Extracting unstable periodic orbits from chaotic time series data

Paul So,[1,2] Edward Ott,[2,*] Tim Sauer,[3] Bruce J. Gluckman,[4,1] Celso Grebogi,[2,†] and Steven J. Schiff[1]
[1]*Center for Neuroscience, Children's Research Institute, 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]*Department of Mathematics, The George Mason University, Fairfax, Virginia 22030*
[4]*Naval Surface Warfare Center, Carderock Division, Bethesda, Maryland 20054-5000*
(Received 27 January 1997)

A general nonlinear method to extract unstable periodic orbits from chaotic time series is proposed. By utilizing the estimated local dynamics along a trajectory, we devise a transformation of the time series data such that the transformed data are concentrated on the periodic orbits. Thus, one can extract unstable periodic orbits from a chaotic time series by simply looking for peaks in a finite grid approximation of the distribution function of the transformed data. Our method is demonstrated using data from both numerical and experimental examples, including neuronal ensemble data from mammalian brain slices. The statistical significance of the results in the presence of noise is assessed using surrogate data. [S1063-651X(97)15205-9]

PACS number(s): 05.45.+b, 87.10.+e

I. INTRODUCTION

Unstable periodic orbits embedded in chaotic attractors are fundamental to an understanding of chaotic dynamics [1,2]. For example, basic ergodic properties such as dimension, Lyapunov exponents, and topological entropy can be determined from periodic orbits. Moreover, the detection of a periodic orbit from experimental data is a test for the presence of determinism. A particularly important application is in the control of chaotic systems [2,3] where the first essential step is often the determination of periodic orbits [4]. For these reasons, detection of periodic orbits in experimental data has become a central issue [4–9].

So *et al.* in Ref. [9] introduced a new method to address this problem. The basic strategy of their method was to transform the experimental time series data by using information of the local linear dynamics along a trajectory such that the transformed data in a suitable phase space (e.g., delay coordinate space) are concentrated on the periodic orbits [10]. Periodic orbits can then be extracted by looking for peaks in a finite grid approximation to the distribution function of the transformed data. In the implementation and examples of this periodic orbit detection method in Ref. [9], only a brief exposition of the theory limited to the problem of extracting period 1 orbits (fixed points) from chaotic time series was presented. In this paper, we elaborate further on the extraction of fixed points, providing detailed derivations and justifications for the method and presenting several variations on it. As in [9], we consider discrete time dynamical systems (maps). We also discuss the extraction of periodic orbits with period greater than 1 and present extensive illustrative examples.

The paper is organized as follows. In Sec. II, we introduce the transform for the case of fixed points, and we provide detailed analytical arguments for the existence of singularities at fixed points in the distribution function of the transformed data. We also discuss various ways of implementing the procedure, as well as the possibility of enhancing the accuracy by use of a nonlinear version of the transformation. Periodic orbits with periods $p>1$ can in principle be extracted by considering fixed points of the pth iterated dynamics, i.e., every pth iterate of the time series. However, to be effective, such a treatment of periodic orbits with $p>1$ requires a dynamical correlation among a relatively long stretch of data points. In the presence of noise or in systems with large Lyapunov exponents, this method would commonly fail. In Sec. III, we introduce a variant of our fixed point detection method suitable to periodic orbits of period greater than 1 such that the method utilizes every iterate of the time series. This method pieces together different *short* portions of the periodic orbits from *nonconsecutive* sections of the data. In Sec. IV, we examine the robustness of our method in the presence of both observational and dynamical noise. The reliability of the method can be objectively assessed by testing the statistical significance of the observed peaks against shuffled surrogate data which approximate the power spectra of the original data [11]. We demonstrate the method by using both numerical examples with added simulated noise and real experimental data collected from a magnetoelastic ribbon system. Lastly, in Sec. V, the possibility of using our method to track periodic orbits in systems with slowly varying parameters is discussed. Again, we use the magnetoelastic ribbon system with changing parameters as our first demonstration. Then, we apply this tracking method to the analysis of a sequence of interburst time intervals collected from a brain slice of the rat hippocampus. In the Appendix, a geometric interpretation and a supplemental analysis of our periodic orbit transform are provided in the special case when the dynamics is described by a one-dimensional map.

*Also at Department of Physics and Electrical Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742.
†Also at Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742.

II. EXTRACTING PERIOD 1 FIXED POINTS

To begin, assume we are given a finite length time series $\{x(n)\}$ from a discrete time dynamical system [13]. With a properly chosen embedding dimension d, one can uniquely represent the state of the system at time n by a delay coordinate vector $\mathbf{z}(n)=(z_1(n),z_2(n),\ldots,z_d(n))^\dagger\equiv(x(n),x(n-1),\ldots,x(n-d+1))^\dagger$ [14,15]. Here, $\mathbf{z}(n)$ is a column vector and $\mathbf{z}^\dagger$ is the transpose of $\mathbf{z}$. With this notation, the evolution of the system state from $\mathbf{z}(n)$ to $\mathbf{z}(n+1)$ can in general be expressed by a d-dimensional nonlinear map $\mathbf{F}$ in the following form:

$$\mathbf{z}(n+1)\equiv\begin{pmatrix} z_1(n+1)\\ z_2(n+1)\\ z_3(n+1)\\ \vdots\\ z_d(n+1)\end{pmatrix}=\mathbf{F}(\mathbf{z}(n))=\begin{pmatrix} F_1(\mathbf{z}(n))\\ F_2(\mathbf{z}(n))\\ F_3(\mathbf{z}(n))\\ \vdots\\ F_d(\mathbf{z}(n))\end{pmatrix}$$
$$=\begin{pmatrix} f(\mathbf{z}(n))\\ z_1(n)\\ z_2(n)\\ \vdots\\ z_{d-1}(n)\end{pmatrix}. \tag{1}$$

Then we ask the following question: without further *a priori* knowledge of the underlying map $\mathbf{F}$, how can we estimate the locations of unstable period p orbits, $\mathbf{z}^*=\mathbf{F}^p(\mathbf{z}^*)$, from the chaotic time series? In this section, we discuss detecting period 1 orbits ($p=1$), also called fixed points. We will extend our discussion to $p>1$ in the section that follows.

One simple approach, as used previously by others, for addressing this question is based on the idea of recurrence (see Refs. [4–8]). In these recurrence based methods, one looks for close encounters (within a ball B_ϵ of a given radius ϵ) of delay vectors reconstructed from time series. There are various implementations of this idea with different discriminating criteria based on the expected dynamical behaviors near fixed points. Fundamentally, the effectiveness of these recurrence based methods depends on the relative frequency that a typical chaotic trajectory visits the ball $B_\epsilon(\mathbf{z}^*)$ centered around the fixed point $\mathbf{z}^*$. Assuming that the chaotic attractor has a natural invariant measure μ, this relative frequency is given by the natural measure of the ball $\mu(B_\epsilon(\mathbf{z}^*))$, which can be small.

In the following, we first introduce our fixed point transform. Then, in the next subsection we describe its functional behaviors and its measure-enhancing effect. By concentrating the natural measure of a chaotic attractor around its fixed points, one can extract these fixed points by simply looking for peaks in a finite-grid approximation to the distribution function of the transformed data.

A. The fixed point transform

Consider a small neighborhood around a fixed point $\mathbf{z}^*\equiv(z^*,\ldots,z^*)^\dagger$. We can describe the local dynamics by a Taylor series expansion of $\mathbf{F}(\mathbf{z}^*)$ around $\mathbf{z}$,

$$\mathbf{F}(\mathbf{z})=\mathbf{z}^*+\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})\cdot(\mathbf{z}-\mathbf{z}^*)+O((\mathbf{z}-\mathbf{z}^*)^2), \tag{2}$$

where $\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})$ is the $d\times d$ Jacobian matrix of $\mathbf{F}(\mathbf{z})$. In index notation, the ijth component of $\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})$ is $\partial_j F_i\equiv\partial F_i/\partial z_j$ with $i,j=1,\ldots,d$ and we use $\boldsymbol{\nabla}\mathbf{F}\cdot(\mathbf{z}-\mathbf{z}^*)$ to denote the vector whose ith component is $\Sigma_j(\partial F_i/\partial z_j)(z_j-z_j^*)$. [Note that this notation is somewhat unconventional since this vector would usually be written $(\mathbf{z}-\mathbf{z}^*)^\dagger\cdot\boldsymbol{\nabla}\mathbf{F}$.] Conversely, one can solve for $\mathbf{z}^*$ in terms of $\mathbf{z}$ from Eq. (2) and consider the resultant equation as a transformation which takes points $\mathbf{z}$ from the trajectory to a new set of transformed variables $\hat{\mathbf{z}}$ (with $\hat{\mathbf{z}}=\mathbf{z}^*$ whenever $\mathbf{z}=\mathbf{z}^*$).

With this motivation, we define our fixed point transform $\hat{\mathbf{z}}=\mathbf{G}(\mathbf{z},\mathbf{R})$ as

$$\hat{\mathbf{z}}\equiv\mathbf{G}(\mathbf{z},\mathbf{R})=[\mathbf{I}-\mathbf{S}(\mathbf{z},\mathbf{R})]^{-1}\cdot[\mathbf{F}(\mathbf{z})-\mathbf{S}(\mathbf{z},\mathbf{R})\cdot\mathbf{z}], \tag{3}$$

where $\mathbf{I}$ is the $d\times d$ identity matrix, and $\mathbf{S}(\mathbf{z},\mathbf{R})$ is a $d\times d$ matrix function of $\mathbf{z}$ and an adjustable $d\times d\times d$ tensoral quantity $\mathbf{R}$,

$$\mathbf{S}(\mathbf{z},\mathbf{R})=\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})+\mathbf{R}\cdot[\mathbf{F}(\mathbf{z})-\mathbf{z}]. \tag{4}$$

In the case $\mathbf{R}=\mathbf{0}$, Eq. (3) together with Eq. (4) is simply the solution of Eq. (2) for $\mathbf{z}^*$ in terms of $\mathbf{z}$ with order $O((\mathbf{z}-\mathbf{z}^*)^2)$ terms neglected. We want to emphasize that local approximations to all terms on the right-hand side of Eq. (3) can be obtained experimentally from time series data (see Sec. II E for details). The inclusion of the additional tensorial parameter $\mathbf{R}$ in the fixed point transform might seem arbitrary at this point, but the role of $\mathbf{R}$ will become transparent in the following subsections. At this point, we consider $\mathbf{R}$ as a fixed tensor with constant elements.

In the special case in which $\mathbf{F}(\mathbf{z})$ is a linear function, i.e., $\mathbf{F}(\mathbf{z})=\mathbf{z}^*+\boldsymbol{\alpha}\cdot(\mathbf{z}-\mathbf{z}^*)$ with $\boldsymbol{\alpha}$ being a constant matrix, and $\mathbf{R}=\mathbf{0}$, we have $\mathbf{S}(\mathbf{z},\mathbf{0})=\boldsymbol{\alpha}$, and

$$\mathbf{G}(\mathbf{z},\mathbf{0})=[\mathbf{I}-\mathbf{S}(\mathbf{z},\mathbf{0})]^{-1}\cdot[\mathbf{F}(\mathbf{z})-\mathbf{S}(\mathbf{z},\mathbf{0})\cdot\mathbf{z}]$$
$$=[\mathbf{I}-\boldsymbol{\alpha}]^{-1}\cdot[\mathbf{z}^*+\boldsymbol{\alpha}\cdot(\mathbf{z}-\mathbf{z}^*)-\boldsymbol{\alpha}\cdot\mathbf{z}]=\mathbf{z}^*,$$

independent of the values of $\mathbf{z}$. In other words, *all* points on the trajectory will be mapped by $\mathbf{G}(\mathbf{z},\mathbf{0})$ to the fixed point $\mathbf{z}^*$.

The case with $\mathbf{R}=\mathbf{0}$ is a special degenerate case for the transformation $\hat{\mathbf{z}}=\mathbf{G}(\mathbf{z},\mathbf{R})$. In this case, even with $\mathbf{F}(\mathbf{z})$ being a nonlinear function, all points in the d-dimensional delay space are mapped to a one-dimensional subspace defined by the line $\hat{z}_1=\hat{z}_2=\cdots=\hat{z}_d$ [16]. In the following discussions, we will consider only the nondegenerate cases with $\mathbf{R}\neq\mathbf{0}$.

In the case of a general nonlinear $\mathbf{F}(\mathbf{z})$ and $\mathbf{R}\neq\mathbf{0}$, most of the points in the linear region around a fixed point $\mathbf{z}^*$ will still be mapped to the vicinity of $\mathbf{z}^*$. Specifically, we will show that the fixed point transform $\mathbf{G}(\mathbf{z},\mathbf{R})$ dramatically enhances the clustering of the transformed data around the fixed point $\mathbf{z}^*$.

B. The stationary points of the fixed point transform

Before describing the effect of the fixed point transform on the natural probability measure on a given attractor, we first examine the functional behavior of $\mathbf{G}(\mathbf{z},\mathbf{R})$ near the fixed point $\mathbf{z}^*$. There are two main properties of $\mathbf{G}(\mathbf{z},\mathbf{R})$.

(i) $\mathbf{G}(\mathbf{z}^*,\mathbf{R})=\mathbf{z}^*$: If $\mathbf{z}^*$ is a fixed point of the underlying dynamics, i.e., if $\mathbf{F}(\mathbf{z}^*)=\mathbf{z}^*$, then $\mathbf{z}^*$ will also be a fixed point of the transformation $\mathbf{G}(\mathbf{z},\mathbf{R})$.

By construction, the matrix $\mathbf{S}(\mathbf{z}^*,\mathbf{R})$ in Eq. (3) evaluated at the fixed point is simply the Jacobian matrix $\nabla\mathbf{F}(\mathbf{z}^*)$ independent of the value of $\mathbf{R}$, i.e.,

$$\mathbf{S}(\mathbf{z}^*,\mathbf{R})=\nabla\mathbf{F}(\mathbf{z}^*). \tag{5}$$

With the above expression for $\mathbf{S}(\mathbf{z}^*,\mathbf{R})$, the transformation $\mathbf{G}(\mathbf{z}^*,\mathbf{R})$ evaluated at the fixed point is

$$\begin{aligned}\mathbf{G}(\mathbf{z}^*,\mathbf{R})&=[\mathbf{I}-\mathbf{S}(\mathbf{z}^*,\mathbf{R})]^{-1}\cdot[\mathbf{F}(\mathbf{z}^*)-\mathbf{S}(\mathbf{z}^*,\mathbf{R})\cdot\mathbf{z}^*]\\&=[\mathbf{I}-\nabla\mathbf{F}(\mathbf{z}^*)]^{-1}\cdot[\mathbf{z}^*-\nabla\mathbf{F}(\mathbf{z}^*)\cdot\mathbf{z}^*]=\mathbf{z}^*,\end{aligned} \tag{6}$$

independent of $\mathbf{R}$.

(ii) Generically, $\nabla\mathbf{G}(\mathbf{z}^*,\mathbf{R})=\mathbf{0}$. That is, $\mathbf{z}^*$ is a stationary point of the transformation $\mathbf{G}(\mathbf{z},\mathbf{R})$.

Multiplying Eq. (3) by the factor $[\mathbf{I}-\mathbf{S}(\mathbf{z},\mathbf{R})]$, we have

$$[\mathbf{I}-\mathbf{S}(\mathbf{z},\mathbf{R})]\cdot\mathbf{G}(\mathbf{z},\mathbf{R})=\mathbf{F}(\mathbf{z})-\mathbf{S}(\mathbf{z},\mathbf{R})\cdot\mathbf{z},$$

which when differentiated gives

$$\begin{aligned}&-\nabla\mathbf{S}(\mathbf{z},\mathbf{R})\circ[\mathbf{G}(\mathbf{z},\mathbf{R})-\mathbf{z}]+[\mathbf{I}-\mathbf{S}(\mathbf{z},\mathbf{R})]\cdot\nabla\mathbf{G}(\mathbf{z},\mathbf{R})\\&\quad=\nabla\mathbf{F}(\mathbf{z})-\mathbf{S}(\mathbf{z},\mathbf{R}).\end{aligned} \tag{7}$$

Here, $\nabla\mathbf{S}$ is a $d\times d\times d$ tensor with its ijkth component given by $\partial S_{ij}/\partial z_k$ and we use $\nabla\mathbf{S}\circ\mathbf{G}$ to denote the $d\times d$ matrix obtained by the contraction of the outer product between the tensor $\nabla\mathbf{S}$ and the vector $\mathbf{G}$, i.e., the ijth component of this matrix is given by $\Sigma_k(\partial S_{ik}/\partial z_j)G_k$.

Evaluating the above expression at $\mathbf{z}^*$ and substituting Eq. (6) $[\mathbf{G}(\mathbf{z}^*,\mathbf{R})=\mathbf{z}^*]$ and Eq. (5) $[\mathbf{S}(\mathbf{z}^*,\mathbf{R})=\nabla\mathbf{F}(\mathbf{z}^*)]$ into Eq. (7), we have

$$[\mathbf{I}-\nabla\mathbf{F}(\mathbf{z}^*)]\cdot\nabla\mathbf{G}(\mathbf{z}^*,\mathbf{R})=\mathbf{0}.$$

The matrix $\mathbf{I}-\nabla\mathbf{F}(\mathbf{z}^*)$ is nonsingular for a typical nonlinear map $\mathbf{F}(\mathbf{z})$ except for $\mathbf{z}^*$ precisely at a bifurcation point. Thus, generically

$$\nabla\mathbf{G}(\mathbf{z}^*,\mathbf{R})=\mathbf{0}. \tag{8}$$

In summary, if $\mathbf{z}^*$ is a fixed point of the map $\mathbf{F}(\mathbf{z})$, it is both a fixed point [Eq. (6)] and a stationary point [Eq. (8)] of the transformation $\mathbf{G}(\mathbf{z},\mathbf{R})$ independent of $\mathbf{R}$. In particular, by Taylor expansion, one arrives at the following form for $\mathbf{G}(\mathbf{z},\mathbf{R})$ near the fixed point $\mathbf{z}^*$:

$$\hat{z}_i-z_i^*\equiv G_i(\mathbf{z},\mathbf{R})-z_i^*\approx\beta_{ijk}(z_j-z_j^*)(z_k-z_k^*), \tag{9}$$

where $i,j,k=1,\ldots,d$, the coefficients β_{ijk} are invariant to interchange of j and k, and summing of repeated indices is implied. Moreover, by differentiating Eq. (9), the Jacobian matrix of the transformation $\nabla\mathbf{G}(\mathbf{z},\mathbf{R})$ $[\partial_jG_i(\mathbf{z},\mathbf{R})$ with $i,j=1,\ldots,d]$ will have the following form near the fixed point $\mathbf{z}^*$:

$$\partial_jG_i(\mathbf{z},\mathbf{R})\approx 2\beta_{ijk}(z_k-z_k^*). \tag{10}$$

We also note that to the lowest order in $(\mathbf{z}-\mathbf{z}^*)$ near a fixed point, both $\mathbf{G}(\mathbf{z},\mathbf{R})$ and $\nabla\mathbf{G}(\mathbf{z},\mathbf{R})$ are *independent* of $\mathbf{R}$.

C. Probability measures on the attractor and the fixed point transform

One way to characterize the clustering of points on an attractor is by its pointwise dimension $D_p(\mathbf{z})$, which is defined by the following limit [17]:

$$D_p(\mathbf{z})=\lim_{\epsilon\to 0}\frac{\ln\mu(B_\epsilon(\mathbf{z}))}{\ln\epsilon}, \tag{11}$$

where $B_\epsilon(\mathbf{z})$ is a d-dimensional ball centered around a point $\mathbf{z}$ in $\mathbb{R}^d$ and μ is the natural measure of the attractor. If the distribution of points is uniform around $\mathbf{z}$, the measure $\mu(B_\epsilon(\mathbf{z}))$ (or the fraction of points inside the ball B_ϵ) will scale as ϵ^d (i.e., the pointwise dimension is the dimension of the embedding space d). On the other hand, if the attractor is a stable fixed point, then all data points are at $\mathbf{z}$ and the measure $\mu(B_\epsilon(\mathbf{z}))$ is one independent of the ball size ϵ. This gives a pointwise dimension of zero, which is the dimension of a point. The pointwise dimension of a fractal object embedded in $\mathbb{R}^d$ typically falls somewhere in between, with the smaller values of $D_p(\mathbf{z})$ indicating a higher degree of clustering around the point $\mathbf{z}$.

The measure enhancing property of the transformation defined in Eq. (3) with $\mathbf{R}\neq\mathbf{0}$ can be stated in the following claim:

Let the pointwise dimension of the attractor at $\mathbf{z}$ *be* $D_p(\mathbf{z})$. *Then, under the transformation* $\mathbf{G}(\mathbf{z},\mathbf{R})$, *the pointwise dimension of the transformed attractor will be unchanged at typical points but it will typically be halved at stationary points of* $\mathbf{G}(\mathbf{z},\mathbf{R})$ [i.e., $\nabla\mathbf{G}(\mathbf{z},\mathbf{R})=\mathbf{0}$]:

$$\hat{D}_p(\hat{\mathbf{z}})=D_p(\mathbf{z})\ \ at\ \ \mathbf{z},\ \ where\ \ \nabla\mathbf{G}(\mathbf{z},\mathbf{R})\ \ is\ nonsingular, \tag{12}$$

$$\hat{D}_p(\hat{\mathbf{z}})=D_p(\mathbf{z})/2\ \ at\ \ \mathbf{z},\ \ where\ \ \nabla\mathbf{G}(\mathbf{z},\mathbf{R})=\mathbf{0}, \tag{13}$$

where $\hat{D}_p(\hat{\mathbf{z}})$ *is the pointwise dimension of the transformed attractor at* $\hat{\mathbf{z}}$. *(We define what we mean by the word ''typically,'' used above, later.)*

The first part of the claim, Eq. (12), can be illustrated using the following argument. For simplicity, say that the origin is a regular point of $\mathbf{G}(\mathbf{z},\mathbf{R})$, i.e., the matrix $\nabla\mathbf{G}(\mathbf{0},\mathbf{R})$ is nonsingular. We wish to show that $\hat{D}_p(\hat{\mathbf{z}}(\mathbf{0}))=D_p(\mathbf{0})$, where $\hat{\mathbf{z}}(\mathbf{0})\equiv\mathbf{G}(\mathbf{0},\mathbf{R})$ is the image of the origin under $\mathbf{G}$ and $\hat{D}_p(\hat{\mathbf{z}})$ denotes the pointwise dimension of the measure in the transformed space. Given a small ball $B_\epsilon(\mathbf{0})$ around the origin, the transformation $\mathbf{G}(\mathbf{z},\mathbf{R})$ is approximately given by the following Taylor expansion of $\mathbf{G}(\mathbf{z},\mathbf{R})$ around the origin:

$$\mathbf{G}(\mathbf{z},\mathbf{R})\approx\hat{\mathbf{z}}(\mathbf{0})+\nabla\mathbf{G}(\mathbf{0},\mathbf{R})\cdot\mathbf{z}.$$

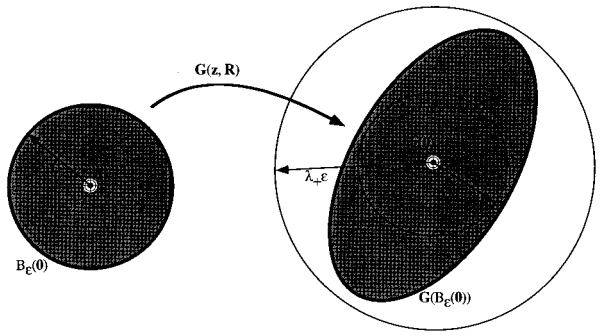

FIG. 1. Schematic illustration showing the action of the fixed point transform $\mathbf{G}(\mathbf{z},\mathbf{R})$ near the origin with $\nabla\mathbf{G}(\mathbf{0},\mathbf{R})$ being nonsingular. The image $\mathbf{G}(B_\epsilon)$ of a ball, B_ϵ, is an ellipsoid centered around $\hat{\mathbf{z}}(\mathbf{0})$. There exist two balls $B_{\lambda_+\epsilon}$ and $B_{\lambda_-\epsilon}$ such that $B_{\lambda_-\epsilon}\subset\mathbf{G}(B_\epsilon)\subset B_{\lambda_+\epsilon}$.

Geometrically, $\mathbf{G}(B_\epsilon(\mathbf{0}))$ will approximately be an ellipsoid centered at $\hat{\mathbf{z}}(\mathbf{0})$ and one can find two constants λ_+ and λ_- and two small balls $B_{\lambda_+\epsilon}(\hat{\mathbf{z}}(\mathbf{0}))$ and $B_{\lambda_-\epsilon}(\hat{\mathbf{z}}(\mathbf{0}))$ such that

$$B_{\lambda_-\epsilon}(\hat{\mathbf{z}}(\mathbf{0}))\subset\mathbf{G}(B_\epsilon(\mathbf{0}))\subset B_{\lambda_+\epsilon}(\hat{\mathbf{z}}(\mathbf{0})) \tag{14}$$

uniformly in ϵ (see Fig. 1 for an illustration of these covering balls in two dimensions). In terms of the natural probability measure μ, Eq. (14) yields

$$\hat{\mu}(B_{\lambda_-\epsilon}(\hat{\mathbf{z}}(\mathbf{0})))\leqslant\hat{\mu}(\mathbf{G}(B_\epsilon(\mathbf{0})))=\mu(B_\epsilon(\mathbf{0}))\leqslant\hat{\mu}(B_{\lambda_+\epsilon}(\hat{\mathbf{z}}(\mathbf{0}))). \tag{15}$$

Then, by utilizing the definition for the pointwise dimension [see Eq. (11)], we arrive at the following relations:

$$\lim_{\epsilon\to0}\frac{\ln\hat{\mu}(\mathbf{G}(B_\epsilon(\mathbf{0})))}{\ln(\lambda_{+,-}\epsilon)}=\lim_{\epsilon\to0}\frac{\ln\mu(B_\epsilon(\mathbf{0}))}{\ln\epsilon}=D_p(\mathbf{0}),$$

$$\lim_{\epsilon\to0}\frac{\ln\hat{\mu}(B_{\lambda_{+,-}\epsilon}(\hat{\mathbf{z}}(\mathbf{0})))}{\ln(\lambda_{+,-}\epsilon)}=\hat{D}_p(\hat{\mathbf{z}}(\mathbf{0})).$$

The inequality in Eq. (15) then gives

$$\hat{D}_p(\hat{\mathbf{z}}(\mathbf{0}))\leqslant D_p(\mathbf{0})\leqslant\hat{D}_p(\hat{\mathbf{z}}(\mathbf{0})),$$

which yields the desired result Eq. (12) [18],

$$\hat{D}_p(\hat{\mathbf{z}}(\mathbf{0}))=D_p(\mathbf{0}).$$

We now turn to the second part of the claim [Eq. (13)]. Let the origin be a stationary point of $\mathbf{G}(\mathbf{z},\mathbf{R})$, i.e., $\nabla\mathbf{G}(\mathbf{0},\mathbf{R})=\mathbf{0}$. Then, near the origin, by Eq. (9), we have

$$G_i(\mathbf{z},\mathbf{R})-G_i(\mathbf{0},\mathbf{R})\approx Q_i(\mathbf{z})=\beta_{ijk}z_jz_k,$$

where $i,j,k=1,\ldots,d$ and summing over repeated indices is assumed.

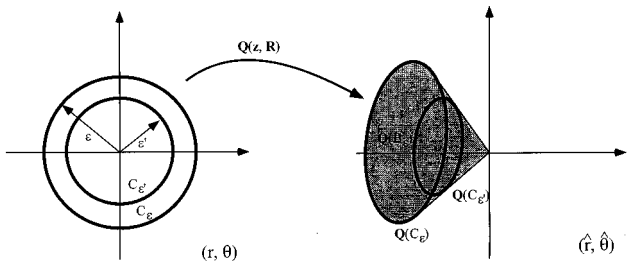

FIG. 2. Schematic illustration showing the action of the quadratic map $\mathbf{Q}(\mathbf{z})$ on a ball B_ϵ. The two circles C_ϵ and $C_{\epsilon'}$ with $\epsilon'<\epsilon$ are mapped by $\mathbf{Q}(\mathbf{z})$ as shown. Because of nonlinearity, the image $C_{\epsilon'}$ is not necessarily entirely contained within the image C_ϵ. However, since $\mathbf{z}\to\mathbf{0}$, $\mathbf{Q}(\mathbf{z})\to\mathbf{0}$, the image $\mathbf{Q}(B_\epsilon)$ of B_ϵ will be a wedged structure as indicated by the shaded region.

Now, for ease of visualization, we will examine the effect of this quadratic transform in two dimensions. Using polar coordinates $\mathbf{z}=(r,\theta)$, the transformation $\mathbf{Q}(\mathbf{z})$ can be written as

$$Q_1(r,\theta)=(\beta_{111}\cos^2\theta+\beta_{122}\sin^2\theta+2\beta_{112}\cos^2\theta\sin^2\theta)r^2=q_1(\theta)r^2,$$

$$Q_2(r,\theta)=(\beta_{211}\cos^2\theta+\beta_{222}\sin^2\theta+2\beta_{212}\cos^2\theta\sin^2\theta)r^2=q_2(\theta)r^2,$$

or, equivalently, the transformed coordinates $(\hat{r},\hat{\theta})$ are given by

$$\hat{r}=\sqrt{Q_1^2+Q_2^2}=r^2\sqrt{q_1^2(\theta)+q_2^2(\theta)},$$

$$\tan(\hat{\theta})=\frac{Q_2}{Q_1}=\frac{q_2(\theta)}{q_1(\theta)}.$$

The form factors $q_1(\theta)$ and $q_2(\theta)$ are functions independent of the radius r and are functions of θ only.

The transformation $\mathbf{Q}(\mathbf{z})$ is degenerate if $\hat{\theta}$ is independent of θ [i.e., the ratio $q_2(\theta)/q_1(\theta)$ is independent of θ]. In this case, all points in the $\mathbf{z}$ plane will be mapped by $\mathbf{Q}(\mathbf{z})$ to a line going through the origin with an inclination θ_0 [19]. We have said that Eq. (13) applied ''typically,'' by which we mean that $q_2(\theta)/q_1(\theta)$ is not independent of θ and that $q_{1,2}(\theta)\neq0$.

Generically, the transformation $\mathbf{Q}(\mathbf{z})$ will map a circle C_ϵ of radius ϵ around the origin to an ellipse $\mathbf{Q}(C_\epsilon)$ in the $(\hat{r},\hat{\theta})$ plane. The ellipse has major and minor radii of order ϵ^2 and may or may not (depending on the β_{ijk}) encircle of the origin ($\hat{r}=0$). A case where it does not encircle the origin is shown schematically in Fig. 2, and though the following discussion is restricted to this case, the case where $\mathbf{Q}(C_\epsilon)$ encircles the origin can be similarly treated. For a value of $\epsilon'<\epsilon$, the circle $C_{\epsilon'}$ becomes an ellipse $\mathbf{Q}(C_{\epsilon'})$ with the same shape and orientation as $\mathbf{Q}(C_\epsilon)$ but is now closer to the origin and smaller. In particular, for fixed $\hat{\theta}$, points on $\mathbf{Q}(C_{\epsilon'})$ are obtained from points on $\mathbf{Q}(C_\epsilon)$ by multiplying $\hat{r}$ by $(\epsilon'/\epsilon)^2$ [this follows because $Q_{1,2}(r,\theta)=q_{1,2}(\theta)r^2$]. As $\epsilon\to0$, the limit of $\mathbf{Q}(C_\epsilon)$ ap-

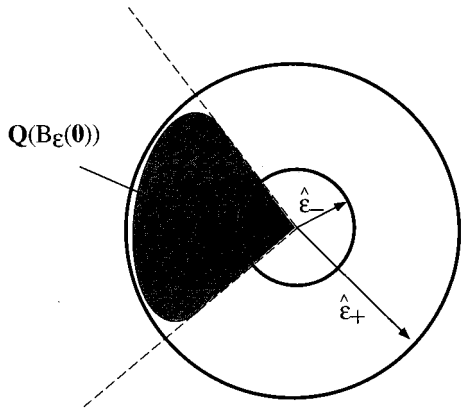

FIG. 3. Illustration of the bounding balls with radii ϵ_+ and ϵ_- for the quadratically transformed ball $\mathbf{Q}(B_\epsilon(\mathbf{0}))$. Note that the entire plane is mapped by $\mathbf{Q}(\mathbf{z})$ into the wedged region between the two dashed lines.

proaches the origin. As a result, the image of a disk B_ϵ under the transformation $\mathbf{Q}(\mathbf{z})$ will be a wedgelike object composed of the union of all the ellipses with $0 \leq \epsilon' < \epsilon$,

$$\mathbf{Q}(B_\epsilon) = \bigcup_{0 \leq \epsilon' < \epsilon} \mathbf{Q}(C_{\epsilon'})$$

(as shown by the shaded region in Fig. 2).

Define $\hat{\boldsymbol{\epsilon}}_+$ and $\hat{\boldsymbol{\epsilon}}_-$ as shown in Fig. 3. From the homogeneous quadratic nature of Q_1 and Q_2, we have that $\hat{\boldsymbol{\epsilon}}_+$ and $\hat{\boldsymbol{\epsilon}}_-$ are of the form

$$\hat{\boldsymbol{\epsilon}}_+ = \lambda_+ \boldsymbol{\epsilon}^2, \tag{16}$$

$$\hat{\boldsymbol{\epsilon}}_- = \lambda_- \boldsymbol{\epsilon}^2, \tag{17}$$

where $\lambda_- < \lambda_+$ are positive constants. Since all points in the neighborhood of $\mathbf{z}=\mathbf{0}$ map to the wedge between the dashed straight lines in Fig. 3, we have that

$$\hat{\mu}(B_{\hat{\epsilon}_+}(\mathbf{0})) \geq \hat{\mu}(\mathbf{Q}(B_\epsilon(\mathbf{0}))) = \mu(B_\epsilon(\mathbf{0})), \tag{18}$$

$$\hat{\mu}(B_{\hat{\epsilon}_-}(\mathbf{0})) \leq \hat{\mu}(\mathbf{Q}(B_\epsilon(\mathbf{0}))) = \mu(B_\epsilon(\mathbf{0})). \tag{19}$$

Inserting Eqs. (16)–(19) into the definition of the pointwise dimension and proceeding as in the steps following Eq. (15), we obtain the previously claimed result, Eq. (13) [20],

$$\hat{D}_p(\hat{\mathbf{z}}(\mathbf{0})) = D_p(\mathbf{0})/2.$$

Although, for illustration, we have restricted consideration to two dimensions, the same reasoning carries over to higher dimensionality, and Eq. (13) still applies.

D. Extracting peaks from the distribution function of the transformed data

From our previous discussion, we have seen that by construction, fixed points $\mathbf{z}^*$ of the discrete dynamical system $\mathbf{F}(\mathbf{z})$ are automatically stationary points of the fixed point transform $\hat{\mathbf{z}} = \mathbf{G}(\mathbf{z},\mathbf{R})$. Thus, by the lowered pointwise dimension [Eq. (13)] at stationary points, one should expect strong clustering (sharp peaks) at fixed points $\mathbf{z}^*$ in a finite grid approximation of the probability measure of the transformed data. In practice, the degree of clustering around fixed points in the transformed data depends on the size of the local regions where Eq. (2) is a good approximation to the dynamics, on how often a typical trajectory visits these local regions, on the accuracy and resolution of the time series data, and on noise. We also find that, since the method samples the local dynamics near a fixed point, it is also possible that fixed points that are very close but not actually on the attractor can be detected from the probability measure of the transformed data [21].

While fixed points of $\mathbf{F}(\mathbf{z})$ are stationary points of the fixed point transform $\mathbf{G}(\mathbf{z},\mathbf{R})$, there exist other points $\mathbf{z}^s \neq \mathbf{z}^*$ that are not fixed points such that $\boldsymbol{\nabla}\mathbf{G}(\mathbf{z}^s,\mathbf{R}) = \mathbf{0}$. With a given fixed value of $\mathbf{R}$, the probability measure of the transformed data will have spurious peaks at the images of these stationary points as well as at true fixed points. However, in the $\hat{\mathbf{z}}$ space of the transformed data, the locations of these spurious peaks given by $\mathbf{G}(\mathbf{z}^s,\mathbf{R})$ will depend upon the tensorial parameter $\mathbf{R}$ while the peaks at true fixed points $\mathbf{G}(\mathbf{z}^*,\mathbf{R}) = \mathbf{z}^*$ will be *independent* of $\mathbf{R}$ [recall property (i) of the fixed point transform in the previous subsection, Sec. II B]. Thus, one way to remove these spurious peaks is to pick many random $\mathbf{R}$'s for the calculation of the fixed point transform $\mathbf{G}(\mathbf{z},\mathbf{R})$ at each point $\mathbf{z}$ of the data set. While the images of the spurious stationary points in $\hat{\mathbf{z}}$ space will be wildly scattered by different choices of $\mathbf{R}$, the true fixed points at $\mathbf{z}^*$ will not be affected. Thus, we remove the clustering around spurious stationary points by smearing them, while preserving the measure enhancing effect around true fixed points.

In practice, to extract unstable fixed points from the set of transformed data, one could simply look for sharp peaks in a finite grid approximation to the distribution function $\hat{\mu}(\hat{\mathbf{z}})$ of $\hat{\mathbf{z}}$. If the dimension of the embedding space is small, it is computationally simple to look for sharp peaks in the bins of a d-dimensional grid. On the other hand, if d is large, it might be preferable to reduce this d-dimensional distribution function to a one-dimensional histogram $\hat{\rho}(\hat{z})$ with the scalar $\hat{z}$ being a function of the d-dimensional vector $\hat{\mathbf{z}}$.

One simple reduction procedure is to utilize the cyclic symmetry in the coordinates of delay time embedding [see Eq. (1)]. With this cyclic symmetry, all fixed points must lie on the diagonal line defined by $\hat{z}_1 = \cdots = \hat{z}_d$. This allows one to construct a one-dimensional histogram $\hat{\rho}(\hat{z})$ from the d-dimensional distribution function $\hat{\mu}(\hat{\mathbf{z}})$ by considering only $\hat{\mathbf{z}}$ within a small cross-section tube about the line $\hat{z}_1 = \cdots = \hat{z}_d$. The relative sharpness of the peaks at fixed points will remain if the width of the cross section of the tube is chosen to be sufficiently small.

In the special case when the pointwise dimension of the attractor at fixed points is not too large [$D_p(\mathbf{z}^*) < 2$], an even simpler procedure is possible. [Here, the restriction is on the pointwise dimension $D_p(\mathbf{z}^*)$ of the attractor at fixed points, while the actual dimensions d needed for time delay embedding can be large.] In this case, one can construct a

one-dimensional histogram by considering all values of $\hat{z}=\hat{z}_i$ $(i=1,\ldots,d)$ for all transformed data $\hat{\mathbf{z}}$. Since this procedure involves projecting the d-dimensional distribution function $\hat{\mu}(\hat{\mathbf{z}})$ onto a one-dimensional set, the relative sharpness of the peaks at fixed points will be reduced.

Projection of the data from the d-dimensional $\hat{\mathbf{z}}$ space to the one-dimensional $\hat{z}$ line preserves the pointwise dimension $\hat{D}_p(\hat{\mathbf{z}}(\mathbf{z}))$ if $\hat{D}_p(\hat{\mathbf{z}}(\mathbf{z}))<1$ (see [22]), while if $\hat{D}_p(\hat{\mathbf{z}}(\mathbf{z}))\geqslant 1$, projection yields a pointwise dimension of 1 at $\hat{z}$. For $\mathbf{z}=\mathbf{z}^*$, a fixed point, $\hat{\mathbf{z}}(\mathbf{z}^*)$, has been made to have a low pointwise dimension, $\hat{D}_p(\hat{\mathbf{z}}(\mathbf{z}))<1$, while, in contrast, at typical points $\mathbf{z}$, the pointwise dimension of the transformed data $\hat{D}_p(\hat{\mathbf{z}}(\mathbf{z}))=D_p(\mathbf{z})$ exceeds 1 for a chaotic attractor. Thus, if we form the density $\hat{\rho}(\hat{z})$ from the projected data, this density has a singularity $\hat{\rho}(\hat{z})\sim|\hat{z}-\hat{z}(\mathbf{z}^*)|^{-\nu}$, where $\nu=[1-D_p(\mathbf{z}^*)/2]>0$. That is, $\hat{\rho}(\hat{z})$ has a singularity at the location of the fixed point. We find that such singularities can often be easily picked out from a histogram plot of $\hat{\rho}(\hat{z})$ [9].

E. Obtaining S(z,R) from time series data

In the definition of the fixed point transform Eq. (3), $\mathbf{S}(\mathbf{z},\mathbf{R})$ is a sum of two terms with the first term given by the derivative of $\mathbf{F}(\mathbf{z})$, $\nabla\mathbf{F}(\mathbf{z})$, and the second perturbative term given by $\mathbf{R}\cdot[\mathbf{F}(\mathbf{z})-\mathbf{z}]$. The value of $\mathbf{R}$ is chosen such that for a typical $\mathbf{z}$, $\mathbf{R}\cdot[\mathbf{F}(\mathbf{z})-\mathbf{z}]$ is roughly comparable in order of magnitude to $\nabla\mathbf{F}(\mathbf{z})$. While $\mathbf{F}(\mathbf{z})$ is simply the next iterate of $\mathbf{z}$ in the delay embedded space, $\nabla\mathbf{F}(\mathbf{z})$ can be estimated using a least square fit procedure.

In delay coordinates, $\nabla\mathbf{F}(\mathbf{z})$ has d free parameters, namely, $\partial_1 f(\mathbf{z}),\ldots,\partial_d f(\mathbf{z})$ [see Eq. (1) for the special structure of $\mathbf{F}(\mathbf{z})$ in delay coordinates],

$$\nabla\mathbf{F}(\mathbf{z})=\begin{pmatrix}\partial_1 f(\mathbf{z}) & \cdots & \partial_{d-1} f(\mathbf{z}) & \partial_d f(\mathbf{z})\\ 1 & \cdots & 0 & 0\\ \vdots & \cdots & \vdots & \vdots\\ 0 & \cdots & 1 & 0\end{pmatrix}.$$

By considering a collection $\{\mathbf{w}^k\}$, $k=1,\ldots,M(\geqslant d)$ of M points which are neighbors of $\mathbf{z}$, the d parameters of $\nabla\mathbf{F}(\mathbf{z}(n))$ can be solved by fitting the following linear equation to the data:

$$w_1^k(n+1)-z_1(n+1)=\nabla f(\mathbf{z}(n))\cdot[\mathbf{w}^k-\mathbf{z}(n)], \quad (20)$$

where $\nabla f(\mathbf{z})=(\partial_1 f(\mathbf{z}),\ldots,\partial_d f(\mathbf{z}))$ is the gradient of the scalar function $f(\mathbf{z})$.

Two types of neighbors are considered in our method. First, since we are interested in capturing the local dynamical behavior near $\mathbf{z}(n)$, the natural choice is the M closest points in the d-dimensional reconstructed phase space. We refer to points whose phase space locations are near $\mathbf{z}(n)$ as ''spatial'' neighbors of $\mathbf{z}(n)$. In the case, however, when the system parameters are not stationary in time, especially in biological systems, the choice of spatial neighbors might not be appropriate. This is because spatial nearby points in general are not necessarily close in time and therefore they might represent states with different system parameters. To partially deal with time series from slowly drifting nonstationary processes, one can alternatively define the collection of neighbors for $\mathbf{z}(n)$ to be its next $M/2$ iterates together with its $M/2$ preimages. [When $\mathbf{z}(n)$ is very near a fixed point, these temporal neighbors are also spatial neighbors.] Thus, in this case, the transformation $\mathbf{G}(\mathbf{z}(n))$ is *solely* defined within a small window of time near n, i.e., within the short sequence of points: $\{\mathbf{z}(n-M/2),\ldots,\mathbf{z}(n+M/2)\}$. By assuming that the typical time scale of the parametric variations is much larger than M, the choice of temporal neighbors might give a better estimate of the local dynamics near $\mathbf{z}(n)$ than the choice of spatial neighbors for nonstationary processes.

Obviously, all points in the temporal sequence, $\{\mathbf{z}(n-M/2),\ldots,\mathbf{z}(n+M/2)\}$, have to be in the local region of $\mathbf{z}(n)$ in order for the linear approximation to the local dynamics $\nabla\mathbf{F}(\mathbf{z}(n))$ [see Eq. (20)] to be meaningful. This will certainly not be true for all points on the trajectory, especially for points far away from the fixed point. On the other hand, near a fixed point, trajectories stay near it for a short period of time. These short sequences of points in the local region of the fixed points are exactly those which will satisfy the fixed point criterion given by Eq. (2) and be transformed to the singularity at the fixed point.

F. A Simple numerical example: Skewed logistic map

As a simple example illustrating the singularities in $\hat{\rho}(\hat{z})$, we generate a time series from the skewed logistic map, $f(z)=\mu e^{-z}z(1-z)$, with $\mu=6.1$ [see the inset of Fig. 4(a) for a graph of $f(z)$]. The exponential factor, e^{-z}, introduces higher order nonlinearities to the dynamics so that the Taylor series expansion of $f(z)$ will not be a trivial quadratic polynomial. A histogram approximation to the distribution function $\rho(z)$ of the *un*transformed data z is plotted in Fig. 4(a) using 8192 iterates of $f(z)$. We see that most of the 8192 data points are primarily clustered around the first and second iterates of the critical point of f [$z^c=0.382$; $f(z^c)=0.983$; $f^2(z^c)=0.0385$]. This one-dimensional map has one fixed point ($z^*=0.667$) on the attractor and a second fixed point $z^*=0$ off the attractor. From the distribution function $\rho(z)$, we also observe that the probability for a typical orbit to visit the region near the fixed point at $z^*=0.667$ (within an interval centered around z^* and with size $\approx 1\%$ of the radius of the attractor) is relatively low (≈ 80 counts/8192 counts$\approx 1\%$). This should later be compared with the much enhanced clustering near $z^*=0.667$ in the distribution function $\hat{\rho}(\hat{z})$ of the transformed data $\hat{z}$ using $G(z)=\hat{z}$ [see Fig. 4(b)].

Using temporal neighbors with $M=d=1$, we can write down the fixed point transform $G(z(n),R)$ in one dimension explicitly,

$$G(z(n),R)=\hat{z}(n)=\frac{z(n+1)-S(z(n),R)z(n)}{1-S(z(n),R)}, \quad (21)$$

where

$$S(z(n),R)=\frac{z(n+2)-z(n+1)}{z(n+1)-z(n)}+R[z(n+1)-z(n)], \quad (22)$$

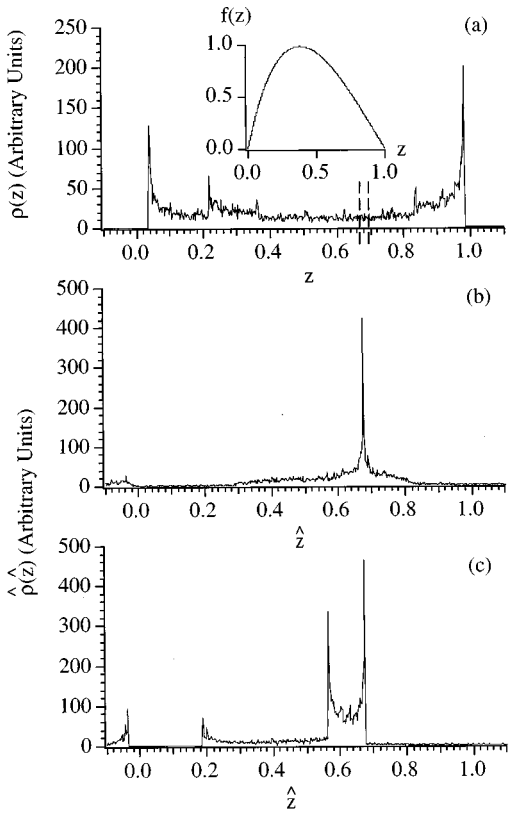

FIG. 4. (a) A histogram approximation to the distribution function of the time series data $\{z(n)\}$ from a one-dimensional chaotic map (see Sec. II F for a description of this map). It has a fixed point on the attractor at $z^*=0.6777$. Using 8192 data points, the number of points that fall within a small interval (indicated by a pair of dotted lines) centered around the fixed point z^* is approximately 80. Inset is a graphical representation of the map f. (b) A histogram plot of $\hat{\rho}(\hat{z})$ averaged over 32 values of $\mathbf{R}$ with $\kappa=10$ using $G(z)$. The number of time series points used here was 256. The parameters are chosen so that the total number of points in the histogram is the same as in (a). (c) A histogram plot of $\hat{\rho}(\hat{z})$ without randomization, i.e., $\kappa=0$. In this case, 8192 data points were used.

with $z(n+1)$ being a temporal neighbor of $z(n)$.

We applied the transformation Eqs. (21) and (22) to 256 iterates of our time series with 32 random realizations of R for each data point $z(n)$. Here, we let $R=\kappa\eta$, where $\kappa=10$ and η is a uniformly distributed random variable in the interval $[-1,1]$. A histogram approximation to the resultant distribution function $\hat{\rho}(\hat{z})$ of the transformed data $\hat{z}$ is plotted in Fig. 4(b). In this graph, the singular peak corresponding to the true fixed point at $\hat{z}=0.677$ can be clearly identified.

As a demonstrative comparison illustrating the efficacy of including the random term $\mathbf{R}$ in the transformation, we recalculated the distribution function $\hat{\rho}(\hat{z})$ for $\hat{z}$ with $\kappa=0$ in Fig. 4(c). In this case, we need a much larger amount of data points (8192) to resolve the peaks. In addition to the true fixed point at $z^*=0.677$, there are three additional spurious peaks. Two of them are from strong singularities of $\rho(z)$ located at the first and second iterates of the critical point z^c: $\hat{z}(f(z^c))=0.189$ and $\hat{z}(f^2(z^c))=-0.0344$, and the third spurious peak at $\hat{z}=0.568$ is from a stationary point of the transformation $G(z,0)$ that is not due to a fixed point. Comparing Fig. 4(c) with the previous result with $\kappa\neq 0$, all spurious peaks are eliminated from $\hat{\rho}(\hat{z})$ in Fig. 4(b) and a true fixed point peak at $\hat{z}=0.677$ remains.

G. Fixed point transform revisited: Second order

In our previous discussion on the fixed point transform (Sec. II A), one needs to estimate the linearized local dynamics along the reconstructed trajectory $\{\mathbf{z}(n)\}$ in delay space. In principle, one can extend this discussion to include higher order approximations to the local dynamics along the trajectory by considering successively higher order derivatives in the Taylor series expansion of $\mathbf{F}(\mathbf{z})$ [see Eq. (2)]. These derivative terms can again be estimated from time series data by a least square fit technique using a collection of neighboring points in the embedding space.

With a sufficiently large data set, one should expect the sensitivity of our method in detecting fixed points to improve as we include higher ordered terms in our series approximation of the local dynamics. In other words, the effective size of the local region which will be mapped by the fixed point transform to a cluster at the fixed point should increase.

As a simple example to illustrate the possible advantage of utilizing the higher order estimates to the local dynamics, we again consider our previous one-dimensional skewed logistic map $f(z)=\mu e^{-z}z(1-z)$. To approximate $f(z)$ by a Taylor series expansion and neglect all terms with $O((z-z^*)^3)$, the local dynamics near a fixed point z^* is approximately given by,

$$f(z)=z^*+f'(z)(z-z^*)-\frac{1}{2}f''(z)(z-z^*)^2. \tag{23}$$

Then, we defined our ''second-order'' fixed point transform $G_2(z,R,R')=\hat{z}$ implicitly by

$$f(z)=G_2+S(z,R)(z-G_2)-\frac{1}{2}P(z,R')(z-G_2)^2, \tag{24}$$

where

$$S(z,R)=f'(z)+R[f(z)-z]^2, \tag{25}$$

$$P(z,R')=f''(z)+R'[f(z)-z]^2. \tag{26}$$

Here, the adjustable parameters R and R' have a similar role to the one which we introduced earlier in the ''first-order'' fixed point transform, Eq. (3). It may be shown for one-dimensional dynamical systems, such as the skewed logistic map treated here, that the density singularity, $\hat{\rho}(\hat{z})\sim|\hat{z}-\hat{z}(\mathbf{z}^*)|^{-\nu}$, is stronger for G_2 than for G with $\nu=[1-D_p(\mathbf{z}^*)/3]$ as compared to $\nu=[1-D_p(\mathbf{z}^*)/2]$.

Figure 5 is a histogram approximation of $\hat{\rho}(\hat{z})$ using the ''second-order'' fixed point transform. All other parameters are the same as in Fig. 4(b). As compared to the first order result, G_2 approximately doubled the number of points mapped to the vicinity near the fixed point at $z^*=0.667$. Furthermore, the histogram $\hat{\rho}(\hat{z})$ in Fig. 5 also indicates the existence of a secondary peak located slightly left of the

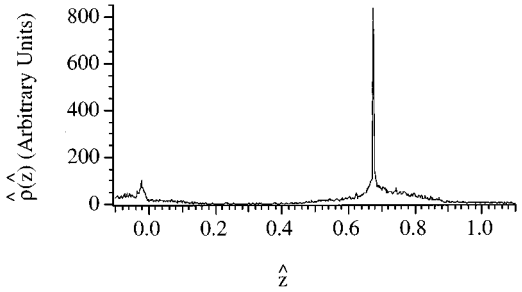

FIG. 5. A histogram plot of $\hat{\rho}(\hat{z})$ using the second-order fixed point transform $G_2(z)$. All parameters are kept the same as in Fig. 4(b).

origin. This secondary peak is due to the existence of another unstable fixed point at $\hat{z}=0$. The estimated location of this unstable fixed point is shifted and the strength of the observed peak is weak because the actual fixed point is located outside the attractor and the trajectory only makes infrequent visits to its local region. As a comparison, by using the ''first order'' transform $G(z)$ as in our previous example [see Fig. 4(b)], the peak near $z^*=0$ is more difficult to distinguish from the background. In other words, the effective *linear* region of the fixed point at $z^*=0$ is not large enough to include much of its natural measure on the attractor.

III. EXTRACTING UNSTABLE PERIODIC ORBITS WITH PERIOD HIGHER THAN 1

In the preceding section, we described a method to extract period 1 fixed points from a finite length discrete time series. In principle, one can simply extend this fixed point detection method to look for periodic orbits with period $p>1$ by applying the method to every pth iterate of the time series. In effect, this corresponds to looking for period 1 fixed points of the pth iterated map $\mathbf{F}^p(\mathbf{z}^*)=\mathbf{z}^*$. Recall that the fixed point transform $\mathbf{G}(\mathbf{z}(n))$ will map $\mathbf{z}(n)$ to the vicinity of a period 1 fixed point $\mathbf{z}^*$ only if both $\mathbf{z}(n)$ and its next iterate $\mathbf{z}(n+1)$ lie in the local region of $\mathbf{z}^*$. In the case of a period p periodic orbit, if we use the fixed point method in the preceding section, then we need to have both $\mathbf{z}(n)$ and $\mathbf{z}(n+p)$ lie in the local region of $\mathbf{z}^*$. In the presence of noise or when the largest Lyapunov exponent of the system is large, the likelihood of finding an appropriate sequence of points which satisfy this requirement is small, especially if p is large. Thus, in many situations, this will not work.

In this section, we introduce a detection method for finding period p periodic orbits with $p>1$ which will take every iterate instead of every pth iterate into consideration. In particular, we describe a technique which utilizes short sequences of points that visit only portions of the desired periodic orbit (see Fig. 6). In the following discussion, we simply refer to this variant method as the ''periodic orbit'' method and the pth iterated extension of the method in Sec. II as the ''fixed point'' method.

A. Periodic orbit transform with $p>1$

To be concrete, assume that we are given a periodic orbit of period p, $\{\mathbf{z}^*(1),\ldots,\mathbf{z}^*(p)\}$, with $\mathbf{F}^p(\mathbf{z}^*(1))=\mathbf{z}^*(p)=\mathbf{z}^*(1)$. There are p elements of this periodic orbit which we label as $\mathbf{z}^*(k)$, where $k=1,\ldots,p$. In delay coordinates, these pd-dimensional vectors have only p independent components, $\{z1^*,z2^*,\ldots,zp^*\}$, and each $\mathbf{z}^*(k)$ along the periodic orbit is structured in the following cyclic fashion:

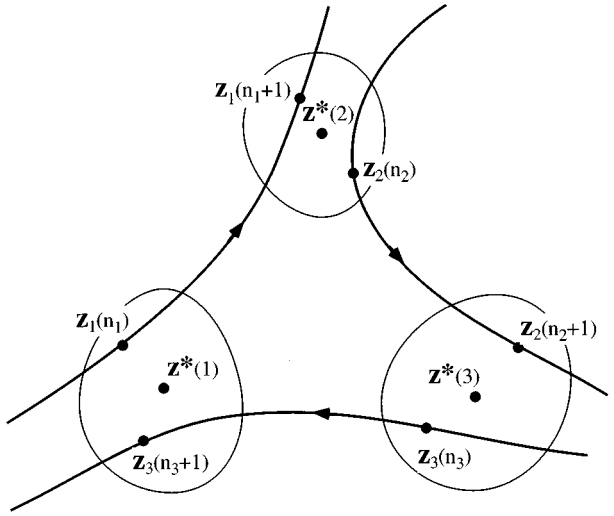

FIG. 6. A schematic representation of three short pieces of trajectories $\{\mathbf{z}_1(n_1),\mathbf{z}_1(n_1+1)\}$, $\{\mathbf{z}_2(n_2),\mathbf{z}_2(n_2+1)\}$, and $\{\mathbf{z}_3(n_3),\mathbf{z}_3(n_3+1)\}$, visiting the local regions of a period 3 orbit, $\{\mathbf{z}^*(1),\mathbf{z}^*(2),\mathbf{z}^*(3)\}$, at three nonconsecutive times n_1,n_2, and n_3.

$$\mathbf{z}^*(k) \equiv (z_1^*, z_2^*, \ldots, z_d^*)^\dagger = (\underbrace{z1^*, z2^*, z3^*, \ldots, zp^*, z1^*, \ldots}_{d})^\dagger, \tag{27}$$

with

$$\mathbf{z}^*(k+1) = (\underbrace{zp^*, z1^*, z2^*, \ldots, zp^*, z1^*, \ldots}_{d})^\dagger. \tag{28}$$

Here, we arbitrarily associate the first component of $\mathbf{z}^*(k)$ with $z1$. At each point $\mathbf{z}^*(k)$ on the periodic orbit, we can again approximate its local dynamics by the following Taylor series expansion:

$$\mathbf{F}(\mathbf{z})=\mathbf{z}^*(k+1)+\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})\cdot[\mathbf{z}-\mathbf{z}^*(k)]+O(|\mathbf{z}-\mathbf{z}^*(k)|^2).$$

Using a collection of temporal or spatial neighbors of $\mathbf{z}$ as described in the preceding section (see Sec. II E), the Jacobian term $\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})$ can again be calculated using least square fit methods.

To get the full description of the local dynamics along the entire periodic orbit, we consider a set of p points $\{\mathbf{z}_k\}$, $k=1,\ldots,p$, chosen from the delay reconstructed data set $\{\mathbf{z}(n)\}$. These points are chosen to be close to the periodic orbit, i.e., $|\mathbf{z}_k-\mathbf{z}^*(k)|$ is small. Note that the points $\{\mathbf{z}_k\}$ are not assumed to be consecutive in time. By going through one cycle along the periodic orbit and neglecting terms $O(|\mathbf{z}_k-\mathbf{z}^*(k)|^2)$, we then have

$$\begin{aligned}
\mathbf{F}(\mathbf{z}_1)&=\mathbf{z}^*(2)+\boldsymbol{\nabla}\mathbf{F}(\mathbf{z}_1)\cdot[\mathbf{z}_1-\mathbf{z}^*(1)],\\
\mathbf{F}(\mathbf{z}_2)&=\mathbf{z}^*(3)+\boldsymbol{\nabla}\mathbf{F}(\mathbf{z}_2)\cdot[\mathbf{z}_2-\mathbf{z}^*(2)],\\
&\vdots \qquad\qquad \vdots\\
\mathbf{F}(\mathbf{z}_p)&=\mathbf{z}^*(1)+\boldsymbol{\nabla}\mathbf{F}(\mathbf{z}_p)\cdot[\mathbf{z}_p-\mathbf{z}^*(p)].
\end{aligned}\tag{29}$$

One should note the cyclic relation in this set of p matrix equations. As in Eq. (2) of the fixed point method, Eq. (29) can be viewed as p simultaneous matrix equations for the periodic orbit $\mathbf{z}^*(1),\mathbf{z}^*(2),\ldots,\mathbf{z}^*(p)$. All collections of points $\{\mathbf{z}_k\}_{k=1,\ldots,p}$ sufficiently close to the periodic orbit will satisfy Eq. (29) to linear order.

Now, we wish to define our ''periodic orbit'' transform $\mathcal{G}(\mathcal{Z},\mathcal{R})\equiv\hat{\mathcal{Z}}$ as a mapping which takes a dp-dimensional vector defined by $\mathcal{Z}\equiv(\mathbf{z}_1,\ldots,\mathbf{z}_p)$ into another dp-dimensional vector $\hat{\mathcal{Z}}$,

$$\hat{\mathcal{Z}}\equiv\mathcal{G}(\mathcal{Z},\mathcal{R})=\begin{pmatrix}
-\mathbf{S}(\mathbf{z}_1,\mathbf{z}_2,\mathbf{R}_1) & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0}\\
\mathbf{0} & -\mathbf{S}(\mathbf{z}_2,\mathbf{z}_3,\mathbf{R}_2) & \mathbf{I} & \cdots & \mathbf{0}\\
\vdots & \vdots & \vdots & \cdots & \vdots\\
\mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{S}(\mathbf{z}_p,\mathbf{z}_1,\mathbf{R}_p)
\end{pmatrix}^{-1}\cdot\begin{pmatrix}
\mathbf{F}(\mathbf{z}_1)-\mathbf{S}(\mathbf{z}_1,\mathbf{z}_2,\mathbf{R}_1)\cdot\mathbf{z}_1\\
\mathbf{F}(\mathbf{z}_2)-\mathbf{S}(\mathbf{z}_2,\mathbf{z}_3,\mathbf{R}_2)\cdot\mathbf{z}_2\\
\vdots\\
\mathbf{F}(\mathbf{z}_p)-\mathbf{S}(\mathbf{z}_p,\mathbf{z}_1,\mathbf{R}_p)\cdot\mathbf{z}_p
\end{pmatrix},\tag{30}$$

where $\mathbf{I}$ is the $d\times d$ identity matrix, $\mathcal{R}=\{\mathbf{R}_1,\ldots,\mathbf{R}_p\}$ is a collection of p adjustable $d\times d\times d$ tensors which serves the same role as the ones in our fixed point transform (see Sec. II A), and the $d\times d$ matrix $\mathbf{S}(\mathbf{z},\mathbf{z}',\mathbf{R})$ is defined as

$$\mathbf{S}(\mathbf{z},\mathbf{z}',\mathbf{R})=\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})+\mathbf{R}\cdot[\mathbf{F}(\mathbf{z})-\mathbf{z}'].$$

Under the map $\mathcal{G}(\mathcal{Z},\mathcal{R})$, each of the p elements of the periodic orbit are mapped to themselves. In other words, if $\mathcal{Z}^*=(\mathbf{z}^*(1),\mathbf{z}^*(2),\ldots,\mathbf{z}^*(p))$, we have $\mathcal{G}(\mathcal{Z}^*,\mathcal{R})=\mathcal{Z}^*$ *independent* of $\mathcal{R}$.

Furthermore, for $\mathcal{Z}^*$ not exactly at a bifurcation point, we can show that $\boldsymbol{\nabla}\mathcal{G}(\mathcal{Z}^*,\mathcal{R})=\mathbf{0}$ *independent* of $\mathcal{R}$ [compare with Eq. (8)]. With a similar argument as in our fixed point transform, the transformed data cluster around the periodic orbit in the distribution function for $\hat{\mathcal{Z}}$. Therefore, an approximation to the periodic orbit can be found from the data set by looking for peaks in the distribution function $\hat{\mu}(\hat{\mathcal{Z}})$. As in the preceding section, the degree of clustering around the periodic orbit in the transformed data $\hat{\mathcal{Z}}$ depends on the effective size of the local region around the periodic orbit and on the frequency with which the periodic orbit is visited.

The issue of spurious peaks due to stationary points $\mathcal{Z}^s$ which are not periodic orbits $\mathcal{Z}^*$ of $\mathbf{F}(\mathbf{z})$ are handled as in our fixed point method (see Sec. II D). In brief, since the locations of the spurious peaks, $\mathcal{G}(\mathcal{Z}^s,\mathcal{R})$, will depend on the set of parameters $\mathcal{R}$ and $\mathcal{G}(\mathcal{Z}^*,\mathcal{R})=\mathcal{Z}^*$ *independent* of $\mathcal{R}$, we can eliminate spurious peaks due to $\mathcal{Z}^s\neq\mathcal{Z}^*$ by smearing using many different random $\mathcal{R}$ for each $\mathcal{Z}$ in our transformation.

So far, we have not specified how to choose the collection of the p test points $\{\mathbf{z}_k\}_{k=1,\ldots,p}$ for our periodic orbit transform. Note that $\{\mathbf{z}_k\}_{k=1,\ldots,p}$ need not be a *consecutive* sequence of points from the original data set. In fact, any combination of p points in the data set approximately satisfying the periodic orbit condition given by Eq. (29) will be mapped to the vicinity near the periodic orbit. This suggests an exhaustive way to generate the periodic orbit density by considering *all* combinations of p delay vectors in the data set. With a total of N delay vectors in our data set, this gives a total of $\Pi_{i=0}^{p-1}(N-i)$ permutations. Even for a moderate value of $N\approx1000$ with $N\gg p$, the number of computations needed to form the distribution function $\hat{\mu}(\hat{\mathcal{Z}})$ can be very large, i.e., $\sim N^p$. In some experimental situations, this relatively heavy computation load for periodic orbits with high periods might not be desirable.

An alternative grouping scheme, which we will refer to as the ''short'' scheme, with a smaller computational load, is possible. First, we begin with a temporally consecutive sequence of p data points as our backbone. Then, for each point on the backbone, we find its K closest spatial neighbors. These give p clusters of test points with $K+1$ members in each consisting of the backbone itself and all its K neighbors. We then apply the periodic orbit transform $\mathcal{G}(\mathcal{Z},\mathcal{R})$ to all combinations of $\{\mathbf{z}_k\}_{k=1,\ldots,p}$ constructed from the p clusters. By sliding the backbone of p consecutive points along the whole data set, the total number of possible sets $\{\mathbf{z}_k\}_{k=1,\ldots,p}$ of test points is $\sim N(K+1)^p$. For small values of K, the computation time in this ''short'' scheme is much less than the exhaustive grouping scheme stated earlier. One regains the exhaustive grouping scheme by increasing the number of neighbors K in each cluster. In the ''short'' scheme, if the sequence of p test points on the backbone is good (i.e., all p points on the backbone stay close to a particular periodic orbit), then all sequences formed by their *closest* neighbors will also be good and all combinations of $\{\mathbf{z}_k\}_{k=1,\ldots,p}$ formed from the p clusters will be mapped by $\mathcal{G}(\mathcal{Z},\mathcal{R})$ to the vicinity near the periodic orbit. Furthermore, even if only a subsequence of the points on the backbone stays close to the periodic orbit, other sequences of test points formed by their neighbors might be close enough to the periodic orbit for the entire cycle to be mapped by $\mathcal{G}(\mathcal{Z},\mathcal{R})$ to the vicinity of the periodic orbit. This situation is illustrated in Fig. 6 for a period three orbit. Intuitively, instead of blindly picking some random collection of p test points for the calculation, the ''short'' scheme selectively chooses the set of p test points that are most likely to be close to the periodic orbit. In the following calculations, we use the ''short'' scheme.

B. Extracting periodic orbits from the transformed data

As in the discussions of our fixed point extraction technique, different variant methods can be applied to the transformed data $\hat{\mathcal{Z}}$ to extract the periodic orbits. Recall that the transformed data $\hat{\mathcal{Z}}$ are in a special form $(\hat{\mathbf{z}}_1, \ldots, \hat{\mathbf{z}}_p)$, where $\hat{\mathbf{z}}_k (1 \leq k \leq p)$ is a d-dimensional vector representing the kth element of the candidate periodic orbit. Here, we will again look for peaks in the distribution function $\hat{\mu}(\hat{\mathcal{Z}})$ for the transformed data.

As in our fixed point method, a reduction procedure is possible for simplifying the distribution function $\hat{\mu}(\hat{\mathcal{Z}})$. Since our analysis is in delay coordinates, the periodic orbits satisfy the cyclic symmetry given by Eqs. (27) and (28). For a periodic orbit with period p, the cyclic symmetry defines a p-dimensional hyperplane which contains the pd-dimensional vector $\mathcal{Z}^*$. As a specific example with $p=2$ and $d=3$, $\mathcal{Z}^*$ is a six-dimensional vector and $\mathcal{Z}^* = (\mathbf{z}^*(1), \mathbf{z}^*(2)) = (z1^*, z2^*, z1^*, z2^*, z1^*, z2^*)$. The two-dimensional hyperplane with the given cyclic symmetry defined by Eqs. (27) and (28) is spanned by the following two vectors: (1,0,1,0,1,0) and (0,1,0,1,0,1). Numerically, for a given point $\hat{\mathcal{Z}}$, we can calculate its perpendicular distance from this hyperplane. If this distance is within a given tolerance, we consider it as a candidate period 2 periodic orbit. [Recall that in the case with period 1 fixed points, the reduction of $\hat{\mu}(\hat{\mathcal{Z}})$ is to restrict data to a small cross section tube along the diagonal with a given width.] In general for a period p orbit, we have a p-dimensional hyperplane within the pd-dimensional space and we consider data falling within a thin slab with a given thickness about this hyperplane.

These candidate periodic orbits in principle will have p independent components, and a p-dimensional distribution function $\tilde{\mu}(\tilde{\mathbf{z}})$ can be constructed by considering a p-dimensional vector in the form $\tilde{\mathbf{z}} \equiv (\hat{z}_1(1), \ldots, \hat{z}_1(p))$, where $\hat{\mathcal{Z}} = (\hat{\mathbf{z}}_1, \ldots, \hat{\mathbf{z}}_p)$ is a point in the thin p dimensional slab described above and $\hat{z}_1(k)$ denotes the first component of the vector $\hat{\mathbf{z}}_k$. Again, we have arbitrary chosen the first component of $\hat{\mathbf{z}}_k$ in this construction of $\tilde{\mathbf{z}}$; other choices can be made. The sharpness of the peak in this reduced distribution function is enhanced as the thickness of the slab is made smaller. However, smaller width requires a larger data set so that enough points fall in the slab.

To uniquely identify a periodic orbit, one needs to determine the values of its p independent components, $\{z1^*, \ldots, zp^*\}$ and the correct sequential ordering of these p elements. To simply look for peaks in the p-dimensional distribution function $\tilde{\mu}(\tilde{\mathbf{z}})$ will serve both of these purposes since the correct ordering of the p sequential pieces of the periodic orbit is preserved in the reduced vector $\tilde{\mathbf{z}}$. One can alternatively form a one-dimensional histogram $\hat{\rho}(\hat{z})$ with $\hat{z}$ being a scalar and $\hat{z} = \tilde{z}_k$ $(k = 1, \ldots, p)$ for all reduced data $\tilde{\mathbf{z}}$. Peaks in this histogram give all possible values of the p components of the periodic orbit. To infer the correct ordering of these p components, one can use the subset of $\mathbf{z}$ with $\mathcal{R} = 0$.

C. Numerical example: Skewed Hénon map

To demonstrate our detection method for unstable periodic orbit with $p > 1$, we use the skewed Hénon map:

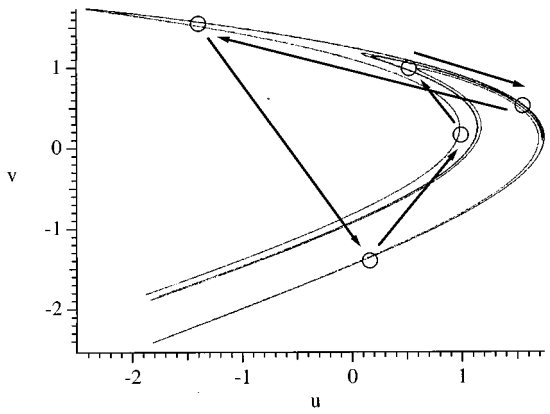

FIG. 7. A picture of the attractor for the skewed Hénon map Eq. (31) with $\beta_1 = 0.965$ and $\beta_2 = 0.25$. The attractor has a Lyapunov dimension of 1.261. The five circles denote the ''almost'' periodic regions and the arrows indicate the order in which these regions are visited.

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \mathbf{F}(u,v) = \begin{pmatrix} f(u,v) \\ u \end{pmatrix} = \begin{pmatrix} \beta_1 e^{\beta_2 u}(1.4 - u^2) + 0.3v \\ u \end{pmatrix}, \tag{31}$$

with $\beta_1 = 0.965$ and $\beta_2 = 0.25$. Again, the exponential term, $e^{\beta_2 u}$, is introduced so that the highest nonlinearity will not be trivially quadratic. The skewed Hénon system illustrated in Fig. 7 has a chaotic attractor with a Lyapunov dimension of 1.261. Using the periodic orbit detection method described above, we found all the periodic orbits on the attractor with periods $p = 1$ to 7.

Note that a complete set of periodic orbits with a given period p contains all cycles with period p and all other cycles with periods which are integral factors of p. As an example, a complete set of period 4 periodic orbits will contain all orbits with exactly four cycles, all orbits with exactly two cycles, and all period 1 fixed points.

Our results are summarized in Table I. In these calculations, we performed the periodic orbit transform $\mathcal{G}$ on a time series $\{x(n)\}$ constructed from 1024 iterates of $\mathbf{F}(u(n), v(n))$ with $x(n) = u(n)$. In the ''short'' scheme, the number of test points used in each cluster, including the backbone and its closest neighbors, was 2. Other operational parameters are as follows: the embedding dimension was 2; the number of random $\mathcal{R}$ used was 100; the magnitude of randomization κ was set at 3; and two spatial neighbors were used in estimating the Jacobian matrices $\nabla \mathbf{F}(\mathbf{z})$. In noisy situations, more neighbors would be needed. One can see from the table that results obtained using the periodic orbit transform compared well with ''exact'' values [calculated using Newton's method on the analytically given map $\mathbf{F}(u,v)$ [23]]. The discrepancy between our numerical results and the ''exact'' values is less than 0.2%.

Figures 8(a)–8(f) show the histogram approximations to $\hat{\rho}(\hat{z})$ used in extracting the periodic orbits for Table I. They correspond, respectively, to the cases with $p = 2,3,4,5,6,7$. As an example, Fig. 8(a) has four peaks in the histogram approximation to $\hat{\rho}(\hat{z})$ corresponding to the four identified cycle points of the periodic orbits. With $p = 2$, there is only one orbit with exactly two cycles, and two period 1 fixed

TABLE I. Table of periodic orbits found from a scalar time series generated by the skewed Hénon map.

Order of PO (p)	Total number of orbits	Number of orbits with exact cycle	Extracted values of $\mathbf{z}^*(1)$ from $\hat{\rho}(\hat{z})$ (± 0.002)	Exact numerical values [23] (± 0.0001)
1	2	2	(0.929,0.929)	(0.9302,0.9302)
			(−1.899,−1.899)	(−1.9030,−1.9030)
2	3	1	(1.300,−0.561)	(1.3014,−0.5604)
3	2	0		
4	3	1	(1.478,0.536)	(1.4785,0.5388)
5	2	0		
6	6	4	(−1.078,1.512)	(−1.0780,1.5125)
			(−1.924,1.667)	(−1.9240,1.6681)
			(−1.689,1.629)	(−1.6886,1.6292)
			(−0.959,1.411)	(−0.9579,1.4115)
7	7	6	(−2.158,1.705)	(−2.1580,1.7060)
			(−2.103,1.698)	(−2.1034,1.6989)
			(−1.539,1.602)	(−1.5386,1.6026)
			(−1.447,1.530)	(−1.4466,1.5322)
			(−1.156,1.527)	(−1.1551,1.5283)
			(−0.412,1.362)	(−0.4124,1.3618)

points. These give a total of $4=2+1\times 2$ peaks in the histogram approximation to $\hat{\rho}(\hat{z})$. The second fixed point peak at $\hat{z}=-1.903$ is small because it is not part of the attractor and a typical chaotic trajectory only intersects a small portion of its effective local region. There are two peaks in Fig. 8(b) for the case $p=3$ indicating the existence of the two period 1 fixed points and the absence of period 3 orbits. Figure 8(c) ($p=4$) has seven peaks corresponding to one orbit of period 1, one orbit of period 2, and one orbit of period 4: $7=1+1\times 2+1\times 4$. The period 1 in Fig. 8(c) occurs at $\hat{z}=0.929$ and the second fixed point off the attractor does not show up in this graph.

D. Detection of ''almost'' periodic orbit

In Fig. 8(d) ($p=5$), there is one strong peak corresponding to the period 1 fixed point at $\hat{z}=0.929$ and there are six additional broad humps in the histogram indicated by six arrows. The one near $\hat{z}=-1.903$ is again the period 1 fixed point located off the attractor. The remaining five humps correspond to an ''almost'' period 5 orbit with

$$\mathbf{F}^5(u,v)-\begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}f_1(u,v)-u\\f_2(u,v)-v\end{pmatrix}\approx\begin{pmatrix}0\\0\end{pmatrix}.$$

Geometrically, all period 5 orbits correspond to the intersection points of the two solution curves defined by $f_1(u,v)=u$ and $f_2(u,v)=v$ in the two-dimensional $(u-v)$ plane. With $\beta_1=0.965$, Fig. 9(a) is a graph of the two solution curves $f_1(u,v)=u$ (thinner curve) and $f_2(u,v)=v$ (thicker curve). They intersect at the two period 1 fixed points of the map at $\hat{z}=0.929$ and -1.903 indicated by open circles in Fig. 9. Although the skewed Hénon map does not possess a true period 5 orbit, the two solution curves nearly touch [see regions indicated by gray ovals in Fig. 9(a)] near the locations (denoted by arrows) where the five ''almost'' periodic humps appear in Fig. 8(d). These regions of close encounter between the two solution curves are indications of a coming saddle-node bifurcation as one or more parameters of the system are varied. A typical trajectory visiting near such ''almost'' periodic regions, nonetheless, for a short while behaves as if there were a periodic orbit. In fact, as the parameter β_1 is increased from 0.965 to 1.0, two period 5 orbits are born through a saddle-node bifurcation. Figure 9(b) is a graph of the two solution curves $f_1(u,v)=u$ (thinner curve) and $f_2(u,v)=v$ (thicker curve) near the bifurcation point at $\beta_1=0.9871$. In addition to the two already existing period 1 fixed points, now there are five other regions where the two solution curves touch (see regions indicated by gray ovals). Finally, Fig. 9(c) shows the crossing of the two solution curves at $\beta_1=1.0$ after the bifurcation. Two period 5 orbits are present and the total number of intersection points becomes 12 (two period 1 fixed points and two orbits with exactly five cycles).

IV. NOISY TIME SERIES

From our noiseless numerical examples in the preceding two sections, we have seen that one should expect singularities in $\hat{z}$ at the locations of the unstable periodic orbits. However, in a real experimental setting, time series are usually contaminated by dynamical and/or observational noise. In these cases, the observed singularities are blurred by small noise into maxima, and can even be completely washed out by large noise.

The technique of surrogate data [11] can be used to assess the reliability of the observed peaks. In this procedure, we produce a truly stochastic time series, surrogate data, with similar statistical properties to the original supposedly deterministic data. We then compare the results of applying our periodic orbit detection method to both the original data set and to the surrogate. In particular, we employ a *Gaussian scaled phase shuffle* [24] to produce the surrogate data. This

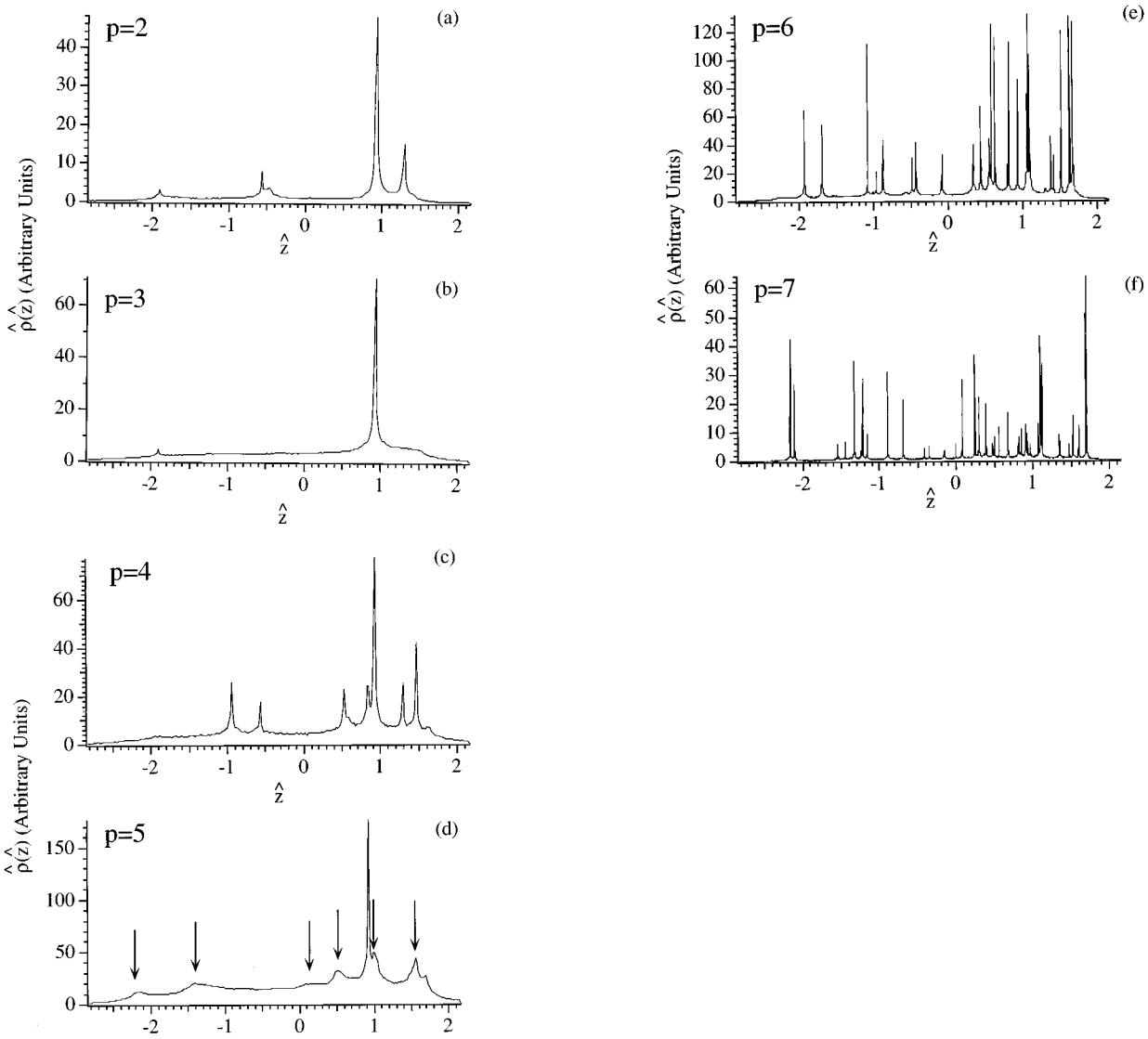

FIG. 8. Histogram plots of $\hat{\rho}(\hat{z})$ for the skewed Hénon map with (a) $p=2$, (b) $p=3$, (c) $p=4$, (d) $p=5$, (e) $p=6$, and (f) $p=7$. These plots are obtained using the second-order periodic orbit transform $\mathbf{G}_2$ with the ''short'' grouping scheme ($K=1$). Other operational parameters are as follows: the embedding dimension was 2, the number of random $\mathbf{R}$ values used was 100, and the magnitude of randomization $\kappa=3$.

technique assumes that the surrogate data arise from a linear stochastic process with the same amplitude distribution and approximately the same power spectrum as the original data. Since the surrogate data are stochastic, we do not expect any unstable periodic orbit structure that may be present in the original data set 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 $\hat{\rho}(\hat{z})$ could be modeled by the surrogates.

Numerically, for each realization of the surrogate data, we apply the same procedure for calculating $\hat{\rho}_{\mathrm{sur}}(\hat{z})$ as for our experimental data. Then, from this collection of $\{\hat{\rho}_{\mathrm{sur}}(\hat{z})\}$, we can estimate the ensemble average $\overline{\rho_{\mathrm{sur}}}(\hat{z})$ for each $\hat{z}$. Similar to our noisy experimental data, $\hat{\rho}_{\mathrm{sur}}(\hat{z})$ from each individual realization of the surrogates will fluctuate, and will consequently have fluctuation peaks which deviate from the mean $\overline{\rho_{\mathrm{sur}}}(\hat{z})$. Denote the deviation from the mean for a given surrogate by $w(\hat{z})=\hat{\rho}_{\mathrm{sur}}(\hat{z})-\overline{\rho_{\mathrm{sur}}}(\hat{z})$ 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'. The numerically determined fraction approaches the probability that W exceeds W' as the number of random surrogates increases. Since we use a large but finite number of surrogates, the actual deviation of the observed peak for a true periodic orbit might fall outside the range of maximum deviations calculated from the surrogates. In this case, one can say that the probability in observing this peak from the surrogates is ''unobservably'' small (for the number of surrogates tested) and we can quantify the distance of this exceptionally large deviation away from the bulk of the distribution $\Xi(W)$ by the ratio, $r_{\Xi}=W/W_0$, where W is the maximum deviation observed in the data and W_0 is a measure of the width of the distribution function $\Xi(W)$ with $\Xi(W_0)=0.5$.

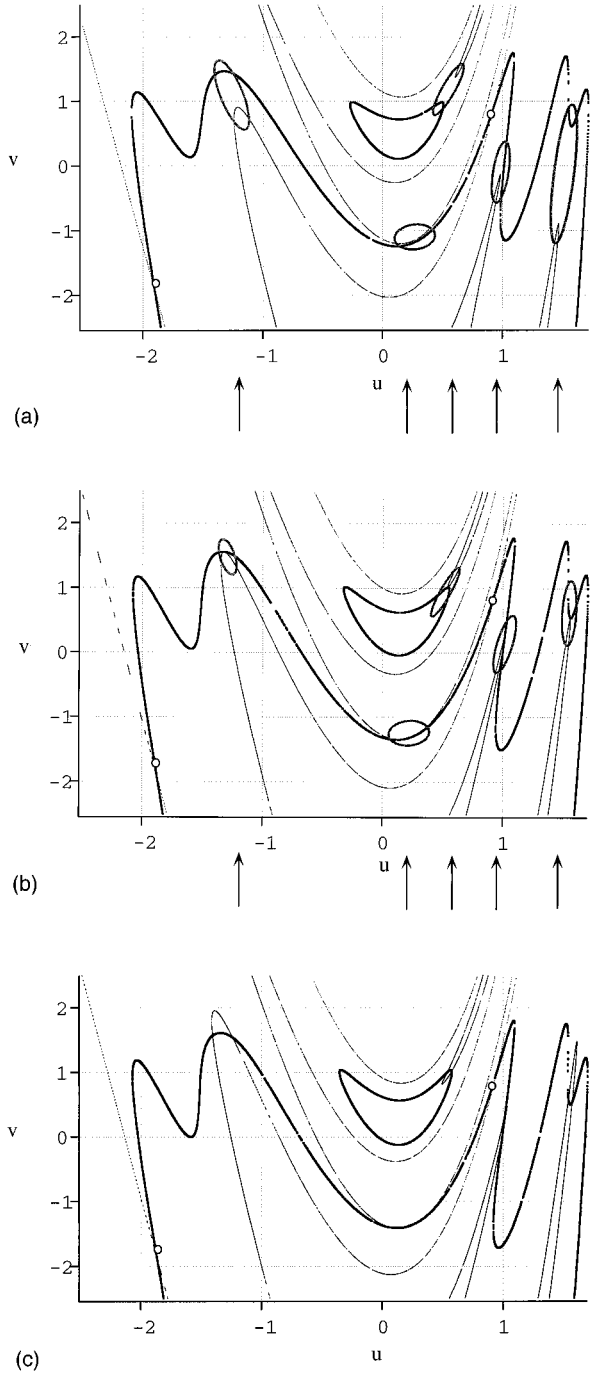

FIG. 9. Solution curves of $\mathbf{F}^5(u,v)-(u,v)=(f_1(u,v)-u, f_2(u,v)-v)=(0,0)$, where $\mathbf{F}(u,v)$ is the skewed Hénon map given by Eq. (31) for the cases (a) $\beta_1=0.965$, (b) $\beta_1=0.9871$, and (c) $\beta_1=1.0$. The thinner curve denotes the set of points on the (u,v) plane which satisfies the equation $f_1(u,v)=u$ and the thicker curve denotes the set of points on the (u,v) plane which satisfies the equation $f_2(u,v)=v$. The two period 1 fixed points of the map are labeled by the two white circles. The five gray ovals in (a) and (b) indicate the region where the two solution curves $f_1(u,v)=u$ and $f_2(u,v)=v$ nearly touch and the five arrows indicate the locations where the five ''almost'' periodic humps appear in Fig. 8(d). In (c), the function $\mathbf{F}^5(u,v)$ had passed through a saddle-node bifurcation and two period 5 orbits were created.

To demonstrate the robustness of our method in a noisy situation, we use the noisy Ikeda map as our numerical example,

$$\begin{pmatrix} u(n+1) \\ v(n+1) \end{pmatrix} = \begin{pmatrix} 1.0+a[u(n)\cos t(n)-v(n)\sin t(n)]+\epsilon_d\delta_d(n) \\ a[u(n)\sin t(n)+v(n)\cos t(n)]+\epsilon_d\delta_d(n) \end{pmatrix},$$

where $t(n)=0.4-b/[1+u^2(n)+v^2(n)]$, ϵ_d is the magnitude of the internal dynamical noise, and $\{\delta_d(n)\}$ is a uniformly distributed random variable in $[-1,1]$. To keep our presentation simple, we will focus only on the detection of period 1 fixed points using the first-order fixed point transform. The noiseless Ikeda map ($\epsilon_d=0$) describes the dynamics of a nonlinear optical cavity and the magnitude and angle of the complex quantity $u(n)+iv(n)$ is related to the amplitude and phase of the magnetic field of the nth light pulse inside the cavity [25]. With $a=0.75$, $b=9.0$, and $\epsilon_d=0$, this system has a chaotic attractor with a Lyapunov dimension of approximately 1.488, and it has an unstable fixed point at $(u^*,v^*)=(0.658,0.537)$. Furthermore, to simulate the effect of external observational noise, we choose $o(n)=v(n)+\epsilon_0\delta_0(n)$ to be our observed scalar output, where ϵ_0 is the magnitude of the external noise and $\{\delta_0(n)\}$ is a uniformly distributed random variable in $[-1,1]$. The delay coordinate vector in d dimensions is then given by $\mathbf{z}(n)\equiv(o(n),o(n-1),\ldots,o(n-(d-1)))^{\dagger}$. The total length of the time series used for this numerical experiments is 2048. A picture of the noiseless Ikeda attractor reconstructed from the delay vector $\mathbf{z}(n)$ in two dimensions is given in Fig. 10(a) along with plots (b) and (c) showing the attractor with two different levels of internal dynamical noise.

In noisy situations, one can often fine tune the statistical significance of the observed peak in $\hat{\rho}(\hat{z})$ with respect to the surrogates by adjusting the embedding dimension d, the number of random tensors $\mathbf{R}$ used, the magnitude of randomization κ, the number of nearest neighbors M used, and the number of surrogates used. In general, one expects improved statistics by using a larger number of random tensors $\mathbf{R}$ and more surrogates.

First, we will examine the effect of internal dynamical noise ($\epsilon_d\neq 0$ and $\epsilon_o=0$). Figure 11 corresponds to our results with four different observational noise levels: (a) $\epsilon_d=0$; (c) $\epsilon_d\sim 5\%$ of the radius of the attractor; (e) $\epsilon_d\sim 10\%$ of the radius of the attractor; and (g) $\epsilon_d\sim 30\%$ of the radius of the attractor. The thin solid curves give the histogram approximations to $\hat{\rho}(\hat{z})$ and the thick curves are the corresponding surrogate averages $\overline{\rho_{\mathrm{sur}}}(\hat{z})$. To quantify the statistical significance of the deviations between the data and the surrogate mean, we plotted the corresponding distribution functions $\Xi(W)$ in Figs. 11(b), 11(d), 11(f), and 11(h) with the arrows indicating the locations of the observed maximum deviations W found in Figs. 11(a), 11(c), 11(e), and 11(g). In all these calculations, the time series data were delay embedded in a four-dimensional space. A collection of $M=7$ spatial neighbors were used for estimating $\nabla\mathbf{F}(\mathbf{z})$ and 50 Gaussian scaled phase shuffled surrogates were em-

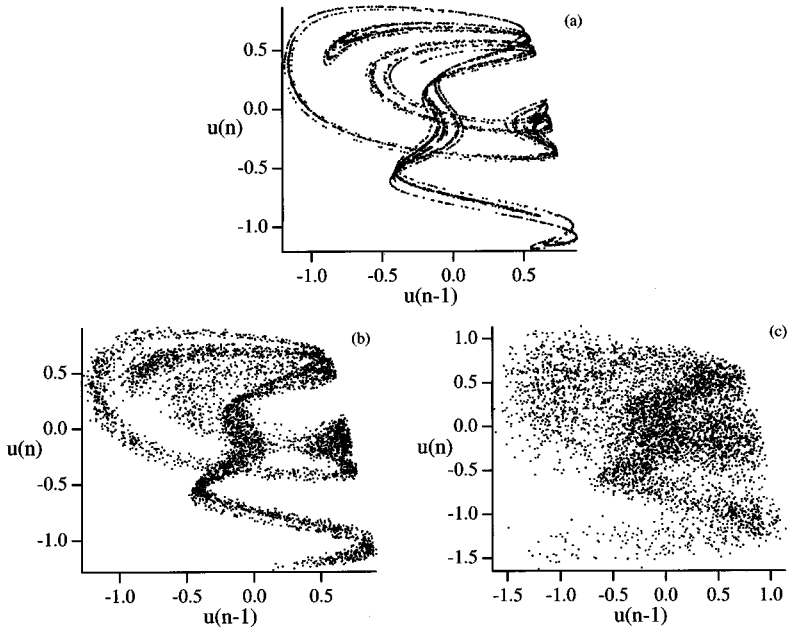

FIG. 10. Ikeda attractor in time delay coordinates: (a) with no noise; (b) with $\epsilon_d = 10\% \times$ (radius of attractor); (c) with $\epsilon_d = 30\% \times$ (radius of attractor).

ployed. Furthermore, in the process of randomization, we have used 500 different random **R** in Eq. (4) for each data point and κ was set at 8. In Fig. 11(a), the fixed point at $\hat{z} = 0.54$ has a strong peak rising sharply above the surrogate mean. As indicated by the arrow in Fig. 11(b), the probability for observing such a large deviation ($W \sim 2300$) from the 50 surrogates is ''unobservably'' small ($r_\Xi = 60.1$). As the amount of noise was increased to $\sim 5\%$ [Figs. 11(c) and 11(d)], the peak at the fixed point broadened. Nonetheless, the deviation of the peak above the surrogate mean with $W \sim 713$ is still significant. The probability of finding such a peak with the same large deviation in the surrogates is ''unobservably'' small with $r_\Xi = 13.2$. With a substantial amount of internal noise ($\sim 10\%$) [see Figs. 11(e) and 11(f)], the probability of finding the fixed point peak is still somewhat significant with $\Xi(W) \sim 6\%$. In the last case with 30% noise [Figs. 11(g) and 11(h)], the indication of the fixed point at $\hat{z} = 0.54$ disappeared and the maximum deviation observed in the data has an approximately 60% probability to be found in the surrogates.

The performance of our periodic orbit detection method is similar in situations with only external observational noise ($\epsilon_d = 0$ and $\epsilon_o \neq 0$). A similar sequence of eight graphs representing our results with four different observational noise levels: (a) $\epsilon_0 = 0$; (c) $\epsilon_0 \sim 5\%$ of the radius of the attractor; (e) $\epsilon_0 \sim 10\%$ of the radius of the attractor, and (g) $\epsilon_0 \sim 30\%$ of the radius of the attractor, are given in Figs. 12(a), 12(c), 12(e), and 12(g). Figures 12(b), 12(d), 12(f), and 12(h) are the corresponding plots of the distributions $\Xi(W)$ for each cases. All operational parameters are the same as in the previous Ikeda example with internal dynamical noise. The observed peaks in the midnoise levels (c) and (e) are statistically significant with ''unobservably'' small probabilities [$r_\Xi = 3.90$ for (d) and $r_\Xi = 2.96$ for (f)] in finding the same large peaks in the 50 random surrogates. In the last set of our graphs [Figs. 12(g) and 12(h)], we demonstrated that the location of the fixed point can again be completely washed out by large external noise ($\epsilon_0 \sim 30\%$).

The last example in this section is from an experiment of a gravitationally buckled, amorphous magnetoelastic ribbon (see Ref. [26] for a detailed description of the experimental setup). In this experiment, the ribbon was periodically driven by a magnetic field $H(t)$,

$$H(t) = H_{\rm dc} + H_{\rm ac}\cos(2\pi\nu t).$$

The time series was obtained optically by measuring the position of the ribbon near its base sampled at a constant phase of the periodic drive. A total of 1024 time series points were collected and a section of the time series is graphed in Fig. 13(c). The dimension of the experimental attractor is approximately 1.2 but in order to unfold all the crossings, delay coordinate data have to be embedded in a three-dimensional delay space. With the following operational parameters: 500 different random **R**, $\kappa = 5$, and 50 surrogates, the histogram for the first-order fixed point density $\hat{\rho}(\hat{z})$ is plotted in Fig. 13(a) as a thin solid curve and the surrogate mean $\overline{\rho_{\rm sur}}(\hat{z})$ is plotted as a thicker curve. Six spatial neighbors were used to estimate the Jacobian matrices $\nabla \mathbf{F}(\mathbf{z})$ at each point **z**. The dominant fixed point is located at approximately 5.40. Again, using 50 surrogates, we calculated the distribution, $\Xi(W)$, and we plotted its histogram approximation in Fig. 13(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 ''unobservably'' small ($r_\Xi = 17.9$). Thus, the observed fixed point structure in the experimental data is highly unlikely to have resulted from a linear stochastic process modeled by the surrogates.

V. QUASISTATIONARY DYNAMICS

With experimental time series, in addition to the complication of noise contamination, the problem of dealing with nonstationary system parameters is also important. This is especially true for biological time series where living tissues are constantly changing and global parameters are difficult to

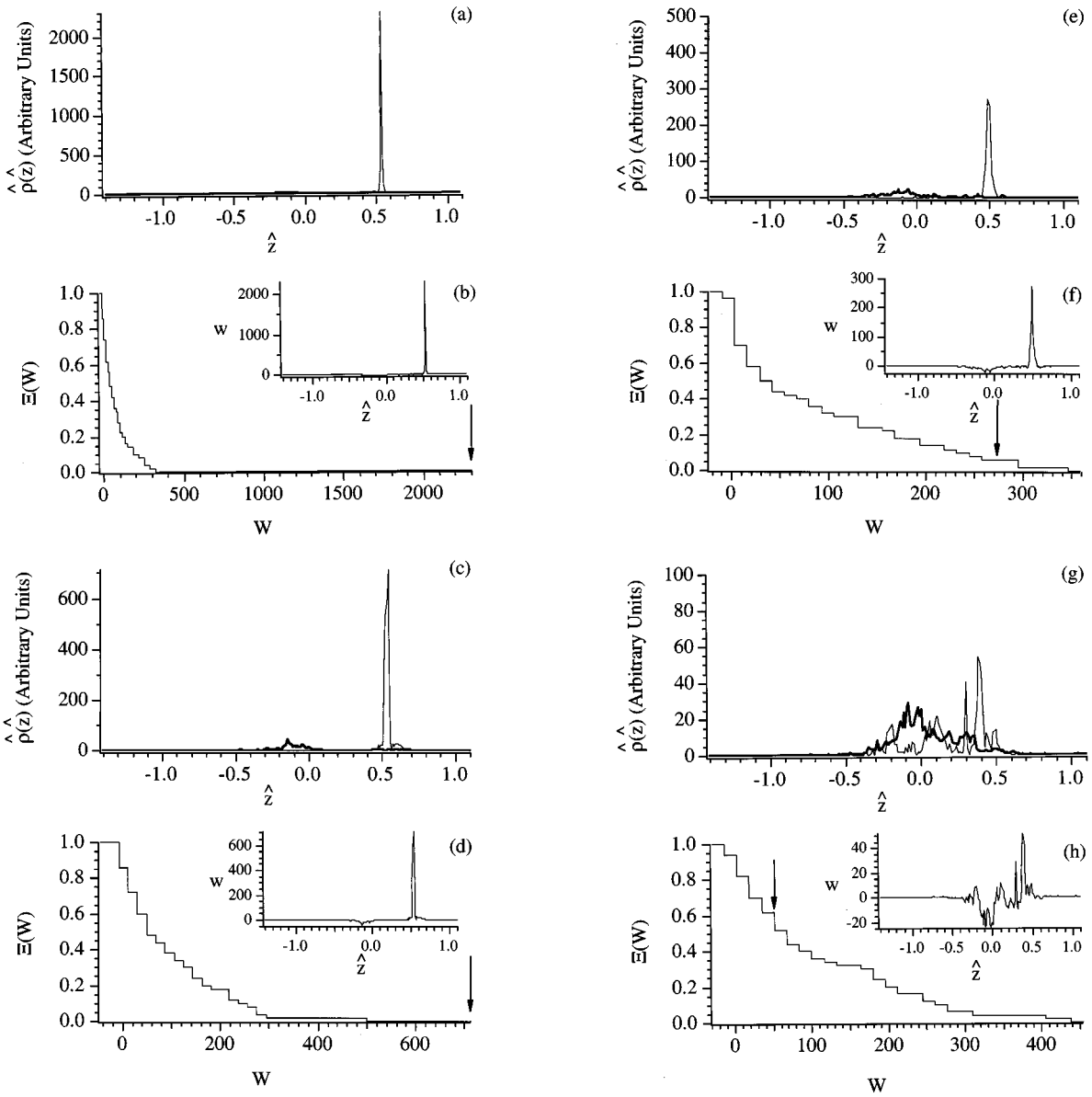

FIG. 11. Histogram plots of $\hat{\rho}(\hat{z})$ (thin curves) for time series generated from the Ikeda map with internal dynamical noise: (a) $\epsilon_d=0$; (c) $\epsilon_d=0.05\times$(radius of attractor); (e) $\epsilon_d=0.1\times$(radius of attractor); and (g) $\epsilon_d=0.3\times$(radius of attractor). The thicker curves are the corresponding surrogate means. 2048 data points were used, 500 different random **R** were used for each data point, 50 surrogates were used, $M=7$, and $\kappa=8$. (b), (d), (f), and (h) are histogram plots of $\Xi(W)$ in the four cases. Insets are the four corresponding plots of $[\hat{\rho}(\hat{z})-\overline{\rho_{\mathrm{sur}}}(\hat{z})]$ vs $\hat{z}$. Arrows indicate the values of maximum deviations between $\hat{\rho}(\hat{z})$ and the surrogate means.

control. In general, changes in system parameters are reflected in changes in the unstable periodic orbits. Here, we assume that the set of system parameters is slowly changing in time with respect to the natural time scale of the dynamics. In this quasistationary condition, one typically expects the periodic orbit structure of the underlining dynamics to vary continuously as a function of time. Obviously, when a system is operating near its bifurcation values, its periodic orbit structure might experience sudden changes. Near these bifurcation points, periodic orbits might be created or destroyed. Furthermore, for existing periodic orbits, their local stability and their associated manifold structures usually go through topological changes in bifurcations. If one blindly bins the *entire* transformed time series $\{\hat{z}(n)\}$ into the periodic orbit density $\hat{\rho}(\hat{z})$, real periodic orbits might be missed because of their constantly shifting locations [27]. A simple solution in dealing with this problem is to divide the entire time series into smaller windows and to perform the periodic orbit transform in each of the windows. The periodic orbit peaks in the histograms from each window should sharpen as the parametric changes within each time slice become small and therefore quasistatic. However, there will be a trade off with the number of data points available in each window; that is, the decreased number of data points in each slice, as the slice duration is decreased, might lower the statistical significance of the observed peaks. We find for the examples below that a good compromise between these factors is possible such that useful results are obtainable.

As stated earlier, as parameters are varied, the system might undergo topological changes through different bifur-

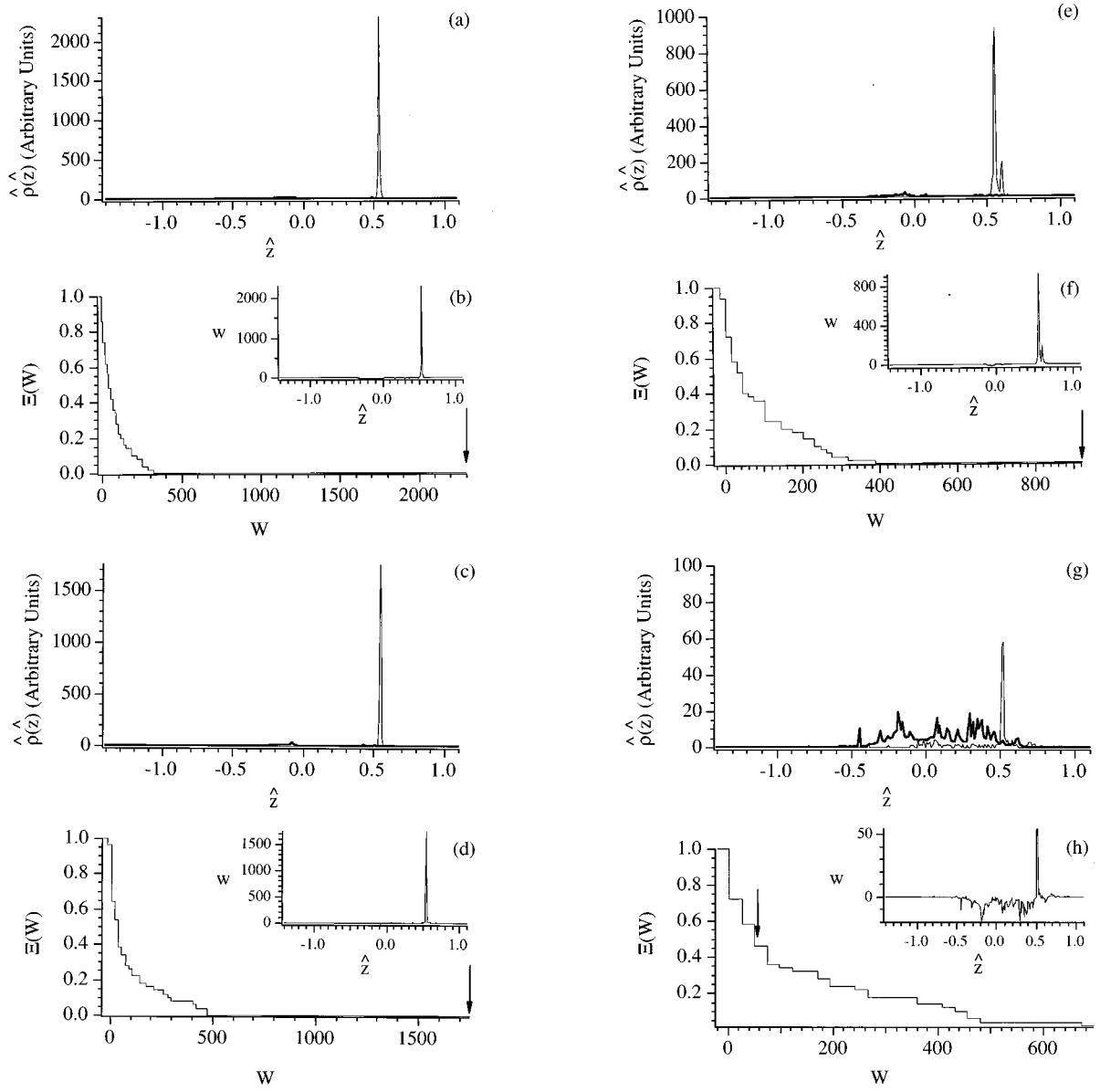

FIG. 12. Histogram plots of $\hat{\rho}(\hat{z})$ (thinner curves) for time series generated from the Ikeda map with external observational noise: (a) $\epsilon_0=0$; (c) $\epsilon_0=0.05\times$(radius of attractor); (e) $\epsilon_0=0.1\times$(radius of attractor); and (g) $\epsilon_0=0.3\times$(radius of attractor). The thicker curves are the corresponding surrogate means. 2048 data points were used, 500 different random **R** were used for each data point, 50 surrogates were used, $M=7$, and $\kappa=8$. (b), (d), (f), and (h) are histogram plots of $\Xi(W)$ in the four cases. Insets are the four corresponding graphs of $[\hat{\rho}(\hat{z})-\overline{\rho_{\rm sur}}(\hat{z})]$ vs $\hat{z}$. Arrows indicate the values of maximum deviations between $\hat{\rho}(\hat{z})$ and the surrogate means.

cations. A knowledge of how the structure of periodic orbits changes as a function of system parameters provides a clearer picture on the possible dynamical ranges of the system. Furthermore, from a practical point of view, the ability to characterize and to track a moving periodic orbit in time is critical in developing a successful control algorithm for the system [28].

To illustrate the functionality of our method in tracking unstable periodic orbits in an experimental setting, we again used the magnetoelastic ribbon system. In this case, the magnitude of the dc magnetic field $H_{\rm dc}$ was slowly varied sinusoidally through one complete cycle during the whole length of the experiment which consisted of approximately 1000 cycles of the ac periodic drive. Figure 14(a) is the histogram approximation to $\hat{\rho}(\hat{z})$ (thin solid curve) over the entire 1000 cycles of the ac drive using 500 random **R** with $\kappa=5$ and 50 surrogates for the entire time series. A section of the time series is plotted in Fig. 14(c). As in the preceding section, the time series was embedded in a three-dimensional delay space. A collection of $M=3$ temporal neighbors was used to determine the Jacobian matrices $\nabla\mathbf{F}(\mathbf{z})$. Again, the thicker curve is a histogram of the surrogate mean $\overline{\rho_{\rm sur}}(\hat{z})$. From this graph, one can see a broad peak with a few sharper features rising above the surrogate mean. The largest peaks near $\hat{z}=5.40$ and 5.36 are labeled as I and II, respectively, in Fig. 14(a). From the probability distribution function $\Xi(W)$, Fig. 14(b), one can see that these two largest peaks are significant with ''unobservably'' small probability [$r_\Xi=4.61$ for the one at 5.40 (I) and $r_\Xi=3.54$ for the one at 5.36 (II)] for them

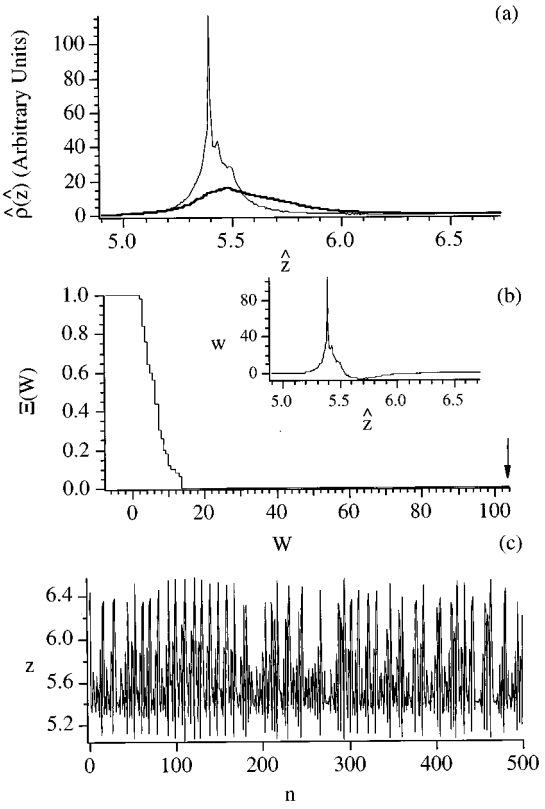

FIG. 13. (a) A histogram plot of $\hat{\rho}(\hat{z})$ (thinner curve) for the magnetoelastic ribbon data. 1024 data points were used, 500 different random **R** were used for each data point, 50 surrogates were used, $M=6$, and $\kappa=5$. The thicker curve corresponds to $\overline{\rho_{\mathrm{sur}}}(\hat{z})$. (b) is a histogram plot of $\Xi(W)$. Inset is the corresponding graphs of $[\hat{\rho}(\hat{z})-\overline{\rho_{\mathrm{sur}}}(\hat{z})]$ vs $\hat{z}$. Arrow indicates the value of maximum deviation between $\hat{\rho}(\hat{z})$ and $\overline{\rho_{\mathrm{sur}}}(\hat{z})$. (c) is a section of the actual time series data.

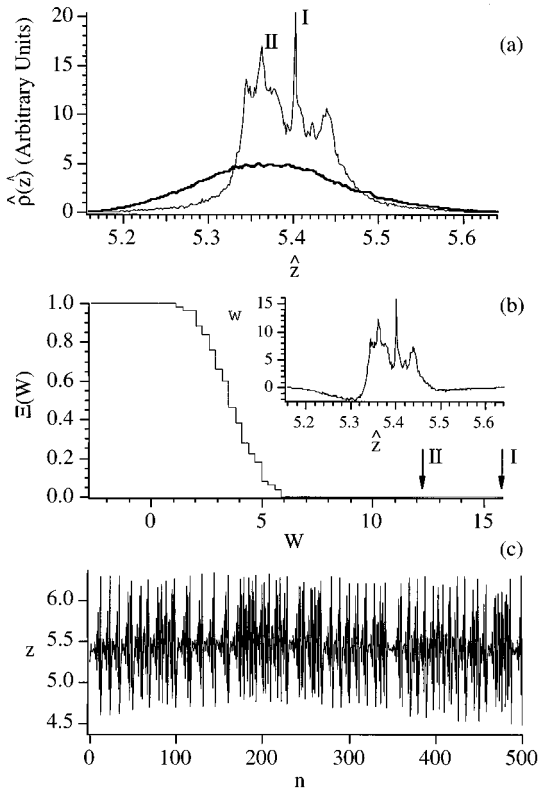

FIG. 14. (a) A histogram plot of $\hat{\rho}(\hat{z})$ (thinner curve) for the magnetoelastic ribbon data with varying H_{dc}. 977 data points were used, 500 different random **R** were used for each data point, 50 surrogates were used, $M=3$, and $\kappa=5$. The thicker curve corresponds to $\overline{\rho_{\mathrm{sur}}}(\hat{z})$. (b) is a histogram plot of $\Xi(W)$ and the inset is the corresponding graphs of $[\hat{\rho}(\hat{z})-\overline{\rho_{\mathrm{sur}}}(\hat{z})]$ vs $\hat{z}$. Arrows (I and II) indicate the value of maximum deviations between $\hat{\rho}(\hat{z})$ and $\overline{\rho_{\mathrm{sur}}}(\hat{z})$ at the two dominant sharp features on the broad peak. (c) is a section of the actual time series data.

to be found in the 50 surrogates. However, the broadness of the background feature prevents one from resolving the true locations of the fixed points further.

In Fig. 15, the whole time series was divided up into a sequence of 107 overlapping windows with 128 data points in each and the fixed point transform **G(z)** was performed on the 128 data points in each time slice. All of the operational parameters were the same as in Fig. 14. Using the probability distribution function Ξ calculated from the 50 surrogates, we assigned a probability to each of the values of the histogram $\hat{\rho}(\hat{z})$. These probabilities are represented by a gray scale in Fig. 15 with black corresponding to ''unobservably'' small probability of observing that histogram value in the surrogates as the maximal value among all the bins, and white corresponding to probability near 1 of observing that histogram value as the maximal value in the surrogates. In this density plot, one can clearly see the track of the period 1 fixed point, indicated by the nearly black streak in the figure. As the dc magnetic field went through one sinusoidal cycle during the experiment, the location of the fixed point swept through the range from 5.35 to 5.45 also approximately in a sinusoidal fashion. In most of the time slices, the observed deviations between the peaks and the surrogate mean are significant with an ''unobservably'' small value of $\Xi(W)$. The peaks in Fig. 14(a) evidently come from averaging the drifting result (Fig. 15) over time. We emphasize, however, that the time plot, Fig. 15, shows that there is only *one* (drifting) peak.

As another example application to real data, we consider a sequence of interburst time intervals [29] obtained from a rat hippocampal slice (see Ref. [30] for a detailed description of the experiment). (A detailed report on the first identification of periodic orbit structures in the rat hippocampus slice and in electroencephalogram signals from a human epileptic focus will appear in Ref. [31].) The 1834 data points collected were obtained during a 25 min recording. The environmental variables such as temperature, oxygen level, potassium concentration, etc. were kept constant. As we will see later from our results, even with all environmental variables being kept constant, the dynamics in this neural ensemble demonstrated both spontaneous changes and slow drifts within the 25 min time span of the experiment. In constructing Fig. 16, we subdivided the entire data set into 54 overlapping windows with 128 data points in each. For each time slice, the 128 data points were delay embedded in a two-dimensional delay space. Using the fixed point transform **G(z)** with 300 random **R** tensors, $M=2$, and $\kappa=0.05$, we generated 54 histo-

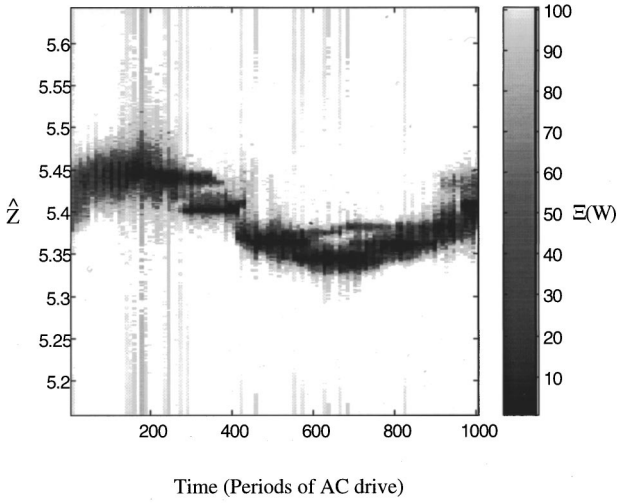

FIG. 15. Results for partitioned time series. The time series was from the magnetoelastic ribbon data with varying H_{dc} (the same as in the previous figure) but we partitioned the entire time series up into 107 overlapping time slices with 128 data points in each. The graph is a gray scaled density plot showing the probabilities of finding the observed deviations in the experimental $\hat{\rho}(\hat{z})$ in each time slice. All operational parameters were the same as in Fig. 14. The black color corresponds to near zero probability in observing that value as the maximum deviation in the surrogates and the white color corresponds to probability near 1 in observing that value as the maximum deviation in the surrogates. The black streak in the figure indicates the track of the period 1 fixed point as H_{dc} varies in time.

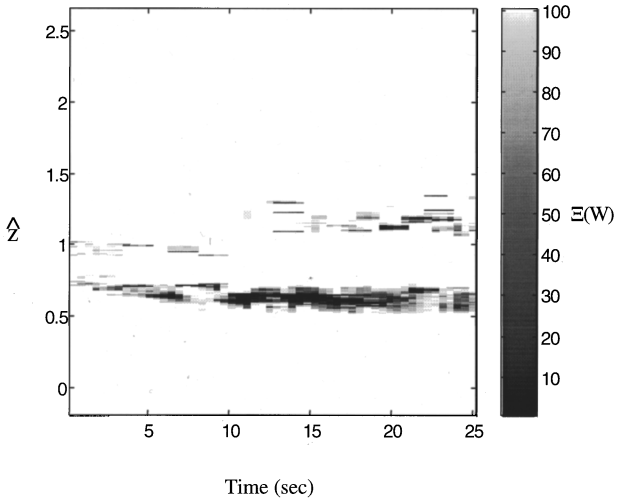

FIG. 16. Results for partitioned time series. The sequence of data points consists of 1834 interburst time intervals collected from a rat hippocampal slice. We partitioned the entire data set up into 54 overlapping time slices with 128 data points in each. The graph is a gray scaled density plot showing the probabilities of finding the observed deviations in the experimental $\hat{\rho}(\hat{z})$ in each time slice. Within each time slice, the fixed point transform was used with the following operational parameters: 300 random **R**, 50 surrogates, $M=2$, and $\kappa=0.05$. The black color corresponds to near zero probability of observing that value as the maximum deviation in the surrogates and the white color corresponds to probability near 1 in observing that value as the maximum deviation in the surrogates. The two black streaks (one continuous and one discontinuous) in the figure correspond to the tracks of two possible period 1 fixed points.

gram approximations to $\hat{\rho}(\hat{z})$ as a function of the experimental time. Furthermore, in each time slice, with 50 Gaussian scaled surrogates, we calculated the distribution function $\Xi(W)$ which characterized the probability of observing a particular maximum deviation W between $\hat{\rho}(\hat{z})$ and the surrogate mean. Then, we used $\Xi(W)$ to assign a probability to each value of the histogram $\hat{\rho}(\hat{z})$. Figure 16 shows a gray scaled density plot of these probabilities as a function of the experimental time with black corresponding to near zero probability and white corresponding to probability near 1. From the density plot in Fig. 16, one can clearly identify a prominent period 1 fixed point near $\hat{z}\approx 0.60$ sec. During the entire duration of the experiment, this period 1 fixed point slowly moved downward from a burst rate of $\hat{z}\approx 0.75$ sec at the beginning of the experiment to a lower burst rate of $\hat{z}\approx 0.55$ sec at the end. Biologically, it is reasonable to speculate that the slow drift toward a slower burst rate is indicative of a gradually depolarizing neural ensemble.

Another interesting feature of this data set was the appearance of a hint of a secondary peak near $\hat{z}\approx 1.20$ at approximately 11 min into the experiment. This peak was less well defined than the one at $\hat{z}\approx 0.60$ but persisted for the remaining duration of the experiment.

There is no surprise that biological systems are nonstationary. However, as in the case when system-wide parameters are varying slowly, we have demonstrated that, using our method, it is still possible to obtain useful structural information on dynamics through their time varying periodic orbit structure.

VI. CONCLUSION

In summary, we have presented a general method to extract unstable periodic orbits from chaotic time series. The method is based on a transformation such that the natural measure in the transformed space is enhanced near the unstable periodic orbits.

Periodic orbits with high periods were found in numerically generated time series. Much to our surprise, the method was able to detect ''almost'' periodic regions in phase space which are precursors to true periodic orbits created by a tangent bifurcation.

The effectiveness of our method in noisy environments was demonstrated using both numerically generated time series and actual experimental data. In the case of the noisy Ikeda system, the method was able to statistically identify periodic orbits with a noise level (dynamical or observational) that was approximately 10% of the size of the attractor.

Lastly, the method was shown to be able to find and track unstable periodic orbits in quasistationary systems. By partitioning the time series into smaller overlapping windows, we were able to identify dynamical changes characterized by different periodic orbit structures in neural activity from the rat hippocampus. While the identification and characterization of unstable periodic orbits in chaotic systems provides a method to analyze their dynamical behaviors, the ability to track unstable periodic orbits in time is an essential step to

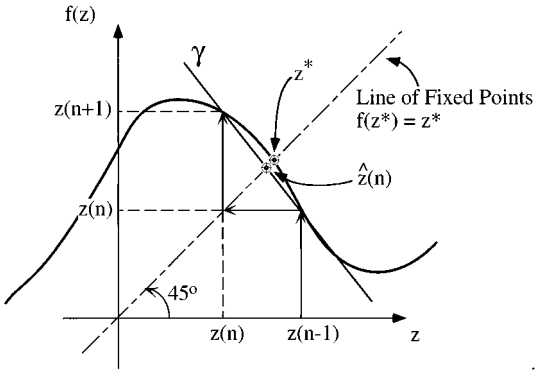

FIG. 17. Geometric interpretation of $G(z(n))$ in one dimension for the case $\kappa=0$.

control nonstationary chaotic systems with time varying system parameters.

ACKNOWLEDGMENTS

We acknowledge the financial support from the Keck Foundation for computing facilities. P.S., S.J.S., E.O., and B.G. received support from Office of Naval Research (N00014-95-1-0138 and N00014-90-J-1190), S.J.S. received support from National Institute of Mental Health (1-R29-MH50006-04) and the Children's Research Institute, and C.G. received support from the Department of Energy (Mathematical, Information, and Computational Sciences Division, High Performance Computing and Communications Program). P.S. is grateful to J. Jacobs and A. Smith for helpful conversations on density estimations.

APPENDIX A: GEOMETRIC INTERPRETATION OF G(z) WITH $\kappa=0$

In this section, we describe a geometric interpretation of the fixed point transform $G(z)$ for a one-dimensional map $z(n+1)=f(z(n))$ (see Fig. 17) with $\kappa=0$. Starting from the value $z(n-1)$ on the horizontal axis, one can obtain the iterates of the map $\{z(n-1),z(n),z(n+1)\}$ by following the straight arrows in Fig. 17. On this graph, the locations of the fixed points are at the intersections of the thick solid line with the diagonal, i.e., $f(z^*)=z^*$.

Our fixed point transform $G(z(n))$ consists of two parts. First, at a given point in the time series, $z(n),S(z(n))$ (with $\kappa=0$),

$$S(z(n))=\frac{z(n+1)-z(n)}{z(n)-z(n-1)},$$

gives the slope of the line segment γ in Fig. 17, which is an estimate of the slope of the function $f(z(n))$ at $z(n)$. Here, we have used a temporal neighbor $z(n-1)$ of $z(n)$ to estimate $f'(z(n))$. Equivalently, a spatial neighbor could have been used. Then, the second step is to substitute this estimated slope $S(z(n))$ into the fixed point transform, Eq. (3),

$$\hat{z}(n)=G(z(n))=\frac{z(n+1)-S(z(n))z(n)}{1-S(z(n))},$$

to evaluate the candidate fixed point $\hat{z}(n)$. One can interpret this operation as an algebraic representation of the geometrical construction of the intersection point $\hat{z}(n)$ between γ and the diagonal line in Fig. 17. In the case of a general rth-order fixed point transform $G_r(z)$, a similar least square fit procedure will give the parameters needed for the construction of an rth-order polynomial approximation to $f(z(n))$ at $z(n)$ and the corresponding rth-order fixed point transform gives the intersection points between this rth-order polynomial curve and the diagonal line.

[1] D. Auerbach *et al.*, Phys. Rev. Lett. **58**, 2387 (1987); P Cvitanović, *ibid.* **61**, 2729 (1988); R. Artuso, E. Aurell, and P. Cvitanović, Nonlinearity **3**, 325 (1990); C. Grebogi, E. Ott, and J. A. Yorke, Phys. Rev. A **37**, 1711 (1988).

[2] B. Hunt and E. Ott, Phys. Rev. Lett. **76**, 2254 (1996); Phys. Rev. E **54**, 328 (1996).

[3] E. Ott, C. Grebogi, and J. A. Yorke, Phys. Rev. Lett. **64,** 1196 (1990); F. Romeiras *et al.*, Physica D **58**, 165 (1992). For reviews see T. Shinbrot *et al.* [Nature **365**, 411 (1993)] and E. Ott and M. L. Spano [Phys. Today **48**(5), 34 (1995)].

[4] Some representative controlling chaos experiments where periodic orbits are determined from data occur in a driven magnetoelastic system [W. Ditto, S. N. Rauseo, and M. L. Spano, Phys. Rev. Lett. **65**, 3211 (1990)], lasers [Z. Gills *et al.*, *ibid.* **69**, 3169 (1992); C. Reyl *et al.*, Phys. Rev. E **47**, 267 (1993)], electric circuits [E. R. Hunt, Phys. Rev. Lett. **67**, 1953 (1991)], chemical reactions [V. Petrov *et al.*, Nature **361**, 240 (1993)], cardiac tissue [A. Garfinkel *et al.*, Science **257**, 1230 (1992)], a neuronal circuit from the brain [S. Schiff *et al.*, Nature **370**, 615 (1994)], and fluidized beds [S. Daw *et al.* (unpublished)].

[5] Periodic orbits are experimentally determined and related to ergodic properties of an experimental process in the papers of D. P. Lathrop and E. J. Kostelich [Phys. Rev. A **40**, 4028 (1989)] and of L. Flepp *et al.* [Phys. Rev. Lett. **67**, 2744 (1991)]. Also, J. C. Sommerer *et al.* used an experimentally determined periodic orbit to analyze a crisis in their system [Phys. Lett. A **153**, 105 (1991)].

[6] Recently D. Pierson and F. Moss [Phys. Rev. Lett. **75**, 2124 (1995)] and X. Pei and F. Moss [Nature **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.

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

[8] T. L. Carroll, Phys. Rev. E **52**, 5823 (1995).

[9] P. So *et al.*, Phys. Rev. Lett. **76**, 4705 (1996).

[10] It has recently come to our attention that a similar approach to

fixed point detection was considered by J. N. Glover in his Ph.D. dissertation (Statistical Laboratory, D.P.M.M.S., The University of Cambridge).

[11] J. Theiler *et al.*, Physica D **58**, 77 (1992).

[12] E. J. Gumbel, *Statistics of Extremes* (Columbia University Press, New York, 1958).

[13] For continuous time systems, a discrete representation of the dynamics can be obtained by a Poincaré surface of section. In the case when the dynamics is driven by a periodic signal, a stroboscopic surface of section can also be applied by sampling the continuous time signal at every period of the drive.

[14] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. Rand and L. S. Young (Springer-Verlag, Berlin, 1981), p. 230.

[15] 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).

[16] With $\mathbf{R}=\mathbf{0}$, we have $\mathbf{S}(\mathbf{z},\mathbf{0})=\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})$, and the transformation, Eq. (3), yields

$$[\mathbf{I}-\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})]\cdot\mathbf{G}(\mathbf{z})=\mathbf{F}(\mathbf{z})-\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})\cdot\mathbf{z}.$$

Because of the special form of the map $\mathbf{F}(\mathbf{z})$ in delay space [see Eq. (1)], the right-hand side of the equation gives a vector with a nonzero element only in the first entry, i.e.,

$$\mathbf{F}(\mathbf{z})-\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})\cdot\mathbf{z}=\begin{pmatrix} f(\mathbf{z}) \\ z_1 \\ z_2 \\ \vdots \\ z_{d-1} \end{pmatrix}-\begin{pmatrix} \partial_j f(\mathbf{z}) z_j \\ z_1 \\ z_2 \\ \vdots \\ z_{d-1} \end{pmatrix}$$

$$=\begin{pmatrix} f(\mathbf{z})-\partial_j f(\mathbf{z}) z_j \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where $j=1,\ldots,d$ and summing of repeated indices is implied. Furthermore, the factor $[\mathbf{I}-\boldsymbol{\nabla}\mathbf{F}(\mathbf{z})]$ on the left-hand side of the first equation will be in the following sparse form:

$$\begin{pmatrix} [1-\partial_1 f(\mathbf{z})] & -\partial_2 f(\mathbf{z}) & -\partial_3 f(\mathbf{z}) & \cdots & -\partial_{d-1} f(\mathbf{z}) & -\partial_d f(\mathbf{z}) \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}.$$

Putting the second and third expressions into the first, we have the following restrictions on the components of the transformed variables:

$$G_i=G_{i-1},$$

for $1<i\leq d$. Thus, the transformed variables $\hat{\mathbf{z}}=\mathbf{G}(\mathbf{z},\mathbf{0})$ live in a one-dimensional subspace spanned by the vector $(1,1,\ldots,1)^{\dagger}$.

[17] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, New York, 1993), p. 87.

[18] There could be other points $\mathbf{a}\neq\mathbf{0}$ mapping to $\hat{\mathbf{z}}(\mathbf{0})=\hat{\mathbf{z}}(\mathbf{a})$. (The mapping $\mathbf{G}:\mathbf{z}\rightarrow\hat{\mathbf{z}}$ need not be invertible.) In this case, $\hat{D}_p(\hat{\mathbf{z}}(\mathbf{0}))$ is $\min(D_p(\mathbf{0}),D_p(\mathbf{a}))$.

[19] $\mathbf{Q}(\mathbf{z})$ is degenerate only in the special case where the two roots of $\det(\boldsymbol{\beta}_1-\lambda\boldsymbol{\beta}_2)=0$ are the same, where $\boldsymbol{\beta}_1$ is the two by two matrix β_{ij1} and $\boldsymbol{\beta}_2$ is the two by two matrix β_{ij2}, $i,j=1,2$.

[20] Again, there could be other points $\mathbf{a}\neq\mathbf{0}$ mapping to $\hat{\mathbf{z}}(\mathbf{0})=\hat{\mathbf{z}}(\mathbf{a})$. In the unlikely case where $\mathbf{a}$ is also a stationary point of $\mathbf{G}(\mathbf{z})$, i.e., $\boldsymbol{\nabla}\mathbf{G}(\mathbf{a})=\mathbf{0}$, we have $\hat{D}_p(\hat{\mathbf{z}}(\mathbf{0}))=\min(D_p(\mathbf{0})/2,D_p(\mathbf{a})/2)$ and when $\mathbf{a}$ is a regular point of $\mathbf{G}(\mathbf{z})$, i.e., $\boldsymbol{\nabla}\mathbf{G}(\mathbf{a})$ being nonsingular, Eq. (13) holds.

[21] An example of these fixed points can be found in systems near a crisis (see J. C. Sommerer *et al.* in Ref. [5]).

[22] T. Sauer and J. Yorke (unpublished).

[23] H. E. Nusse and J. A. Yorke, *Dynamics: Numerical Explorations* (Springer-Verlag, New York, 1994).

[24] In Ref. [11], the *Gaussian scaled phase shuffle* is called the *amplitude adjusted Fourier transform* (see Ref. [11] for a detailed description of its construction procedure).

[25] S. Hammel, C. K. Jones, and J. Maloney, J. Opt. Soc. Am. B **2**, 552 (1985).

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

[27] With the same reasoning, in systems with slowly varying parameters, using temporal neighbors to estimate the values of $\boldsymbol{\nabla}\mathbf{F}$ for the Taylor expansion might be more appropriate than using spatial neighbors. In all the examples in Sec. V, we used temporal neighbors.

[28] B. J. Gluckman *et al.* (unpublished).

[29] T. Sauer, Phys. Rev. Lett. **72**, 3811 (1994).

[30] S. J. Schiff, K. Jerger, T. Chang, T. Sauer, and P. G. Aitken, Biophys. J. **67**, 684 (1994).

[31] P. So *et al.* (unpublished).