Quantifying the Dynamical Complexity of Chaotic Time Series

Antonio Politi

Institute for Complex Systems and Mathematical Biology, SUPA, University of Aberdeen, AB24 3UE, Aberdeen, United Kingdom

A powerful approach is proposed for the characterization of chaotic signals. It is based on the combined use of two classes of indicators: (i) the probability of suitable symbolic sequences (obtained from the ordinal patterns of the corresponding time series); (ii) the width of the corresponding cylinder sets. This way, much information can be extracted and used to quantify the complexity of a given signal. As an example of the potentiality of the method, I introduce a modified permutation entropy which allows for quantitative estimates of the Kolmogorov-Sinai entropy in hyperchaotic models, where other methods would be unpractical. As a by-product, estimates of the fractal dimension of the underlying attractors are possible as well.

Since the discovery of deterministic chaos, the problem of distinguishing irregular deterministic from stochastic dynamics has attracted the interest of many scientists who have thereby proposed different approaches. An effective method consists in quantifying the diversity of the trajectories of a given length that are generated by a dynamical system. Bandt and Pompe proposed a simple procedure, based on the classification of sequentially sampled time series according to their relative order [1]. The growth rate k_P of the corresponding *permutation entropy* can be indeed easily computed without any prior work. In fact, the permutation entropy has been widely used in many contexts: see, e.g., Refs. [2–6]. The weakness of this approach is the lack of a quantitative correspondence with rigorously defined observables such as the Kolmogorov-Sinai (KS) entropy h_{KS}, which expresses the growth rate of the entropy required to characterize trajectories of a given length [7]. In 1D and 1D-like maps, it has been proved that k_P is equal to h_{KS} [8,9]. Unfortunately, k_P is affected by so strong finite-size corrections as to make extrapolations questionable. In fact, it is generally difficult to obtain directly reliable estimates of h_{KS}. Its computation requires partitioning the phase space into cells so that any trajectory can be encoded as a symbolic sequence. However, only generating partitions ensure a correct estimate of h_{KS} [10]: generic partitions give lower bounds, whose quality is unknown. Effective procedures to construct generating partitions have been developed for two-dimensional maps (or, equivalently, for three-dimensional continuous-time attractors). They are based on the so-called primary homoclinic tangencies which have to be connected in a suitable order [11] (when the dynamics is dissipative) and symmetry lines which allow splitting the stability islands [12] (when the dynamics is Hamiltonian). In any case the procedure requires much work, including an accurate identification of stable and unstable manifolds. Even worse, extensions to higher dimensions are not available. Alternative approaches

have been proposed, based on various types of symbolic encoding (see, e.g., Refs. [13–15]), none of which, goes, however, beyond two-dimensional maps. Given such difficulties, some researchers have proposed various combinations of different indicators [16,17]. However, such strategies do not go that far, as they rely on the same background information: the probability of different symbolic sequences.

In this Letter, I show that substantial progress can be made by taking into account information so far overlooked: "dispersion" among trajectories characterized by the same ordinal sequence. I illustrate the power of the idea by presenting a method which allows obtaining quantitative estimates of h_{KS} even in dynamical systems characterized by more than one positive Lyapunov exponent (LE).

Models.— I start introducing four dynamical systems of increasing complexity used here as a test bed: (i) the Hénon map $[z(t + 1) = a - z(t)^2 + bz(t - 1)$, for $a = 1.4$ and $b = 0.3]$; (ii) the Rössler model $[\dot{z}_1 = z_2 - z_3; \dot{z}_2 = z_1 + az_2; \dot{z}_3 = b + z_3(z_1 - c)$, with $a = 1/2$, $b = 2$, $c = 4]$; (iii) a generalized 3D Hénon (GH) map $[z(t + 1) = a - z(t - 1)^2 - bz(t - 2)$, with $a = 1.5$ and $b = 0.29]$ [18]; (iv) the Mackey-Glass model $\{\dot{z} = 2z(t - t_d)/[1 + z(t - t_d)^{10}] - z$, with $t_d = 3.3\}$.

Let $z(t)$ denote a scalar variable; it is sampled every T time units, so that $z_n = z(t = Tn)$. Next, consider a moving time window of length m. An example is given in Fig. 1: the top red curve is a trajectory of the Mackey-Glass system, while the green dots represent the sampled points (a window of length $m = 6$ is there considered, with $T = 1$). Such a specific sequence is encoded as (3,2,1,5,4,6): each single number denotes the ordinal position (from the smallest to the largest point within the window itself), so that the "1" in position 3 means that the third element within the window is the smallest one. The trajectories can be grouped according to their symbolic sequence. Given the probabilities p_i of all the sequences of length m, one
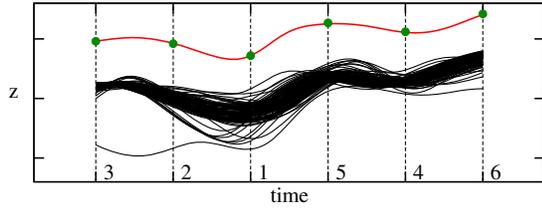
FIG. 1.   100 trajectories of the Mackey-Glass model characterized by the same ordinal pattern (3,2,1,5,4,6) (see the text for more details). The dashed lines identify the sampling times. One trajectory is arbitrarily shifted to better show the ordering of the sampled points (green dots).

can determine the "permutation entropy" $K_P(m) = -\langle \log p_i \rangle$. Its growth rate $k_P = [K_P(m+1) - K_P(m)]/T$ is often used as a proxy for h_{KS}.

In Fig. 1, I have plotted 100 randomly sampled trajectories all encoded as (3,2,1,5,4,6) (the scale is irrelevant). They are basically grouped in a bundle of variable width: the dispersion of $z_i(j)$ ($1 \leq j \leq m$) measures the uncertainty in the position j, for the given ith symbolic sequence. Here below I show that the dispersion of z can be used to characterize a time series. I propose to consider the standard deviation $\sigma_i(j)$ of the variable z measured in the jth position of the window corresponding to the ith sequence. Simulations show that $\sigma_i(j)$ strongly varies with i; it is therefore necessary to to average this observable over the elements of the partition: for reasons that will become clear later, I propose to average its logarithm. The results for the Hénon map are reported in Fig. 2 for different values of the window length m (after rescaling m to unit length). The average $\langle \ln \sigma_i(j) \rangle$ is computed by weighting each sequence according to its probability p_i. It varies along the window and, more important, it decreases upon increasing m.

This is not accidental. In purely stochastic signals, the dispersion decreases as a power law with m. Given, for
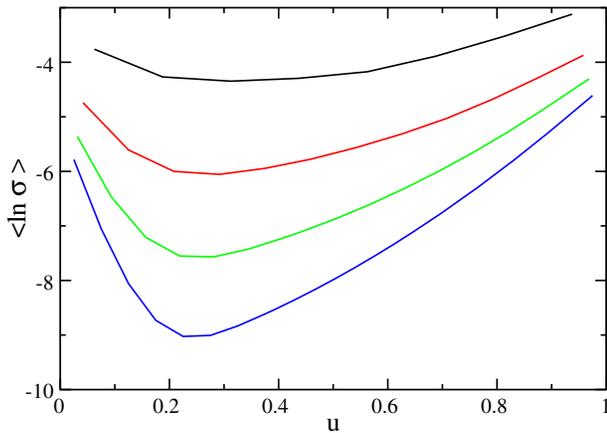
example, a sequence of m random variables uniformly distributed in the unit interval, $\sigma_i(j)$, i.e., the fluctuations of the jth variable (after reordering) is at most of order $1/\sqrt{m}$. In Fig. 3, I plot the dependence of $\langle \ln \sigma_i(m) \rangle$ on the window length m for the above mentioned dynamical systems. Straight lines correspond to power laws; the dashed line exemplifies a $1/m^2$ dependence, a behavior approximately followed in all cases.

The different scaling behavior of $\langle \ln \sigma_i(m) \rangle$ provides a first evidence of the usefulness of the concept of dispersion, as it alone allows distinguishing deterministic from stochastic signals. Going beyond this preliminary observation, in this Letter, I show that the combined use of σ together with the standard permutation entropy makes it also possible to obtain accurate estimates of h_{KS} in purely deterministic contexts.

In this perspective, it is necessary to recall the definition of Kolmogorov-Sinai entropy. Consider an N-dimensional variable $\mathbf{x}(t)$ which evolves in time and denote with $\mathbf{C}_i$ a cylinder in $\mathbb{R}^N \times \mathbb{R}$ of width ε_i and time length τ, centered around some trajectory $\mathbf{x}(t')$ for $t < t' < t + \tau$; let also $p_i(\varepsilon_i, \tau)$ denote the probability that a generic trajectory of length τ is fully contained in $\mathbf{C}_i$. By then covering $\mathbb{R}^N \times [t, t + \tau]$ with nonoverlapping cylinders of width ε_i, one can introduce the entropy $H(\varepsilon, \tau) = -\langle p_i \ln p_i(\varepsilon_i, \tau) \rangle$ (ε without subscripts denotes a generic average width) and thereby the Kolmogorov-Sinai entropy as

$$h_{KS} = -\lim_{\varepsilon \to 0} \lim_{\tau \to \infty} \frac{H(\varepsilon, \tau)}{\tau}, \qquad (1)$$

where the infinite-time limit is to be taken first. The $\varepsilon \to 0$ limit is needed to avoid underestimations of h_{KS} (as mentioned in the introduction, this is not required in the case of generating partitions). From a computational point of view, it is convenient to determine the *derivative*

FIG. 2.   Average dispersion of the Hénon maps for $m = 8, 12, 16$, and 20 (from top to bottom). The window length is rescaled to allow for a clearer comparison [$u = (j - 1/2)/m$].
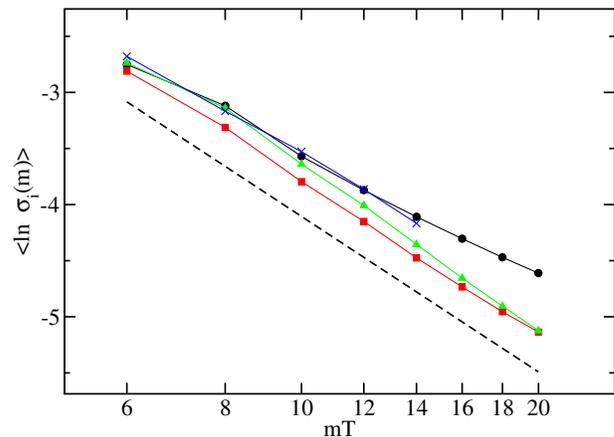
FIG. 3.   The average logarithm of the largest dispersion $\sigma_i(m)$ along a window of length m for various dynamical systems. Circles, squares, triangles, and crosses refer to the Hénon map, Rössler attractor, GH map, and Mackey-Glass model, respectively. The dashed line illustrates a $1/m^2$ decrease.

$$h = [H(\varepsilon, \tau + \Delta) - H(\varepsilon, \tau)]/\Delta$$

as it converges faster than H/τ for increasing τ.

The Pesin relationship between h_{KS} and the Lyapunov exponents is based on the formula (see Refs. [19,20])

$$p_i(\varepsilon_i, \tau) \approx \varepsilon_i^{D_i} \exp[-\Lambda_i \tau],$$

where D_i and Λ_i are the fractal dimension and the sum of the positive finite-time Lyapunov exponents associated with the ith symbolic sequence [21]. Upon averaging over all cylinders, one obtains

$$H(\varepsilon, \tau) = -D\langle \ln \varepsilon_i \rangle + \Lambda \tau \qquad (2)$$

where D is the information dimension and Λ the sum of the positive Lyapunov exponents. The limit $\varepsilon \to 0$ has been implicitly taken since the LEs, by definition, refer to infinitesimal perturbations. Equation (2) implies that $h = \Lambda$, which is nothing but the Pesin formula.

By now going back to the permutation entropy, we can identify $H(\varepsilon, \tau)$ with $K_P(\sigma, mT)$, provided that a meaningful mapping between ε_i and σ_i is established. Rigorously speaking, ε_i is determined in the original N-dimensional phase space, while σ_i refers to the measured, scalar, variable. However, we can safely identify the two observables, since the embedding theorem proved by Takens [22], ensures the existence of a one-to-one mapping between the original and the embedding variables (at least when the window length m is sufficiently large). In the context of the permutation entropy, σ_i is not given *a priori*, but self-determined by the ordering procedure and depends on the symbolic sequence: this is the reason why I have introduced a subscript to denote the width ε_i. In order to complete the identification of ε_i with $\sigma_i(j)$, one should notice that the latter indicator depends on j, i.e., on the position along the window where it is determined. The theoretical argument invoked to derive the Pesin formula requires that ε_i is the maximal distance between two trajectories characterized by the same symbolic sequence. Accordingly, I propose the identification of $\sigma_i(m)$ with ε_i, since in all cases I have investigated, the maximal cylinder width is attained in the mth position.

The most important property of $\sigma_i(m)$ is that it decreases upon increasing the window length (see Fig. 3). This implies that ε_i in Eq. (2) does depend on τ (or, equivalently, on m) and this invalidates the direct connection between k_P and Λ (and thereby with h_{KS}). A clean relationship can be reestablished by introducing the relative permutation entropy

$$\tilde{K}_P(m) = K_P(m) + D\langle \log \sigma_i(m) \rangle,$$

where the dependence of σ on time is spelled out. The structure of Eq. (3) justifies the choice of averaging the
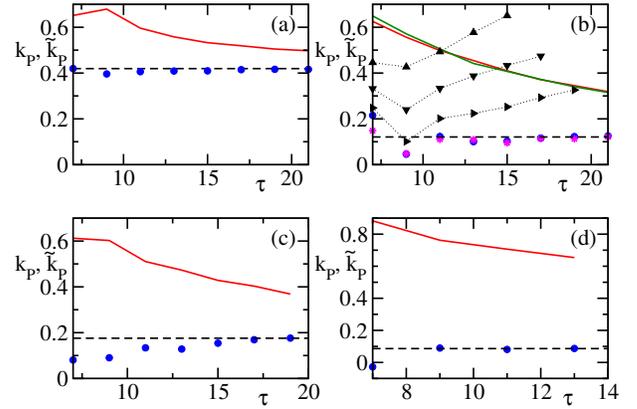
FIG. 4. Finite-size estimates of the Kolmogorov Sinai entropy. The solid curves correspond to the derivative k_P of the permutation entropy; symbols to the derivative $\tilde{k}_P$ of the relative entropy Eq. (3). The four panels report the results for: (a) Hénon map, (b) Rössler attractor, (c) GH map, and (d) Mackey-Glass model. In all cases the horizontal black dashed line corresponds to the KS entropy estimated as the sum of the positive Lyapunov exponents. The triangles in panel (b) correspond to $\tilde{k}_P$ for three different levels of observational noise: 0.01, 0.005, 0.002 (from top to bottom).

logarithm of σ_i. A reliable estimate of h_{KS} is obtained by taking the derivative $\tilde{k}_P = [\tilde{K}_P(m + 1) - \tilde{K}_P(m)]/T$ that is cleansed of the time dependence of σ.

Validation.— The results obtained for four different models are summarized in Fig. 4. I start from the Hénon map, whose analysis is plotted in panel (a). In this case, there is only one positive LE, $\lambda_1 \approx 0.4192$, so that h_{KS} coincides with λ_1. The fractal (information) dimension D, is equal to 1.258…, as obtained from the Kaplan-Yorke formula. The derivatives k_P and $\tilde{k}_P$ reported in Fig. 4(a) are both determined by referring to a time interval Δ equal to 2. The results indicate that $\tilde{k}_P$, provides relatively accurate estimates already for $\tau = 9$ (here $m = \tau$). The second model I have studied is the Rössler attractor, selecting the parameter values in such a way that the dynamics is not phase coherent, to make the attractor as different as possible from that of the Hénon map. Since time is continuous, it is necessary to fix the sampling interval T. I have chosen $T = 1$, which is about 1/8th of the main periodicity, but allows for an appreciable variation of z_1 (up to 1/5 of its whole range). In this case, there is again not only one positive LE ($\lambda_1 \approx 0.1208$), but also a vanishing one, which does not contribute to the KS-entropy, but indirectly to the correction term, affecting the dimension which is $D \approx 2.05$. The results are plotted in Fig. 4(b). The red curve and the blue circles have been obtained by sampling $z_1(t)$, while the green curve and the stars correspond to $z_3(t)$. The agreement proves the approach is robust: there is no problem of variable selection. The asymptotic value of the KS entropy is achieved for $m = \tau = 10$, where the

permutation entropy estimates are still five times larger than expected.

After validating the method in well known setups, I turned my attention to models characterized by two positive Lyapunov exponents, where no direct information is available. The first is another discrete-time system, the GH map, where $\lambda_1 \approx 0.1179$, and $\lambda_2 \approx 0.0577$, so that the KS entropy is 0.1756…, while the dimension is $D \approx 2.12$; $\tilde{k}_P$ exhibits a convincing agreement with the theoretical expectations, although for $m = 13$ there are still deviations of order 25% (to be, however, compared with the 6-times larger overestimation of the traditional method). Finally, I have studied the Mackey-Glass equation. This is a model with delayed interactions; i.e., the phase space is infinite dimensional. For the parameter value I have selected $\lambda_1 \approx 0.0617$ and $\lambda_2 \approx 0.0234$, so that $h_{KS} \approx 0.0862$, and the Kaplan-Yorke dimension is $D \approx 3.73$. By looking at Fig. 4(d), one can see that a remarkable agreement is found already for $\tau = 9$ (the sampling time T has been fixed equal to 1).

Since the goal of the simulations was to validate the approach, all studies have been performed using enough data points to ensure sufficiently small errors. In the perspective of future applications, it is important to have an idea of the required number of points. In the Mackey-Glass attractor (the model affected by the largest deviations), 3×10^5 points are sufficient to ensure a 1% accuracy for $\tau = 10$ of both the entropy and the logarithm of the corresponding box size. This number grows exponentially with τ, becoming of the order of 10^7 for $\tau = 14$, but it should be noticed that $\tau = 10$ is sufficient to ensure a reliable estimate of h_{KS}. Furthermore, the convergence of both quantities is monotonic (upon increasing the length of the time series); it follows a power law so that extrapolation methods can be used to improve the accuracy of the asymptotic value (see, e.g., Ref. [23]).

Experimental data are typically affected by observational noise. As a result, the Kolmogorov-Sinai entropy becomes, strictly speaking, infinite. Therefore, it is important to understand how noise affects the scaling of entropy over different (temporal and observational) resolutions. As a minimal test I have simulated the Rössler attractor by adding a uniform noise distributed within the interval $[-\delta, +\delta]$ [for $\delta = 0.01, 0.005$, and 0.002—see the triangles in Fig. 4(b)]. The finite-size estimates of h_{KS} increase with the size of the noise (as expected), but the resulting values are still much below the maximal values obtained for random signals (around 2.74 for $\tau = 15$).

Altogether, the time dependence of the cylinder width helps to resolve an ostensible paradox: finite partitions typically tend to underestimate the KS entropy, because they are unable to discriminate all of the different trajectories. Nevertheless, the permutation entropy overestimates h_{KS}: this is because part of the entropy increase is a spurious effect induced by the implicit refinement of the phase-space partition. The modified definition herein proposed gets rid of such a contribution.

The results for $\langle \log \sigma_i(m) \rangle$ reported in Fig. 3 and theoretical arguments for purely random signals show that σ decreases as a power law, $\sigma \approx m^{-\gamma}$. By combining this observation with the assumption that this is the major source of finite-size corrections (at least in a suitable range of m values), one can claim that $\tilde{K}_P(m) = K_0 + \Lambda m$ and thereby write

$$K_P(m) = K_0 + \Lambda m + D\gamma \ln m. \qquad (3)$$

This equation suggests, at the same time, that the derivative of the standard permutation entropy eventually converges to Λ, but also that it is affected by strong (logarithmic in the window length) corrections. They make the estimation of the asymptotic value prohibitive.

So far I have shown that accurate estimates of h_{KS} can be obtained without the need of explicitly partitioning the phase space, but this requires the knowledge of the fractal dimension D. Now I show that this obstacle can be overcome. With reference to Eq. (3), I replace D with an unknown parameter d and thereby introduce $\tilde{K}_P(m, d)$ [$\tilde{K}_P(m, 0)$ is the usual permutation entropy]. So long as $d < D$, the derivative $\tilde{k}_P(m, d)$ converges to the asymptotic value from above, while a convergence from below is expected when $d > D$. Therefore, D can be estimated as the d value such that $\tilde{k}_P(m, d)$ is independent of m. With this idea in mind, one can go to the initial data and determine $\tilde{k}_P(m, d)$ in a suitable range of m values for different d values. A linear fit of the last seven points for the Hénon and Roessler attractors shows that the average derivative of $\tilde{k}(m, d)$ changes sign for $d \approx 1.13$ and $d \approx 1.9$, respectively. For the GH map, the change of sign occurs for $d \approx 1.56$, while for the Mackey Glass model, I obtain $d \approx 3.8$. All values are close to the expected estimates of the dimension, with the exception of the generalized Hénon map, whose dimension is underestimated by about 0.6. This is understandable, since from Fig. 4(c) one can see that such a dynamical system is the only one where the convergence of $\tilde{k}_P$ is not perfect. The reason is probably a slow convergence of the dimension itself to its asymptotic value: in other words, it is reasonable to interpret the value $d = 1.56$ as the finite-size dimension of the attractor on the scales that are accessed by the numerical analysis.

Altogether, I have shown that the information contained in the permutation entropy can be complemented by the dispersion σ of trajectories characterized by the same ordinal sequence to provide a more complete description of the time series. In the case of deterministic signals, the two notions are combined into a a single indicator, which provides reliable estimates of the Kolmogorov-Sinai entropy even in the case of multiple positive Lyapunov exponents. Preliminary studies suggest that this approach

can be used also as a zero-knowledge tool for the estimation of the fractal dimension. Such achievements are possible because the increase of the window length corresponds to the simultaneous increase of both the embedding dimension [24] and of the resolution in phase space. I am confident that σ can be profitably used also for a better characterization of (partially) stochastic signals. The results reported in Fig. 3 provide encouraging evidence of quantitative differences between stochastic and deterministic signals. Further progress can be made by introducing proper indicators which take into account the scaling dependence of σ on noise.

[1] C. Bandt and B. Pompe, Phys. Rev. Lett. **88**, 174102 (2002).
[2] Y. Cao, W. W. Tung, J. B. Gao, V. A. Protopopescu, and L. M. Hively, Phys. Rev. E **70**, 046217 (2004).
[3] M. Staniek and K. Lehnertz, Int. J. Bifurcation Chaos Appl. Sci. Eng., **17**, 3729 (2007).
[4] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, Entropy, **14**, 1553 (2012).
[5] C. Quintero-Quiroz, S. Pigolotti, M. C. Torrent, and C. Masoller, New J. Phys. **17**, 093002 (2015).
[6] P. J. Weck, D. A. Schaffner, M. R. Brown, and R. T. Wicks, Phys. Rev. E **91**, 023101 (2015).
[7] Ya. Sinai, Scholarpedia **4**, 2034 (2009).
[8] C. Bandt, G. Keller, and B. Pompe, Nonlinearity, **15**, 1595 (2002).
[9] J. M. Amigó, B. Kennel, and L. Kocarev, Physica (Amsterdam) **210D**, 77 (2005).
[10] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[11] P. Grassberger and H. Kantz, Phys. Lett. A **113**, 235 (1985).
[12] P. Christiansen and A. Politi, Physica **109D**, 32 (1997).
[13] M. B. Kennel and M. Buhl, Phys. Rev. Lett. **91**, 084102 (2003).
[14] Y. Hirata, K. Judd, and D. Kilminster, Phys. Rev. E **70**, 016215 (2004).
[15] M. Buhl and M. B. Kennel, Phys. Rev. E **71**, 046213 (2005).
[16] R. López-Ruiz, H. L. Mancini, and X. Calbet, Phys. Lett. A **209**, 321 (1995).
[17] O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, and M. A. Fuentes, Phys. Rev. Lett. **99**, 154102 (2007).
[18] Initial conditions which lie in the (small) basin of attraction are $z(0) = -0.07$, $z(1) = 0$, $z(2) = 0.07$.
[19] Ya. B. Pesin, Russ. Math. Surv. **32**, 55 (1977).
[20] A. Pikovsky and A. Politi, *Lyapunov Exponents* (Cambridge University Press, Cambridge, 2016).
[21] Generic systems are "multifractal," i.e., they are characterized by a spectrum of different dimensions and LEs.
[22] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. A. Rand and L.-S Young (Springer, London, 1981), Vol. 898, pp. 1–366.
[23] T. Schürmann and P. Grassberger, Chaos **6**, 414 (1996).
[24] H. Kantz and Th. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, England, 2004).