Dynamic stability of high dimensional dynamical systems

D. J. Albers[1,2,*] and J. C. Sprott[1,†]

[1] *Physics Department, University of Wisconsin, Madison, WI 53706*
[2] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501*

(Dated: May 7, 2004)

We investigate the dynamical stability conjectures of Palis and Smale, and Pugh and Shub from the stand point of numerical observation and put forth a stability conjecture of our own. We find that as the dimension of a dissipative dynamical system is increased, the number of positive Lyapunov exponents increases monotonically, the number of observable periodic windows decreases at least below numerical precision, and we observe a subset of parameter space such that topological change is very common with small parameter perturbation. However, this seemingly inevitable topological variation is never catastrophic (the dynamic type is preserved) if the dimension of the system is high enough.

PACS numbers: 05.45.-a, 89.75.-k, 05.45.Tp, 89.70.+c, 89.20.Ff
Keywords: Partial hyperbolicity, dynamical systems, structural stability, stability conjecture, Lyapunov exponents, complex systems

Since the days of Poincaré, mathematicians and physicists have worried about the stability of solutions of differential equations and dynamical systems in the space of functions used to approximate physical phenomena. If mathematical models are to represent the real world, and the real world is dynamically stable [22] (e.g. turbulence as a dynamical phenomena is quite resistant to perturbations), our models must also have persistent dynamical behavior. We will discuss two of the leading stability conjectures of mathematicians (the Smale-Palis conjecture has been proven; the Pugh-Shub conjecture remains an open question in its full generality), as well as posing one of our own. Our conjecture will be rooted in numerical results of a statistical study of a very general set of dynamical systems, the set of feed-forward neural networks. Whereas the stability conjectures of Smale and Palis, and Pugh and Shub are often difficult to observe or relate to physical phenomena, ours, being based in computation, will yield insight into what a numerical or experimental scientist is likely to observe. The key notions of all stability conjectures are of dynamics and of equivalence between dynamics. Our notions of dynamics and equivalence are defined in a more practical way for numerical scientists than those of Smale and Palis, and Pugh and Shub. We claim that despite a possible lack of strict dynamical stability as it is often defined in rigorous mathematical frameworks, dynamic instability in large dynamical systems is likely benign.

As the background for his report is not necessarily well known by the physics audience, we will provide a brief summary if the results relevant to this report. In the now famous volume 14 [1], Smale and Palis put forth a stability conjecture for C^1 diffeomorphisms from compact manifolds to themselves. The proof of one direction of the conjecture was provided by Robbin [2] and Robinson [3], while the other direction was shown by Mañé [4] much later. The notion Palis and Smale used to distinguish dynamic stability between dynamical systems is that of topological conjugacy. Two dynamical systems, f and g, are said to be C^0 conjugate, or topologically conjugate, if there exists a homeomorphism h such that $f = h^{-1}g \circ h$. The diffeomorphism f is called structurally stable if and only if there exists a C^0 neighborhood V of f, such that for all $g \in V$, g is topologically conjugate to f. Let us now state the stability theorem:

Theorem 1 (A structural stability theorem — Mañé [4], Robbin [2], Robinson [3]) *A C^1 diffeomorphism (on a compact, boundaryless manifold) is structurally stable if and only if it satisfies axiom A and the strong transversality condition.*

Recall that axiom A says the diffeomorphism is hyperbolic with dense periodic points on its non-wandering set Ω ($p \in \Omega$ is non-wandering if for any neighborhood U of x, there is an $n > 0$ such that $f^n(U) \cap U \neq 0$). The diffeomorphism f satisfies the strong transversality condition if and only if, for every $x \in M$, $M_x = E^s_x + E^u_x$ where E^s_x and E^u_x are the stable and unstable manifolds for f at x.

The requirement for structural stability was quite strict; famous examples such as the Lorenz equations [5] [6] are not strictly structurally stable. Nevertheless, variation of parameters of the Lorenz equations during numerical simulation yields a sizable set of parameters for which chaos is seemingly robust. We will address a possible mechanism (not the mechanism existent in the Lorenz equations) for the failure of numerically (or experimentally) observable structural stability along a parameterized curve in high-dimensional dynamical systems.

To circumvent some of the problems with the Smale and Palis conjecture, as well as to provide a more general stability conjecture, Pugh and Shub put forth a stability conjecture based on a more inclusive notion of dynamics,

*Electronic address: albers@santafe.edu
†Electronic address: sprott@physics.wisc.edu

ergodicity. We will state the most recent version of their stability theorem which we will comment on later:

Theorem 2 (Pugh-Shub theorem (theorem A [7])) *If $f \in Diff^2_\mu(M)$ is a center bunched, partially hyperbolic, dynamically coherent diffeomorphism with the essential accessibility property, then f is ergodic.*

Recall, a measure-preserving diffeomorphism $f : M \to M$ (M is compact, $Diff^2_\mu(M)$ is the space of C^2 diffeomorphisms of M that are measure preserving) is partially hyperbolic if the tangent space of M splits into the Tf invariant sum $E^u \oplus E^c \oplus E^s$ [8] [23]. Ergodic behavior implies that, upon breaking the attractor into measurable sets, V_i, for f applied to each measurable set for enough time, $f^n(V_i)$ will intersect every other measurable set, V_j. This implies a weak sense of recurrence; for instance, quasi-periodic orbits, chaotic orbits, and some random processes, are at least colloquially ergodic. More formally, a dynamical system is ergodic if and only if almost every point of each set visits each set with positive measure. Ergodicity is a precise mathematical notion meant to capture Boltzmann's ergodic hypothesis. The accessibility property simply formalizes a notion of one point being able to reach another point. Given a partially hyperbolic dynamical system, $f : M \to M$ such that there is a splitting on the tangent bundle $TM = E^u \oplus E^c \oplus E^s$, and $x, y \in M$, y is accessible from x if there is a C^1 path from x to y whose tangent vector lies in $E^u \cap E^s$ and vanishes finitely many times. The diffeomorphism f is center bunched if the spectra of Tf corresponding to the stable ($T^s f$), unstable ($T^u f$), and ($T^c f$) central directions lie in thin, well separated annuli (see [9], page 131 for more detail — the radii of the annuli is technical and is determined by the Holder continuity of the diffeomorphism.) Lastly, let us note that a dynamical system is called stably ergodic if, given $f \in Diff^2_\mu(M)$ (again M compact), there is a neighborhood, $f \in U \subset Diff^2_\mu(M)$, such that every $g \in U$ is ergodic with respect to μ. We will refrain from divulging an explanation of dynamical coherency; it is a very crucial characteristic for the proof of theorem 2, but we have little to say in its regard.

For our stability conjecture, we will be considering C^r ($r > 1$) discrete-time dynamical systems with a single parameter (in R^1) of compact subsets of R^d to themselves. Let us begin with a definition:

Definition 1 (Robust chaos of degree k) *Given a C^r ($r > 0$) discrete-time map from a compact set to itself, the map is robustly chaotic of degree k on a open subset $U \in R$ in parameter space if, for all $s \in U$, and for all initial conditions of the map f at $s \in U$, f retains at least k positive Lyapunov exponents.*

Robust chaos of degree k is our notion of dynamical equivalence on an open set of parameter space. We would like to note a difference between this notion and the notion of a robust chaotic attractor put forth in [10] [11]. In definition 1, we do not require that the attractor be unique on the subset U. This is an important distinction; the uniqueness of an attractor is significantly more difficult to demonstrate, and there is little evidence demonstrating that such uniqueness is present in many complicated physical systems [12] [13].

We claim that in very high dimensions (high for our dynamical systems will be $d \geq 30$), there will exist a sizable subset of parameter space such that very small variations in parameter space will cause subtle topological change, i.e., a change in the number of positive Lyapunov exponents. Nevertheless, this inevitable topological change will usually not be catastrophic, as only a small proportion of Lyapunov exponents will undergo a sign change. Thus, there will exist sizeable subsets of parameter space such that there will exist robust chaos of degree $k > 1$. For many scientists working on dynamically complicated physical experiments (such as fluid dynamics or plasma physics), dynamic stability is not a novel concept, for often these scientists spend much effort attempting to eliminate this chaotic dynamic stability. We are proposing the mechanism under which dynamic stability of high-dimensional systems remains under parameter perturbation.

Lyapunov exponents will be our chief tool of analysis and identification for dynamic types (for how to compute them, see [14]). There are many reasons why we have chosen Lyapunov exponents, but the foremost reason is that Ruelle [15] has shown that there is an equivalence between the number of negative and positive Lyapunov exponents and the number of global stable and unstable manifolds, respectively. Note that when we refer to topological variation in the context of our dynamical systems, we mean a change in the number of positive Lyapunov exponents.

The single-layer feed-forward neural networks we will consider are of the form

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

which is a map from R^d to R. In eq.(1), n represents the number of hidden units or neurons, d is the input or embedding dimension of the system which functions simply as the number of time lags, and s is a scaling factor on the weights.

The parameters are all real ($\beta_i, w_{ij}, x_j, s \in R$), the β_i's and w_{ij}'s are elements of weight matrices (which we hold fixed for each case), $(x_0, x_1, \ldots, x_d)$ represent initial conditions, and $(x_t, x_{t+1}, \ldots, x_{t+d})$ represents the current state of the system at time t. We assume that the β's are *iid* uniform over $[0, 1]$ and then re-scaled to satisfy $\sum_{i=1}^{n} \beta_i^2 = n$. The w_{ij}'s are *iid* normal with zero mean and unit variance. The s parameter is a real number, and it can be interpreted as the standard deviation of the w matrix of weights. The initial x_j's are chosen *iid* uniform on the interval $[-1, 1]$.

We would like to make a few comments with respect to our squashing function, tanh(). First, $\tanh(x)$, for $|x| \gg 1$ will tend to behave much like a binary function.

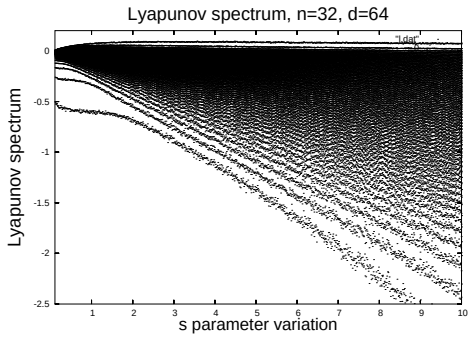

FIG. 1: LE spectrum: 32 neurons, 64 dimensions.

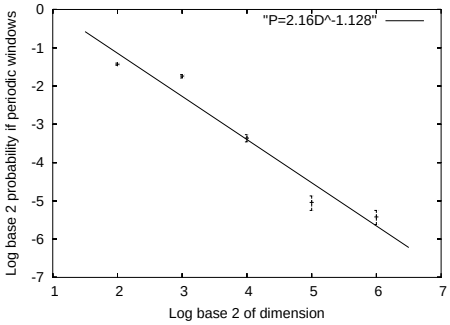

FIG. 2: $\log_2$ of the probability of periodic or quasi-periodic windows versus $\log_2$ of dimension. The line $P_w = 2.16D^{-1.128}$ is the least squares fit of the plotted data.

Thus, the states of the neural network will tend toward the finite set $\{\pm\beta_1, \pm\beta_2, \ldots, \pm\beta_n\}$, or each x_t can have 2^n different states. In the limit where the arguments of tanh() become infinite, the neural network will have periodic dynamics. The other extreme of our squashing function also yields a very specific behavior. For x very near 0, the $\tanh(x)$ function is nearly linear. Thus choosing s to be small will force the dynamics to be mostly linear, again yielding fixed point and periodic behavior (no chaos). Thus the scaling parameter s will provide a unique bifurcation parameter that will sweep from linear ranges to highly non-linear ranges, to binary ranges - fixed points to chaos and back to periodic phenomena.

Scalar neural networks are "universal approximators," meaning they can approximate many very general spaces of mappings (e.g. any C^1 mapping and it's derivatives to arbitrary accuracy given enough neurons). That scalar neural networks are able to approximate the mappings we are interested in is a topic addressed in [17]. Combining the approximation theorems of [17] and the time series embedding results of Takens [18] shows the equivalence between the neural networks of this section and the dynamical systems per our conjecture (for specific arguments along these lines, see [16]).

In this report we will present two figures demonstrating why we believe our conjecture is reasonable. Two much more detailed and exhaustive reports are in preparation [16], [19].

Figure (1) represents the typical scenario for our high dimensional neural networks, there is a lack of periodic windows and the Lyapunov exponents look quite continuous with parameter variation and have a single maximum. Based on observing of hundreds of such cases, we also wish to report the other observations. The maximum number of positive Lyapunov exponents is roughly one quarter the number of dimensions and the Kaplan-Yorke dimension was roughly one half of d. Most of the exponent sign changes occur over an increasingly short subsets as the dimension is increased (e.g. the U_i sets in Fig. (3)) of the parameter space as d is increased. This behavior is contrasted by the dynamical behavior at lower dimensions where there are many periodic windows and the Lyapunov exponents vary in a discontinuous manner with parameter variation.

Of the aforementioned general characteristics of the high dimensional dynamical systems we have observed, we wish to present results regarding the disappearance of periodic and quasi-periodic windows as the dimension is increased. The lack of dense periodic and quasi-periodic windows is a necessary conditions for robust chaos of $k-$degree for $k > 0$. However, if periodic windows exist but are not dense, then our conjecture will remain in tact. Figure (2) is a $\log_2 - \log_2$ plot of the probability of a window existing for a given s value of a given network for $s \in [2, 4]$ with s increments of 0.001. Considering Fig. (2), it should be clear that the probability of a window existing decreases exponentially with increasing dimension. Moreover, the error bars, which represent the standard error given the number of independent windows, are increasing in size due to the fact that the number of independent windows is decreasing with dimension also. If periodic windows were to disappear entirely, which we do not claim we can show here, we would be done. Never-the-less, the fact that the probability and number of windows is decreasing is strong evidence that our dynamical systems will likely be, as the dimension becomes sufficiently high, robustly chaotic.

Based on observing 400 four-dimensional dynamical systems and 200 64-dimensional dynamical systems (as well as many dynamical systems of intermediate dimensions), we would like to put forward what we believe is the mechanism for dynamics variation upon parameter variation in many high-dimensional, chaotic dynamical systems. In our networks, all the exponents that are positive are negative for very small and very large values of s. As the dimension is increased, the Lyapunov exponents become more continuous with respect to parameter variation; the positive Lyapunov exponents have a single maximum, and are monotonically increasing on one side of the maximum, and monotonically decreasing on the other side of the maximum. Also, as the dimension is increased the number of positive exponents increases monotonically. Lastly, the distance between exponent zero crossings after the exponent reaches a maximum will decrease with dimension (see Fig. (3)). Therefore, in the limit of very high dimensions, there will exist as many

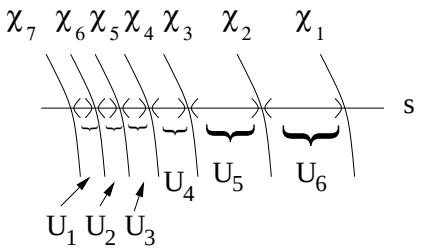

FIG. 3: Drawn close-up of s versus Lyapunov spectrum. The U_i's are the open sets upon which structural stability is believed to persist. $|U_i|$ shrinks as the dimension is increased.

positive exponents as desired; the Lyapunov exponents behave relatively continuously with parameter change; and thus parameter change required to alter the number of positive exponents will decrease, but the perturbation required for all the positive exponents to vanish will be considerable. For instance, if one considers the set $U = [0.2, 5]$, the variation in the number of positive exponents runs from 16 to 3. Nevertheless, the chaotic dynamics is persistent over a considerably larger portion of parameter space. The 64 dimensional dynamical system in Fig. (1) exhibits robust chaos of degree 3 over the subset $[0.1, C]$ where $C > 10$.

For a graphical understanding of what we claim, see Fig. (3), a plot of the s axis transversally intersected by various Lyapunov exponents. We claim that the sequential subsets U_i of hyperbolic behavior fall below any numerical resolution with a sufficiently high number of dimensions. What will be observed in such a situation will be continuous topological change (bifurcations) with parameter variation. This topological change however will never be catastrophic when the dimension is sufficiently high. This is the motivation for $k-$degree robust chaos.

Our results seem to disagree somewhat with the windows conjecture of Barreto et. al. [20] (for other possible counter-examples see [11]). However, much work remains to fully understand the relationship our conjecture and results and that of Barreto et. al.

We do not claim counter-examples to either the Smale-Palis stability results nor the Pugh-Shub conjecture. Rather we claim the effects of the Smale-Palis results may not be observable in high dimensional dynamical systems. Our results yield support of and agreement with the Pugh-Shub conjecture.

Pugh and Shub have been attempting to remove the dynamical coherency and center bunching conditions from their stability conjecture. Our results show a large set of dynamical systems that are seemingly stably ergodic without center bunching, supporting the claims of Pugh and Shub. One reason in physicists would be interested in riding the current version of the Pugh-Shub of the center-bunching condition excludes many important physical examples.

D. J. Albers would like to thank J. R. Albers, R. A. Bayliss, K. Burns, W. D. Dechert, D. Feldman, J. Robbin, C. R. Shalizi, and J. Supanich for many helpful discussions and advice. D. J. Albers would like to give special thanks to J. P. Crutchfield for much guidance, many fruitful discussions, and support. The computing for the project was done on the Beowulf cluster at the Santa Fe Institute and was partially supported at the Santa Fe Institute under the Networks, Dynamics Program funded by the Intel Corporation under the Computation, Dynamics, and Inference Program via SFI's core grant from the National Science and MacArthur Foundations. Direct support for D. J. Albers was provided by NSF grants DMR-9820816 and PHY-9910217 and DARPA Agreement F30602-00-2-0583.

[1] S. Chern and S. Smale, eds., *Global Analysis*, vol. 14 of *Proc. Sympos. Pure Math.* (A.M.S., Bekeley, Ca, 1970).
[2] J. Robbin, Annals math. **94**, 447 (1971).
[3] C. Robinson, in *Dynamical Systems* (Academic Press, Salvador 1971, 1973), pp. 443–449.
[4] R. Mané, Publ. Math. I.H.E.S. **66**, 161 (1988).
[5] E. N. Lorenz, J. Atmosph. Sci **20**, 130 (1963).
[6] R. F. Williams, Publ. Math. IHES (1980).
[7] C. Pugh and M. Shub, J. Eur. Math. Soc. **2**, 1 (2000).
[8] M. I. Brin and J. B. Pesin, English transl. Math. Ussr Izv **8**, 177 (1974).
[9] C. Pugh and M. Shub, J. of Complexity **13**, 125 (1997).
[10] S. Banerjee, J. A. Yorke, and C. Grebogi, Phys. Rev. Lett. **80**, 3049 (1998).
[11] M. Andrecut and M. K. Ali, Phys. Rev. E (2001).
[12] K. Kaneko, Phys. Rev. E (2002).
[13] U. Feudel and C. Grebogi, Phys. Rev. Lett. (2003).
[14] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, **15**, 9 (1979).
[15] D. Ruelle, Ann. of math. **115**, 243 (1982).
[16] D. J. Albers and J. C. Sprott, in preparation.
[17] K. Hornik, M. Stinchocombe, and H. White, Neural Networks **3**, 535 (1990).
[18] 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.
[19] D. J. Albers, J. C. Sprott, and J. P. Crutchfield, in preparation.
[20] E. Barreto, B. Hunt, C. Grebogi, and J. A. Yorke, Phys. Rev. Lett. **78**, 4561 (1997).
[21] J. Palis and M. Viana, Ann. of math. **140**, 207 (1994).
[22] By dynamic stability we mean the persistence of a type of dynamics (e.g. fixed points, strange attractors) with functional or parameter variation, not orbital stability.
[23] The major point of partial hyperbolicity is it allows for zero Lyapunov exponents - or neutral directions in the derivative