Detecting dynamical changes in time series using the permutation entropy

Yinhe Cao,[1,3,*] Wen-wen Tung,[2,†] J. B. Gao,[3,‡] V. A. Protopopescu,[4,§] and L. M. Hively[4,‖]
[1]*BioSieve, San Jose, California 95117, USA*
[2]*National Center for Atmospheric Research, P.O. Box 3000, Boulder, Colorado 80307-3000, USA*
[3]*Department of Electrical and Computer Engineering, University of Florida, Gainesville, Florida 32611, USA*
[4]*Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6418, USA*
(Received 5 April 2004; published 27 October 2004)

Timely detection of unusual and/or unexpected events in natural and man-made systems has deep scientific and practical relevance. We show that the recently proposed conceptually simple and easily calculated measure of permutation entropy can be effectively used to detect qualitative and quantitative dynamical changes. We illustrate our results on two model systems as well as on clinically characterized brain wave data from epileptic patients.

DOI: 10.1103/PhysRevE.70.046217 PACS number(s): 05.45.Tp, 02.50.Fz, 87.19.La

I. INTRODUCTION

Detection of dynamical changes in complex systems is one of the most important problems in physical, medical, engineering, and economic sciences. Indeed, in meteorology, quantitative description of time and location of weather changes is crucial for accurate weather forecasting; in physiology and medicine, accurate detection of transitions from a normal to an abnormal state may improve diagnosis and treatment; in communications networks, robust and timely detection of anomalies, either due to hardware or software failure, or due to hacking, is crucial to maintain the network's integrity and functionality. Other important applications include earthquake prediction and detection of anomalous events leading to power outages in power grids or financial crashes.

During the last two decades, a number of interesting methods have been proposed to detect dynamical changes. They include, among others, recurrence plots [1] and recurrence quantification analysis [2,3], recurrence time statistics based approaches [4,5], space-time separation plots [6] and their associated probability distributions [7], metadynamical recurrence plot [8], statistical tests using discretized invariant distributions in the reconstructed phase space [9,10], cross-correlation sum analysis [11], and nonlinear cross prediction analysis [12]. Most of these methods are based on quantifying certain aspects of the nearest neighbors in phase space, and, as a result, are computationally expensive. Recently, Bandt and Pompe introduced the interesting concept of permutation entropy (PE), as a complexity measure for time series analysis [13]. The PE is conceptually simple and computationally very fast. These two features motivate us to explore whether this concept can be effectively used to detect dynamical changes in complex time series. We use two model systems, namely, a transient logistic map and a transient Lorenz system, and a number of clinically characterized EEG (brain wave) data to show that the PE can indeed be effectively used to detect bifurcationlike transitions from model-generated data as well as epileptic seizures from EEG data. We note that although the PE is expected to be closely related to the Kolmogorov and topological entropy [13], in certain dynamical systems they may not be equivalent [14]. Thus in general, PE based methods for detecting transitions may yield different results than previously used methods, such as Kolmogorov entropy [15] or Lyapunov exponents [16]. Moreover, our preliminary studies on seizure detection indicate that the PE based method is up to 100 times faster than a Lyapunov exponent based method [16], due to the fact that neighborhood searching is not needed. We postpone a systematic comparison among these different methods for a future study, and focus here on presenting and applying the PE based method for event detection.

The reminder of the paper is organized as follows. In Sec. II, we review the concept of PE and present our algorithm for detecting dynamical changes in time series. In Sec. III we apply the PE to detect state transitions in the two model systems mentioned above. In Sec. IV, we report epileptic seizure detection from EEG data. Section V contains the concluding remarks.

II. PERMUTATION ENTROPY: DEFINITION AND ALGORITHM FOR DETECTING DYNAMICAL CHANGES

As is well known, a dynamical system can be suitably represented and analyzed by using a symbolic sequence. Recently, permutation was introduced by Bandt and Pompe [13] as a convenient means of mapping a continuous time series onto a symbolic sequence. To illustrate the idea,

*Electronic address: contact@biosieve.com
†Author to whom correspondence should be addressed. Electronic address: wwtung@ucar.edu
‡Electronic address: gao@ece.ufl.edu
§Electronic address: protopopesva@ornl.gov
‖Electronic address: hivelylm@ornl.gov

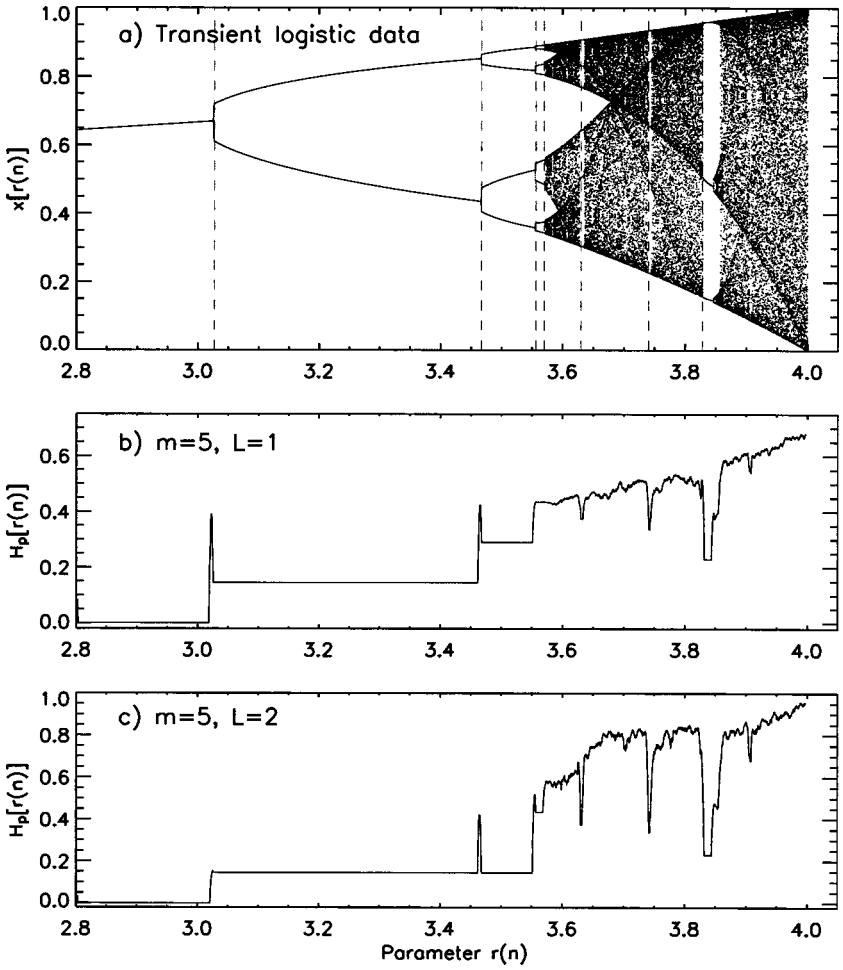

FIG. 1. (a) The transient logistic map data; (b),(c) variations of $H_p(r)$ with r for $m=5$, $L=1$ and $m=5$, $L=2$.

let us first embed a scalar time series $\{x(i), i=1,2,\ldots\}$ to a m-dimensional space [17]: $X_i=[x(i),x(i+L),\ldots, x(i+(m-1)L)]$, where m is called the embedding dimension and L the delay time. In their original paper, Bandt and Pompe [13] chose $L=1$. Since in practice the optimal L may be different from 1, we shall present the idea for any m and L. For a given, but otherwise arbitrary i, the m number of real values $X_i=[x(i),x(i+L),\ldots,x(i+(m-1)L)]$ can be arranged in an increasing order: $[x(i+(j_1-1)L) \leqslant x(i+(j_2-1)L) \leqslant \cdots \leqslant x(i+(j_m-1)L)]$. When an equality occurs, e.g., $x[i+(j_{i1}-1)L]=x[i+(j_{i2}-1)L]$, we order the quantities x according to the values of their corresponding j's, namely if $j_{i1}<j_{i2}$, we write $x(i+(j_{i1}-1)L) \leqslant x(i+(j_{i2}-1)L)$. For example, a vector whose components are all equal (i.e., $X_0=[x_0,x_0,\ldots,x_0]$) is mapped onto $[1,2,\ldots,m]$. Hence, any vector X_i is uniquely mapped onto $(j_1,j_2,\ldots,j_m)$, which is one of the $m!$ permutations of m distinct symbols $(1,2,\ldots,m)$. It is clear that each point in the m-dimensional embedding space, indexed by i, can be mapped onto one of the $m!$ permutations. When each such permutation is considered as a symbol, then the reconstructed trajectory in the m-dimensional space is represented by a symbol sequence. The number of distinct symbols can be at most $m!$. Let the probability distribution for the distinct symbols be $P_1,P_2,\ldots,P_K$, where $K \leqslant m!$. Then the PE for the time series $\{x(i), i=1,2,\ldots\}$ is defined [13] as the Shannon entropy for the K distinct symbols

$$H_p(m) = -\sum_{j=1}^{k} P_j \ln P_j. \tag{1}$$

When $P_j=1/m!$, then $H_P(m)$ attains the maximum value $\ln(m!)$. For convenience, we always normalize $H_p(m)$ by $\ln(m!)$, and denote

$$0 \leqslant H_p = H_p(m)/\ln(m!) \leqslant 1. \tag{2}$$

Thus H_p gives a measure of the departure of the time series under study from a complete random one: the smaller the value of H_p, the more regular the time series is. It is clear that if m is too small, such as 1 or 2, the scheme will not work, since there are only very few distinct states. In principle, using a large value of m is fine, as long as the length of a stationary time series under study can be made proportional to $m!$. However, since the purpose of the study is to *detect changes in signals*, too large a value of m, such as 12 or 15, is inappropriate. In their paper [13], Bandt and Pompe recommend $m=3,\ldots,7$. We often found that $m=3$ and 4 may

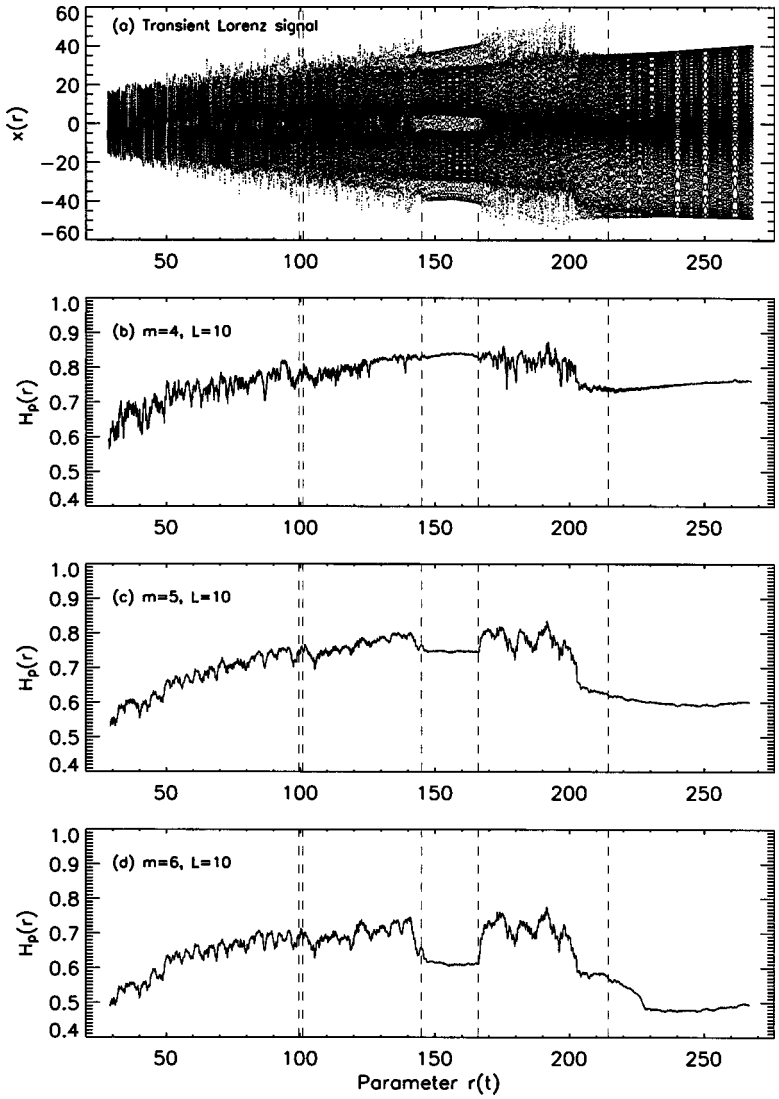

FIG. 2. (a) The transient Lorenz data; (b)–(d) variations of $H_p(r)$ with r for three different m. L is always 10.

still be too small, and a value of $m=5$, 6, or 7 seems to be the most suitable.

Our algorithm for the detection of dynamical changes in a time series can be described as follows: Partition a long time series into (overlapping or nonoverlapping) blocks of data sets of short length w, and compute H_p for each data subset. In the examples presented below, maximal overlapping (window shift by 1 time step) is used. We expect that the variation of H_p as a function of time or certain time-varying parameter can accurately indicate interesting dynamical changes in a time series. This is indeed so, as shall be shown by the examples in the following two sections.

Before proceeding, we comment on the selection of the window size w. Intuitively it is obvious that w should not be small. Otherwise, the statistics are not conclusive. However, if one's purpose is to accurately find transitional signals, then w should not be too large either. Upon using $w=512$, 1024, and 2048, we found consistently similar results. Hence, it appears that, within this range, the precise choice of the window size w is not critical. The examples presented in the next section are all analyzed for $w=1024$.

III. DETECTING DYNAMICAL CHANGES IN MODEL SYSTEMS

Our first example is the transient logistic map,

$$x_{n+1} = r(n)x_n(1 - x_n). \tag{3}$$

Following Trulla *et al.* [2], we first generate a transient time series $x(i)$, consisting of 120 001 points, by starting from $x_0=0.65$, $r(0)=2.8$ and consistently incrementing r in steps of 10^{-5} at each iteration. Figure 1(a) shows the resulting time series. To emphasize the fact that Fig. 1(a) is not the usual (asymptotically stabilized) bifurcation diagram, but the entire transient time series [one point for each $r(n)$ at the discrete time step n], we have denoted $r(n)$ instead of r under the x axis. We notice that although the overall pattern is very similar to the familiar bifurcation diagram for the logistic map, the parameter positions where transitions occur are slightly different from the well-known bifurcation points. For example, the period 2 bifurcation occurs at $r\approx 3.025$ instead of $r=3$. Such minute differences can be attributed to the transient character and "memory" effect of the system: the entire

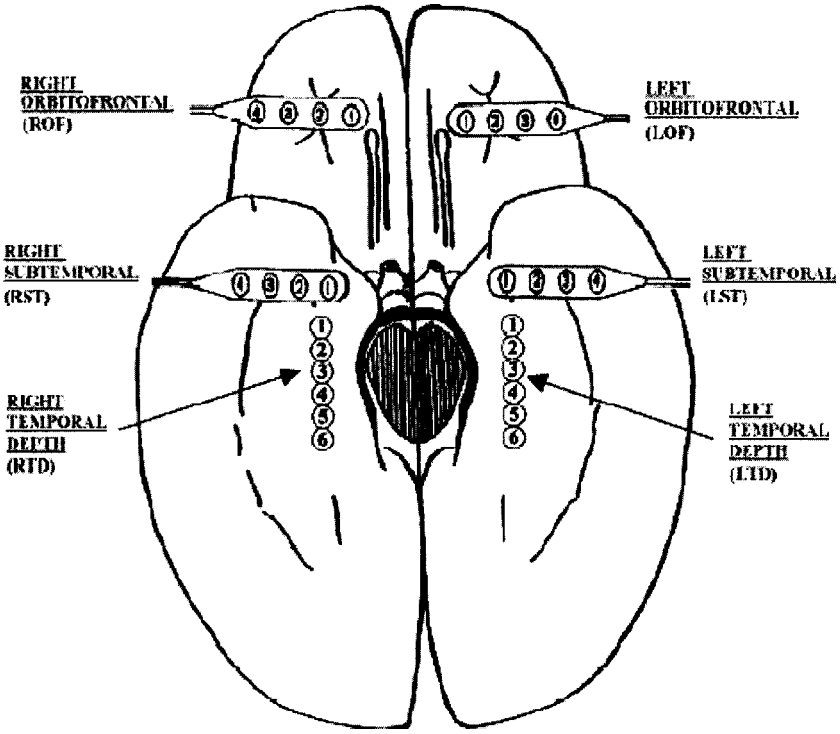

FIG. 3. Depth electrode placement diagram.

time series is generated by iterating Eq. (3) without discarding any points, hence, adjacent points in the time series are not only parameter value, but also time correlated. This correlation somewhat delays the onset of the actual bifurcation transitions. As illustrated in Figs. 1(b) and 1(c), this memory effect is responsible for the fact that the variation of the permutation entropy with the parameter r captures the transitions better when $L=2$ than when $L=1$.

Now let us check how well the PE can detect dynamical changes from the x_i time series. Figure 1(b) shows the variation of H_p with $r(n)$ when $m=5$, $L=1$. It is interesting to note that the $H_p(r)$ versus $r(n)$ curve correlates very well with the original time series, except that it misses the period-8 to period-16 bifurcation. When we choose $L=2$ instead, even this bifurcation is accurately identified, as is shown by Fig. 1(c). This can be clearly seen by comparing Figs. 1(a)–1(c) around $r \approx 3.56$.

Next we examine a continuous time system described by the following transient Lorenz equations [3]:

$$dx/dt = -10(x-y),$$

$$dy/dt = -xz + r(t)x - y, \tag{4}$$

$$dz/dt = xy - 8z/3.$$

The system is solved using a fourth-order Runge-Kutta method with a time step $\Delta t=0.01$. The parameter $r(t)$ is incremented from 28.0 to 268.0 by 0.002 at each integration step. The signal, of total length 120 001, one for each time step Δt, is shown simply by dots in Fig. 2(a). When Eq. (4) is solved with the parameter r fixed, theoretically, there are three periodic windows: $99.524 < r < 100.795$, $145 < r < 166$, and $r > 214.4$, as indicated by dashed vertical lines in Fig. 2(a). However, we emphasize again that the whole data set shown in Fig. 2(a) is a transient signal. In fact, due to transient nature of the system, the window $99.524 < r < 100.795$ is no longer periodic, hence, a good discriminating measure should not identify it as such.

To compute the PE, we chose $L=10$. Figures 2(b)–2(d) show the variations of $H_p(r)$ with $r(t)$ for $m=4$, 5, and 6, respectively. We note that all these graphs, especially those with $m=5$ and 6, capture the periodic windows very well. The reason that $m=4$ is not as good as $m=5$ or 6 lies in the fact that the state space with $m=4$ is characterized by $4!=24$ distinct states, hence, is not very large. Thus, for practical applications, we recommend $m=5$, 6 or 7 [the $H_p(r)$ versus $r(t)$ curve for $m=7$ is similar to those for $m=5$ and 6, and hence, not shown here]. It is clear that—to a certain extent—the choices of both m and L are related to the time step Δt.

Before we end this section, we comment that the PE can do more than detect bifurcationlike transitions. This is clearly indicated in Figs. 1 and 2 that the PE can vary considerably in chaotic windows. As we shall further show below, PE can vary significantly between epileptic seizures.

IV. EPILEPTIC SEIZURE DETECTION FROM EEG SIGNALS

Epilepsy is one of the most common disorders of the brain. Although epilepsy can be treated effectively in many instances, severe side effects have frequently resulted from constant medication. Even worse, patients may become drug resistant not long after being treated. To make medication more effective, timely detection of seizure is very important. In the past several decades, considerable efforts have been made to detect/predict seizure through analysis of continuous

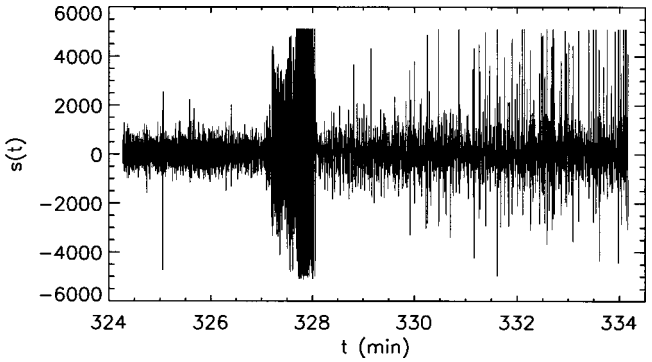

FIG. 4. An example of 10-minute long depth EEG signal.

EEG measurements. Representative nonlinear methods proposed include approaches based on correlation dimension [18–21], entropy [15], short time largest Lyapunov exponent [22,23], and dissimilarity measures [10,24].

We analyzed EEG signals recorded intracranially with approved clinical equipment by the Shands Hospital at the University of Florida. Such EEG signals are also called depth EEG, in contrast to scalp EEG. Depth EEG signals are less contaminated by noise or motion artifacts. Typically, a measurement is made with 28 electrodes (see Fig. 3).

Figure 4 shows a 10-minute duration EEG signal from one electrode. Signals with small amplitudes are considered normal background activities. The clinical equipment used to measure the data has a pre-set, unadjustable maximal amplitude, which is around 5300 μV. This causes clipping of the signals when the signal amplitude is higher than this threshold. This is often the case during seizure episodes, especially for certain electrodes. This is evident in Fig. 4 around minute 328 [which corresponds to the second seizure in Figs. 5(c)–5(e)]. To a certain extent, this clipping complicates seizure detection, since certain seizure signatures are not captured by the measuring equipment.

We studied multiple channel EEG signals of three patients. Each signal is more than 5 hours long, with a sampling frequency of 200 Hz. The PE was computed from each data set, with embedding dimension $m=5$, delay time $L=3$, and a window size of 2048 points. The value of L was determined according to an optimal embedding procedure described in Ref. [16]. The precise time of seizure onset was

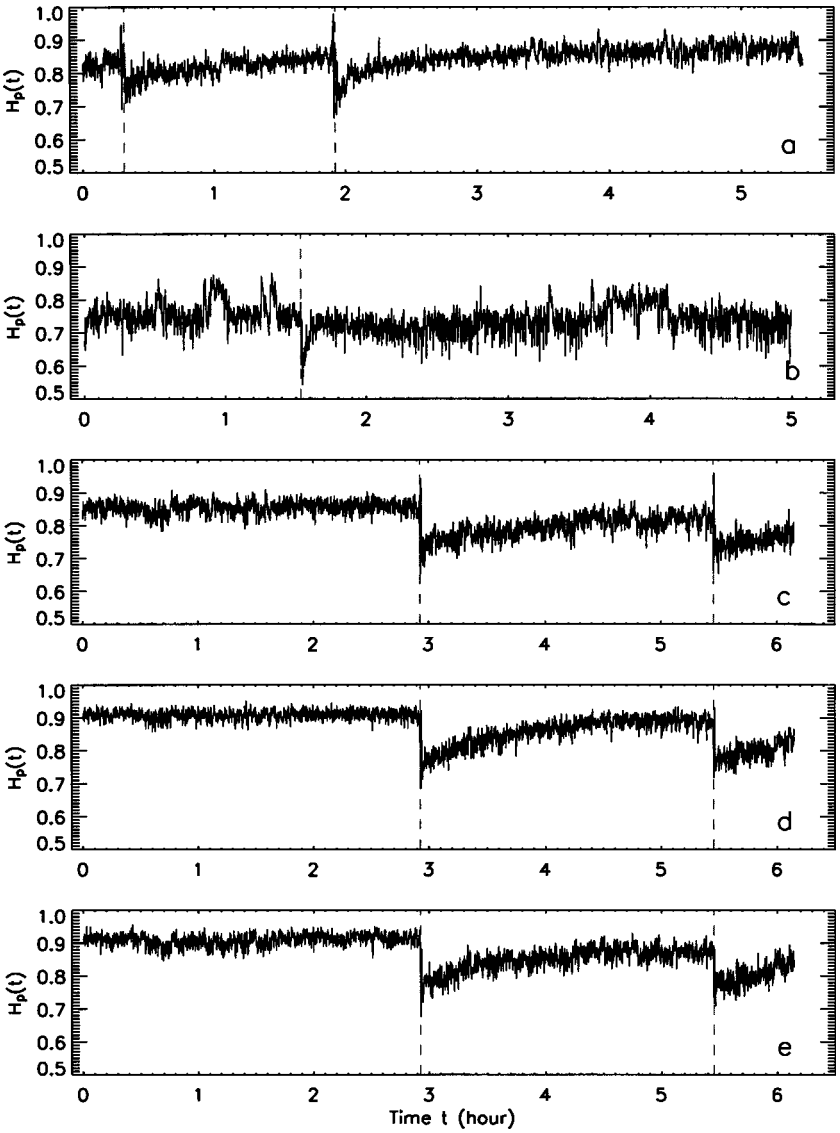

FIG. 5. Variations of H_p with time for EEG signals of (a) patient 1, channel LTD 1, (b) patient 2, channel LTD 1, and (c)–(e) patient 3, channels LTD 1–3.

determined by medical experts by viewing video tapes as well as the EEG signals and is indicated by dashed vertical lines in Figs. 5(a)–5(e). Seizures were associated with either abnormal running/bouncing fits, clonus of face and forelimbs, or tonic rearing movement as well as with simultaneous occurrence of transient EEG signals such as spikes, spike and slow wave complexes or rhythmic slow wave bursts. The variations of H_p vs time are shown as solid curves in Figs. 5(a)–5(e). We notice that slightly after the seizure, the PE has a sharp drop, followed by a gradual increase. This indicates that the dynamics of the brain first becomes more regular right after the seizure, then its irregularity increases as it approaches the normal state. We also note that the sharp drop in the PE is often preceded with an abrupt increase in the magnitude either slightly before the seizure [see Fig. 5(a), the first seizure], or slightly after the seizure [see Fig. 5(a), second seizure, and Fig. 5(c), both seizures]. This indicates that the dynamics of the brain may momentarily become very irregular associated with the occurrence of each seizure. We note that these features reflect well-known phenomena associated with seizures. Based on the clinical characterization, we conclude that the PE has indicated all the seizures present in the analyzed data. Similar results have been found when the method is applied to analyze EEG signals measured by other electrodes. In Figs. 5(c)–5(e), we represent the H_p time series registered by channels LTD 1–3, respectively, from patient 3. For this patient, it was documented clinically that the seizure was localized to electrodes RTD 1–3. This channel consistency suggests that it could be sufficient to analyze EEG signals measured by one channel alone.

We make two final comments: (i) Here we have focused only on seizure detection. Other methods, such as that based on the Lyapunov exponents [22], have been primarily used for seizure prediction. In the future we intend to investigate the use of PE for seizure prediction. (ii) Much of the driving force behind the research on seizure detection/prediction is the perspective on clinical real-time on-line monitoring of seizures. Hence, computational efficiency of a method is of paramount importance and must be given serious consideration. It is this very feature that distinguishes the PE approach from others. Indeed, we have found that a single channel 6-hour duration EEG signal, with sampling frequency of 200 Hz, can be processed in less than 1 minute. Hence processing 30 channel EEG data can be done in less than one-half hour. This means a simple PC is more than sufficient for on-line processing of multiple channel EEG signals.

V. CONCLUDING REMARKS

In this paper, we have explored the possibility of using the PE to detect dynamical changes in a complex time series. By analyzing two simple models, namely a transient logistic map and transient Lorenz system, as well as a number of clinical EEG data, we have shown that the PE can indeed be effectively used to detect bifurcations in model systems as well as the onset of epileptic seizures in intracrannial EEG data. Certainly there is no reason to expect that the PE is universally and indiscriminately applicable. Most likely, no such measure exists; instead, various measures would have to be used in a complementary fashion, to take best advantage of their respective merits within their ranges of applicability. We conclude though by emphasizing that the most attractive features of the PE, namely its conceptual simplicity and computational efficiency make it an excellent candidate for a fast, robust, and useful screener and detector of unusual patterns in complex time series.

ACKNOWLEDGMENTS

The authors thank Professor Sackellares of the Shands Hospital at the University of Florida for kindly providing them with the EEG data and Ms. Hui Liu for preparing the data. V.P. was partially supported by the Division of Material Sciences and Engineering, DOE Office of Basic Sciences. ORNL is operated for the DOE under Contract No. DE-AC05–00OR22725 with UT-Battele, LLC.

[1] J.P. Eckmann, S.O. Kamphorst, and D. Ruelle, Europhys. Lett. **4**, 973 (1987).
[2] L.L. Trulla, A. Giuliant, J.P. Zbilut, and C.L. Webber, Phys. Lett. A **223**, 255 (1996).
[3] J.B. Gao and H.Q. Cai, Phys. Lett. A **270**, 75 (2000).
[4] J.B. Gao, Phys. Rev. Lett. **83**, 3178 (1999); Phys. Rev. E **63**, 066202 (2001); J.B. Gao, Yinhe Cao, Lingyun Gu, J.G. Harris, and J.C. Principe, Phys. Lett. A **317**, 64 (2003); J.B. Gao, Hui Liu, J.C. Principe, and Li Zhang (unpublished).
[5] C. Rieke, K. Sternickel, R.G. Andrzejak, C.E. Elger, P. David, and K. Lehnertz, Phys. Rev. Lett. **88**, 244102 (2002).
[6] A. Provenzale, L.A. Smith, R. Vio, and G. Murante, Physica D **58**, 31 (1992).
[7] D.J. Yu, W.P. Lu, and R.G. Harrison, Phys. Lett. A **250**, 323 (1998); Chaos **9**, 865 (1999).
[8] R. Manuca and R. Savit, Physica D **99**, 134 (1996).
[9] M.B. Kennel, Phys. Rev. E **56**, 316 (1997).
[10] L.M. Hively, P.C. Gailey, and V.A. Protopopescu, Phys. Lett. A **258**, 103 (1999); L.M. Hively, V.A. Protopopescu, and P.C. Gailey, Chaos **10**, 864 (2000); V.A. Protopopescu, L.M. Hively, and P.C. Gailey, J. Clin. Neurophysiol. **18**, 223 (2001); L.M. Hively and V.A. Protopopescu, IEEE Trans. Biomed. Eng. **50**, 584 (2003).
[11] H. Kantz, Phys. Rev. E **49**, 5091 (1994).
[12] T. Schreiber, Phys. Rev. Lett. **78**, 843 (1997).
[13] C. Bandt and B. Pompe, Phys. Rev. Lett. **88**, 174102 (2002); C. Bandt, G. Keller, and B. Pompe, Nonlinearity **15**, 1595 (2002).
[14] M. Misiurewicz, Nonlinearity **16**, 971 (2003).
[15] W. van Drongelen, S. Nayak, D.M. Frim, M.H. Kohrman, V.L. Towle, H.C. Lee, A.B. McGee, M.S. Chico, and K.E. Hecox, Pediatr. Neurol. **29**, 207 (2003).

[16] J.B. Gao and Z.M. Zheng, Phys. Lett. A **181**, 153 (1993); Europhys. Lett. **25**, 485 (1994); Phys. Rev. E **49**, 3807 (1994).
[17] N.H. Packard, J.P. Crutchfield, J.D. Farmer, and R.S. Shaw, Phys. Rev. Lett. **45**, 712 (1980); F. Takens, in *Dynamical Systems and Turbulence*, edited by D.A. Rand and L.S. Young, Lecture Notes in Mathematics, Vol. 898, (Springer-Verlag, Berlin, 1981), p. 366.
[18] K. Lehnertz and C.E. Elger, Electroencephalogr. Clin. Neurophysiol. **95**, 108 (1995).
[19] K. Lehnertz and C.E. Elger, Electroencephalogr. Clin. Neurophysiol. **103**, 376 (1997).
[20] J. Martinerie, C. Adam, M. Le Van Quyen, M. Baulac, S. Clemenceau, B. Renault, and F.J. Varela, Nat. Med. **4**, 1173 (1998).
[21] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H.U. Voss, J. Timmer, and A. Schulze-Bonhage, Brain **126**, 2616 (2003).
[22] L.D. Iasemidis, J.C. Sackellares, H.P. Zaveri, and W.J. Williams, Brain Topogr **2**, 187 (1990).
[23] Y.C. Lai, M.A.F. Harrison, M.G. Frei, and I. Osorio, Phys. Rev. Lett. **91**, 068102 (2003).
[24] M. Le Van Quyen, J. Martinerie, M. Baulac, and F. Varela, NeuroReport **10**, 2149 (1999); M. Le Van Quyen, J. Martinerie, V. Navarro, P. Boon, M. D'Have, C. Adam, B. Renault, F. Varela, and M. Baulac, Lancet **357**, 183 (2001).