Detection and classification of nonlinear dynamic switching events

Christian Storm and Walter J. Freeman
Graduate Group in Biophysics, Life Sciences Addition, University of California, Berkeley, California 94720-3200
(Received 15 May 2002; published 21 November 2002)

A method is proposed for detecting chaotic switching events. Switching events are classified by the time of the event and by parameter value. Classifications are based on the density of localized dynamics about a test trajectory. This method is shown to be successful in tracking short-time parameter modulation and hyperchaotic key shifting used in otherwise secure communications.

DOI: 10.1103/PhysRevE.66.057202 PACS number(s): 05.45.Tp

The parametrization of a system abruptly changes to a new parametrization during a chaotic switching event. This reveals a sudden change in the system dynamics. Such changes can be subtle, as with chaotic shift key cryptographic systems [1], or substantial with the possibility of bifurcation or phase transition, as in the case of biological [2], electronic [3], or mechanical systems [4]. The detection and classification of switching events is integral to the understanding of nonstationary systems and deriving applications from those systems. For example, classifying pathological cardiac rhythms relies on the detection of distinctive changes in electrocardiographic signals [2], and the onset of a mechanical malfunction can be detected from changes in a machine's dynamic performance [4].

A wide variety of techniques have been developed to classify the dynamics of a system. A small cross section of those techniques includes unstable periodic orbits [5], clustering of dynamical similarity measures [6], chaotic synchronization [1], fuzzy logic [7], and spectral analysis [8]. Classification methods are often based on calculating a small number of statistical quantities that summarize the global behavior of the system thereby eliminating any chance of temporal resolution. Only a small subset of such classification techniques are suitable for the temporal localization of a switching event and for the classification of new parametrizations. This subset of techniques is successful at localizing switching events, but they have not been demonstrated to track a wide range of parameters. They also lack the ability to detect switching in high-dimensional systems.

A method is proposed which lends itself to the detection of a wide range of possible parametrizations and to the detection of slight dynamical differences. Rather than relying on summary statistical information, all of the available data is utilized to develop a model of an attractor's localized dynamics. Two characteristic features of an attractor are modeled with this method: (1) the density of state space points in a bounded region of state space and (2) the vector field which threads these states together into dynamical trajectories. We refer to this model as a weighted density of state model. Hively *et al.* utilized visitation frequencies of a system's symbolic dynamics to detect dissimilarities between parameterizations [9]. Multidimensional probability distributions of delay vector sets were used by Diks *et al.* to determine if two time series had been generated by the same system [10]. One drawback of these approaches concerns the need for long time series segments to measure a dissimilarity or distance between attractors thereby making them unsuitable for chaotic switching detection. The methodologies also do not utilize the information about the dynamics contained in the vector field. This hinders comparisons between attractors which essentially fill the same bounded region of state space but may differ dynamically.

The approach used in this study models the localized dynamics within a collection of attractors with switching between attractors. The main objective behind a weighted density of state model is to use all available system data to capture the behavior of the localized dynamics on the attractor. In particular, the model captures the continuous variation of density across the states comprising the attractor in a bounded region of state space and estimates the vector field that links these states together into trajectories.

Such a model forms the foundation for the ability to measure similarities between the localized dynamics of two systems. There are two steps in creating a weighted density of state model. The first step requires the formulation of a normalized density model of an attractor using either the full state vectors, or reconstructed state vectors, as the n-dimensional points in a kernel density estimate. There are a number of ways to create reconstructed vectors from scalar or multivariate time series data. Typically, time delayed values of a sampled continuous scalar signal x_t are used to form a state vector $\mathbf{a}_i=[x_i,x_{i-\tau},\ldots,x_{i-\tau(n-1)}]$ where the delay τ is some positive integer number. The time t and sample index i are related by $i=(t-t_0)\Delta s+1$ where Δs is the sampling rate and t_0 is the sampling start time. The dimension n has to be large enough to provide a proper reconstruction, or unfolding, of the dynamics. General rules for dimension are $n\geqslant 2d_b+1$ [11] and $n>d_c$ [12] where d_b and d_c are the box counting and correlation dimension, respectively.

In the second step the vector field $\mathbf{v}_{\mathbf{a}_i}$ is estimated at each state space point $\mathbf{a}_i$. A simple approximation takes the line segment connecting the point in question to the next point on the trajectory translated to zero $\mathbf{v}_{\mathbf{a}_i}=\mathbf{a}_{i+1}-\mathbf{a}_i$. The magnitude is calculated as the distance between the two points $\|\mathbf{v}_{\mathbf{a}_i}\|=\|\mathbf{a}_{i+1}-\mathbf{a}_i\|$ and the direction is defined by the coordinates of $\mathbf{v}_{\mathbf{a}_i}$.

Given M samples of a test signal y_t, a trajectory $\mathbf{b}_{j=1+\tau(n-1)}^{M}$ is constructed with the same reconstruction parameters, τ and n, used in forming the weighted density of state model for x_t. Note that it is assumed that the sampling rates for x_t and y_t are the same. Weighted densities can be determined at each point along the trajectory by using a modified kernel density estimate

$$\hat{f}(\mathbf{b}_j)=\frac{1}{[N-\tau(n-1)]h^n}\sum_{i=1+\tau(n-1)}^{N} w_{ij}K(u_{ij}), \quad (1)$$

where h is the kernel bandwidth parameter and $u_{ij}=(\mathbf{b}_j-\mathbf{a}_i)^T(\mathbf{b}_j-\mathbf{a}_i)/h^2$. The weight w_{ij} is defined as the normalized dot product of the vector fields

$$w_{ij}=\frac{\mathbf{v}_{\mathbf{a}_i}}{\|\mathbf{c}_{ij}\|}\cdot\frac{\mathbf{v}_{\mathbf{b}_j}}{\|\mathbf{c}_{ij}\|}=\frac{\|\mathbf{v}_{\mathbf{a}_i}\|}{\|\mathbf{c}_{ij}\|}\frac{\|\mathbf{v}_{\mathbf{b}_j}\|}{\|\mathbf{c}_{ij}\|}\cos\theta, \quad (2)$$

where the normalization $\|\mathbf{c}_{ij}\|=\max(\|\mathbf{v}_{\mathbf{a}_i}\|,\|\mathbf{v}_{\mathbf{b}_j}\|)$ making the weight bounded, $w_{ij}\leqslant\pm 1$. The two attributes of a vector, length and direction, appear in the weighting factor in Eq. (2). Each is an independent way of filtering out neighboring points which would otherwise contribute to the density. Weighting the individual density contributions reinforces the notion that each point in the test trajectory measures the local density of similar dynamics, and not just the presence of nearby state space points.

For this study the radially symmetric multivariate Epanechnikov joint probability function is used as the kernel [13]. Because it has finite support the resulting density distribution is bounded. The kernel is defined as

$$K(u)=\begin{cases}(2c_n)^{-1}(n+2)(1-u) & \text{if } u<1\\ 0 & \text{otherwise,}\end{cases} \quad (3)$$

where c_n is the volume of the unit sphere in n space.

A variety of techniques have been developed for selecting an ‘‘optimal’’ bandwidth h. Despite their different approaches, each method shares the same trade-off of under- or oversmoothing the density estimate when too small or too large a bandwidth h is used. This trade-off also applies to weighted density of state models. For weighted density of state models the objective is to find the localized density around each point in a trajectory such that, on average, there is a nonzero density between nearby orbits. Too small a bandwidth h can leave the density contributions of nearby and dynamically relevant trajectories out of range. Too large a bandwidth h will cause distant and less dynamically relevant trajectories to contribute to a query of an attractor's localized density. To accommodate these competing concerns the bandwidth h is calculated as the average Euclidean distance between k nearest neighbors which are not neighbors on the same trajectory segment. The value k is an empirically derived number. The best models use a small k value thereby minimizing the overlap with one another if two models have differing support in reconstruction space. However, if the dynamics are switched at relatively short time intervals, the trajectories will often reside in attractor basins rather than on the attractor itself. In this case, too small a bandwidth h, or number of nearest neighbors k, will result in a large fraction of zero density values. The ‘‘optimal’’ choice for the number of neighbors k is therefore application dependent and is derived heuristically.

Weighted density of state models were developed as a means of detecting switching events induced by changes in a known set of parametrizations. Let it be given that a scalar or multivariate signal from a system is known to be operating within a bounded region of parameter space comprised of a set of parametrizations $\{\mathbf{p}\}$. The goal is to detect transitions between parametrizations, $\mathbf{p}_n\rightarrow\mathbf{p}_m$, by finding which sequence of weighted density of state models best match the sequence of dynamical behaviors exhibited by the test signal. Detection and classification of a switching event relies on gathering evidence of which reconstruction space, i.e., parametrization, has the highest total density of similar dynamics to a test trajectory. The measurement is relative because the outcome relies on comparing the densities found for each weighted density of state model. Parameter classifications are made by taking the maximum likelihood of a moving average of m density values or the center of mass of density values. The center of mass, for a distribution of density values in parameter space for a point $\mathbf{b}_j$ on the trajectory,

$$\hat{\mathbf{p}}_j=\frac{\sum^{\mathbf{p}}\hat{f}_{\mathbf{p}}(\mathbf{b}_j)\mathbf{p}}{\sum^{\mathbf{p}}\hat{f}_{\mathbf{p}}(\mathbf{b}_j)}, \quad (4)$$

is used when a large set of parametrizations $\{\mathbf{p}\}$ is being explored and when the distance between parametrizations is small. The dynamics exhibited by nearby parametrizations can be very similar in the structurally stable sense, i.e., attractors within the set are topologically conjugate [14].

The time of the switching event is centered about the point where half of the time series data used in the moving average would be halfway in each parametrization. For example, with a delay coordinate spanning $(n-1)\tau$ sample points and a moving average of m density values, the moving average would be centered around the point $i'=i+[(n-1)\tau+m]/2$ where i is the earliest time sample used in the delay coordinate.

As a first demonstration of this methodology, we studied the Lorenz equation with a time-varying parametrization. The Lorenz equation [15]

$$\dot{u}=-10(u-v),$$

$$\dot{v}=c_t u-v-20uw,$$

$$\dot{w}=uv-2.666\,666\,666\,6w, \quad (5)$$

was made nonstationary by a stepped gain coefficient $c_t=[65+20\sin(4\pi\lfloor t/2\rfloor/80)]$. The unknown signal y_t was taken as the scalar signal u sampled at $\Delta s=100$ samples per second with $t\in[0,80]$ and initial conditions $\{1,1,1\}$ (see Fig. 1). Weighted density of state models were built for the parametrizations $\mathbf{p}=\{c\}=\{25,26,\ldots,90\}$ using the last 12 000 points of u from $t\in[0,130]$ from each initial condition $\{3,1,k\}$, $k=1,2,\ldots,10$. The same reconstruction parameters, $\tau=7$, $n=14$, and $h=6.0$, were used for each model. Because only local densities were measured, the ‘‘curse of

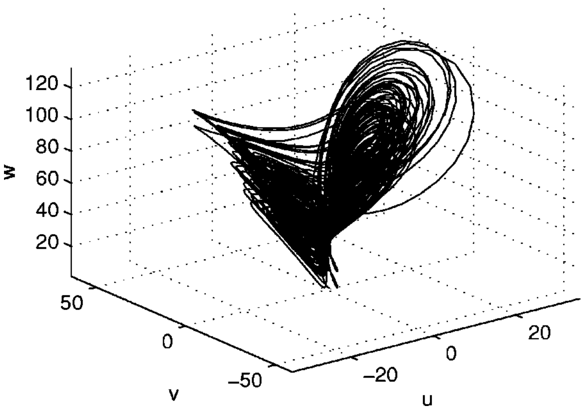

FIG. 1. State space of Lorenz system with sinusoidally varying parameter c_t. The state variable u was used as the unknown signal y_t.

dimensionality'' does not apply. The attractor and, at smaller scales, trajectory bundles form a compact subspace within this high dimension. A high dimension insured that similar dynamics remained close where dissimilar dynamics, i.e., false neighbors, diverged. The bandwidth h was chosen heuristically as the mean distance of the twelfth nearest neighbor for $c=90$, the largest attractor in terms of support. For comparison the mean distance for $c=25$ was approximately 2.

Using the same reconstruction parameters, the test signal y_t was transformed into a trajectory $\mathbf{b}_j$ in this reconstruction space. Using Eq. (1), weighted densities were determined for each point in the reconstructed trajectory for each model resulting in 66 time series of density values. A moving average of 50 center of mass estimates [Eq. (4)], is displayed in Fig. 2. Figure 2 demonstrates that the estimation $c_{t'}$ closely follows the true parametrization c_t with estimation errors closely scattered about c_t. These errors are caused by the trajectory taking time to converge to a new attractor after each switching event. With a relatively small shifting period

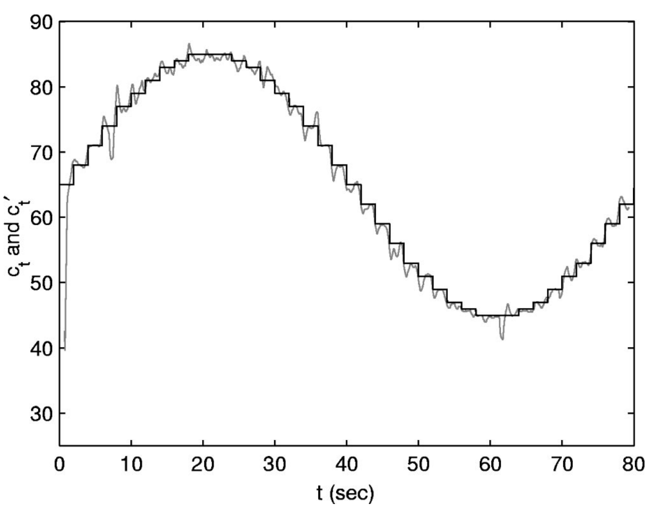

FIG. 2. Weighted density of state models were used to estimate the sinusoidally varying parameter c_t using the center of mass approach in Eq. (4). c_t and $c_{t'}$ are denoted by black and gray lines, respectively.

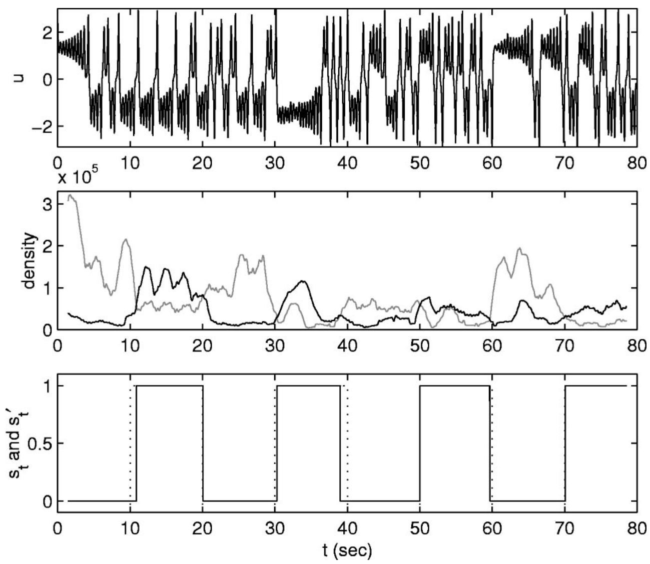

FIG. 3. Weighted density of state models broke a chaotic key shift code with a unidirectionally coupled Lorenz system as the transmitter. (a) The transmitted signal. (b) The average density of a 200-point (2 s) reconstructed trajectory segment compared against two weighted density of state models. $\{b\}=\{4.0\}$ and $\{b\}=\{4.4\}$ are denoted by gray and black lines, respectively. (c) The recovered (black line) binary message signal $s_{t'}$ found with maximum likelihood. The message signal s_t (dotted line) is included for comparison.

of 2 s the trajectory spent a relatively long time in the basin of the attractors it converged to after each switching event. Errors arise because neighboring attractors can occupy each others basin of attraction. This was best illustrated in the beginning with the estimations generated after the initialization of the trajectory with an initial condition far from the attractor. As the trajectory converged to the attractor, it moved through high density areas in a number of models.

In chaotic key shift code cryptography a binary message signal s_t is used to modulate a chaotic transmitter between two nearby parametrizations. To transmit the message one of the state variables is sent. The message signal is then decoded at the receiver through synchronization. Hyperchaotic chaotic systems are generally considered more secure since the geometric structure of an attractor is more complex [8,16–18].

To illustrate this, the authors of Ref. [8] demonstrated an unmasking technique in the frequency domain which failed for the unidirectionally coupled Lorenz system. The equation of state of the coupled Lorenz systems was given by

$$\dot{u}_1=-16(u_1-v_1),$$

$$\dot{v}_1=45.6u_1-v_1-20u_1w_1,$$

$$\dot{w}_1=5(0.9u_1+0.1u_2)v_1-bw_1,$$

$$\dot{u}_2=-16(u_2-v_2),$$

$$\dot{v}_2=45.6u_2-v_2-20u_2w_2,$$

$$\dot{w}_2 = 5u_2v_2 - bw_2, \quad (6)$$

where binary 0 and 1 were represented by the parametrizations $\mathbf{p}_0 = \{b\} = \{4.0\}$ and $\mathbf{p}_1 = \{b\} = \{4.4\}$, respectively, with a clock rate of 10 s. The state variable u_1 was sent as the transmission signal. This hyperchaotic system was chosen to demonstrate the ability to detect chaotic switching between nearby parametrizations in higher dimensional systems.

The test signal y_t with initial conditions $\{1,1,3,1,1,3\}$ and $t \in [0,80]$ was taken as the scalar time series u_1 sampled at $\Delta s = 100$ samples per second where the binary message is set to an alternating sequence of $\mathbf{p}_0$ and $\mathbf{p}_1$'s. The last 25 000 points of u_1 from $t \in [0,260]$, with initial conditions $\{3,1,k,3,1,k\}$, $k = 1,2,\ldots,10$, and the reconstruction parameters $\tau = 7$, $n = 14$, and $k = 4$ or $h = 0.39$, were used to create a weighted density model for each parametrization $\mathbf{p}_0$ and $\mathbf{p}_1$, using Eq. (1) the density at each point along the test trajectory for each model was calculated. The moving average density is shown in Fig. 3(b). The result of the maximum likelihood approximation, $s_{t'}$, is shown in Fig. 3(c). Since a reconstruction state space point can contain time series information from both attractors, the estimation of the time of the switching event tended to be blurred but generally near the actual switching time. Despite this, the hidden message is reliably intercepted.

The detection and classification of chaotic switching events is an important aspect of understanding nonstationary nonlinear systems. Many industrial, medical, and electronic applications rely on the detection of these changes in near real time. The methodology presented here is intended as a tool suitable for detecting such changes.

[1] U. Parlitz, L.O. Chua, L. Kocarev, K.S. Halle, and A. Shang, Int. J. Bifurcation Chaos Appl. Sci. Eng. **2**, 973 (1992).

[2] S. Sinha, Physica A **270**, 204 (1999); W.J. Freeman and L.J. Rogers, J. Neurophysiol. **87**, 937 (2002); K. Narayanan, R.B. Govindan, and M.S. Gopinathan, Phys. Rev. E **57**, 4594 (1998).

[3] J.Y. Ko, K. Otsuka and T. Kubota, Phys. Rev. Lett. **86**, 4025 (2001); T. Tsukamoto, M. Tachikawa, T. Sugawara, and T. Shimizu, Phys. Rev. A **52**, 1561 (1995).

[4] X. Wang and S. Noah, J. Vibr. Acoust. **120**, 596 (1998).

[5] D. Pingel, P. Schmelcher, F.K. Diakonos, and O. Biham, Phys. Rev. E **62**, 2119 (2000).

[6] T. Schreiber and A. Schmitz, Phys. Rev. Lett. **79**, 1475 (1997).

[7] W.J. Freeman, H.J. Chang, B.C. Burke, P.A. Rose, and J. Badler, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **44**, 989 (1997).

[8] T. Yang, L. Yang, and C. Yang, Physica D **124**, 248 (1998).

[9] L.M. Hively, P.C. Gailey, and V.A. Protopopescu, Phys. Lett. A **258**, 103 (1999).

[10] C. Diks, W.R. van Zwet, F. Takens, and J. DeGoede, Phys. Rev. E **53**, 2169 (1996).

[11] F. Takens, Lect. Notes Math. **898**, 366 (1981).

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

[13] B.W. Silverman, *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, London, 1986).

[14] M. Hirsch and S. Smale, *Differential Equations, Dynamical Systems and Linear Algerbra* (Academic, New York, 1974).

[15] E.N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).

[16] G. Grassi and S. Mascolo, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **46**, 1135 (1999).

[17] C. Zhou and T. Chen, Phys. Lett. A **234**, 429 (1997).

[18] T. Yang, L. Yang, and C. Yang, IEEE Trans. Circuits Syst., I: Fundam. Theory Appl. **45**, 1062 (1998).