Journal of Statistical Physics, Vol. 114, Nos. 3/4, February 2004 (© 2004)

Complexity of Dynamics as Variability of Predictability

Ruedi Stoop,[1] Norbert Stoop,[2] and Leonid Bunimovich[3]

Received February 20, 2003; accepted August 20, 2003

We construct a complexity measure from first principles, as an average over the "obstruction against prediction" of some observable that can be chosen by the observer. Our measure evaluates the variability of the predictability for characteristic system behaviors, which we extract by means of the thermodynamic formalism. Using theoretical and experimental applications, we show that "complex" and "chaotic" are different notions of perception. In comparison to other proposed measures of complexity, our measure is easily computable, non-divergent for the classical 1-d dynamical systems, and has properties of non-overuniversality. The measure can also be computed for higher-dimensional and experimental systems, including systems composed of different attractors. Moreover, the results of the computations made for classical 1-d dynamical systems imply that it is not the nonhyperbolicity, but the existence of a continuum of characteristic system length scales, that is at the heart of complexity.

KEY WORDS: Complex systems; thermodynamic formalism; complexity measures; entropy; prediction.

The notion of complexity of dynamical systems is common and widely used. However, there exist many such notions.[(1–8)] On the one hand, this variety reflects the fact that dynamics of systems contain many aspects, each of which can be characterized by a complexity of its own. On the other hand, the still growing number of complexity measures demonstrates the lack of a notion that could be considered universal, and yet practical at the same time. We argue, that this problem can be effectively addressed and solved by a new measure of complexity for dynamical systems.

[1] INI, Physics Departement, University/ETH Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland; e-mail: ruedi@ini.phys.ethz.ch

[2] Physics Department ETH Zürich, and International School of Scientific Computing, Rand Afrikaans University, ZA-2000 Johannesburg, South Africa.

[3] Southeast Applied Analysis Center, Georgia Institute of Technology, Atlanta, Georgia 30332.

Our measure is based on the concept of neighborhoods of orbits rather than on individual trajectories. It measures the complexity of the predictability of the temporal behavior of some variable on these sets. As such, our measure integrates the complexity of the dynamics itself with the complexity of the time evolution of an observable. The classical Kolmogorov complexity[1,2] deals with the prediction (computability) of individual orbits. Our measure, in contrast, is statistical in nature; it is extracted by means of the thermodynamic formalism. It does not require explicit hierarchical analysis and is non-divergent by construction. These two properties particularly distinguish our measure from previous approaches. We demonstrate its applicability and usefulness in different examples, which include some of the most studied families of dynamical systems, as well as experimental data.

The most important practical problem in dealing with any system is to predict its evolution, i.e., the next state. By posing this problem, one already assumes that the current state of the system is known, i.e., can be measured (or computed). This aspect deals with the Kolmogorov (algorithmic) complexity. However, the current state of a natural system is practically never exactly known (e.g., there is always noise present). This suggests considering the problem of prediction on whole sets of neighboring trajectories. The prediction of the next value of an observable for given exact initial states, is a different problem. This complexity is related to the degree of chaoticity of the underlying dynamics. The most chaotic dynamical systems are considered to be the ones that behave like the most random stochastic processes, i.e., sequences of independent identically distributed (i.i.d.) random variables. However, i.i.d. random processes are not the most complex random processes, but rather the simplest ones (although non-predictable). Such processes correspond to uniformly hyperbolic dynamical systems, which are also, in a sense, the simplest among all chaotic systems.

Complexity of dynamics is reflected as a variability of predictability in the following sense: In order to obtain the next value of a process, one needs to perform some experiment. It is the complexity of this experiment that should be taken as the definition of the complexity of the system. For instance, to obtain the next value of an i.i.d. random process, one needs to throw a generalized die. However, if we consider a slightly more general random process, which is a mixture of i.i.d. random variables, then to obtain the next value one needs to throw *two* dice: The first determines which random variable is chosen, whereas the second die determines the next value of the process. From this, one can see that the key feature of complex processes (both random and deterministic ones) is their nonhomogeneity. This nonhomogeneity means that the behavior of a dynamical system is quite different in different parts of the phase space, and therefore predictability should be averaged over an appropriate measure in the phase space. Such an appropriate measure is obtained via the thermodynamic formalism.[9–11]

For the derivation of such a measure, let us consider a dynamical system with discrete time, defined by a map f on some set M in the Euclidean space $\mathbb{R}^n$. Pick an arbitrary point x_0 in the phase space, take some neighborhood $U = U(x_0)$ and consider the orbits $\{f^{(n)}U\}$, $n \in \mathbb{N}$. We are interested in observables that relate to measures that are multiplicative along the orbit, i.e., for which the n-step average is evaluated as $(\prod_{k=0}^{n-1} \mu(f^{(k)}(x)))^{1/n}$, where $x \in U$ is the initial state of a particular orbit. Examples of such measures are derivatives, probabilities, etc. Take such a measure $\mu(x)$ and define our observable v as $\mu(x) =: \exp(v(x))$. Our goal is to study the problem of prediction of the next values $v(f^{(r)}(x)), r > n$, along the orbits, and to evaluate the complexity of this task. We claim that this prediction problem is the one that defines an intrinsic complexity of dynamics. Because the initial conditions of the orbits are not perfectly known, the outcome of a measurement is not constant (when we deal with chaotic processes) and the exact value of the observable is not known (no measurement is exact), it is clear that this prediction should be probabilistic in nature.

For the decay of the probability P of retaining a particular measurement value of the observable during a system evolution of n steps, we employ the large deviation ansatz[9]

$$P(v, n)\, dv \sim e^{-ng(v)}\, dv. \tag{1}$$

The thermodynamic formalism implies[10] that

$$g(v) = v - S(v), \tag{2}$$

where $S(v)$ is an entropy function. In more detail, the thermodynamic formalism departs from a partition function $Z(n, \beta, v)$, where n is the level or depth of the partition and β can be viewed as an inverse temperature. With $Z(n, \beta, v)$, a free energy $F(\beta) = \lim_{n \to \infty} \frac{1}{n} \log(Z(n, \beta, v))$ is associated, where in $F(\beta)$ we suppressed the dependence on the observable. In the absence of phase transitions, an entropy function is obtained by means of the Legendre transform

$$S(v) = v\beta - F(\beta). \tag{3}$$

Requirements that apply to entropy functions are strict convexity with infinite derivatives at the two end-points of the curve (in the absence of phase transition effects).

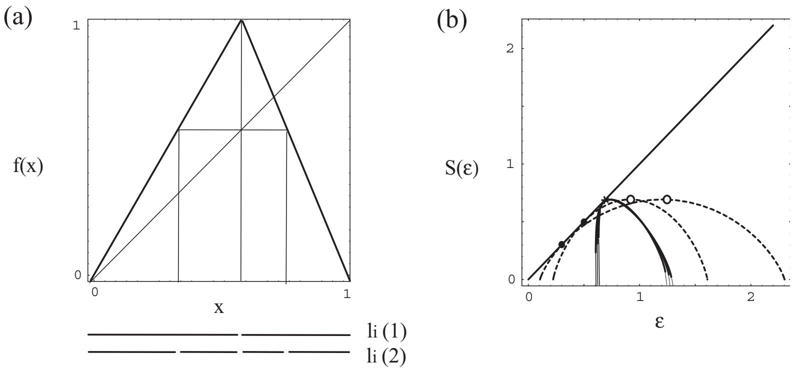

Fig. 1. (a) Unit interval partition into lengths $l_i(n)$, for increasing partition levels $n = 1, 2$. (b) Dashed lines: Entropy functions of the asymmetric tent map, for two values of asymmetry. Filled circles: results for the natural entropy measure (Lyapunov exponent), open circles: topological entropy measure (topological entropy). Solid lines: Entropy functions of the parabola at $a = 4$, for increasing n, where asymptotically, the Lyapunov and the topological characterization coincide (star).

Entropy functions $S(v)$ are at the heart of the construction of our measure of complexity. To provide simple leading examples, we will essentially restrict ourselves to 1-d systems and focus on the finite-time Lyapunov exponents of n steps $\frac{1}{n}\log|(f^{(n)})'(x)|$. Consider, for the map shown in Fig. 1a, a partition element $l_j(n)$. Then, for all $x \in l_j(n)$, we have

$$\frac{1}{n}\log|(f^{(n)})'(x)| = -\frac{1}{n}\log l_j(n) =: \varepsilon_j(n). \tag{4}$$

This definition justifies calling ε the scaling exponent of the support of the associated probability measure.[9–11] Letting $n \to \infty$, generally creates a continuum of values ε. Corresponding relationships can be established for more general nonlinear maps. The entropy function $S(\varepsilon)$ of scaling exponents can easily be calculated by writing the partition function in terms of the lengths $l_i(n)$ of the partition of depth n, obtaining $Z(n, \beta, \varepsilon) = \sum_i l^{\beta}{}_i(n)$, where the summation index i runs over the members of the partition.[9–13, 15–16]

As examples, we show the calculated entropy functions for two asymmetric tent maps of varying skewness (see Fig. 1b, dashed lines). Two particular function points deserve explicit mentioning. The point with the property $S(\varepsilon) = \varepsilon$ (filled circles) characterizes the long-time properties of the natural invariant measure (e.g., ref. 9); its abscissa value gives the Lyapunov exponent of the system. Similarly, the point of maximal entropy, denoted by $(\varepsilon_0, S(\varepsilon_0))$ characterizes the topological entropy (open circles), where ε_0 is called the topological length scale exponent of the system. The results obtained for the fully developed parabola (i.e., $a = 4$) are also displayed. Solid lines indicate finite-partition-level entropy functions of this system. They slowly converge towards a triangular graph with corner points $\{\log 2, 0\}$, $\{\log 2, \log 2\}$, $\{2 \log 2, 0\}$. In this case, the topological and the Lyapunov characterization of the system coincide (marked by a star), and strict convexity is violated.

It is important to realize that the obtained entropy functions do not only depend on the chosen system, but also on the chosen observable. If another observable is taken for the same system, the associated entropy functions will usually be different. When, e.g., the local finite-resolution fractal dimension $\alpha(x)$[11] is the observable, for symmetric as well as for asymmetric tent maps, the corresponding entropy functions are trivial (i.e, for $\alpha = 1$, $S(\alpha) =: f(\alpha) \equiv 1$, and zero otherwise). For the fully developed parabola, a nontrivial entropy function $f(\alpha) = 2\alpha - 1$ is obtained. It will be shown that the entropy function is a sufficient basis for the calculation of our complexity measure. Unfortunately, the historical notations of entropy functions associated with different observables hide their common origin (Eq. (3)) (see refs. 12, 13, and 15). The generation of fractals by means of strange repellers[9] provides explicit examples of how local dimensions can be related to dynamical processes. The task of predicting $\alpha(x)$ upon partition refinement, defines the "fractal complexity."

1. MEASURE CONSTRUCTION

Our complexity measure is defined as the variability of the predictability of the observable, averaged over all system behaviors. The choice of the particular observable ε, instead of the dummy variable v, will simplify the illustration of the general construction of our measure. Equation (2) implies that the probability for observing trajectories with a specific value of ε, as a function of n behaves as

$$P(\varepsilon, n)\, d\varepsilon \sim e^{-n(\varepsilon - S(\varepsilon))}\, d\varepsilon. \tag{5}$$

As $\varepsilon \geqslant S(\varepsilon)$, the smaller $\varepsilon - S(\varepsilon)$, the better the prediction based on the past of the orbits will be. Orbits with $\varepsilon = S(\varepsilon)$ will yield perfect long-time prediction. Indeed, this situation characterizes the long-time average of the natural invariant measure (expressed by the Lyapunov exponent of the system in the case of the observable ε). However, we seek a measure able to quantify the *difficulty* of making correct predictions, over all length scales. How this task relates to the Eq. (5) is motivated by the following thought experiment. Suppose—as in the previous discussion on throwing dice—that $l > 1$ invariant measures satisfy $\varepsilon_j = S(\varepsilon_j)$, $j = 1,\dots, l$. Then, firstly, the system's future is obviously more difficult to predict, and, secondly, the difficulty added by a particular measure j should be independent of the size ε_j. The difficulty of making good predictions is also substantially increased by invariant measures for which $\varepsilon - S(\varepsilon)$ is sufficiently close to zero. Inversely, the larger this difference, the lesser the extent to which the corresponding invariant measures will contribute to the difficulty of making good predictions. The simplest candidate for a complexity measure reflecting these properties is the ratio $S(\varepsilon)/\varepsilon$.[15] To account for all system orbits, the average

$$\int \frac{S(\varepsilon)}{\varepsilon}\, d\varepsilon \tag{6}$$

is defined as the measure of complexity. In order to facilitate the comparison of systems with different topological entropies, we may rescale ε and $S(\varepsilon)$ as $\tilde{\varepsilon} = \varepsilon/\varepsilon_0$ and $\tilde{S}(\tilde{\varepsilon}) = S(\varepsilon)/\varepsilon_0$, where ε_0 is the topological length scale exponent. Geometrically, this corresponds to a similarity transformation of the entropy function's graph at (0, 0), which maps the topological length scale exponent ε_0 to unity. In this case, our complexity measure assumes the form

$$C_s(\varepsilon) := \varepsilon_0^2 \int \frac{\tilde{S}(\tilde{\varepsilon})}{\tilde{\varepsilon}}\, d\tilde{\varepsilon} \tag{7}$$

where on the left-hand side, ε refers to the chosen observable.

Extensions. The measure can be modified to embrace repellers. To chaotic repellors the same complexity should be attributed as to the corresponding chaotic attractors. Compensating for the escape rate κ[9] which distinguishes between the two system classes, yields for our measure the form

$$C_s(\varepsilon) = \varepsilon_0^2 \frac{\varepsilon_1}{\varepsilon_1 - \kappa} \int \frac{\tilde{S}(\tilde{\varepsilon})}{\tilde{\varepsilon}}\, d\tilde{\varepsilon}. \tag{8}$$

Note that for all non-fixedpoint systems, we have automatically $\varepsilon_1 > \kappa$. To refine the distinction of dynamical systems according to their complexity, we may exponentiate the front factor and the integrand independently. Then the most general form of our measure is obtained as

$$C_s(\gamma, \beta)(\varepsilon) := \varepsilon_0^{2\beta} \frac{\varepsilon_1}{\varepsilon_1 - \kappa} \int \left(\frac{\tilde{S}(\tilde{\varepsilon})}{\tilde{\varepsilon}}\right)^{\gamma} d\tilde{\varepsilon}, \tag{9}$$

where γ and β are weightening exponents. To avoid divergence, we require $\gamma > -1$.

2. MEASURE APPLICATIONS

To retain simplicity and focus, we will mainly restrict our discussion to $C_s(1,0)$. From this quantity, the equally significant quantitiy $C_s(1,1)$ is obtained by multiplication with the square of the topological length scale exponent (i.e., in case of the scaling of the support, by ε_0^2). This quantity is generally slowly varying, as a function of the system parameters.

Exact Results. In special cases, $C_s(1,0)$ (and $C_s(1,1)$) can be evaluated analytically. If the support of the specific entropy function only consists of one point (trivial spectrum), then complexity 0 is obtained. As a prominent example, zero complexity is obtained for the dynamically generated $1/3$ Cantor set, for observables α and ε alike. As another system, the fully developed parabola has a "fractal complexity" $C_s(1,0)(\alpha) = 1-\log(2)$, whereas its "dynamical complexity" evaluates to $C_s(1,0)(\varepsilon) = 1/2$. Intermittent systems with intermittency exponent $z \in I_z = (\frac{3}{2}, 2)$, constitute the most complex class of 1-d dynamical systems, since in this case the contribution by length scale exponents $\varepsilon \leqslant \varepsilon_0$ is already close to 1. Their fractal complexity can be evaluated from the fact that the whole family is characterized by a singularity at the intermittent point of the order of $1/z$. This leads to the result that the fractal complexity increases monotonously on I_z, from $1-2\log(3/2)$ to $1-\log(2)$, i.e., $C_s(1,0)(\alpha)$ is considerably smaller than $C_s(1,0)(\varepsilon)$.

Numerical Results. In most cases, the complexity measure needs to be calculated numerically. In Fig. 2 we show the dynamical complexities $C_s(1,0)(\varepsilon)$ obtained for typical representatives of classical classes of 1-d dynamical maps: (a) hyperbolic systems (tent map, order parameter a: skewness), (b) nonhyperbolic systems with slopes bounded away from zero (bungalow tent map,[16] a: position of the point on the diagonal), (c) nonhyperbolic systems with slopes not bounded away from zero (parabola, a: opening), (d) intermittent maps (see ref. 14, order parameter: z). The obtained results are based on finite Markov partitions[17] (a and b), on numerical interval partitions (c, partition level $n = 15$), and on grand-canonical partitions (d, partition level $n = 2000$). The first two classes nicely demonstrate how the complexity vanishes when the maps become symmetric ($a = 1/2$ and $a = 1/6$, respectively). For the parabola, the dependence of $C_s(1,0)(\varepsilon)$ on the order parameter a is fractal. At fully developed chaos, our numerical result based on an $n = 15$ partition level yields $C_s(1,0)(\varepsilon) = \frac{1}{2}$, which can be verified by an analytical approach. At parameter values leading to attracting periodic orbits, slow convergence to zero complexity is observed. The numerical results for the intermittent map family corroborate our expectations that this family should have maximal complexity.

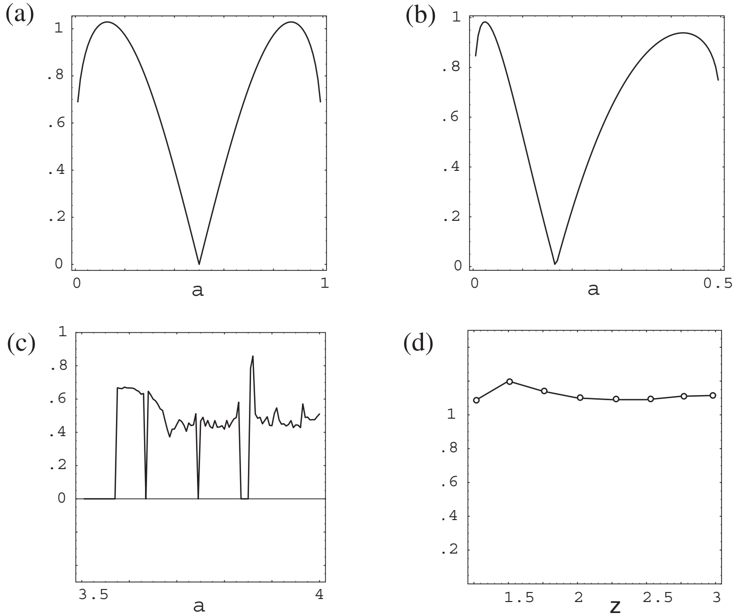

Fig. 2. Complexity $C_s(0, 1)(\varepsilon)$. (a) Tent map paradigm, parameter a: location of peak. (b) Bungalow tent map paradigm, parameter a: location of right diagonal point. (c) Logistic parabola paradigm, parameter a: opening of parabola. (d) Intermittent map family $(1+\varepsilon)\, x + x^z$ Mod 1, where ε is small. The result for $z = 1.25$ is affected by numerical inaccuracy.

We also evaluated our complexity measure for experimental neuron data. In an in vivo experiment of cat visual cortex V1, inter-spike intervals ("ISI") between firing events were recorded and analyzed, for two distinct stimulation paradigms[18] (stimulation by noisy patterns moving into the neuron's preferred direction and square-wave stimulation as the optimal stimulus). Well-separated density humps characterize the ISI distribution of the latter, which are also detectable in the much smoother ISI distribution of the noise-driven neuron. This fact allows a natural partition in the space of embedded ISI. Using a matrix representation, the thermodynamic analysis could be performed (similarly to ref. 16).

3. DISCUSSION

Thus, we are able to measure the difficulty of making predictions of system behavior in a way directly applicable to the most studied classes of dynamical system. Our measure takes into account that the difficulty of the prediction task is increased if the number of distinct persistent system behaviors is higher (in addition to the infinite-time limit behavior). The objects of the prediction are sets of neighboring trajectories, rather than individual trajectories. The rescaling $\frac{S(\varepsilon)}{\varepsilon}$ of the entropy S (see Eq. 5) is the essential ingredient for our measure's boundedness on classes for which other measures diverge. From the measure construction, it follows that the system is more complex with respect to the observable v, if the measure $C_s(1,1)(v)$ is increased. $C_s(1,1)(v)=0$ or $C_s(1,0)(v)=0$ implies a non-complex system. Infinite complexity could only be expected if $S(v)=S(v_0)$ for $v\to\infty$. However, due to the factor $1/v$, our measure $C_s(1,0)(v)=0$ will be finite even in this case.

As another example of the fundamental difference between Lyapunov exponents and our complexity measure, let us compare the fully developed parabola with the symmetric tent map. These systems can be transformed one into the other, by means of a conjugacy preserving two points of $S(\varepsilon)$ (the natural measure and the topological measure). For an observer making predictions on ε, they appear as very distinct instead. This is captured by our measure, which yields complexities $C_s(1,0)(\varepsilon)=\frac{1}{2}$ and $C_s(1,0)(\varepsilon)=0$, respectively,

Among the numerically investigated systems, the parabola has a lower overall complexity than may be intuitively expected. However, this expectation is caused by the property of the system to vary its complexity as a function of the order parameter in a non-smooth way. Our measure is not designed to account for this aspect. As nonhyperbolicity contributes to a nonhomogeneity of the phase space of systems, it could be expected, that nonhyperbolicity is a major source of complexity. This view is only partially supported by our results. Nonhyperbolicity generally gives rise to phase-transition phenomena in dynamical systems.[19–20] Written in terms of the general observable v, if over a given range $[v_{\min}, v_{\max}]$ a convex entropy function is replaced by a straight line, the complexity will decrease. Often, however, phase transitions are generated from the superposition of separated systems. This case is described by a convex-hull construction of the individual entropy functions that correspond to the different phases. For the combined system, $C_s(1,1)(v)$ necessarily increases. $C_s(1,0)(v)$ also increases, except when an initially nearly pure intermittent system is "diluted" in a specific manner by a hyperbolic phase.

In an application of our measure to experimental neuron data we obtained a more coherent characterisation of the biological data than what is provided by traditional entropy measures. In particular, optimal stimulation of the neuron led to smaller complexity of the response if compared to non-optimal stimulation. This can be interpreted as follows:

Optimal stimulation leads to an improved computation that is expressed by a reduction of the associated complexity. The details of this analysis will be reported elsewhere.

To conclude, starting from the variability of predictability, using the thermodynamic formalism of dynamical systems, we constructed an observable-dependent measure of complexity. The range of applications of our method extends to a very broad class of systems. The evaluation is particularly simple, if a generating or approximate generating partition is available. More generally, our measure can be calculated, whenever an entropy function of scaling exponents can be evaluated. These cases include experimental time series (see, e.g., ref. 9). Our measure provides three main insights. First, the numerical results for 1-d maps point out that nonhyperbolicity *per se* is not a strong ingredient of complexity. Second, the logistic parabola family shows that our perception of complexity often includes two separate aspects: Complexity as defined in our approach (at fixed order parameter a), and the difficulty to infer such a value from the complexity at neighboring values of a. Third, at the significant parameter values $\beta = 0$ and $\gamma = 1$, our measure emerges to be largest for intermittent systems, at the border between chaos and order. This finding is in agreement with insights from theoretical biology.[21–22]

ACKNOWLEDGMENTS

The first two authors strongly benefitted from discussions with W.-H. Steeb. R.S. acknowledges the support by a KTI-contract with Phonak AG Hearing Systems.

REFERENCES

1. R. J. Solomonoff, Rep. ZTB-138 (ZATOR Co., Cambridge, Massachusetts, 1960).
2. A. N. Kolmogorov, *Probl. Inf. Theory* **1**:3 (1965).
3. A. Lempel and J. Ziv, *IEEE Trans. Inf. Theory* **IT-22**:75 (1976).
4. P. Grassberger, *Internat. J. Theoret. Phys.* **25**:907 (1986).
5. R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, and H. Scheingraber, *Chaos Solitons Fractals* **4**:133 (1994).
6. J. E. Bates and H. K. Shepard, *Phys. Lett. A* **172**:416 (1993).
7. J. P. Crutchfield and C. R. Shalizi, *Phys. Rev. E* **59**:275 (1999).
8. J. S. Shiner, M. Davison, and P. T. Landsberg, *Phys. Rev. E* **59**:1459 (1999).
9. J. Peinke, J. Parisi, O. E. Rössler, and R. Stoop, *Encounter with Chaos* (Springer, Berlin, 1992).
10. M. Kohmoto, *Phys. Rev. A* **37**:1345 (1988).
11. Y. Oono and Y. Takahashi, *Progr. Theor. Phys.* **63**:1804 (1980).
12. T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman, *Phys. Rev. A* **33**:1141 (1986).

13. P. Szépfalusy, T. Tél, A. Csordas, and Z. Kovacs, *Phys. Rev. A* **36**:3525 (1987).
14. Note that $S(\varepsilon)/\varepsilon$ can be interpreted as the dimension of the fractal set of the observable assuming the value ε, which yields a with our derivation consistent interpretation.
15. R. Stoop, *Europhys. Lett.* **25**:99 (1994).
16. R. Stoop and W.-H. Steeb, *Phys. Rev. E* **55**:6589 (1997).
17. Ya. G. Sinai, *Anal. and Appl.* **2**:70–80 (1968).
18. R. Stoop, D. Blank, A. Kern, J.-J. v.d. Vyver, M. Christen, S. Lecchini, and C. Wagner, *Cog. Brain Res.* **13**:293 (2002).
19. R. Stoop and J. Parisi, *Phys. Rev. A* **43**:1802 (1991).
20. R. Stoop, *Phys. Lett. A* **173**:369 (1993).
21. C. G. Langton, *Physica D* **22**:120 (1986).
22. S. A. Kauffman, *The Origin of Order* (Oxford University Press, Oxford, 1993).