Persistent chaos in high dimensions

D. J. Albers,[1,2,3,*] J. C. Sprott,[4,†] and J. P. Crutchfield[2,3,‡]

[1]*Max Plank Institute for Mathematics in the Sciences, Leipzig 04103, Germany*

[2]*Center for Computational Science & Engineering and Physics Department, University of California-Davis, One Shields Avenue, Davis, California 95616, USA*

[3]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

[4]*Physics Department, University of Wisconsin, 1150 University Avenue, Madison, Wisconsin 53706, USA*

(Received 19 April 2005; published 2 November 2006)

An extensive statistical survey of universal approximators shows that as the dimension of a typical dissipative dynamical system is increased, the number of positive Lyapunov exponents increases monotonically and the number of parameter windows with periodic behavior decreases. A subset of parameter space remains where noncatastrophic topological change induced by a small parameter variation becomes inevitable. A geometric mechanism depending on dimension and an associated conjecture depict why topological change is expected but not catastrophic, thus providing an explanation of how and why deterministic chaos persists in high dimensions.

DOI: 10.1103/PhysRevE.74.057201 PACS number(s): 05.45.Jn, 05.45.Tp, 89.70.+c, 89.75.−k

Physical theory attempts to describe and predict the natural world by expressing observed behavior and the governing balance of forces formally in mathematical models—models that can only be approximate representations. Empirically, many natural phenomena persist even when control parameters and external conditions vary. For example, the essential character of fully developed fluid turbulence is little affected if one slightly changes the energy flux that drives it or if a small dent is made in the containing vessel's wall. In building a theory of a system exhibiting this kind of dynamical persistence, one hopes that, despite its approximations, one's model also has this persistence.

A century of analyzing nonlinear dynamical systems, however, has led to an apparent inconsistency with this goal. Since the days of Poincaré's development of qualitative dynamics, mathematicians and physicists have probed differential equations to test their solutions for different kinds of stability. Poincaré's discovery of deterministic chaos [1] demonstrated that at the most detailed level, there was inherent instability of system solutions: change the initial condition only slightly and one finds a different state-space trajectory develops rapidly. Later studies showed that there was also an instability in behavior if the equations or parameters were changed only slightly [2–4]. Even arbitrarily small functional perturbations to the governing dynamic leads to radical changes in behavior—from unpredictable to predictable behavior, for example. The conclusion has been that nonlinear, chaotic systems are exquisitely sensitive, amplifying arbitrarily small variations in initial and boundary conditions and parameters to macroscopic scales.

How can one reconcile this with the observed fact of dynamical persistence in large-scale systems [5]? We take *dynamical persistence* to mean that a behavior type—e.g., equilibrium, oscillation, chaos—does *not* change with functional perturbation or parameter variation. Here we describe the results of a statistical survey which empirically demonstrate that chaos is dynamically persistent if the dimension of a nonlinear system is sufficiently high. More importantly, we argue that a particular geometric *mechanism* is responsible for persistent chaos.

Specifically, the survey shows that in large-scale systems dynamical sensitivity—when defined as breaking topological equivalences associated with structural stability [6], ergodicity [7], and statistical stability [8]—is typically benign and does not affect behavior types. Naturally, drastic changes in a system's invariant measure yield different observed dynamics, but our results indicate that this becomes increasingly less probable. Moreover, the instability associated with deterministic chaos dominates high-dimensional dynamical systems, except at extreme parameter settings.

Much of the intuition and motivation for our investigation comes from the analytical results found in abstract dynamical systems theory, but our construction and conclusions highlight a distinct difference. Said most simply, the number of dimensions of the dynamical system matters. That is, there is a qualitative difference between common behaviors in high- and low-dimensional dynamical systems, subject to how the given space of dynamical systems is stratified, e.g., topologically [9], algebraic geometrically [10], or in our case, measure-theoretically. Beyond giving empirical evidence (relative to a measure on a function space) to support these conclusions, we introduce a definition of persistent chaos that suggests an alternative approach to longstanding questions of dynamical stability and offer a conjecture detailing a geometric mechanism underlying persistent chaos in high dimensions.

Assuming the existence of a Sinai-Ruelle-Bowen (SRB) measure [11], the spectrum of Lyapunov characteristic exponents (LCEs) [12] for a d-dimensional system consists of d LCEs: $\chi_1 \geq \chi_2 \geq \cdots \geq \chi_d$, where indexing gives a monotonic ordering. The LCEs will be our primary tool for analyzing and identifying behavior types since there is an equivalence between the number of negative and positive Lyapunov exponents and the number of global stable and unstable mani-

*Electronic address: albers@cse.ucdavis.edu

†Electronic address: sprott@physics.wisc.edu

‡Electronic address: chaos@cse.ucdavis.edu

folds, respectively—structures that organize the state space and constrain behavior [13]. Therefore in referring to topological variation here we mean a change in the number of positive LCEs.

In order to give a complete representation of the space of all systems, we investigate typical behaviors in high dimensions using a class of dynamical systems that are known to be *universal function approximators*. They are universal in two equivalent senses: (i) in the limit that they have infinitely many parameters they are dense [14] in C^r on compacta; and (ii) they can approximate arbitrarily closely any C^r mapping and its derivatives on compacta [14]. These are single-layer recurrent neural networks of the form

$$x_t = \beta_0 + \sum_{i=1}^{n} \beta_i \tanh s\left(\omega_{i0} + \sum_{j=1}^{d} \omega_{ij} x_{t-j}\right), \qquad (1)$$

which are maps from R^d to R and denoted $f_{s,\beta,\omega}$. Here n is the number of hidden units (neurons), d the number of time lags which determines the system's input (embedding) dimension, and s a scaling factor for the connection weights ω_{ij}. The initial condition is $(x_1, x_2, \ldots, x_d)$ and the state at time t is $(x_t, x_{t+1}, \ldots, x_{t+d-1})$. The approximation theorems of Ref. [14] and time-series embedding of Ref. [15] establish an equivalence between these neural networks and general dynamical systems [16].

In the statistical survey we sample the $k=[n(d+2)+1]$-dimensional parameter space taking (i) $\beta_i \in [0,1]$ uniformly distributed and rescaled to satisfy $\sum_{i=1}^{n}\beta_i^2 = n$, (ii) ω_{ij} normally distributed with zero mean and unit variance, and (iii) the initial $x_j \in [-1,1]$ as uniform. These distributions—denoted m_β, m_ω, and m_I—form a product measure on the space of parameters and initial conditions. The survey's results are statistical estimates with respect to this product measure.

To perform the experiment, each network has fixed weights, and thus forms a standard discrete-time, time-delay dynamical system. We then use the s parameter as the primary control as it gives the magnitude of the argument of $\tanh(x)$. When $x \approx 0$ Eq. (1) is linear and one finds fixed points and limit cycles; when $|x| \gg 1$ the output is binary and one finds 2^n different periodic states; and for $|x|$ between these extremes we find the nonlinear behavior we will focus on.

The study of a general function space should be contrasted with investigations of restricted mappings—e.g., map lattices [17,18]—for which the number of parameters is a bounded function of the number of state variables, since these cannot systematically approximate C^r function space.

We now define *persistent chaos* (*p-chaos*) of degree p for a dynamical system as follows:

Definition. Assume a map $f_\xi : X \to X (X \subset R^d)$ that depends on a parameter $\xi \in R^k$. The map f_ξ has chaos of degree-p on an open set $O \subset X$ that is persistent for $\xi \in \mathcal{U} \subset R^k$ if there is a neighborhood $\mathcal{N}$ of $\mathcal{U}$ such that for all $\xi \in \mathcal{N}$, the map f_ξ retains at least $p \geq 1$ positive LCEs Lebesgue almost every in O.

The choice of defining p is flexible. For example, fixing p to be the number of positive LCEs is a very strict constraint; specifying a minimum p or ratio of p to the maximum number of positive exponents are weaker. Flexibility allows one to analyze (say) systems with 10^6 unstable directions in which a change in 1% of the geometry is undetectable, but a 50% change is. This notion differs from that of a robust chaotic attractor [19,20] in several ways. Importantly, we do not require the attractor to be unique on U since, physically, there is little evidence indicating that such strict forms of uniqueness are present in many complex physical systems (Ref. [21] presents a low-d example) and, technically, uniqueness is markedly more difficult to establish.

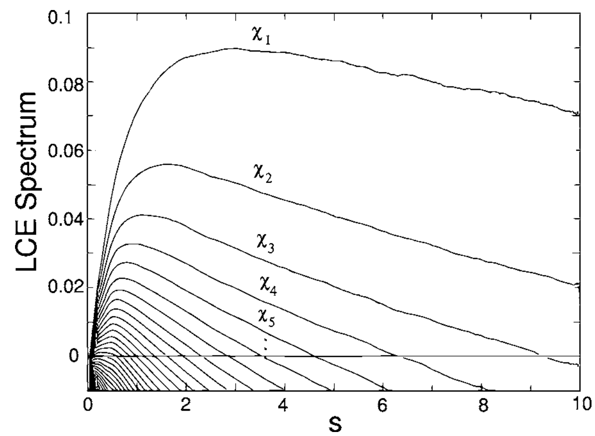

FIG. 1. LCE spectrum as a function of gain s for a network of $n=32$ neurons and $d=64$ dimensions. (15 000 total time steps; 5000 initial time steps to arrive on the attractor; LCE renormalization is performed at each time step.)

Figure 1 presents the typical scenario for the LCE spectra of the high-dimensional systems as a function of s. Here *typical* refers to what was observed in more than 99% of the 15 800 networks with $n \geq 32$ and $d > 32$. Important properties to notice include (a) lack of periodic windows with respect to (s, β, ω), (b) LCEs vary continuously with s, (c) they have a single maximum (up to statistical fluctuations), and (d) $f_{s,\beta,\omega}$ has SRB measures that yield a distribution of LCEs whose variance obeys $\sigma^2_{\chi_i} < \inf_{j=\pm 1}(|\chi_i - \chi_j|)$ at fixed s. Other previously documented properties [16,22] include (e) as d increases, the length of the s-intervals U_i between LCE zero crossings are asymptotically dense, $|U_i| \sim d^{-1.92}$, and (f) the maximum number of positive LCEs increases monotonically as $d/4$ and the attractor's Kaplan-Yorke dimension scales as $d/2$. These properties contrast sharply with familiar low-dimensional bifurcation scenarios where one typically encounters a preponderance of stable behavior and periodic windows and the LCEs vary in a discontinuous manner with control parameters.

The observations complement those from a previous study of chaos in neural-network continuous-time differential equations [23]. There, a mean-field analysis, which assumed that inputs are statistically independent (and which does not apply in the present case), also suggested that chaos should be common in high dimensions; cf. Refs. [24,25].

In light of Fig. 1 and the fact that LCE zero crossings become dense, we propose the following *geometric mechanism* for persistent chaos. For a finite but arbitrarily large number of segments along an s interval—e.g.,

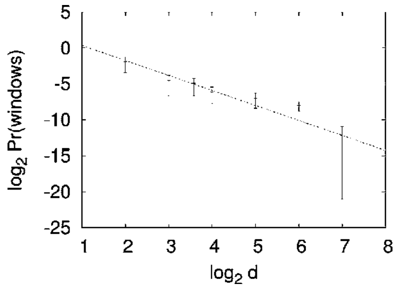

FIG. 2. Log probability of periodic behavior vs log dimension for 700 cases per d. Each case has all the weights perturbed on the order of 10^{-3}; 100 times per case. The best-fit line is $\sim 1/d^2$.

$s \in (0.1, 8)$—there is an asymptotically dense, always countable sequence of parameter values that have an LCE transversally crossing through zero. Thus a continuous path along an s-interval yields inevitable, but noncatastrophic (i.e., $p > 1$) topological change. This implies that when varying parameters, periodic and quasiperiodic windows will not exist in chaotic regions of parameter space of dynamical systems with a sufficiently large number of positive exponents. The lack of dense periodic and quasiperiodic windows is a necessary condition for p-chaos.

To test for this mechanism, we analyzed the existence of periodic and quasiperiodic windows along $s \in (1, 4)$ in networks with $n=32$ and d ranging from 8 to 128 and with an ensemble of 700 networks per n and d. We observed that (i) the mean fraction of networks with periodic and quasiperiodic windows decreases like $\sim d^{-1.3}$, (ii) the mean number of windows decreases like $\sim d^{-2}$, and (iii) the window lengths increase linearly with increasing d. These observations are insensitive to increments in s as long as $\Delta s \leq 0.005$. As the dimension increased above 64 the only networks with periodic windows had windows that persisted for most of the s-interval under consideration. That is, as dimension was increased, periodic windows became increasingly rare. When they were observed, however, they were neither small nor intermittent, but instead dominated the parameter space.

To explore the full parameter space systematically, one can fix s and vary the weights with random perturbations of a given size. We surveyed networks with parameters varied in a k-ball with its center fixed at s. Figure 2 shows how the probability of observing periodic windows decreases as the dimension increases. Each data point corresponds to the probability of finding a system with a periodic orbit among a set of 700 networks at a given n and d and each perturbed 100 times. The range of weight perturbations was 10^{-3} with $s=3$. We found that the probability of periodic networks decreases as d^{-2}. Thus as dimension increases the systems are far less likely to display periodic windows and, as a consequence, become more persistently chaotic. Our findings are robust for $s \in (0.1, 8)$ and to perturbation sizes ranging from 10^{-10} to 0.1.

While this is strong evidence for the disappearance of periodic windows in parameter space with increasing dimension, a stronger argument follows from our observation that the fraction of networks with windows decreases less quickly ($\sim d^{-1.3}$) than the overall probability of windows ($\sim d^{-2}$).

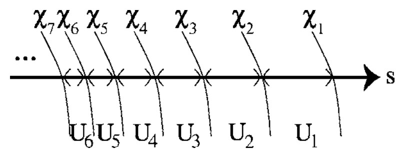

FIG. 3. Lyapunov spectrum vs network nonlinearity s: U_i's are the open sets in parameter space where structural stability is believed to persist. The $|U_i|$ parameter intervals shrink like $\sim d^{-1.92}$ as the dimension increases.

Thus periodic windows which do exist are concentrated in an ever-decreasing fraction of networks and those with one periodic window are more likely to have many periodic windows.

That the exponents are continuous with s variation can be seen in Fig. 1; quantitative analysis is performed in Ref. [16]. The claim that $f_{s,\beta,\omega}$ has SRB measures obeying property (d) can be seen by noting that for every s value in Fig. 1 a different initial condition with respect to m_I was used for computation. For s values below the onset of chaos ($s < 0.1$), this condition does not apply.

These observations, Fig. 1, and detailed analysis of 400 four-dimensional dynamical systems and 200 64-dimensional dynamical systems, as well as many of intermediate dimension, leads to the following view of the geometric persistence mechanism. All of the LCEs that become positive are negative for very small and very large values of s—the LCE spectra are unimodal. As the dimension d is increased, the LCE s-dependence becomes smoother. Moreover, with increasing dimension the number of positive exponents increases monotonically [22]. Finally, the distance between LCE zero crossings, above the maximum, decreases with dimension as shown schematically in Fig. 3. In sufficiently high dimensions, the subsets U_i shrink and eventually fall below resolution. The result, then, is twofold: one observes continuous topological change (bifurcations), but this is never catastrophic. The claim is that the geometry depicted in Fig. 1 is persistent to parameter variation.

Conjecture. Given $f_{s,\beta,\omega}$, if k and d are large enough, the probability with respect to $m_\beta \times m_\omega$ of the set (β, ω) with the properties (a)–(f) is large and approaches 1 as $k, d \to \infty$.

Since our networks are universal function approximators, this behavior should be observed in many nonlinear high-dimensional dynamical systems. The onset of sufficiently high dimension for this to occur for our dynamical systems was observed to be $d \sim 30$.

The conjecture can be quantified for particular s-intervals with the notion of p-chaos of degree p for an ensemble of mappings with our construction. For example, for an s-interval centered at $s=3$ (far from the s-interval containing the maximum number of positive LCEs), if one considers p to be the mean number of positive LCEs minus three standard deviations, then $p > 0$ at $d=32$ and increases like $p \sim 4 \log d$ for $d \geq 32$ [22]. We will refrain from arguing for a best definition of p and simply note that p_{min} increases with d monotonically on $s \in [0.1, 10]$ and thus there exists an open set for which $f_{s,\beta,\omega}$ has p-chaos of degree $p > 0$.

Two comments are in order. First, the existence of chaos as a persistent behavior type depends on dimension. The subset of parameter space in which chaos becomes persistent increases in size (with respect to $m_\beta \times m_\omega \times m_I$) as the dimension of the dynamical system increases. This is due both to the increase in the number of positive LCEs (given a sufficient increase in n) and to a decrease in the appearance of periodic windows. Second, persistence is related to the number of (linearly independent) parameters in the dynamical system. The number n of neurons in the network effectively controls the entropy rate [26]—that is, increasing the number of neurons increases the entropy rate, number of positive exponents, and the maximum of the largest exponent. Moreover, increasing n increases the degree (p) of the persistent chaos, but the mechanism for persistent chaos remains, due to the decreasing probability of periodic windows. Conversely, networks with few parameters exhibit considerably less persistent chaos.

In this way high entropy-rate systems are more persistent with respect to functional *and* parameter perturbations. This is in accord with a wide range of experimental observations of such systems. Indeed, dynamical persistence is not a novel experience; often hydrodynamic engineers and plasma experimentalists expend much effort in attempts to eliminate persistent chaos [5]. Here we described a mechanism in which the dynamical persistence of high-dimensional systems is retained under parameter perturbation, despite the fact that stricter notions of dynamical equivalence are violated. This sets the stage for more specific investigations of the statistical topology of stable and unstable manifolds in high-dimensional systems—investigations that, one hopes, will lead to predictive scaling theories for observed macroscopic properties that are grounded in microscopic dynamics.

We thank J. R. Albers, R. A. Bayliss, W. D. Dechert, D. Feldman, J. Robbin, C. R. Shalizi, J. Supanich, and K. Burns for helpful discussions. This work was partially supported at the Santa Fe Institute under the Networks Dynamics Program funded by the Intel Corporation and under the Computation, Dynamics, and Inference Program via SFI's core grants from the National Science and MacArthur Foundations. Direct support for D.J.A. was provided by NSF Grants DMR-9820816 and PHY-9910217 and DARPA Agreement F30602-00-2-0583. The Center for Computational Science & Engineering provided much needed computing support.

[1] H. Poincaré, *Oeuvres* (Gauthir-Villars, Paris, 1929).
[2] J. D. Farmer, Phys. Rev. Lett. **55**, 351 (1985).
[3] S. Newhouse, Publ. Math., Inst. Hautes Etud. Sci. **50**, 101 (1979).
[4] J. Graxzyk and G. Światek, Ann. Math. **146**, 1 (1997).
[5] P. Manneville, *Dissipative Structures and Weak Turbulence: Perspectives in Physics* (Academic Press, Boston, 1990).
[6] R. Mañé, Publ. Math., Inst. Hautes Etud. Sci. **66**, 161 (1988).
[7] C. Pugh and M. Shub, Bull., New Ser., Am. Math. Soc. **41**, 1 (2003).
[8] J. F. Alves and M. Viana, Ergod. Theory Dyn. Syst. **22**, 1 (2002).
[9] J. Palis, Asterisque **261**, 339 (2000).
[10] E. Barreto, B. R. Hunt, C. Grebogi, and J. A. Yorke, Phys. Rev. Lett. **78**, 4561 (1997).
[11] L.-S. Young, J. Stat. Phys. **108**, 733 (2002).
[12] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica **15**, 21 (1979).
[13] D. Ruelle, Ann. Math. **115**, 243 (1982).
[14] K. Hornik, M. Stinchocombe, and H. White, Neural Networks **3**, 551 (1990).
[15] T. Sauer, J. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).
[16] D. J. Albers and J. C. Sprott, Nonlinearity **19**, 1801 (2006).
[17] P. Grassberger, Phys. Scr. **40**, 346 (1989).
[18] K. Kaneko, *Collapse of Tori and Genesis of Chaos in Dissipative Systems* (World Scientific, Singapore, 1986).
[19] M. Andrecut and M. K. Ali, Phys. Rev. E **64**, 025203(R) (2001).
[20] S. Banerjee, J. A. Yorke, and C. Grebogi, Phys. Rev. Lett. **80**, 3049 (1998).
[21] K. Kaneko, Phys. Rev. E **66**, 055201(R) (2002).
[22] D. J. Albers, Ph.D. thesis, University of Wisconsin in Madison, 2004 (unpublished).
[23] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).
[24] B. Doyon, B. Cessac, M. Quoy, and M. Samuelides, Int. J. Bifurcation Chaos Appl. Sci. Eng. **3**, 279 (1993).
[25] O. Moynot and M. Samuelides, Probab. Theory Relat. Fields **123**, 41 (2002).
[26] E. F. Manffra, Ph.D. thesis, Univesitat Gesamthochschule Wuppertal, 2002 (unpublished).