Detecting Unstable Periodic Orbits in Chaotic Experimental Data

Paul So,[1,2] Edward Ott,[2,*] Steven J. Schiff,[1] Daniel T. Kaplan,[3] Tim Sauer,[4] and Celso Grebogi,[2,†]

[1]*Department of Neurosurgery, Children's National Medical Center and the George Washington University, NW, Washington, D.C. 20010*

[2]*Institute for Plasma Research, University of Maryland, College Park, Maryland 20742*

[3]*Center for Nonlinear Dynamics, McGill University, Montreal, Quebec, Canada H3G 1Y6*

[4]*Department of Mathematics, the George Mason University, Fairfax, Virginia 22030*

(Received 8 December 1995)

A new method is proposed for detecting unstable periodic orbits and their linear stability properties from chaotic experimental time series. Illustrative examples are presented for both numerically and experimentally generated time series. The statistical significance of the results is assessed using surrogate data. [S0031-9007(96)00488-7]

PACS numbers: 05.45.+b

Unstable periodic orbits embedded in chaotic attractors are fundamental to an understanding of chaotic dynamics. For example, basic ergodic properties such as dimension, Lyapunov exponents, and topological entropy can be determined from periodic orbits. Moreover, the detection of such an orbit from experimental data is a test for the presence of determinism. A particularly important application is in the control of chaotic systems [1] where the first essential step is often the determination of periodic orbits. For these reasons, detection of periodic orbits in experimental data has become a central issue [2–4].

In this paper, we introduce new techniques for addressing this problem. Our method utilizes a transformation of the experimental time series data, such that the transformed series is concentrated on the periodic orbits. Histograms of the transformed data thus have sharp peaks at the locations of periodic orbits which can then be read off. The reliability of the method can be objectively assessed by testing the statistical significance of these peaks against surrogate data that is random but preserves statistical properties of the original data [5].

Although higher-dimensional cases are of most interest, in order to get the ideas of our method across with the least technical complication, we first discuss in some detail the case of a one-dimensional map. We also limit the discussion to period one orbits (fixed points) [6,7]. Following our one-dimensional map discussion, we then state a result for higher dimensions (details appear in Ref. [7]). Using data generated from a noisy two-dimensional map, we demonstrate the robustness of our method. Then we apply it to data collected from an elastic mechanical experiment [8].

To begin, assume we are given a finite length time series from a one-dimensional map $f(x)$, and we desire to estimate the locations of the fixed points $x^* = f(x^*)$. Consider the transformation

$$\hat{x}_n = [x_{n+1} - s_n(k)x_n]/[1 - s_n(k)], \quad (1)$$

where [9]

$$s_n(k) = (x_{n+2} - x_{n+1})/(x_{n+1} - x_n) + k(x_{n+1} - x_n) \quad (2)$$

[see Fig. 1 for a geometrical interpretation of (1) and (2) for the case $k = 0$]. In the case where $k = 0$ and $f(x)$ is a linear function $f(x) = x^* + \alpha(x - x^*)$, we have $s_n(0) \equiv \alpha$ and $\hat{x}_n \equiv x^*$ independent of n. Thus, in this case, all the data are transformed to the fixed point.

In the case of a general nonlinear $f(x)$ and $k \neq 0$, all points that lie in the linear region of the fixed point x^* will be transformed to the vicinity near x^*. In particular, as shown below, the density function for $\hat{x}$, denoted $\hat{\rho}(\hat{x})$, has inverse square root type singularities at the fixed points $\hat{\rho}(\hat{x}) \sim |\hat{x} - x^*|^{-1/2}$. Thus if we plot a histogram approximation to $\hat{\rho}(\hat{x})$ using a finite amount of data, then there will be a sharp peak at $\hat{x} = x^*$. This is potentially a way of estimating the fixed points. By construction, the transformation in Eqs. (1) and (2) utilizes all appropriate points in the linear region of a fixed point to form the singularity. This is in contrast to typical "recurrence" based methods (see Refs. [2–4]), in which one often has to determine the appropriate ball size in phase space to decide for close encounters. In practice, the degree of clustering around the fixed point in our method depends

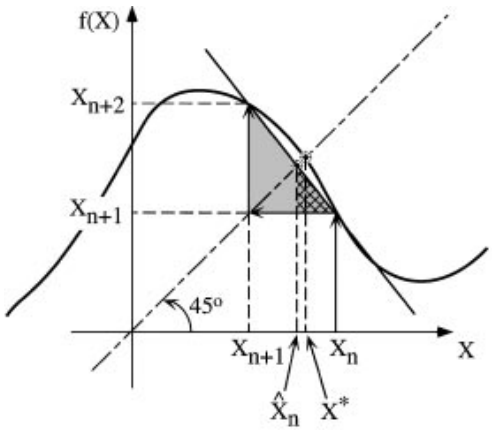

FIG. 1. Geometric interpretation of Eqs. (1) and (2) with $k = 0$. Given a sequence of points $\{x_{n+2}, x_{n+1}, x_n\}$, Eq. (2) with $k = 0$ gives the slope of the hypotenuse of the shaded triangle while Eq. (1) is the construction of the estimated fixed point $\hat{x}_n$ using the smaller cross-hatched triangle.

on the size of the linear region and on how often a typical trajectory visits the linear region [10].

To demonstrate the inverse square root singularities, we write the transformation from x to $\hat{x}$ as

$$\hat{x} = g(x,k) = [f(x) - s(x,k)x]/[1 - s(x,k)], \quad (3)$$

where $s(x,k) = [f(f(x)) - f(x)]/[f(x) - x] + k[f(x) - x]$. Note that $g(x^*,k) = x^*$ *independent* of k. If $\rho(x)$ denotes the distribution function for x, then $\hat{\rho}(\hat{x}) = \sum_{i=1}^{N} \rho(x_i(\hat{x}))|g'(x)|^{-1}_{x=x_i(\hat{x})}$, where $x_i(\hat{x}), i = 1,\ldots,N$, denote the values of x satisfying $\hat{x} = g(x)$ and $g'(x) \equiv dg/dx$. We see that $\hat{\rho}(\hat{x})$ will in principle be singular at zeros of g' and at singularities of ρ. Differentiating $g(x,k)$ with respect to x, one finds that $g' = 0$ at x values for which $f(x) = x$ (i.e., at fixed points). Thus near $x = x^*$ we have by Taylor expansion that $\hat{x} = g(x,k) \cong x^* + \beta(x - x^*)^2$ and $|g'|^{-1} \sim |x - x^*|^{-1} \sim |\hat{x} - x^*|^{-1/2}$; i.e., $\hat{\rho}(\hat{x}) \sim |\hat{x} - x^*|^{-1/2}$. Thus $\hat{\rho}(\hat{x})$ is singular at $\hat{x} = x^*$.

However, we do not wish to mistake other potential singularities not due to fixed points with true fixed points. These spurious peaks result from singularities in $\rho(x)$ and from zeros of $g'(x)$ that occur at x values not at fixed points. To distinguish the spurious peaks, we note that their locations depend on the parameter k. One method is to randomly pick many different k values for each x value and form the average $\langle\hat{\rho}(\hat{x})\rangle$ of the resulting distributions $\hat{\rho}(\hat{x})$. Since the spurious peaks occur at different $\hat{x}$ values for each k, the spurious peaks will be eliminated from the average $\langle\hat{\rho}(\hat{x})\rangle$ by smearing. Only the true peaks will remain sharply defined.

Figure 2(a) shows $\hat{\rho}(\hat{x})$ obtained from 5000 iterates of the logistic map $f(x) = rx(1 - x)$, with $r = 3.92$ and $k = 0$. There are four sharp peaks, one at the true fixed point on the attractor [$x = (r - 1)/r \cong 0.745$], two at strong singularities of $\rho(x)$ located at the first and second iterates of the critical point [$\hat{x}(f(1/2)) = 0.241$ and $\hat{x}(f^2(1/2)) = -0.0548$], and one at a zero of g' that is not due to a fixed point ($\hat{x} = 0.627$). Figure 2(b) shows the average $\langle\hat{\rho}(\hat{x})\rangle$ calculated using 500 randomly chosen values of k for each data point x. To generate these random k values, we set $k = \kappa R$ with R chosen randomly in $[-1,1]$ and $\kappa = 5$. Since for each data point we have 500 different calculated $\hat{x}$ corresponding to the 500 random k values, we were able to reduce the number of data points used from 5000 to 100 without degrading the counting statistic in the histogram approximation of the average $\langle\hat{\rho}(\hat{x})\rangle$. In Fig. 2(b), all spurious peaks are absent and only the true fixed point peak at $\hat{x} = 0.745$ remains.

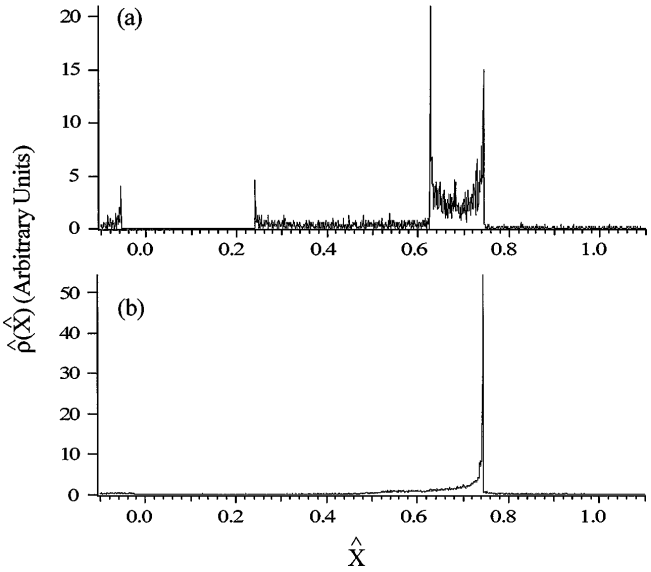

FIG. 2. (a) A histogram plot of $\hat{\rho}(\hat{x})$ for the map $f(x) = 3.92x(1 - x)$. 5000 data points were used and $k = 0$. (b) A histogram plot of $\langle\hat{\rho}(\hat{x})\rangle$ averaged over 500 values of k, where $k = 5R, R \in [-1,1]$. The number of data points used in (b) was 100.

The previous discussion on one-dimensional systems can be naturally extended to systems in higher dimensions (see Ref. [7]). Here, we state a result and give an algorithm for the construction of $\langle\hat{\rho}(\hat{x})\rangle$ in systems with arbitrary dimensions, for the example where the data are delay coordinate vectors $\{\mathbf{z}_n\}$ reconstructed from a scalar time series $\{x_n\}$ (see Ref. [11]). With a properly chosen embedding dimension d, $\mathbf{z}_n = (z_n^1, z_n^2, \ldots, z_n^d)^\dagger \equiv (x_n, x_{n-1}, \ldots, x_{n-d+1})^\dagger$. Here, $\mathbf{z}_n$ is a column vector and $\mathbf{z}^\dagger$ is the transpose of $\mathbf{z}$. With this notation, the transformation from $\mathbf{z}$ to $\hat{\mathbf{z}}$ is given by $\hat{\mathbf{z}}_n = (\mathbf{1} - \mathbf{S}_n)^{-1}(\mathbf{z}_{n+1} - \mathbf{S}_n\mathbf{z}_n)$, where

$$\mathbf{S}_n = \begin{pmatrix} a_n^1 & a_n^2 & \cdots & a_n^{(d-1)} & a_n^d \\ & & \mathbf{1} & & \mathbf{0} \end{pmatrix} + \kappa\mathbf{R}\|\mathbf{z}_{n+1} - \mathbf{z}_n\|, \quad \begin{pmatrix} a_n^1 \\ \vdots \\ a_n^d \end{pmatrix} = \begin{pmatrix} (\mathbf{z}_n - \mathbf{z}_{n-1})^\dagger \\ \vdots \\ (\mathbf{z}_{n-(d-1)} - \mathbf{z}_{n-d})^\dagger \end{pmatrix}^{-1} \begin{pmatrix} z_{n+1}^1 - z_n^1 \\ \vdots \\ z_{n-(d-2)}^1 - z_{n-(d-1)}^1 \end{pmatrix}, \quad (4)$$

κ is the magnitude of the randomization and $\mathbf{R}$ is a $d \times d$ random matrix with each element chosen independently with uniform distribution in $[-1,1]$ (other choices could be used). The norm in Eq. (4) is chosen to be the L_1 norm, $\|\mathbf{z}\| = \sum_{i=1}^{d} |z^i|$.

To find the possible fixed points from our data, we again look for peaks in $\langle\hat{\rho}(\hat{z})\rangle$. In addition, for each identified peak $\hat{z}_0$, we can simply obtain and estimate the local Jacobian matrix by collecting the $\mathbf{z}_n$ which give values of $\hat{z}$ in the cluster near $\hat{z}_0$, and then averaging the corresponding $\mathbf{S}_n$ with $\kappa = 0$.

Although ideally one should expect inverse square root singularities in $\hat{z}$ at the locations of the fixed points, in a real experimental setting, these singularities are blurred by small noise into maxima, and can even be completely washed out by large noise. We can use the technique of

surrogate data [5] to assess the reliability of the observed peaks. That is, we can produce a truly stochastic time series (surrogate data) with similar statistical properties to the original (supposedly deterministic) data, and then compare the results of applying our periodic orbit detection method to the original data set and to the surrogate. Since the surrogate data are random, we do not expect any nonlinear fixed point structure that may be present in the original data to survive. Most importantly, using many different realizations of the surrogate data, we can estimate the statistical probability that the observed peaks in our experimental $\langle\hat{\rho}(\hat{z})\rangle$ could arise from a linear stochastic process modeled by the surrogates. Numerically, for each realization of the surrogate data, we apply the same procedure for calculating $\langle\hat{\rho}_{\rm sur}(\hat{z})\rangle$ as for our experimental data. A mean value $\overline{\rho}_{\rm sur}(\hat{z})$ can be estimated from this collection of $\{\langle\hat{\rho}_{\rm sur}(\hat{z})\rangle\}$. Similar to our noisy experimental data, $\langle\hat{\rho}_{\rm sur}(\hat{z})\rangle$ from each individual realization of the surrogates will fluctuate, and will consequently have fluctuation peaks which deviate from the mean $\overline{\rho}_{\rm sur}(\hat{z})$. Denote the deviation from the mean for a given surrogate by $w(\hat{z}) = \langle\hat{\rho}_{\rm sur}(\hat{z})\rangle - \overline{\rho}_{\rm sur}(\hat{z})$ and let $W = \max_{\hat{z}}[w(\hat{z})]$. Using many surrogates we can determine the fraction $\Xi(W')$ of surrogates with maximum deviations W exceeding W'. This gives an estimate of the probability that W exceeds W' for any chosen surrogate.

To demonstrate the robustness of our method in a noisy situation, we use the Ikeda map with additive observational noise [12]. The Ikeda map describes the dynamics of a nonlinear optical cavity and is given by the two-dimensional map $(u_{n+1} = 1.0 + a(u_n \cos t_n - v_n \sin t_n), v_{n+1} = a(u_n \sin t_n + v_n \cos t_n))$, where $t_n = 0.4 - b/(1 + u_n^2 + v_n^2)$. With $a = 0.9$ and $b = 6.0$, this system has a chaotic attractor with a Lyapunov dimension of approximately 1.71, and it has an unstable fixed point at $(u^*, v^*) = (0.53, 0.25)$. To generate the noisy time series for our numerical experiment, we choose $o_n = u_n + \epsilon\delta_n$ to be our scalar output. The delay coordinate vector in two dimensions is then given by $\mathbf{z}_n \equiv (o_n, o_{n-1})^\dagger$. Here, ϵ is the magnitude of the external noise and $\{\delta_n\}$ is a uniformly distributed random variable in $[-1, 1]$.

Figure 3 corresponds to our results with three different noise levels: (a) $\epsilon = 0$; (b) $\epsilon \sim 15\%$ of the radius of the attractor; (c) $\epsilon \sim 50\%$ of the radius of the attractor. The solid curves give the histogram approximations to $\langle\hat{\rho}(\hat{z})\rangle$ and the dashed curves are the corresponding surrogate averages $\overline{\rho}_{\rm sur}(\hat{\hat{z}})$. To quantify the statistical significance of the deviations between the data and the surrogate mean, we plotted the distribution $\Xi(W)$ in Figs. 3(d), 3(e), and 3(f) with the arrows indicating the locations of the observed maximum deviations. In all these calculations, the time series data were delay embedded in a two-dimensional space, 20 surrogates were used and in the process of randomization, we have used 10^4 different random matrices for each data point [13] and κ was 5 [14]. In Fig. 3(a), the fixed point at $\hat{z} = 0.53$ has a strong peak rising sharply above the surrogates mean. As indicated by the arrow in Fig. 3(d), the probability for observing such a large deviation ($W \sim 20$)

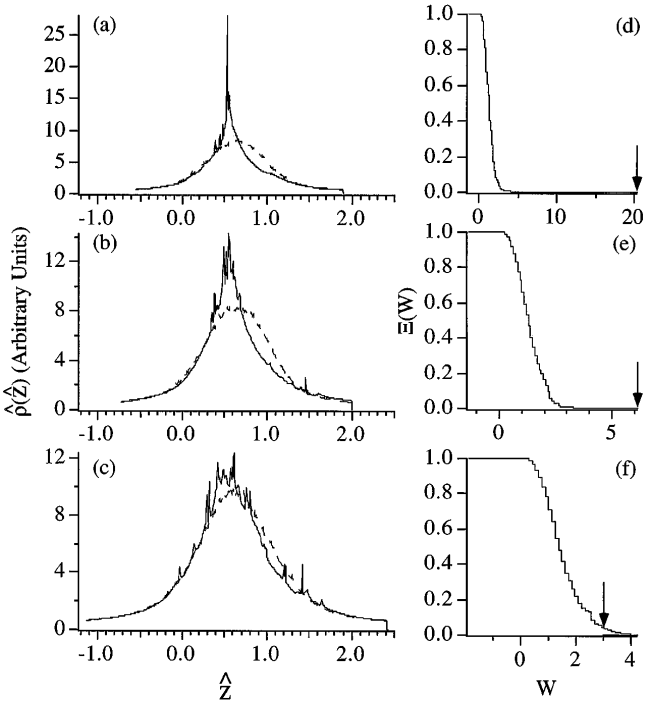

FIG. 3. Histogram plots of $\langle\hat{\rho}(\hat{z})\rangle$ (solid curves) for the Ikeda map: (a) $\epsilon = 0$; (b) $\epsilon = 0.15\times$ (radius of attractor); and (c) $\epsilon = 0.5\times$ (radius of attractor). The dashed curves are the corresponding $\overline{\rho}_{\rm sur}(\hat{z})$. 1024 data points were used, 10^4 different random matrices were used for each data point, and $\kappa = 5$. (d), (e), and (f) are histogram plots of $\Xi(W)$ in the three cases. Arrows indicate the values of maximum deviations between $\langle\hat{\rho}(\hat{z})\rangle$ and $\overline{\rho}_{\rm sur}(\hat{z})$.

from the 20 surrogates is unobservably small ($\ll 10^{-5}$). As the amount of noise increases [Figs. 3(b) and 3(e)], the peak at the fixed point broadens. However, its deviation above the surrogate mean ($W \sim 6$) is still significant. The probability of finding a peak with the same large deviation in the surrogates is still unobservably small. In the last case with 50% noise [Figs. 3(c) and 3(f)], the distinctiveness of the peak in the solid curve disappears and the maximum deviation observed in the data has an approximately 5% probability to be found in the surrogates.

We also applied our method to an experimental time series from a periodically driven, gravitationally buckled, amorphous magnetoelastic ribbon. The 10^3 data points used are the measured curvature of the ribbon at its base sampled at the period of the drive (see Ditto, Rauseo, and Spano, Ref. [15], for details on this experiment). The inset of Fig. 4(a) is a section of the time series used. The dimension of the experimental attractor is approximately 1.2 but in order to unfold all the crossings, the data have to be embedded in a three-dimensional delay space. In the process of randomization, we again used 500 different random matrices for each data point and κ was 15 [14]. The histogram for $\langle\hat{\rho}(\hat{z})\rangle$ is plotted in Fig. 4(a) as a solid curve and the surrogate mean $\overline{\rho}_{\rm sur}(\hat{z})$ is plotted as a dashed curve. The dominant fixed point is located at approximately 5.40 [16]. Again, using 20 surrogates, we calculated the distribution $\Xi(W)$, and we

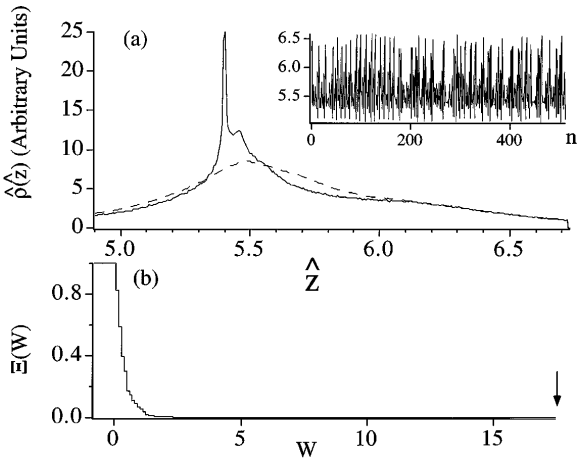

FIG. 4. (a) A histogram plot of $\langle\hat{\rho}(\hat{z})\rangle$ (solid curve) for magnetoelastic ribbon data. 1024 data points were used, 10^4 different random matrices were used for each data point, and $\kappa = 15$. The dashed curve corresponds to $\overline{\rho}_{\text{sur}}(\hat{z})$. Inset is a section of the actual time series data; (b) is a histogram plot of $\Xi(W)$. Arrow indicates the value of maximum deviation between $\langle\hat{\rho}(\hat{z})\rangle$ and $\overline{\rho}_{\text{sur}}(\hat{z})$.

plotted its approximation in Fig. 4(b). The arrow indicates the location of the maximum deviation calculated at the fixed point and the probability for observing this value in the surrogates is again unobservably small. At this fixed point, the eigenvalues of the estimated Jacobian matrix (average of $\mathbf{S}_n$) gives an expansion rate (Lyapunov number) of $8.51(\pm 3.47)$ per time unit and a contraction rate (Lyapunov number) of $0.749(\pm 0.118)$ per time unit.

In this Letter, we present a statistical method to detect unstable periodic orbits from a chaotic data set. The method utilizes the linear dynamics around an unstable periodic point to produce a statistical measure which, in theory, is singular at the periodic point. By construction, all points that lie within the linear regions of the periodic orbits are utilized. There is no need to search for an optimal neighborhood size as in other "recurrence" methods. Using this method, unstable fixed points were reliably identified in noisy numerically generated data as well as in real experimental data from a magnetoelastic ribbon system.

We thank B. Gluckman and M. Spano for the ribbon data, and D. Scott for implemening the Pierson-Moss method. We thank the Keck Foundation for financial support for computing facilities. P. S., S. J. S., and E. O. received support from ONR (N00014-95-1-0138), S. J. S. received support from NIH (1-R29-MH50006-04) and the Children's Research Institute, and C. G. was supported by DOE.

*Department of Physics and Electrical Engineering and Institute for Systems Research.

†Department of Mathematics and Institute for Physical Science and Technology.

[1] E. Ott, C. Grebogi, and J. A. Yorke, Phys. Rev. Lett. **64**, 1196 (1990); F. Romeiras *et al.*, Physica (Amsterdam) **58D**, 165 (1992). For reviews, see T. Shinbrot *et al.* [Nature (London) **365**, 411 (1993)] and E. Ott and M. L. Spano [Phys. Today **48**, No. 5, 34 (1995)].

[2] D. P. Lathrop and E. J. Kostelich, Phys. Rev. A **40**, 4028 (1989); L. Flepp *et al.*, Phys. Rev. Lett. **67**, 2744 (1991); J. C. Sommerer *et al.*, Phys. Lett. A **153**, 105 (1991); see references in T. Shinbrot *et al.* (Ref. [1]).

[3] Recently D. Pierson and F. Moss [Phys. Rev. Lett. **75**, 2124 (1995)] and X. Pei and F. Moss [Nature (London) **379**, 619 (1996)] have introduced a method for locating fixed points in data. This method is based on recurrence together with criteria derived from the local behavior near a fixed point.

[4] D. Christini and J. J. Collins, Phys. Rev. Lett. **75**, 2782 (1995).

[5] J. Theiler *et al.*, Physica (Amsterdam) **58D**, 77 (1992). In our examples, we used the *Gaussian Scaled Phase Shuffle* (or the *Amplitude Adjusted Fourier Transform*) method. It gives a stochastic surrogate with the same amplitude distribution and an approximated power spectrum of the original data.

[6] One way to obtain period p periodic orbits is to consider fixed points of the pth iterate of the map. Another method (less sensitive to noise) which is an extension of the fixed point method but utilizes all iterates of the data set is discussed in [7].

[7] P. So *et al.* (to be published).

[8] The application of this method to data collected from a neuronal ensemble will appear in S. Schiff *et al.* (to be published).

[9] Other choices in place of $k(x_n - x_{n-1})$ are discussed in Ref. [7]. The main requirement for this term is that it vanishes at the fixed point.

[10] Since the method samples the linear dynamics near a fixed point, it is also possible that fixed points that are near but not actually on the attractor might show up in $\hat{\rho}(\hat{x})$. An example of these fixed points can be found in systems near a crisis [see J. C. Sommerer *et al.* (Ref. [2])].

[11] F. Takens, in *Dynamical Systems and Turbulence,* edited by D. Rand and L. S. Young (Springer-Verlag, Berlin, 1981), p. 230; T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991); E. Ott, T. Sauer, and J. A. Yorke, *Coping with Chaos* (John Wiley, New York, 1994).

[12] Using dynamical noise instead of observational noise, we were able to pick out statistical significant fixed points in the Ikeda system with $a = 0.75, b = 9.0$, and a noise level of 20%.

[13] In the case with low noise, $\epsilon < 15\%$, 512 data points with 100 random matrices are enough to detect statistically significant peaks in the Ikeda system. A higher number of data points and random matrices are needed if the system is much noisier or if the periodic orbits are less frequently visited by a typical trajectory.

[14] It might be useful to choose the value of κ to optimize the statistical significance of the candidate peaks.

[15] W. Ditto, S. N. Rauseo, and M. L. Spano, Phys. Rev. Lett. **65**, 3211 (1990).

[16] There exists a statistically significant secondary peak in the experimental $\langle\hat{\rho}(\hat{z})\rangle$ indicating the possibility of a second fixed point in the ribbon data.