Estimating the Shannon Entropy: Recurrence Plots versus Symbolic Dynamics

Christophe Letellier

CORIA UMR 6614—Université de Rouen, Avenue de l'Université, Boîte Postale 12, F-76801 Saint-Etienne du Rouvray cedex, France

(Received 6 January 2006; published 29 June 2006)

Recurrence plots were first introduced to quantify the recurrence properties of chaotic dynamics. A few years later, the recurrence quantification analysis was introduced to transform graphical representations into statistical analysis. Among the different measures introduced, a Shannon entropy was found to be correlated with the inverse of the largest Lyapunov exponent. The discrepancy between this and the usual interpretation of a Shannon entropy is solved here by using a new definition—still based on the recurrence plots—and it is verified that this new definition is correlated with the largest Lyapunov exponent, as expected from the Pesin conjecture. A comparison with a Shannon entropy computed from symbolic dynamics is also provided.

DOI: 10.1103/PhysRevLett.96.254102 PACS numbers: 05.45.−a

Chaotic dynamics is usually defined as low dimensional deterministic dynamics with (i) a great sensitivity to initial condition and (ii) some recurrence properties. This allows an aperiodic solution bounded in the phase space to be obtained. The recurrence properties are therefore not necessarily trivial and Eckmann, Hamphorst, and Ruelle [1] introduced a graphical representation of them. A *recurrence plot* R_{ij} is a square array built as follows. Every point of the phase space trajectory $\{\boldsymbol{x}_i\}_{i=1}^N$ is tested to see whether it is close to another point $\boldsymbol{x}_j$ of the trajectory, that is, whether the distance between these two points is less than a specified threshold ϵ. In this case, the point is said to be recurrent and is represented by a black dot. Otherwise, the point is not recurrent and is represented by a white dot. This can be described as a $N \times N$ array

$$R_{ij} = \theta(\epsilon - \| \boldsymbol{x}_i - \boldsymbol{x}_j \|) \quad (1)$$

where $\theta(\boldsymbol{x}_i)$ is the Heaviside function.

A few years later, Trulla *et al.* [2] coupled the recurrence plots with various measures, helpful for transforming graphical representations into statistical analysis. With this quantification, the recurrence plots have been used to investigate more and more varied topics such as heart beats, neuron signals, hydrophobicity in prions, exchange rates, chemical reactions, ecosystems, epileptic seizures, etc., [3–10]. Among the different measures used was the Shannon entropy which was correlated with the inverse of the largest Lyapunov exponent. This is quite opposite to the usual meaning of a Shannon entropy, which is a measure of the complexity of the dynamics. We propose here a new definition of the Shannon entropy, still computed from the recurrence plot, to obtain a measure which increases when the chaotic dynamics is developed. We start by determining some parameters which can affect the results.

It has been shown that a recurrence plot analysis is optimal when the trajectory is embedded in a phase space reconstructed with an appropriate dimension d_E [11]. Such a dimension can be well estimated using a false nearest neighbors technique as introduced by Kennel *et al.* [12] or improved by Cao [13]. The d_E-dimensional phase space is then reconstructed using delay coordinates. The time delay τ can be estimated using mutual information [14] or the first zero of an autocorrelation function [15], but most of the time a visual inspection works well too. Basically, the time delay has to be as small as possible and always less than a quarter of the pseudoperiod. A parameter specific to the recurrence plot is the threshold ϵ. Many trials lead us to choose (for all the dynamics investigated) a threshold equal to $\sqrt{d_E} \times 10\%$ of the fluctuations of the signal. The threshold is therefore automatically computed from the time series investigated. Thus, only two parameters need to be determined: the embedding dimension, d_E, and the time delay, τ. When not extracted from a Poincaré section, the time series used will be sampled at τ. This appears to be a good balance between covering each oscillation and covering the whole attractor.

Among the quantifiers introduced by Trulla *et al.* [2], the Shannon entropy is defined as

$$S = -\sum_{n=1}^{H} P_n \log(P_n) \quad (2)$$

where H is the length of the maximum recurrent segment, $P_n \neq 0$ is the relative frequency of the periodic segment with length $n > 0$. Since the Shannon entropy quantifies the complexity of the dynamics, it should increase when the chaotic behavior is developed, for instance, when the bifurcation parameter μ of the logistic map $x_{n+1} = \mu x_n(1 - x_n)$ is increased. Unfortunately, this is quite opposite to the above definition [10]. This was originally pointed out by Eckmann *et al.*, since they claimed that line lengths on recurrence plots are directly related to the inverse of the largest positive Lyapunov exponent [1]. On the other hand, it is expected that the Kolmogorov-Sinaï entropy should be proportional to the largest Lyapunov exponent according to the Pesin conjecture [16].

To obtain a better estimation of the Shannon entropy from recurrence plots, denoted as $S_{\rm RP}$, the key point is to replace P_n with the relative frequency of the occurrence of the diagonal segments of *nonrecurrent* points (made of white dots). Indeed, a white dot representing a nonrecurrent point is nothing more than a signature of complexity within the data. With this definition, S_{RP} increases as the bifurcation parameter increases [as shown for the logistic map in Fig. 1(b)]. There is a one-to-one correspondence between the new definition of S_{RP} and the positive largest Lyapunov exponent. In particular, it is possible to identify periodic windows (at least the largest ones). By definition, a Shannon entropy is positive. Consequently, when the Lyapunov exponent becomes negative, S_{RP} saturates to 0.

Another common way to estimate a Shannon entropy is based on symbolic dynamics, here denoted as S_{SD} [17,18]. The advantage of S_{SD} is that it can be quickly (in terms of time computation) estimated over long time series. The Shannon entropy S_{SD} is obtained by replacing P_n with the relative frequency of the nth sequence made of k symbols. For the logistic map, the partition is not dependent on the bifurcation parameter and is given by the critical point located at the maximum of the parabola, that is, at $x_c = 0.5$. Thus, the time series $\{x_i\}_{i=1}^N$ is mapped into a sequence of symbols such as $s_i = 0$ if $x_i < 0.5$ and $S_i = 1$ if $x_i > 0.5$. The Shannon entropy S_{SD} thus estimated [Fig. 1(c)] is related to the largest Lyapunov exponent [Fig. 1(a)] as S_{RP}.

Both of these Shannon entropies depend on the parameters. From recurrence plots, computations monotonically depend on the threshold ϵ in Eq. (1) as shown in Fig. 2(a). Indeed, increasing the threshold can be seen as replacing nonrecurrent points by recurrent points; it thus decreases the entropy. A threshold around $\sqrt{d_E} \times 10\% \approx 0.17$ avoids the fluctuations observed when the chosen thresholds are too low. Safe relative comparisons can therefore be performed. Using symbolic dynamics, computations depend on the partition—not investigated here—and the sequence length k. With $N = 3000$ data points, the estimation starts to saturate around $k \approx 12$ [Fig. 2(b)]. Such a saturation results from poor statistics (3000 data points distributed over more than 3000 possible sequences). As a result, k should be chosen such that $2^k \approx N$. When S_{SD} is computed, a data set larger than for computing S_{RP} is thus required. Computing S_{RP} avoids the explicit construction of generating partitions as required when symbolic dynamics is used. Indeed, partitions require very sophisticated algorithms as detailed by Plumecoq and Lefranc [19] among others. However, the computational time for S_{RP} with a long data set ($N > 3000$) becomes significantly greater than for computing S_{SD} over the same data set.

Another advantage of the Shannon entropy S_{RP} over S_{SD} is when the noise-contaminated logistic map $x_{n+1} = \mu x_n(1 - x_n) + \zeta_n$ is investigated (Fig. 3). While S_{RP} estimated using the recurrence plots [Fig. 3(b)] presents behavior similar to that of the largest Lyapunov exponent [Fig. 3(a)], S_{SD} appears to be less robust against noise contamination. In particular, there is no longer a trace of the period-3 window, although it is still identifiable with

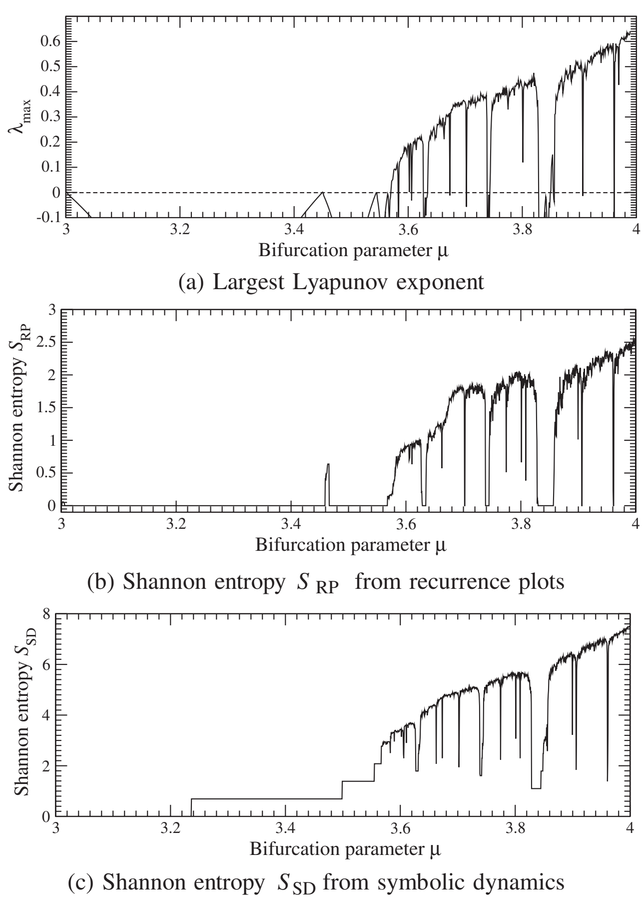


FIG. 1. Comparison between the largest Lyapunov exponent and the Shannon entropy for the logistic function. One thousand data points are used for computing S_{RP} (b) and 3000 for S_{SD} (c).

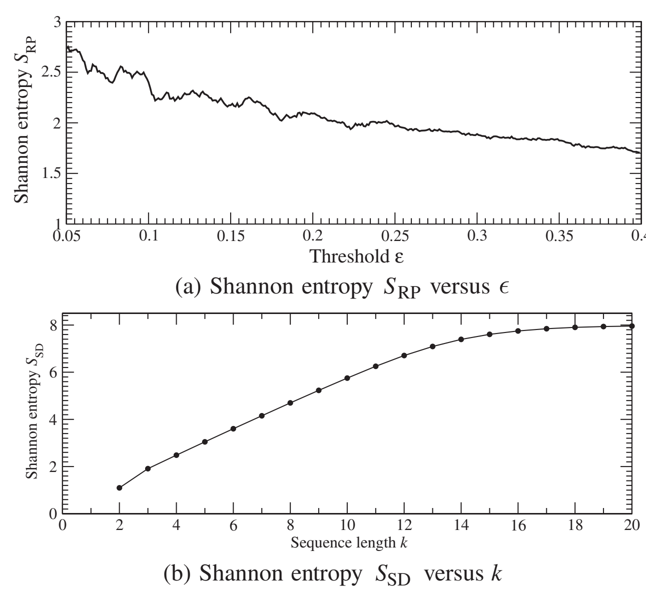


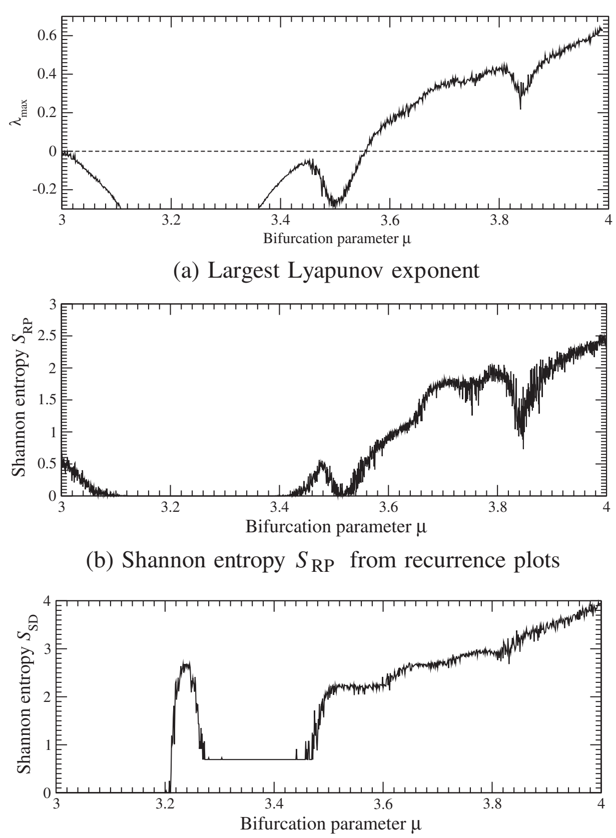
FIG. 2. Shannon entropy estimations depend on different parameters, according to the approach. ($\mu = 3.94$)

the largest Lyapunov exponent as well as with S_{RP}. Moreover, there is a strong sensitivity to the noise contamination near the bifurcation, as exemplified around $\mu = 3.25$ [Fig. 3(c)]. In fact, the noise contamination strongly affects the symbolic sequence for all points located around the critical point. This is sufficient to blurr the details of the bifurcation diagram. There is thus great advantage in computing Shannon entropy using recurrence plots since (i) there is no partition to define and (ii) the estimation is more robust against noise contamination.

Recurrence plot analysis can be used for continuous time series as well as for time series formed from the intersections of the phase trajectory with a Poincaré section. In order to check whether the analyses from continuous and discrete time series are equivalent, the Shannon entropy is now estimated for the Rössler system [$\dot{x} = -y - z$, $\dot{y} = x + ay$, and $\dot{z} = b + z(x - c)$] used with parameters $b = 2$ and $c = 4$. Parameter a will be used as a bifurcation parameter between 0.38 and 0.432, that is, over the interval corresponding to unimodal dynamics such as the logistic map [20]. The bifurcation diagram for the Rössler system [Fig. 4(a)] is thus similar to the bifurcation diagram for the logistic map. Direct comparisons between the logistic map and a Poincaré section of the Rössler system (with $a \in [0.325; 0.432]$) are thus possible.

We start by estimating Shannon entropies of the Rössler system using a Poincaré section defined as $x_n = x_c$ with $\dot{x}_n > 0$ where x_c is the coordinate of its inner fixed point. Estimations are performed with the set $\{y_n\}$ of the y coordinate of the intersections as with the data generated from the logistic map. Results are in agreement with those obtained for the logistic map [compare Figs. 4(b) and 4(c) with Figs. 1(b) and 1(c)].

The Shannon entropy from recurrence plots is now computed from the different variables of the Rössler system. It is denoted $\bar{S}_{RP}$ to avoid any confusion with S_{RP} computed in a Poincaré section. Delay coordinates with a time delay $\tau = 1.5$ s (to compare to the pseudoperiod equal to 6.2 s) are used to reconstruct a 3D phase space ($D_E = 3$ is confirmed by an estimation with a false nearest neighbor technique). For these computations, $N = 1500$ data points are retained. In spite of the data set being slightly larger compared with the 1000 points used in the

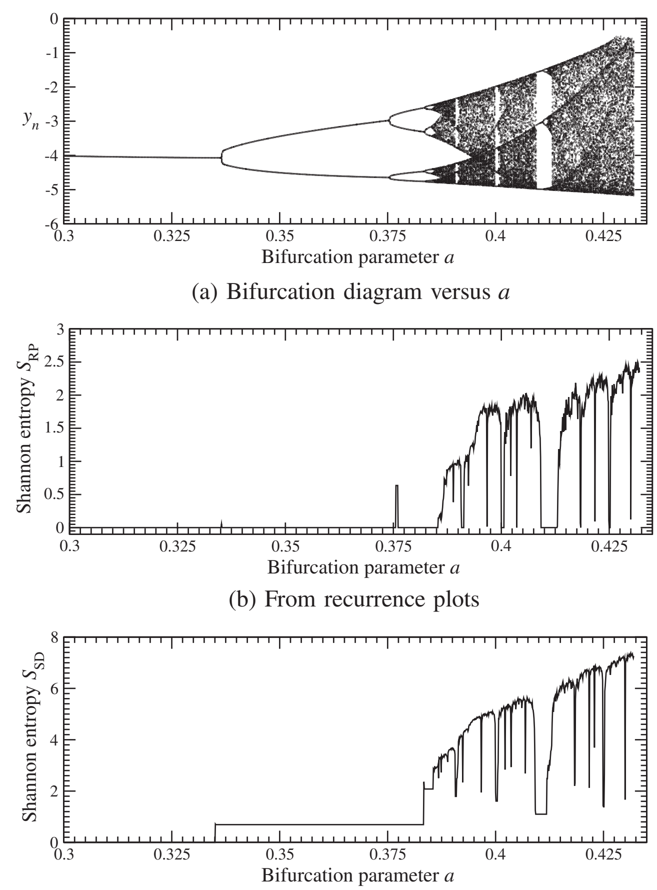


(a) Largest Lyapunov exponent

(b) Shannon entropy S_{RP} from recurrence plots

(c) Shannon entropy S_{SD} from symbolic dynamics

FIG. 3. Comparison between the largest Lyapunov exponent and the Shannon entropy for the logistic function contaminated by a multiplicative noise (0.5%).



(a) Bifurcation diagram versus a

(b) From recurrence plots

(c) From symbolic dynamics

FIG. 4. Bifurcation diagram (a) of the Rössler system computed over the interval where the dynamics is unimodal. Comparison between the Shannon entropy estimated in a Poincaré section (b) from the recurrence plots and (c) from symbolic dynamics.

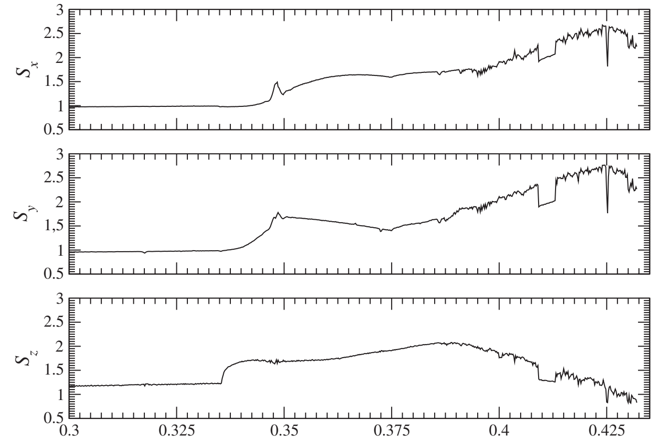


FIG. 5. Shannon entropy computed using recurrence plots from the different variables of the Rössler system. A 3D phase space is reconstructed using delay coordinates with a time delay $\tau = 1.5$ s. Fifteen hundred data points sampled at τ are used.

Poincaré section, estimations of $\bar{S}_{RP}$ from the x or the y variables (Fig. 5) are less accurate than S_{RP} computed in a Poincaré section. The period-3 window is not clearly distinguished and the $\bar{S}_{RP}$ tends to decrease for $a > 0.425$, a tendency observed neither in the Lyapunov exponent nor in S_{RP}. Moreover, the quality of the estimation of the Shannon entropy clearly depends on the choice of the variable used. The poorest estimation is obtained from the z variable of the Rössler system in agreement with the observability of the dynamics; this observability is lower when computed from the z variable than from the two others as discussed elsewhere [21,22]. It would be therefore much more efficient to estimate the Shannon entropy working within a Poincaré section than from a "continuous" time series.

Recurrence plots can be widely used to analyze the properties of dynamics. A new definition of the Shannon entropy has been proposed. It gives a measure which increases with the complexity and which is strongly correlated to the largest Lyapunov exponent. Estimating the Shannon entropy using recurrence plots is more robust against noise contamination than when symbolic dynamics is used. Moreover, the former does not require a partition of the attractor to be constructed, not always easy to do. When estimated with recurrence plots computed from time series (not within a Poincaré section), the Shannon entropy is not very sensitive to the complexity of the dynamics. In addition to that, there is a strong dependence on the choice of the observable which results from the observability of the dynamics. We would thus recommend using estimations of Shannon entropies from time series only when a Poincaré section is too difficult to obtain.

I wish to thank W. Hamadene, L. Peyrodie, and H. Rabarimanantsoa for stimulating discussions about recurrence plots analysis. I also thank Dilys Moscato for helpful comments on the manuscript.

[1] J.-P. Eckmann, S. Oliffson Kamphorst, and D. Ruelle, Europhys. Lett. **4**, 973 (1987).
[2] L. L. Trulla, A. Giuliani, J. P. Zbilut, and C. L. Webber, Jr., Phys. Lett. A **223**, 255 (1996).
[3] J. P. Zbilut, M. Koebbe, H. Loeb, and G. Mayer-Kress, in *Proceedings of the IEEE Conference on Computers in Cardiology* (IEEE Computer Society Press, New York, 1991), p. 263.
[4] P. Faure and H. Korn, Physica (Amsterdam) **122D**, 265 (1998).
[5] J. P. Zbilut, C. L. Webber, Jr., A. Colosimo, and A. Giuliani, Protein Eng. **13**, 99 (2000).
[6] J. Belaire-Franch, D. Contreras, and L. Tordera-Lledó, Physica (Amsterdam) **171D**, 249 (2002).
[7] P. Babinec, L. Zemanová, and M. Babincová, Physica Medica **18**, 63 (2002).
[8] H. Castellini and L. Romanelli, Physica (Amsterdam) **342A**, 301 (2004).
[9] L. Parrott, Ecological Complexity **1**, 111 (2004).
[10] W. Hamadene, Ph.D. thesis, Lille University, 2005.
[11] J. Zbilut and C. Webber, Phys. Lett. A **171**, 199 (1992).
[12] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, Phys. Rev. A **45**, 3403 (1992).
[13] L. Cao, Physica (Amsterdam) **110D**, 43 (1997).
[14] A. M. Fraser and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
[15] W. Liebert and H. G. Schuster, Phys. Lett. A **142**, 107 (1989).
[16] Ya B. Pesin, *Dimension Theory in Dynamical Systems* (University of Chicago Press, Chicago, 1998).
[17] J. Kurths, A. Voss, P. Saparin, A. Witt, H. J. Kleiner, and N. Wessel, Chaos **5**, 88 (1995).
[18] A. Voss, J. Kurths, H. J. Kleiner, A. Witt, N. Wessel, P. Saparin, K. J. Osterziel, R. Scharath, and R. Dietz, Cardiovasc. Res. **31**, 419 (1996).
[19] J. Plumecoq and M. Lefranc, Physica (Amsterdam) **144D**, 259 (2000).
[20] C. Letellier, P. Dutertre, and B. Maheu, Chaos **5**, 271 (1995).
[21] C. Letellier and L. A. Aguirre, Chaos **12**, 549 (2002).
[22] C. Letellier, L. A. Aguirre, and J. Maquet, Phys. Rev. E **71**, 066213 (2005).