Shadowing breakdown and large errors in dynamical simulations of physical systems

Timothy D. Sauer[a]

[a]*Department of Mathematical Sciences, George Mason University, Fairfax, VA 22030*

Abstract

Simulations play a crucial role in the modern study of physical systems. A major open question for long dynamical simulations of physical processes is the role of discretization and truncation errors in the outcome. For the first time, a general mechanism is described that can cause extremely small noise inputs to result in errors in simulation statistics that are several orders of magnitude larger. A scaling law for the size of such errors in terms of the noise level and properties of the dynamics is given. This result brings into question trajectory averages that are computed for systems with particular dynamical behaviors, in particular systems that exhibit fluctuating Lyapunov exponents or unstable dimension variability.

Typeset using REVTEX

The advent of high-performance computing has put a valuable tool in the hands of the physical scientist. Studies of long-time celestial mechanics, meterological and climate phenomena, molecular and protein dynamics, and countless other areas have been revolutionized by the ability to simulate dynamical processes over many time steps. When chaotic dynamics enters the picture, as is known to happen in the above examples, the question of the validity of the simulation results becomes complicated. Computer simulations by their very nature make small truncation errors as the model is advanced in time. When small one-step errors are made, chaotic dynamics destroys the possibility of following individual trajectories accurately, due to sensitivity to initial conditions. In such a case, scientists often abandon the goal of studying individual trajectories, and resort to estimating ergodic averages, say for estimating the mean CO_2 level in a climate model or temperature in a molecular simulation. A large literature on shadowing theory [1,2] has developed to try to justify this approach, essentially by locating or proving the existence of true model trajectories, with possibly different initial conditions, that stay near the simulation trajectory for long times. If successful, this can be a justification of the validity of the statistic being computed by the simulation. However, it has become clear from recent research that shadowing fails for certain systems, in a readily quantifiable way [3,4]. In other words, even when the model is well-specified, and the approximately correct chaotic attractor is formed by the simulation, *there is no fundamental reason for computer-simulated long-time statistics to be even approximately correct.*

In this article we complete this sequence of observations by exhibiting a family of examples for which large errors in long-time statistics indeed occur, even though the dynamics is restricted to an approximately correct attractor and only tiny one-step errors are made. The reason for the large simulation error is that shadowing breakdown causes a systematic, macroscopic bias in the natural ergodic measure of the attractor which is several orders of magnitude greater than the one-step error size. The examples have only two degrees of freedom, and are made to be as simple as possible, to show the likely pervasiveness of this phenomenon. More complicated, higher-dimensional examples will have correspondingly larger effects. In addition, we propose a scaling formula for the expected error of the computed statistic, and identify the exponent in the scaling formula with dynamical quantities from the underlying deterministic system. It is apparent from the formula that when this so-called hyperbolicity exponent is very small, the potential exists for large errors in measured statistics.

For chaotic systems, sensitivity to initial data makes computer simulations especially problematic. The distance between nearby trajectories grows exponentially as a function of time. When running a computer simulation of a chaotic system, continual small errors on the order of machine roundoff are made, making the computed trajectory distinct from the true trajectory that is the goal of the simulation. The distance between the computed trajectory and the original true trajectory therefore grows exponentially. However, it is possible that another true trajectory, with a slightly different initial condition from the original true trajectory, remains relatively close to the computed trajectory, or *shadows* the computed trajectory, for long computing times. For ideal (hyperbolic) chaotic dynamics such shadowing trajectories were shown to exist by Anosov and Bowen [1], leading to a great deal of research in that area. A dynamical system is hyperbolic if phase space can be spanned locally by a fixed number of independent stable and unstable directions which are consistent under the operation of the dynamics. Methods of shadowing have shown that for chaotic systems that are nearly hyperbolic [2], locally sensitive trajectories are often *globally insensitive*, in that there exist (true) shadowing trajectories, very close to long computer-generated pseudo-trajectories.

The situation was clarified further by the development of rough criteria [3] for when a dynamical system was too far from hyperbolicity to have long shadowing trajectories. In [4], a scaling law was developed for the expected length of shadowing trajectories. The scaling exponent was expressed in terms of the distance of the finite-time Lyapunov exponents (FTLE) to zero. If a FTLE moves back and forth through zero during a trajectory, it represents a tangential direction that is uncertain between expansion and contraction, which precludes the existence of a hyperbolic structure on the attractor. This effect, also called unstable dimension variability, has been the focus of much recent work [4,5].

In spite of the improved understanding of shadowing failure in the case of fluctuating Lyapunov exponents and unstable dimension variability, the connection to computing long-term averages has not been clear. In fact, it has been theorized that perhaps long term averages can be computed correctly in the absence of long correct trajectories [6]. However, while this case may be common, in this article we show that the possibility exists for the reverse to be true. We define a family of examples that exhibits shadowing failure because of unstable dimension variability, and for which microscopic one-step errors cause macroscopic errors in long-term averages.

A typical interesting computed quantity in a simulation of a chaotic dynamical system can be expressed as a long term average of a function, called the observation function, over the trajectory of the system. We will be concerned with ergodic systems for which the probability distribution of trajectory points forms a *natural measure* of the global attractor. If the observation function is the characteristic (or indicator) function of a subset of phase space, then the asymptotic average is the natural measure of that subset. We will make use of this fact in the heuristic argument below.

The shadowing results of [4] turned on the development of a scaling law for shadowing distance, defined to be the distance between a given point on a computed trajectory and its corresponding point on the shadowing true trajectory, assuming it exists. The scaling law says that the distribution of the log shadowing distances along the trajectory is approximately an exponential distribution

$$p(y) = he^{-h(y-\ln\delta)} \tag{1}$$

for $y \geq \ln\delta$, where δ is the one-step error size and h is a scaling exponent expressing the severity of the fluctuation of Lyapunov exponents:

$$h = \frac{2|m|}{\sigma^2} \tag{2}$$

Here m and σ represent the mean and standard deviation, respectively, of the distribution of the finite-time Lyapunov exponent closest to zero [7].

FIGURES

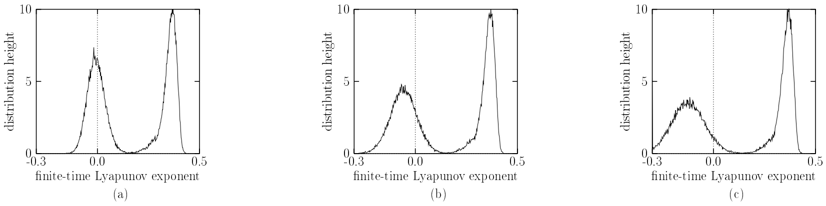

FIG. 1. Distribution of time-100 Lyapunov exponents of planar map (9), for parameters $a = 0.9, c = 0.9, d = 0.0$ and various b. (a) $b = 1.0$ (b) $b = 1.1$ (c) $b = 1.2$

In light of the connection of the quantity h to various aspects of nonhyperbolicity, we will call h the *hyperbolicity exponent* for the remainder of this article. Fig. 1 shows an example set of FTLE distributions for a two-dimensional map that we examine more closely below. (The time-t Lyapunov exponents of a chaotic trajectory are the averages λ_i of the logarithm of local expansion rates along the trajectory of length t, so that an infinitesimal sphere of radius dr at the beginning of the trajectory would evolve to an ellipsoid with axes $\lambda_i^t dr$ after t time units.) The shadowing distance scaling law (1) is closely related to intermittency, and may be thought of as "intermittency in miniature". The exponential distribution is the result of tiny excursions that periodically move the computed trajectory away from the true trajectory, and then return toward it. The cumulative distribution function of the log shadowing distances, obtained by integrating the probability distribution function (1), is $P(y) = 1 - e^{-h(y - \ln \delta)}$. The probability that the shadowing distance is greater than D is therefore

$$e^{-h(\ln D - \ln \delta)} = (\frac{\delta}{D})^h. \tag{3}$$

A heuristic argument allows us to develop a scaling formula for the error in a trajectory average due to small one-step errors. Our goal will be to show that for a fixed observation function $g(x)$ defined on the phase space of the system, the difference between the computed trajectory average of g and the true trajectory average of g scales with the one-step error δ with a scaling exponent of $h = 2|m|/\sigma^2$, or in terms of expectation over the computed natural measure and the true natural measure:

$$\langle g\rangle_{\text{computed}} - \langle g\rangle_{\text{true}} = K\delta^h. \tag{4}$$

First we will reduce to the case where g is the characteristic function of a disk S in the phase space. If we can show (4) to hold for all disk subsets, then we know it holds for the natural measure itself, and therefore for any function integrated over the natural measure. The trajectory average of $g(x) = \chi_S(x)$ is simply the proportion of time spent in the disk S by the trajectory. We will calculate the probability that $\chi_{S,computed}$ and $\chi_{S,true}$ differ, for a fixed time t. Let U, T respectively denote the corresponding points on the computed and true trajectories. Assume first that T lies in the set S and we calculate the probability that U lies outside the set. Consider the set of half-lines L radiating from the point T, as in Fig. 2. If U lies on the half-line L, then it lies outside S if and only if the distance $d(T, U)$ between T and U is at least $d(T, L_T)$. Using (3), the probability that U lies outside of the set S is

$$P(U \text{ not in } S|T \text{ in } S) = \int_L (\frac{\delta}{d(T, L_T)})^h p_{U,L} dL, \tag{5}$$

where $p_{U,L}$ denotes the probability that U lies on the half-line L. Assuming that this probability is independent of δ, the terms relying on δ factor out of the integral to give

$$P(U \text{ not in } S|T \text{ in } S) = K_1\delta^h. \tag{6}$$

A similar argument shows that

$$P(T \text{ not in } S|U \text{ in } S) = K_2\delta^h, \tag{7}$$

and we may conclude that

$$P\{\chi(\text{computed orbit}) \neq \chi(\text{true orbit})\} \propto \delta^h, \tag{8}$$

which is the claim (4).

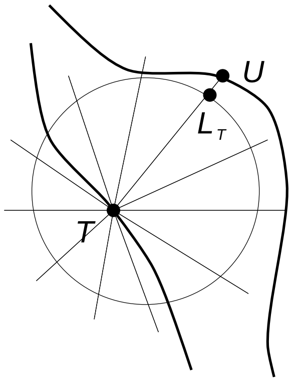

FIG. 2. Schematic view of heuristic argument. Under the condition that the trajectory point U lies on half-line L, the probability that U lies outside of the disk S shown can be expressed in term of the probability distribution of distances between T and U.

The role of roundoff errors of size δ in computing trajectory averages is now clear from (4). If the hyperbolicity exponent h is near 1, little propagation of error will occur from the small scale to the large. On the other hand, if $h \approx 0$, there may be a large bias, or difference in expected values, of the observation function $g(x)$ under the computed and true probability measures. Note that the proportionality constant K in (4) is dependent on the observation function g. In fact, K may be zero or small enough to make the bias zero or undetectably small for some choices of g, and large for others.

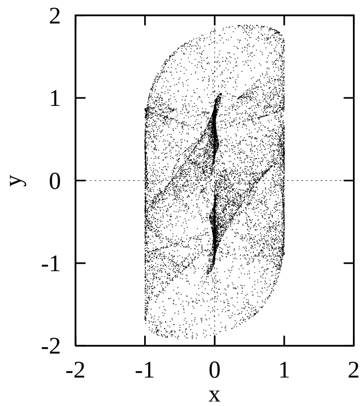

FIG. 3. One trajectory of the chaotic attractor generated by the planar map (9).

We illustrate a case for which the bias in natural measure is easily detectable with the two-dimensional example

$$f(x, y) = (\sin x(a\pi \cos \pi y + b), c(\sin(\pi y + d) - \sin d) + \sin \pi x). \quad (9)$$

The chaotic attractor for this map, shown in Fig. 3 for a particular choice of parameters, has fractal dimension between one and two. Fig. 1 showed the distributions of finite-time Lyapunov exponents for this planar discrete dynamical system, with parameters set to $a = 0.9, c = 0.9, d = 0.0$, and for various b. The smaller Lyapunov exponent fluctuates around zero in the three cases shown in Fig. 1, but most strongly in the case $b = 1.0$. The mean and standard deviation [8] for the time-100 Lyapunov exponents in this case are $m = 0.006$ and $\sigma_{100} = 0.044$, giving $\sigma = 0.44$ and a hyperbolicity exponent of $h = 0.058$, from formula (2). For the case $b = 1.2$, on the other hand, the smaller Lyapunov exponent has moved somewhat away from zero, $m = .124$ and $\sigma_{100} = 0.081$, leading to a much higher exponent $h = 0.377$. According to (4), we expect that computing accurate trajectory averages will be much more difficult for $b = 1.0$ than for $b = 1.2$.

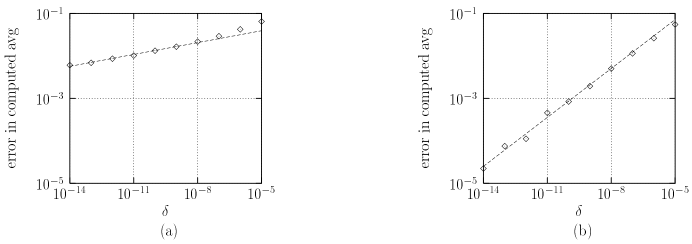

FIG. 4. Plot of the difference in the average value of $g(x, y) = y^2$, under a true ($\delta = 0$) versus a noisy ($\delta > 0$) trajectory, as a function of δ. Parameter values for the map (9) are $a = 0.9, c = 0.9, d = 0.0$ and (a) $b = 1.0$ (b) $b = 1.2$. Power law scaling is clearly evident; the slopes (a) 0.073 and (b) 0.39 agree with the hyperbolicity exponents $h = 2m/\sigma^2$.

To test (4) quantitatively, we computed trajectory averages of a variety of observation functions $g(x, y)$, for various noise levels δ. We computed the trajectory average in high precision (used to represent the "true" value) for a long trajectory, and compared it to the trajectory average from a trajectory computed with one-step random error of size δ. For $g(x, y) = y^2$, the resulting differences in average $\langle g \rangle_{\text{computed}}$ - $\langle g \rangle_{\text{true}}$ are plotted in Fig. 4 as a function of δ. They follow a power law as a function of δ, in agreement with the heuristic argument above. The exponent from the power law can be estimated by the least squares lines shown in Fig. 4. In Fig. 5, we plot the results of estimating the power law exponent as done in Fig. 4, for several parameter values b. Also plotted in Fig. 5 are the corresponding values of $h = 2m/\sigma^2$, computed from the Lyapunov exponent distributions as in Fig. 1. They are in reasonably good agreement [9], as predicted by formula (4).

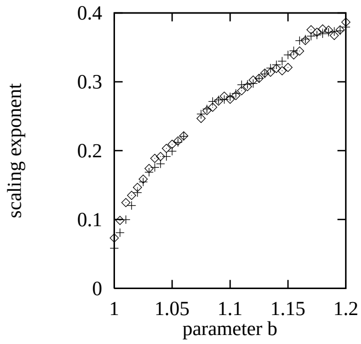

FIG. 5. Comparison of hyperbolicity exponent h derived from finite-time Lyapunov exponents (2) (small crosses) and the scaling exponent (diamonds) from long simulations of trajectory averages, using (4). They should agree, according to the heuristic argument of the text. Results are omitted near $b = 1.07$ because the attractor disappears due to an interior crisis.

It is important to note that the differences between true and computed averages are not due to the length of trajectory over which the averages were taken (about 10^7 steps, in this case). The difference is one of bias, not variance, and persists as the length of test trajectory is made arbitrarily large.

This family of examples shows the unexpected effects that very small noise levels can have on the computation of global averages in a dynamical setting. In Fig. 4, these effects are evident. When the hyperbolicity exponent h is around 0.3, as in frame (b) of the figure, one-step error of size 10^{-14} propagates to errors of size 10 to the power $-(14)(0.3) \approx -4$, a magnification of 10 orders of magnitude. In the more serious case of Fig. 4(a), one can expect errors of macroscopic size.

Fluctuating Lyapunov exponents and unstable dimension variability are expected to be common in high dimensional systems. When these effects are present, physical observables computed from long-term simulations may converge to a biased, erroneous value. The argument is often made that such statistical averages can be computed accurately by amalgamating many short correct trajectories, consisting of the shadowing trajectories corresponding to short pieces of the computed trajectory. However, while these short shadowing trajectories may be correct, their initial conditions suffer from the same bias (4) and offer no improvement in computing a correct average. Similarly, Monte Carlo simulations, achieved by starting at many different initial conditions and averaging the results, will have the same communal bias. In such cases no method for eliminating this large error in the observed physical quantity, beyond computing in even higher precision, is presently known.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation.

REFERENCES

[1] D. V. Anosov, Proc. Steklov Inst. Math. **90**, 1 (1967). R. Bowen, J. of Differential Equations **18**, 333 (1975). See also the more general Theorem S.4.14 in A. Katok, B. Hasselblatt, Introduction to the Modern Theory of Dynamical Systems, Cambridge Univ. Press (1995).

[2] C. Grebogi, S. Hammel, and J. Yorke, J. Complexity **3**, 136 (1987). C. Grebogi, S. Hammel, and J. Yorke, Bull. Amer. Math. Soc. **19**, 465 (1988). C. Grebogi, S. Hammel, J. A. Yorke, T. Sauer, Phys. Rev. Lett. **65**, 1527 (1990). S-N. Chow and K. Palmer, Dynamics and Diff. Eqns. **3**, 361 (1991). J. M. Sanz-Serna and S. Larsson, Appl. Num. Math. **13**, 181 (1993). G. D. Quinlan, S. Tremaine, Mon. Not. R. Astron. Soc. **259**, 505 (1992). E.S. Van Vleck, SIAM J. Sci. Comput. **16**, 1177 (1995).

[3] S. Dawson, C. Grebogi, T. Sauer, J.A. Yorke. Phys. Rev. Lett. **73**, 1927 (1994).

[4] T. Sauer, C. Grebogi, and J.A. Yorke, Phys. Rev. Lett. **79**, 59 (1997).

[5] R. Abraham and S. Smale, Proc. Symp. Pure Math. (AMS) **14**, 5 (1970). S.P. Dawson, Phys. Rev. Lett. **76**, 4348 (1996). E. Kostelich, I. Kan, C. Grebogi, E. Ott, and J.A. Yorke, Physica D **109**, 81 (1997). Y.C. Lai, D. Lerner, K. Williams, C. Grebogi. Phys. Rev. E **60**, 5445 (1999). R. Viana, C. Grebogi, Phys. Rev. E **62**, 462 (2000). E. Barreto, P. So, Phys. Rev. Lett. **85**, 2490 (2000).

[6] Y.-C. Lai, C. Grebogi, and J. Kurths, Phys. Rev. E **59**, 2907 (1999).

[7] The scaling formula is an asymptotic formula, which is true as $\delta \to 0$ and to the extent that one FTLE is closer to zero than all others. If two FTLE's are equally distant from zero, the formula (4) may not hold. The correct formula remains to be worked out for such special case.

[8] The mean m is computed as the mean of the second-largest time-100 Lyapunov exponent, which is an approximation to the (infinite-time) Lyapunov exponent. The standard

deviation σ was computed as ten times the standard deviation of the distribution of the second-largest time-100 Lyapunov exponent.

[9] Fig. 5 shows a small overestimation of the scaling slope for values of b near 1, possibly due to using the same scaling region, 10^{-14} to 10^{-5}, for all values of b. Since the scaling formula is an asymptotic one, valid as $\delta \to 0$, the agreement in Fig. 5 could be improved by more careful monitoring of the scaling region, and in particular moving the scaling region lower, as is evident from Fig. 4(b). In this example, the scaling region was kept constant for the sake of consistency and clarity.