Persistent Chaos in High Dimensions

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

[1] *Max Plank Institute for Mathematics in the Sciences, Leipzig 04103, Germany*
[2] *Computational Science and Engineering Center and Physics Department, University of California, Davis, One Shields Ave, Davis CA 95616*
[3] *Physics Department, University of Wisconsin, Madison, WI 53706*
[4] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501*

(Dated: April 19, 2005)

As the dimension of a typical dissipative dynamical system is increased, the number of positive Lyapunov exponents increases monotonically and the number parameter windows with periodic behavior decreases. A subset of parameter space remains in which topological change induced by small parameter variation is very common. It turns out, however, that if the system's dimension is sufficiently high, this seemingly inevitable (and expected) topological change is never catastrophic, in the sense that the behavior type is preserved. One concludes that deterministic chaos is persistent in high dimensions.

PACS numbers: 05.45.-a, 89.75.-k, 05.45.Tp, 89.70.+c, 89.20.Ff
Keywords: Chaos, high dimensions, dynamical systems, structural stability, Lyapunov exponents

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. The character of many natural phenomena persists even when control parameters and external conditions vary. For example, the essential properties of fully developed fluid turbulence are 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 that exhibits this kind of dynamical persistence, one hopes that one's model also has this persistence.

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 substantially 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, 3]. Even arbitrarily small functional perturbations to the governing dynamic lead to radical changes in behavior—from unpredictable to predictable behavior, for example. The overall conclusion has been that nonlinear, chaotic systems are exquisitely sensitive, amplifying arbitrarily small variations in initial conditions and parameters to macroscopic scales.

How can one reconcile this with the observed fact of dynamical persistence in many large-scale systems? We will take *dynamical persistence* to mean that a behavior type —e.g., fixed point, limit cycle, or chaotic attractor— does *not* change with functional perturbation or parameter variation.

Here we show that in large-scale systems dynamical sensitivity—when defined as breaking topological equivalences associated with structural stability [4], ergodicity [5], and statistical stability [6]—is typically benign and does not affect behavior types. Moreover, the instability associated with deterministic chaos dominates high-dimensional dynamical systems. 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. Beyond giving empirical evidence to support these conclusions, we introduce a definition of persistent chaos that suggests an alternative approach to the long-standing questions of dynamical stability and offer a mathematical conjecture on persistent chaos in high dimensions.

The spectrum of Lyapunov characteristic exponents (LCEs) [7] will be our primary tool of analysis and identification for behavior types. For a d-dimensional system the spectrum consists of d LCEs: $\chi_1 \geq \chi_2 \geq \ldots \geq \chi_d$, where indexing is chosen to give a monotonic ordering. The foremost reason to use this tool is that there is an equivalence between the number of negative and positive Lyapunov exponents and the number of global stable and unstable manifolds, respectively—structures that organize the state space and constrain behavior [8]. Therefore, in referring to topological variation here we mean a change in the number of positive Lyapunov exponents.

In order to give a complete representation of the space of all systems, we will investigate typical behaviors in high dimensions using a class of dynamical systems that

*Electronic address: albers@cse.ucdavis.edu
†Electronic address: sprott@physics.wisc.edu
‡Electronic address: chaos@cse.ucdavis.edu

are known to be *universal function approximators* [9], single-layer 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 is a map from R^d to R. 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 on the connection weights w_{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. [9] and well known time-series embedding results Ref. [10] together establish the equivalence between this class neural networks and the general dynamical systems of interest here [11].

We sample the $(n(d+1)+1)$-dimensional parameter space taking (i) $\beta_{ij} \in [0,1]$ uniformly distributed and rescaling them to satisfy $\sum_{i=1}^{n} \beta_i^2 = n$, (ii) w_{ij} as normally distributed with zero mean and unit variance, and (iii) the initial $x_j \in [-1,1]$ as uniform. We will focus largely on behavior types as a function of the parameter s, which can be interpreted as the standard deviation of the w weight matrix, and the embedding dimension d.

A few simple observations are in order to set limits on the behavior types one expects at the extreme regions of the parameter space. The *squashing function* $\tanh(x)$, for $|x| \gg 1$, behaves like a binary function. The neural network states will tend toward the finite set $x_t = \beta_0 \pm \beta_1 \cdots \pm \beta_n$; that is, each x_t can have 2^n different states. The result is that in the limit where the arguments of $\tanh(\cdot)$ become infinite, the neural network has periodic behavior. For $x \approx 0$, however, $\tanh(x)$ is nearly linear. Thus, choosing s to be small forces the dynamics to be mostly linear, again yielding fixed-point and periodic behavior (and no chaos). Due to this, s provides a unique bifurcation parameter that sweeps from linear to highly nonlinear parameter regimes: from discretized binary behavior—fixed points—to chaos and back to periodic phenomena.

We define *persistent chaos* in terms of the LCE spectrum as follows:

Definition 1 (Degree-p Persistent Chaos) *Assume a discrete-time map f that takes a compact set to itself. The map has persistent chaos of degree-p if there exists an open subset U of parameter space, such that, for all $\xi \in U$ and a given open set $\mathcal{O}$ of initial conditions, $f|_\xi$ retains $p \geq 1$ positive Lyapunov exponents.*

Persistent chaos of degree p (*p-chaos*) is our notion of dynamical equivalence on an open set of parameter space. A important difference between this notion and those of a robust chaotic attractor put forth in Refs. [12] and [13], for example, is that we do not require the attractor is unique on the subset U. This is an important distinction since, on the one hand, uniqueness is significantly more difficult to demonstrate and, on the other, there is little

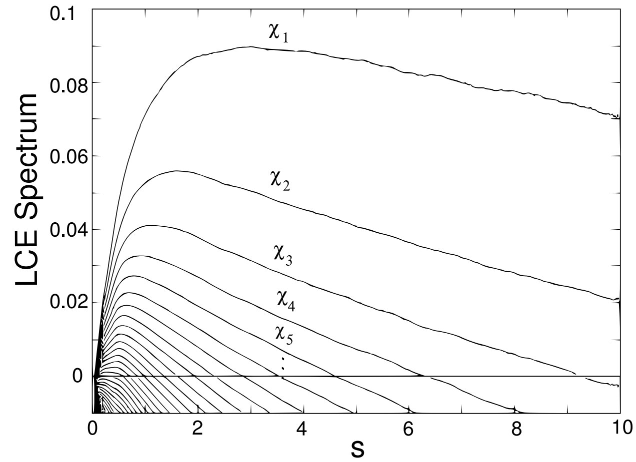

FIG. 1: LCE spectrum as a function of scale factor s for a network of 32 neurons and 64 dimensions. (15000 total time-steps, 5000 initial time-steps to arrive on the attractor.)

evidence demonstrating that such strict forms of uniqueness are present in many complex physical systems [14]. Indeed, alternative dynamical equivalence frameworks—such as, nonuniform partial hyperbolicity—were invented to circumvent the problems with nonuniqueness [15].

Figure 1 represents the typical scenario for the LCE spectra of the high-dimensional neural networks as a function of s. First, there is a lack of periodic windows, where $\chi_1 < 0$. Second, up to noise induced by initial condition variation and numerical errors, the LCEs vary continuously with s and have a single maximum. When we say *typical* here we refer to what is seen in the overwhelming fraction of 100s of such studies. In addition, we have also observed that the maximum number of positive Lyapunov exponents is approximately $d/4$ and the attractors' Kaplan-Yorke dimension is roughly $d/2$ [16]. When an LCE changes sign it happens over increasingly smaller s-intervals as d increases. These properties contrast sharply with the familiar low-dimensional scenario where one usually encounters a preponderance of stable behavior and periodic windows and the LCEs vary in a discontinuous manner with control parameters. (A more complete analysis of these observations is found in Ref. [11].)

These observations complement those from a previous study of chaos in neural-network continuous-time differential equations [17]. There a mean-field analysis, which assumes that inputs are statistically independent (and which does not apply in the present case), also suggested that chaos should be common.

In light of Fig. 1 we propose the following heuristic scenario. For a finite but arbitrarily high number of dimensions along an s interval—e.g., $s \in (2,8)$—there is a nearly dense, always countable set of points that have a LCE transversally crossing through zero. Thus, a continuous path along an s-interval will yield inevitable but noncatastrophic (i.e. $p > 1$) topological change. This implies that upon variation of parameters, periodic and quasiperiodic windows will not exist in chaotic regions of parameter space of dynamical systems with sufficiently high number of positive exponents (and entropy rate). The lack of dense periodic and quasiperiodic windows is a necessary condition for p-chaos with $p > 0$.

The claim is that high-dimensions provide a mechanism not afforded low-d dynamical systems, one which increases p with d and, more to the point, increases dynamical persistence. In support of this, we examine two sets of evidence, the first from along an s-interval and the second in the full $(n(d+1)+1)$-dimensional parameter space.

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 was increased above 64 the only networks with periodic windows had windows that persisted for most of the s-interval under consideration. Thus, as dimension was increased, periodic windows became increasingly rare but when they were observed, they were neither small nor intermittent, but instead dominated the dynamics.

To explore the full parameter space systematically, one can fix s and vary the weights with random perturbations of a given size. Doing this for all s values and perturbation sizes samples the entire parameter space and so probes the entire function space under consideration, but this is computationally very expensive. Instead we will consider data for networks with parameters varied in a $(n(d+1)+1)$-ball with its center fixed at s. The results are insensitive to s variation in the chaotic portion of the s-interval; i.e., nearby s values yield identical results. Similarly, the results are insensitive to perturbation size—perturbations on scales ranging from 10^{-10} to 1 yield very similar results.

Figure 2 shows the decrease in the probability of observing periodic windows in perturbed systems as the dimension is increased. Each data point corresponds to the probability of finding a system with a periodic orbit among a set of 500 networks at a given n and d and each perturbed 100 times. The range of random perturbations was 10^{-3} and $s = 3$. 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 so become more persistently chaotic. The final data point at $d = 128$ represents a single network (of the 500) that was found to be nonchaotic for one of the 100 perturbations—its largest LCE for this perturbation was on the order of 10^{-6}.

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. Thus, periodic windows which do exist are concentrated in an ever-decreasing fraction of networks. In other words, networks with one periodic window are more likely to have many periodic windows. (Detailed results and analysis along these lines will be presented elsewhere.)

These observations 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 dynamic (topological) variation with parameter change. All of the LCEs that become positive are negative for very small and very large values of s—exhibiting a single maximum. As the dimension d is increased, their variations with s decrease and they become smoother; recall the bands in which they fall in Fig. 1. Moreover, with increasing dimension the number of positive exponents increases monotonically [16]. Finally, the distance between LCE zero-crossings, above the maximum, decreases with dimension (Fig. 3).

This is all to say that, in the limit of very high dimension there are as many positive exponents as desired and they behave smoothly with parameter change. While the parameter change required to alter the absolute number of positive exponents decreases, the perturbation required for *all* the positive exponents to vanish grows substantially. Thus, the chance of falling into a periodic window vanishes in proportion and chaos becomes persistent. For instance, if one considers s values in the set $U = [0.2, 5]$, the variation in the number of positive exponents runs from 16 to 3. Nevertheless, chaotic dynamics are persistent over a considerably larger portion of parameter space. The 64-dimensional dynamical system of Fig. 1 exhibits persistent chaos of degree 3 (3-chaos) over at least the subset $[0.1, 10]$ [11].

Figure 3 is a graphical depiction of the hypothesized

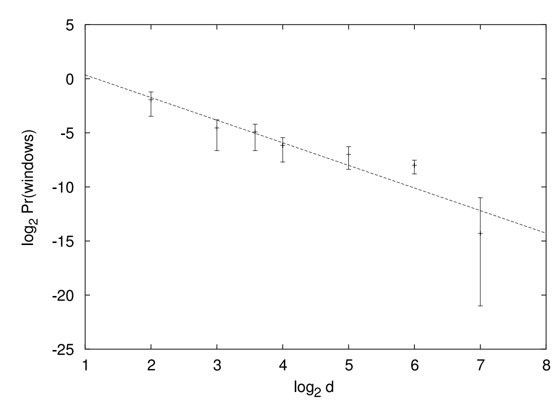

FIG. 2: Log probability of the existence of periodic windows versus log dimension for 500 cases per d. Each case has all the weights perturbed on the order of 10^{-3}; 100 times per case. The line of best fit is $\sim 1/d^2$.

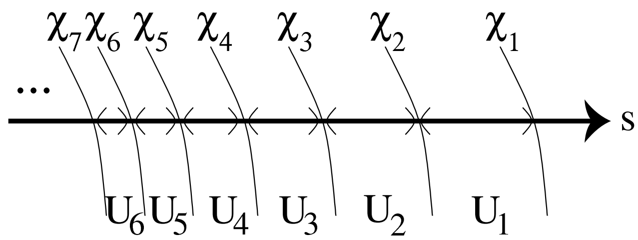

FIG. 3: Lyapunov spectrum versus s: U_i's are the open sets in parameter space where structural stability is believed to persist. The $|U_i|$ parameter intervals shrink with increasing i and as dimension increases.

mechanism—a plot of the s axis transversally intersected by LCEs. In sufficiently high dimensions, the subsets U_i of hyperbolic behavior shrink and eventually fall below any resolvability. The result, then, is twofold: one observes continuous topological change (bifurcations), but this is never catastrophic. One sees persistent chaos of varying degrees. The onset of "sufficiently high" dimension for this to occur for our dynamical systems was observed when $d \geq 30$. These investigations lead to the following conjecture for persistent chaos in high dimensions:

Conjecture 1 *Assume f is a network with a sufficiently large number d of dimensions and number of parameters $k = n(d+2)+1$. There exists an open set of significant positive Lebesgue measure in parameter space R^k for which chaos will be degree-p persistent, with $p \to \infty$ as $d \to \infty$.*

Since the networks are universal function approximators, then this state of affairs should be observed in typical nonlinear high-dimensional dynamical systems [19].

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 Lebesgue measure) 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, the persistence is related to the number of (linearly independent) parameters in the dynamical system. The number of neurons in the network effectively controls the entropy rate [18]—that is, increasing the number of neurons increases the entropy rate, number of positive exponents, and the maximum of the largest exponent. Increasing n simply increases the degree (p) of the persistent chaos, but the mechanism for persistent chaos remains, due to the decreasing probability of periodic windows. In fact, networks of very few parameters have considerably less persistent chaos.

In this way high entropy-rate systems are more persistent with respect to functional *and* parameter perturbations, and this accords 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. Here we described a mechanism in which the dynamical persistence of high-dimensional systems is retained upon 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 founded on dynamics.

We thank J. R. Albers, R. A. Bayliss, K. Burns, W. D. Dechert, D. Feldman, J. Robbin, C. R. Shalizi, and J. Supanich 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 DJA was provided by NSF grants DMR-9820816 and PHY-9910217 and DARPA Agreement F30602-00-2-0583.

[1] H. Poincaré, Oeuvres (Gauthir-Villars, Paris, 1929).
[2] J. D. Farmer, Phys. Rev. Lett. **55**, 351 (1985).
[3] S. Newhouse, Publ. Math. IHES **50**, 101 (1979).
[4] R. Mañé, Publ. Math. IHES **66**, 161 (1988).
[5] C. Pugh and M. Shub, Bull. Amer. Math. Soc. **41**, 1 (2003).
[6] J. F. Alves and M. Viana, Ergod. Th. and Dynam. Sys. **22**, 1 (2002).
[7] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica **15**, 21 (1979).
[8] D. Ruelle, Ann. Math. **115**, 243 (1982).
[9] K. Hornik, M. Stinchocombe, and H. White, Neural Networks **3**, 535 (1990).
[10] F. Takens, in *Lecture Notes in Mathematics*, edited by D. Rand and L. Young (Springer-Verlag, Berlin, Dynamical Systems and Turbulence, Warwick, 1981), vol. 898, pp. 366–381.
[11] D. J. Albers and J. C. Sprott, submitted.
[12] M. Andrecut and M. K. Ali, Phys. Rev. E **64**, 025203 (2001).
[13] S. Banerjee, J. A. Yorke, and C. Grebogi, Phys. Rev. Lett. **80**, 3049 (1998).
[14] K. Kaneko, Phys. Rev. E **66**, 055201 (2002).
[15] L. Barreira and Y. Pesin, *Lyapunov exponents and smooth ergodic theory.* (AMS, 2002).
[16] D. J. Albers, Ph.D. thesis, University of Wisconsin in Madison (2004).
[17] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).
[18] E. F. Manffra, Ph.D. thesis, Univesitat Gesamthochschule Wuppertal (2002).
[19] Our use of *typical* here refers to randomly sampling high-dimensional vector fields (ODEs) or functions (maps).