Statistical properties of the maximum Lyapunov exponent calculated via the divergence rate method

Matteo Franchi and Leonardo Ricci*
Dipartimento di Fisica, Università di Trento, I-38123 Trento, Italy
(Received 8 August 2014; published 29 December 2014)

The embedding of a time series provides a basic tool to analyze dynamical properties of the underlying chaotic system. To this purpose, the choice of the embedding dimension and lag is crucial. Although several methods have been devised to tackle the issue of the optimal setting of these parameters, a conclusive criterion to make the most appropriate choice is still lacking. An accepted procedure to rank different embedding methods relies on the evaluation of the maximum Lyapunov exponent (MLE) out of embedded time series that are generated by chaotic systems with explicit analytic representation. The MLE is evaluated as the local divergence rate of nearby trajectories. Given a system, embedding methods are ranked according to how close such MLE values are to the true MLE. This is provided by the so-called standard method in a way that exploits the mathematical description of the system and does not require embedding. In this paper we study the dependence of the finite-time MLE evaluated via the divergence rate method on the embedding dimension and lag in the case of time series generated by four systems that are widely used as references in the scientific literature. We develop a completely automatic algorithm that provides the divergence rate and its statistical uncertainty. We show that the uncertainty can provide useful information about the optimal choice of the embedding parameters. In addition, our approach allows us to find which systems provide suitable benchmarks for the comparison and ranking of different embedding methods.

DOI: 10.1103/PhysRevE.90.062920 PACS number(s): 05.45.Pq, 05.45.Tp

I. INTRODUCTION

Embedding data from a time series is an important tool to investigate the dynamical properties of the underlying chaotic system. As stated by the Takens-Mañè embedding theorem [1], these properties can be inferred starting from a sufficiently large sample of the time series, provided a suitable pair of integer parameters, the dimension m and the lag L, is chosen [2]. Regrettably, the Takens-Mañè theorem gives no clues on how to choose the embedding parameter pair (m,L).

Several different methods have been devised to address the issue of the parameter choice. A review of these methods is given by Cellucci *et al.* [2]. In their work, methods were compared by analyzing two dynamical properties: The maximum Lyapunov exponent (MLE) and the robustness to noise. Robustness to noise is determined in terms of impact on the cumulative distribution of interpoint distances in the embedding space. However, the embedding parameters (m,L) are found to vary when noise is added to the time series and it is not clear whether a wrong embedding choice produces a lesser degree of robustness.

The determination of the MLE provides a more decisive way to rank the quality of different embedding methods. The reason is the existence of the so-called standard method [3–5], a technique that, independently of any embedding, allows for determining the standard MLE $\chi_{\mathcal{S}}$ of a dynamical system $\mathcal{S}$ by directly exploiting its analytic description. Given an (m,L)-embedded time series, i.e., embedded by using a particular parameter pair (m,L), the MLE can be estimated by determining the local divergence rate of nearby trajectories [6–8] (see [9] for a recent comprehensive review on this topic). Henceforth, this estimate is referred to as $\Gamma_{\mathcal{S}}(m,L)$. So, according to Cellucci *et al.*, method A is considered to be superior to method B if $\Gamma_{\mathcal{S}}(m_A,L_A)$, calculated by using the embedding pair assessed via A, approaches $\chi_{\mathcal{S}}$ better than $\Gamma_{\mathcal{S}}(m_B,L_B)$, calculated by using the embedding pair assessed via B.

The key parameter that determines the validity of the ranking procedure described above is the uncertainty $\sigma_{\Gamma,\mathcal{S}}(m,L)$ of the calculation of $\Gamma_{\mathcal{S}}(m,L)$: If the variation of $\Gamma_{\mathcal{S}}(m,L)$ *between* different embedding points does not significantly exceed the uncertainty affecting the MLE calculations *within* each single embedding point, namely, $\sigma_{\Gamma,\mathcal{S}}(m,L)$, the dynamical system turns out to be unsuitable for the sake of comparing different embedding methods. The origin of the statistical uncertainty, or noise, that affects finite-time MLE evaluations was first discussed by Grassberger *et al.* [10]. By using an argument based on the central limit theorem, they showed that, at least for attractors with short correlation time, noise is normally distributed and has an amplitude that depends on the duration of the time sequence on which MLE is evaluated or, equivalently, on the number of time steps considered. For intermittent attractors, the noise distribution departs from normal behavior and typically shows positive skewness due to exponential tails [11]. Consequently, provided the embedding choice is correct, the noise behavior should be reflected in the distribution of the MLE estimated by determining the local divergence rate $\Gamma_{\mathcal{S}}(m,L)$. In addition, the mean of distributions corresponding to different, correct embedding choices is necessary to approach the same value $\chi_{\mathcal{S}}$. The uncertainty $\sigma_{\Gamma,\mathcal{S}}(m,L)$ should show a similar behavior, i.e., uncertainties corresponding to different, correct embedding choices are to tend to the same value and follow the same distribution.

In this paper we consider time series generated by four reference systems widely discussed in the scientific literature and for which the standard method is applicable. We first calculate $\Gamma_{\mathcal{S}}(m,L)$ and its uncertainty $\sigma_{\Gamma,\mathcal{S}}(m,L)$ for the embedding points (m,L) belonging to the lattice $2 \leqslant m \leqslant 10$, $1 \leqslant L \leqslant 10$. The calculation of the MLE relies on an algorithm that automatically identifies the longest straight segment

*leonardo.ricci@unitn.it

on the growing section of the time-dependent divergence exponent. Second, we study the dependence of both Γ_S and $\sigma_{\Gamma,S}$ on the embedding parameters m and L.

We find that, below a suitable threshold, squared uncertainty values $\sigma^2_{\Gamma,S}(m,L)$ are approximately normally distributed. In addition, embedding points showing an uncertainty lying significantly outside this distribution are more likely to provide $\Gamma_S(m,L)$ values far from χ_S. Thus, uncertainty can indeed be used to sort out (m,L) pairs that are likely to provide a reliable embedding. Finally, by using an approach based on statistical hypothesis testing, we show that, for example, the Rössler system is less suitable than the Mackey-Glass system as a benchmark to compare different embedding methods.

The paper is organized as follows. The reference dynamical systems used in the present paper are the topic of Sec. II. Our implementation of the divergence rate method to determine the MLE is discussed in Sec. III. The assessment of the distribution of the squared uncertainty is presented in Sec. IV. Consequences with regard to optimal choice of the embedding parameters and to the suitability of a dynamical system as a benchmark for the comparison and ranking of different embedding methods are discussed in Sec. V.

II. REFERENCE SYSTEMS

In the present paper we used sample time sequences of 10^6 points generated by calculating the evolution of four dynamical systems [6,12]: The discrete Hénon map and the Lorenz, Mackey-Glass, and Rössler continuous systems. The Hénon map is governed by the following system of equations [13]:

$$\begin{aligned} x_{n+1} &= 1 - ax_n^2 + y_n, \\ y_{n+1} &= bx_n. \end{aligned} \tag{1}$$

Here we used $a = 1.4$ and $b = 0.3$. The Lorenz attractor is governed by the following system of equations [14]:

$$\begin{aligned} \frac{dx}{dt} &= \sigma(y - x), \\ \frac{dy}{dt} &= x(r - z) - y, \\ \frac{dz}{dt} &= xy - bz. \end{aligned} \tag{2}$$

Here we used $\sigma = 10$, $r = 45.92$, $b = 8/3$, and a sampling time $\delta t = 0.03$. The Rössler attractor is governed by the following system of equations [15]:

$$\begin{aligned} \frac{dx}{dt} &= -y - z, \\ \frac{dy}{dt} &= x + ay, \\ \frac{dz}{dt} &= b + z(x - c). \end{aligned} \tag{3}$$

Here we used $a = 0.15$, $b = 0.2$, $c = 10$, and $\delta t = 0.125$. The Mackey-Glass attractor is governed by the following equation [16]:

$$\frac{dx(t)}{dt} = a\frac{x(t-\tau)}{1 + [x(t-\tau)]^c} - bx(t), \quad a,b,c > 0. \tag{4}$$

Here we used $a = 0.2$, $b = 0.1$, $c = 10$, $\tau = 30$, and $\delta t = 1.5$.

For the multidimensional systems (1)–(3), the time sequences to be embedded were generated by sampling the x component only. The integration of the systems of differential equations (2)–(4) was carried out by using a sixth-order Runge-Kutta algorithm and taking the sampling time δt as the integration time step. The first point of each sequence was assigned by means of a random generator.

III. IMPLEMENTATION OF THE DIVERGENCE RATE METHOD

Starting from a sample time sequence $\{x_1,x_2,x_3,\ldots,x_n\}$ of a scalar, real variable x, the embedding procedure consists in constructing a set of m-dimensional vectors $\mathbf{X}_i$ such that the lth component of $\mathbf{X}_i$ $(0 \leqslant l < m)$ is given by x_{i+lL}. Given an embedding pair (m,L), we consider a set of $N = 500$ randomly selected pairs $\mathbf{X}_i,\mathbf{X}_j$ of m-dimensional vectors. The selection occurs without replacement. Each $\mathbf{X}_i,\mathbf{X}_j$ pair must satisfy the two following conditions [2]: (i) $\|\mathbf{X}_i - \mathbf{X}_j\| \leqslant r$, with r corresponding to a given percentile of the distribution of all Euclidean distances $\|\mathbf{X}_i - \mathbf{X}_j\|$, and (ii) $|i - j| \geqslant c_0$, where c_0 is a constraint linked to the autocorrelation function. Theiler [17,18] suggests to take c_0 equal to the autocorrelation time, whereas Gao and Zheng [6] suggest to take $c_0 = (m - 1)L$. Here we set c_0 to the first zero of the autocorrelation function multiplied times $m(L + 1)$.

The time-dependent divergence exponent $\Lambda(k)$ is then defined as

$$\Lambda(k) \equiv \left\langle \ln \frac{\|\mathbf{X}_{i+k} - \mathbf{X}_{j+k}\|}{\|\mathbf{X}_i - \mathbf{X}_j\|} \right\rangle,$$

where k is a non-negative integer delay and the average $\langle\cdots\rangle$ is taken on the N-dimensional set of $\mathbf{X}_i,\mathbf{X}_j$ pairs. From the definition it follows that $\Lambda(0) = 0$.

An important step in our algorithm is that, rather than considering a single calculation for each value of k, we determine each point of $\Lambda(k)$ and the respective uncertainty $\sigma_\Lambda(k)$ as the pointwise sample mean and sample standard deviation of a set of M different calculations, respectively,

$$\Lambda(k) \equiv \frac{1}{M}\sum_{i=1}^{M} \Lambda_i(k),$$

$$\sigma_\Lambda(k) \equiv \frac{1}{M-1}\sum_{i=1}^{M} [\Lambda_i(k) - \Lambda(k)]^2.$$

In the following $M = 50$ and the maximum value of the delay k was set to 100 for the Hénon map, 200 for the Lorenz attractor, and 400 both for the Mackey-Glass and the Rössler attractors.

Figure 1 shows the time-dependent divergence exponent $\Lambda(k)$ for the Lorenz attractor and different values of r. Typically, there is a k range within which the dependence of Λ on k turns out to be linearly growing and independent of the boundary conditions. The corresponding value of the MLE is then given by the slope of $\Lambda(k)$ divided by δt ($\delta t = 1$ for maps). The slopes of $\Lambda(k)$ tend to increase for progressively smaller r. The reason is that the smaller r is, the smaller the probability that two close points belong to trajectories that

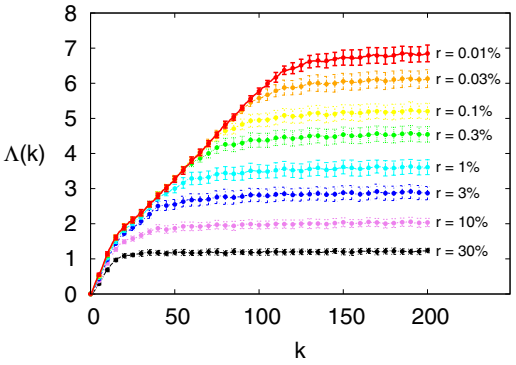

FIG. 1. (Color online) Time-dependent divergence exponent $\Lambda(k)$ for the Lorenz attractor, obtained by using the embedding pair (3,3). The different plots correspond to eight different values of r: 0.01%, 0.03%, 0.1%, 0.3%, 1%, 3%, 10%, and 30%. For the sake of clarity, only one point every five is plotted.

are still approaching each other. The analysis discussed below refers to data obtained by setting $r = 0.01\%$.

Finally, to determine $\Gamma_{\mathcal{S}}(m,L)$, the following automatic procedure is used.

(i) The plateau Λ_∞ of $\Lambda(k)$ at $k \gg 1$ is estimated by averaging the ten largest available k values.

(ii) Thereupon, the lowest k, referred to as k_{slope}, is determined such that $\Lambda(k) > 0.9\Lambda_\infty$.

(iii) Given the growing section of $\Lambda = \Lambda(k)|k \leqslant k_{\text{slope}}$, we determine the point P of abscissa k_P that allows for the longest straight-line segment, centered on P, that fits the data with a reduced χ^2 not larger than $1 + \sqrt{8/(2\ell-1)}$, where 2ℓ is the length of the segment (in other words, a fit is assumed to be valid if the corresponding χ^2 does not exceed its expected value by twice its standard deviation).

(iv) If more adjacent points satisfy the requirement of step (iii), point P is selected by considering the minimum χ^2.

(v) To increase accuracy, a final straight-line fit is carried out; the fitting segment is centered in P and has a length equal to $2\ell'$, where $\ell' \equiv \lceil \ell/2 \rceil$; the fit result is taken to be valid if $\ell' \leqslant 2$, i.e., if the new length is at least 4.

(vi) The estimated MLE $\Gamma_{\mathcal{S}}(m,L)$ is set to the straight-line slope divided by the time step δt of the sample time sequence; accordingly, the uncertainty $\sigma_{\Gamma,\mathcal{S}}(m,L)$ on $\Gamma_{\mathcal{S}}(m,L)$ is set as the error on the slope divided by δt:

$$\Gamma_{\mathcal{S}}(m,L) = \frac{1}{\delta t}\frac{\overline{k\Lambda(k)} - \overline{k}\,\overline{\Lambda(k)}}{\overline{k^2} - \overline{k}\,\overline{k}}, \tag{5a}$$

$$\sigma_{\Gamma,\mathcal{S}}(m,L) = \frac{1}{\delta t}\left[(\overline{k^2} - \overline{k}\,\overline{k})\sum_{k=k_P-\ell'}^{k_P+\ell'}\frac{1}{\sigma_\Lambda^2(k)}\right]^{-1/2}, \tag{5b}$$

where, for a generic $f(k)$,

$$\overline{f(k)} \equiv \left[\sum_{k=k_P-\ell'}^{k_P+\ell'}\frac{f(k)}{\sigma_\Lambda^2(k)}\right]\left[\sum_{k=k_P-\ell'}^{k_P+\ell'}\frac{1}{\sigma_\Lambda^2(k)}\right]^{-1}. \tag{6}$$

The reference MLE $\chi_{\mathcal{S}}$ was calculated for each of the four dynamical systems by implementing the standard method [3–5]. The four values are reported in Table I. Each value corresponds to the sample mean and the sample standard deviation of the results of ten runs, each made of 10^7 integration steps: The result of a run corresponds to the very last integration step.

TABLE I. Standard value $\chi_{\mathcal{S}}$ of the MLE, calculated for each of the four dynamical systems of Sec. II by means of the standard method. The digit in parentheses corresponds to the uncertainty $\sigma_{\chi,\mathcal{S}}$ on the respective least significant digit.

Dynamical system $\mathcal{S}$	$\chi_{\mathcal{S}}$
Hénon	0.41924(9)
Lorenz	1.2346(6)
Rössler	0.08889(9)
Mackey-Glass	0.00742(2)

In the remainder of the paper, results are discussed in terms of $\Gamma_{\mathcal{S}}(m,L)$ and its uncertainty $\sigma_{\Gamma,\mathcal{S}}(m,L)$ normalized to the related standard value $\chi_{\mathcal{S}}$, namely,

$$\mu_{\mathcal{S}}(m,L) \equiv \frac{\Gamma_{\mathcal{S}}(m,L)}{\chi_{\mathcal{S}}}, \tag{7a}$$

$$\sigma_{\mathcal{S}}(m,L) \equiv \frac{1}{\chi_{\mathcal{S}}}\left[\sigma_{\Gamma,\mathcal{S}}^2(m,L) + \frac{\Gamma_{\mathcal{S}}^2(m,L)\sigma_{\chi,\mathcal{S}}^2}{\chi_{\mathcal{S}}^2}\right]^{1/2}. \tag{7b}$$

For each of the four dynamical systems, the MLE was calculated on the lattice $2 \leqslant m \leqslant 10$, $1 \leqslant L \leqslant 10$. The results are shown in Fig. 2. Apparently, points for which $\mu_{\mathcal{S}}(m,L) \approx 1$ are characterized by lower uncertainty. The investigation of this correlation is the topic of the following section.

IV. DISTRIBUTION OF UNCERTAINTY

Because $\sigma_{\Gamma,\mathcal{S}}^2(m,L) \gg \sigma_{\chi,\mathcal{S}}^2$, from Eq. (7b) it follows that $\sigma_{\mathcal{S}}(m,L) \approx \sigma_{\Gamma,\mathcal{S}}(m,L)$, so the statistical properties of $\sigma_{\mathcal{S}}(m,L)$ are mainly given by those of $\sigma_{\Gamma,\mathcal{S}}(m,L)$. As a result of the straight-line fit [see Eq. (5b)], the squared uncertainty is essentially given by the sum of contributions stemming from

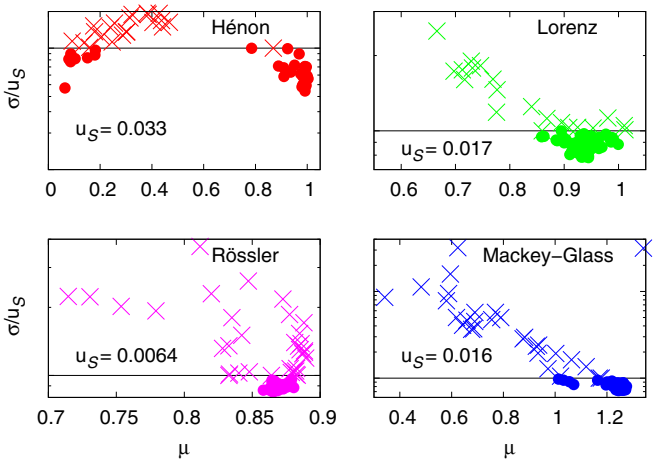

FIG. 2. (Color online) Diagram of the uncertainty $\sigma_{\mathcal{S}}(m,L)$, normalized to the respective $u_{\mathcal{S}}$ (see Sec. IV), vs the normalized MLE $\mu_{\mathcal{S}}(m,L)$. Each point corresponds to an embedding pair. Black lines correspond to $\sigma = u_{\mathcal{S}}$. Points below (above) the black lines are represented with dots (crosses). The ordinate axes are logarithmically scaled.

the variances of the points belonging to the time-dependent divergence exponent $\Lambda(k)$. By virtue of the central limit theorem, for each embedding pair (m,L), the squared uncertainty is expected to follow a normal distribution provided the number of degrees of freedom ν, which is of order ℓ, is sufficiently large. This is the case for Lorenz, Mackey-Glass, and Rössler continuous systems, for which $\ell \gtrsim 50$. The normal approximation should also work satisfactorily with regard to the Hénon map, for which $\ell \lesssim 10$. The mean value of the uncertainty can be estimated by assuming homoscedasticity with respect to k in Eqs. (5b) and (6):

$$\sigma_{\Gamma,\mathcal{S}}(m,L) \approx \frac{\sigma_\Lambda}{\Delta}\sqrt{\frac{12}{\ell}}, \tag{8}$$

where $\Delta \equiv \ell \delta t$ is the length of the time interval on which the linear fit is carried out and $\sigma_\Lambda = \sigma_\Lambda(k)$ for the k values belonging to the $\Lambda(k)$ slope.

As discussed in the Introduction (Sec. I), the intrinsic distribution of a finite-time MLE is approximately normal, so its squared uncertainty is expected to be χ^2 distributed. If one considers only the "good" embedding pairs, namely, those whose $\Lambda(k)$ plots deliver MLE values close to the standard one, both the MLE values and the corresponding squared uncertainties are expected to follow the same intrinsic distributions, independently of the values of m and L. The reason is that those $\Lambda(k)$ plots approximately have the same slope and the same length Δ. Consequently, once observational conditions are fixed (e.g., the sampling time δt), a dynamical system is characterized by a typical MLE distribution as well as by a typical uncertainty distribution and thus by a typical average uncertainty. To verify this crucial statement we proceeded as follows.

Given the system $\mathcal{S}$, let $V_\mathcal{S}(u_\mathcal{S})$ be the set of the sample variances (the squared uncertainties) that do not exceed the upper constraint $u_\mathcal{S}^2$: $V_\mathcal{S}(u_\mathcal{S}) \equiv \{\sigma_\mathcal{S}^2(m,L) | \sigma_\mathcal{S}^2(m,L) \leqslant u_\mathcal{S}^2\}$. By exploiting the Shapiro-Wilk test [19] and using a significance threshold $\alpha = 0.01$, we determined (if any) the maximum value $u_\mathcal{S}^2$ such that the set $V_\mathcal{S}(u_\mathcal{S})$ is compatible with a normal distribution. The value of $u_\mathcal{S}$ for each of the four dynamical systems is reported in Table II along with the root-mean-square value $\Sigma_\mathcal{S}$ of the uncertainties belonging to $V_\mathcal{S}(u_\mathcal{S})$.

The root mean square $\Sigma_\mathcal{S}$ turns out to be a parameter typical of the dynamical system that, despite being dependent on the observational conditions, is independent of the embedding choice. It describes the uncertainty with which the MLE can be estimated by means of the divergence rate method. The distribution of the variance for each of the four dynamical systems is shown in Fig. 3. As predicted, the red (dark gray) histograms, which correspond to the sets $V_\mathcal{S}(u_\mathcal{S})$, are bell shaped.

TABLE II. Maximum value $u_\mathcal{S}$ (column 2) such that the set $V_\mathcal{S}(u_\mathcal{S})$ of squared uncertainties that do not exceed $u_\mathcal{S}^2$ is compatible with a normal distribution according to the Shapiro-Wilk test ($p \geqslant 0.01$) and (column 3) root-mean-square value $\Sigma_\mathcal{S}$ of the uncertainties belonging to $V_\mathcal{S}(u_\mathcal{S})$.

Dynamical system $\mathcal{S}$	$u_\mathcal{S}$	$\Sigma_\mathcal{S}$
Hénon	0.033	0.023
Lorenz	0.017	0.015
Rössler	0.0064	0.0057
Mackey-Glass	0.016	0.013

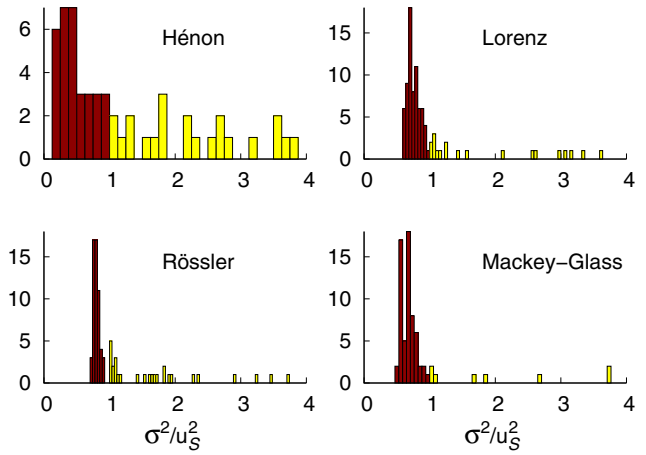

FIG. 3. (Color online) Distribution of the variance for each of the four dynamical systems. Red (dark gray) bins refer to $\sigma_\mathcal{S}(m,L)$ values that do not exceed the respective threshold $u_\mathcal{S}$. Conversely, yellow (light gray) bins refer to $\sigma_\mathcal{S}(m,L)$ values that exceed the respective threshold $u_\mathcal{S}$. For the sake of clarity, only data corresponding to $\sigma_\mathcal{S}(m,L)/u_\mathcal{S} \leqslant 4$ are shown.

V. DISCUSSION

The distribution of the normalized MLE $\mu_\mathcal{S}(m,L)$ for each of the four dynamical systems is shown in Fig. 4. Clearly, values of $\mu_\mathcal{S}(m,L)$ whose corresponding $\sigma_\mathcal{S}(m,L)$ belong to $V_\mathcal{S}(u_\mathcal{S})$ [histograms marked in red (dark gray)] tend to cluster about 1 or in regions nearby.

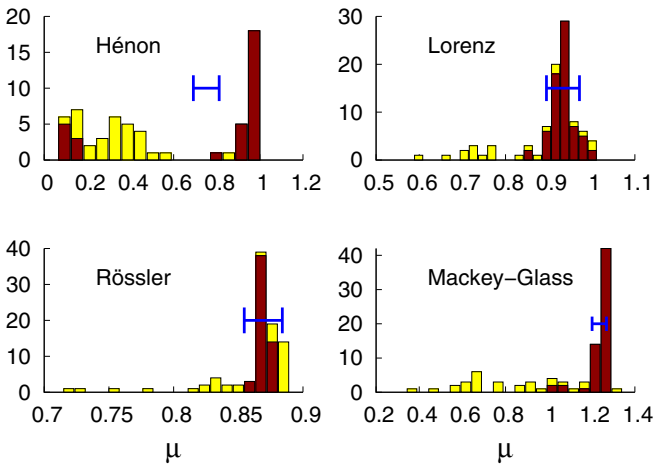

FIG. 4. (Color online) Distributions of the normalized MLE for each of the four dynamical systems: The yellow (light gray) histogram refers to all available data; the superimposed, red (dark gray) histogram refers to normalized MLE values whose corresponding $\sigma_\mathcal{S}(m,L)$ does not exceed the respective threshold $u_\mathcal{S}$ [i.e., $\sigma_\mathcal{S}^2(m,L) \in V_\mathcal{S}(u_\mathcal{S})$]. Each horizontal error bar corresponds to $5.15\Sigma_\mathcal{S}$, i.e., the width of the two-tailed 99% nonrejection region (level of significance equal to 0.01). The error bars are centered about the mean values of the red (dark gray) histograms.

Actually, the Hénon map produces an additional cluster of low uncertainty points that lie far from the optimal value $\mu = 1$. This behavior can be explained, first, by considering that most of the points in the cluster about $\mu = 1$ have $L \leqslant 2$, whereas most of the points belonging to the other "spurious" cluster correspond to lag values $L \geqslant 3$, and, second, by taking into account the discrete nature of this chaotic system: Higher-lag embeddings sample the trajectories too slowly and thus describe an aliased system.

Figure 4 also shows how potentially good points, namely, embedding points such that $\mu_{\mathcal{S}}(m,L) \approx 1$, are distributed in comparison with the typical uncertainty $\Sigma_{\mathcal{S}}$. For example, in the case of the Rössler attractor, potentially good MLE values are peaked about 0.87 and have a standard deviation of 0.005. The normalized MLE μ of the Rössler attractor is clearly slightly underestimated, probably because of its intermittency, i.e., the extremely long correlation time of this attractor. More importantly, the standard deviation is very close to the $\Sigma_{\mathcal{S}}$ for the Rössler dynamical system (see Table II) and therefore well within the corresponding 99% nonrejection region, given by $5.15\Sigma_{\mathcal{S}}$. Consequently, the Rössler attractor is not useful to compare between different embedding methods by using the MLE calculation as a gauge: Different embedding choices essentially yield the same result. On the other hand, the other three dynamical systems, and especially the Mackey-Glass attractor, appear to be more appropriate to be used as benchmarks.

Our approach can be used to evaluate the MLE of time sequences generated by unknown dynamical systems. The investigation of the distribution of the MLE calculated via divergence rate method and of the related uncertainty, for example, by identifying normally distributed clusters of values, can provide clues to optimally tune the embedding parameters and thus to reconstruct the dynamics of the underlying dynamical system.

Finally, it is worth noting that the statistical properties of the finite-time MLE evaluated by means of our algorithm comply with the theory discussed by Grassberger *et al.* [10]. In the case of continuous systems with short correlation times, if the sampling frequency δt^{-1} is sufficiently high, the time interval $\Delta = \ell \delta t$ is independent of the sampling time δt and thus of the number of points ℓ. Consequently, according to Eq. (8), the uncertainty scales as $\ell^{-1/2}$, in agreement with the results of these authors. On the other hand, in the case of intermittent systems, where correlations are long-lived, the length Δ is expected to be dependent on the sampling frequency and therefore on ℓ. This dependence changes the power-law scaling of uncertainty $\mathcal{S}(m,L)$ vs the number of points ℓ, namely, $\mathcal{S}(m,L) \sim \ell^{-p}$, so the exponent p departs from the standard value 0.5 [11,20].

The investigation of the scaling laws in combination with the divergence rate method, e.g., by changing the sampling frequency, could be exploited to provide new tools to identify optimal embedding parameters.

ACKNOWLEDGMENT

The authors thank an anonymous referee for valuable suggestions.

[1] F. Takens, in *Dynamical Systems and Turbolence, Warwick 1980*, edited by D. A. Rand and L. S. Young, Lecture Notes in Mathematics Vol. 898 (Springer, Berlin, 1981), pp. 366–381.
[2] C. J. Cellucci, A. M. Albano, and P. E. Rapp, Phys. Rev. E **67**, 066210 (2003).
[3] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica **15**, 9 (1980).
[4] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica **15**, 21 (1980).
[5] C. Skokos, Lect. Notes Phys. **790**, 63 (2010).
[6] J. Gao and Z. Zheng, Phys. Lett. A **181**, 153 (1993).
[7] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, Physica D **65**, 117 (1993).
[8] H. Kantz, Phys. Lett. A **185**, 77 (1994).
[9] H. Kantz, G. Radons, and H. Yang, J. Phys. A: Math. Theor. **46**, 254009 (2013).
[10] P. Grassberger, R. Badii, and A. Politi, J. Stat. Phys. **51**, 135 (1988).
[11] A. Prasad and R. Ramaswamy, Phys. Rev. E **60**, 2761 (1999).
[12] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Physica D **16**, 285 (1985).
[13] M. Hénon, Commun. Math. Phys. **50**, 69 (1976).
[14] E. N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
[15] O. E. Rössler, Phys. Lett. A **57**, 397 (1976).
[16] M. C. Mackey and L. Glass, Science **197**, 287 (1977).
[17] J. Theiler, Phys. Rev. A **34**, 2427 (1986).
[18] J. Theiler, J. Opt. Soc. Am. A **7**, 1055 (1990).
[19] S. S. Shapiro and M. B. Wilk, Biometrika **52**, 591 (1965).
[20] K. Kanno and A. Uchida, Phys. Rev. E **89**, 032918 (2014).