Karhunen-Loève Decomposition of Extensive Chaos

Scott M. Zoldi*,† and Henry S. Greenside*
Department of Physics, Duke University, Durham, North Carolina 27708-0305
(Received 11 October 1996)

We show that the number of KLD (Karhunen-Loève decomposition) modes D_{KLD} needed to capture a fraction f of the total variance of an extensively chaotic state scales extensively with subsystem volume V. This allows a correlation length ξ_{KLD} to be defined that is easily calculated from spatially localized data. We show that ξ_{KLD} has a parametric dependence similar to that of the dimension correlation length and demonstrate that this length can be used to characterize high-dimensional inhomogeneous spatiotemporal chaos. [S0031-9007(97)02528-3]

PACS numbers: 05.45.+b, 05.70.Ln, 47.27.Cn

Spatiotemporal states that are nonperiodic both in space and time abound in nature and are often important technologically, e.g., for lasers, fibrillating hearts, and convective transport of heat [1]. Experiments and simulations raise questions that are presently poorly understood: Are there different kinds of spatiotemporal nonperiodic states? What kinds of bifurcations lead to such states? How do inhomogeneities and boundary conditions affect the dynamics? And how does the transport of energy and matter depend on the details of the spatiotemporal disorder? An essential first step towards answering these questions is to develop methods to quantify spatiotemporal dynamic states, so that one state can be distinguished from another and so that theory can be compared with experiment and with simulation. Data analysis and theoretical progress are presently limited by a scarcity of concepts and of computational methods for analyzing spatiotemporal disorder.

In the absence of a fundamental theory of sustained nonequilibrium systems that could indicate appropriate quantities to measure, researchers have used primarily two approaches for quantifying spatiotemporal disorder: two-point correlation functions and dynamical invariants such as Lyapunov exponents and fractal dimensions [1]. Temporal correlation functions have been effective for distinguishing periodic and quasiperiodic dynamical states from each other and from chaotic ones [2] while spatial correlation functions have played an important role in discovering and demonstrating the absence of long range spatial order [3]. Both correlation functions have been less useful for distinguishing one chaotic state from another or for comparing experimental with computational chaotic data. Dynamical invariants have been somewhat useful for ordering the chaotic states of low-dimensional dynamical systems but severe difficulties remain in calculating these quantities for higher-dimensional systems because of the demanding computational effort, the slow and often ambiguous convergence of time-series-based algorithms, and the need for large amounts of noise-free data [4].

In this Letter, we propose and analyze a new measure of spatiotemporal disorder, a correlation length ξ_{KLD} defined below, that seems promising especially for the analysis of large, high-dimensional, nontransient, driven-dissipative systems. This quantity has some of the flavor of correlation functions and also of dynamical invariants and is straightforward to compute with moderate amounts of spatiotemporal data (unlike global dynamical invariants like the fractal dimension) since, as we show below, it is a *local* quantity that can be calculated from spatiotemporal data associated with a finite region of space. This last feature suggests that ξ_{KLD} will be useful for studying spatially inhomogeneous dynamics arising from slowly changing external parameters, from broken symmetries, or from the influence of boundaries. We investigate the properties of ξ_{KLD} numerically for two idealized mathematical models—the one-dimensional Kuramoto-Sivashinsky (KS) equation [1] and the two-dimensional Miller-Huse model [5]—whose spatiotemporal chaotic solutions have been thoroughly studied and for which inhomogeneities can be introduced in a controlled manner. In later work, applications to experimental data will be reported.

Our motivation for defining and studying the correlation length ξ_{KLD} comes from three different ideas. First is the idea of extensive chaos, that a sufficiently big *homogeneous* spatiotemporal chaotic system has the property that its fractal dimension D is extensive, growing in proportion to the system volume V [1,6,7]. This extensive scaling suggests that bounded *intensive* quantities, such as a dimension density $\delta = \lim_{V\to\infty} D/V$ or the equivalent dimension correlation length $\xi_\delta = \delta^{-1/d}$ [1,8] (where d is the number of asymptotically large spatial dimensions, e.g., $d = 2$ for a large-aspect-ratio convection experiment) are more appropriate for characterizing large nonequilibrium systems. It was shown recently that the length ξ_δ varies independently of the two-point and mutual-information correlation lengths [8]. This suggests that certain measures of spatiotemporal dynamics should be sensitive to structure in phase space, not just to instantaneous or time-averaged measures of spatial disorder in configuration space. The quantity ξ_{KLD} turns out to have this property.

The second idea is to extend the concept of local thermodynamic equilibrium [9, p. 13] to slowly varying inhomogeneous driven dissipative systems, with the implication that intensive dynamical quantities can be defined locally and will be slowly varying in space. As an illustration, assume that a large sustained nonequilibrium system has a parameter $p(x)$ that varies slowly with position x and consider a subsystem of size L centered at position x_0. Then over a certain time scale that decreases with decreasing subsystem size L (not necessarily diffusively), this subsystem will be approximately nontransient even if the system containing the subsystem is not. Further, the values of intensive parameters associated with this approximately nontransient subsystem should correspond closely in value to those of an infinite homogeneous nontransient system with the parameter value $p(x_0)$. Unfortunately, there are no reliable algorithms that can estimate intensive quantities such as the dimension density δ from information localized to some region of space; calculations of intensive quantities such as the Lyapunov dimension density δ have so far relied on the expensive calculation of global extensive quantities followed by taking the limit of some intensive ratio [8]. This is impractical for the analysis of experimental data or for the evaluation of local measures.

The third idea is to make use of the Karhunen-Loève decomposition (KLD), which has been used by researchers in many disciplines to analyze spatiotemporal data [10], although not in the context of extensive chaos or of inhomogeneous systems. The KLD is a statistical method for compressing spatiotemporal data by finding the largest linear subspace that contains substantial statistical variations of the data. To illustrate the idea in the discrete case and also to introduce some notation, we consider a one-dimensional zero-mean field $u(t,x)$ on a spatial interval $[0,L]$ whose values are measured on a finite space-time mesh of T uniformly sampled time points $t_i = i\Delta t$ and of S uniformly sampled spatial points $x_j = j\Delta x$. Then a $T \times S$ rectangular data matrix $A_{ij} = u(t_i, x_j)$ can be defined from which a $S \times S$ symmetric positive semidefinite scatter matrix $\mathbf{M} = \mathbf{A}^T\mathbf{A}$ can be calculated, where $\mathbf{A}^T$ denotes the matrix transpose of $\mathbf{A}$. The scatter matrix can be diagonalized to obtain its nonnegative eigenvalues σ_i^2 which can be further ordered in decreasing size $\sigma_1^2 \geq \sigma_2^2 \cdots \geq \sigma_S^2 \geq 0$.

Since the ordered eigenvalues σ_i^2 often decrease rapidly in magnitude with increasing index i, researchers [11] have introduced a positive integer D_{KLD}:

$$D_{\mathrm{KLD}} = \max\left\{ p : \sum_{i=1}^{p} \sigma_i^2 \Big/ \sum_{i=1}^{S} \sigma_i^2 \leq f \right\}, \qquad (1)$$

which represents the largest number of KLD modes p needed to capture some specified fraction $f \leq 1$ of the total variance $\sum_{i=1}^{S} \sigma_i^2$ of the data. Researchers have suggested using D_{KLD} like a fractal dimension D to measure the complexity of spatiotemporal data [11] although care is needed when interpreting D_{KLD}. The T S-dimensional vectors defined by the rows of the data matrix $\mathbf{A}$ constitute an embedding of the dynamics into a S-dimensional phase space. The quantity D_{KLD} indicates the dimension of a linear subspace that includes most of the statistical variation of this embedding, and is generally quite different from the attractor's fractal dimension D, e.g., a limit cycle with fractal dimension $D = 1$ in a N-dimensional phase space could have a value of D_{KLD} between 1 and N depending on how the limit cycle is folded in different orthogonal directions.

Although D_{KLD} need have no particular relation to the data's fractal dimension D, we argue that for *extensive* chaos, the rate of increase of D_{KLD} with volume V will generally be similar to the rate of increase of fractal dimension D with V. This follows from extensivity, that the addition of an independent subsystem (as we increase the system volume) adds a new linear subspace (increasing D_{KLD}) that contains the new degrees of freedom that lead to an increase in D. This suggests that D_{KLD} is related to D for extensive systems and so we can then try to use the more readily calculated quantity D_{KLD} to estimate intensive correlation lengths like the length ξ_δ discussed above, and to study the dependence of such lengths on system parameters.

We now present some results that illustrate these observations. When evaluated for spatiotemporal data of a large, approximately homogeneous, sustained nonequilibrium system of volume V, the KLD dimension D_{KLD} of Eq. (1) grows extensively with V as shown in Fig. 1 where V is either the volume of the entire system or the volume of a sufficiently large subsystem in a fixed volume. Here we have used $T \times S$ data matrices $\mathbf{A}$ (with $10^4 \leq T \leq 2 \times 10^4$ and $100 \leq S \leq 800$) derived from the spatiotemporal field $u(t,x)$ of the one-dimensional

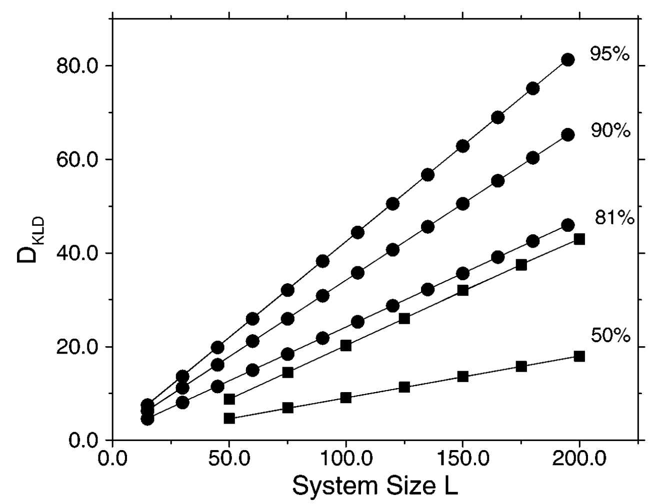

FIG. 1. KLD dimension D_{KLD} [Eq. (1)] versus size L for subsystems (circles) and for full systems (squares) of Eq. (2). Lines connecting points are to guide the eye. The labels indicate the value of f. The $f = 0.81$ system size line corresponds to the extensive scaling of the Lyapunov dimension D [7]. Spatial and temporal sampling intervals of $\Delta x = 0.25$ and $\Delta t = 2$ were used, respectively.

Kuramoto-Sivashinsky equation

$$\partial_t u + \partial_x^2 u + \partial_x^4 u + u\partial_x u = 0, \qquad x \in [0, L], \quad (2)$$

with rigid boundary conditions $u = \partial_x u = 0$ at $x = 0$ and at $x = L$. Equation (2) was integrated numerically with a semi-implicit finite-difference method that was first- and second-order accurate in time and space, respectively. For $L > 50$, most initial conditions yield spatiotemporal chaotic states that were previously shown to be extensively chaotic [7]. Figure 1 shows that D_{KLD} is extensive for system sizes $L \geq 50$ (plotted in squares), growing in proportion to the system volume L with a slope that depends on the fraction f. Figure 1 also shows that D_{KLD} is extensive for sufficiently large open subsystems (plotted in circles) centered on the middle of a system of fixed size 400 over the range $15 \leq L \leq 200$. The slope of D_{KLD} for subsystems for a given fraction f is the *same* as the slope of D_{KLD} for full systems (although the intercepts differ). This implies the important point that the intensive density $\lim_{V\to\infty} D_{\mathrm{KLD}}/V$ can be estimated from information localized to a certain region of space.

The extensivity of the KLD dimension, for both the entire system and for subsystems, suggests introducing an intensive KLD correlation length ξ_{KLD}:

$$\xi_{\mathrm{KLD}} = \left(\lim_{V\to\infty} D_{\mathrm{KLD}}/V\right)^{-1/d}, \quad (3)$$

by analogy to the definition of the dimension correlation-length ξ_δ (where again d is the spatial dimensionality of the system). Based on the data in Fig. 1 for the KS equation, Fig. 2 shows how the length ξ_{KLD} varies with the fraction f. The dependence is nonlinear, with the magnitude of ξ_{KLD} changing by a factor of 10 over the range $0.3 \leq f \leq 1$. Contrary to an earlier claim by Ciliberto and Nicolaenko [11], Fig. 2 suggests that the fractal dimension of a high-dimensional system cannot generally be estimated from a knowledge of ξ_{KLD} since the fraction f corresponding to the dimension correlation length will not be known in advance and because ξ_{KLD} can vary substantially with f. However, the onset of extensivity for D_{KLD} does accurately predict the onset of extensivity for the Lyapunov dimension D with increasing volume V which provides a way to determine if an experimental system is extensively chaotic.

Figure 3 shows how the length ξ_{KLD} (for $f = 0.95$) compares with the dimension correlation length ξ_δ (derived from the extensive Lyapunov fractal dimension [8]) for a nonequilibrium Ising-like phase transition of a mathematical model invented by Miller and Huse [5]. (The functional dependence of ξ_{KLD} on parameter g depends only weakly on the fraction f which therefore is not an important parameter here.) The model is a 2D coupled-map lattice for which a 1D chaotic map of odd symmetry is placed at each node of a periodic $L \times L$ square lattice. Each site is coupled diffusively to nearest neighbors with a strength g that acts as the bifurcation parameter for this system. For increasing values of g, Miller and Huse found that a quantity analogous to a lattice magnetization M bifurcated from a zero to nonzero value at a critical value $g_c = 0.205$ at which point also a two-point correlation length ξ_2 diverged to infinity. O'Hern

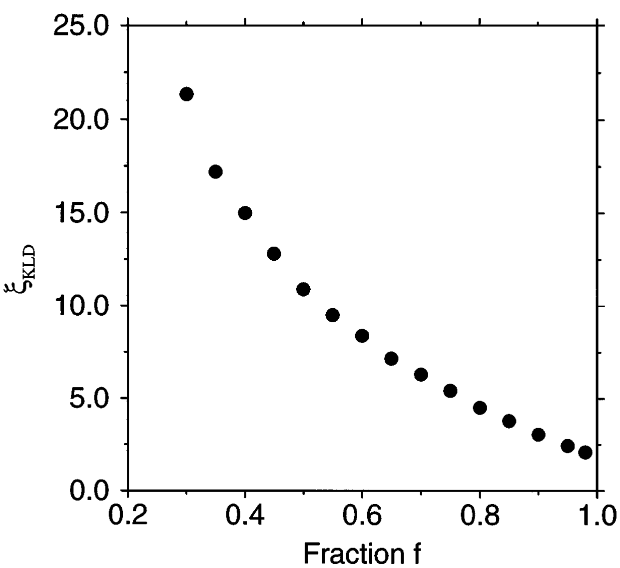

FIG. 2. KLD correlation length ξ_{KLD} versus fraction f for the data of Fig. 1, showing a nonlinear dependence over an order of magnitude in ξ_{KLD}.

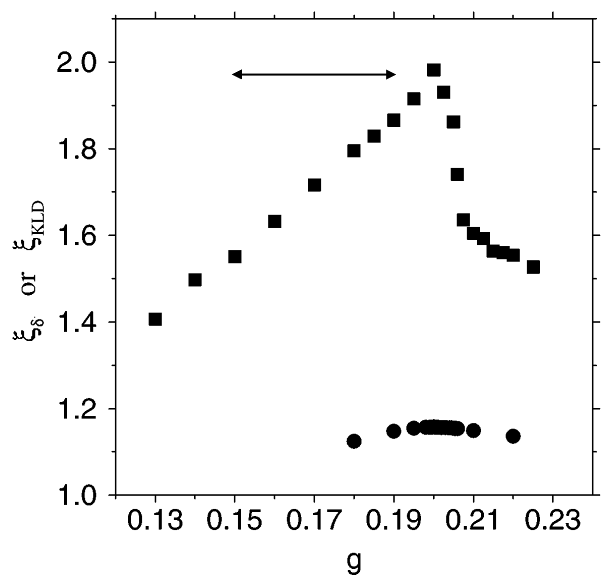

FIG. 3. Comparison of the KLD and Lyapunov dimension correlation lengths [8] (squares and circles, respectively) for a nonequilibrium transition [5]. The values for ξ_{KLD} were calculated for the fraction $f = 0.95$ using spatiotemporal data of the Miller-Huse model on a 2D periodic square lattice of size $L = 30$. $T = 10^4$ time samples of $S = L^2$ latice sites over the range $12^2 \leq S \leq 20^2$ were used to define the data matrix **A**.

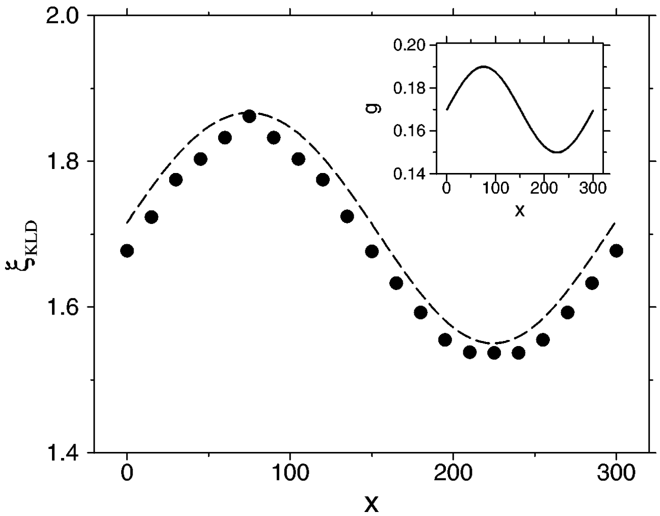

FIG. 4. KLD correlation length ξ_{KLD} with $f = 0.95$ (circles) versus spatial coordinate x for a weakly inhomogeneous 2D Miller-Huse model on a square lattice with periodic boundary conditions, with rectangular geometry $L_x = 300$ and $L_y = 30$. The scatter matrices were calculated from $T = 10^4$ time samples and $S = 9^2$, 11^2, 13^2, and 15^2, respectively. The inset indicates the periodic spatial variation of $g(x) = 0.17 + 0.02\sin(2\pi x/300)$, whose range corresponds to the interval indicated by the horizontal arrow in Fig. 3. The dashed curve represents the value of ξ_{KLD} for infinite homogeneous lattices with constant coupling constant $g(x)$.

et al. [8] showed that the dimension correlation length ξ_δ did not diverge near g_c but instead was a quantity of order one that smoothly reached a local maximum at a value $g = 0.200$ distinctly less than the critical value g_c. Figure 3 shows that the length ξ_{KLD} is somewhat larger than, but comparable in magnitude with, ξ_δ and has a qualitatively similar dependence on g in that it increases to a local maximum at the same g value. Unlike the dimension correlation length, ξ_{KLD} does not seem to vary smoothly near its maximum but we lack sufficient numerical resolution to determine unambiguously whether there is a finite jump in value, in analogy to the dependence of specific heat on temperature for a second-order equilibrium phase transition.

Finally, in Fig. 4 we demonstrate how the KLD correlation length ξ_{KLD} can be used to characterize inhomogeneous spatiotemporal chaos, a result that opens up interesting possibilities for the future analysis of experimental data. For Fig. 4, we introduced a spatial inhomogeneity into a 300×30 periodic Miller-Huse lattice by allowing the coupling constant $g = g(x)$ to vary periodically in the x lattice direction as shown in the inset. At each of several x coordinates, we calculated the KLD dimension Eq. (1) for subsystems centered on x and of increasing width L with $9 \le L \le 15$. From these data, local extensive scaling was identified from which a length ξ_{KLD} was calculated from Eq. (3). In Fig. 4, the lengths ξ_{KLD} are given by the circles which can be compared with the dashed curve representing the corresponding value of ξ_{KLD} that would be obtained from Fig. 3 for an infinite homogeneous system with constant value $g = g(x)$. The agreement is good to about 4% throughout, which is sufficient to determine the quantitative spatial dependence of the inhomogeneity.

In conclusion, the correlation length Eq. (3) obtained by studying the extensive scaling of the Karhunen-Loève decomposition with increasing subsystem volume provides an easily calculated and novel way to characterize the spatiotemporal disorder of an extensively chaotic system, including the case of slowly varying spatial inhomogeneities. We believe that the ideas on which our analysis is based—namely, extensivity, local stationarity, and the Karhunen-Loève decomposition—will be important ingredients in the future analysis of large nonequilibrium systems.

This work was supported by a DOE Computational Science Fellowship, by NSF Grants No. NSF-DMS-93-07893 and No. NSF-CDA-92123483-04, and by DOE Grant No. DOE-DE-FG05-94ER25214.

*Also at Center for Nonlinear and Complex Systems, Duke University, Durham, NC.
†Electronic address: zoldi@phy.duke.edu

[1] M. C. Cross and P. C. Hohenberg, Rev. Mod. Phys. **65**, 851 (1993).
[2] G. Ahlers and R. P. Behringer, Prog. Theor. Phys. Suppl. **64**, 186 (1978).
[3] J. P. Gollub and R. Ramshankar, in *New Perspectives in Turbulence*, edited by L. Sirovich (Springer-Verlag, Berlin, 1991), pp. 165–194; M. C. Cross, D. Meiron, and Y. Tu, Chaos **4**, 607 (1994).
[4] H. D. Abarbanel, R. Brown, and L. S. Tsimring, Rev. Mod. Phys. **65**, 1331 (1993).
[5] J. Miller and D. A. Huse, Phys. Rev. E **48**, 2528 (1993).
[6] D. Ruelle, Commun. Math. Phys. **87**, 287 (1982).
[7] P. Manneville, in *Macroscopic Modeling of Turbulent Flows*, edited by O. Pironneau, *Lecture Notes in Physics* (Springer-Verlag, New York, 1985), pp. 319–326.
[8] C. O'Hern, D. Egolf, and H. Greenside, Phys. Rev. E **53**, 3374 (1996).
[9] L. Landau and E. Lifshitz, *Statistical Physics* (Pergamon, New York, 1980), 3rd ed., Vol. **5**, Part 1.
[10] R. W. Preisendorfer, *Principal Component Analysis in Meteorology and Oceanography, Developments in Atmospheric Science* Vol. **17** (Elsevier, Amsterdam, 1988); K. Fukunaga, *Introduction to Statistical Pattern Recognition* (Academic Press, Boston, 1990), 2nd ed.; G. Berkooz, P. Holmes, and J. L. Lumley, Annu. Rev. Fluid Mech. **25**, 539 (1993).
[11] L. Sirovich and A. E. Deane, J. Fluid Mech. **222**, 251 (1991); S. Ciliberto and B. Nicolaenko, Europhys. Lett. **14**, 303 (1991); R. Vautard and M. Ghil, Physica (Amsterdam) **35D**, 395 (1989).