Identification of suitable embedding dimensions and lags for time series generated by chaotic, finite-dimensional systems

Alessio Perinelli
Department of Physics, University of Trento, I-38123 Trento, Italy

Leonardo Ricci*
Department of Physics, University of Trento, I-38123 Trento, Italy
and CIMeC, Center for Mind/Brain Sciences, University of Trento, I-38068 Rovereto, Italy

(Received 6 July 2018; revised manuscript received 5 October 2018; published 27 November 2018)

In the field of nonlinear dynamics, many methods have been proposed to tackle the issue of optimally setting embedding dimension and lag in order to analyze sampled scalar signals. However, intrinsic statistical uncertainties due to the finiteness of input sequences severely hinder a general solution to the problem. A more achievable approach consists of assessing sets of dimension and lag pairs that are equivalently suitable to embed a time series. Here we present a method to identify these sets of embedding pairs, under the hypothesis that the time series of interest is generated by a chaotic, finite-dimensional dynamical system. We first introduce a "distance gauge transformation" based on the analytical forms of correlation integrals corresponding to a Gaussian white noise source. We show that in this new distance gauge, correlation integrals generated by chaotic, finite-dimensional systems are characterized by distinctive features, whose absence is a marker of the unsuitability of the underlying embedding choice. By means of a new estimator of the correlation dimension that relies on the new distance gauge, sets of suitable embedding pairs are finally identified by looking at the uniformity of the estimation. The method is completely automatic and was successfully tested on both synthetic and experimental time series. It also provides a tool to estimate the redundance and irrelevance timescales of the system that underlie the time series as well as a lower constraint to the sampling rate. The method is suitable for applications in research fields where a chaotic behavior has to be identified, such as neuroscience, geophysics, and economics.

DOI: 10.1103/PhysRevE.98.052226

I. INTRODUCTION

As a result of the pioneering works by Packard *et al.* [1] and Takens [2], properties of a dynamical system can be inferred by suitably embedding a sample, scalar time series generated by the system itself (see also [3] for a recent review on the topic). Embedding has since become an essential tool to investigate the dynamical properties of a system, although a definitive method to set optimal embedding dimension m and lag L is still lacking [4].

The search for optimal embedding was traditionally conceived as the assessment of an embedding pair (m, L) that allows for the maximization of a given measure: above all, the robustness to noise and the accuracy in the determination of the maximum Lyapunov exponent (MLE). More recently, it has been shown [5] that, especially for continuous dynamical systems (flows), this goal is thwarted by intrinsic statistical uncertainties due to the finiteness of the time series. A more achievable goal consists of assessing a set of embedding pairs that are statistically equivalent in order to optimize the chosen measure.

Conventional optimal embedding techniques depend on both "the time series and the applied measure" [4]. However, theory provides constraints that are independent of time series and measures. So it is well known that embedding must satisfy Takens' requirement $m > 2d$, where d is the box-counting dimension. The condition can be relaxed to $m > \nu$ when the correlation dimension ν is to be evaluated [6,7]. In addition it is widely accepted (see, for example, Kugiumtzis [8]) that the *embedding window* $(m-1)L$ is constrained by the necessity of avoiding, on the one side, redundance and, on the other side, irrelevance [9,10]: this corresponds to $\tau_R/T \lesssim (m-1)L \lesssim \tau_I/T$, where T is the sampling time while τ_R and τ_I represent the so-called *redundance* and *irrelevance time*, respectively. Furthermore, because a dynamical system is characterized by fixed, embedding-independent observables like MLE and correlation dimension, pairs that are suitable for embedding are to provide consistent estimates of these observables.

The topic of the present paper is a method to identify sets of embedding pairs that are possibly suitable to embed a time series generated by a chaotic, finite-dimensional dynamical system. This identification is carried out both by ruling out pairs unfit to embedding and by assessing others that comply with basic requirements posited by finite-dimensional chaos. Our approach consists of four steps. First, upon determining the correlation integral by using a Monte Carlo integration, we transform it by means of a "distance gauge" based on the analytical forms of the correlation integrals corresponding to

*leonardo.ricci@unitn.it

a Gaussian white noise (GWN) source. The gauge transformation allows us to describe any correlation integral within the square $[0, 1] \times [0, 1]$. The second step consists of carrying out a Kolmogorov-Smirnov (K-S) test [11] on a finite set of embedding pairs, henceforth referred to as a lattice. If the hypothesis of a non-GWN source holds, embedding pairs that nevertheless generate correlation integrals compatible with a GWN source are deemed to be unsuitable to embed the time series. As far as the third step is concerned, we first show that a divergent behavior of the correlation integral at small scales in the new gauge is an efficient marker of a chaotic, finite-dimensional system. Thereupon, we look at embedding pairs in which such divergent behavior shows up. Only these pairs are further dealt with in the fourth and final step, in which the correlation dimension ν is evaluated—or "mapped"—by means of a novel estimator in the new distance gauge. Embedding pairs on which the correlation dimension is essentially uniform are finally deemed to be suitable to embed the time series.

Our method does not require any subjective evaluation of parameters. By generating synthetic time series out of well known attractors, we show that the values of the correlation dimension ν estimated on sets of embedding-suitable pairs are in good agreement with data reported in previous works. In addition, our method provides estimates of the redundance and irrelevance times τ_R, τ_I as well as of an upper constraint to the sampling time T.

The paper is organized so that a section is devoted to each of the steps outlined above. The determination of the correlation integrals via Monte Carlo integration as well as the distance gauge transformation based on the analytical form of correlation integrals of a GWN source are the topic of Sec. II. The K-S test carried out by analyzing the K-S statistic between embedding-dependent correlation integrals built on an input sequence and analytical GWN correlation integrals is discussed in Sec. III. The behavior of the correlation integrals at small scales in the new gauge and its making up a marker of a finite-dimensional system is analyzed in Sec. IV. A new estimator of the correlation dimension ν, the identification of embedding regions in which this estimator is uniform, and the consequences in terms of sampling time and embedding are the topic of Sec. V. The application of our method to synthetic and experimental time series is discussed in Sec. VI, while Sec. VII is devoted to possible applications in different research fields. Finally, the Appendix is devoted to the case of quasiperiodic systems.

II. CORRELATION INTEGRALS AND GAUSSIAN NOISE GAUGE TRANSFORMATION OF DISTANCE

A. Sample and asymptotic correlation integrals

A sample, scalar time series $\{x_i\}$, consisting of ℓ points and henceforth referred to as a sequence, is given. Upon evaluation of their sample mean $\bar{x}$ and sample standard deviation s_x, the elements of an input sequence are standardized via $(x_i - \bar{x})/s_x \to x_i$. The standardized sequence is then embedded with dimension m and lag L; henceforth we refer to this operation by saying that the sequence is embedded on the pair (m, L). The embedding procedure produces a sequence of vectors where the kth component $(k = 0, \ldots, m-1)$ of the ith vector $\mathbf{X}_i$ is given by x_{i+kL}. The embedded sequence's length is $\ell - (m-1)L$.

Instead of evaluating the sample correlation integral via the standard Grassberger-Procaccia sum [6], a Monte Carlo integration is used. A number N of vector pairs $\mathbf{X}_i$, $\mathbf{X}_j$ is randomly chosen without replacement. To avoid artifactual contributions from close points, the selected time indexes i, j must satisfy the condition $|i - j| \geqslant c_0$. Theiler [12] and Albano *et al.* [11] suggest taking c_0 equal to the first zero of the autocorrelation function; Gao and Zheng [13] instead suggest taking c_0 equal to the embedding window $(m-1)L$. Here a more restrictive condition is applied: c_0 is set to the largest of the second zero of the autocorrelation function and the product $2m(L+1)$. To determine the distance $d_{i,j}$ between the two vectors $\mathbf{X}_i, \mathbf{X}_j$, the Euclidean norm normalized by $\sqrt{m}$ is used:

$$d_{i,j} = \left[\frac{1}{m} \sum_{k=0}^{m-1} (x_{i+kL} - x_{j+kL})^2 \right]^{1/2}. \tag{1}$$

The normalizing term is introduced to provide an m-independent distance in the case $L = 0$. A sample correlation integral $\hat{C}_{m,L;N}(r)$ is finally computed as the empirical cumulative distribution function of N sample distances:

$$\hat{C}_{m,L;N}(r) = \frac{1}{N} \sum_{k=1}^{N} \Theta\big(r - d_{i_k,j_k}\big), \tag{2}$$

where i_k, j_k are the time indexes of the kth pair, and $\Theta(x)$ is the Heaviside step function. Because the number of available pairs is of order $\ell^2/2$ (provided that $\ell \gg c_0$), a reliable integration requires $\ell^2/2$ being at least 2–3 orders of magnitude greater than N. In the present paper, the input sequences consisted of $\ell = 10^5$ points and the maximum value of c_0 that occurred was of order 10^3. Thereupon, N was set to 4000.

Figure 1 shows examples of sample correlation integrals built on a sequence generated by integrating the following Lorenz attractor and taking the x component [14]:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x), \\ \frac{dy}{dt} &= x(r - z) - y, \\ \frac{dz}{dt} &= xy - bz. \end{aligned} \tag{3}$$

In these equations, $\sigma = 10$, $r = 28$, $b = 8/3$. Throughout this work, integrations of differential equations were carried out via a Runge-Kutta Prince-Dormand (8,9) algorithm, by randomly setting the starting point; sampling times T were taken equal to the integration times dt. In the case of Eqs. (3), $T = dt = 0.03$.

As shown in the figure, the Monte Carlo approach used in the present work correctly reproduces the small-distance behavior of $\hat{C}_{m,L;N}(r)$ as predicted by Grassberger and Procaccia [6], thus underpinning the reliability of the method.

By virtue of the Glivenko-Cantelli theorem, if the number N of pairs tends to infinity, a sample correlation integral $\hat{C}_{m,L;N}(r)$ tends to the asymptotic limit $C_{m,L}(r)$, henceforth referred to as the "asymptotic correlation integral."

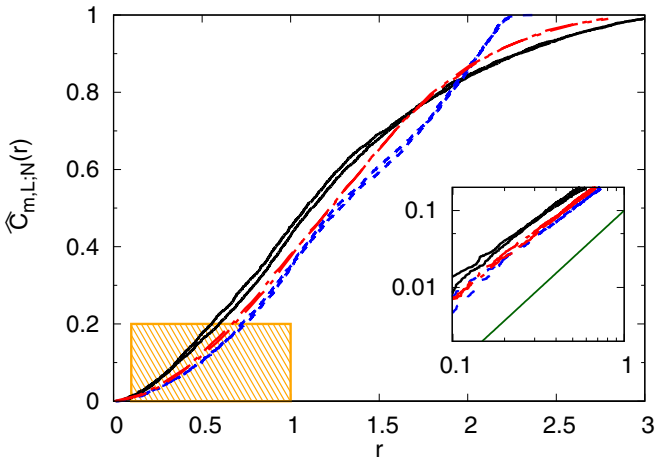

FIG. 1. Examples of sample correlation integrals evaluated by randomly selecting $N = 4000$ vector pairs of a sequence generated by sampling the x component of a Lorenz attractor. The sequence was embedded on the pairs (3,3) (black solid), (4,5) (blue dashed), (2,19) (red dash-dotted). For each embedding pair, two sample correlation integrals are shown. The inset magnifies in log-log scale the orange, diagonally patterned region at small distances. The linearity of the graphs is a marker of the Grassberger-Procaccia scaling relationship $C_m(r) \propto r^{\nu}$, where ν is the correlation dimension [6] (also commonly referred to as D_2). The green, solid straight line corresponds to $\nu = 2.05$, namely the accepted value for the correlation dimension of the Lorenz attractor [14].

For most dynamical systems, it is virtually impossible to analytically determine the shape of asymptotic correlation integrals. A remarkable exception is provided by GWN sources. A GWN sequence is a set of independent elements, each drawn from a normal probability density function $\mathcal{N}(\mu, \sigma^2)$. Because the sequence is discrete, the second zero of the autocorrelation function occurs at a lag equal to 2. According to the rule expressed above, c_0 is then set to $2m(L+1)$. By virtue of the constraint $|i-j| \geqslant c_0$, all x terms occurring within Eq. (1) turn out to be independent, so that each difference $x_{i+kL} - x_{j+kL}$ is distributed as $\mathcal{N}(0, 2\sigma^2)$. Thus, provided that $L > 0$, the square of the distance $d_{i,j}$ is a random variable distributed according to a χ^2 distribution with m degrees of freedom: $d_{i,j}^2 \sim (2\sigma^2/m)\,\chi_m^2$. If $L = 0$, because of the metric choice of Eq. (1), and independently of m, one has $d_{i,j}^2 \sim 2\sigma^2 \chi_1^2$. In the case of a GWN sequence, the analytic form of the asymptotic correlation integral is then

$$C_{\mathrm{GWN};m}(r) = \begin{cases} \frac{1}{\Gamma(\frac{m}{2})}\,\gamma\left(\frac{m}{2}, \frac{mr^2}{4\sigma^2}\right), & \text{if } L > 0, \\ \frac{1}{\Gamma(\frac{1}{2})}\,\gamma\left(\frac{1}{2}, \frac{r^2}{4\sigma^2}\right), & \text{if } L = 0,\ \forall m, \end{cases} \tag{4}$$

where $\gamma(s, x)$ is the lower incomplete Gamma function.

B. Gaussian noise gauge transformation of distance

Equation (4) shows that a correlation integral for a GWN source depends on the embedding dimension m. This m dependence can be removed by means of a suitable "gauge" transformation of distance. As a result, sample correlation integrals that are built on GWN sequences turn out to be drawn from the same distribution, independently of the embedding used.

The gauge transformation consists of mapping the distance $d_{i,j}$ of Eq. (1) onto a new distance $\delta_{i,j}$, as follows:

$$\delta_{i,j} = C_{\mathrm{GWN};m}(d_{i,j};\, \sigma = 1). \tag{5}$$

In this expression, the function $C_{\mathrm{GWN};m}$ is the same as in Eq. (4). The setting $\sigma = 1$ is coherent with the standardization procedure that is applied to the input sequence.

Given the new metric, a sample correlation integral $\hat{C}'_{m,L;N}(\rho)$—the prime indicates the new gauge—is evaluated by replacing the distance $d_{i,j}$ with the distance $\delta_{i,j}$ in the correlation sum of Eq. (2):

$$\hat{C}'_{m,L;N}(\rho) = \frac{1}{N}\sum_{k=1}^{N} \Theta\left(\rho - \delta_{i_k,j_k}\right). \tag{6}$$

As a result of Eqs. (2), (5), (6) and of the strict monotonicity of the $C_{\mathrm{GWN};m}$ functions, setting

$$\rho = C_{\mathrm{GWN};m}(r;\, \sigma = 1) \tag{7}$$

implies

$$\hat{C}'_{m,L;N}(\rho) = \hat{C}_{m,L;N}(r). \tag{8}$$

Because $C_{\mathrm{GWN};m}$ is a cumulative distribution function, the former d domain given by $\mathbb{R}_{\geqslant 0}$ is squeezed into the new one $[0, 1]$. Therefore, any correlation integral can be now plotted within the unitary square $[0, 1] \times [0, 1]$.

The probability integral transform theorem [15] states that if a random variable x has a cumulative distribution function $F_x(x)$, then the random variable $y = F_x(x)$ is uniformly distributed in the interval $[0, 1]$ and therefore has a cumulative distribution $F_y(y)$ given by $\xi(y) \equiv y$ ($y \in [0, 1]$). Consequently, under the gauge transformation, an asymptotic correlation integral $C'_{\mathrm{GWN};m,L}(\rho)$ that is computed on an infinitely long GWN time series turns out to be given by $\xi(\rho)$ independently of the embedding choice. If the GWN sequence is instead finite, the sample correlation integral $\hat{C}'_{m,L;N}(\rho)$ of Eq. (6) is expected to be statistically compatible with $\xi(\rho)$ according to the K-S test. The K-S statistic $\hat{K}_{m,L;N}$, given by

$$\hat{K}_{m,L;N} = \sup_{\rho\in[0,1]} |\hat{C}'_{m,L;N}(\rho) - \rho|, \tag{9}$$

is expected to be drawn from the following cumulative distribution function (see for example [16]):

$$F(K) = 1 - 2\sum_{k=1}^{\infty} (-1)^{k-1} \exp(-2k^2 N' K^2),$$

where $N' \equiv (\sqrt{N} + 0.12 + 0.11/\sqrt{N})^2$. We note that the population mean and standard deviation of the distribution are

$$\mu_K = \sqrt{\frac{\pi}{2N'}}\ln 2, \tag{10a}$$

$$\sigma_K = \sqrt{\frac{\pi}{2N'}}\sqrt{\frac{\pi}{6} - \ln^2 2} \simeq 0.3\mu_K. \tag{10b}$$

Both parameters asymptotically tend to zero.

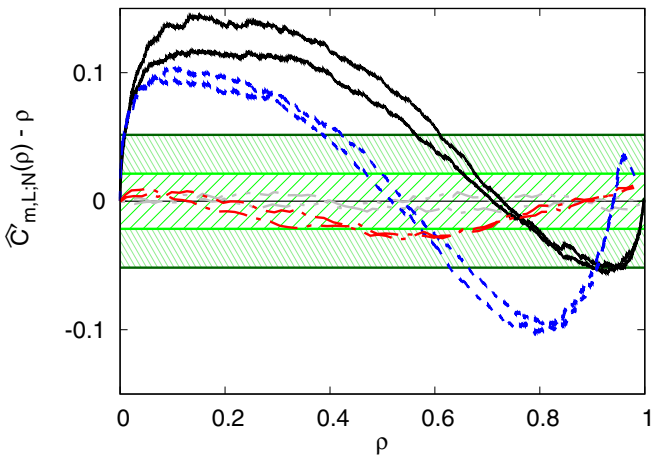

FIG. 2. Correlation bridges $\hat{B}_{m,L;N}(\rho)$ [see Eq. (11)] evaluated by means of the gauge-transformed vector distance ρ and using $N = 4000$ vector pairs. The black solid, blue dashed, and red dash-dotted pairs of graphs correspond to the standard sample correlation integrals $\hat{C}_{m,L;N}(\rho)$ shown in Fig. 1 and built on a Lorenz sequence embedded on the pairs (3,3), (4,5), (2,19), respectively. The two gray dash-double-dotted lines represent the Brownian bridges generated out of a GWN source. The K-S statistic $\hat{K}$, i.e., the gap from noise, for these last two sequences is well within the inner diagonally patterned region bounded by the $\hat{K}$ values that correspond to the $\alpha = 0.05$ significance level. On the other hand, the Lorenz-based sample correlation bridges embedded on the pairs (3,3), (4,5) exceed the outer diagonally patterned region, bounded by the $\hat{K}$ values that correspond to the $\alpha = 10^{-9}$ significance level. The Lorenz sequence embedded on the pair (2,19) yields a sample correlation bridge that only slightly overcomes the 0.05 significance level.

C. Correlation bridges and gap from noise

For the purposes of the following discussion, it is worth defining a "sample correlation bridge" $\hat{B}_{m,L;N}(\rho)$ as the deviation of the sample correlation integral $\hat{C}'_{m,L;N}(\rho)$ from the "reference line" $\xi(\rho)$:

$$\hat{B}_{m,L;N}(\rho) \equiv \hat{C}'_{m,L;N}(\rho) - \xi(\rho). \tag{11}$$

Each correlation bridge is then a function defined in [0,1], with codomain $[-1, 1]$ and tied to 0 in $\rho = 0, 1$. Similarly to the asymptotic correlation integral, the "asymptotic correlation bridge" $B_{m,L}(\rho)$ is the limit for $N \to \infty$ of the sample correlation bridge $\hat{B}_{m,L;N}(\rho)$.

With this notation, a sample correlation bridge $\hat{B}_{m,L}(\rho)$ computed on a finite GWN sequence turns out to be a Brownian bridge, which becomes identically zero in the case of an infinitely long time series. Furthermore, independently of the input sequence, the K-S statistic of Eq. (9), is simply given by

$$\hat{K}_{m,L;N} = \sup_{\rho\in[0,1]} |\hat{B}_{m,L;N}(\rho)|.$$

Henceforth, $\hat{K}_{m,L;N}$ ($\hat{K}_{m,L}$ if $N \to \infty$) is referred to as the sample (asymptotic) "gap from noise" for the input sequence.

Figure 2 shows the result of the gauge transformation for the six sample correlation integrals of Fig. 1 and for two GWN sequences.

III. TESTING A GWN SOURCE NULL HYPOTHESIS

Given a sequence originating from a generic dynamical system, as well as a "lattice" of embedding pairs $\{(m, L) \,|\, m \in [2, m_{\max}], L \in [1, L_{\max}]\}$, we start evaluating a sample correlation bridge on each pair of the lattice. Thereupon, a "map" of the gap from noise on the given lattice is determined, namely a set $\{\hat{K}_{m,L;N} \mid m \in [2, m_{\max}], L \in [1, L_{\max}]\}$. In the present work $m_{\max} = L_{\max} = 20$. By repeating a number M times the generation of a sample correlation bridge and the subsequent evaluation of the related gap from noise, and by averaging the results, a map of the averaged gap from noise is obtained: $\{\langle\hat{K}\rangle_{m,L;N} \mid m \in [2, m_{\max}], L \in [1, L_{\max}]\}$. It is worth noting that maps of different figures of merit and built on an embedding lattice were also recently used by Krakovská *et al.* [17] in order to test a modification of the false-nearest-neighbor method.

Under a null hypothesis $H_{0,\mathrm{GWN}}$ that the elements of the input sequence are independent and drawn from the same Gaussian distribution and provided that $M \gg 1$, each sample mean $\langle\hat{K}\rangle_{m,L;N}$ is approximately distributed as $\mathcal{N}(\mu_K, \frac{\sigma_K}{\sqrt{M}})$ as a consequence of the central limit theorem, with μ_K, σ_K given by Eq. (10). The p value corresponding to each averaged gap from noise can be then evaluated by using the z distribution.

The gauge transformation thus provides a straightforward method to assess the compatibility of a sample sequence with a GWN source. As an example, Fig. 3 shows the maps of the p value that, under $H_{0,\mathrm{GWN}}$, corresponds to the gap from noise evaluated out of four different sequences: a GWN sequence; a GWN sequence filtered by means of a first-order low-pass filter with cutoff frequency $0.05T^{-1}$; a sequence generated by sampling the y component of the Ueda oscillator [18],

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = -x^3 - ky + A\sin(t),$$

where $A = 7.5$, $k = 0.05$, $T = dt = 0.5$ [14]; a sequence generated by the Lorenz attractor of Eq. (3). In the case of the GWN sequence, the gap from noise is always so small that the corresponding p value largely overcomes the Bonferroni-corrected 0.01 significance level, i.e., 0.01/380 (380 being the number of embedding pairs shown within each map). With regard to the filtered GWN sequence, the p value map's pairs with $L \lesssim 7$ are incompatible with a pure GWN source, while the contrary occurs for pairs with $L \gtrsim 6$. The map corresponding to the Ueda oscillator—obtained with $M = 10$ rather than the value 100 used for the other sequences—shows most embedding pairs leading to a violation of the null hypothesis $H_{0,\mathrm{GWN}}$, and only a few embedding pairs on which $H_{0,\mathrm{GWN}}$ is accepted. Finally, the Lorenz sequence generated a violation of $H_{0,\mathrm{GWN}}$ in each point of the lattice.

In the case of an unknown source of the sequence, a behavior similar to the filtered GWN sequence as well as the Ueda and Lorenz systems is to be interpreted as being due to a non-GWN source because of the presence of incompatible embedding pairs. Apparently compatible embedding pairs are then to be deemed as unsuitable to embedding and therefore excluded from any further evaluation. In the examples of Fig. 3, such pairs are marked with a cross.

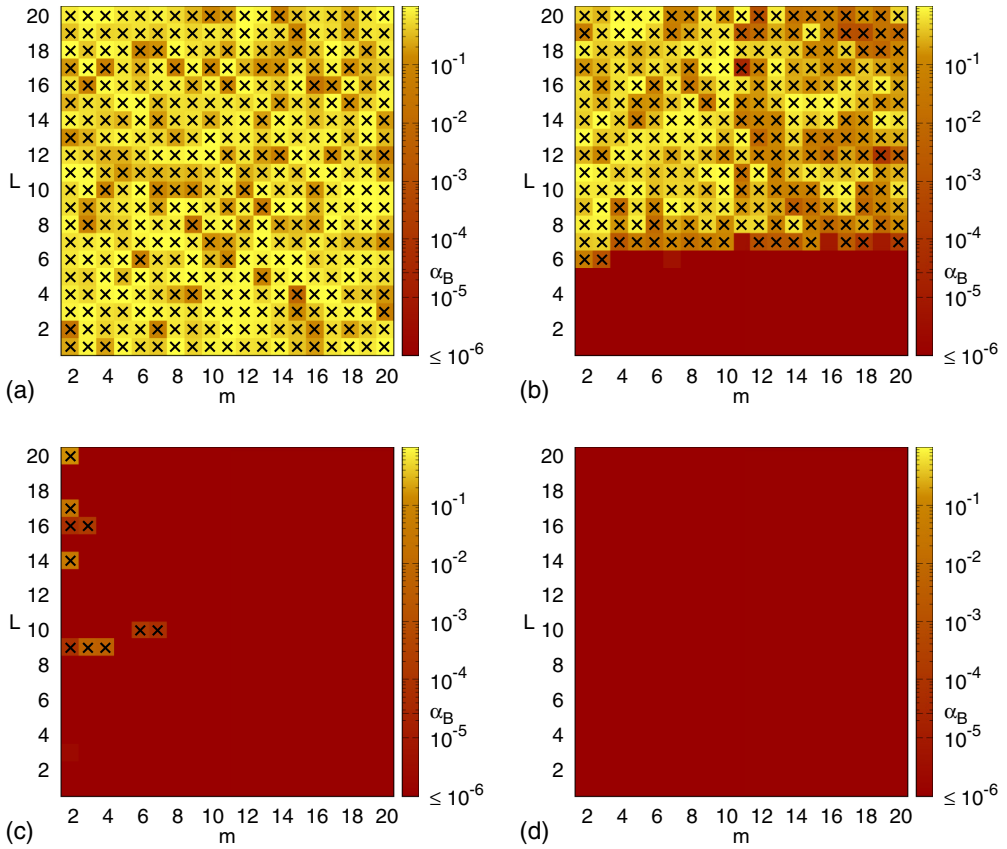

FIG. 3. Maps of p value evaluated, under $H_{0,\mathrm{GWN}}$, out of the averaged gap from noise for four different sequences: (a) GWN sequence; (b) filtered GWN sequence (first-order low-pass filter with $f_{3\mathrm{dB}} = 0.05T^{-1}$); (c) Ueda oscillator; (d) Lorenz attractor. The color (gray scale) corresponding to the $p = 0.01$ significance level, Bonferroni-corrected to $\alpha_B = 0.01/380$ to keep into account all displayed embedding pairs, is highlighted in the palette. Pairs whose gap from noise overcomes the significance threshold, and are thus compatible with $H_{0,\mathrm{GWN}}$ but not with the hypothesis of a chaotic, finite-dimensional dynamical system, are marked with a cross. Using $M = 100$ also for the Ueda oscillator would lead to a map like the Lorenz one.

IV. SMALL-SCALE BEHAVIOR OF CORRELATION BRIDGES

The GWN gauge transformation of distances provides a way to investigate the chaoticity and dimensionality of the dynamical system that underlies an input sequence by evaluating the behavior at small distances of the correlation bridges. An example is shown in Fig. 2, in which the correlation bridges built out of a sequence generated by a Lorenz system grow in the region $\rho \ll 1$ up to a local maximum (not necessarily corresponding to the gap from noise). The goal of the following discussion is to show that this behavior is typical of any chaotic, finite-dimensional dynamical system. Henceforth, the hypothesis that the dynamical system under investigation is chaotic and finite-dimensional, and has in addition a correlation dimension ν, is referred to as $H_{0,\mathrm{cfd}}(\nu)$.

The derivative of the asymptotic correlation bridge $B_{m,L}(\rho)$ can be expressed as follows:

$$\frac{dB_{m,L}(\rho)}{d\rho} = \frac{dC'_{m,L}(\rho)}{d\rho} - 1 = \frac{\partial C_{m,L}(r)}{\partial r}\frac{dr(\rho)}{d\rho} - 1, \quad (12)$$

where $r = r(\rho)$ is the inverse of the function $\rho = \rho(r)$ defined in Eq. (7), and the asymptotic limits of both Eq. (8) and Eq. (11) were used. By applying the inverse function theorem and using Eq. (4), the derivative $\frac{dr(\rho)}{d\rho}$ can be expressed as (henceforth we assume $m > 1$)

$$\frac{dr(\rho)}{d\rho} = \left(\frac{d\rho}{dr}\right)^{-1} = \frac{1}{2}\Gamma\left(\frac{m}{2}\right)\left(\frac{m}{4\sigma^2}\right)^{-\frac{m}{2}} \exp\left(\frac{mr^2}{4\sigma^2}\right) r^{1-m}. \quad (13)$$

Under $H_{0,\mathrm{cfd}}(\nu)$ and in the limit $r \to 0$, according to the seminal work by Grassberger and Procaccia [6] the correlation integral $C_{m,L}(r)$ is characterized by a power-law scaling $C_{m,L}(r) = \beta_{m,L} r^{\nu}$, where $\beta_{m,L}$ is an embedding-dependent constant. Consequently, in the small-scale regime we have

$$\frac{dC_{m,L}(r)}{dr} = \nu\beta_{m,L} r^{\nu-1} = \frac{\nu}{r} C_{m,L}(r). \quad (14)$$

By replacing both Eq. (13) and Eq. (14) into Eq. (12) it follows

$$\frac{dB_{m,L}(\rho)}{d\rho} = \frac{\nu\beta_{m,L}}{2}\Gamma\left(\frac{m}{2}\right)\left(\frac{m}{4\sigma^2}\right)^{-\frac{m}{2}} \exp\left(\frac{mr^2}{4\sigma^2}\right) r^{\nu-m} - 1. \quad (15)$$

The last expression shows that, under $H_{0,\mathrm{cfd}}(\nu)$, the derivative $\frac{dB_{m,L}(\rho)}{d\rho}$ has to diverge to $+\infty$ as $\rho, r \to 0$ and $m > \nu$. The previous expression and statement are valid provided that m is

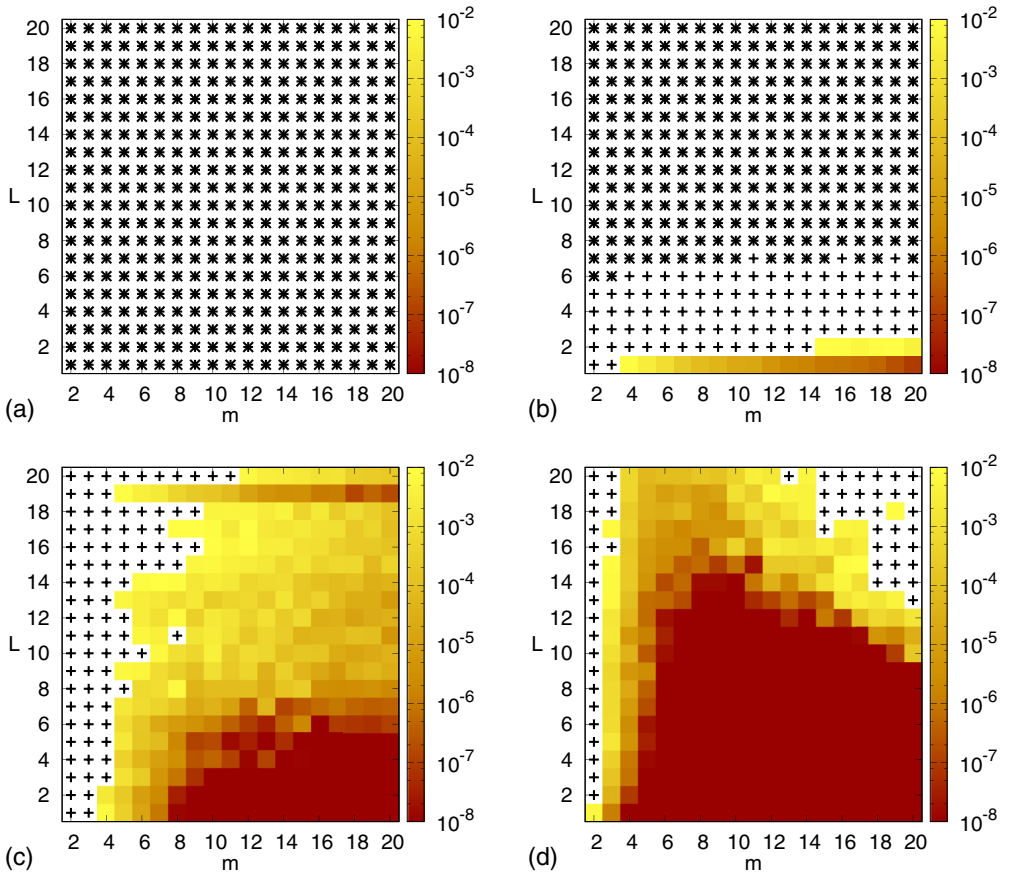

FIG. 4. Maps of the probability given by Eq. (16) for (a) GWN sequence; (b) filtered GWN sequence (first-order low-pass filter with $f_{3\mathrm{dB}} = 0.05T^{-1}$); (c) Ueda oscillator (with $M = 100$ instead of $M = 10$ as in Fig. 3); (d) Lorenz attractor. Pairs for which the small-scale behavior of the corresponding correlation bridges is not divergent, and are thus incompatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$, are marked with a plus (+) sign. Pairs in which $H_{0,\mathrm{GWN}}$ holds are marked with a cross and lead, as expected, to a rejection of $H_{0,\mathrm{cfd}}(\nu)$.

less than the irrelevance-dependent embedding dimension m_I, which can be estimated as $m_I \approx \tau_I/(LT)$. Consequently, due to the boundary condition $B_{m,L}(0) = 0$, in a neighborhood of the origin and in the case $\nu < m < m_I$, each correlation bridge has to be positive and growing, and its first stationary point thus has to be a maximum point.

To sum up, the simultaneous fulfillment of three conditions—validity of $H_{0,\mathrm{cfd}}(\nu)$, $\nu < m$, $m \lesssim \tau_I/(LT)$—makes up a sufficient condition to observe a divergent behavior of the correlation bridges at small distance, though not a necessary one (see examples at the end of this section). On the contrary, if no divergent behavior is observed, either $H_{0,\mathrm{cfd}}(\nu)$ does not hold, or $\nu > m$, or $m > \tau_I/(LT)$; this last situation might, for example, occur if too a large value of L is chosen. In the last two cases, the embedding pair is clearly unsuitable, whereas in the first one it is useless. Divergent behavior of the correlation bridges can be therefore used as an additional tool to rule out unsuitable embedding pairs.

The argument above holds for an asymptotic correlation bridge. The assessment of the divergent behavior at small distances in a sample correlation bridge built on N samples and for a given embedding pair relies on the following statistical approach. First, the number N must be large enough so that the minimum sampled distance r_1 lays within the Grassberger-Procaccia power-law regime. Let then ρ_1 be the corresponding, gauge-transformed distance. The sample distance ρ_1 provides an estimate of a sample correlation integral in the GWN gauge, namely $\hat{C}'_{m,L;N}(\rho_1) = 1/N$, and thus of a sample correlation bridge, namely $b_1 \equiv \hat{B}_{m,L;N}(\rho_1) = 1/N - \rho_1$.

By determining a number M of correlation bridges and thereupon a set of size M of b_1 values, we assume a divergent behavior being immediately disproved if a single vanishing or negative value of b_1 occurs. The probability that a divergent behavior exists also if all M samples are positive can be determined by using Chebyshev's inequality. Although other approaches, like extreme value analysis [19], might provide a more precise assessment, the chosen approach is the safest one, because Chebyshev's inequality provides an upper limit to the p value. Let $\bar{b}_1$, s_{b_1} be the sample mean and sample standard deviation of the M values of b_1, respectively. If M is large enough, these two statistics can be taken as estimates of the respective population parameters. By virtue of Chebyshev's inequality, it follows

$$P(b_1 \leqslant \bar{b}_1 - k s_{b_1}) \leqslant P(|b_1 - \bar{b}_1| \geqslant k s_{b_1}) \leqslant \frac{1}{k^2},$$

where k is a positive number. By setting $k = \bar{b}_1/s_{b_1}$ and assuming $\bar{b}_1$ to be positive (otherwise, a divergent behavior

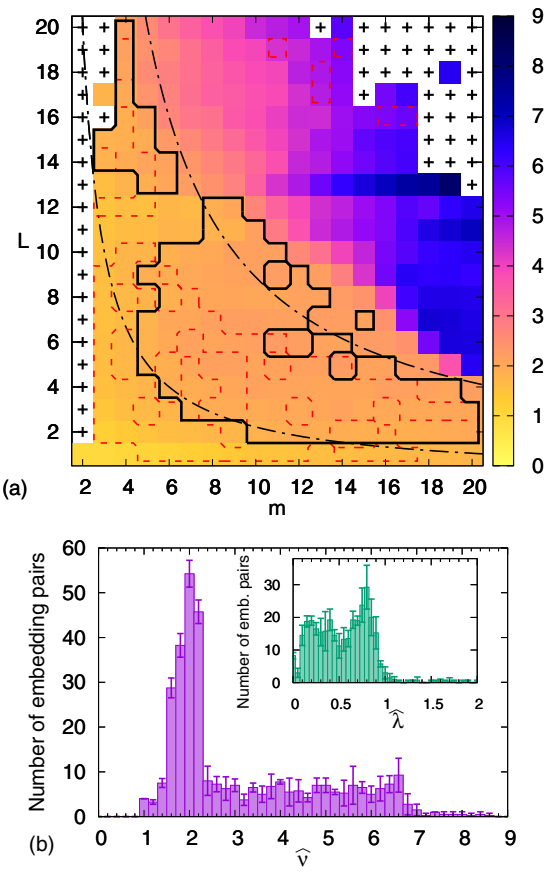

FIG. 5. Analysis of the estimated correlation dimension $\hat{\nu}$ given by Eq. (18) and computed out of a Lorenz sequence. (a) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ (see Fig. 4). Embedding pairs incompatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ are marked with a + sign. The black, bold line bounds the plateau made of the embedding pairs whose $\hat{\nu}$ lies within the range $\hat{\nu} = 2.0 \pm 0.2$ (see histogram below). The red, dashed line bounds the set of embedding pairs whose estimated MLE $\hat{\lambda}$ lies within the range 0.8 ± 0.1 (see inset below). The two black, dash-dotted lines correspond to hyperbolae approximatively enclosing the $\hat{\nu}$ plateau and described by $L = 20/(m-1)$, $L = 80/(m-1)$. (b) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.2 wide bins. Inset: Histogram of the number of embedding pairs whose estimated MLE $\hat{\lambda}$ falls within 0.05 wide bins. In both histograms, bin heights and error bars correspond to the sample mean and sample standard deviation, respectively, evaluated on a set of four calculations.

would be trivially disproved), one gets

$$P(b_1 \leqslant 0) \leqslant \left(\frac{s_{b_1}}{\bar{b}_1}\right)^2. \tag{16}$$

Assuming $H_{0,\mathrm{cfd}}(\nu)$ to hold, we conservatively consider the cases in which $(s_{b_1}/\bar{b}_1)^2$ overcomes the significance level of 0.01 as disproving the divergent behavior and thus assessing the unsuitability of the corresponding embedding pair. Due to the mutual exclusiveness of $H_{0,\mathrm{cfd}}(\nu)$ and $H_{0,\mathrm{GWN}}$, if $H_{0,\mathrm{GWN}}$ holds, $H_{0,\mathrm{cfd}}(\nu)$ is rejected. Consequently, the divergent behavior is a more stringent marker than the GWN compatibility in order to identify a suitable embedding pair.

Figure 4 shows the maps of the probability given by Eq. (16) for the four sequences used to generate Fig. 3. The maps show the suitable embedding pairs for each one of the input sequences: in the case of the GWN sequence, no such pair is available.

V. ESTIMATING THE CORRELATION DIMENSION AND ITS UNIFORMITY WITHIN AN EMBEDDING LATTICE

A. Estimator of the correlation dimension ν in the Gaussian noise gauge

In this section we show that, under $H_{0,\mathrm{cfd}}(\nu)$ and provided that a small-scale divergent behavior is assessed, the correlation dimension ν can be estimated out of the position and the amplitude of the first maximum of a correlation bridge. This approach requires that the small-scale power-law scaling survives at sufficiently large values of r and thus of ρ. However, by noting that the diverging behavior of the derivative $\frac{dB_{m,L}(\rho)}{d\rho}$ at the origin increases with m, and knowing that any correlation bridge cannot overcome the boundaries ± 1 and must eventually end up in 0 (at $\rho = 1$), we can assume that the larger m, the smaller the position of the first maximum (see, for example, Fig. 2). Consequently, the requirement of a small-scale power-law regime can be satisfied by taking a sufficiently large value of m.

Let ρ_0 be the position of the first zero of the derivative of the asymptotic correlation bridge $B_{m,L}(\rho)$, given by Eq. (15), and r_0 the corresponding Euclidean distance normalized by m. By again exploiting the inverse function theorem, it is straightforward to show that imposing ρ_0 to be a zero of the derivative of Eq. (15), and thus of Eq. (12), is equivalent to state that r_0 satisfies the following equation in r:

$$\frac{d\rho(r)}{dr} = \frac{dC_{m,L}(r)}{dr}. \tag{17}$$

The ratio between the amplitude $B_{m,L}(\rho_0)$ of the first maximum of the correlation bridge and its position ρ_0 is given by

$$\frac{B_{m,L}(\rho_0)}{\rho_0} = \frac{C'_{m,L}(\rho_0)}{\rho_0} - 1 = \frac{C_{m,L}(r_0)}{\rho_0} - 1.$$

However, by exploiting Eq. (14) to express $C_{m,L}(r_0)$ as $\frac{r_0}{\nu}\frac{dC_{m,L}(r_0)}{dr}$, and subsequently using Eq. (17), the previous expression can be rewritten as

$$\frac{B_{m,L}(\rho_0)}{\rho_0} = \frac{1}{\nu}\frac{r_0}{\rho_0}\frac{d\rho(r_0)}{dr} - 1.$$

By defining the variable $x \equiv \frac{mr^2}{4\sigma^2}$ and the related constant $x_0 \equiv \frac{mr_0^2}{4\sigma^2}$, we get

$$\frac{B_{m,L}(\rho_0)}{\rho_0} = \frac{2}{\nu}\frac{x_0}{\rho_0}\frac{d\rho(x_0)}{dx} - 1.$$

Finally, by knowing the GWN gauge transformation for $m > 1$, namely $\rho = \gamma(\frac{m}{2}, x)/\Gamma(\frac{m}{2})$, and that $\frac{d\gamma(s,x)}{dx} = x^{s-1}e^{-x}$, the ratio $\frac{B_{m,L}(\rho_0)}{\rho_0}$ can be expressed as follows:

$$\frac{B_{m,L}(\rho_0)}{\rho_0} = \frac{2}{\nu\rho_0}\frac{x_0^{m/2}e^{-x_0}}{\Gamma\left(\frac{m}{2}\right)} - 1,$$

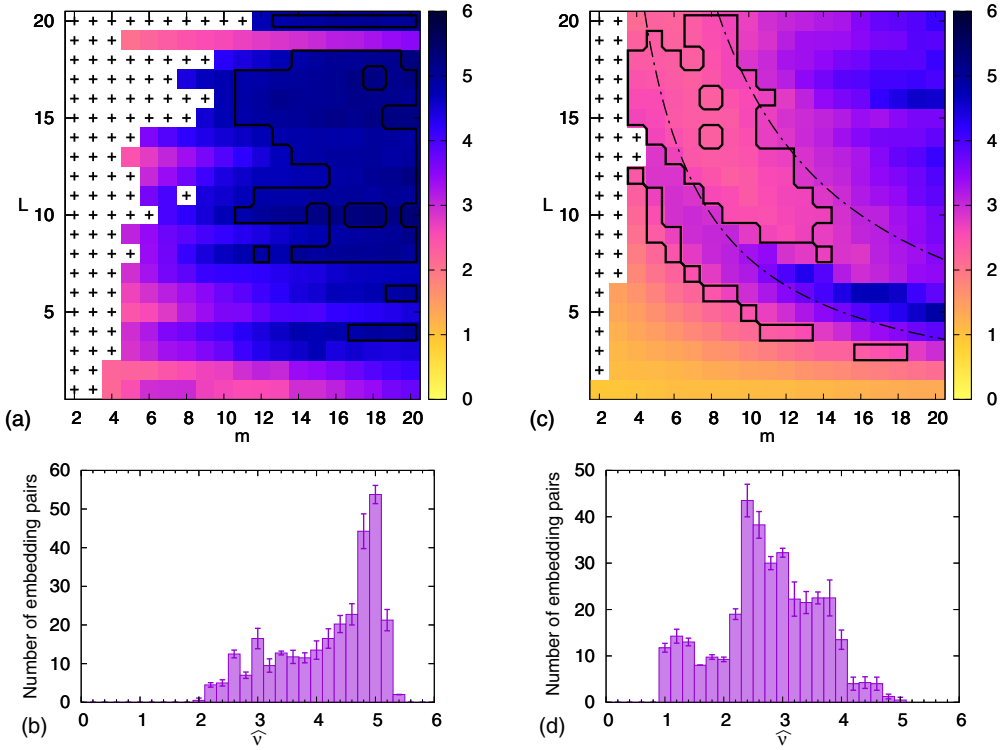

FIG. 6. Analysis of the estimated correlation dimension $\hat{\nu}$ given by Eq. (18) and computed out of a Ueda sequence. (a) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ (see Fig. 4), for the sequence sampled with $T = 0.5$. (b) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.2 wide bins. No significant inner peak can be identified (for the peak at $\hat{\nu} = 5$, see main text). (c) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$, for the Ueda sequence sampled 10 times faster, i.e., with $T = 0.05$. The black, bold line bounds the embedding pairs that make up the peak at $\hat{\nu} = 2.5 \pm 0.2$ appearing in the histogram below. The two black, dash-dotted lines correspond to hyperbolae approximatively enclosing the $\hat{\nu}$ plateau and described by $L = 70/(m-1)$, $L = 150/(m-1)$. (d) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.2 wide bins. In both maps, embedding pairs incompatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ are marked with a + sign, while, in both histograms, bin heights and error bars correspond to the sample mean and sample standard deviation, respectively, evaluated on a set of four calculations.

By inverting the previous expression we get

$$\nu = \frac{2}{B_{m,L}(\rho_0) + \rho_0} \frac{x_0^{m/2} e^{-x_0}}{\Gamma\left(\frac{m}{2}\right)}. \tag{18}$$

Equation (18) is valid in the asymptotic limit $N \to \infty$. However, it is reasonable to assume that, in the case of a sample correlation bridge with finite N, replacing in Eq. (18) the three parameters x_0, ρ_0, and $B_{m,L}(\rho_0)$ with the respective sample values provides an estimate $\hat{\nu}$ of the correlation dimension.

It is worth remarking that although the first maximum is a local characteristic of the correlation bridge, its value is made up by all preceding bridge points that are constrained, in the small-scale regime, to build up a growing curve. Consequently, the estimator of Eq. (18) relies on contributions integrated on the whole small-scale region, and therefore has an intrinsically *global* character.

B. Uniformity of the estimated correlation dimension within an embedding lattice

As mentioned in Sec. I, the estimated correlation dimension $\hat{\nu}$ is expected not to depend on the embedding pair (m, L) provided that the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ holds and

$$\tau_R/T \lesssim (m-1)L \lesssim \tau_I/T. \tag{19}$$

Accordingly, the set of suitable embedding pairs identified via the analysis discussed in Secs. III, IV can be further *purified* by assessing sets of embedding pairs on which $\hat{\nu}$ shows uniformity, and ruling out the others. On a $\hat{\nu}$ map, a similar set is expected to correspond to a $\hat{\nu}$ plateau, bound by two right hyperbolae.

Figure 5(a) shows the map of $\hat{\nu}$ for the same Lorenz sequence used above. Here and henceforth, $\hat{\nu}$ corresponds to the sample mean of $M = 100$ calculations. The evaluation of $\hat{\nu}$ regards only the points that show a divergent behavior compatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$. Figure 5(b) shows the histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.2 wide bins. The histogram has a peak at 2.0 ± 0.2. As highlighted in Fig. 5(a) by means of a black, bold boundary, the corresponding embedding pairs form a $\hat{\nu}$ *plateau*. As a consequence of the expected uniformity of the estimates of an embedding-independent observable on a set of suitable embedding pairs, this $\hat{\nu}$ plateau can be deemed as a possible optimal set to embed the input sequence. Remarkably, the plateau level of 2.0 ± 0.2 is in agreement with the

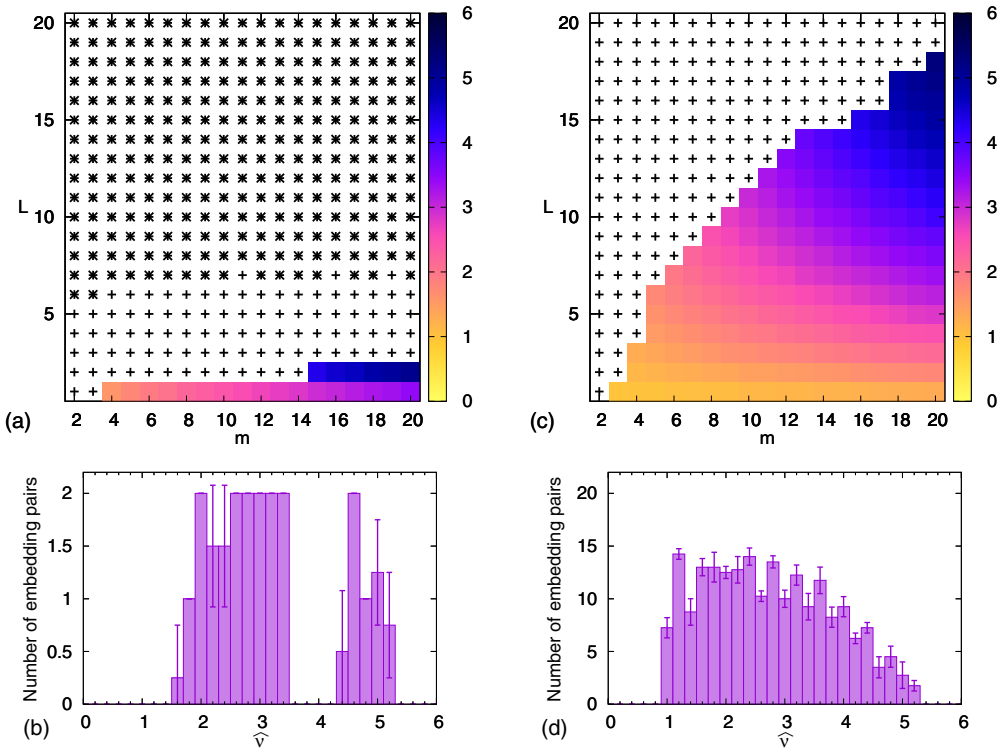

FIG. 7. Analysis of the estimated correlation dimension $\hat{\nu}$ given by Eq. (18) and computed out of a filtered GWN sequence. (a) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ (see Fig. 4), for the sequence sampled at $f_s = 1$ kHz. (b) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.2 wide bins. No significant peak can be identified. (c) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$, for the sequence sampled 10 times faster, i.e., at $f_s = 10$ kHz. (d) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.2 wide bins. Again, no significant peak can be identified. In both maps, embedding pairs compatible with $H_{0,\mathrm{GWN}}$ are marked with a cross while those incompatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ are marked with a + sign. In both histograms, bin heights and error bars correspond to the sample mean and sample standard deviation, respectively, evaluated on a set of four calculations.

accepted value of 2.05 for the Lorenz attractor's correlation dimension [14].

Figure 5(b) also shows the histogram of the estimated MLE $\hat{\lambda}$ calculated via the divergence rate method [13,20–22]. Despite the statistical noise, the histogram shows a local maximum at 0.8 ± 0.1 superimposed on a slowly decreasing distribution. The set of points corresponding to the maximum is highlighted in the map by means of a red, dashed boundary. Approximately half of the points belonging to this set lie within the $\hat{\nu}$ plateau. Again, this behavior reflects the expected uniformity of estimates of an embedding-independent observable. In addition, the observed MLE value of 0.8 ± 0.1 is in agreement with the value 0.9057(8) for the Lorenz attractor of Eq. (3) obtained by using the standard method [23,24]. However, due to the fact that the divergence rate method tends to underestimate $\hat{\lambda}$ especially in unsuitable embedding pairs—which accounts for the shape of the leftover part of the histogram—the $\hat{\lambda}$ plateau is less regular than the $\hat{\nu}$ plateau. Consequently, to select an optimal set of embedding pairs the correlation dimension $\hat{\nu}$ of Eq. (18) appears to be more appropriate than the MLE estimated via divergence rate method.

Crucially, the shape of the $\hat{\nu}$ plateau resulting from Fig. 5(a) appears to be bounded by two hyperbolae, in agreement with the constraints of Eq. (19). The two hyperbolae displayed in the figure correspond to $\tau_R/T \approx 20$ and $\tau_I/T \approx 80$, or equivalently $\tau_R \approx 0.6$, $\tau_I \approx 2.4$. Moreover, the fact that the two hyperbolae are sufficiently separated suggests that the sampling time T was correctly chosen: besides aliasing, too a long time T leads to a compression of the $\hat{\nu}$ plateau towards the axes. This compression is visible in Fig. 6(a), which shows the map of $\hat{\nu}$ for the Ueda sequence used above.

Considering the histogram of $\hat{\nu}$ in Fig. 6(b), a growing profile appears, which can be due to two alternative mechanisms: a peak—and thus a plateau on the map—at $\hat{\nu}$ values higher than those accessible by the current analysis and whose resolution would require wider lattices; alternatively, m overcoming the boundary set by the irrelevance time, so that the sequence becomes noiselike and thus, apparently, very high-dimensional. (The fact that the histogram contains two occupied bins beyond the peak at $\hat{\nu} = 5$ is due to the variability intrinsic in the evaluation of $\hat{\nu}$, which, in terms of sample standard deviation on the $M = 100$ calculations, ranges from 0.1 to 0.4.) While tackling the first mechanism is computationally demanding, the second one can be conveniently addressed by increasing, if possible, the sampling frequency. Figure 6(c) shows the map of $\hat{\nu}$ for a sequence generated by the same Ueda oscillator as above, but with a 10 times higher sampling frequency. The resulting $\hat{\nu}$ histogram, shown in Fig. 6(d), reveals a peak at $\hat{\nu} = 2.5 \pm 0.2$, and the corresponding plateau is shown in the map by means of a

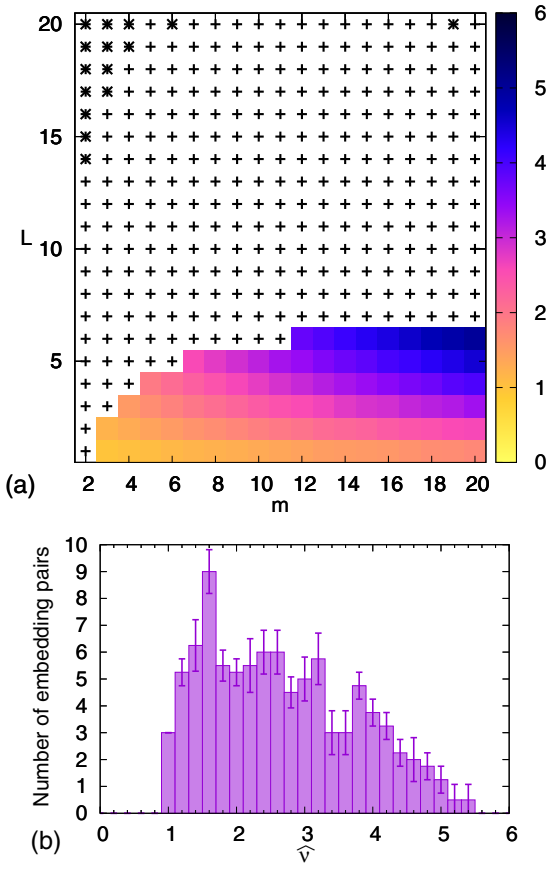

FIG. 8. Analysis of the estimated correlation dimension $\hat{\nu}$ given by Eq. (18) and computed out of a surrogate sequence of the Lorenz sequence analyzed in the paper. (a) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$. Embedding pairs compatible with $H_{0,\mathrm{GWN}}$ are marked with a cross while those incompatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ are marked with a + sign. (b) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.2 wide bins. Bin heights and error bars correspond to the sample mean and sample standard deviation, respectively, evaluated on a set of four calculations. No significant peak can be identified.

black boundary. Two hyperbolae, corresponding to $\tau_R \approx 3.5$, $\tau_I \approx 7.5$ and approximately enclosing the $\hat{\nu}$ plateau, are also shown. The estimated correlation dimension is in agreement with the value reported in the literature for the same system [14], namely $\nu = 2.68 \pm 0.13$.

In the case of the filtered GWN sequence, the increase of the sampling frequency leads to no significant results, as displayed in Fig. 7. The left part of the figure shows the map and the related histogram evaluated for the same sequence used above and sampled with $f_s = 1$ kHz, while the right part shows the map and histogram in the case of a sequence sampled with $f_s = 10$ kHz. Besides an obvious expansion of the lattice region that contains, according to the gap from noise and small-scale behavior, suitable embedding pairs and a related increase in the population of the histogram, there is no significant change in the morphology of both the region and the histogram.

VI. APPLICATION OF THE METHOD TO SYNTHETIC AND EXPERIMENTAL SEQUENCES

In this section, the method discussed above is applied to several synthetic and experimental sequences. The goal is to highlight both its reliability and possible weaknesses.

Figure 8 shows the result of the analysis carried out on a Fourier-based surrogate of the Lorenz sequence used in the previous sections. In compliance with a standard procedure [25] (see also [26] for a more recent review on this topic) the surrogate was generated by randomizing the phase of the Fourier components while conserving the power spectral density and thus the autocorrelation. The resulting map is similar to that obtained by using a filtered GWN sequence: in the set of embedding pairs that are compatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ no plateau structure is present. This result provides evidence of the reliability of the method to distinguish between a true finite-dimensional chaotic source and a surrogate, noisy one.

Figure 9 shows the results of the analysis carried out on both a sequence generated by integrating a Rössler attractor and a related Fourier-based surrogate. The equations describing the attractor are [27]

$$\frac{dx}{dt} = -y - z, \quad \frac{dy}{dt} = x + ay, \quad \frac{dz}{dt} = b + z(x - c),$$

where $a = b = 0.2$, $c = 5.7$, namely the archetypical setting by Rössler, and $T = dt = 0.125$. The sequence corresponds to the x component. The two histograms show pronounced peaks at 1.6 ± 0.1 and 1.9 ± 0.1, respectively. However, the corresponding $\hat{\nu}$ plateaus are definitely not bounded by any pairs of hyperbolae. While in the case of the surrogate sequence this lack of evidence of a finite-dimensional chaotic behavior is expected and welcome, the same is not true for the original Rössler sequence, as the attractor is known to have a correlation dimension of 1.99 ± 0.08 [14].

The failure in the identification of the source of the sequence as a finite-dimensional chaotic system reflects well-known difficulties to deal with the Rössler attractor, both with regard to the determination of the correlation dimension [14] and in terms of a relative insensitivity towards optimal embedding approaches [4,5]. A main reason of this failure is the quasiperiodic behavior of the Rössler attractor [28] that entails an extremely long autocorrelation time. As discussed in the Appendix, the $\hat{\nu}$ plateau of Fig. 9(a), which corresponds to a rectangular region defined by $m \gtrsim 10$, is typical of such behavior. Besides showing that the Rössler system is similar to a quasiperiodic system, the analysis discussed in the Appendix also provides an estimate of a signal-to-noise (SNR) ratio characteristic of the Rössler system: on the basis of the plateau value of 1.55 ± 0.05 this SNR can be estimated to be 2.1 ± 0.2. Remarkably, this estimate very well matches an SNR value estimated by looking at the Fourier spectrum of a Rössler sequence and singling out the fundamental frequency and its harmonics.

Further evidence of the sensitivity of the method to the autocorrelation and thus the spectral properties of the underlying dynamical system is provided by the analysis carried out on a sequence generated by integrating the Rössler attractor with the parameter setting $a = 0.343$, $b = 1.82$, $c = 9.75$,

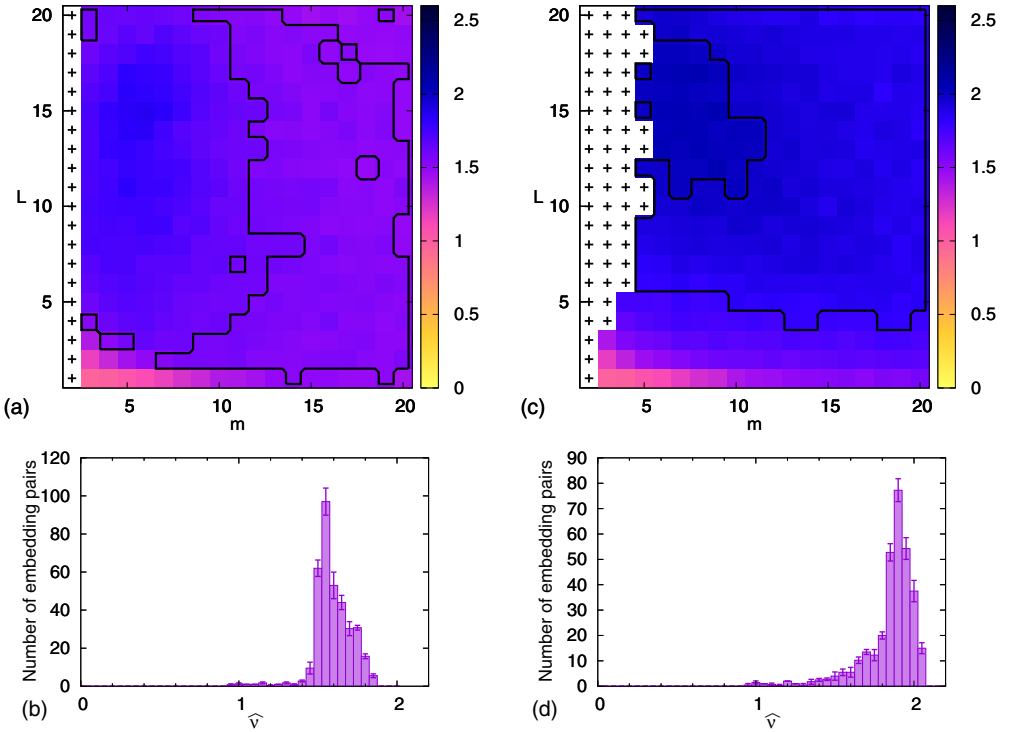

FIG. 9. Analysis of the estimated correlation dimension $\hat{\nu}$ given by Eq. (18) and computed out of a Rössler sequence and a related surrogate sequence. (a) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ in the case of the Rössler sequence. The black, bold line bounds the embedding pairs that make up the peak at $\hat{\nu} = 1.55 \pm 0.05$ appearing in the histogram below. (b) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.05 wide bins. (c) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ in the case of a surrogate sequence. The black, bold line bounds the embedding pairs that make up the peak at $\hat{\nu} = 1.9 \pm 0.1$ appearing in the histogram below. (d) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.05 wide bins. In both maps, embedding pairs incompatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ are marked with a + sign. In both histograms, bin heights and error bars correspond to the sample mean and sample standard deviation, respectively, evaluated on a set of four calculations. Despite the peaks in the histograms, the two $\hat{\nu}$ plateaus do not correspond to any region bounded by fixed values of the embedding window and therefore by any pairs of hyperbolae.

with $T = dt = 0.05$. This choice of the parameters a, b, c corresponds to the so-called *funnel* attractor [29], which is known to yield time series with a smoother spectrum than the archetypical choice $a = b = 0.2$, $c = 5.7$ used above. The results are shown in Fig. 10. While no significant $\hat{\nu}$ plateau shows up in the case of the Fourier-based surrogate, in the case of the Rössler sequence the $\hat{\nu}$ histogram of Fig. 10(b) reveals a peak at $\hat{\nu} = 1.7 \pm 0.2$. The related plateau, shown in the map of Fig. 10(a) by means of a black boundary, is approximately enclosed by two hyperbolae corresponding to $\tau_R \approx 1$, $\tau_I \approx 4$.

The method was finally applied to the analysis of an experimental sequence, namely the intensity data of an 81.5 μm NH_3 laser [30], which is publicly available at [31]. The results are shown in Fig. 11. A first selection of suitable embedding pairs was identified assuming the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ and analyzing the gap from noise (see Sec. III) and the small-scale divergent behavior (see Sec. IV). Thereupon the correlation dimension was evaluated via the estimator of Eq. (18). The resulting map of Fig. 11(a) shows a $\hat{\nu}$ plateau whose level can be clearly identified by means of the $\hat{\nu}$ histogram of Fig. 11(b). The plateau level of $\hat{\nu} = 2.0 \pm 0.2$ is in agreement with results of a previous work [32] concerning the same sequence. The redundance and the irrelevance times are estimated to be $\tau_R/T \approx 10$, $\tau_I/T \approx 40$, respectively.

Finally, the MLE of the 81.5 μm NH_3 laser system was evaluated on the 72 embedding pairs belonging to the $\hat{\nu}$ plateau shown in Fig. 11(a) by using the divergence rate method [5]. By considering the 48 embedding pairs on which the divergence rate method provided a nonzero value and averaging the corresponding results we got $\hat{\lambda} = (0.025 \pm 0.007)T^{-1}$. This result is in agreement with the assessment by Kantz and Schreiber [28].

VII. DISCUSSION

In this paper we discussed an approach to tackle the enduring issue of determining optimal parameters to embed time series generated by unknown dynamical systems. The method relies on statistically testing the hypothesis that a given sequence originates from a Gaussian white noise source and, provided that the previous hypothesis is discarded, on the assessment of features that are typical of chaotic, finite-dimensional systems. These features, above all the correlation dimension, are mapped on a lattice of embedding pairs. Rather than assigning a specific pair of the embedding dimension and lag—as generally purported by conventional embedding methods—our method delivers a set of suitable pairs. In this sense, our method is not alternative to—or even competitive

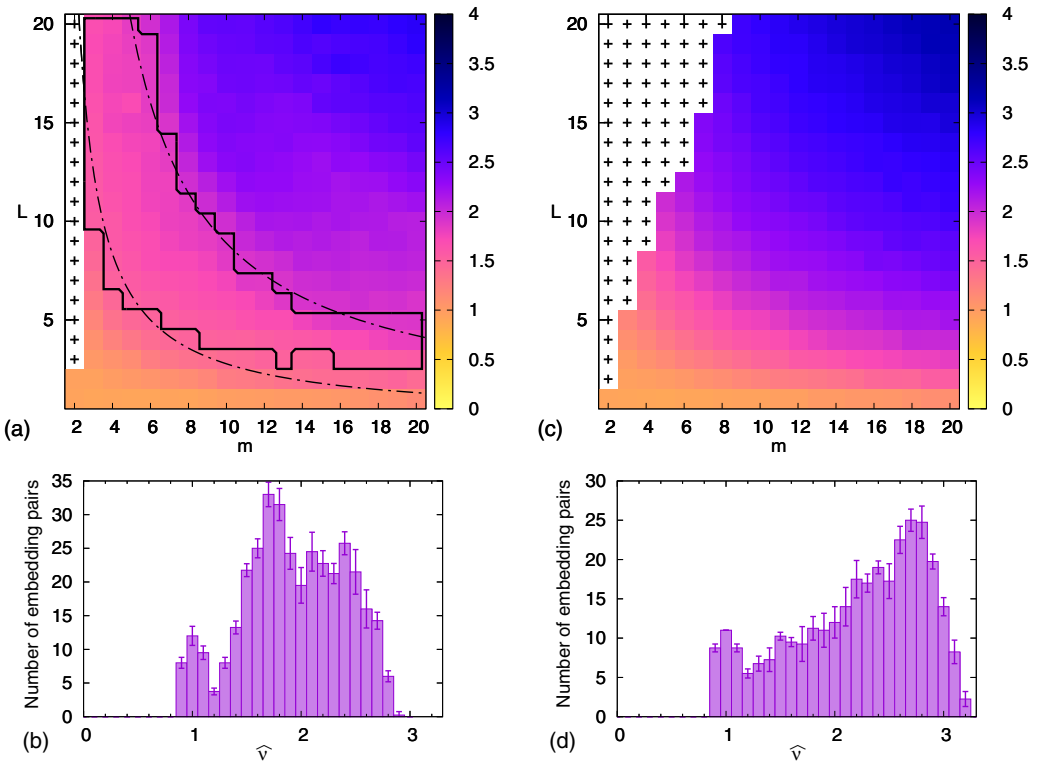

FIG. 10. Analysis of the estimated correlation dimension $\hat{\nu}$ given by Eq. (18) and computed out of a Rössler sequence generated by setting the flow parameters so as to yield a *funnel* attractor. (a) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ in the case of the Rössler sequence. The black, bold line bounds the embedding pairs that make up the peak at $\hat{\nu} = 1.7 \pm 0.2$ appearing in the histogram below. (b) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.1 wide bins. (c) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ in the case of a surrogate sequence. (d) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.1 wide bins. No significant peak can be identified. In both maps, embedding pairs incompatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ are marked with a + sign. In both histograms, bin heights and error bars correspond to the sample mean and sample standard deviation, respectively, evaluated on a set of four calculations. While the surrogate case does not show any significant $\hat{\nu}$ plateau, the peak in the histogram shown in (b) corresponds to a $\hat{\nu}$ plateau bounded by the pair of hyperbolae described by $L = 25/(m-1)$, $L = 80/(m-1)$.

towards—conventional approaches but provides a complementary tool to be used together with, rather than instead of, existing methods. In addition, our approach is in line with previous findings [5] concerning the intrinsic statistical uncertainty of the embedding parameters. The method also provides estimates for the irrelevance and the redundance times, as well as indications on the suitability of the sampling rate.

The method was tested on both synthetic and experimental sequences. Similarly to other approaches described in the scientific literature, our method is vulnerable when applied to sequences originating from quasiperiodic or intermittent systems like the Rössler attractor. However, our approach allows us to recognize systems of this kind on the basis of a characteristic feature of the maps of the estimated correlation dimension. In all other cases taken into account in the present work, also in the case of relatively short sequences, the results are in very good agreement with data reported in the scientific literature. Our method correctly distinguishes a sequence generated by a finite-dimensional chaotic source from a surrogate one that has the same power spectrum as the original sequence and randomly chosen phases. This feature can be particularly useful in experimental investigations that implement surrogate time series analysis, for example in neuroscience and geoscience [33–36].

Our method can be also useful in research fields where a chaotic behavior has to be identified and characterized. In physiology studies, for example, the assessment of different dynamical regimes by relying on recorded signals such as heart rate variability can help improving clinical diagnostics [37]. In geophysics, the possible role of low-dimensional chaotic sources or stochastic ones in the generation of seismic sequences is still debated [38,39]. Similar issues are also investigated in economics with regard to the analysis of stock exchange time series [40,41]. Finally, lasers and photonic systems exhibiting a chaotic behavior are receiving increasing attention, due to applications that range from secure communications [42] to random number generation [43] (see [44] for a recent review on the topic). Also within this research framework, our method can deliver useful information for a correct sampling and embedding of experimentally recorded time series.

An improved identification of the plateaulike regions bounded by hyperbolae within maps of the estimated correlation dimension makes up a possible, further development of the method described in this work.

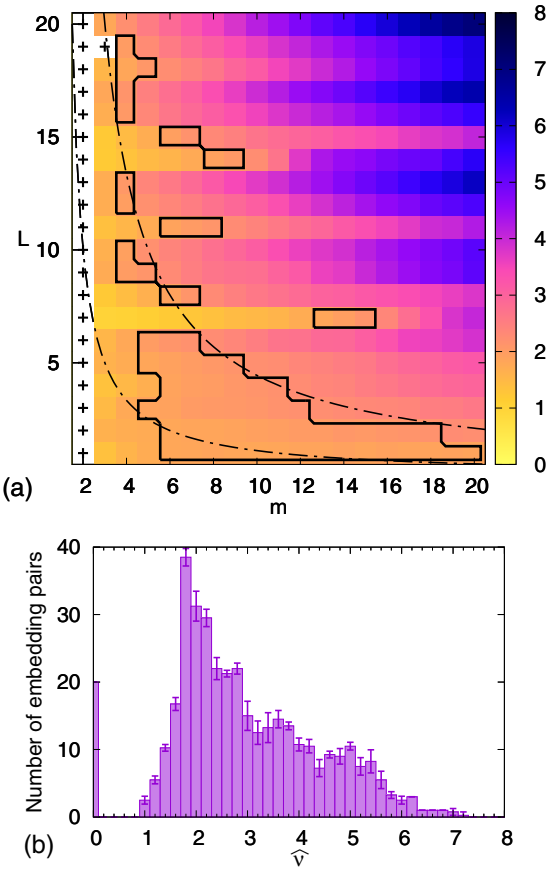

FIG. 11. Analysis of the estimated correlation dimension $\hat{\nu}$ given by Eq. (18) and computed out of the experimental sequence corresponding to the intensity of a 81.5 μm NH_3 laser. (a) Map of $\hat{\nu}$ estimated on the embedding pairs that exhibit compatibility with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$. Embedding pairs incompatible with the null hypothesis $H_{0,\mathrm{cfd}}(\nu)$ are marked with a + sign. The black, bold line bounds the embedding pairs that make up the peak between $\hat{\nu} = 1.8$ and $\hat{\nu} = 2.2$ appearing in the histogram below. The two black, dash-dotted lines correspond to hyperbolae approximatively enclosing the $\hat{\nu}$ plateau and described by $L = 10/(m-1)$, $L = 40/(m-1)$. (b) Histogram of the number of embedding pairs whose $\hat{\nu}$ falls within 0.2 wide bins. Bin heights and error bars correspond to the sample mean and sample standard deviation, respectively, evaluated on a set of four calculations.

APPENDIX: ESTIMATED CORRELATION DIMENSION IN THE CASE OF NOISY, PERIODIC SIGNALS

The goal of this appendix is to investigate the behavior of the correlation dimension, estimated by means of the method of Sec. V, in the case of quasiperiodic systems. As a model we consider here a sinusoidal signal affected by an additive noise component.

Let the sequence $\{x_i\}$ be given by

$$x_i = \sin(\phi_0 i) + \frac{\eta_i}{\sqrt{2}}, \tag{A1}$$

where the terms η_i are supposed to be i.i.d. and to have vanishing mean and standard deviation η. By virtue of the factor $1/\sqrt{2}$, the standard deviation η corresponds to the reciprocal of the signal-to-noise ratio.

Inserting the sequence of Eq. (A1) into Eq. (1) yields

$$d_{i,j}^2 = \frac{1}{m}\sum_{k=0}^{m-1}\left\{2\cos\left[\left(\frac{i+j}{2}+kL\right)\phi_0\right] \times \sin\left(\frac{i-j}{2}\phi_0\right) + \frac{\eta_{i+kL}-\eta_{j+kL}}{\sqrt{2}}\right\}^2.$$

The previous expression can be written as the sum of three contributions:

$$\begin{aligned} d_{i,j}^2 = {} & 4\sin^2\left(\frac{i-j}{2}\phi_0\right)\frac{1}{m}\sum_{k=0}^{m-1}\cos^2\left[\left(\frac{i+j}{2}+kL\right)\phi_0\right] \\ & + 2\sqrt{2}\sin\left(\frac{i-j}{2}\phi_0\right)\frac{1}{m}\sum_{k=0}^{m-1}(\eta_{i+kL}-\eta_{j+kL}) \\ & \times\cos\left[\left(\frac{i+j}{2}+kL\right)\phi_0\right] + \frac{1}{2m}\sum_{k=0}^{m-1}(\eta_{i+kL}-\eta_{j+kL})^2. \end{aligned} \tag{A2}$$

With regard to the first, noise-free contribution, which is henceforth referred to as y, provided that m is sufficiently large—namely, larger than a suitable value m_0—the probability that all the cosine terms within the sum are simultaneously small is negligible. Consequently, small-scale contributions to the correlation integral primarily occur through the squared sine term external to the sum, when the argument is close to an integer multiple of π. The fact that the indexes i, j are uniformly distributed has two consequences: first, the sum of the squared cosines turns out to be approximately 1/2; second, the noise-free contribution y is distributed according to the cumulative distribution $C_y(y) = \frac{2}{\pi}\arcsin\sqrt{\frac{y}{2}}$. Henceforth the terms η_i are assumed to be normally distributed, i.e., $\eta_i \sim \mathcal{N}(0, \eta^2)$. The second contribution is therefore also normally distributed and can be expressed as $2z\eta/\sqrt{m}$ with $z \sim \mathcal{N}(0, 1)$. In addition, the third contribution is distributed as a χ^2 variable with m degrees of freedom, i.e., as $\eta^2\chi_m^2/m$.

In the limit $m \to \infty$, the second, normal contribution of Eq. (A2) becomes negligible because of the factor $1/\sqrt{m}$. In the same limit, the third contribution tends to a constant because the expected value of χ_m^2/m tends to 1 and its standard deviation to 0. Consequently, in the limit of large m, the squared distances $d_{i,j}^2$ can be written as the sum of the noise-free contributions y and the noise variance, i.e.,

$$d_{i,j}^2 \simeq y + \eta^2.$$

The cumulative distribution of the distance $d_{i,j}$—or, as in Sec. II, $r = d_{i,j}$—is then given by

$$C(r) = \frac{2}{\pi}\arcsin\sqrt{\frac{r^2-\eta^2}{2}}.$$

According to the method discussed in Sec. V, the corresponding estimated correlation dimension ν is given

by

$$\nu = \left.\frac{d\ln[C(r)]}{d\ln r}\right|_{r=r_0} = \frac{r_0^2}{\sqrt{r_0^2-\eta^2}\sqrt{2+\eta^2-r_0^2}\,\arcsin\sqrt{\frac{r_0^2-\eta^2}{2}}}, \quad \text{(A3)}$$

where r_0 is the value r such that $\rho_0 = \rho(r_0)$ corresponds to the first maximum of the correlation bridge [see Eq. (17)]. In the limit $m \to \infty$, r_0 can be estimated as follows. The gauge transformation $\rho = \rho(r)$ corresponds to the cumulative distribution of the χ^2_m-distributed random variable $\frac{mr^2}{2\sigma^2}$ [see Eq. (4)]. By virtue of the central limit theorem, this cumulative distribution can be approximated by a normal cumulative distribution, namely

$$\rho(r) = \frac{1}{\Gamma\left(\frac{m}{2}\right)}\gamma\left(\frac{m}{2}, \frac{1}{2}\frac{mr^2}{2\sigma^2}\right) \approx \Phi\left(\frac{\frac{mr^2}{2\sigma^2}-m}{\sqrt{2m}}\right) = \Phi\left(\frac{r^2-2\sigma^2}{2\sigma^2\sqrt{2/m}}\right), \quad \text{(A4)}$$

where $\Phi(z)$ is the cumulative standard normal distribution. The term $r^2 - 2\sigma^2$ can be expanded as $(r-\sqrt{2}\sigma)(r+\sqrt{2}\sigma)$ and, because the distribution is significantly larger than zero only for $r \approx \sqrt{2}\sigma$, it can be replaced by $2\sqrt{2}\sigma(r-\sqrt{2}\sigma)$. Consequently, Eq. (A4) can be further approximated as

$$\rho(r) \approx \Phi\left(\frac{r-\sigma\sqrt{2}}{2\sigma^2\sqrt{2/m}}2\sqrt{2}\sigma\right) = \Phi\left(\frac{r-\sigma\sqrt{2}}{\sigma/\sqrt{m}}\right).$$

In the limit $m \to \infty$, the previous expression becomes a step function centered in $\sigma\sqrt{2}$ while its derivative becomes a Dirac delta function, i.e.,

$$\frac{d\rho}{dr} \approx \delta(r-\sigma\sqrt{2}). \quad \text{(A5)}$$

Finally, if $m \to \infty$, the root r_0 of Eq. (17) is $\sigma\sqrt{2}$ independently of the shape of the correlation integral $C_{m,L}(r)$.

In the case of the sinusoidal signal affected by an additive noise component as described by Eq. (A1), due to the contributions of both the sine and the superimposed noise, the sequence variance is given by $\sigma^2 = (1+\eta^2)/2$. In combination with Eq. (A5), this result leads to $r_0^2 = 1+\eta^2$. By inserting this value within Eq. (A3), the estimated ν becomes

$$\nu = \frac{4}{\pi}(1+\eta^2) \quad (\text{as } m \to \infty). \quad \text{(A6)}$$

This last expression provides the height of a rectangular $\hat{\nu}$ plateau that spans any value of L and sufficiently large values of m ($m > m_0$). Because of the central limit theorem, Eq. (A6) is valid not only for a Gaussian noise but for any kind of noise having vanishing mean and variance η^2.

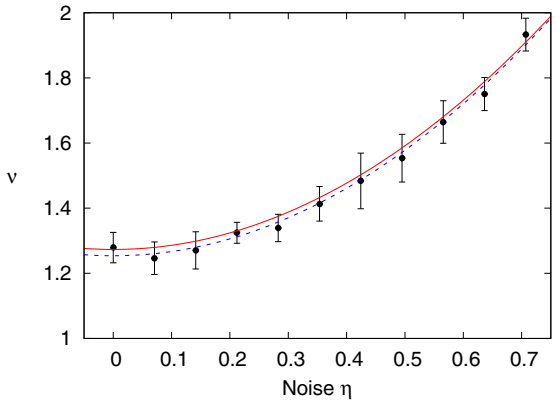

FIG. 12. Estimated ν as a function of the inverse signal-to-noise ratio η for sinusoidal sequences with superimposed Gaussian white noise as in Eq. (A1). Data points and the corresponding error bars are obtained as the sample mean and the sample standard deviation of a set of 11 $\hat{\nu}$ values estimated by setting $m = 5000$ and with L varying from 5 to 15. The red, solid line is the curve corresponding to Eq. (A6), i.e., the asymptotic prediction for $m \to \infty$ of the estimated ν. The blue, dashed line is the result of a fitting procedure described in the main text.

Figure 12 shows the numerical results of ν estimations computed by applying the method of Sec. V to synthetic time series described by Eq. (A1) and for different values of η. For each single value η, the point and the related error bar respectively correspond to the sample mean and the sample standard deviation of a set of 11 $\hat{\nu}$ values computed by setting $m = 5000$ and with L varying from 5 to 15. The red, solid line is the analytical curve of Eq. (A6). The blue, dashed line is the curve resulting from fitting the parabola $\nu(\eta) = (4/\pi)(a + b\eta^2)$ to the data points. The best-fit parameters are $a = 0.985 \pm 0.006$ and $b = 1.02 \pm 0.03$, thus proving the reliability of the model described above.

[1] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Phys. Rev. Lett. **45**, 712 (1980).

[2] F. Takens, in *Dynamical Systems and Turbulence (Warwick 1980)*, edited by D. A. Rand and L. S. Young (Springer, Berlin, 1980), pp. 366–380.

[3] E. Bradley and H. Kantz, Chaos **25**, 097610 (2015).

[4] C. J. Cellucci, A. M. Albano, and P. E. Rapp, Phys. Rev. E **67**, 066210 (2003).

[5] M. Franchi and L. Ricci, Phys. Rev. E **90**, 062920 (2014).

[6] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).

[7] M. Ding, C. Grebogi, E. Ott, T. Sauer, and J. A. Yorke, Phys. Rev. Lett. **70**, 3872 (1993).

[8] D. Kugiumtzis, Physica D **95**, 13 (1996).

[9] M. Casdagli, S. Eubank, J. D. Farmer, and J. Gibson, Physica D **51**, 52 (1991).

[10] H. S. Kim, R. Eykholt, and J. D. Salas, Physica D **127**, 48 (1999).

[11] A. M. Albano, P. E. Rapp, and A. Passamante, Phys. Rev. E **52**, 196 (1995).

[12] J. Theiler, Phys. Rev. A **34**, 2427 (1986).

[13] J. Gao and Z. Zheng, Phys. Lett. A **181**, 153 (1993).
[14] J. C. Sprott and G. Rowlands, Int. J. Bifurcation Chaos **11**, 1865 (2001).
[15] J. E. Angus, SIAM Rev. **36**, 652 (1994).
[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, UK, 1997).
[17] A. Krakovská, K. Mezeiová, and H. Budáčová, J. Complex Syst. **2015**, 932750 (2015).
[18] Y. Ueda, J. Stat. Phys. **20**, 181 (1979).
[19] S. Coles, *An Introduction to Statistical Modeling of Extreme Values* (Springer-Verlag, London, UK, 2001).
[20] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, Physica D **65**, 117 (1993).
[21] H. Kantz, Phys. Lett. A **185**, 77 (1994).
[22] H. Kantz, G. Radons, and H. Yang, J. Phys. A: Math. Theor. **46**, 254009 (2013).
[23] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica **15**, 9 (1980).
[24] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica **15**, 21 (1980).
[25] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Physica D **58**, 77 (1992).
[26] M. Paluš, Contemp. Phys. **48**, 307 (2007).
[27] O. E. Rössler, Phys. Lett. **57A**, 397 (1976).
[28] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, UK, 2004).
[29] D. Farmer, J. Crutchfield, H. Froehling, N. Packard, and R. Shaw, Ann. N.Y. Acad. Sci. **357**, 453 (1980).
[30] U. Hübner, N. B. Abraham, and C. O. Weiss, Phys. Rev. A **40**, 6354 (1989).
[31] See https://www.comp-engine.org/#!visualize/4057f818-3874-11e8-8680-0242ac120002 for public research data. The time series contains 9093 points.
[32] F. Camastra and M. Filippone, Neural Comput. Applic. **18**, 1021 (2009).
[33] T. Schreiber and A. Schmitz, Physica D **142**, 346 (2000).
[34] R. Quian Quiroga, T. Kreuz, and P. Grassberger, Phys. Rev. E **66**, 041904 (2002).
[35] G. C. O'Neill, E. L. Barratt, B. A. E. Hunt, P. K. Tewarie, and M. J. Brookes, Phys. Med. Biol. **60**, R271 (2015).
[36] A. Perinelli, D. E. Chiari, and L. Ricci, Chaos **28**, 063127 (2018).
[37] N. Wessel, H. Malberg, R. Bauernschmitt, and J. Kurths, Int. J. Bifurcation Chaos **17**, 3325 (2007).
[38] W. Marzocchi, F. Mulargia, and G. Gonzato, J. Geophys. Res.: Solid Earth **102**, 3195 (1997).
[39] A. C. Iliopoulos, G. P. Pavlos, E. E. Papadimitriou, D. S. Sfiris, M. A. Athanasiou, and V. G. Tsoutsouras, Int. J. Bifurcation Chaos **22**, 1250224 (2012).
[40] W. A. Barnett and A. Serletis, J. Econ. Dyn. Control **24**, 703 (2000).
[41] W. A. Barnett, A. Serletis, and D. Serletis, Macroecon. Dyn. **19**, 1749 (2015).
[42] A. Argyris, D. Syvridis, L. Larger, V. Annovazzi-Lodi, P. Colet, I. Fischer, J. García-Ojalvo, C. R. Mirasso, L. Pesquera, and K. A. Shore, Nature **438**, 343 (2005).
[43] A. Uchida, K. Amano, M. Inoue, K. Hirano, S. Naito, H. Someya, I. Oowada, T. Kurashige, M. Shiki, S. Yoshimori, K. Yoshimura, and P. Davis, Nat. Photon. **2**, 728 (2008).
[44] M. Sciamanna and K. A. Shore, Nat. Photon. **9**, 151 (2015).