

Time series irreversibility: a visibility graph approach

L. Lacasa^{1,a}, A. Nuñez¹, É. Roldán², J.M.R. Parrondo², and B. Luque¹

¹ Departamento de Matemática Aplicada y Estadística, ETSI Aeronáuticos, Universidad Politécnica de Madrid, 28040 Madrid, Spain

² Departamento de Física Atómica, Molecular y Nuclear and *GLSC*, Universidad Complutense de Madrid, 28040 Madrid, Spain

Abstract. We propose a method to measure real-valued time series irreversibility which combines two different tools: the horizontal visibility algorithm and the Kullback-Leibler divergence. This method maps a time series to a directed network according to a geometric criterion. The degree of irreversibility of the series is then estimated by the Kullback-Leibler divergence (i.e. the distinguishability) between the *in* and *out* degree distributions of the associated graph. The method is computationally efficient and does not require any ad hoc symbolization process. We find that the method correctly distinguishes between reversible and irreversible stationary time series, including analytical and numerical studies of its performance for: (i) reversible stochastic processes (uncorrelated and Gaussian linearly correlated), (ii) irreversible stochastic processes (a discrete flashing ratchet in an asymmetric potential), (iii) reversible (conservative) and irreversible (dissipative) chaotic maps, and (iv) dissipative chaotic maps in the presence of noise. Two alternative graph functionals, the degree and the degree-degree distributions, can be used as the Kullback-Leibler divergence argument. The former is simpler and more intuitive and can be used as a benchmark, but in the case of an irreversible process with null net current, the degree-degree distribution has to be considered to identify the irreversible nature of the series.

1 Introduction

A stationary process $X(t)$ is said to be statistically time reversible (hereafter time reversible) if for every N , the series $\{X(t_1), \dots, X(t_N)\}$ and $\{X(t_N), \dots, X(t_1)\}$ have the same joint probability distributions [1]. 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 [2,3]. 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 [2,3], recent work uses the concept of time series irreversibility to derive information about the entropy production of the physical mechanism generating the series, even if one ignores any detail of such mechanism [4,5]. 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 [6–8], render-

ing 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. Recently, several methods to measure time irreversibility have been proposed [6,7,9–15]. The majority of them perform a time series symbolization, typically making an empirical partition of the data range [9] (note that such a transformation does not alter the reversible character of the output series [10]) and subsequently analyze the symbolized series, through statistical comparison of symbol strings occurrence in the forward and backwards series or using compression algorithms [5,10,16]. 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 [7,8].

Motivated by these facts, here we explore the usefulness of the horizontal visibility algorithm in such context. This is a time series analysis method which was proposed recently [17]. It makes use of graph theoretical concepts,

^a e-mail: lucas_lacasa@yahoo.es

and it is based on the mapping of a time series to a graph and the subsequent analysis of the associated graph properties [17–20]. Here we propose a *time directed* version of the horizontal visibility algorithm, and we show that it is a simple and well defined tool for measuring time series irreversibility. More precisely, we show that the Kullback-Leibler divergence [16] between the *out* and *in* degree distributions, $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$, is a simple measure of the irreversibility of real-valued stationary stochastic series. Analytical and numerical results support our claims, and the presentation is as follows: the method is introduced in Section 2. In Section 3 we analyze reversible time series generated from linear stochastic processes, which yield $D[P_{\text{out}}(k)||P_{\text{in}}(k)] = 0$. As a further validation, in Section 4 we report the results obtained for irreversible series. We first analyze a thermodynamic system (a discrete flashing ratchet) which shows time irreversibility when driven out of equilibrium. Its amount of irreversibility can be increased continuously tuning the value of a parameter of the system, and we find that the method can, not only distinguish, but also quantify the degree of irreversibility. We also study the effect of applying a stalling force in the opposite direction of the net current of particles in the ratchet. In this case the benchmark measure fails predicting reversibility whereas a generalized measure based on degree-degree distributions $D[P_{\text{out}}(k, k')||P_{\text{in}}(k, k')]$ goes beyond the phenomenon associated to physical currents and still detects irreversibility. We extend this analysis to chaotic signals, where our method distinguishes between dissipative and conservative chaos, and we analyze chaotic signals polluted with noise. Finally, a discussion is presented in Section 5.

2 The method

2.1 The horizontal visibility graph

The family of visibility algorithms is a collection of methods that map series to networks according to specific geometric criteria [17,18]. The general purpose of such methods is to accurately map the information stored in a time series into an alternative mathematical structure, so that the powerful tools of graph theory may eventually be employed to characterize time series from a different viewpoint, bridging the gap between nonlinear time series analysis, dynamical systems, and graph theory [19,21–24].

We focus here on a specific subclass called horizontal visibility algorithm, firstly proposed in [17], and defined as follows: let $\{x_t\}_{t=1,\dots,N}$ be a real-valued time series of N data. The algorithm assigns each datum of the series to a node in the horizontal visibility graph (HVg). Then, 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 Fig. 1). Hence, i and j are two connected nodes if the following geometrical criterion is fulfilled within the time series:

$$x_i, x_j > x_n, \quad \forall n \mid i < n < j. \quad (1)$$

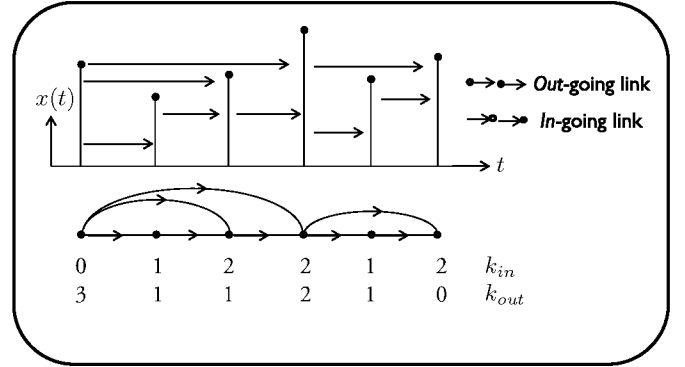


Fig. 1. Graphical illustration of the method. In the top we plot a sample time series $\{x(t)\}$. Each datum in the series is mapped to a node in the graph. Arrows, describing allowed directed visibility, link nodes. The associated directed horizontal visibility graph is plotted below. In this graph, 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.

Some results regarding the characterization of stochastic and chaotic series through this method have been put forward recently [17,20], and the first steps for a mathematically sound characterization of horizontal visibility graphs have been established [25]. Interestingly, a very recent work suggests that the method can be used in practice to characterize not only time series but generic nonlinear discrete dynamical systems, sharing similarities with the theory of symbolic dynamics [21].

2.2 Directed HVg

So far in the literature the family of visibility graphs are undirected, as visibility did not have a predefined temporal arrow. However, as conjectured in a previous work [18], such a directionality can be made explicit by making use of directed networks or digraphs [26]. We address such directed version, defining a *Directed Horizontal Visibility graph* (DHVg) as a HVg, where the degree $k(t)$ of the node t is now split in an *ingoing* degree $k_{\text{in}}(t)$, and an *outgoing* degree, such that $k(t) = k_{\text{in}}(t) + k_{\text{out}}(t)$. The ingoing degree $k_{\text{in}}(t)$ is defined as the number of links of node t with other *past* nodes associated with data in the series (that is, nodes with $t' < t$). Conversely, the outgoing degree $k_{\text{out}}(t)$, is defined as the number of links with *future* nodes.

For a graphical illustration of the method, see Figure 1. The degree distribution of a graph describes the probability of an arbitrary node to have degree k (i.e. k links) [26]. 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.

2.3 Quantifying irreversibility: DHVg and Kullback-Leibler divergence

The main conjecture of this work is that the information stored in the *in* and *out* distributions takes 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 ($P_{\text{in}}(k)$ and $P_{\text{out}}(k)$). If needed, higher order measures can be used, such as the corresponding distance between the *in* and *out* degree-degree distributions ($P_{\text{in}}(k, k')$ and $P_{\text{out}}(k, k')$). These are defined as the *in* and *out* joint degree distributions of a node and its first neighbors [26], describing the probability of an arbitrary node whose neighbor has degree k' to have degree k .

We make use of the Kullback-Leibler divergence [16] 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 to distinguish 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:

$$D(p||q) \equiv \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}, \quad (2)$$

which vanishes if and only if both probability distributions are equal $p = q$ and it is bigger than zero otherwise. Unlike other measures used to estimate time irreversibility [6,9,10,13], the KLD is statistically significant, as it is proved by the Chernoff-Stein lemma: the probability of incorrectly guessing (via hypothesis testing) that a sequence of n data is distributed as p when the true distribution is q tends to $e^{-nD(p||q)}$ when $n \rightarrow \infty$. The KLD is then related to the probability to fail when doing an hypothesis test, or equivalently, it is a measure of “distinguishability”: the more distinguishable are p and q with respect to each other, the larger is $D(p||q)$.

In statistical mechanics, the KLD can be used to measure the time irreversibility of data produced by nonequilibrium processes but also to estimate the average entropy production of the physical process that generated the data [2,4,12,27,28]. Irreversibility can be assessed by the KLD between probability distributions associated to observables in the process and in its time reversal. These measure gives lower bounds to the entropy production, whose accuracy increases as the observables contain a more detailed description of the system. The measure that we present in this work has this limitation: it takes the information from the degree, which is a partial description of the process. Consequently, our technique does not give a tight bound to the entropy production.

Nevertheless, as we will show in several examples, the information of the outgoing degree distribution k_{out} is sufficient to distinguish between reversible and irreversible stochastic stationary series which are real-valued but discrete in time $\{x_t\}_{t=1,\dots,N}$. 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

$$D[P_{\text{out}}(k)||P_{\text{in}}(k)] = \sum_k P_{\text{out}}(k) \log \frac{P_{\text{out}}(k)}{P_{\text{in}}(k)}. \quad (3)$$

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 the majority of 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 [14,15]. 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 $\{k_{\text{in}}(t)\}_{t=1,\dots,N}$ (note that the network has indeed more structure than the degree series). 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) in principle, it takes into account *global* information: each coarse-graining $x_t \rightarrow (k_{\text{in}}(t), k_{\text{out}}(t))$ is performed using information from the whole series, according to the mapping criterion (1). Hence, this symbolization may in principle take into account multiple scales, which is desirable if we want to tackle complex signals [7,8].

3 Reversibility

3.1 Uncorrelated stochastic series

For illustrative purposes, in Figure 2 we have plotted the *in* and *out* degree distributions of the visibility graph associated to an uncorrelated random series $\{x_t\}_{t=1,\dots,N}$ of size $N = 10^6$: the distributions cannot be distinguished and KLD vanishes (the numerical value of KLD is shown in Tab. 1) which is indicative of a reversible series. In what follows we provide an exact derivation of the associated outgoing and ingoing degree distributions associated to this specific process, showing that they are indeed identical in the limit of infinite size series.

Theorem 1. *Let $\{x_t\}_{t=-\infty,\dots,\infty}$ 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*

$$P_{\text{in}}(k) = P_{\text{out}}(k) = \left(\frac{1}{2}\right)^k, \quad k = 1, 2, 3, \dots \quad (4)$$

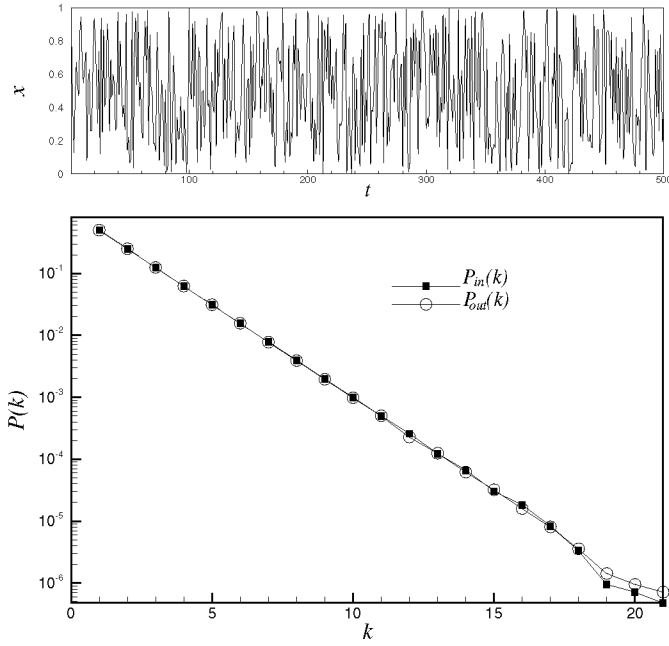


Fig. 2. Top: a sample uncorrelated random time series (500 data points) extracted from a uniform distribution $U[0, 1]$. Bottom: the in and out degree distributions of the DHVg associated to the random series of 10^6 data points. The process is reversible and the graph degree distributions are, besides finite size effects, equivalent. The deviation is measured through their KLD (see Tab. 1). Note that for computing the KLD, in all numerical simulations we have assumed $0 \log 0 = 0$ [5].

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 $f(x)$,

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

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

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

$$P(k) = Q(k)\Phi_1 = \frac{1}{2}Q(k), \quad (5)$$

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

$$Q(k) = Q(k-1)(1 - \Phi_1) = \frac{1}{2}Q(k-1), \quad (6)$$

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

$$Q(k) = \left(\frac{1}{2}\right)^{k-1}, \quad (7)$$

which together with equation (5) concludes the proof. An analogous derivation holds for the *in* case.

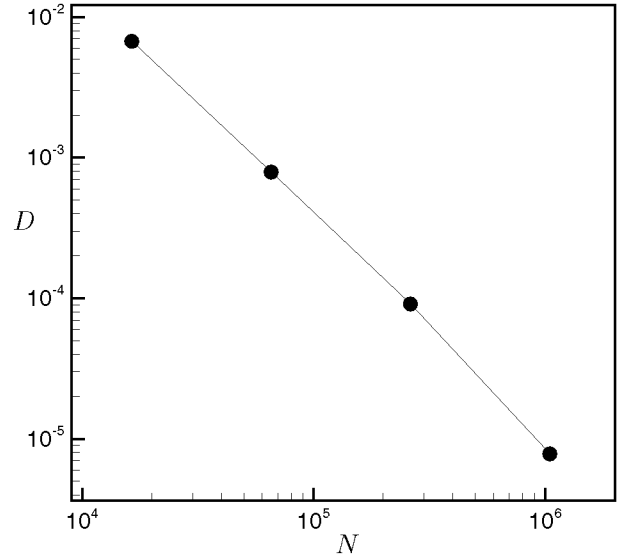


Fig. 3. Log-log plot of $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ of the graph associated to an Ornstein-Uhlenbeck process as a function of the series size N (dots are the result of an ensemble average over several realizations). Note that KLD decreases with series size and tends to zero asymptotically.

Note that 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 $f(x)$. A trivial corollary of this theorem is that the KLD between the *in* and *out* degree distributions associated to a random uncorrelated process tends asymptotically to zero with the series size, which correctly suggests that the series is time reversible.

3.2 Correlated stochastic series

In the last section we considered uncorrelated stochastic series which are our first example of a reversible series with $D[P_{\text{out}}(k)||P_{\text{in}}(k)] = 0$. As a further validation, here we focus on linearly correlated stochastic processes as additional examples of reversible dynamics [1]. We use the *minimal subtraction procedure* [20] to generate such correlated series (details are depicted in an appendix). This method is a modification of the standard Fourier filtering method, which consists in filtering a series of uncorrelated random numbers in Fourier space. We study time series whose correlation is exponentially decaying $C(t) \sim \exp(-t/\tau)$ (akin to an Ornstein-Uhlenbeck process) and power law decaying $C(t) \sim t^{-\gamma}$. In Table 1 we show that the KLD of these series (for $\tau = 1.0$ and $\gamma = 2.0$) are all very close to zero, and its deviation from zero is originated by finite size effects, as it is shown in Figure 3.

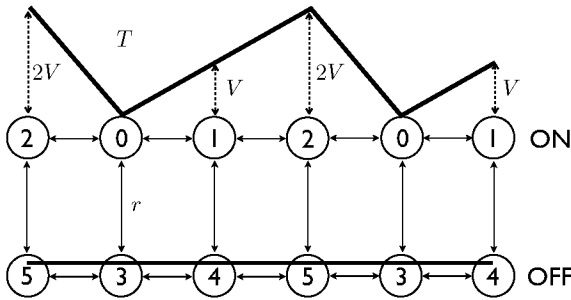


Fig. 4. Discrete flashing ratchet scheme. Particles are at temperature T moving in a periodic linear asymmetric potential of height $2V$. The potential is switched on and off at a constant rate r , which originates a net current of particles to the left. If the potential is ON, the state of the potential is represented by its position $x = \{0, 1, 2\}$, and if it is OFF by $x + 3 = \{3, 4, 5\}$.

4 Irreversibility

4.1 Discrete flashing ratchet

We now study a thermodynamic system which can be smoothly driven out of equilibrium by modifying the value of a physical parameter. We make use of the time series generated by a discrete flashing ratchet model introduced in [4]. The ratchet consists of a particle moving in a one dimensional lattice. The particle is at temperature T and moves in a periodic asymmetric potential of height $2V$, which is switched on and off at a rate r (see Fig. 4 for details). The switching rate is independent of the position of the particle, breaking detailed balance [4,5]. Hence, switching the potential drives the system out of equilibrium resulting in a directed motion or net current of particles. When using full information of the process, trajectories of the system are described by two variables: the position of the particle $x = \{0, 1, 2\}$ and the state of the potential, $y = \{\text{ON}, \text{OFF}\}$. The time series are constructed from x and y variables as follows: $(x, y) = x$ if $y = \text{ON}$ and $(x, y) = x + 3$ if $y = \text{OFF}$.

The dynamics of the system is described by a six-state Markov chain with transition probabilities $p_{i \rightarrow j} = \Gamma_{i \rightarrow j} / \sum_j \Gamma_{i \rightarrow j}$, where $\Gamma_{i \rightarrow j}$ is the transition rate from i to j and the sum \sum_j runs over the accessible states from i (see Fig. 4). All transition rates satisfy the detailed balance condition with respect to the thermal bath at temperature T , except the switches between ON and OFF. When the potential is on, $i, j = \{0, 1, 2\}$ and $\Gamma_{i \rightarrow j} = \exp[-(V_j - V_i)/kT]$. When it is off, $i, j = \{3, 4, 5\}$ and $\Gamma_{i \rightarrow j} = 1$. On the other hand, switches are implemented with rates that do not depend on the position of the particle and therefore do not satisfy detail balance condition [5]: $\Gamma_{i \rightarrow i+3} = \Gamma_{i+3 \rightarrow i} = r$, for $i = \{0, 1, 2\}$ ¹.

¹ Note that in this concrete subsection the series under study are discrete, and in this sense the lack of symbolization that our approach provides is not relevant in this case. However it should be stressed that this in this subsection the aspect under study is not the absence of symbolization, but the degree up to which the method can not only distinguish but quantify

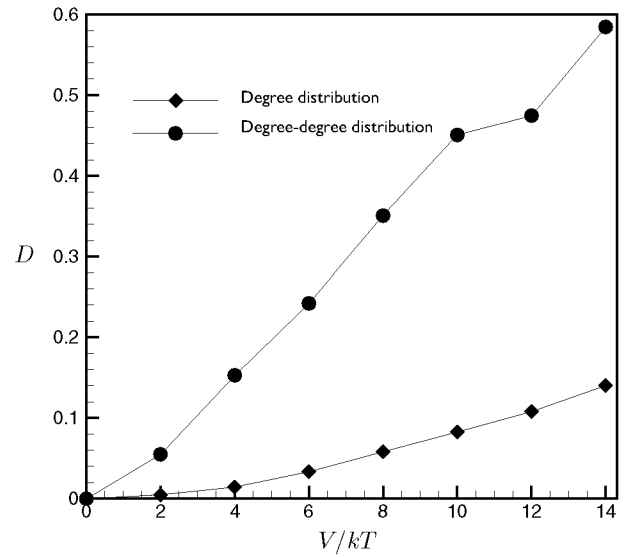


Fig. 5. $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ and $D[P_{\text{out}}(k, k')||P_{\text{in}}(k, k')]$ for a discrete flashing ratchet ($r = 1$) as a function of V/kT . For each value of V we generate a stationary time series of $N = 10^6$ steps described with full information (position and state of the potential). The system is in equilibrium for $V = 0$, and it is driven out of equilibrium for $V > 0$.

In Figure 5 we depict the values of $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ and $D[P_{\text{out}}(k, k')||P_{\text{in}}(k, k')]$ as a function of V/kT , for 6-state time series of 2^{19} data. Note that for $V = 0$ detailed balance condition is satisfied, the system is in equilibrium and trajectories are statistically reversible. In this case both KLD 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 [4]. The amount of irreversibility estimated with KLD increases with V for both measures, therefore the results produced by the method are qualitatively correct. 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 [4]. However the values of the KLD that we find are far below the KLD per step between the forward and backward trajectories, which is equal to the dissipation as reported in [4]. The degree distributions capture the irreversibility of the original series but it is difficult to establish a quantitative relationship between (3) and the KLD between trajectories.

On the other hand, the 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 [16], $D[P_{\text{out}}(k, k')||P_{\text{in}}(k, k')] \geq D[P_{\text{out}}(k)||P_{\text{in}}(k)]$. The improvement is significant in some

the amount of irreversibility, something that can be analyzed within this model, where the amount of irreversibility can be fine tuned.

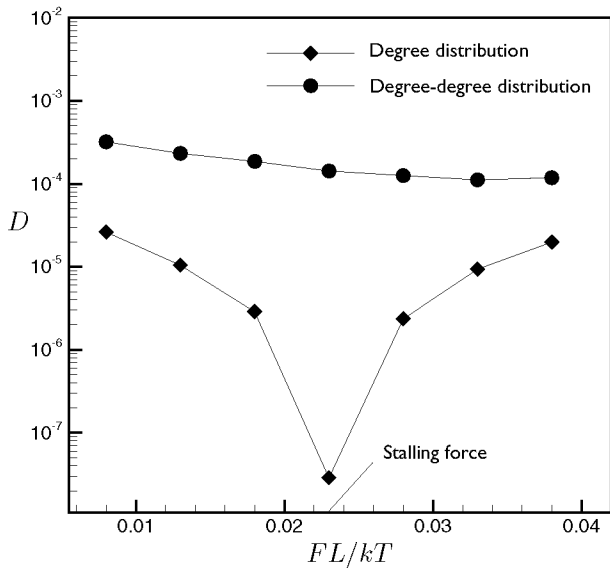


Fig. 6. Irreversibility measures $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ and $D[P_{\text{out}}(k, k')||P_{\text{in}}(k, k')]$ in the flashing ratchet ($r = 2, V = 2kT$) as a function of FL/kT . Here, F is the applied force and L is the spatial period of the ratchet, which in this case is equal to 1. 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).

situations. Consider for instance the flashing ratchet with a force opposite to the net current on the system [4]. 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 only with *partial information* (that is, if the series under study are generated by the successive positions of the particle $x = \{0, 1, 2\}$).

In Figure 6 we address this situation, evaluating our method for series with only partial information. We show how $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ tends to zero when the force approaches to the stalling value (situation with null net current). Therefore, our measure of irreversibility (3) fails in this case, as do other KLD estimators based on local flows or currents [4]. 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 (note that when addressing higher order statistics, convergence of KLD values with system size is slower [5]).

4.2 Chaotic series

We have applied our method 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

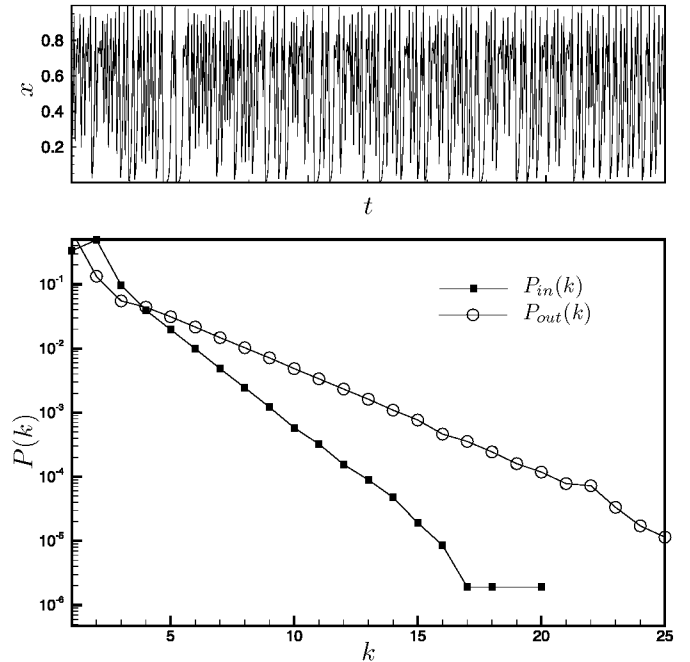


Fig. 7. Top: a sample chaotic time series (500 data points) extracted from a fully chaotic Logistic map $x_{t+1} = 4x_t(1 - x_t)$. Bottom: the *in* and *out* degree distributions of the DHVg associated to the chaotic series of 10^6 data points. The process is irreversible and the graph degree distributions are clearly different. The deviation is measured through the KLD, which is positive in this case (see Tab. 1).

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 [10]. In this section we analyze several chaotic maps and estimate the degree of reversibility of their associated time series using our measure, showing that for dissipative chaotic series it is positive while it vanishes for an example of conservative chaos.

4.2.1 The logistic map at $\mu = 4$ is irreversible: analytical derivations

For illustrative purposes, in Figure 7 we have plotted the *in* and *out* degree distributions of the DHVg associated to a paradigmatic dissipative system: the Logistic map at $\mu = 4$. There is a clear distinction between both distributions, as it is quantified by the KLD, which in this case is positive both for degree and degree-degree cases (see Tab. 1). Furthermore, in Figure 8 we make a finite size analysis in this particular case, showing that our measure quickly converges to an asymptotic value which clearly deviates from zero, at odds with reversible processes.

Recall that in Section 3 we proved analytically that for a random uncorrelated process $D[P_{\text{out}}(k)||P_{\text{in}}(k)] = 0$, since $P_{\text{in}}(k) = P_{\text{out}}(k)$. Proving a similar result for a generic irreversible process is a major challenge, since

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.

Series description	$D[P_{\text{out}}(k) P_{\text{in}}(k)]$	$D[P_{\text{out}}(k,k') P_{\text{in}}(k,k')]$
<i>Reversible stochastic processes</i>		
$U[0,1]$ uncorrelated	3.88×10^{-6}	2.85×10^{-4}
Ornstein-Uhlenbeck ($\tau = 1.0$)	7.82×10^{-6}	1.52×10^{-4}
Long-range (power law) correlated stationary process ($\gamma = 2.0$)	1.28×10^{-5}	2.0×10^{-4}
<i>Dissipative chaos</i>		
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
<i>Conservative chaos</i>		
Arnold Cat map	1.77×10^{-5}	4.05×10^{-4}

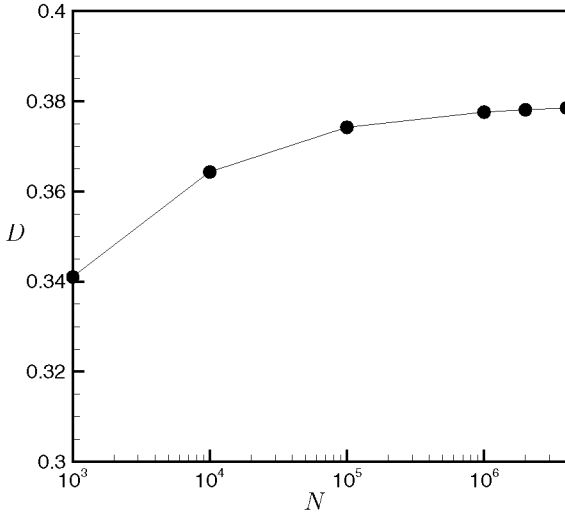


Fig. 8. Semi-log plot of $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ of the graph associated to a fully chaotic Logistic map $x_{t+1} = 4x_t(1 - x_t)$, as a function of the series size N (dots are the result of an ensemble average over different realizations). Our irreversibility measure converges with series size to an asymptotical nonzero value.

finding out exact results for the entire degree distributions is in general difficult [20]. However, note that the KLD between two distributions is zero if and only if the distributions are the same in the entire support. Therefore, if we want to prove that this measure is strictly positive, it is sufficient to find that $P_{\text{in}}(k) \neq P_{\text{out}}(k)$ for some value of the degree k . Here we take advantage of this fact to provide a rather general recipe to prove that a chaotic system is irreversible.

Consider a time series $\{x_t\}_{t=1,\dots,N}$ with a joint probability distribution $f(x_1, x_2, \dots, x_N)$ and support (a, b) , and denote x_{t-1}, x_t, x_{t+1} three (ordered) generic data of

the series. By construction,

$$\begin{aligned}
 P_{\text{out}}(k=1) &= P(x_t \leq x_{t+1}) \\
 &= \int_a^b dx_t \int_{x_t}^b dx_{t+1} f(x_t, x_{t+1}), \\
 P_{\text{in}}(k=1) &= P(x_{t-1} > x_t) \\
 &= \int_a^b dx_{t-1} \int_a^{x_{t-1}} dx_t f(x_{t-1}, x_t). \quad (8)
 \end{aligned}$$

The probability that $k_{\text{out}} = 1$ ($k_{\text{in}} = 1$) is actually the probability that the series increases (decreases) in one step. This probability is independent of time, because we consider stationary series. If the chaotic map is of the form $x_{t+1} = F(x_t)$, it is Markovian, and the preceding equations simplify:

$$\begin{aligned}
 P_{\text{out}}(k=1) &= \int_a^b dx_t \int_{x_t}^b dx_{t+1} f(x_t) f(x_{t+1}|x_t), \\
 P_{\text{in}}(k=1) &= \int_a^b dx_t \int_a^{x_t} dx_{t-1} f(x_{t-1}) f(x_t|x_{t-1}). \quad (9)
 \end{aligned}$$

For chaotic dynamical systems whose trajectories are in the attractor, there exists an invariant probability measure that characterizes the long-term fraction of time spent by the system in the various regions of the attractor. In the case of the Logistic map

$$F(x_t) = \mu x_t(1 - x_t) \quad (10)$$

with parameter $\mu = 4$, the attractor is the whole interval $[0, 1]$ and the probability measure $f(x)$ corresponds to

$$f(x) \equiv \rho(x) = \frac{1}{\pi \sqrt{x(1-x)}}. \quad (11)$$

Now, for a deterministic system, the transition probability is simply

$$f(x_{t+1}|x_t) = \delta(x_{t+1} - F(x_t)), \quad (12)$$

where $\delta(x)$ is the Dirac delta distribution. Equations (9) for the Logistic map with $\mu = 4$ and $x \in [0, 1]$ read

$$\begin{aligned} P_{\text{out}}(k=1) &= \int_0^1 dx_t \int_{x_t}^1 dx_{t+1} f(x_t) \delta(x_{t+1} - F(x_t)), \\ P_{\text{in}}(k=1) &= \int_0^1 dx_t \int_{x_t}^1 dx_{t-1} f(x_{t-1}) \delta(x_t - F(x_{t-1})). \end{aligned} \quad (13)$$

Notice that, using the properties of the Dirac delta distribution, $\int_{x_t}^1 \delta(x_{t+1} - F(x_t)) dx_{t+1}$ is equal to one iff $F(x_t) \in [x_t, 1]$, what happens iff $0 < x_t < 3/4$, and it is zero otherwise. Therefore the only effect of this integral is to restrict the integration range of x_t to be $[0, 3/4]$. The first equation in (13) reduces to

$$P_{\text{out}}(k=1) = \int_0^{3/4} dx_t f(x_t) = 2/3. \quad (14)$$

On the other hand,

$$\int_{x_t}^1 dx_{t-1} f(x_{t-1}) \delta(x_t - F(x_{t-1})) = \sum_{x_k^* | F(x_k^*) = x_t} f(x_k^*) / |F'(x_k^*)|, \quad (15)$$

that is, the sum over the roots x^* of the equation $F(x) = x_t$, iff $F(x_{t-1}) > x_0$. But since $x_{t-1} \in [x_t, 1]$ in the latter integral, it is easy to see that again, this is verified iff $0 < x_t < 3/4$ (as a matter of fact, if $0 < x_t < 3/4$ there is always a *single* value of $x_{t-1} \in [x_t, 1]$ such that $F(x_{t-1}) = x_t$, so the sum restricts to the adequate root). It is easy to see that the particular value is $x^* = (1 + \sqrt{1 - x_t})/2$. Making use of these piecewise solutions and equation (11), we finally have

$$P_{\text{in}}(k=1) = \int_0^{3/4} dx_t \frac{f(x^*)}{4\sqrt{1-x_t}} = 1/3.$$

We conclude that $P_{\text{out}}(k) \neq P_{\text{in}}(k)$ for the Logistic map and hence the KLD measure based on degree distributions is positive. Recall that $P_{\text{out}}(k=1) = 2/3$ is the probability that the series exhibits a positive jump ($x_t > x_{t-1}$) once in the attractor. These positive jumps must be smaller in size than the negative jumps because, once in the attractor, $\langle x_t \rangle$ is constant. The irreversibility captured by the difference between $P_{\text{out}}(k=1)$ and $P_{\text{in}}(k=1)$ is then the asymmetry of the probability distribution of the slope $x_t - x_{t-1}$ of the original time series. The KLD of the degree distributions given by (3) clearly goes beyond this simple signature of irreversibility and can capture more complex and long-range traits.

4.2.2 Other chaotic maps

For completeness, we consider other examples of dissipative chaotic systems analyzed in [29]:

- (1) *the α -map*: $x_{t+1} = 1 - |2x_t - 1|^\alpha$, which reduces to the Logistic and tent maps in their fully chaotic region for

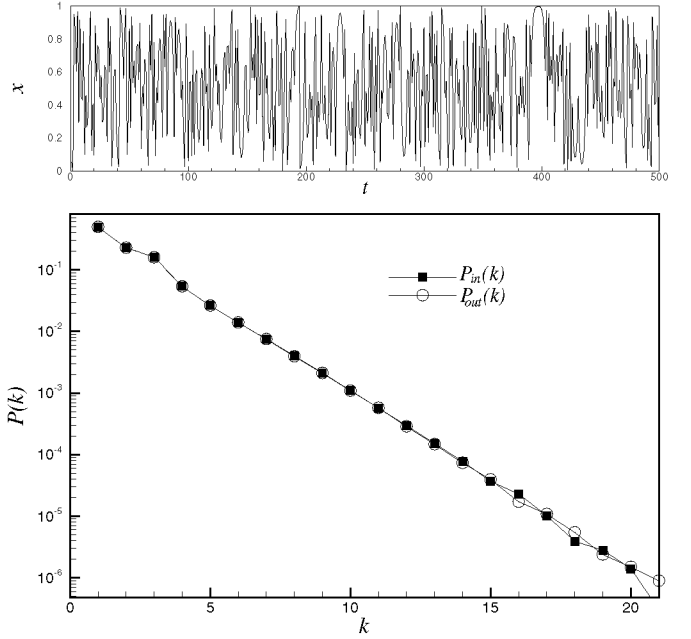


Fig. 9. Top: a sample chaotic time series (500 data points) extracted from the (chaotic and conservative) Arnold cat map. Bottom: the *in* and *out* degree distributions of the DHVg associated to the chaotic series of 10^6 data points. Albeit chaotic, the process is reversible (see the text) and the *in* and *out* degree distributions are, besides finite size effects, equivalent. The deviation is measured through their KLD (see Tab. 1).

$\alpha = 2$ and $\alpha = 1$ respectively. We analyze this map for $\alpha = 3, 4$;

- (2) *the 2D Hénon map*: $x_{t+1} = 1 + y_t - ax_t^2$, $y_{t+1} = bx_t$, in the fully chaotic region ($a = 1.4$, $b = 0.3$);
- (3) *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 $b = 0.5$);
- (4) *the Kaplan-Yorke map*: $x_{t+1} = 2x_t \bmod(1)$, $y_{t+1} = \lambda y_t + \cos(4\pi x_t) \bmod(1)$.

We generate stationary time series with these maps and take data once the system is in the corresponding attractor. In Table 1 we show the value of the KLD between the *in* and *out* degree and degree-degree distributions for these series. 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*: $x_{t+1} = x_t + y_t \bmod(1)$, $y_{t+1} = x_t + 2y_t \bmod(1)$.

At odds with previous dissipative maps, this is an example of a *conservative* (measure-preserving) chaotic system with integer Kaplan-Yorke dimension [29]. 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. In Figure 9 we show a sample series generated by the Arnold cat map, and the *in* and *out* degree distributions

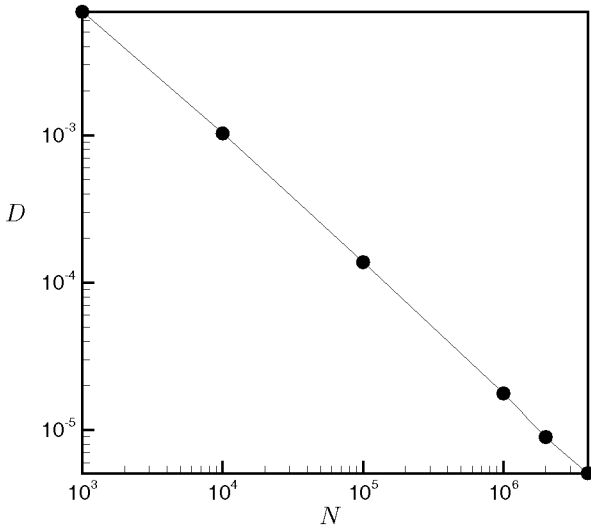


Fig. 10. Log-log plot of $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ of the graph associated to the Arnold cat map as a function of the series size N (dots are the result of an ensemble average over different realizations). Note that the irreversibility measure decreases with series size, and asymptotically tends to zero, which suggests that this chaotic map is reversible.

of its associated DHVg, for a time series of 10^6 data (their KLD is depicted in Tab. 1), and in Figure 10 we show that $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ asymptotically tends to zero with series size, and the same happens with the degree-degree distributions (see Tab. 1). This correctly suggests that albeit chaotic, the map is statistically time reversible.

4.3 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 [30]. 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 ($\mu = 4.0$). In Figure 11 we plot $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ of its associated visibility graph as a function of the noise amplitude (the value corresponding to a pure random signal is also plotted for comparison). 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. Therefore our method correctly predicts that the signal is irreversible even when adding noise.

5 Discussion

In this paper we have introduced a new method to measure time irreversibility of real valued stationary stochastic time series. The algorithm proceeds by mapping the

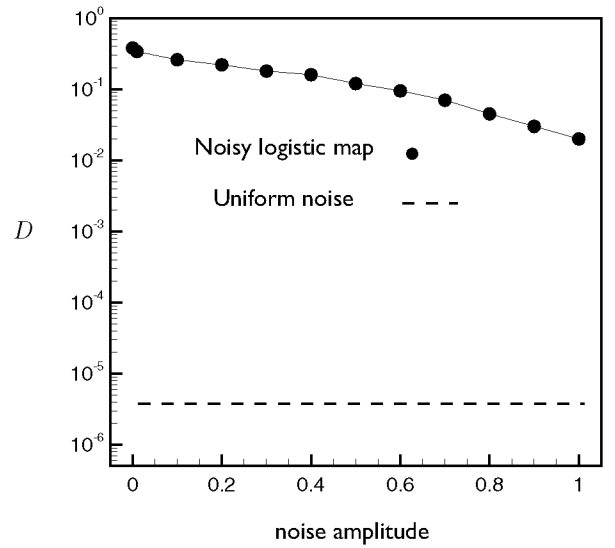


Fig. 11. Semi-log plot of $D[P_{\text{out}}(k)||P_{\text{in}}(k)]$ of the graph associated to series of 10^6 data extracted from a fully chaotic Logistic map $x_{t+1} = 4x_t(1 - x_t)$ polluted with extrinsic white uniform noise $U[-0.5, 0.5]$, as a function of the noise amplitude. The corresponding KLD value of a uniform series is plotted for comparison, which is five orders of magnitude smaller even when the chaotic signal is polluted with an amount of noise of the same amplitude. This suggests that our measure is robust against noise.

series into an alternative representation, the directed horizontal visibility graph. We have shown that the Kullback-Leibler divergence (KLD) between the *in* and *out* degree distributions calculated on this graph is a measure of the irreversibility of the series.

The method has been validated by studying both reversible (uncorrelated and linearly correlated stochastic processes as well as conservative chaotic maps) and irreversible (out-of-equilibrium physical processes and dissipative chaotic maps) series. The method not only discriminates but also quantifies the amount of irreversibility present in the series, as shown in the case study of the discrete flashing ratchet. When the dissipative process happens to show null net current, higher-order statistics of the visibility graph (namely, the joint degree-degree distribution) need to be addressed to detect the irreversible character of the process.

We have also shown that the method is robust against noise, in the sense that irreversible signals are well characterized even when these signals are polluted with a significant amount of (reversible) noise. While the results of our measure for reversible and irreversible dynamics quantitatively differ in several orders of magnitude, a statistical test [11,31,32] can be easily built as follows: one first proceeds to shuffle the series under study in order to generate a randomized resampled data set with the same underlying probability density. This resampled series, whose irreversibility measure is asymptotically null in the light of Theorem 1, is considered as the null hypothesis of the test.

It is also worth emphasizing that it lacks an ad hoc symbolization process, and hence it can be applied directly

to any kind of real-valued time series. While a detailed comparison of the performance of this approach to classical time series symbolization techniques is left for future investigation, the current results suggest that this technique can be of potential interest for several communities. This includes for instance biological sciences, where there is not such a simple tool to discriminate between time series generated by active (irreversible) and passive (reversible) processes. In further work this proposed measure will indeed be used to study empirical data of such kind.

Appendix: Generating correlated series through the minimal subtraction procedure

In what follows we explain the method [20] we have used in Section 3 to generate series of correlated Gaussian random numbers x_i of zero mean and correlation function $\langle x_i x_j \rangle = C(|i - j|)$. The classical method for generating such correlated series is the so-called Fourier filtering method (FFM). This method proceeds by filtering the Fourier components of an uncorrelated sequence of random numbers with a given filter (usually, a power-law function) in order to introduce correlations among the variables. However, the method presents the drawback of evidencing a finite cut-off in the range where the variables are actually correlated, rendering it useless in practical situations. An interesting improvement was introduced some years ago by Makse et al. [33] in order to remove such cut-off. This improvement was based on the removal of the singularity of the power-law correlation function $C(t) \sim t^{-\gamma}$ at $t = 0$ and the associated aliasing effects by introducing a well defined one $C(t) = (1+t^2)^{-\gamma/2}$ and its Fourier transform in continuous-time space. Accordingly, cut-off effects were removed and variables present the desired correlations in their whole range.

We use here an alternative modification of the FFM that also removes undesired cut-off effects for generic correlation functions and takes in consideration the discrete nature of the series. Our modification is based on the fact that not every function $C(t)$ can be considered to be the correlation function of a Gaussian field, since some mathematical requirements need to be fulfilled, namely that the quadratic form $\sum_{ij} x_i C(|i - j|) x_j$ be positive definite. For instance, let us suppose that we want to represent data with a correlation function that behaves asymptotically as $C(t) \sim t^{-\gamma}$. As this function diverges for $t \rightarrow 0$ a regularization is needed. If we take $C(t) = (1+t^2)^{-\gamma/2}$, then the discrete Fourier transform $S(k) = N^{1/2} \sum_{j=1}^N \exp(i \frac{jk}{N}) C(j)$ turns out to be negative for some values of k , which is not acceptable. To overcome this problem, we introduce the *minimal subtraction procedure*, defining a new spectral density as $S_0(k) = S(k) - S_{min}(k)$, being $S_{min}(k)$ the minimum value of $S(k)$ and using this expression instead of the former one in the filtering step. The only effect that the minimal subtraction procedure has on the field correlations is that $C(0)$ is no longer equal to 1 but adopts the minimal

value required to make the previous quadratic form positive definite. The modified algorithm is thus the following:

- generate a set $\{u_j\}$, $j = 1, \dots, N$, of independent Gaussian variables of zero mean and variance one, and compute the discrete Fourier transform of the sequence, $\{\hat{u}_k\}$;
- correlations are incorporated in the sequence by multiplying the new set by the desired spectral density $S(k)$, having in mind that this density is related with the correlation function $C(r)$ through $S(k) = \sum_r N^{1/2} \exp(irk) C(r)$. Make use of $S_0(k) = S(k) - S_{min}(k)$ (minimal subtraction procedure) rather than $S(k)$ in this process. Concretely, the correlated sequence in Fourier space \hat{x}_k is given by $\hat{x}_k = N^{1/2} S_0(k)^{1/2} \hat{u}_k$;
- calculate the inverse Fourier transform of \hat{x}_k to obtain the Gaussian field x_j with the desired correlations.

We acknowledge interesting comments from anonymous referees and financial support from Grants MODELICO, Comunidad de Madrid; FIS2009-13690 (LL, AN and BL) and MO-SAICO (ER and JMRP), Ministerio de Educación.

References

1. G. Weiss, J. Appl. Prob. **12**, 831 (1975)
2. R. Kawai, J.M.R. Parrondo, C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007)
3. J.M.R. Parrondo, C. Van den Broeck, R. Kawai, New. J. Phys. **11**, 073008 (2009)
4. E. Roldan, J.M.R. Parrondo, Phys. Rev. Lett. **105**, 15 (2010)
5. E. Roldan, J.M.R. Parrondo, Entropy production and Kullback-Leibler divergence between stationary trajectories of discrete systems, <http://arxiv.org/abs/1201.5613>
6. A.C. Yang, S.S. Hseu, H.W. Yien, A.L. Goldberger, C.-K. Peng, Phys. Rev. Lett. **90**, 10 (2003)
7. M. Costa, A.L. Goldberger, C.-K. Peng, Phys. Rev. Lett. **95**, 198102 (2005)
8. M.D. Costa, C.K. Peng, A.L. Goldberger, Cardiovasc. Eng. **8**, (2008)
9. C.S. Daw, C.E.A. Finney, M.B. Kennel, Phys. Rev. E **62**, 2 (2000)
10. M.B. Kennel, Phys. Rev. E **69**, 056208 (2004)
11. C. Diks, J.C. van Houwelingen, F. Takens, J. DeGoede, Phys. Lett. A **201**, 221 (1995)
12. P. Gaspard, J. Stat. Phys. **117**, (2004)
13. C. Cammarota, E. Rogora, Chaos Solitons Fractals **32**, 1649 (2007)
14. D. Andrieux, P. Gaspard, S. Ciliberto, N. Garnier, S. Joubaud, A. Petrosyan, Phys. Rev. Lett. **98**, 150601 (2007)
15. Q. Wang, S.R. Kulkarni, S. Verdú, IEEE Transactions on Information Theory **51**, 9 (2005)
16. T.M. Cover, J.A. Thomas, *Elements of Information Theory* (Wiley, New Jersey, 2006)
17. B. Luque, L. Lacasa, J. Luque, F. Ballesteros, Phys. Rev. E **80**, 046103 (2009)
18. L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J.C. Nuno, Proc. Natl. Acad. Sci. USA **105**, 4973 (2008)

19. L. Lacasa, B. Luque, J. Luque, J.C. Nuno, *Europhys. Lett.* **86**, 30001 (2009)
20. L. Lacasa, R. Toral, *Phys. Rev. E* **82**, 036120 (2010)
21. B. Luque, L. Lacasa, F.J. Ballesteros, A. Robledo, *PLoS One* **6**, 9 (2011)
22. J.B. Elsner, T.H. Jagger, E.A. Fogarty, *Geophys. Res. Lett.* **36**, 16 (2009)
23. C. Liu, W.-X. Zhou, W.-K. Yuan, *Physica A* **389**, 13 (2010)
24. Y. Yang, J. Wang, H. Yang, J. Mang, *Physica A* **388**, 4431 (2009)
25. G. Gutin, T. Mansour, S. Severini, *Physica A* **390**, 12 (2011)
26. M.E.J. Newmann, *SIAM Rev.* **45**, 167 (2003)
27. P. Gaspard, *Physica A* **369**, 1 (2006)
28. A. Porporato, J.R. Rigby, E. Daly, *Phys. Rev. Lett.* **98**, 9 (2007)
29. J.C. Sprott, G. Rowlands, *Int. J. Bifurc. Chaos* **11**, 1865 (2001)
30. E.J. Kostelich, T. Schreiber, *Phys. Rev. E* **48**, 1752 (1993)
31. M.J. Hinich, P. Rothman, *Macroecon. Dyn.* **2**, 1 (1998)
32. Y.T. Chen, R.Y. Chou, C.M. Kuan, *J. Econom.* **95**, 199 (2000)
33. H.A. Makse, S. Havlin, M. Schwartz, H.E. Stanley, *Phys. Rev. E* **53**, 5 (1996)