

Timely detection of dynamical change in scalp EEG signals

L. M. Hively and V. A. Protopopescu

Oak Ridge National Laboratory, P.O. Box 2009, Oak Ridge, Tennessee 37831-8066

P. C. Gailey

*Oak Ridge National Laboratory, P.O. Box 2009, Oak Ridge, Tennessee 37831-8066
and Department of Physics and Astronomy, Ohio University, Athens, Ohio 45701*

(Received 22 February 2000; accepted for publication 21 July 2000)

We present a robust, model-independent technique for quantifying changes in the dynamics underlying nonlinear time-serial data. After constructing discrete density distributions of phase-space points on the attractor for time-windowed data sets, we measure the dissimilarity between density distributions via L_1 -distance and χ^2 statistics. The discriminating power of the new measures is first tested on data generated by the Bondarenko “synthetic brain” model. We also compare traditional nonlinear measures and the new dissimilarity measures to detect dynamical change in scalp EEG data. The results demonstrate a clear superiority of the new measures in comparison to traditional nonlinear measures as robust and timely discriminators of changing dynamics. © 2000 American Institute of Physics. [S1054-1500(00)00504-8]

One of the most important problems encountered in nonlinear time-series analysis is the appropriate characterization of changes in the system’s dynamics. This problem is particularly vexing in physiological systems, which are more often than not: complex, nonlinear, nonstationary, and affected by noise. It is generally accepted that, in relation to epileptic phenomena, the brain behaves like a reasonably low-dimensional dynamical system whose dynamics may vary between (quasi-)periodic and completely irregular (chaotic). Thus, to a certain extent, global aspects of brain dynamics may be legitimately quantified by traditional nonlinear descriptors such as Lyapunov exponents, Kolmogorov entropy, and correlation dimension. While these descriptors are adequate for discriminating between clear-cut regular and chaotic dynamics, they are not sufficiently sensitive to distinguish between slightly different chaotic regimes, especially when data are limited and/or noisy. Most brain dynamics prior to, during, and following an epileptic seizure fall within the latter regime. Therefore, robust and timely forewarning of epileptic seizures has remained an outstanding medical challenge, especially for nonhospitalized patients. To address this problem, we introduce four new measures of dissimilarity that are much more sensitive than the traditional nonlinear measures. Following standard techniques, our method converts time-serial data to a geometric representation, that describes the dynamics of the underlying nonlinear system, namely, the brain, on the corresponding attractor in the phase space. The frequency and sequence of visitation of various points of the attractor are described by a distribution function (DF), which does not change if the dynamics remain unchanged. If the dynamics change, the attractor and the DF will change as well. To compare a test case DF to the base case DF, we define various distances between the DFs. A significant distance signifies that the system has departed from the base case and can be inter-

preted as a forewarning of an impending unusual event, possibly a seizure. Our method combines several original advances to achieve sensitivity which is at least two orders of magnitude larger than that obtained to date by competing methods. First, before constructing the phase space distribution, we remove confounding artifacts, such as eye-blinks, with a new zero-phase quadratic filter. Artifact removal allows detection of dynamical change from single-channel, scalp (as opposed to multichannel, subdural) EEG, thereby enabling noninvasive, ambulatory, long-term, nonclinical monitoring. Second, we use differential measures of dissimilarity which preserve a much higher content of dynamical information than the traditional measures that average out dynamical changes by integration of large amounts of data. Third, our technique applies to various seizure types as opposed to present approaches that focus on temporal lobe epilepsy only. Finally, our technique has provided robust forewarning of seizures for a variety of clinical data: digital and analog from various sites; 200 Hz and 512 Hz sampling; raw data precision between 10 and 12 bits; presence of substantial noise; and use of a variety of channels in the 10/20 montage, as well as a single channel (13) in the bipolar montage.

I. INTRODUCTION

Physiological systems in either normal or pathologic conditions display a very rich variety of dynamical behaviors. These behaviors manifest themselves in signals that can be interpreted at various levels, namely, clinical, physiological, chemical, physical, etc. In the last two decades, since the advent of chaotic dynamics on the scientific stage, there has been strong re-energized interest in reading and interpreting physiological data within a physical framework. The dynamical approach is motivated by several features that are

shared by physiological and complex physical systems, namely, multiple time scales, quasiperiodicity, chaos, and self-organization. Typically, these systems: (i) comprise many components; (ii) have hierarchical structure; (iii) are driven by competing forces; and (iv) interact strongly with noisy and/or nonstationary environments. It is therefore reasonable to assume that, under certain circumstances, one can use the same framework to analyze and interpret physical as well as physiological time series; this approach would complement traditional medical diagnostics, warning, prevention, and cure, with more precisely quantified assessments.

Quantitative analysis of physiologic time series has been a difficult and frustrating problem. The most important issues include:

(i) Lack of proper (physical) modeling for physiological phenomena. As a result, signals have to be considered as generated by a black box whose internal mechanism is either poorly understood, or not understood at all.

(ii) Signals are usually nonstationary, i.e., statistical properties of the signal may change significantly over the observation period. Usually this change is not known *a priori* and not explicitly advertised.

(iii) Usually, physiological time series are nonlinear, belying the nonlinear structure of various organ dynamics and their complex, intricate interconnection, rich in feedbacks and hysteresis.

(iv) Physiological systems rarely function at steady state. On the contrary, living processes typically occur far from equilibrium, and use continuous feedback and control to adjust to changing conditions.

One of the most important problems encountered in nonlinear time-series analysis is the appropriate characterization of changes in the system's dynamics. As mentioned before, these changes may have various origins, namely, nonstationarity, nonlinearity, and nonequilibrium. The presence of any one of these factors in the dynamical equation frequently introduces erratic fluctuations, patchiness, lack of obvious structure, or other irregularities with a multiplicity of widely disparate length- and time-scales. Often times, these irregularities have been neglected as noise without much structure and meaning. However, recent advances in chaotic dynamics facilitate the interpretation of these intermediate and small scale details as structure with significant information about the underlying dynamics. Accounting for this structure would enable a deeper understanding of basic dynamical features of the heart, brain, and lungs, and result in more efficient assessment, prediction, prevention, control, and treatment.

The aim of this paper is to describe a new method for detecting dynamical change in scalp EEG signals as an additional quantitative means to complement timely clinical forewarning of a possible epileptic seizure. Epilepsy afflicts many millions in the U.S. alone. Epilepsy can be effectively treated and many patients are indeed under constant medication. However, constant medication frequently has severe side effects. Moreover, between 10% and 30% of the cases cannot be controlled by medication. In addition, some extreme seizures are accompanied by heart failure or breathing

interruption that require immediate medical intervention. On the other hand, many seizures are not life threatening or even serious medical events; they are simply a social nuisance and embarrassment. Consequently, a robust, reliable, and unobtrusive *seizure warning system* would provide a new treatment paradigm whereby patients would be constantly monitored rather than medicated. When pre-seizure activity is detected, the patient can be forewarned to take timely preventive steps such as interrupting one's activity and lying down in a quiet place, taking medicine, requesting emergency responders, or contacting one's physician.

Several related problems have been recently pursued by both the medical and scientific community. We mention among others: (i) modeling and analysis of the brain as a dynamical system, (ii) chaotic analysis of the EEG data,^{1,2} (iii) seizure onset prediction,^{1,3} and (iv) seizure mitigation by chaos anticontrol.⁴ Similar problems have been studied for the heart in connection with ECG data, heart attack prevention, and control of fibrillation, and, more recently, for the lungs.

Our approach to the warning problem is purely pragmatic. We set aside deep and difficult questions about the nature of the brain and brain dynamics or the accurate description thereof. Indeed, current literature provides evidence both for⁵⁻⁷ and against⁸⁻¹¹ the description of the brain as a low-dimensional dynamical system. We think that the jury is still out on this issue, and will be for some time to come. Consequently, here we *assume* that the brain behaves like a reasonably low dimensional dynamical system and, therefore, global aspects of the brain dynamics can be conveniently characterized by nonlinear descriptors such as Lyapunov exponents, Kolmogorov entropy, correlation dimension, etc.⁶

Moreover, we make no attempt to answer questions about nonstationarity or nonequilibrium per se. Indeed, statistical tests for stationarity produce a binary result, namely, they indicate whether or not a change occurred but provide no information about the extent of departure from one state to another. Stationarity tests also have limited value for inherently nonstationary processes that undergo frequent or continual changes in dynamics (e.g., physiological data, like EEG). For such nonstationary processes, a measure of dissimilarity that quantifies the "distance" between attractors turns out to be more useful.^{12,13} Straightforward methods exist¹⁴⁻¹⁶ for discriminating between regular and chaotic motion, or for detecting the transition between these regimes. However, distinguishing different chaotic regimes can be very difficult, especially when data are limited and noisy. We describe four sensitive measures of dissimilarity, applying them first to model data and then to scalp EEG data. Our method is useful for (1) discriminating between different and possibly close chaotic regimes, and (2) monitoring the extent of departure of a system from a given dynamical state.

The paper is organized as follows: Sec. II discusses some traditional nonlinear measures for time series analysis. Section III presents our indicators of dynamical change by comparison of the phase-space distribution functions via dissimilarity measures. Section IV describes the discriminating power of these measures on a "synthetic brain" model pro-

posed by Bodnarenko, in which changes are easily monitored and controlled. Finally, Sec. V explains the use of these measures on experimental EEG data to detect the transition between nonseizure and epileptic brain activity. While the dissimilarity measures reflect the inherent noise in the EEG data, detection of pre-seizure condition change is clearly superior to that realized using traditional nonlinear measures.

II. NONLINEAR MEASURES FROM TIME SERIES

Our analysis begins with a process-indicative scalar signal, x , from a dynamical system whose dimensionality and structure are usually unknown. This signal is sampled at equal time intervals, τ , starting at time, t_0 , yielding a sequence of N points, $x_i = x(t_0 + i\tau)$. Dynamical process reconstruction¹⁴ uses d -dimensional time-delay vectors, $y(i) = [x_i, x_{i+\lambda}, \dots, x_{i+(d-1)\lambda}]$, for a system with d active variables and time lag, λ . The choice of lag and embedding dimension, d , determine how well the phase space reconstruction unfolds the dynamics for a finite amount of noisy data. Takens found that, for a d -dimensional system, $2d + 1$ dimensions generally results in a smooth, nonintersecting reconstruction.¹⁷ Sauer *et al.*¹⁸ showed that, under ideal conditions, the first integer greater than the correlation dimension is often sufficient to reconstruct the system dynamics. This last statement has been confirmed by computing the embedding dimension via the false nearest-neighbors method.¹⁹⁻²¹ However, real data have finite length and are affected by noise, implying that too high an embedding dimension may result in overfitting. We further note that different observables of a system contain unequal amounts of dynamical information,²² implying that phase space reconstruction could be easier from one variable, but more difficult (or even impossible) from another. Our subsequent analysis is mindful of the balance between these caveats and the constraints imposed by the limited amount of noisy data.

Based on the phase space reconstruction, various nonlinear measures have been defined to characterize process dynamics. We choose three of these nonlinear measures, against which we compare the new metrics. In particular, we use (i) the first minimum in the mutual information function as a measure of decorrelation time, (ii) the correlation dimension as a measure of dynamic complexity, and (iii) the Kolmogorov entropy as a measure of predictability. We briefly describe these three measures in the following paragraphs.

(i) The mutual information function (MIF) is a nonlinear version of the (linear) autocorrelation and cross-correlation functions, and was originally developed by Shannon and Weaver²³ with subsequent application to time series analysis by Fraser and Swinney.²⁴ The MIF measures the average information (in bits) that can be inferred from one measurement about a second measurement, and is a function of the time delay between the measurements. Univariate MIF measures predictability within the same data stream at different times. Bivariate MIF measures predictability of one data channel, based on measurements in a second signal at different times. For the present analysis, we use the first minimum in the univariate MIF, M_1 , to indicate the average time lag

that makes x_i independent of x_j . The MIF, $I(q,r)$, and system entropy, H , are defined by

$$I(q,r) = I(r,q) = H(q) + H(r) - H(q,r), \tag{1}$$

$$H(q) = - \sum_i P(q_i) \log[P(q_i)], \tag{2}$$

$$H(q,r) = - \sum_{i,j} P(q_i,r_j) \log[P(q_i,r_j)]. \tag{3}$$

For a window of w points, we denote one set of data measurements by q_1, q_2, \dots, q_w , with associated occurrence probabilities $P(q_1), P(q_2), \dots, P(q_w)$. R denotes a second set of data measurements, r_1, r_2, \dots, r_w , with a time delay relative to the q_i values, having associated occurrence probabilities $P(r_1), P(r_2), \dots, P(r_w)$. The function $P(q_i, r_j)$ denotes the joint probability of both states occurring simultaneously. H and I are expressed in units of bits if the logarithm is taken in base two.

(ii) The maximum-likelihood correlation dimension, D , is defined by^{25,26}

$$D = \left\{ (-1/M) \sum_{i,j} \ln[(\delta_{ij}/\delta_0 - \delta_n/\delta_0)/(1 - \delta_n/\delta_0)] \right\}^{-1}, \tag{4}$$

where M is the number of randomly sampled point pairs; δ_{ij} is the maximum-norm distance between the (randomly chosen) $i-j$ point pairs, as defined in Eq. (6) (below). The distance (scale length) δ_n is associated with noise as measured from the time serial data. Note that the distances are normalized with respect to a nominal scale length δ_0 , which is chosen as a balance between sensitivity to local dynamics (typically at $\delta_0 \leq 5a$) and avoidance of excessive noise (typically at $\delta_0 \geq a$). Here, the symbol a denotes the absolute average deviation as a robust indicator of variability²⁶ in the time serial data,

$$a = (1/w) \sum_{i=1}^w |x_i - \bar{x}|, \tag{5}$$

where \bar{x} is the mean of x_i over a window of w points. The distances δ_{ij} are defined by

$$\delta_{ij} = \max_{0 \leq k \leq m-1} |x_{i+k} - x_{j+k}|, \tag{6}$$

where m is the average number of points per cycle.

(iii) The Kolmogorov entropy, K , measures the rate of information loss per unit time, or (alternatively) the degree of predictability. A positive, finite entropy generally is considered to be a clear demonstration that the time series and its underlying dynamics are chaotic. A large entropy indicates a stochastic, nondeterministic (totally unpredictable) phenomenon. One estimates the entropy from the average divergence time for pairs of initially-close orbits. More precisely, the entropy is obtained from the average time for two points on an attractor to go from an initial separation ($\delta < \delta_0$), to a separation of more than a specific distance ($\delta > \delta_0$). The maximum-likelihood entropy is calculated from the method by Schouten, Takens, and van den Bleek,²⁷

$$K = -f_s \log(1 - 1/b), \tag{7}$$

$$\underline{b} = (1/M) \sum_{i=1}^M b_i, \tag{8}$$

with b_i as the number of timesteps for two points, initially within $\delta < \delta_0$, to diverge to $\delta > \delta_0$. The symbol f_s denotes the data sampling rate.

Entropy and correlation dimension usually are defined in the limit of zero scale length. However, all real data have noise and even noiseless model data is limited by the finite precision of computer arithmetic. Thus, we choose a finite scale length that is slightly larger than the noise ($\delta_0 = 2a$), at which to report the values of K and D , corresponding to finite-scale dynamic structure. Consequently, the values of K and D that we report do not capture the full dynamical complexity and have smaller values than expected for the zero-scale-length limit ($\delta_0 \rightarrow 0$).

III. DEFINITION AND USE OF THE NEW MEASURES

The traditional nonlinear measures described in the previous section characterize global features of the nonlinear dynamics and distinguish sufficiently clearly between, say, regular and chaotic dynamics. However they do not easily reveal *slight* dissimilarities between dynamical states. The same is true for other global indicators, such as fractal dimension, Lyapunov exponents, etc. This lack of discrimination occurs because the traditional measures are based on averaged or integrated system properties of the attractor, which provide a global picture of long-term dynamical behavior. Traditional nonlinear measures ultimately provide only one or a few scalar measures as summary descriptors of large data segments.

Greater discrimination is possible by more detailed analysis of the reconstructed dynamics. The natural (or invariant) measure on the attractor provides a more refined representation of the reconstruction, describing the visitation frequency of the system dynamics over the phase space. To obtain a useful discrete representation of the invariant measure from time serial data, we proceed as follows. We first represent each signal value, x_i , as a symbolized form, s_i , that is, one of S different integers, $0, 1, \dots, S-1$,

$$0 \leq s_i = \text{INT}[S(x_i - x_{\min}) / (x_{\max} - x_{\min})] \leq S-1. \tag{9}$$

Here, the function (INT) converts a decimal number to the closest lower integer, and x_{\min} and x_{\max} denote the minimum and maximum values of x_i , respectively, over the base case (reference data). We previously used³¹ the minimum and maximum values over both the basecase and testcase (data to be tested for departure from the basecase). However, in real- or near-real-time analyses, only basecase extrema are actually known. We require that $s_i(x_i = x_{\max}) = S-1$ in order to maintain exactly S distinct symbols. Thus, the phase space is partitioned into S^d hypercubes or bins. By counting the number of phase-space points occurring in each bin, we obtain the distribution function as a discretized density on the attractor. We denote the population of the i th bin of the distribution function, Q_i , for the base case, and R_i for a test case, respectively. For infinitely precise data, this representation has been used in Ref. 28. The choice of parameters (S ,

N , and d) depends on the specific data under consideration. In the preliminary phase of the analysis, we iteratively varied each parameter with the others fixed, to obtain optimum sensitivity of the measures to changes in system dynamics for each class of data evaluated. After realizing optimal sensitivity, the values of the parameters were kept fixed. A systematic method to determine optimal values for these parameters is the subject of future work.

We use an embedding window, $M_1 = (d-1)\lambda$, based on the first minimum in the mutual information function, M_1 .²⁴ This choice of time delay provides maximal information for the reconstruction of the phase space dynamics. Then, we set $\lambda = \text{INT}[0.5 + M_1 / (d-1)]$ to obtain an integer value for the reconstruction lag when M_1 is not evenly divisible by $d-1$. The reconstruction requires that $\lambda \geq 1$, thus constraining the largest value of dimensionality to $d \leq 2M_1 + 1$ from the above formula.

We next compare the distribution function of a testcase process state to that of a basecase. Diks *et al.*²⁹ measured differences between delay vector distributions by the square of the distance between two distribution functions. Schreiber¹² measures dissimilarity via the Euclidean distance between phase-space points of the attractor. This measure *only* accounts for the geometrical shape and location of the attractor. Manuca and Savit³⁴ measure dissimilarity via ratios of the correlation integral over the DF. This is essentially the correlation dimension discussed in the previous section. Moreover, these papers discuss dissimilarity measures from the perspective of nonstationarity, while our focus is on condition change, as explained in the Introduction. Thus, here we measure the difference between Q_i with R_i by the χ^2 statistics and L_1 distance,

$$\chi^2 = \sum_i (Q_i - R_i)^2 / (Q_i + R_i), \tag{10}$$

$$L = \sum_i |Q_i - R_i|, \tag{11}$$

where the summations in both equations run over all of the populated cells in the phase space. The choice of these measures is dictated by the following considerations. The χ^2 statistic is one of the most powerful, robust, and widely-used statistical tests to measure discrepancies between observed and expected frequencies. The L_1 distance is the natural metric for distribution functions since it is directly related to the total invariant measure on the attractor. We note that these measures account for changes in the geometry of the attractor and for changes in the DF as well. To apply these measures properly we scale the total population of the unknown distribution function (sum over all the domain populations in R_i) to be the same as the total population of the basecase. The sum in the denominator of Eq. (10) is based on a test for equality of two multinomial distributions.³⁰

The previous analysis can be extended in a natural manner that is inherently compatible with the underlying dynamics. By connecting successive phase space points as indicated by the dynamics, $y(i) \rightarrow y(i+1)$, one obtains a discrete representation of the process flow.¹⁵ We thus form a $2d$ -dimensional vector, $Y(i) = [y(i), y(i+1)]$, in the con-

nected phase space. As before, Q and R denote the distribution functions for the basecase and testcase, respectively, in the connected phase space. We define the measure of dissimilarity between these two connected phase space states, as before, via the L_1 -distance and χ^2 statistic,³¹

$$\chi_c^2 = \sum_{ij} (Q_{ij} - R_{ij})^2 / (Q_{ij} + R_{ij}), \tag{12}$$

$$L_c = \sum_{ij} |Q_{ij} - R_{ij}|. \tag{13}$$

The subscript c indicates the connected distribution function measure. We note that the value $\lambda = 1$ results in $d - 1$ components of $y(i + 1)$ being redundant with those of $y(i)$, but we allow this redundancy to accommodate other data such as discrete points from two dimensional maps. Using pairwise connectivity between successive d -dimensional states, this approach captures even more dynamical information. This additional information results in a higher discriminating power of the connected measures as compared with their nonconnected counterparts. Indeed, we can prove that the measures defined in Eqs. (10)–(13) satisfy the following inequalities:

$$\chi^2 \leq L, \tag{14}$$

$$\chi_c^2 \leq L_c, \tag{15}$$

$$L \leq L_c, \tag{16}$$

$$\chi^2 \leq \chi_c^2. \tag{17}$$

To prove Eq. (14), we note that for any non-negative numbers, Q_i and R_i , we have $(Q_i + R_i) \geq |Q_i - R_i|$. Dividing the equality $(Q_i - R_i)^2 = |Q_i - R_i|^2$ by the preceding inequality yields $(Q_i - R_i)^2 / (Q_i + R_i) \leq |Q_i - R_i|$, which completes the

proof when summed over all values of i . One can prove Eq. (15) by the same argument. The proof for Eq. (16) follows directly from the definition of the L_1 -norm,

$$L = \sum_i |Q_i - R_i| = \sum_i \left| \sum_j (Q_{ij} - R_{ij}) \right| \leq \sum_i \sum_j |Q_{ij} - R_{ij}| = L_c. \tag{18}$$

Proving Eq. (17) amounts to showing

$$\sum_i \frac{[\sum_j (Q_{ij} - R_{ij})]^2}{\sum_j (Q_{ij} + R_{ij})} \leq \sum_i \sum_j \frac{(Q_{ij} - R_{ij})^2}{Q_{ij} + R_{ij}}, \tag{19}$$

where $Q_{ij}, R_{ij} \geq 0$, not all $Q_{ij} = 0$ and not all $R_{ij} = 0$. Inequality (19) is proven if it holds independently for each term in the i -sum.

By noting that $(A - B)^2 = (A + B)^2 - 4AB$, we can rewrite the inequality between the i th terms as

$$\begin{aligned} & \sum_j \left(Q_{ij} + R_{ij} - \frac{4Q_{ij}R_{ij}}{Q_{ij} + R_{ij}} \right) - \sum_j (Q_{ij} + R_{ij}) \\ & + 4 \frac{\sum_j Q_{ij} \sum_j R_{ij}}{\sum_j Q_{ij} + \sum_j R_{ij}} \\ & = 4 \left(\frac{\sum_{j=1}^n Q_{ij} \sum_{j=i}^n R_{ij}}{\sum_{j=i}^n Q_{ij} + \sum_{j=i}^n R_{ij}} - \sum_{j=i}^n \frac{Q_{ij} R_{ij}}{Q_{ij} + R_{ij}} \right) \geq 0. \end{aligned} \tag{20}$$

We shall prove inequality (20) by complete induction. We denote the expression in the last brackets by $E_n \equiv [NM / (N + M)] = S$, where $N = \sum_{j=i}^n Q_{ij}, M = \sum_{j=i}^n R_{ij}$, and $S = \sum_{j=i}^n [Q_{ij} R_{ij} / (Q_{ij} + R_{ij})]$.

For $n = 1, E_1 = 0$. Suppose now that $E_n \geq 0$ for some $n \geq 2$ and evaluate

$$\begin{aligned} E_{n+1} &= \frac{(N + Q_{n+1})(M + R_{n+1})}{N + Q_{n+1} + M + R_{n+1}} - S - \frac{Q_{n+1}R_{n+1}}{Q_{n+1} + R_{n+1}} \\ &= \frac{MN + NR_{n+1} + MQ_{n+1} + Q_{n+1}R_{n+1} - NS - Q_{n+1}S - MS - R_{n+1}S}{N + Q_{n+1} + M + R_{n+1}} - \frac{Q_{n+1}R_{n+1}}{Q_{n+1} + R_{n+1}} \\ &\geq \frac{NR_{n+1} + MQ_{n+1} + Q_{n+1}R_{n+1} - Q_{n+1}S - R_{n+1}S}{N + Q_{n+1} + M + R_{n+1}} - \frac{Q_{n+1}R_{n+1}}{Q_{n+1} + R_{n+1}}, \end{aligned} \tag{21}$$

where we have used that $MN - NS - MS \geq 0$ since $E_n \geq 0$.

After some algebra and taking into account that (i) the denominators are positive and (ii) $-S \geq -[MN / (M + N)]$ we transform Eq. (22) into

$$\begin{aligned} & MQ_{n+1}^2 - SQ_{n+1}^2 - 2Q_{n+1}R_{n+1}S + NR_{n+1}^2 - SR_{n+1}^2 \\ & \geq MQ_{n+1}^2 - \frac{MN}{M+N}Q_{n+1}^2 - 2Q_{n+1}R_{n+1}\frac{MN}{M+N} \\ & \quad + NR_{n+1}^2 - \frac{MN}{M+N}R_{n+1}^2. \end{aligned}$$

By multiplying with $M + N$ and canceling terms we finally obtain $(MQ_{n+1} - NR_{n+1})^2$ which is obviously non-negative.

Equation (17) follows without difficulty from a similar complete induction argument. These inequalities show that, as expected, the measures of condition change for the connected phase space are stronger than those for the phase space representation.

In the subsequent application of the new phase space measures to discriminate condition change, we note that distribution function values depend on one another due to phase

space construction from time delay vectors with dynamical structure.²⁹ The resulting statistical bias is avoidable by averaging contributions to Eqs. (10) and (12) over values of $y(j)$ or $Y(j)$ which satisfy $|i-j| < \Lambda$,²⁹ where Λ is some largest typical correlation scale length in the time series. We tested the bias in typical data by sampling every Λ -th connected phase space point for $4 \leq \Lambda \leq 23$, resulting in Λ different samples for the base case (Q_i) and for each cutset (R_i). We then averaged the sampled χ^2 values over the Λ^2 different combinations of distribution functions for the base case and test case cutsets. As expected, a decrease proportional to $1/\Lambda$ occurs in the sampled χ^2 values, because the number of data points contributing to χ^2 decreases in the same proportion. The trend over time in sampled χ^2 values remains the same as in χ^2 values without sampling, showing that no unexpected bias is present. Thus, we use unsampled χ^2 values for the remainder of this work as a *relative* measure, rather than as an unbiased statistic for accepting or rejecting a null statistical hypothesis.

IV. APPLICATION TO THE BONDARENKO MODEL

We assess the discriminating power of the new measures by testing them on the Bondarenko “synthetic brain” model,³² which is described by a coupled set of time-delayed ordinary differential equations,

$$\dot{u}_i(t) = -u_i(t) + \sum_{j=1}^M a_{ij} f(u_j(t - \tau_j)),$$

$$i, j = 1, 2, \dots, M. \quad (22)$$

This model is the generalization of the Hopfield model for the electronic circuit realization of a neural network by adding a time delay τ_j . Here $u_i(t)$ is the output signal of the i th neuron. The matrix a_{ij} denotes the coupling coefficients between the neurons, with randomly chosen values, $-2 \leq a_{ij} \leq 2$. The indices i and j run from 1 to $M = 10$ (ten neurons). The time delay of the j th neuron output, τ_j , is constant and equal to 10. The nonlinear response function, $f(x) = c \tanh(x)$, simulates nonlinear neural response to signals from neighboring neurons. The coefficient c is used in order to change the values of the coupling coefficients between the neurons a_{ij} simultaneously.

As mentioned before, some traditional nonlinear measures are good indicators of a bifurcation or transition to chaos. However, transitions between two chaotic regimes are not readily detected by these same measures, especially for relatively small changes in the parameter that underlies the transition. Therefore, the present work concentrates on measuring dissimilarity within a region where the Bondarenko system is known to behave chaotically,³² $5 \leq c \leq 18$. We integrate the model using a standard fourth-order Runge–Kutta method³³ with a time step of $h = 0.3$. We allow a time of $4 \times 10^8 h$ for the solution to achieve stationarity after initiating the integration with random impulses, $u_j(t=0) = \rho_j$ with ρ_j having uniformly random values, $-2 \leq \rho_j \leq 2$. We calculated 100 000 data values of u_i at fixed time intervals of $\Delta t = 60$ for each value of c . We obtained the (connected) phase space

measures by partitioning each 100 000-point Bondarenko dataset into four nonoverlapping subsets of 25 000 points each, for comparison to each of the 25 000-point subsets of base case at $c = 5$. We compared each of the four test case subsets to each of the four base case subsets, yielding 16 values for each of the four measures of dissimilarity, from which we obtain a mean and the standard deviation of the mean. We use one of the ten neuron signals for dissimilarity detection. For example, Fig. 1 shows various nonlinear measures vs c , by analyzing only the one neuron signal of the Bondarenko system. We obtain similar results for other neuron signals. The correlation dimension [Fig. 1(a)] varies erratically between 3.5 and 8.5. The Kolmogorov entropy [Fig. 1(b)] rises almost monotonically from 0.025 to 0.16. Figure 1(c) shows the location of the first minimum in the mutual information function, M_1 , with erratic variation as c increases. In sharp contrast, the (connected) phase space measures [Figs. 1(d) and 1(e)] increase almost monotonically from zero to more than 8×10^4 as c rises from 5 to 18. The values of L and χ^2 essentially coincide over the whole range because the measures are dominated by phase space bins that are populated only for the base case $P_i > 0$ for $Q_i = 0$ and only the test case $P_i > 0$ for $Q_i = 0$, for which the two measures become analytically equivalent. The curves in Figs. 1(d) and 1(e) correspond to the average measure of dissimilarity, while the error bars indicate the standard deviation of the means. We show error bars for the nonconnected phase space metrics only, because error bars for the connected phase measures are comparable. As expected from Eqs. (14)–(17), the connected phase-space measures are stronger than their nonconnected counterparts.

V. APPLICATION TO EEG DATA

We turn next to analysis of brain wave data, which have been described in terms of nonlinear dynamics.⁶ Nonlinear EEG measures are not stationary,³⁴ displaying instead marked transitions between normal and epileptic states. EEG data display low-dimensional features^{6,35} with at least one positive Lyapunov exponent,^{6,36} and hence positive Kolmogorov entropy. EEG data also display clear phase space structure,^{6,36} on which our analysis relies for measuring condition change in the Bondarenko model (see previous section) and the Lorenz attractor.³¹ We find that phase space measures are useful for nonlinear detection of condition changes in brain wave data.³¹ We emphasize that our work relies on scalp EEG which measures the synchrony in cortical neurons on an area of roughly 6 cm². Our analysis assumes that both the traditional and the new measures are sensitive to changes in this synchrony.

Nine data sets with 16 channels of analog scalp data in the bipolar montage were obtained from archival VHS tapes.³⁷ We used only channel 13, closest to the patient’s right eye. We note that channel 13 in the bipolar montage is an old designation for the difference between channels FP2 and F4 in the 10/20 montage. We used channel 13 in initial analyses [31] to demonstrate the robustness of the zero-phase quadratic filter for removing the eyeblink artifact (discussed below). We digitized this data at a sampling rate of 512 Hz

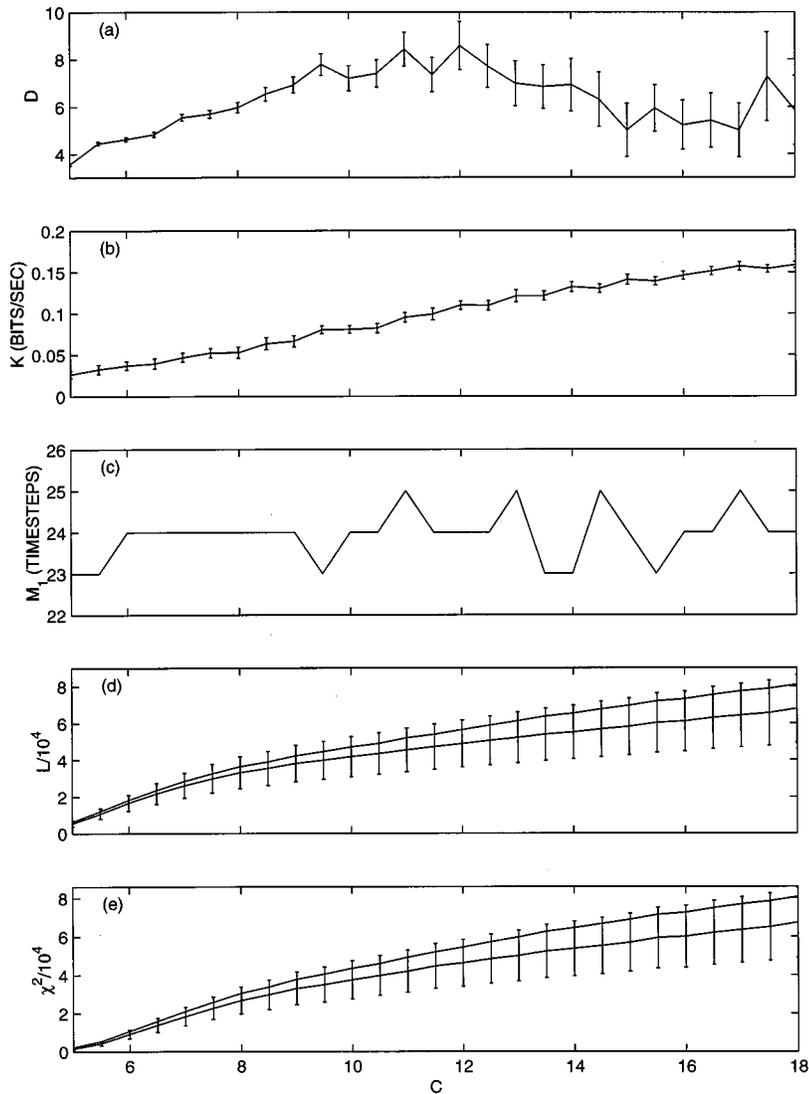


FIG. 1. Nonlinear measures versus c for the Bondarenko system, for one neuron channel: (a) correlation dimension, D , (b) Kolmogorov entropy, K , (c) location of the first minimum in the MIF, M_1 , (d) $\chi^2/10^5$, and (e) $L/10^5$. The error bars on D and K correspond to 95% confidence intervals. The phase space reconstruction parameters are $S=34$, $d=3$, $N=50\,000$, and $\lambda=4$. The (connected) phase space measures are the top (bottom) curves, respectively, in (d) and (e).

with 12-bit precision, corresponding to integers between -2048 and $+2047$. Table I summarizes these nine datasets with monitoring periods of 1380–3115 s. Seizures begin at times ranging from 966 and 2775 s. The seizure times in Table I were indicated on the records by the attending clinicians.

We also examined digital EEG scalp data from other clinical sites in the 10/20 International System of electrode placement, sampled at 200 Hz. These data have 10–12 bits of precision, with signal amplitudes between 0 and 3000, depending on the dataset. They were collected from a number of channels, varying between 23 to 32, with monitoring periods between 2217 and 20000 s. The clinical seizures begin at times that range between 1930 and 15750 s. We examined only one clinically interesting channel in each of these eleven datasets, as shown in Table I. In all cases the data were obtained by physicians under their own Human Studies IRB approval which included informed consent by the patients.

All scalp EEG are obscured by muscular activity due to eye blinks, facial twitches, etc. These artifacts are avoidable by obtaining EEG data from depth or subdural electrodes,

but such methods are invasive and nonambulatory. We remove most of the low frequency artifacts from the scalp EEG data with a novel zero-phase quadratic filter, which—unlike standard linear filters—retains the nonlinear amplitude and phase relationships.³⁷

This filter uses a moving window of $2n+1$ points of raw EEG data, e_i , with the same number of data points, n , on either side of a central point. We fit the data to a quadratic equation of the form, $F(t)=a_1T^2+a_2T+a_3$, with $T=t-t_c$, and t_c the time at the central point of the moving window. We fit this quadratic form to the data, by minimizing $\Psi=\sum[F(t)-e_i]^2$, where the sum is over the $2n+1$ points in the moving window. The minimum in Ψ is found from the condition $\partial\Psi/\partial a_k=0$, which yields three linear equations in three unknowns. The window-averaged artifact at the central point is given by the fitted value of the central point, $F(t_c=t_i)=a_3$. We note that the sums over odd powers of T_i are zero and that symmetric sums over even powers of T_i (over i from $-n$ to n) can be converted to sums from 1 to n , giving a window-averaged solution for the artifact signal,

TABLE I. Summary of EEG data (see text for discussion).

Dataset No.	Number of channels	Time of seizure (seconds)	Dataset length (seconds)	Channel analyzed	Sample rate (Hertz)	Preseizure activity	Seizure description ^a
109310	16	2775	3115.3	13	512 ^b	Asleep	*
109314	16	2480	2742.4	13	512 ^b	Drinking	fixed gaze
119230	16	2491	2917.4	13	512 ^b	Sponge bathing	tonic
119234	16	2560	2649.6	13	512 ^b	Asleep	*
62723t	16	2620	3060.8	13	512 ^b	*	*
69212	16	2356	2547.8	13	512 ^b	Moving in bed	subclinical
73305d	16	1245	1380	13	512 ^b	Talking	*
c8492d	16	966	1603.6	13	512 ^b	Awake/in bed	tonic
wm12sd	16	1041	1428.6	13	512 ^b	Hangs up phone	tonic
szprec	32	1930	2217	F7	200	*	
szpr00	23	5236	5401	Fp2	200	*	G, TC
szpr03	32	1932	2217	T4	200	*	2' G
szpr04	23	3794	3963	T4	200	*	G
szpr05	23	4888	6000.2	T4	200	*	G
emu002	27	4320	15 066	F4	200	*	LH onset, dorsal convexity outside the temporal lobe
emu003	27	13 200	16 228	C3	200	*	A, LF focus
emu004	27	15 750	18 423	C4	200	*	focal with partial seizures arising from R anterior mesial temporal
emu14	27	4080	20 222.2	F4	200	*	symptomatic G
emu18	27	4200	18 000.2	T3	200	*	focal arising from L anterior to mid- temporal area
emu26	26	13.987	16 224	Fp1	200	*	focal arising from R anterior temporal area

^a2' G = secondarily generalizing seizure; L=left; A=atypical; R=right; F=frontal; T/C=tonic/clonic seizure; G=generalized seizure; H=hemisphere.

^bThe 512 Hz datasets were acquired before 1995, and subsequent personnel changes make the information for these entries unavailable. Preseizure activity for the other datasets was not provided.

$$F(t=t_c) = [3(3n^2 + 3n - 1)(\sum e_{i+c}) - 15(\sum i^2 e_{i+c})] / (4n^2 + 4n - 3)(2n + 1).$$

The sums in this last equation are over i from $-n$ to n , with sums over even powers of i explicitly evaluated with standard formulas. The effort to evaluate this equation can be reduced further by computing the sums initially from the above equation with $c=n+1$, and then using recursions thereafter. Application of this filter to the N -point set of raw EEG data, e_i , yields $N-2n$ points of artifact data, f_i , that contains the low frequency artifact signal. The residual (artifact-filtered) signal, $g_i = e_i - f_i$, is free of low-frequency activity.

The filter window width corresponds to eye blink activity at 2 Hz, for which $n=128$ in the nine datasets with a 512 Hz sampling rate, and $n=50$ in the eleven datasets with a 200 Hz sampling rate. All subsequent EEG analysis uses this artifact-filtered data.

We choose $N=22\,000$ data points (43 s) for each cutset of the nine datasets, sampled at 512 Hz. This value balances the improvement in forewarning time discrimination at smaller N , with the statistical power to measure dissimilarity at larger N . For this same reason, we also choose N

$=22\,000$ data points (110 s) for each cutset of the eleven datasets, sampled at 200 Hz.

Our previous analysis of EEG data^{37,40} found correlation dimension values ranging between 2 and 6, consistent with values found by other groups.^{6,35} These results suggest a choice of $d \leq 7$ for the connected phase space reconstruction. However, we find that $d=7$ overfits the EEG data due to noise, modest cutset size, and the finite precision. We iteratively varied each phase-space construction parameter with the others fixed, to obtain optimum sensitivity of the phase-space measures to EEG changes. We subsequently performed all EEG analyses with the single best choice for S, d , and N . In particular, we found the best phase-space reconstruction parameters using the nine 512 Hz datasets, and then fixed the parameters $d=3$, $N=22\,000$, and $S=22$ for all the subsequent analysis of the other datasets. For the nine datasets, sampled at 512 Hz, the value of M_1 is taken from the first 430 s of (nonseizure) data; in the eleven datasets, sampled at 200 Hz, the value of M_1 is taken from the first 1100 s of data.

We use the first ten nonoverlapping cutsets in each of the datasets as basecases.³¹ This choice is a balance between a reasonably short basecase period to capture quasistationary

nonseizure activity and a sufficiently long period for statistical significance. However, a few of these basecase cutsets are very different from typical non-seizure activity, causing a severe bias in the detection of condition change. Thus, we statistically test the data for outlier cutsets as follows. Comparisons among the ten basecase cutsets yields 45 ($=9 \times 10/2$) unique pairs, from which we obtain an average, \bar{V} , and sample standard deviation, σ for each of the dissimilarity measures, $V=L$, L_c , χ^2 , and χ_c^2 . We calculate a χ^2 statistic, $\Sigma(V_{ij}-\bar{V})^2/\sigma^2$, for each of these four dissimilarity measures. The index j is fixed, to test the j th cutset against the other nine cutsets, thereby giving nine degrees of freedom in the χ^2 statistic. The null statistical hypothesis allows a random outlier in these 45 unique comparisons with a probability of $\leq 1/45$, corresponding to less than one out of the 45 unique pairs. Thus, we identify an outlier cutset as having the largest χ^2 statistic greater than 19.38 over the four dissimilarity measures, which corresponds to a probability larger than $1/45$. If this analysis does not identify any outlier, then the previous values of V and are used for subsequent renormalization, as described below. If this analysis identifies an outlier, we remove it and repeat this analysis for the remaining nine basecase cutsets. Repeated application of this analysis identifies any additional outliers when the largest chi-squared statistic exceeds the below threshold, corresponding to a random probability of greater than $2/B(B-1)$, as interpolated from standard statistical tables for $B-1$ degrees of freedom.³⁸ Here, B is the number of nonoutlier basecase cutsets. Thus, rejection of the null hypothesis corresponds to a χ^2 statistic greater than 19.38, 17.24, 15.03, 12.74, and 10.33, for $B=10, 9, 8, 7,$ and 6 , respectively.

This approach dramatically improves the robustness of the condition change detection. If the analysis identifies five (or more) outliers, we would have to reject all ten basecases as unrepresentative, and acquire a new set of ten cutsets as basecases. However, the present analysis never finds more than four outliers. Subsequently, we compare the nonoutlier basecase cutsets to each nonoverlapping testcase cutset, and obtain average values for the dissimilarity measures for each testcase.

The disparate range and variability of the various nonlinear measures are difficult to interpret, so we need a consistent means of comparison. Thus, we convert the nonlinear measures to a renormalized form.³¹ For each nonlinear measure, $V=D, K, M_1, L, L_c, \chi^2$, and χ_c^2 , we define V_i as the value of nonlinear measure for the i th cutset. As before, \bar{V} is the mean value of that nonlinear measure over the nonoutlier basecases, with a corresponding sample standard deviation σ , as described above. For the new measures of dissimilarity, namely L, L_c, χ^2 , and χ_c^2 , V_i is averaged over the nonoutlier cutsets. No such averaging is done for D, K , and M_1 since the calculation of these measures involves only one cutset at the time. The renormalized form is then $U(V) = |V_i - \bar{V}|/\sigma$, which measures the number of standard deviations that the testcase deviates from the basecase mean. For a positive indication of change, we use $U \geq U_c = 3.09$, corresponding to a false positive probability of less than 10^{-3} in Gaussian random data. To avoid spurious false positives, we require two or more consecutive occurrences for indication

of condition change; this corresponds to a joint false positive probability of less than 10^{-6} in Gaussian data. Figure 3 shows a typical example in which single excursions of the nonlinear measures above the threshold occur at 55 s and 715 s. Consequently, we require two excursions above threshold as the criterion for condition change. Using this "twice-in-succession" criterion, Fig. 3 shows clear forewarning at 1100 s without false positives. We also note that the forewarning times in Table II are determined by using the full length of available data in all cases. We next use these renormalized forms for measuring changes in EEG.

Figure 2 shows the renormalized nonlinear measures for dataset #szprec. The vertical lines in these plots indicate onset of the clinical seizure at 1930 s and subsequent postseizure period. The nonlinear measures are plotted at the center of the time window for each cutset. All of the measures show low to modest variability during the period of nonseizure brain activity (less than 900 s). Figure 2(a) shows the minimum, e_{\min} , and maximum, e_{\max} , in the raw EEG signal with no clear pre-seizure features in the signal envelop. Other subplots [Figs. 2(b)–2(f)] show the renormalized nonlinear measures, with a horizontal line indicating the threshold for condition change detection, $U_c = 3.09$. The correlation dimension, D [Fig. 2(b)] rises above the threshold from 1250 to 1500 s, subsequently falls below the threshold, and then rises above threshold at 1925 s through the seizure. The Kolmogorov entropy, K [Fig. 2(c)] provides no pre-seizure indication, and rises above U_c only during the seizure. The first minimum in the mutual information function, M_1 [Fig. 2(d)] exceeds U_c from 1250 to 1500 s, then falls below the threshold without any seizure indication. In sharp contrast to these weak pre-seizure indications, the renormalized phase-space measures [Figs. 2(e)–2(f)] all rise above the threshold at 1155 s, rising still further near and immediately following the seizure.

TABLE II. Difference(s) between seizure onset and change detection (in s). Entries with an asterisk (*) show no positive indication of change. For each dataset, bold entries denote the earliest time of change.

Dataset No.	D	K	M_1	L_c	L	χ_c^2	χ^2
109310	1 099	*	*	-61	-61	1 142	-61
109314	1 921	1 406	1 835	1 878	1 921	1 921	1 921
119230	901	386	-216	471	-44	471	514
119234	1 915	*	*	1 915	1 915	1 915	1 915
62723t	1 374	*	-44	2 233	1 675	2 233	2 233
69212	*	165	637	1 626	1 497	1 626	1 626
73305d	600	600	*	343	772	-87	772
c8492d	-22	321	364	193	193	193	193
wm12sd	*	*	*	-76	10	10	10
szprec	500	-160	500	610	610	610	610
szpr00	*	*	1 496	726	-154	836	1 716
szpr03	-158	-158	172	502	502	502	502
szpr04	-166	*	-166	384	384	384	384
szpr05	3 568	3 348	3 568	3 678	3 568	3 678	3 568
emu002	*	-190	-410	2 230	2 780	1 900	2 780
emu003	*	*	*	12 760	12 760	12 760	12 760
emu004	*	6 950	*	13 660	13 550	14 540	13 660
emu014	*	*	-540	670	670	-210	670
emu018	-90	-1 630	-310	3 650	2 200	3 650	2 220
emu026	11 127	11 237	4 747	11 237	11 237	11 237	11 237

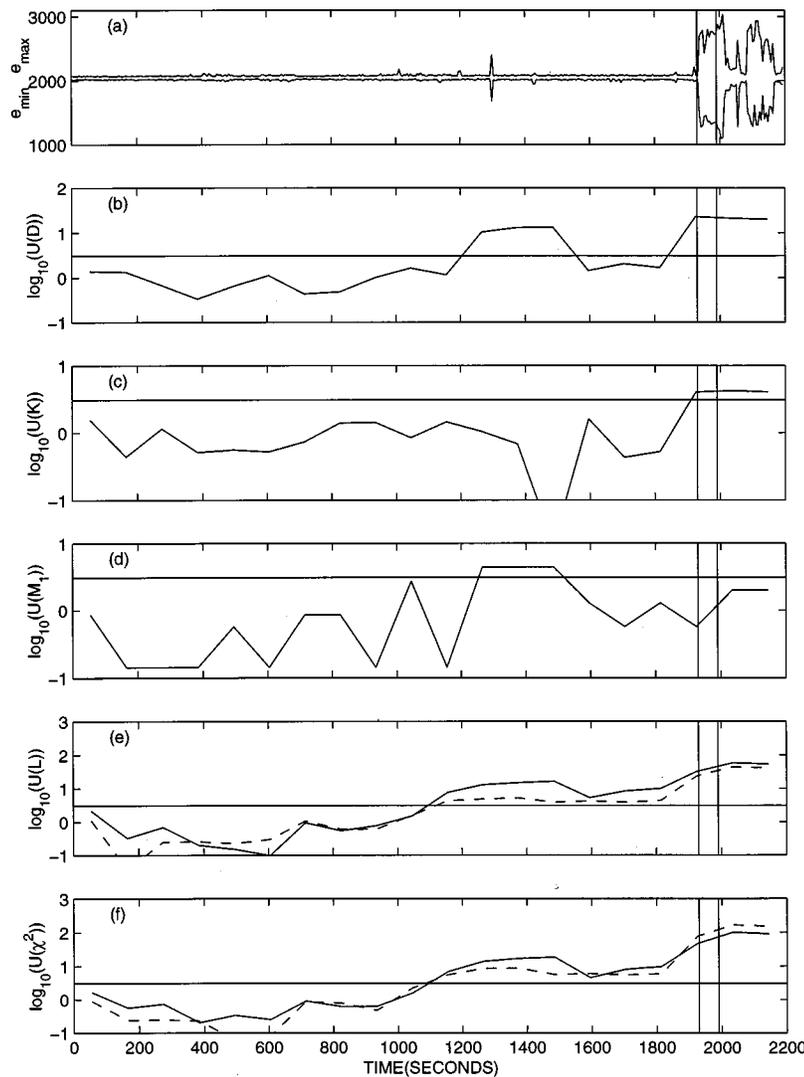


FIG. 2. Renormalized nonlinear measures versus time for dataset #szprec: (a) correlation dimension, (b) Kolmogorov entropy, (c) first minimum in the MIF (in time steps), (d) L_1 measure for the connected phase space (solid) and phase space (dotted), and (e) χ^2 measure for the connected phase space (solid) and phase space (dotted). The ordinate values of the change metric U are in units of standard deviations from the mean.

Figure 3 displays the renormalized nonlinear measures for dataset #szpr03, for which the onset of clinical seizure occurs at 1932 s. Figure 3(a) shows the minimum, e_{\min} , and maximum, e_{\max} , in the raw EEG signal with little pre-seizure amplitude variability, except for four positive spikes between 450 and 750 s. D [Fig. 3(b)] and K [Fig. 3(c)] give essentially no pre-seizure warning, exceeding the threshold for condition change immediately prior to and during the seizure. M_1 exceeds U_c at 1375 s and at 1595–1815 s, but does not indicate the seizure. The connected phase-space measures (solid lines) in Figs. 3(e)–3(f) show a single excursion above threshold at 715 s. Subsequently, all of the phase-space measures rise and stay above U_c , beginning at 1265 s through seizure.

Table II summarizes the forewarning times for each measure over all 20 EEG datasets. A negative value of forewarning time corresponds to an indication after seizure onset. Starred (*) values indicate that no condition change was detected by this measure. Analysis of normal EEG shows no positive indication of change. We assess these results as follows. The phase space measures provide the earliest seizure forewarning in 11, 10, 14, and 13 datasets for L , L_c , χ^2 , and χ_c^2 , respectively. Moreover, the phase-space measures pro-

vide pre-seizure indications in all 20 cases. In sharp contrast, the traditional nonlinear measures only give the earliest forewarning of a seizure in 1, 1, and 3 instances for K , M_1 , and D , respectively. These same traditional measures provide no forewarning of a seizure in 7, 8, and 6 cases, respectively. The sum of the earliest-forewarning times exceeds twenty, because more than one measure can simultaneously detect condition change. A topic of future research is improvement of the method to provide the earliest possible indication of condition change. We also note that the forewarning time (10 s) for dataset #wm12sd is too short to be clinically useful. In addition, the forewarnings of more than 1 h (datasets #emu003, emu004, emu026) are too long to be clinically useful. We conclude that the phase space measures are much superior to the conventional nonlinear measures as pre-seizure indicators of condition change for a single channel of scalp EEG.

Our approach differs markedly from earlier work. First, previous studies^{2,3,35,36} used multichannel data from subdural and depth electrodes to avoid low-frequency (e.g., eyeblink) artifacts. Instead, we use one channel of scalp EEG data that allows for noninvasive, ambulatory, long-term, nonclinical monitoring. Second, we remove the low-frequency artifacts

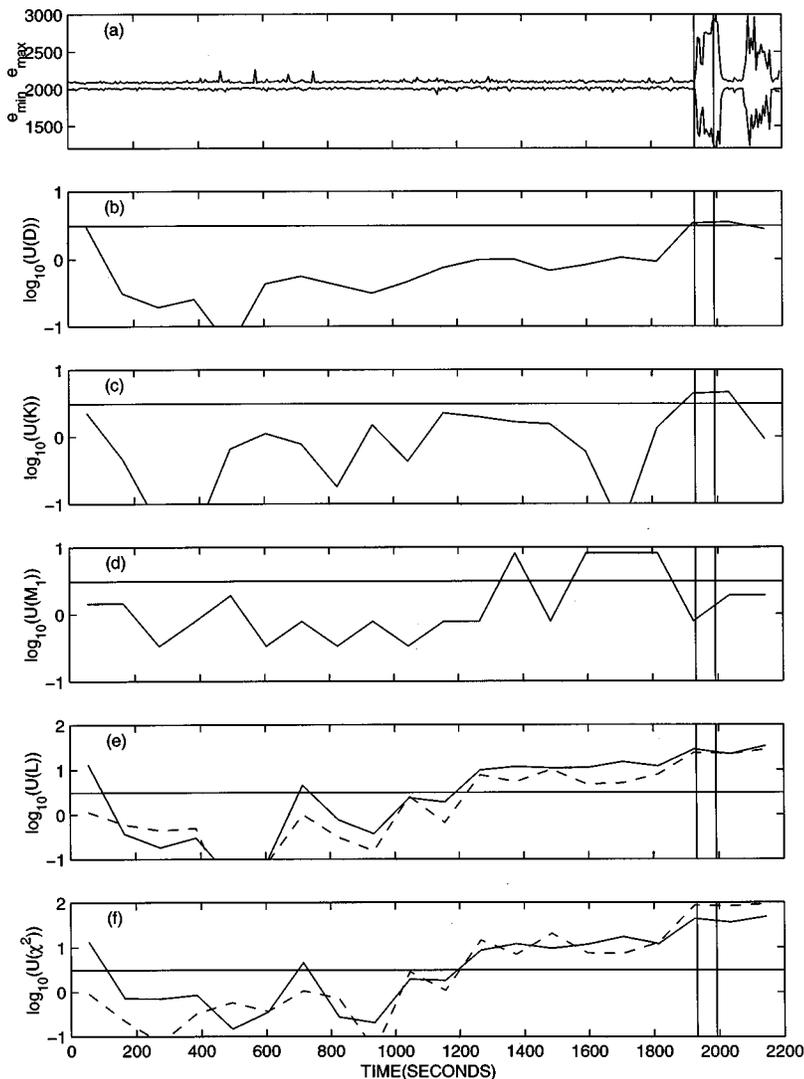


FIG. 3. Renormalized nonlinear measures versus time for dataset #szpr03: (a) correlation dimension, (b) Kolmogorov entropy, (c) first minimum in the MIF, (d) L_1 measure for the connected phase space (solid) and phase space (dotted), and (e) χ^2 measure for the connected phase space (solid) and phase space (dotted). The ordinate values of the change metric U are in units of standard deviations from the mean.

from scalp EEG with a novel zero-phase quadratic filter. Third, prior investigations^{2,3,35,36} forewarn of temporal lobe epilepsy via changes in conventional nonlinear measures. In particular, other groups use the correlation integral,^{2,3} correlation dimension, and largest Lyapunov exponent.³⁶ Earlier we determined no consistent trends in such conventional nonlinear measures for various seizure types.³¹ Instead, we focus on phase-space dissimilarity measures without regard to seizure type. Fourth, our studies demonstrate methodology robustness over a variety of clinical conditions: digital and analog EEG from several clinical sites; data sampling at 200 and 512 Hz; raw EEG data precision between 10–12 bits; presence of substantial noise in the raw EEG, as well as periods of constant signal; and use of a fixed channel, namely, 13, in the bipolar montage, and use of a variety of clinically interesting channels in the 10/20 montage. The robustness of our approach suggests that this methodology could allow convenient electrode placement by a patient in a nonclinical setting and be used as a complementary quantitative method in conjunction with clinical assessment. Future use of this approach as a stand-alone method for epilepsy forewarning will require analysis of several seizures for each patient, and detailed determination of detection criteria. Fu-

ture work will involve: (i) statistical evaluation of false positives and negatives in epileptic patients, and of false positives in normal patients, (ii) clinical monitoring of each patient to determine optimal phase-space reconstruction parameters, which subsequently would be fixed for ambulatory monitoring, (iii) the specific nonlinear feature(s) for seizure forewarning in the EEG, and (iv) the response of our approach to psychiatric states, such as hysterical seizures.

VI. DISCUSSION

We presented model-independent indicators to detect condition change (dissimilarity) in nonlinear time series. The phase-space indicators of condition change measure the difference between these density functions for a basecase and a testcase, as χ^2 statistics and L_1 distance. Thus, these indicators integrate (and magnify) the differences between the process dynamics, and avoid the inner cancellation effects due to averaging over many orbits (as one does, for instance, when calculating the correlation dimension and Kolmogorov entropy). Changes in the Bondarenko attractor are clearly detected by these phase-space measures and by Kolmogorov entropy, as c increases from 5 to 18 (Fig. 1). On the other

hand, these changes are either undetected or barely detected by the correlation dimension and mutual information measures. The phase-space measures also indicate significant preseizure changes in all 20 EEG datasets (Table II). However, the correlation dimension, Kolmogorov entropy, and mutual information fail to detect any change at all in 7, 8, and 6 EEG datasets, respectively. These results show that the phase-space measures are superior to traditional nonlinear measures for detection of condition change. The connected phase space distribution function also contains information about the dynamical flow from one attractor state to the next.

We find that indications of condition change exist over a range of embedding parameters (d , S , N , and λ), for the following reason. If the full distribution function in the actual phase space is Z , then a partially reconstructed distribution function, Y , is the sum over regions of Z that map into Y , multiplied by the Jacobian determinant for the change in variables. Thus, the information in every component of Z will affect the Y space, unless the projection, $Z \rightarrow Y$, is chosen carefully to avoid dependency, which we have not done. Kennel³⁹ used this approach to test for dynamical nonstationarity in experimental data without resolving the dimensionality in the time series. The important question of optimal embedding parameters for measuring dissimilarity with our measures will be addressed elsewhere.

An early version⁴⁰ of this approach was successfully applied to detect dynamical change in various physical processes. Examples include distinguishing different drilling conditions from spindle motor current of a machining center; detecting balanced and unbalanced centrifugal pump conditions from motor power; and predicting failure of a bellows coupling in a rotating drive train from motor current. More recent analyses include discerning the difference in microcantilever vibrations with and without mercury on the sensor; and forewarning of cardiac fibrillation. Success for such diverse applications suggests that this technique can be reliably used for measuring condition change in nonlinear processes.

ACKNOWLEDGMENTS

This work was sponsored by the U. S. Department of Energy, Office of Energy Management. V.A.P. was partially supported by the U. S. Department of Energy, Office of Basic Energy Sciences. Oak Ridge National Laboratory is managed for the United States Department of Energy by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725. We are grateful to members of the staff at the Knoxville Neurology Clinic at St. Mary's Medical Center (Knoxville, TN), including Dr. M. L. Eisenstadt, Dr. R. E. Leppanen, W. W. Holland, and H. L. Leach, who provided expert assistance with the EEG data acquisition and interpretation. We thank N. E. Clapp, Jr. (ORNL) for converting the data from archival VHS tapes to digital form, and for determining the times of the clinical seizures from the paper records.

¹I. Osorio, M. G. Frei, and S. B. Wilkinson, *Epilepsia* **39**, 615 (1998).

²J. Martinerie, C. Adam, M. Le Van Quyen, M. Baulac, S. Clemenceau, B.

Renault, and F. J. Varela, *Nat. Med. (N.Y.)* **4**, 1173 (1998).

³M. Le Van Quyen, J. Martinerie, M. Baulac, and F. Varela, *NeuroReport* **10**, 2149 (1999).

⁴S. J. Schiff, K. Jerger, D. H. Duong, T. Chang, M. L. Spano, and W. L. Ditto, *Nature (London)* **370**, 615 (1994).

⁵L. Pezard, J. Martinerie, F. Breton, J.-C. Bourzeix, and B. Renault, *Electroencephalogr. Clin. Neurophysiol.* **91**, 383 (1994).

⁶T. Elbert, W. J. Ray, Z. J. Kowalik, J. E. Skinner, E. E. Graf, and N. Birbaumer, *Physiol. Rev.* **74**, 1 (1994).

⁷D. E. Lerner, *Physica D* **97**, 563 (1996).

⁸G. W. Frank, T. Lookman, M. A. H. Nerenberg, C. Essex, J. Lemieux, and W. Blume, *Physica D* **46**, 427 (1990).

⁹D. K. Ivanov and H. A. Posch, *Chaos* **6**, 243 (1996).

¹⁰N. Pradhan and P. K. Sadasivan, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **7**, 173 (1997).

¹¹J. Jeong, M. S. Kim, and S. Y. Kim, *Phys. Rev. E* **60**, 831 (1999).

¹²A. T. Schreiber, *Phys. Rep.* **308**, 1 (1999); *Phys. Rev. Lett.* **78**, 843 (1997).

¹³B. R. Moeckel and B. Murray, *Physica D* **102**, 187 (1997).

¹⁴J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* **57**, 617 (1985).

¹⁵H. D. I. Abarbanel, *Analysis of Observed Chaotic Data* (Springer, New York, 1996).

¹⁶A. Cover, J. Reneke, S. Lenhart, and V. Protopopescu, *Math. Models and Meth. in Appl. Sci.* **7**, 823 (1997).

¹⁷F. Takens, *Lecture Notes in Mathematics* (Springer, New York, 1981), Vol. 898, p. 366.

¹⁸T. Sauer, J. Yorke, and M. Casdagli, *J. Stat. Phys.* **65**, 579 (1991).

¹⁹H. D. I. Abarbanel and M. B. Kennel, *Phys. Rev. E* **47**, 3057 (1993).

²⁰H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. Sh. Tsimring, *Rev. Mod. Phys.* **65**, 1331 (1993).

²¹L. Cao, *Physica D* **110**, 43 (1997).

²²C. Letellier, J. Maquet, L. Le Sceller, G. Gouesbet, and L. A. Aguirre, *J. Phys. A* **31**, 7913 (1998).

²³C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949).

²⁴A. M. Fraser and H. L. Swinney, *Phys. Rev. A* **33**, 1134 (1986).

²⁵F. Takens, *Lecture Notes in Mathematics* (Springer-Verlag, Berlin, 1984), Vol. 1125, pp. 99–106.

²⁶J. C. Schouten, F. Takens, and C. M. van den Bleek, *Phys. Rev. E* **50**, 1851 (1994).

²⁷J. C. Schouten, F. Takens, and C. M. van den Bleek, *Phys. Rev. E* **49**, 126 (1994).

²⁸C. Grebogi, E. Ott, and J. A. Yorke, *Phys. Rev. A* **37**, 1711 (1988).

²⁹C. Diks, W. R. van Zwet, F. Takens, and J. DeGoede, *Phys. Rev. E* **53**, 2169 (1996).

³⁰A. M. Mood, S. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics* (McGraw-Hill, New York, 1974).

³¹L. M. Hively, P. C. Gailey, and V. A. Protopopescu, *Phys. Lett. A* **258**, 103 (1999); in *Proceedings of the International Workshop: "Chaos in Brain?," Bonn, Germany, 10–12 March, 1999*, edited by K. Lehnertz *et al.* (World Scientific, Singapore, 2000), p. 333; P. C. Gailey, L. M. Hively, and V. A. Protopopescu, in *Proceedings of the Fifth Experimental Chaos Conference (ECC5), Orlando, Florida, June 28–July 1, 1999* (to be published).

³²V. E. Bondarenko, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **7**, 1133 (1997).

³³W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN* (Cambridge University Press, Cambridge, 1992).

³⁴R. Manuca, M. C. Casdagli, and R. S. Savit, *Math. Biosci.* **147**, 1 (1998); R. Manuca and R. Savit, *Physica D* **99**, 134 (1996).

³⁵K. Lehnertz and C. E. Elger, *Phys. Rev. Lett.* **80**, 5019 (1998).

³⁶L. D. Iasemidis and J. C. Sackellares, *Neuroscientist* **2**, 118 (1996).

³⁷L. M. Hively, N. E. Clapp, C. S. Daw, and W. F. Lawkins, *ORNL/TM-12961* (Oak Ridge National Laboratory, Oak Ridge, TN, 1995).

³⁸M. Abramovitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1964).

³⁹M. B. Kennel, *Phys. Rev. E* **56**, 316 (1997).

⁴⁰L. M. Hively, in *Proceedings of the Maintenance and Reliability Conference*, edited by T. E. Shannon *et al.* (University of Tennessee, Knoxville, 1997), Vol. 1, p. 16.01.