Automatic estimation of attractor invariants

Juan F. Restrepo & Gastón Schlotthauer
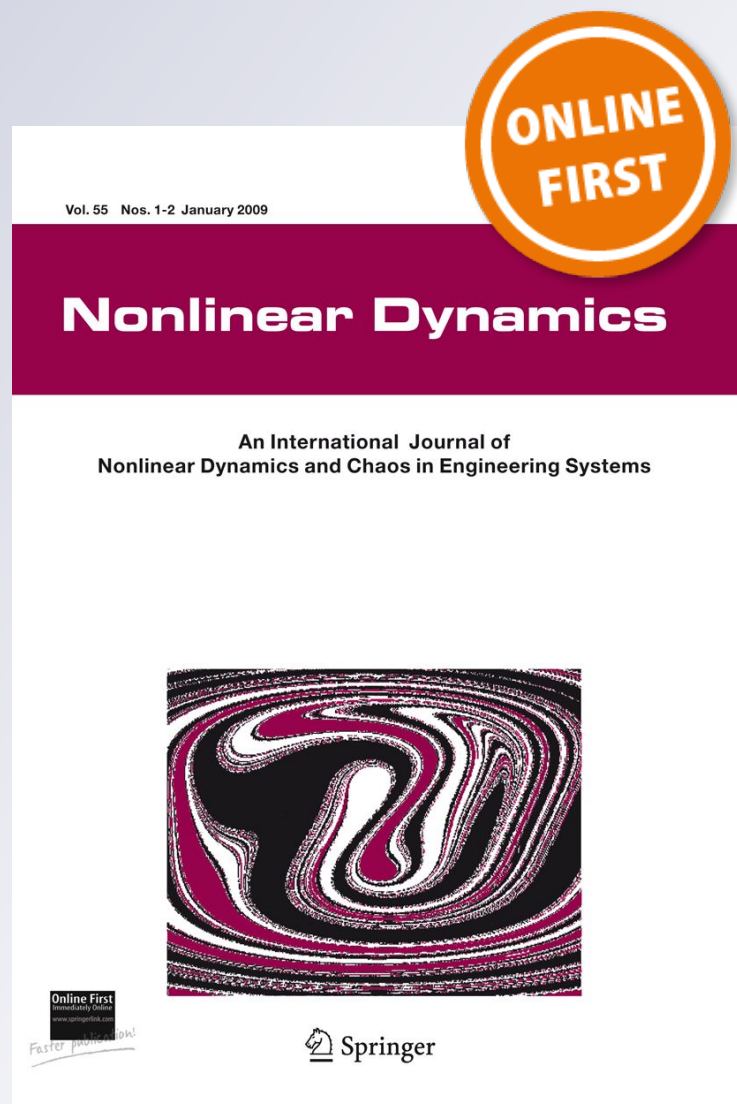
Vol. 55   Nos. 1-2 January 2009

Nonlinear Dynamics

An International Journal of
Nonlinear Dynamics and Chaos in Engineering Systems

Online First
Immediately Online
www.springerlink.com
Faster publication!

ONLINE FIRST

Springer

Springer

CrossMark

ORIGINAL PAPER

Automatic estimation of attractor invariants

Juan F. Restrepo · Gastón Schlotthauer

Abstract The invariants of an attractor have been the most used resource to characterize a nonlinear dynamics. Their estimation is a challenging endeavor in short-time series and/or in presence of noise. In this article, we present two new coarse-grained estimators for the correlation dimension and for the correlation entropy. They can be easily estimated from the calculation of two U-correlation integrals. We have also developed an algorithm that is able to automatically obtain these invariants and the noise level in order to process large data sets. This method has been statistically tested through simulations in low-dimensional systems. The results show that it is robust in presence of noise and short data lengths. In comparison with similar approaches, our algorithm outperforms the estimation of the correlation entropy.

J. F. Restrepo (✉) · G. Schlotthauer
Signals and Nonlinear Dynamics Laboratory - Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Ruta prov 11 km 10, Oro Verde, Entre Ríos, Argentina
e-mail: jrestrepo@bioingenieria.edu.ar

G. Schlotthauer
e-mail: gschlotthauer@conicet.gov.ar

1 Introduction

Correlation dimension (D) and correlation entropy (K_2) are quantities that characterize the natural measure of an attractor. The first one can be interpreted as the dimension of the attractor. It is a rough measure of the effective number of degrees of freedom (number of variables) involved in the dynamical process. The correlation entropy is a measure of the rate at which pairs of nearby orbits diverge. In other words, it gives us an estimate of the unpredictability of the system [6].

The correlation dimension and entropy allow us to measure the complexity of a dynamical system. Typically, more complex systems have higher dimensions and larger values of entropy. More important, these quantities help us to identify changes in the system's complexity. For this reason, they have become very popular and widely used in the biomedical field [27], not only to characterize the dynamics of physiological systems, but also to detect different types of pathologies. For example, Choi et al. used the correlation dimension calculated from voice signal to characterize the dynamics of the phonatory system in normal and pathological conditions. They reported a mean value of $D = 1.57$ for normal voices and an increased dimension value for pathological ones [3]. In Hosseinifard et al. [15] used D, among other nonlinear

Springer

measures, extracted from electroencephalographic signals to classify between depressed and normal patients. Combining these features with machine learning techniques, they achieved a 90% classification accuracy. On the other hand, in [8,9] the authors illustrated how K_2 and other related entropy measures have been used to characterize different biological phenomena.

Since the natural measure of an attractor is invariant under the dynamical evolution, both D and K_2 can be easily estimated from indirect time measurements of one of the system variables [19]. For an observed stationary time series of length N, $\{x_n\}_{n=1}^N$, the reconstructed m-dimensional delay vectors $\boldsymbol{x_i} = \left(x_i, x_{i+\tau} \ldots, x_{i-(m-1)\tau}\right)$, $i = 1, 2, \ldots, L = N - (m-1)\tau$, must be formed. Here m is the embedding dimension and τ is the embedding lag. As a collection, these delay vectors constitute the reconstructed attractor of the system in $\mathbb{R}^m$. From another point of view, this collection can be thought as samples drawn from the natural measure of the system. These samples can be used to estimate the invariants D and K_2 using the correlation integral $G_m(h)$. This last quantity is defined as the probability that the distance $\tilde{z} = \|\boldsymbol{x_i} - \boldsymbol{x_j}\|$ between two randomly selected m-dimensional delay vectors $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$ is smaller than a value h [6]:

$$G_m(h) = \int g(h, \tilde{z}) \, f_{\tilde{z}}^m(\tilde{z}) \, \mathrm{d}\tilde{z}, \tag{1.1}$$

where g is a kernel function and $f_{\tilde{z}}^m(\tilde{z})$ is the probability density function (pdf) of the distance between pairs of delay vectors. In this article, the Euclidean distance is used, but others measures of distance can be employed. The Grassberger and Procaccia correlation integral $C_m(h)$ uses as kernel function $g(h, \tilde{z}) = H(1 - \tilde{z}/h)$, where $H(\cdot)$ is the Heaviside step function [11]. On the other hand, the Gaussian kernel correlation integral (GCI) $T_m(h)$ proposed by Diks et al. [5] adopts $g(h, \tilde{z}) = \exp\left(-\tilde{z}^2/4h^2\right)$. For deterministic time series and in absence of noise, both correlation integrals scale as:

$$C_m(h) = T_m(h) \sim h^D e^{-m\Delta t \tau K_2},$$
$$\text{for} \quad m \to \infty \quad \text{and} \quad h \to 0.$$

where $\Delta t = 1/f_s$, being f_s the sampling frequency of the time series. This behavior is due to that, in the attractor of a chaotic dynamical system, the expected

density of points within a ball of small radius h scales as h^D. Thus, the greater the correlation dimension, the more complex the attractor, and the longer it takes to revisit the same region in it [2]. On the other hand, the $e^{-m\Delta t \tau K_2}$ factor is related to the exponential divergence of nearby trajectories. The higher the entropy values, the faster those trajectories diverge [7,12].

The main advantage of the GCI over the Grassberger and Procaccia correlation integral is that the former allows us to model the influence of additive noise on the scaling law. For a time series with added white Gaussian noise of variance σ^2, the scaling rule is [5, 22,28]:

$$T_m(h) \sim h^m \left(h^2 + \sigma^2\right)^{\frac{D-m}{2}} e^{-m\Delta t \tau K_2}$$
$$\text{for} \quad m \to \infty \quad \text{and} \quad \sqrt{h^2 + \sigma^2} \to 0. \tag{1.2}$$

In presence of noise, it is very important to be able to quantify its level σ in order to correct the estimation of D and K_2. Moreover, an imprecise estimate of these invariants can lead to mistaken conclusions about the studied phenomena. In this sense, an estimate of the noise level must be reported allowing other researchers to be aware of the limitations of the data.

The main problem of the GCI is that it requires high values of m to obtain a convergent estimate of K_2 [22, 25]. This is a major drawback, since high values of m imply much longer time series in order to achieve good estimations.

The cause of this problem is related with the kernel function of the GCI. It can be seen in Eq. (1.1) that the pdf of the interpoint distance depends on the embedding dimension. However, the Gaussian kernel function does not change with m, and, for this reason, the convergence of the GCI to the K_2 is slowed down [24].

As a solution, we have recently proposed the U-correlation integral (UCI) [24]:

$$U_m^\beta(h) = \int \frac{\Gamma\left(\beta/2, z/h^2\right)}{\Gamma(\beta/2)} \hat{f}_m(\sigma; z) \, \mathrm{d}z, \tag{1.3}$$

where $\Gamma(a, t)$ is the upper incomplete Gamma function, $\Gamma(a)$ is the Gamma function and $\hat{f}_m(\sigma; z)$ is the pdf of the squared interpoint distance. There are two important issues about this correlation integral that deserve further analysis. First, note that the UCI's kernel function, given by:

$$g\left(h, \beta, z\right) = \frac{\Gamma\left(\beta/2, z/h^2\right)}{\Gamma\left(\beta/2\right)}$$

has a parameter β that is used to incorporate information about the embedding dimension. In other words, this kernel function is able to change according to m. Second, we are now working with the square of the interpoint distance, i.e., $z = \tilde{z}^2$, to reduce the computational cost.

In order to find the scaling behavior of the UCI, we need an expression for the pdf of the squared distances between pairs of noisy delay vectors $\hat{f}_m\left(\sigma; z\right)$ [23, see Eq. (18)]:

$$\hat{f}_m\left(\sigma; z\right) = \frac{z^{m/2-1}}{2\left(2\sigma\right)^{m-D}} \frac{\Gamma\left(D/2\right)}{\Gamma\left(m/2\right)}$$
$$\times \,_1F_1\left(, \frac{m-D}{2}; \frac{m}{2}\right) - \frac{z}{4\sigma^2}, \qquad (1.4)$$

where $\,_1F_1(a, b; t)$ is the Kummer's confluent hypergeometric function. This distribution arises under the assumption that, in absence of noise, the pdf of the pairwise distances between m-dimensional delay vectors (points in the reconstructed attractor) behaves as h^{D-1} for a certain range of h values [23]. If each of the coordinates of these vectors is perturbed with white Gaussian noise of variance σ^2, then the pdf of the perturbed distances will follow Eq. (1.4).

The scaling law for the UCI can be obtained by substituting Eq. (1.4) in Eq. (1.3) and integrating over z [24]:

$$U_m^\beta\left(h\right) = \frac{\widehat{\phi}}{2}\left(2\sigma\right)^D \frac{\Gamma\left(D/2\right)\Gamma\left(\left(\beta+m\right)/2\right)}{\Gamma\left(\beta/2\right)\Gamma\left(m/2+1\right)} e^{-m\Delta t\tau K_2}$$
$$\times \left(\frac{h^2}{4\sigma^2}\right)^{m/2} F\left(\frac{\beta+m}{2}, \frac{m-D}{2}; \frac{m+2}{2}; -\frac{h^2}{4\sigma^2}\right),$$
$$(1.5)$$

where $\hat{\phi}$ is a normalization constant and $F\left(a, b; c; t\right)$ is the Gauss hypergeometric function.

In practical applications, once the scaling law has been selected and the correlation integral has been calculated, there are two main options to estimate D, K_2 and σ. The first one is to estimate these invariants through a nonlinear fitting of the scaling rule [5,28]. This method is highly dependent on the range of scale selected to fit the model, and there is not a consensus about the correct way to choose it [19]. The second option is to use coarse-grained estimators. These are

explicit expressions for D, K_2, and σ as functions of m and h [6]. For example, in [22] Nolte et al. have proposed a set of coarse-grained functions based on the GCI. The most important aspect of their approach is that the calculation of these functions only depends on the estimation of $T_m(h)$ for two consecutive values of the embedding dimension m. Nevertheless, given that these estimators are based on the GCI, the convergence to K_2 needs high values of m and, therefore, large amounts of data.

In [24] we have proposed a set of coarse-grained estimators for D, K_2 and σ based on the UCI. However, these estimators are highly dependent on a precise estimation of σ.

In this article, we present a new set of coarse-grained estimators for D and K_2, based on the UCI, that only needs the calculation of two U-correlation integrals i. e. they do not require the a priori estimation of any quantity. Moreover, we propose a methodology to automatically estimate these invariants and the noise level of the time series.

In Sect. 2, we derive the here proposed coarse-grained estimators. Section 3 is devoted to describe the algorithm to automatically estimate the invariants. In Sect. 4, we discuss these results and give some suggestions for the practical of the algorithm. Finally, the conclusions are presented in Sect. 5.

2 Theory

Following a similar approach to one given by Nolte et al. [22], the here proposed coarse-grained estimators are based on a noise level functional. This quantity, closely related to the noise level of the time series, can be estimated using two U-correlation integrals. In this work, the noise level functional will be employed to correct the deviation of the coarse-grained estimators from the scaling law in presence of noise.

2.1 Noise level functional

The noise level functional based on the UCI can be defined as [24]:

$$\Delta_m^U\left(h\right) = \frac{1}{2}\left(\dot{D}_{m+2}^{\beta=m}\left(h\right) - \dot{D}_m^{\beta=m}\left(h\right)\right)$$
$$\approx \frac{4\sigma^2}{h^2 + 4\sigma^2} \quad \text{for } m \gg D, \qquad (2.1)$$

where

$$\dot{D}_m^\beta(h) = \frac{d \ln U_m^\beta(h)}{d \ln h}.$$

As it can be observed from Eq. (2.1), $\Delta_m^U(h)$ is a function that decreases monotonically from 1 to 0 on a scale that is proportional to the value of σ. This functional depends on the logarithmic derivative of two U-correlation integrals: $U_m^{\beta=m}(h)$ and $U_{m+2}^{\beta=m}(h)$. Note that the parameter β is set equal to m in both correlation integrals.

In order illustrate the behavior of this noise level functional, we will use the Henon map:

$$\begin{cases} x_{n+1} = 1 - ax_n^2 + y_n \\ y_{n+1} = bx_n \end{cases},$$

where $a = 1.4$ and $b = 0.3$. This map has been widely used in the literature to numerically characterize different methods devoted to estimate attractor's invariants [4,6,16,19,20,22,23,25,26,28]. For this map, and the given parameters values, the correlation dimension and correlation entropy are $D = 1.22$ and $K_2 = 0.3$, respectively [28].

The noise level functionals for both a clean and a noisy ($\sigma = 0.05$) realizations of the Henon map of length $N = 10000$ are shown in Fig. 1. Note that in absence of noise (Fig. 1a), this functional is close to zero for a wide range of h values regardless the value of m. On the other hand, in Fig. 1b it can be observed that $\Delta_m^U(h)$ falls of from 1 to 0 and the curves for different m values are very close to the theoretical curve for $\sigma = 0.05$. We must point out that the closest curve of this functional to the theoretical value is the one corresponding to $m = 2$. For larger values of m, $\Delta_m^U(h)$ slightly deviates from this behavior.

2.2 Noise level coarse-grained estimator

Starting from Eq. (2.1), we have derived the coarse-grained noise level estimator [24]:

$$\sigma_m^U(h) = \frac{h}{2} \sqrt{\frac{\Delta_m^U(h)}{1 - \Delta_m^U(h)}}. \tag{2.2}$$

We must point out that all time series used in this work were rescaled to have unitary standard deviation. In consequence, σ is the noise level after rescaling, i.e.,
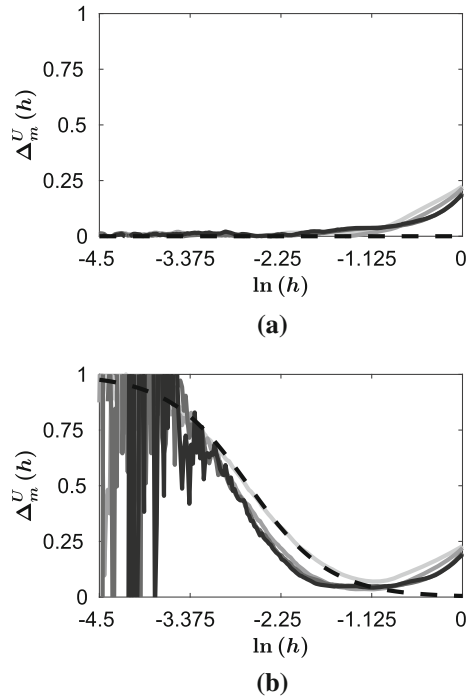

Fig. 1 Henon map: Noise level functional for $m = \{2, 4, 6, 8\}$, the curves for different m values are color-coded in grayscale where the lightest gray corresponds to $m = 2$. **a** Noiseless. **b** $\sigma = 0.05$. The theoretical curve [right-hand side of Eq. (2.1)] is shown in dashed black line

$$\sigma = \frac{\sigma_n}{\sqrt{\sigma_c^2 + \sigma_n^2}},$$

where σ_n^2 is the noise variance and σ_c^2 is the variance of the clean time series. In this sense, $\sigma \to 0$ corresponds to clean time series and $\sigma \to 1$ implies only noise. The signal-to-noise ratio (SNR) can be calculated as:

$$\text{SNR} = 10 \log_{10} \left(\frac{1 - \sigma^2}{\sigma^2} \right) \text{ dB}.$$

In Fig. 2, it is shown the coarse-grained estimator for noise level as a function of h for $m = \{2, 4, 6, 8, 10\}$ for both a clean and a noisy ($\sigma = 0.05$) realization of the Henon map with length $N = 10000$. It can be observed in Fig. 2a that, for the noiseless case, $\sigma_m^U(h)$ approaches to zero as $h \to 0$. On the other hand, Fig. 2b shows that, when the noise level is increased, $\sigma_m^U(h)$ underestimates the value of σ, but it is not too far from its real value. Note that the curve of $\sigma_m^U(h)$ closest to the real σ is the one calculated for $m = 2$ (top curve). Moreover, the range of h values where σ can be esti-
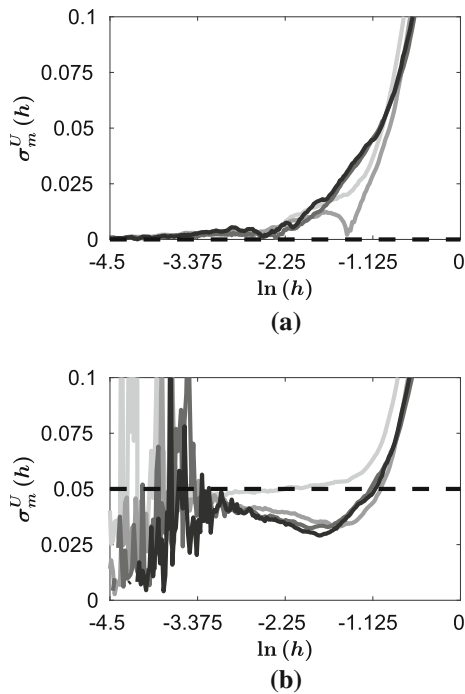
Fig. 2 Henon map: Noise level coarse-grained estimator for $m = \{2, 4, 6, 8\}$, the curves for different m values are color-coded in grayscale where the lightest gray corresponds to $m = 2$. **a** Noiseless. **b** $\sigma = 0.05$. The true value of the noise level is shown in dashed black line
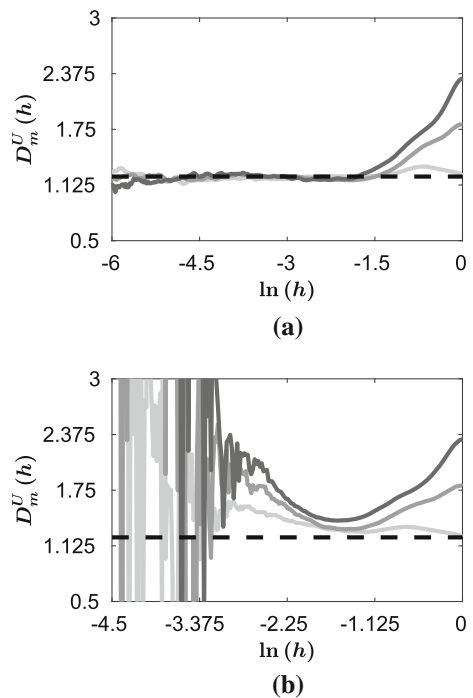
Fig. 3 Henon map: Correlation dimension coarse-grained estimator for $m = \{4, 6, 8\}$, the curves for different m values are color-coded in grayscale where the lightest gray corresponds to $m = 4$. **a** Noiseless. **b** $\sigma = 0.05$. The reported value of the correlation dimension is shown in dashed black line

mated is larger for $m = 2$ than for the other m values. This suggests that a coarse-grained estimator based on the GCI could be better than one based on the UGI for noise level estimation.

2.3 Correlation dimension coarse-grained estimator

Here we present a new coarse-grained estimator for D (see Appendix A for deduction) which is given by:

$$D_m^U(h) = \dot{D}_m^{\beta=m}(h) + \frac{\Delta_m^U(h)}{1 - \Delta_m^U(h)} \left[\dot{D}_m^{\beta=m}(h) + 2(m-1) \left(\frac{U_m^{\beta=m-2}(h)}{U_m^{\beta=m}(h)} - 1 \right) \right]. \quad (2.3)$$

The calculation of $D_m^U(h)$ requires the estimation of $\Delta_m^U(h)$ and two correlation integrals: $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$. In [24] we have shown that if $\sigma \ll h$, then $\dot{D}_m^{\beta=m}(h) \to D$. Moreover, under the same condition $\Delta_m^U(h) \to 0$. Replacing in Eq. (2.3) we have that:

$$D_m^U(h) \to \dot{D}_m^{\beta=m}(h) \to D, \quad \text{for } \sigma \ll h.$$

This means that, in absence of noise, this coarse-grained estimator is the logarithmic derivative of the UCI, which tends to D. On the other hand, if noise is present, the second term in the right-hand side of Eq. (2.3) helps to correct the estimation.

Since the GCI is a special case of the UCI for $\beta = 2$. The coarse-grained estimator of D proposed by Nolte et al. in [22] can be easily derived from Eq. (A.7) (see Appendix A.1).

In Fig. 3, the estimations of $D_m^U(h)$ for both a clean and a noisy ($\sigma = 0.05$) realizations of the Henon map are shown for $m = \{4, 6, 8\}$. In absence of noise, it can be seen in Fig. 3a that the estimator $D_m^U(h)$ oscillates near the reported value of correlation dimension ($D = 1.22$) for a wide range of scales, regardless the value of m. In contrast, Fig. 3b shows that $D_m^U(h)$ overestimates D. Furthermore, the best estimation is reached with the lowest value of the embedding dimension, $m = 4$. This is related with the previously discussed scaling

behavior of $\Delta_m^U(h)$. However, it can be observed that the values of $D_m^U(h)$ are close to the reported value of D.

2.4 Correlation entropy coarse-grained estimator

The here proposed coarse-grained entropy estimator is defined as (see Appendix B):

$$
\begin{aligned}
K_m^U(h) = & -\frac{1}{2\Delta t\tau}\ln\left(\frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)}\right) \\
& -\frac{1}{2\Delta t\tau}\ln\left(\frac{D_m^U(h)}{m}+1\right) \\
& +\frac{1}{2\Delta t\tau}\ln\left[\Delta_m^U(h)\left(\frac{m-\dot{D}_m^{\beta=m}(h)}{m-D_m^U(h)}\right)\right. \\
& \left.+\left(1-\Delta_m^U(h)\right)\left(\frac{\dot{D}_m^{\beta=m}(h)}{m}+1\right)\right].
\end{aligned} \tag{2.4}
$$

This coarse-grained estimator depends on the correlation integrals $U_{m+2}^{\beta=m+2}(h)$ and $U_m^{\beta=m}(h)$, $\Delta_m^U(h)$, the logarithmic derivative $\dot{D}_m^{\beta=m}(h)$, and the coarse-grained estimator $D_m^U(h)$. When $\sigma \ll h$ the quantities $D_m^U(h) \to \dot{D}_m^{\beta=m}(h) \to D$, and taking into account that for $\sigma \ll h$ [24]:

$$
\frac{1}{2\Delta t\tau}\ln\left[\frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)}\right] \approx \frac{1}{2\Delta t\tau}\ln\left(\frac{D}{m}+1\right)-K_2,
$$

then it can be easily shown that:

$$
K_m^U(h) \to K_2, \quad \text{for } \sigma \ll h \text{ and } m \gg D.
$$

The estimations of $K_m^U(h)$ from both a clean and a noisy realization ($\sigma = 0.05$) of the Henon map for $m = \{4, 6, 8\}$ are presented in Fig. 4. It can be observed that in both cases (Fig. 4a, b) $K_m^U(h)$ is near to the reported value of $K_2 = 0.3$ for all m values. This is the main advantage of the UCI in comparison with other correlation integrals [24].

The last statement can be confirmed by comparing $K_m^U(h)$ with two similar coarse-grained estimators based on different correlation integrals. The first one was proposed by Jayawardena et al. [17], and it is based on the Grassberger and Procaccia correlation integral ($C_m(h)$). The second one was proposed by Nolte et
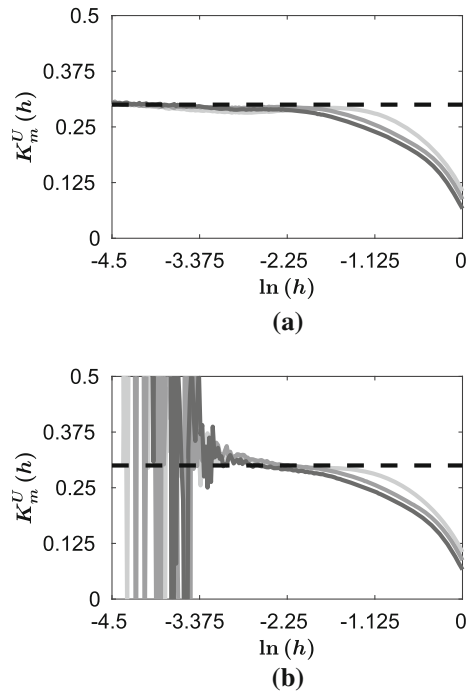

Fig. 4 Correlation entropy coarse-grained estimator for $m = \{4, 6, 8\}$, the curves for different m values are color-coded in grayscale where the lightest gray corresponds to $m = 4$. **a** Noiseless. **b** $\sigma = 0.05$. The reported value of the correlation entropy is shown in dashed black line

al., and it is calculated from the GCI ($T_m(h)$). These approaches, as well as the here proposed, $K_m^U(h)$ only requires the estimation of two correlation integrals of the same kind.

In Fig. 5, the estimations of K_2 for the Henon map through the estimators proposed by Jayawardena et al., Nolte et al. and $K_m^U(h)$ can be compared. They were calculated from a time series of length $N = 5000$ with added white Gaussian noise at a level $\sigma = 0.05$. The correlation sums $C_m(h)$, $T_m(h)$, $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$ were calculated for $m = \{4, 6, \ldots, 22\}$ with $\tau = 1$. Figure 5a shows the results for the estimator proposed by Jayawardena et al. It can be observed that, as m is increased, this estimator slowly approaches to the reported value of K_2. However, it must be pointed out that, regardless that the dimension of this map is low ($D = 1.22$), high values of m are needed in order to converge. On the other hand, it can be seen in Fig. 5b that the results of the coarse-grained estimator proposed by Nolte et al. also converge slowly to the reported value for K_2. In contrast, in Fig. 5c it can be observed that the estimator $K_m^U(h)$ approaches faster to
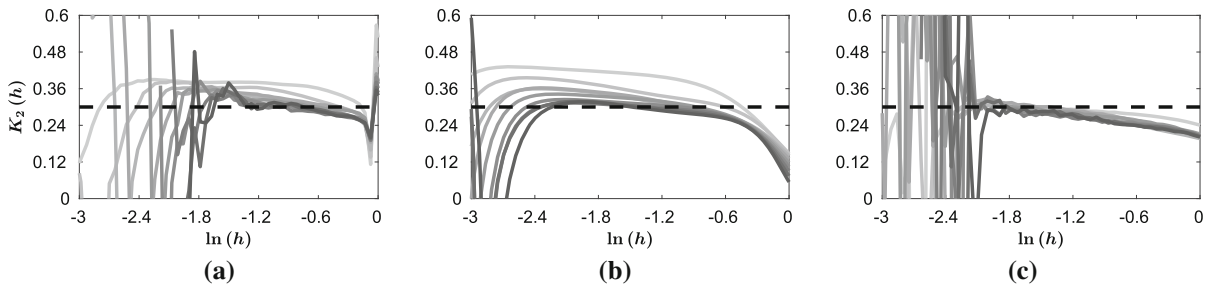
Fig. 5 Henon map with noise ($\sigma = 0.05$). Estimation of K_2 for $m = \{4, 6, \ldots, 22\}$ through different coarse-grained estimators, the curves for different m values are color-coded in grayscale where the lightest gray corresponds to $m = 4$. **a** Jayawardena et al. **b** Nolte et al. **c** $K_m^U(h)$. The reported value $K = 0.3$ is shown in dashed black line

Fig. 6 Rössler system. Estimation of K_2 for $m = \{4, 6, \ldots, 22\}$ through different coarse-grained estimators, the curves for different m values are color-coded in grayscale where the lightest gray corresponds to $m = 4$. **a** Jayawardena et al. **b** Nolte et al. **c** $K_m^U(h)$. The reported value $K = 0.066$ is shown in dashed black line
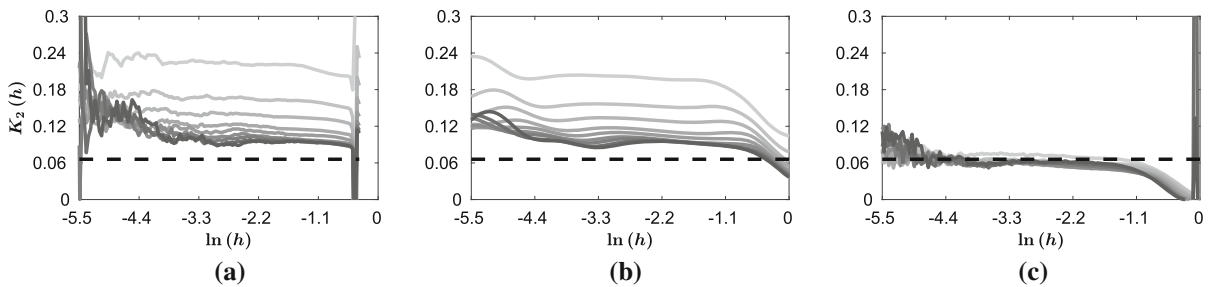
the value of K_2 than the other two studied estimators. Note that the curves for different m values of $K_m^U(h)$ are very close to each other, meaning that this estimator is more robust to changes in the parameter m. Please note that the curves corresponding to $m = \{4, 6, 8\}$ of Fig. 5c are also shown in Fig. 4b.

The same methodology was applied to a clean time series obtained from the Rössler system:

$$\begin{cases} \dot{x} &= -y - z \\ \dot{y} &= x + ay \\ \dot{z} &= b + z(x - c) \end{cases},$$

where $a = 0.15$, $b = 0.2$ and $c = 10$. The system was solved using the Runge–Kutta method with time step 0.5 and initial condition $x(0) = 0.1$, $y(0) = 0.1$, $z(0) = 0.1$. The first 20,000 data points were discarded, and the next $N = 10,000$ data points were taken to estimate the attractor's invariants. The correlation sums $C_m(h)$, $T_m(h)$, $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$ were calculated for $m = \{4, 6, \ldots, 22\}$ with $\tau = 3$.

The results for the estimators proposed by Jayawardena, Nolte, and $K_m^U(h)$ are shown in Figs. 6a–c, respectively. In all three cases, a typical scaling behavior can be observed. Once again, the estimators proposed by Jayawardena and Nolte need high values of m, in contrast with $K_m^U(h)$. Interestingly, all estimators converge to values of correlation entropy that differs from the one reported ($K_2 = 0.066$) [6]. Both the Jayawardena's and Nolte's estimators converge to $\hat{K}_2 \approx 0.085$, while $K_m^U(h)$ converges to $\hat{K}_2 \approx 0.052$. From these results, we can conclude that the estimator $K_m^U(h)$ converges by using low values of m, in comparison with the previously proposed estimators [17,22].

3 Automatic invariants estimation

In practical applications, it is important to verify the existence of a scaling behavior for a range of h. The determination of a stable scaling region, which is required to estimate the invariants, is usually done by visual inspection. However, that is not feasible when
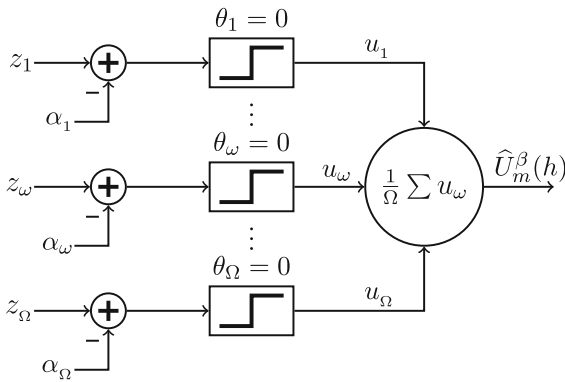
Fig. 7 Noise-assisted correlation algorithm. The ω-th comparator ($\omega = 1, 2, \ldots, \Omega$) with threshold $\theta_\omega = 0$ receives as input the squared distance between two m-dimensional delay vectors minus an i.i.d. realization $\alpha_\omega \sim \left(h^2/2\right) \chi_\beta^2$. All binary outputs μ_ω are averaged to calculate the U-correlation sum $\widehat{U}_m^\beta(h)$

large sets of data are studied Additionally, the subjective judgment should be avoided. The necessity of an algorithm for the automatic estimation of invariants arises, once the scaling regime has been confirmed. In this section, we propose an algorithm (see Algorithm 2) based on the coarse-grained estimators $\sigma_m^U(h)$, $D_m^U(h)$ and $K_m^U(h)$, that is able to automatically estimate these invariants. Before describing this algorithm, we need to detail some aspects about the calculation of the U-correlation integral.

3.1 Noise-assisted correlation algorithm

The UCI is approximated with U-correlation sum $\widehat{U}_m^\beta(h)$. The calculation of $\widehat{U}_m^\beta(h)$ requires a computationally expensive numerical integration due to the kernel of the UCI. However, we can use the noise-assisted correlation algorithm instead [24] (see Fig. 7). This algorithm begins by calculating the squared distances between each pair of delay vectors $z_\omega = \left\| x_i - x_j \right\|^2$, where $\omega = \{(i, j) \,/\, i = 1, 2, \ldots L, \ j = 1, 2, \ldots, L, \ i \neq j\}$ and $\Omega = L(L-1)$. Then, for each squared distance z_ω an i.i.d. noise realization $\alpha_\omega \sim \text{Gamma}\left(\beta/2, h^2\right)$ is subtracted. Here, $\beta/2$ and h^2 are the shape and scale parameters, respectively. The result is then compared with a threshold $\theta_\omega = 0$ to produce a binary output u_ω. Then $\widehat{U}_m^\beta(h)$ is the average of all binary outputs. These steps must be repeated for each value of h and m. Algorithm 1 describes all the steps that must be followed in order to esti-

---

Algorithm 1 Noise-assisted U-correlation algorithm.

1: Form the m-dimensional delay vectors from the temporal series and calculate all pairwise squared distances z_ω with $1 \leq \omega \leq \Omega = L(L-1)$.
2: Fix the value of the parameter h and obtain Ω realization of noise α_ω from the distribution $\left(h^2/2\right) \chi_\beta^2$.
3: Calculate the binary variable:

$$u_\omega(h) = \begin{cases} 0 & \text{if} \quad z_\omega - \alpha_\omega \geq 0 \\ 1 & \text{otherwise} \end{cases}.$$

4: Calculate the U-correlation sum as:

$$\widehat{U}_m^\beta(h) = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} u_\omega.$$

5: Repeat steps 1-4 for all values of h and m.

---

mate the noise-assisted U-correlation sum. The noise realizations α_ω were generated using χ_β^2 realizations given that, if $X \sim \chi_\beta^2$, then $\left(h^2/2\right) X \sim \text{Gamma}\left(\beta/2, h^2\right)$.

As an example, the correlation integral $U_m^{\beta=m}(h)$ can be approximated by $\widehat{U}_m^{\beta=m}(h)$ which in turn is calculated from the squared distances z_ω of m-dimensional delay vectors, and noise realizations α_ω that follow a Gamma distribution with shape parameter equal to $m/2$. Similarly, the correlation sum $\widehat{U}_{m+2}^{\beta=m-2}(h)$ can be obtained from the squared distances z_ω of $(m+2)$-dimensional delay vectors and noise realizations α_ω following a Gamma distribution with shape parameter equal to $m/2 - 1$.

3.2 Algorithm for automatic estimation of invariants

The automatic algorithm to estimate invariants begins by approximating the U-correlation integrals $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$, for different m values ($m > 2$) using the noise-assisted correlation algorithm (Algorithm 1). This step does not requires a narrow range of h as initial guess. For example, for a normalized time series the range $10^{-8} \leq h \leq 10$ is a good choice. On the other hand, note that the expression $U_m^{\beta=m-2}(h)$ requires values of $m > 2$ since Gamma distribution's shape parameter must be greater than zero.

Next, the logarithmic derivatives $\dot{D}_m^{\beta=m}(h)$, $\dot{D}_{m+2}^{\beta=m}(h)$ are computed and $\Delta_m^U(h)$ is calculated using Eq. (2.1). To calculate the logarithmic derivatives, we make use of a wavelet transform approach to approximate the derivative [21]. This allows us to achieve better

estimations of the invariants because it implements a low-pass filter reducing the high-frequency oscillations that are naturally present in estimators derived from the noise-assisted algorithm [24]. Please observe that correlation integral $U_{m+2}^{\beta=m}(h)$ is equal to $U_{\hat{m}}^{\beta=\hat{m}-2}(h)$ correlation integral evaluated at $\hat{m} = m + 2$. In order to obtain σ, the coarse-grained estimator $\sigma_m^U(h)$ must be calculated [Eq. (2.2)].

To obtain a suitable range of scale values where the invariants can be estimated, one must look for a range of h where the coarse-grained estimator is nearly constant and its variation across the different m values is the smallest. In this sense, for the noise level we define the functions:

$$A_\sigma(h) = \frac{1}{M} \sum_{i=1}^{M} \frac{d\,\sigma_{m_i}^U(h)}{d\ln h},$$

$$V_\sigma(h) = \frac{1}{M-1} \sum_{i=1}^{M} \left(\sigma_{m_i}^U(h) - \hat{\sigma}^U(h) \right)^2,$$

and

$$F_\sigma(h) = A_\sigma(h)\,V_\sigma(h),$$

where $m \in \{m_1, m_2, \ldots, m_i, \ldots, m_M\}$, $\hat{\sigma}^U(h)$ is the average of $\sigma_m^U(h)$ across m. $A_\sigma(h)$ is the average over m of the derivative of $\sigma_m^U(h)$ with respect to $\ln h$, V_σ gives the variation of $\sigma_m^U(h)$ across m and $F_\sigma(h)$ is the product of the aforementioned functions. We propose to estimate σ within a range of h centered at the h value at which $F_\sigma(h)$ is minimum. In this way, σ is estimated in a range of h centered in a plateaued region (scaling region) of $\sigma_m^U(h)$, and its value is consistent through the parameter m.

The correlation dimension is determined using the coarse-grained estimator $D_m^U(h)$ [Eq. (2.3)]. To find a range of h values to estimate D, we define the functions $A_D(h)$, $V_D(h)$ and $F_D(h)$ similarly to the ones above but using the estimator $D_m^U(h)$ instead of $\sigma_m^U(h)$. The estimation of D must be done within a range of h centered at the h value at which $F_D(h)$ is minimum.

Finally, the correlation entropy can be estimated using the coarse-grained estimator $K_m^U(h)$. As it can be seen in Eq. (2.4), this estimator depends on $\Delta_m^U(h)$, $\dot{D}_m^{\beta=m}(h)$, the coarse-grained estimator $D_m^U(h)$ and the ratio between $U_m^{\beta=m}(h)$ and $U_{m+2}^{\beta=m+2}(h)$. K_2 must be

Algorithm 2 Automatic estimation of attractors' invariants.

1: Calculate $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$ using Algorithm 1 for $m > 2$.
2: Calculate $\Delta_m^U(h)$ using Eq. (2.1) and the UGI obtained in step 1.
3: Compute $\sigma_m^U(h)$ with Eq. (2.2) and obtain the function $F_\sigma(h)$. Estimate σ within a range of h centered at the value h where $F_\sigma(h)$ is minimum.
4: Use $U_m^{\beta=m}(h)$, $U_m^{\beta=m-2}(h)$ and $\Delta_m^U(h)$ to calculate $D_m^U(h)$ Eq. (2.3) and obtain the function $F_D(h)$. Estimate D within a range of h centered at the value h where $F_D(h)$ is minimum.
5: Obtain $K_m^U(h)$ (Eq. 2.4) using $U_m^{\beta=m}(h)$, $\Delta_m^U(h)$ and $D_m^U(h)$. Calculate the function $F_{K_2}(h)$. Estimate K_2 within a range of h centered at the value h where $F_D(h)$ is minimum.

---

determined within a range of h centered at the h value at which $F_{K_2}(h)$ is minimum.[1]

3.3 Simulations

The aforementioned methodology was applied to a set of 128 realization of the Henon map with different noise levels $\sigma = \{0, 0.05, 0.2\}$ and data lengths which are proportional to the number L of available delay vectors to estimate the invariants σ, D and K_2. To calculate the UCIs for $m = \{4, 5, \ldots, 8\}$ all realizations were normalized (unitary standard deviation), and the nearest 15 temporal neighbors of each delay vector were discarded [19]. Figure 8 shows the boxplots of the invariants σ, D and K_2 for $L = \{500, 1000, 3000, 5000, 10,000\}$. Unfortunately, there is no other automatic algorithm for the estimation of invariants. Due to this reason, we only can compare the results of our estimations with previously reported values calculated by selecting the scaling range by visual inspection. For the noise level estimation (first row of Fig. 8), it can be observed that, in absence of noise, our algorithm overestimates the value of σ (dashed black line). On the contrary, for higher noise levels ($\sigma = \{0.05, 0.2\}$) the proposed method slightly underestimates σ. It can be noticed that, regardless the noise level, the median of each box is very close to the true value of σ and the variance of the estimation decreases as the number of data points increases. The estimation of the correlation dimension is shown in the

---

[1] The functions $F_D(h)$ and $F_{K_2}(h)$ can be computed similarly to $F_\sigma(h)$ but using the coarse-grained estimators $D_m^U(h)$ and $K_m^U(h)$, respectively.
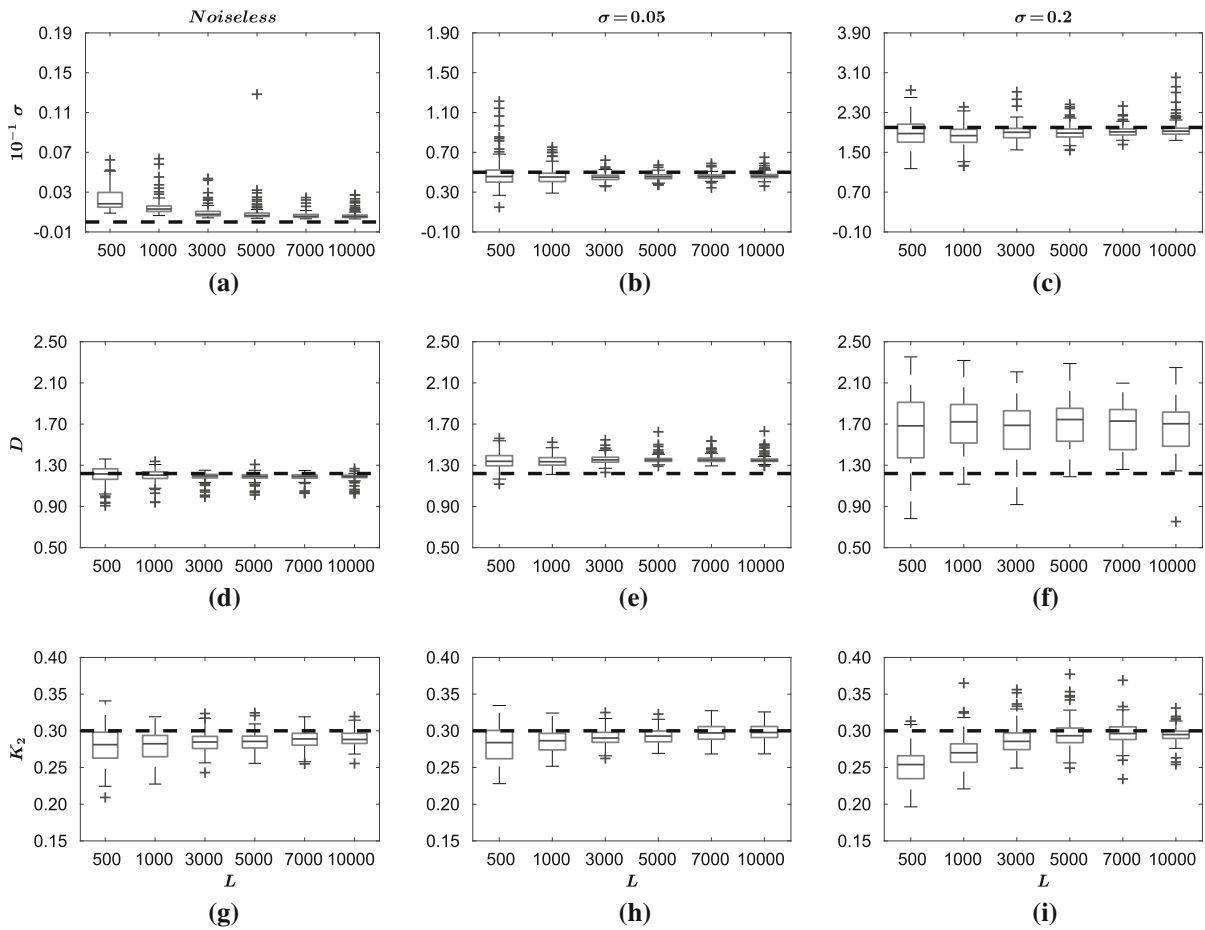
Fig. 8 Henon map. Box plot of the estimation of σ, D, and K_2 for different numbers of available delay vectors $L = \{500, 1000, 3000, 5000, 10,000\}$. First row: estimation of σ. **a** noiseless, **b** $\sigma = 0.05$, and **c** $\sigma = 0.2$. Second row: estima- tion of D. **d** noiseless, **e** $\sigma = 0.05$, and **f** $\sigma = 0.2$. Third row: estimation of K_2. **g** noiseless, **h** $\sigma = 0.05$, and **i** $\sigma = 0.2$. The reported values are shown in dashed black line

second row of Fig. 8. It can be seen in Fig. 8d that for noiseless time series the estimation of D is very close to the reported value ($D = 1.22$) and its variance decreases as the length of the time series is increased. For higher noise levels, the correlation dimension is slightly biased, but the estimations are still close to the previously reported value.

The results for the correlation entropy are shown in the third row of Fig. 8. In general, the here pro- posed method yields correlation entropy values con- verging to values lesser than the reported for this map ($K_2 = 0.3$).

The proposed approach was also tested on the Mackey–Glass time series, which is generated by the following nonlinear time-delay differential equation:

$$\dot{x}(t) = \frac{ax(t-\lambda)}{1 + [x(t-\lambda)]^{10}} - bx(t),$$

where $a = 0.2$, $b = 0.1$, and $\lambda = 23$. We estimate the invariants σ, D, and K_2 from 128 realizations with dif- ferent initial conditions and normalized to have unitary standard deviation. The embedding dimensions were $m = \{4, 6, \ldots, 10\}$ and the embedding lag τ was set to the first local minimum of the mutual information function ($\tau = 20$) [18]. Finally, the nearest 15 tempo- ral neighbors of each delay vector were discarded.

The results of these simulations are presented in Fig. 9, and the estimations of the noise level are pre- sented in Fig. 9a–c. In absence of noise, the algorithm estimates small positive values of σ. This behavior is
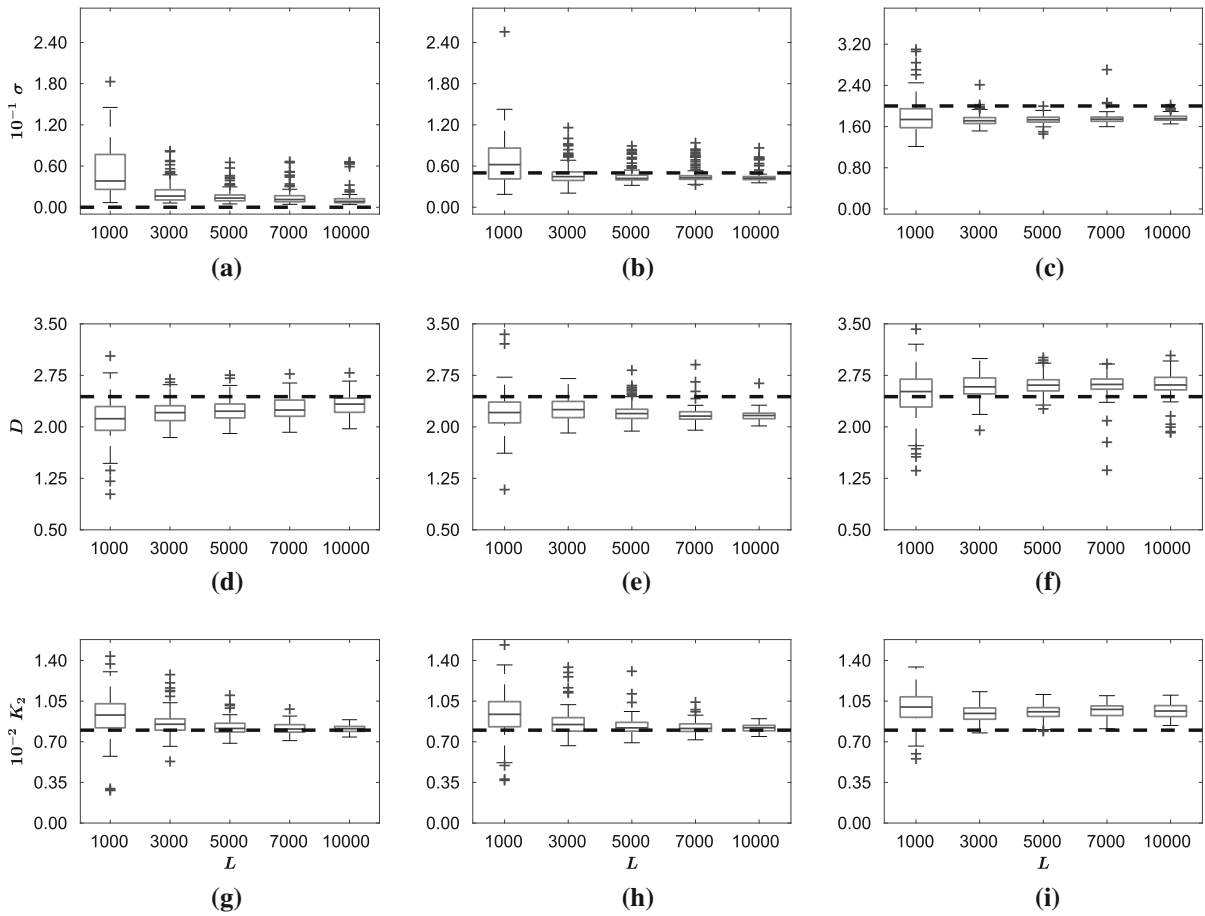
Fig. 9 Mackey–Glass box plot of the estimation of σ, D, and K_2 for different numbers of available delay vectors, $L = \{1000, 3000, 5000, 10{,}000\}$. First row: estimation of σ. **a** noiseless, **b** $\sigma = 0.05$, and **c** $\sigma = 0.2$. Second row: esti- mation of D. **d** noiseless, **e** $\sigma = 0.05$, and **f** $\sigma = 0.2$. Third row: estimation of K_2. **g** noiseless, **h** $\sigma = 0.05$, and **i** $\sigma = 0.2$. The reported values are shown in dashed black line

explained because the function $\Delta_m^U(h)$ tends to zero as the length of the time series increases, producing more accurate estimations with less variance. In presence of noise ($\sigma = 0.05$ and $\sigma = 0.2$), the here proposed algorithm underestimates the noise level, but in Fig.8 it can be seen how the median values tend to the real value of σ as L increases.

The estimations of the correlation dimension are shown in the second row of Fig. 9. For clean time series, our algorithm provides estimations of D very close to the reported value $D = 2.44$ [13]. As the data length is increased, the estimations approach the reported value. In case of noisy time series, the estimations of the cor- relation dimension are greater than those obtained in

the noiseless case. Nevertheless, the estimation of D is still very good.

In Fig. 9g–i, the estimations for the correlation entropy are presented. Our algorithm achieves sat- isfactory estimations of K_2 (reported value $K_2 = 0.008$) [14], even in presence of noise. It can be observed that regardless the value of L the estimation of K_2 is consistent, but its variance decreases as the data length increases.

4 Discussion

In this section, we review the main contributions of this article and give some practical recommendations

for its implementation. In this work, we introduce a new methodology to estimate the correlation dimension, the correlation entropy and the noise level of time series. This approach is based on the recent proposed U-correlation integral (Eq. 1.3), introduced in [24], which uses a kernel function that depends on the embedding dimension. From the scaling function of the UCI (Eq. 1.5), we have developed a set of coarse-grained estimators for the correlation dimension (Eq. 2.3), the correlation entropy (Eq. 2.4) and the noise level (Eq. 2.2). The advantage of these coarse-grained estimators is that they only depend on the estimation of two U-correlation integrals. In other words, they do not need of the tuning of any external parameter and/or the estimation of the noise level, as it does other approaches [6,16,24].

From the results presented in the previous section, we can conclude that the here proposed algorithm (Algorithm 2) is able to automatically select the scaling region and to estimate the invariants. As far as we know, there is no other algorithms with these features reported in the literature. For this reason, our results were compared against reported estimations that were obtained selecting the scaling region by visual inspection.

Regarding the coarse-grained estimators, the most outstanding result was obtained with $K_m^U(h)$. It is more accurate than previously proposed estimators and needs lesser values of m to converge.

Finally, it is important to recall that the implementation of our method only requires the calculation of two correlation integrals: $U_m^{\beta=m}(h)$ and $U_m^{\beta=m-2}(h)$, that can be obtained through the noise-assisted correlation algorithm (Algorithm 1). It is also important to say that the time series must be normalized to have unitary standard deviation. As it was reported in [24], the correlation integrals computed using the noise-assisted correlation algorithm present oscillations for small values of h. This phenomenon makes difficult to look for scaling regions over small h values. In order to avoid this difficulty, three strategies can be followed: The first one is to increase the number of comparators Ω in the noise-assisted correlation algorithm by increasing the data length. This will make coarse-grained curves smoother and also will improve the convergence of these estimators. The second one is to increase Ω by making copies of the squared distances between delay vectors z_ω and adding different noise realizations μ_ω to them. This will not improve the convergence of any coarse-grained

estimators to the invariant since no new information is added, but the ripple in the coarse-grained curves will be reduced. The third one, and the one we have found as the more adequate, is to use the algorithm proposed in [21] to calculate the logarithmic derivatives of the UCI. This algorithm calculates a smoothed version of the derivative of a function using the wavelet transform.

5 Conclusions

The original contributions of this document are twofold: both a new method for the estimation of invariants based on the recently proposed U-correlation integral, and an algorithm for the automatic selection of the scaling region. The combined use of these two ideas allows to automatically estimate the noise level, correlation dimension and correlation entropy of a dynamical system. This approach have been statistically tested using synthetic data coming from low-dimensional systems with different data lengths and noise levels. The results suggest that our algorithm provides robust estimations of the correlation dimension, the correlation entropy and the noise level.

A Analytic deduction for the coarse-grained D estimator

Let:

$$a = \frac{\beta + m}{2}; \quad b = \frac{m - D}{2}; \quad c = \frac{m + 2}{2};$$

$$y = -\frac{h^2}{4\sigma^2}; \quad \frac{dy}{dh} = -\frac{h}{2\sigma^2}, \quad \text{(A.1)}$$

and

$$P = (-1)^{m/2} \frac{\hat{\phi}}{2} e^{-m\Delta t \tau K_2} \frac{\Gamma(D/2)\,\Gamma((\beta+m)/2)}{\Gamma(\beta/2)\,\Gamma(m/2+1)}$$

then Eq. (1.5) can be rewritten as:

$$U_m^\beta(y) = P y^{m/2} F(a, b; c; y) \quad \text{(A.2)}$$

We need to find the logarithmic derivative of Eq. (A.2). For this end we can find its first derivative as:

$$\frac{d\left[U_m^\beta(y)\right]}{dh} = \frac{d\left[Py^{m/2}F(a,b;c;y)\right]}{dy}\frac{dy}{dh}$$

using Eq. 15.2.1 from Ref. [1]:

$$\frac{d\left[U_m^\beta(y)\right]}{dh} = -\frac{Ph}{2\sigma^2}y^{m/2}\left[\frac{m}{2}F(a,b;c;y)\right.$$
$$\left. + \frac{ab}{c}F(a+1,b+1;c+1;y)\right]$$

Then the logarithmic derivative can be found as:

$$\frac{d\left[\ln U_m^\beta(y)\right]}{d\ln h} = \frac{h}{U_m^\beta(y)}\frac{dU_m^\beta(y)}{dh}$$
$$= m + \frac{2aby}{c}\frac{F(a+1,b+1;c+1;y)}{F(a,b;c;y)}, \qquad (A.3)$$

and for simplicity in the notation:

$$\dot{D}_m^\beta(y) = \frac{d\left[\ln U_m^\beta(y)\right]}{d\ln h}.$$

Using the identity [10, Eq. 9.137.9]:

$$\frac{aby}{c}F(a+1,b+1;c+1;y)$$
$$= \frac{1}{1-y}\left[(c-a)F(a-1,b;c;y)\right.$$
$$\left. + (a-c+by)F(a,b;c;y)\right] \qquad (A.4)$$

we can rewrite Eq. (A.3) as:

$$\dot{D}_m^\beta(y) = m + \frac{2}{1-y}$$
$$\left[(a-c+by) + (c-a)\frac{F(a-1,b;c;y)}{F(a,b;c;y)}\right] \qquad (A.5)$$

From Eq. (1.5) and using the definitions in (A.1) it is possible to demonstrate that:

$$\frac{U_m^{\beta-2}(y)}{U_m^\beta(y)} = \frac{\beta-2}{2(a-1)}\frac{F(a-1,b;c;y)}{F(a,b;c;y)}. \qquad (A.6)$$

Substituting Eq. (A.6) into (A.5):

$$\dot{D}_m^\beta(y) = m + \frac{2}{1-y}$$
$$\left[(a-c+by) + \frac{2(c-a)(a-1)}{\beta-2}\frac{U_m^{\beta-2}(y)}{U_m^\beta(y)}\right]$$

Restoring the values of β, m and D and clearing for D

$$D = \dot{D}_m^\beta(y) - \frac{1}{y}\left[\dot{D}_m^\beta(y) + (m+\beta-2)\left(\frac{U_m^{\beta-2}(y)}{U_m^\beta(y)} - 1\right)\right]$$

notice that $-1/y = 4\sigma^2/h^2$ and from Eq. (2.1)

$$\frac{h^2}{4\sigma^2} = \frac{\Delta_m^U(h)}{1-\Delta_m^U(h)},$$

then

$$D = \dot{D}_m^\beta(h) + \frac{\Delta_m^U(h)}{1-\Delta_m^U(h)}\left[\dot{D}_m^\beta(h)\right.$$
$$\left. + (m+\beta-2)\left(\frac{U_m^{\beta-2}(h)}{U_m^\beta(h)} - 1\right)\right]. \qquad (A.7)$$

Making $\beta = m$, the coarse-grained estimator for D can be defined as:

$$D_m^U(h) = \dot{D}_m^{\beta=m}(h) + \frac{\Delta_m^U(h)}{1-\Delta_m^U(h)}\left[\dot{D}_m^{\beta=m}(h)\right.$$
$$\left. + 2(m-1)\left(\frac{U_m^{\beta=m-2}(h)}{U_m^{\beta=m}(h)} - 1\right)\right]. \qquad (A.8)$$

A.1 Relation with Nolte's et al. coarse-grained D estimator:

It was proved in [24] that the Gaussian correlation integral is a particular case of the U-correlation integral that arises when β is constant and equal to 2. Let $\beta = 2$, then from Eq. (A.6) it can be proved that:

$$\frac{U_m^{\beta-2}(h)}{U_m^\beta(h)} = 0,$$

for all m values. Thus, Eq. (A.7) can be rewritten as:

$$D = \dot{D}_m^{\beta=2}(h) - \frac{\Delta_m^U(h)}{1 - \Delta_m^U(h)}\left(m - \dot{D}_m^{\beta=2}(h)\right). \tag{A.9}$$

Finally, given that the Gaussian correlation integral is a particular case of the U-correlation integral for $\beta = 2$, i.e., $U_m^{\beta=2}(h) = T_m(h)$, we can write:

$$\dot{D}_m^{\beta=2}(h) = \frac{\mathrm{d}\left[\ln T_m(h)\right]}{\mathrm{d}\ln h}.$$

Then, Eq. (A.9) is the same coarse-grained estimator for the correlation dimension that was derived by Nolte et al. in [22].

B Analytic deduction for the coarse-grained K_2 estimator

The analytic deduction of the coarse-grained estimator for K_2 begins by defining the quantity:

$$M = \frac{\hat{\phi}}{2}\left(2\sigma^2\right)^D \Gamma(D/2).$$

Substituting this equation and Eq. (A.1) into Eq. (1.5) we can write:

$$U_m^\beta(y) = M\frac{\Gamma(a)}{\Gamma(\beta/2)\,\Gamma(c)}e^{-m\Delta t\tau K_2}(-y)^{m/2}\,F(a, b; c; y)$$

Then the ratio $U_{m+2}^{\beta+2}(y)/U_m^\beta(y)$ can be written as:

$$\frac{U_{m+2}^{\beta+2}(y)}{U_m^\beta(y)} = \frac{2a(a+1)}{\beta c}e^{-2\Delta t\tau K_2}(-y)$$
$$\frac{F(a+2, b+1; c+1; y)}{F(a, b; c; y)}.$$

Making $\alpha = a + 1$:

$$\frac{U_{m+2}^{\beta+2}(y)}{U_m^\beta(y)} = \frac{2(\alpha-1)\alpha}{\beta c}e^{-2\Delta t\tau K_2}$$
$$(-y)\frac{F(\alpha+1, b+1; c+1; y)}{F(\alpha-1, b; c; y)}. \tag{B.1}$$

The identity in Eq. (A.4) can be rewritten as:

$$\frac{\alpha(-y)}{c}\frac{F(\alpha+1, b+1; c+1; y)}{F(\alpha-1, b; c; y)} = \frac{-1}{b(1-y)}\left[(c-\alpha)\right.$$
$$\left. + (\alpha-c+by)\frac{F(\alpha, b; c; y)}{F(\alpha-1, b; c; y)}\right]. \tag{B.2}$$

Substituting Eq. (B.2) into Eq. (B.1) and restituting a:

$$\frac{U_{m+2}^{\beta+2}(y)}{U_m^\beta(y)} = \frac{-2a}{\beta b(1-y)}e^{-2\Delta t\tau K_2}\left[(c-a-1)\right. +$$
$$\left. + (a+1-c+by)\frac{F(a+1, b; c; y)}{F(a, b; c; y)}\right]. \tag{B.3}$$

Using the identity [10, Eq. 9.137.12]:

$$\frac{by}{c}\frac{F(a+1, b+1; c+1; y)}{F(a, b; c; y)} + 1 = \frac{F(a+1, b; c; y)}{F(a, b; c; y)}. \tag{B.4}$$

Equation (B.3) can be written as:

$$\frac{U_{m+2}^{\beta+2}(y)}{U_m^\beta(y)} = \frac{e^{-2K_2\Delta t\tau}}{2b\beta(y-1)}\left[4a(c-a-1)\right. +$$
$$\left. + 2(a+1-c+by)\left(\frac{2aby}{c}\frac{F(a+1, b+1; c+1; y)}{F(a, b; c; y)} + 2a\right)\right]. \tag{B.5}$$

Using Eq. (A.5)

$$\frac{U_{m+2}^{\beta+2}(y)}{U_m^\beta(y)} = \frac{e^{-2K_2\Delta t\tau}}{2b\beta(y-1)}\left[4a(c-a-1)\right. +$$
$$\left. + 2(a+1-c+by)\left(\dot{D}_m^\beta(y) - m + 2a\right)\right]. \tag{B.6}$$

Replacing the values of a, c and b:

$$\frac{U_{m+2}^{\beta+2}(y)}{U_m^\beta(y)} = \frac{e^{-2K_2\Delta t\tau}}{2b\beta(y-1)}\left[-\beta(m+\beta)+(\beta+2by)\left(\dot{D}_m^\beta(y)+\beta\right)\right]$$
$$= \frac{e^{-2K_2\Delta t\tau}}{2b(y-1)}\left[\dot{D}_m^\beta(y)-m+2by\left(\frac{\dot{D}_m^\beta(y)}{\beta}+1\right)\right]$$
$$= \frac{e^{-2K_2\Delta t\tau}}{(y-1)}\left[\frac{\dot{D}_m^\beta(y)-m}{m-D} + y\left(\frac{\dot{D}_m^\beta(y)}{\beta}+1\right)\right]. \tag{B.7}$$

It can be shown that $\Delta_m^U(h) = 1/(1-y)$ and $1 - \Delta_m^U(h) = y/(y-1)$ so:

$$\frac{U_{m+2}^{\beta+2}(y)}{U_m^\beta(y)} = e^{-2K_2 \Delta t \tau} \left[\Delta_m^U(h) \left(\frac{m - \dot{D}_m^\beta(y)}{m - D} \right) \right.$$
$$\left. + \left(1 - \Delta_m^U(h) \right) \left(\frac{\dot{D}_m^\beta(y)}{\beta} + 1 \right) \right].$$

Taking natural logarithm, subtracting the term $\ln(D/\beta + 1)$ on both sides and dividing by $2\Delta t \tau$:

$$K_2 - \frac{1}{2\Delta t \tau} \ln \left(\frac{D}{\beta} + 1 \right)$$
$$= \frac{1}{2\Delta t \tau} \left\{ \ln \left[\Delta_m^U(h) \left(\frac{m - \dot{D}_m^\beta(y)}{m - D} \right) \right.\right.$$
$$\left. + \left(1 - \Delta_m^U(h) \right) \left(\frac{\dot{D}_m^\beta(y)}{\beta} + 1 \right) \right]$$
$$\left. - \ln \left[\frac{U_{m+2}^{\beta+2}(y)}{U_m^\beta(y)} \right] - \ln \left(\frac{D}{\beta + 1} \right) \right\}. \qquad (B.8)$$

Let call the left-hand side of Eq. (B.8) as $K_m^U(h)$:

$$K_m^U(h) = \frac{1}{2\Delta t \tau} \left\{ \ln \left[\Delta_m^U(h) \left(\frac{m - \dot{D}_m^\beta(h)}{m - D} \right) \right.\right.$$
$$\left. + \left(1 - \Delta_m^U(h) \right) \left(\frac{\dot{D}_m^\beta(y)}{\beta} + 1 \right) \right]$$
$$\left. - \ln \left[\frac{U_{m+2}^{\beta+2}(h)}{U_m^\beta(h)} \right] - \ln \left(\frac{D}{\beta} + 1 \right) \right\}. \quad (B.9)$$

Note that this estimator still depends on the correlation dimension D. However, this dependency can be avoided using $D_m^U(h)$ as estimator for D. Making $\beta = m$ we can define the coarse-grained entropy estimator as:

$$K_m^U(h) = \frac{1}{2\Delta t \tau} \left\{ \ln \left[\Delta_m^U(h) \left(\frac{m - \dot{D}_m^{\beta=m}(h)}{m - D_m^U(h)} \right) \right.\right.$$
$$\left. + \left(1 - \Delta_m^U(h) \right) \left(\frac{\dot{D}_m^{\beta=m}(y)}{m} + 1 \right) \right]$$
$$\left. - \ln \left[\frac{U_{m+2}^{\beta=m+2}(h)}{U_m^{\beta=m}(h)} \right] - \ln \left(\frac{D_m^U(h)}{m} + 1 \right) \right\}.$$
$$(B.10)$$

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables. Applied Mathematics Series. Dover Publications, Mineola (1964)
2. Chan, K.S., Tong, H.: Chaos: A Statistical Perspective. Springer, Berlin (2013)
3. Choi, S.H., Zhang, Y., Jiang, J.J., Bless, D.M., Welham, N.V.: Nonlinear dynamic-based analysis of severe dysphonia in patients with vocal fold scar and sulcus vocalis. J. Voice **26**(5), 566–576 (2012)
4. Çoban, G., Büyüklü, A.H., Das, A.: A linearization based non-iterative approach to measure the Gaussian noise level for chaotic time series. Chaos Solitons Fractals **45**(3), 266–278 (2012)
5. Diks, C.: Estimating invariants of noisy attractors. Phys. Rev. E **53**(5), R4263–R4266 (1996)
6. Diks, C.: Nonlinear Time Series Analysis: Methods and Applications. Nonlinear Time Series and Chaos. World Scientific, Singapore (1999)
7. Frank, M., Blank, H.R., Heindl, J., Kaltenhäuser, M., Köchner, H., Kreische, W., Müller, N., Poscher, S., Sporer, R., Wagner, T.: Improvement of K2-entropy calculations by means of dimension scaled distances. Phys. D **65**, 359–364 (1993)
8. Gao, J., Hu, J., Liu, F., Cao, Y.: Multiscale entropy analysis of biological signals: a fundamental bi-scaling law. Front. Comput. Neurosci **9**(June), 1–9 (2015)
9. Gao, J., Hu, J., Tung, W.W.: Entropy measures for biological signal analyses. Nonlinear Dyn. **68**(3), 431–444 (2012)
10. Gradshteyn, I.S., Ryzhik, I.M.: Table of Integrals, Series, and Products, 5th edn. Academic Press, Cambridge (1994)
11. Grassberger, P., Procaccia, I.: Characterization of strange attractors. Phys. Rev. Lett. **50**, 346–349 (1983)
12. Grassberger, P., Procaccia, I.: Estimation of the Kolmogorov entropy from a chaotic signal. Phys. Rev. A **28**(4), 2591–2593 (1983)
13. Grassberger, P., Procaccia, I.: Measuring the strangeness of strange attractors. Phys. D **9**, 189–208 (1983)
14. Grassberger, P., Procaccia, I.: Dimensions and entropies of strange attractors from a fluctuating dynamics approach. Phys. D **13**, 34–54 (1984)
15. Hosseinifard, B., Moradi, M.H., Rostami, R.: Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. Comput. Methods Progr. Biomed. **109**(3), 339–345 (2013)
16. Jayawardena, A.W., Xu, P., Li, W.K.: A method of estimating the noise level in a chaotic time series. Chaos **18**, 023115 (2008)
17. Jayawardena, A.W., Xu, P., Li, W.K.: Modified correlation entropy estimation for a noisy chaotic time series. Chaos (Woodbury, N.Y.) **20**(2), 023104 (2010)
18. Kaffashi, F., Foglyano, R., Wilson, C.G., Loparo, K.A.: The effect of time delay on approximate and sample entropy calculations. Phys. D **237**(23), 3069–3074 (2008)
19. Kantz, H., Schreiber, T.: Nonlinear Time Series Analysis. Cambridge University Press, Cambridge (2004)

20. Kugiumtzis, D.: Correction of the correlation dimension for noisy time series. Int. J. Bifurc. Chaos **07**(06), 1283–1294 (1997)
21. Luo, J.W., Bai, J., Shao, J.H.: Application of the wavelet transforms on axial strain calculation in ultrasound elastography. Prog. Nat. Sci. **16**(9), 942–947 (2006)
22. Nolte, G., Ziehe, A., Müller, K.R.: Noise robust estimates of correlation dimension and K2 entropy. Phys. Rev. E **64**, 1–10 (2001)
23. Oltmans, H., Verheijen, P.: Influence of noise on power-law scaling functions and an algorithm for dimension estimations. Phys. Rev. E **56**(1), 1160–1170 (1997)
24. Restrepo, J.F., Schlotthauer, G.: Noise-assisted estimation of attractor invariants. Phys. Rev. E **94**, 012–212 (2016)
25. Small, M.: Applied Nonlinear time Series Analysis: Applications in Physics, Physiology and Finance. World Scientific, Singapore (2005)
26. Smith, R.L.: Estimating dimension in noisy chaotic time series. J. R. Stat. Soc. Ser. B (Stat. Methodol.) **54**(2), 329–351 (1992)
27. Strogatz, S.H.: Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering. Westview press, Boulder (2014)
28. Yu, D., Small, M., Harrison, R.G., Diks, C.: Efficient implementation of the Gaussian kernel algorithm in estimating invariants and noise level from noisy time series data. Phys. Rev. E **61**(4), 3750–3756 (2000)