Chaos and experimental resolution

P.-M. Binder* and Milena C. Cuéllar
Departamento de Física, Universidad de Los Andes, Apartado Aéreo 4976, Bogotá, Colombia

(Received 27 October 1999)

We study systematically what levels of finite-resolution in measurements significantly affect the calculation of time delays, embedding dimensions, and Lyapunov exponent for two well-known chaotic systems. We find a tradeoff between the information contained in the measured time series and what is lost. Moreover, the ''noise'' series is low-dimensional and highly correlated.

PACS number(s): 05.45.Tp, 06.20.Dk, 07.05.Kf

INTRODUCTION

A finite-resolution (coarse) measurement of a chaotic dynamical system leads to limited prediction of its future through the butterfly effect [1,2]. In this paper we explore systematically the effects of coarse measurements in the time-series analysis [3–6] of chaotic systems. Such measurements can originate in systems with an inherently discrete variable (e.g., population number), the counting of rare events, a continuous-variable measurement which is digitized by a computer, or a process in which interspike intervals [7] are discretized through the sampling process.

We study noiseless time series of the Lorenz [8] and Hénon [9] systems, measured with different degrees of resolution, as well as the measurement noise series (MNS), or difference between the original and measured series, which can be interpreted as lost information. Our main findings are as follows. (1) Coarse measurements affect the analysis of time series: we find thresholds for the appropriate estimation of time delays and embedding dimensions, crucial for attractor reconstruction and further analysis. Incorrect estimations of these quantities can propagate to the calculation of measures of chaos such as Lyapunov exponents. (2) The reconstruction of the MNS and the time dependence of its average mutual information (AMI) show that the MNS contains valuable low-dimensional information about the time series which is lost in the measurement process. Information about the dynamical system is split between the coarse measurement and the MNS, with more of it going to the former as the measurement becomes finer.

Our work spans the area between symbolic dynamics and time-series analysis of continuous-valued data. The former, obtainable through a suitable encoding of measured data [10], can be analyzed by methods reviewed in detail in Ref. [11]. We only know of one previous systematic study of resolution effects [12], very different from ours. Our work also differs from previous studies [13] of finite-state dynamical systems. We preserve the underlying real-number dynamics, and discretize only *the measurement* itself. Therefore, we do not run into the problem of artificially short limit cycles [13].

*Author to whom correspondence should be addressed. Electronic address: p@faoa.uniandes.edu.co

THE DATA

For the Hénon map we used 9800 points, and for the Lorenz system, we generated 33 760 points for the variable x with a fourth-order Runge-Kutta method and time step $h = 0.01$ (about 400 maxima). Neither series has transients. In both cases we normalized the series between 0 and 1, and rounded it off to the nearest (below) multiple of the discretization step $\delta = 2^{-m}$, with m integer [14]. We denote the MNS with a Δ symbol preceding the variable name. In Fig. 1 we show both a discretized and a MNS series for the Lorenz map and $m = 4$.

RESEARCH PROTOCOL

We used the csp package (chaotic signal processor) [4,15] for our time-series analysis. The initial step is finding an optimal delay time for reconstructing trajectories; we used the first minimum of AMI [16] vs time. Too short a delay will not allow the variables to decorrelate, resulting in a flattened attractor, while too long a time will produce a random-looking object. The first minimum of AMI corresponds roughly to a time sufficient for a small region of the attractor to stretch, but not long enough to fold.

With the optimum delay time τ, one can construct d_E-dimensional vectors $(x(t), x(t+\tau), \ldots, x(t+(d_E - 1)\tau))$. The optimal embedding dimension d_E [17] needs to be determined, for example, with the calculation of false

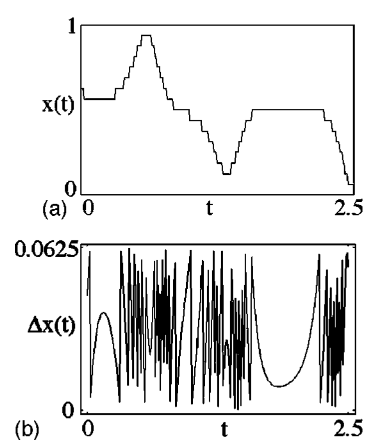

FIG. 1. (a) Discretized time series, and (b) MNS time series, for the Lorenz map with discretization 2^{-4}. Δx is the difference between the discretized and the original series.

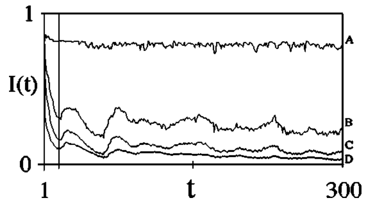

FIG. 2. Average mutual information $[I(t)]$ for discretized series as function of time (t). A, $m=4$; B, $m=6$; C, $m=7$; D, full machine precision.

nearest neighbors [18]. Projections of the attractor onto lower dimensions than d_E will result in points that appear very close, when in reality they are not. The fraction of false nearest neighbors decreases as one embeds the attractor in larger and larger dimensions, thus unfolding the attractor; d_E has been reached when the fraction goes to zero.

IDENTIFYING CHAOS IN MEASURED SERIES

We describe the results obtained for the Lorenz variable x, for $1 \leqslant m \leqslant 20$. In Fig. 2 we show the AMI, $I(t)$ vs time, for several values of m. Identification of the first minimum is quite difficult for small values of m, as shown in the figure. The wrong choice can affect measures of chaos (e.g., exponents) that are calculated based on this measure. We see that already for $m=7$ the AMI curve is starting to resemble the AMI curve obtained with full machine precision. In fact, for $m \geqslant 6$ we start identifying the correct minimum, $\tau \sim 16-17$, given in the figure by a vertical line. Since AMI is a function of the form $I \sim \Sigma p \log p$, we expect that with coarser resolution it will yield *larger* values. This is observed in Fig. 2, and confirmed in our studies of the Hénon map.

In Fig. 3 we show $P(d)$, the percentage of false nearest neighbors for reconstructions of the attractor in d dimensions, and several values of m. We see that for $m=5,6$ embedding dimensions of 1 and 2 are (incorrectly) identified. For $m \geqslant 7$ the correct result, $d_E=3$ is given; we show $m=15$. We note that $P(d)$ does not start to increase again for large dimension $d \sim 8$, which would be a typical signature of a system with noise. This suggests further study of the nature of the MNS, which we report below. The results for small m can be seen as an artifact of projecting a set of lattice points onto a lower dimension. With the Hénon map similar results are obtained. For intermediate values of m, $P(d)$ increases slightly above zero for $5 \leqslant d \leqslant 8$. The effect disappears with increasing m and d. We have no explanation for this fact. Finally, we have studied resolution effects in a direct test for

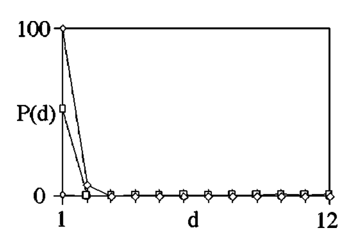

FIG. 3. Percentage of false near neighbors $P(d)$ as function of dimension d for discretized series. Circles, $m=5$; squares, $m=6$; diamonds, $m=15$. For $d \geqslant 2$ the circles are hidden behind the other two symbols.

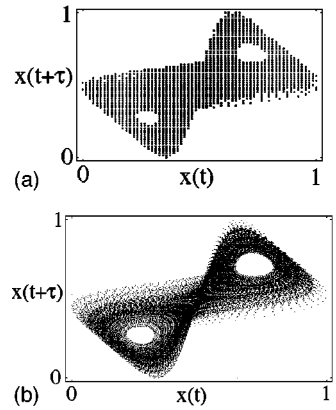

FIG. 4. Two-dimensional reconstruction of time series, with discretization (a) $m=6$ and (b) $m=15$.

determinism which is independent of $P(d)$. We have found results that are consistent with the findings of this paragraph [19].

THE MEASUREMENT NOISE SERIES

We have verified that the MNS is not ordinary noise. Figures 4 and 5 show, respectively, reconstructions of the measured series and Δx. In each case the first minimum of $I(t)$ has been used for the reconstruction. We see that as one series of data gains resolution and information, the other loses it. Figure 5 is particularly telling: it demonstrates that the information thrown away in an imperfect measurement is highly correlated, and can contain valuable information about the dynamical system. (For larger m, not shown, the reconstructions rapidly lose structure.) For $m=1$, a symbolic dynamics which codes turns around each half of the attractor with ones and zeros, respectively, requires both the discretized series and the MNS. The former provides information about which lobe of the Lorenz attractor is being orbited, and the latter about the number of turns; see also Fig. 1 for $m=4$. Details of this work will be presented elsewhere [20].

RESOLUTION EFFECTS ON LYAPUNOV EXPONENTS

We studied the Lyapunov exponents obtained from the Lorenz x series with varying m. We fixed $d_E=3$, the correct result, to guarantee the appropriate Lyapunov spectrum, regardless of the results given by the package. We used both the value of τ obtained from each AMI plot, and the correct result ($\tau \sim 16-17$). We show the latter in Fig. 6 as a function of m. The horizontal line is the result obtained with full machine precision. Region A corresponds to negative expo-

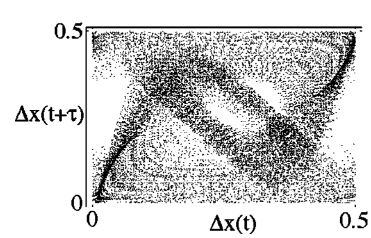

FIG. 5. Two-dimensional reconstruction of measurement noise series (MNS), Δx for $m=1$.

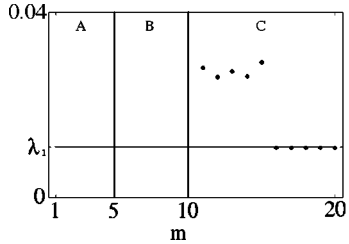

FIG. 6. Largest Lyapunov exponent vs resolution (m) for the Lorenz system. Horizontal line: result for undiscretized series. Region A: the largest exponent is incorrectly calculated to be negative. region B: the largest exponent is 7–300 times larger than the correct result.

nents, and region B to results 7–300 times larger than the correct exponent; these are due to the coarseness of the measurement. For $11 < m < 16$ the results become better, but can be quite sensitive to the choice of τ (for example, for $m = 13$, using the package-supplied time delay results in a large negative exponent). Only for $m \geq 16$ the results are within 1% of the undiscretized time series. Preliminary results [20] obtained with a series of 3000 points of the Lorenz system indicate that the effects of finite resolution on both the identification and characterization of chaos worsen for shorter series [21].

DISCUSSION

In this paper we have explored the effects of finite-resolution measurements on the identification and quantification of chaos from time series. The consistent use of a commercial package (csp) which yields reasonable results for data with machine precision lends support to our conclusions. Based on Fig. 2, we find that the information lost by these effects cannot be considered as high-dimensional noise. This observation sheds new light on the interaction between signal and noise that takes place in a classical measurement: while noise intrinsic to the system (e.g., thermal) or resulting from nonsystematic measurement errors [22] is often high dimensional and can be reduced, resolution error is of the same dimensionality (see Fig. 5) as the original data, and therefore the measured data cannot be improved by standard noise-reduction techniques [3–6]. The information contained in the measurement ''noise'' is simply lost. Note that this mechanism is distinct from the filtering of small, *high-dimensional* components of a signal through finite-resolution measurements discussed in Abarbanel's book [3], and in no way contradicts it. Compare our results, however, with the discussion of resolution noise on page 55 of Ref. [6].

Perhaps not many physical observations are subject to the severe discretization, $m \leq 7$, for which we observe the worst problems; the counting of rare events is the most likely candidate. However, our results are relevant to a recent report [23] that so far no unambiguous observations of chaos in wild animal populations have been made: this may be caused in part by the limited resolution (± 1 individual) of the observations. Moreover, our results also bring out the difficulty of identifying chaotic behavior in few-agent (typically 100) simulations of social systems [24] of recent interest to physicists, as well as in traffic-flow and stock market time series.

ACKNOWLEDGMENTS

This work was partly supported by Colciencias (Contract No. CT-259-96). We thank D. H. Campos and D. Laverde for useful discussions, and T. Dittrich for comments on the manuscript.

[1] E.N. Lorenz, *The Essence of Chaos* (Washington University Press, Seattle, 1993), pp. 181–184.

[2] S.H. Kellert, *In the Wake of Chaos* (University of Chicago Press, Chicago, 1994), pp. 29–48.

[3] *Nonlinear Modeling and Forecasting*, edited by M. Casdagli and S. Eubank (Addison-Wesley, Reading, MA, 1992); *Time Series Prediction*, edited by A.S. Weigend and N.A. Gershenfeld (Addison-Wesley, Reading, MA, 1993).

[4] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, and L.Sh. Tsimring, Rev. Mod. Phys. **65**, 1331 (1993); H.D.I. Abarbanel, *Analysis of Observed Chaotic Data* (Springer, New York, 1996).

[5] *Coping with Chaos*, edited by E. Ott, T. Sauer, and J.A. Yorke (Wiley, New York, 1994).

[6] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, U.K., 1999).

[7] T. Sauer, Phys. Rev. Lett. **72**, 3811 (1994); R. Castro and T. Sauer, Phys. Rev. E **59**, 2911 (1999). Numerical integration can introduce finite resolution effects for interspike reconstructions through the finite integration step.

[8] E.N. Lorenz, J. Atmos. Sci. **20**, 130 (1963). We use $r = 45.92$, $b = 4$, and $\sigma = 16$.

[9] M. Hénon, Commun. Math. Phys. **50**, 69 (1976). We use $a = 1.4$, $b = 0.3$.

[10] V.M. Alekseev and M.V. Yakobson, Phys. Rep. **75**, 287 (1981); A proper encoding is often difficult to obtain: M. Lehrman, A.B. Rechester, and R.B. White, Phys. Rev. Lett. **78**, 54 (1997).

[11] R. Badii and A. Politi, *Complexity: Hierarchical Structure and Scaling in Physics* (Cambridge University Press, Cambridge, U.K., 1997).

[12] C. Schittenkopf and G. Deco, Physica D **94**, 57 (1996); G. Deco, C. Schittenkopf, and B. Schürmann, Int. J. Bifurcation Chaos Appl. Sci. Eng. **7**, 97 (1997). The authors study the information flow step by step, and find that information about initial conditions is lost linearly in time for infinite symbols.

[13] P.-M. Binder, Physica D **57**, 31 (1992); P. Diamond, P. Kloeden, A. Pokrovskii, and A. Vladimirov, *ibid.* **86**, 559 (1995); P.-M. Binder and J.C. Idrobo, Phys. Rev. E **58**, 7987 (1998), and references therein.

[14] Previous studies indicate that other discretization schemes do not change significantly the results: see P.-M. Binder, Comput. Math. Appl. **21**, 133 (1991).

[15] *Tools for Dynamics: csp for Unix* (Randle, Inc., Great Falls, VA, 1997); http://www.cquest.com/chaos.html.

[16] A.M. Fraser and H.L. Swinney, Phys. Rev. A **33**, 1134 (1986).

[17] M. Casdagli, T. Sauer, and J.A. Yorke, J. Stat. Phys. **65**, 579 (1991).

[18] A. Cenys and K. Pyragas, Phys. Lett. A **129**, 227 (1988); Th. Buzug, T. Reimers, and G. Pfitzer, Europhys. Lett. **13**, 605 (1990); M.B. Kennel, R. Brown, and H.D.I. Abarbanel, Phys. Rev. A **45**, 3403 (1992); H.D.I. Abarbanel and M.B. Kennel, Phys. Rev. E **47**, 3057 (1994); see, however, D.R. Fredkin and J.A. Rice, *ibid.* **51**, 2950 (1995).

[19] D.T. Kaplan and L. Glass, Phys. Rev. Lett. **68**, 427 (1992). We find that for very small resolutions, $m \sim 4$, the average directional vectors are very close to zero, so the result is ambiguous. For $m \sim 10$ the flow vectors do not cross, although they show lower resolution than in the continuous case.

[20] M.C. Cuéllar and P.-M. Binder (unpublished).

[21] This question follows: Can a very long series of low-resolution data (e.g., symbolic dynamics) make up for the information lost in the MNS? So far [11], we are far from achieving this: even quantities like τ and d_E cannot be calculated from symbolic dynamics data yet.

[22] See, for example, V. Thomsen, Phys. Teach. **35**, 15 (1997).

[23] S. Ellner and P. Turchin, Am. Nat. **145**, 343 (1995); C. Zimmer, Science **284**, 83 (1999). See also R.F. Costantino, R.A. Desharnais, and B. Dennis, *ibid.* **275**, 389 (1997), for data on confined populations.

[24] R. Savit, R. Manuca, and R. Riolo, Phys. Rev. Lett. **82**, 2203 (1999); W.B. Arthur, Science **284**, 107 (1999); N.F. Johnson, P.M. Hui, and T.S. Lo, Phys. Rev. Lett. **82**, 3360 (1999).