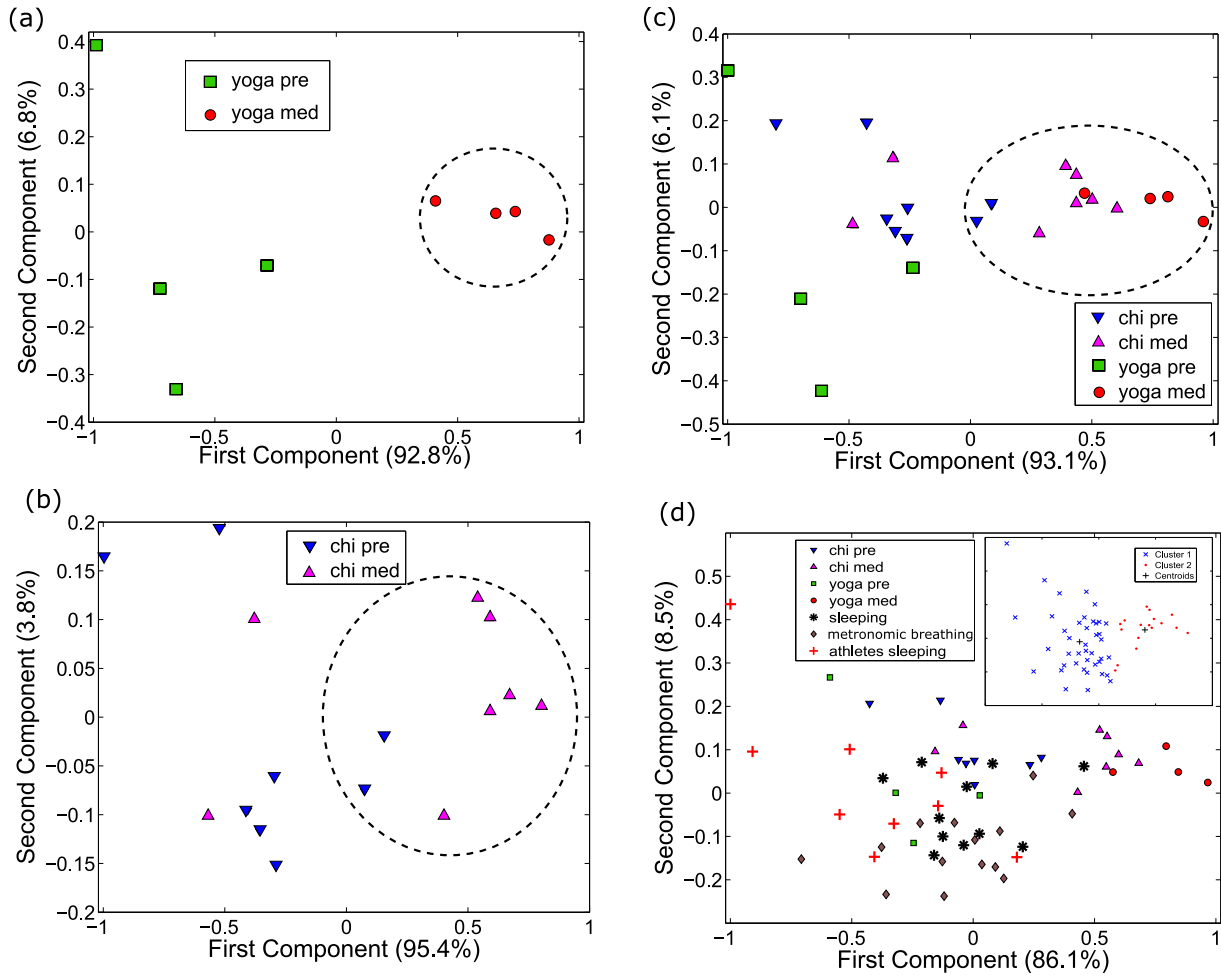
FIG. 10. Two-dimensional principal component space of $\mathbf{Z}^4$ extracted from heart rate time series of subjects performing different tasks. (a) Four yoga meditators recorded during meditation (red dots) and during premeditation (green squares). The k-means algorithm (black dotted line) correctly assigns each of the 8 observations to the correct cluster. (b) Eight chi meditators recorded during meditation (magenta triangles) and during premeditation (blue reverse triangles). The k-means algorithm correctly assigns 12 out of 16 observations; however in this case subjects were novice meditators, hence clusters are not that well defined (see the text). (c) The two clusters found by the k-means algorithm correctly clusters the points related to yoga meditation and yoga premeditation, and 12 out of 16 points related to different chi meditators (black dotted line). (d) Although the k-means clustering (small panel on the top right) fails to precisely distinguish a cluster related to meditation from a cluster related to nonmeditation, all the meditation points are surprisingly well separated from all the remaining points, on the right side of the plane.

correctly clusters together most of the premeditative states and distinguishes them from the meditative states (yoga and chi style), with purity = 0.75. There are two clear "false nonmeditators" which seem to correspond to two novice chi meditators that falsely fall in the nonmeditation state despite that they were supposedly performing meditation. The two "false meditators" are not mixed among the meditators but placed in the boundary of the cluster, meaning that a refined clustering algorithm would very likely do a better job. On the other hand, it is worth highlighting that meditators show lower scattering than nonmeditators, and are placed at the right-hand side of the plane. Among these, chi meditators (the experienced subjects) appear even more towards the right-hand side in the PCA plane. According to the motif scores (Appendix B), one can conclude that meditation promotes the onset of type-I motifs, that is to say, generates a relative decrease of high-frequency heart rate fluctuations.

Finally, in panel (d) of Fig. 10 we show the results for the analysis of the whole data set (the projection in PCA space still gathers more than 94% of the data variability). Here we have highly heterogeneous subjects performing totally different tasks, which somehow can be classified into "meditative" and "nonmeditative" states. In the inset panel of the same figure, each observation is labeled according to the result of the k-means algorithm (crosses for nonmeditative and dots for meditative states). Despite the heterogeneity of subjects, the purity of the partition obtained is high ($\simeq 0.81$), and most of the observations associated with the meditative state concentrate towards the right-hand side of the PCA plane (which, again according to the scores, corresponds to an overcontribution of the type-I motif). We conclude that meditative practices leave a unique physiological fingerprint in the heart rate time series of its practitioners, which can be distinguished from other relaxation techniques and states such as metronomic breathing

TABLE IV. Purity measures [41] of the *k*-means clustering analysis depicted in the four panels of Fig. 10: yoga meditators [panel (a)], chi meditators [panel (b)], yoga and chi meditators [panel (c)], and all states [panel (d)].

Panel (a)	Panel (b)	Panel (c)	Panel (d)
1	0.83	0.75	0.81

or sleeping by using the HVG motif profile of each time series. This is a remarkable result, taking into account that this profile only consists of a vector of 6 numbers (actually 5 as $\mathbb{P}_2^4 = 0$) per observation.

VII. CONCLUSIONS

The theory of visibility graphs (VG and HVG) allows us to describe and characterize time series and dynamics using the powerful machinery of graph theory and network science. Here we have introduced the concept of horizontal visibility graph (HVG) motifs, substructures present in the HVG of a time series, whose statistics have been shown to be informative about the time series structure and its underlying dynamics (comparison with VG motifs will be published elsewhere [31]). We have advanced a mathematically sound theory by which the motif profile of large classes of stochastic and deterministic dynamics can be computed exactly. Interestingly, under the HVG framework, graph motifs are in direct correspondence with ordinal patterns [27–30]. This means, for instance, that the theory developed here can be exported to find rigorous results on the permutation entropy [28] and permutation spectra [29] of different dynamical systems. In the same vein, one could import concepts and ideas from ordinal patters to the context of visibility graphs. For instance, one can define an HVG motif entropy $S_n = -\frac{1}{n} \sum \mathbf{Z}_i^n \ln(\mathbf{Z}_i^n)$ and explore its similarities with permutation entropy. More generally, the relation (and possible equivalences) between ordinal pattern analysis (so-called permutation complexity [28,30]) and horizontal visibility graph analysis should be studied in more depth. We have found that this graph feature is surprisingly robust, in the sense that it is still able to distinguish among different dynamics even when the signals are polluted with large amounts of measurement noise, which enables its use in practical problems. Despite the apparently difficult combinatorial interpretation of the visibility criteria, these latter results further suggest that HVG motifs are more than just an arbitrary partition on the set of ordinal patterns. As an application, we have tackled the problem of disentangling meditative from general relaxation states from the HVG motif profiles of heartbeat time series of different subjects performing different tasks. We have been able to provide a positive, unsupervised solution to this question by applying standard clustering algorithms on this simple feature.

To conclude, HVG motifs provide a mathematically sound, computationally efficient, and highly informative simple feature (a few numbers per time series) which can be extracted from any kind of time series and used to describe complex signals and dynamics from a new viewpoint. In direct analogy with the role played by standard motifs in biological networks,

further work should evaluate whether HVG graph motifs can be seen as the building blocks of time series. In this sense a study of standard network motifs on visibility graphs can be of interest, especially in the case of directed and weighted HVGs where the edge weights describe temporal relations between nodes. Potential applications of visibility-graph analysis pervades the biological, financial, and physical sciences. Finally, other questions for future work include to assess which motifs are more informative for a given class of dynamics, and to extend this analysis to the realm of multivariate time series [17].

APPENDIX A: EXPLICIT COMPUTATION OF Z⁴ FOR THE FULLY CHAOTIC LOGISTIC MAP

(1) $\mathbb{P}_1^4$.

$$
\begin{aligned}
\mathbb{P}_1^4 &= \int_0^1 f(x_0)dx_0 \int_0^1 \delta(x_1 - \mathcal{H}(x_0))dx_1 \\
&\quad \times \int_0^{x_1} \delta(x_2 - \mathcal{H}^2(x_0))dx_2 \int_0^{x_2} \delta(x_3 - \mathcal{H}^3(x_0))dx_3 \\
&\quad + \int_0^1 f(x_0)dx_0 \int_{x_0}^1 \delta(x_1 - \mathcal{H}(x_0))dx_1 \\
&\quad \times \int_{x_1}^1 \delta(x_2 - \mathcal{H}^2(x_0))dx_2 \int_0^1 \delta(x_3 - \mathcal{H}^3(x_0))dx_3.
\end{aligned}
$$

The first integral on the right gives the following conditions:
$\mathcal{H}^3(x_0) < \mathcal{H}^2(x_0)$,
$\mathcal{H}^2(x_0) < \mathcal{H}(x_0)$,
which are never satisfied. The second integral gives
$\mathcal{H}^2(x_0) > \mathcal{H}(x_0)$,
$\mathcal{H}(x_0) > x_0$,
which are satisfied for $x_0 \in [0, 1/4]$. Thus

$$
\mathbb{P}_1^4 = \frac{1}{\pi} B_{[0, \frac{1}{4}]} \left(\frac{1}{2}, \frac{1}{2} \right) = \frac{1}{3} \quad (=8/24).
$$

(2) $\mathbb{P}_2^4 = 0$ since the probability of having $\mathcal{H}^2(x_0) = \mathcal{H}(x_0)$ is of zero measure.

(3) $\mathbb{P}_3^4$.

$$
\begin{aligned}
\mathbb{P}_3^4 &= \int_0^1 f(x_0)dx_0 \int_0^{x_0} \delta(x_1 - \mathcal{H}(x_0))dx_1 \\
&\quad \times \int_{x_1}^{x_0} \delta(x_2 - \mathcal{H}^2(x_0))dx_2 \int_0^{x_2} \delta(x_3 - \mathcal{H}^3(x_0))dx_3 \\
&\quad + \int_0^1 f(x_0)dx_0 \int_0^{x_0} \delta(x_1 - \mathcal{H}(x_0))dx_1 \\
&\quad \times \int_{x_0}^1 \delta(x_2 - \mathcal{H}^2(x_0))dx_2 \int_0^1 \delta(x_3 - \mathcal{H}^3(x_0))dx_3.
\end{aligned}
$$

In the first term,
$\mathcal{H}(x_0) < x_0 \Rightarrow x_0 \in [3/4, 1]$,

$\mathcal{H}^2(x_0) > \mathcal{H}(x_0) \cap \mathcal{H}^2(x_0) < x_0 \cap [3/4,1] \Rightarrow x_0 \in [\frac{5+\sqrt{5}}{8}, 1]$,

$\mathcal{H}^3(x_0) < \mathcal{H}^2(x_0) \cap [\frac{5+\sqrt{5}}{8}, 1] \Rightarrow x_0 \in [\frac{5+\sqrt{5}}{8}, \frac{1}{2} + \frac{\sqrt{3}}{4}]$.

Analogously for the second term,

$\mathcal{H}(x_0) < x_0 \Rightarrow x_0 \in [3/4, 1]$,

$\mathcal{H}^2(x_0) > x_0 \cap [3/4, 1] \Rightarrow x_0 \in [3/4, \frac{5+\sqrt{5}}{8}]$.

Altogether,

$$\mathbb{P}_3^4 = \frac{1}{\pi} B_{[3/4, \frac{1}{2} + \frac{\sqrt{3}}{4}]}(1/2, 1/2) = \frac{1}{6} (=4/24).$$

(4) $\mathbb{P}_4^4$.

$$\mathbb{P}_4^4 = \int_0^1 f(x_0) dx_0 \int_{x_0}^1 \delta(x_1 - \mathcal{H}(x_0)) dx_1$$

$$\times \int_0^{x_1} \delta(x_2 - \mathcal{H}^2(x_0)) dx_2 \int_{x_2}^1 \delta(x_3 - \mathcal{H}^3(x_0)) dx_3$$

$$+ \int_0^1 f(x_0) dx_0 \int_0^{x_0} \delta(x_1 - \mathcal{H}(x_0)) dx_1$$

$$\times \int_0^{x_1} \delta(x_2 - \mathcal{H}^2(x_0)) dx_2 \int_{x_2}^{x_1} \delta(x_3 - \mathcal{H}^3(x_0)) dx_3.$$

The first integral on the right gives the following conditions:

$\mathcal{H}^3(x_0) > \mathcal{H}^2(x_0)$,

$\mathcal{H}^2(x_0) < \mathcal{H}(x_0)$,

$\mathcal{H}(x_0) > x_0$,

which are satisfied for $x_0 \in [1/2, 3/4]$. The second integral gives

$\mathcal{H}^2(x_0) < \mathcal{H}^3(x_0) < \mathcal{H}(x_0)$,

$\mathcal{H}^2(x_0) < \mathcal{H}(x_0)$,

$\mathcal{H}(x_0) < x_0$,

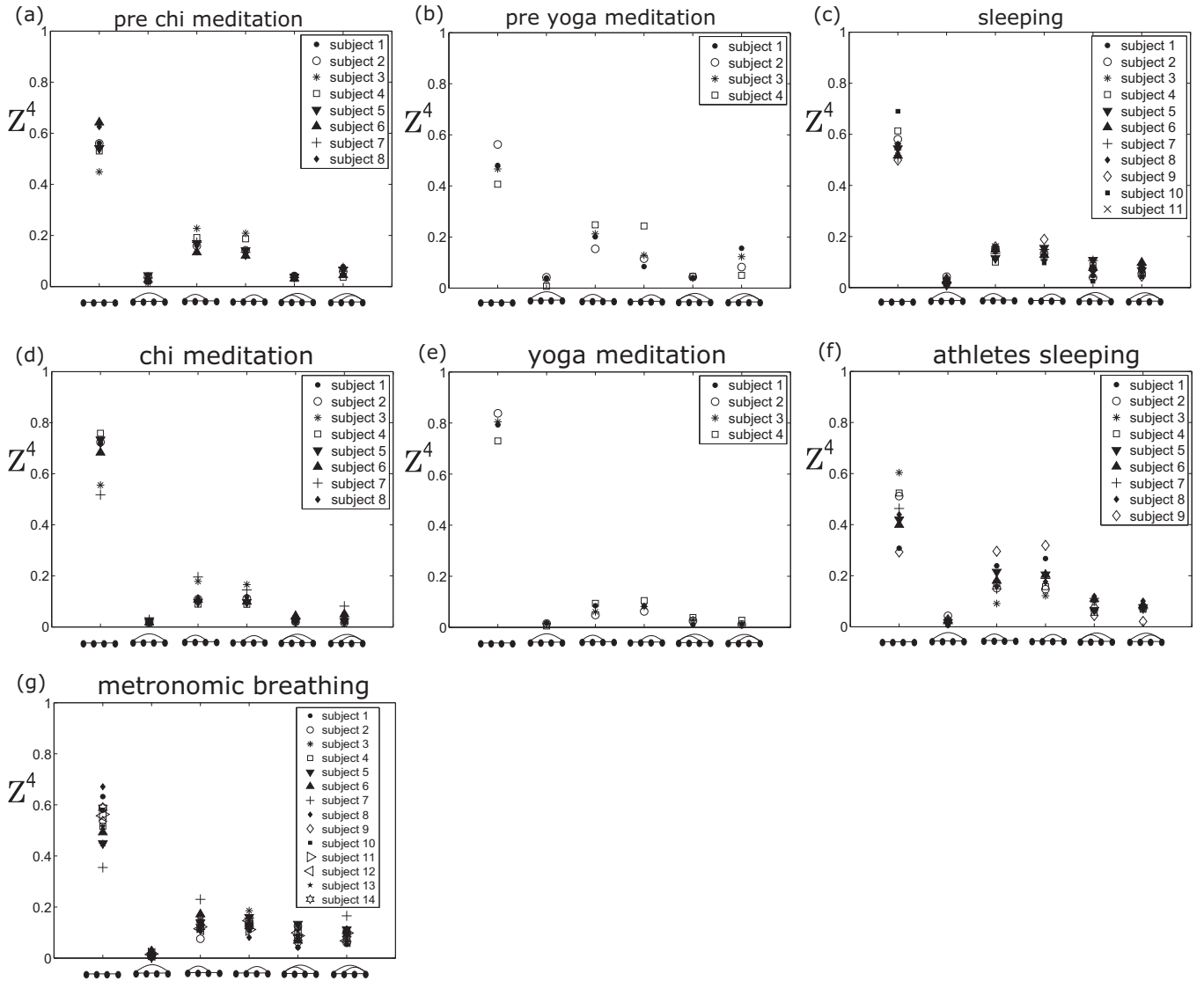FIG. 11. HVG motif significance profile $\mathbf{Z}^4$ obtained by analyzing heart rate time series form different groups of subjects in different states: (a) 8 chi meditators before the meditation practice; (b) 4 yoga meditators before the meditation practice; (c) 11 subjects during sleeping; (d) same 8 chi meditators of (a) during the meditation practice; (e) same 4 yoga meditators of (b) during the meditation practice; (f) 9 elite athletes during sleeping; (g) 14 subjects during metronomic breathing at 0.25 Hz.

TABLE V. Principal component scores obtained from PCA considering the yoga meditators data subset (left) and the chi meditators data subset (right).

	Yoga		Chi	
	First component	Second component	First component	Second component
●━●━●━●	0.874	0.0346	0.871	0.171
●⌒●⌒●⌒●	−0.029	−0.203	−0.023	−0.239
●⌒●⌒●━●	−0.379	0.074	−0.376	0.207
●━●⌒●━●	−0.204	0.731	−0.272	0.666
●⌒●━●━●	−0.043	0.01	−0.048	−0.173
●⌒●━●━●	−0.219	−0.647	−0.152	−0.631

TABLE VI. Principal component scores obtained from PCA considering the subset data of yoga and chi meditators together (left) and considering all the data set (right).

	Chi and Yoga		All States	
	First component	Second component	First component	Second component
●━●━●━●	0.874	0.08	0.881	0.133
●⌒●⌒●⌒●	−0.027	−0.181	0.007	0.073
●⌒●⌒●━●	−0.378	0.089	−0.315	0.477
●━●⌒●━●	−0.235	0.714	−0.294	0.4
●⌒●━●━●	−0.045	−0.039	−0.13	−0.58
●⌒●━●━●	−0.188	−0.664	−0.15	−0.503

which are satisfied for $x_0 \in [1/2 + \sqrt{3}/4, 1]$. Thus

$$\mathbb{P}_4^4 = \frac{1}{\pi}\left[B_{\left[\frac{1}{2},\frac{3}{4}\right]}\left(\frac{1}{2},\frac{1}{2}\right) + B_{\left[\frac{1}{2}+\frac{\sqrt{3}}{4},1\right]}\left(\frac{1}{2},\frac{1}{2}\right) \right] = 8/24.$$

(5) $\mathbb{P}_5^4$.

$$\mathbb{P}_5^4 = \int_0^1 f(x_0)dx_0 \int_0^{x_0} \delta(x_1 - \mathcal{H}(x_0))dx_1$$
$$\times \int_{x_1}^{x_0} \delta(x_2 - \mathcal{H}^2(x_0))dx_2 \int_{x_2}^1 \delta(x_3 - \mathcal{H}^3(x_0))dx_3$$

gives the following conditions:

$\mathcal{H}^3(x_0) > \mathcal{H}^2(x_0)$,
$\mathcal{H}(x_0) < \mathcal{H}^2(x_0) < x_0$,

which are satisfied for $x_0 \in [1/4 + \sqrt{3}/4, 1]$ and

$$\mathbb{P}_5^4 = \frac{1}{\pi} B_{\left[\frac{1}{4}+\frac{\sqrt{3}}{4},1\right]}\left(\frac{1}{2},\frac{1}{2}\right) = \frac{1}{6} \quad (=4/24).$$

(6) $\mathbb{P}_6^4$.

$$\mathbb{P}_6^4 = \int_0^1 f(x_0)dx_0 \int_0^{x_0} \delta(x_1 - \mathcal{H}(x_0))dx_1$$
$$\times \int_0^{x_1} \delta(x_2 - \mathcal{H}^2(x_0))dx_2 \int_{x_1}^1 \delta(x_3 - \mathcal{H}^3(x_0))dx_3$$

gives the following conditions:

$\mathcal{H}^3(x_0) > \mathcal{H}(x_0) > \mathcal{H}^2(x_0)$,
$\mathcal{H}(x_0) < x_0$,

which are never satisfied for the $\mathcal{H}(x)$ map [this is indeed based on the fact that the pattern $x_i > x_{i+1} < x_{i+2}$ is indeed a forbidden pattern in the orbit of $\mathcal{H}(x)$]. Hence

$$\mathbb{P}_6^4 = 0.$$

APPENDIX B: MOTIF PROFILES FOR ALL SUBJECTS IN THE EMPIRICAL STUDY

In Fig. 11 we give an overview of the 4-node motif profiles, measured for the different subjects in the different states. Interestingly, the motif that shows more variability in each of the given states is the one related to the type-I motif, which we have seen to play a minor role in the case of the chaotic dynamics polluted with noise.

1. Scores

The scores of the two components in terms of motifs are reported in Tables V and VI, and as expected the highest contribution to the first component (0.874) is given by the motif of type I.

[1] R. Albert and A.-L. Barabasi, Statistical mechanics of complex networks, Rev. Mod. Phys. **74**, 47 (2002).

[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. **424**, 175 (2006).

[3] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).

[4] M. E. J. Newman, The structure and function of complex networks, SIAM Rev **45**, 167 (2003).

[5] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Network motifs: Simple building blocks of complex networks, Science **298**, 824 (2002).

[6] U. Alon, Network motifs: Theory and experimental approaches, Nat. Rev. Genet. **8**, 450 (2007).

[7] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, Superfamilies of evolved and designed networks, Science **303**, 1538 (2004).

[8] A. Masoudi-Nejad, F. Schreiber, and Z. R. M. Kashani, Building blocks of biological networks: A review on major network motif discovery algorithms, IET Systems Biology **6**, 164 (2012).

[9] J. Zhang and M. Small, Complex Network from Pseudoperiodic Time Series: Topology versus Dynamics, Phys. Rev. Lett. **96**, 238701 (2006).

[10] F. Kyriakopoulos and S. Thurner, Directed network representations of discrete dynamical maps, Lect. Notes Comput. Sci. **4488**, 625 (2007).

[11] X. Xu and J. Zhang, Small M superfamily phenomena and motifs of networks induced from time series, Proc. Natl. Acad. Sci. USA **105**, 19601 (2008).

[12] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths, Recurrence networks: A novel paradigm for nonlinear time series analysis, New J. Phys. **12**, 033025 (2010).

[13] R. V. Donner *et al.*, Geometry of chaotic dynamics: A complex network perspective, Eur. Phys. J. B **84**, 653 (2011).

[14] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, From time series to complex networks: The visibility graph, Proc. Natl. Acad. Sci. USA **105**, 4972 (2008).

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

[16] L. Lacasa, A. Nuñez, E. Roldan, J. M. R. Parrondo, and B. Luque, Time series irreversibility: A visibility graph approach, Eur. Phys. J. B **85**, 217 (2012).

[17] L. Lacasa, V. Nicosia, and V. Latora, Network structure of multivariate time series, Sci. Rep. **5**, 15508 (2015).

[18] L. Lacasa, B. Luque, J. Luque, and J. C. Nuño, The visibility graph: A new method for estimating the Hurst exponent of fractional Brownian motion, Europhys. Lett. **86**, 30001 (2009).

[19] L. Lacasa and R. Flanagan, Time reversibility of visibility from nonstationary processes, Phys. Rev. E **92**, 022817 (2015).

[20] R. Flanagan and L. Lacasa, Irreversibility of financial time series: A graph-theoretical approach, Phys. Lett. A **380**, 1689 (2016).

[21] S. Severini, G. Gutin, and T. Mansour, A characterization of horizontal visibility graphs and combinatorics on words, Physica A **390**, 2421 (2011).

[22] L. Lacasa, On the degree distribution of horizontal visibility graphs associated with Markov processes and dynamical systems: Diagrammatic and variational approaches, Nonlinearity **27**, 2063 (2014).

[23] B. Luque, L. Lacasa, F. J. Ballesteros, and A. Robledo, Feigenbaum graphs: A complex network perspective of chaos, PLoS ONE **6**, e22411 (2011).

[24] A. M. Nunez, B. Luque, L. Lacasa, J. P. Gomez, and A. Robledo, Horizontal visibility graphs generated by type-I intermittency, Phys. Rev. E **87**, 052801 (2013).

[25] A. M. Nunez, L. Lacasa, and J. P. Gomez, Horizontal visibility graphs generated by type-II intermittency, J. Phys. A: Math. Theor. **47**, 035102 (2014).

[26] B. Luque, L. Lacasa, F. Ballesteros, and A. Robledo, Analytical properties of horizontal visibility graphs in the Feigenbaum scenario, Chaos **22**, 013109 (2012).

[27] C. Bandt and B. Pompe, Permutation Entropy: A Natural Complexity Measure for Time Series, Phys. Rev. Lett. **88**, 174102 (2002).

[28] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, Permutation entropy and its main biomedical and econophysics applications: A review, Entropy **14**, 1553 (2012).

[29] C. W. Kulp and L. Zunino, Discriminating chaotic and stochastic dynamics through the permutation spectrum test, Chaos **24**, 033116 (2014).

[30] J. Amigo, *Permutation Complexity in Dynamical Systems: Ordinal Patterns, Permutation Entropy and All That* (Springer, Berlin, 2010).

[31] J. Iacovacci and L. Lacasa (unpublished).

[32] See http://www.wolfram.com/mathematica.

[33] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006).

[34] A. Lutz, L. L. Greischar, N. B. Rawlings, M. Ricard, and R. J. Davidson, Long-term meditators self-induce high-amplitude gamma synchrony during mental practice, Proc. Natl. Acad. Sci. USA **101**, 16369 (2004).

[35] A. Lutz, H. A. Slagter, N. B. Rawlings, A. D. Francis, L. L. Greischar, and R. J. Davidson, Mental training enhances attentional stability: Neural and behavioral evidence, J. Neurosci. **29**, 13418 (2009).

[36] Yi-Yuan Tang, B. K. Holzel, and M. I. Posner, The neuroscience of mindfulness meditation, Nat. Rev. Neurosci. **16**, 213 (2015).

[37] S. Jiang, C. Bian, X. Ning and Q. D. Y. Ma, Visibility graph analysis on heartbeat dynamics of meditation training, Appl. Phys. Lett. **102**, 253702 (2013).

[38] A. Sarkar and P. Barat, Effect of meditation on scaling behavior and complexity of human heart rate variability, Fractals **16**, 199 (2008).

[39] C.-K. Peng, J. E. Mietus, Y. Liu, G. Khalsa, P. S. Douglas, H. Benson, and A. L. Goldberger, Exaggerated heart rate oscillations during two meditation techniques, Int. J. Cardiol. **70**, 101 (1999).

[40] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, Circulation **101**, 2909 (2000).

[41] K. Hyunsoo and H. Park, Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis, Bioinformatics **23**, 1495 (2007).

[42] According to numerical experiments, HVG has linear complexity on aperiodic dynamics (http://www.maths.qmul.ac.uk/~lacasa/Software.html).