Correlation dimension and integral do not predict epileptic seizures

Mary Ann F. Harrison[a)]
Flint Hills Scientific L.L.C., 5020 Bob Billings Parkway, Suite A, Lawrence, Kansas 66049

Ivan Osorio
Flint Hills Scientific L.L.C., 5020 Bob Billings Parkway, Suite A, Lawrence, Kansas 66049 and Department of Neurology, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, Kansas 66160

Mark G. Frei and Srividhya Asuri
Flint Hills Scientific L.L.C., 5020 Bob Billings Parkway, Suite A, Lawrence, Kansas 66049

Ying-Cheng Lai[b)]
Department of Electrical Engineering and Department of Mathematics, Arizona State University, Tempe, Arizona 85287

(Received 9 December 2004; accepted 21 April 2005; published online 26 July 2005)

Reports in the literature have indicated potential value of the correlation integral and dimension for prediction of epileptic seizures up to several minutes before electrographic onset. We apply these measures to over 2000 total hours of continuous electrocortiogram, taken from 20 patients with epilepsy, examine their sensitivity to quantifiable properties such as the signal amplitude and autocorrelation, and investigate the influence of embedding and filtering strategies on their performance. The results are compared against those obtained from surrogate time series. Our conclusion is that neither the correlation dimension nor the correlation integral has predictive power for seizures.

[DOI: 10.1063/1.1935138]

Recent years have witnessed an increasing use of nonlinear-dynamics based measures in biomedical signal analysis. In the area of epilepsy, where an important goal is to predict seizures, the use of such measures as the correlation dimension, the correlation integral, and Lyapunov exponents has been practiced by various research groups for about 15 years, with claims that seizures can be predicted minutes or even hours in advance. Considering the enormous implication of such claims in terms of improving the quality of life for patients with epilepsy (in the United States, approximately 1% of the population are influenced by this disease), it is of paramount interest to perform systematic and extensive validation tests based on clinical data to generate evidence either for or against these claims. The aim of this paper is to present results of validation tests for the correlation dimension and integral using extensive recordings of electrocortiogram over 2000 hours from 20 patients with intractable epilepsy. In order to carry out the required tests, it is necessary to develop efficient algorithms for computing the correlation dimension and integral in an automated manner. We will present our new algorithms that allow us to accomplish this task. In addition, we will address technical issues such as signal amplitude normalization, use of autocorrelation, effects of filtering and embedding, and surrogate tests. To our knowledge, previous studies of the correlation dimension and integral used very limited numbers of cases and, in fact, no prior efforts were attempted to conduct validation tests to the extent that we have accomplished. The results of our tests suggest strongly that the correlation dimension and integral have no predictive power for epileptic seizures.

I. INTRODUCTION

Nonlinear dynamical measures such as the correlation integral and dimension, and Lyapunov exponents are commonly used in medical signal analysis for purposes of characterization, prediction, and system control.[1] In the context of epilepsy, these techniques have been utilized to analyze the electroencephalogram (EEG) or electrocortiogram (ECoG)[2–19] with various claims that seizures can be predicted up to several minutes in advance.[6,8,10,13,16] Reliable seizure prediction would reduce morbidity and greatly improve the quality of life for subjects with epilepsy, approximately 1% of the population of the United States.

Because ECoG is nonstationary and has a stochastic component, there are questions about the utility of nonlinear-dynamics based methods, and some studies have reported no significant difference in the seizure detection or prediction performance between linear and nonlinear measures.[20–22] In addition, there has been a lack of large-scale studies consisting of long time series necessary to validate prediction claims. Other significant factors that may hamper the validation of prediction claims include lack of standardization in the choice of method parameters, digital precision, filtering

[a)]Present address: The Institute for Scientific Research, Inc., 320 Adams Street, P.O. Box 2720, Fairmont, WV 26555.
[b)]Electronic mail: yclai@chaosl.la.asu.edu

methods, amplitude normalization schemes, and window length. Most existing reports[8–10,23] have been based on very short (10–60 min) time series, an insufficient time to test reliability and specificity of the methods. Though the dependence of fractal dimension computation on method parameters and window length has been discussed at length in the literature,[1] the effects of filtering methods and amplitude normalization have received much less attention but they have at least as much if not more impact on the value of the fractal dimension.[19] And, the intensive computational requirement of these algorithms, a barrier to real-time application and miniaturization, remains an under-studied problem.

To address these current limitations, in this paper we (1) implement the correlation integral computation in an efficient manner; (2) compute correlation integral-based measures for 2347 h of continuous ECoG data from 20 subjects with a median of 124 h (mean 117 h) per subject; and (3) assess the interplay between time-frequency changes in the signal and the responses of the measures. These systematic approaches suggest strongly that neither the correlation dimension nor the correlation integral is useful for seizure prediction.

In Sec. II, we present our efficient method for computing the correlation integral and dimension. In Sec. III, we describe how our data were collected. Important computational issues such as amplitude normalization, autocorrelation, filtering, embedding, and surrogate data analysis are addressed in Sec. IV. Section V gives results on validation using long time series. A discussion is presented in Sec. VI.

II. EFFICIENT METHOD FOR COMPUTING CORRELATION INTEGRAL

The correlation dimension, one definition of a fractal dimension, is often computed from experimental time series due to its relative ease of computation with this type of data. A generalized definition of the fractal dimension is[24–26]

$$D_q = \lim_{\epsilon \to 0} \frac{1}{q-1} \frac{\log \sum_i I(q,\epsilon)}{\log \epsilon}, \tag{1}$$

where a grid of size ϵ is used to cover the attractor,

$$I(q,\epsilon) = \sum_{i=1}^{N(\epsilon)} \mu_i^q,$$

and μ_i^q is the fraction of time a typical trajectory spent in the ith cell, or its natural measure. The sum over the natural measure can be written as a weighted average,[24–26]

$$\sum_{i=1}^{N(\epsilon)} \mu_i^q = \sum_{i=1}^{N(\epsilon)} \mu_i(\mu_i^{(q-1)}) = \langle \mu_i^{(q-1)} \rangle, \tag{2}$$

and if $q=2$, the expectation value is the arithmetic average. Typically, this average is expressed through the correlation integral, which computes the fraction of pairs of points on the attractor within a hypersphere of radius ϵ. The correlation integral can be approximated by the correlation sum

$$C(N,\epsilon) = \frac{1}{N(N-1)} \sum_{i \neq j} \Theta(\epsilon - \|\mathbf{x}_i - \mathbf{x}_j\|), \tag{3}$$

where N is the number of data points $\mathbf{x}$ in the reconstructed phase space. The definition for the dimension with $q=2$ thus becomes

$$D_2 = \lim_{N \to \infty} \lim_{\epsilon \to 0} \frac{\log C(N,\epsilon)}{\log \epsilon}. \tag{4}$$

For scalar time-series data, we perform computations of the correlation integrals by using the standard time-delay embedding[27] in order to recover hidden degrees of freedom

$$\mathbf{y}_i = [x_i, x_{i+\tau}, x_{i+2\tau}, \ldots, x_{i+(m-1)\tau}], \tag{5}$$

where m is the embedding dimension, and τ is the time delay. In order to remove possible autocorrelations due to oversampling of the data,[28] we remove delay vector points with a temporal separation of less than $t=W/F_s$, where F_s is the sampling rate of the data, and W is an integer threshold called the Theiler correction.[29] Incorporating this correction and exploiting the fact that $\|\mathbf{x}_i - \mathbf{x}_j\| = \|\mathbf{x}_j - \mathbf{x}_i\|$, we obtain the standard Theiler-corrected form of the correlation sum:

$$C(N,\epsilon) = F \sum_{k=W}^{N-1} \sum_{i=1}^{N-k} \Theta(\epsilon - \|\mathbf{x}_i - \mathbf{x}_{i+k}\|), \tag{6}$$

where $k=j-i$, and

$$F = 1 \Bigg/ \left(\sum_{k=W}^{N-1} \sum_{i=1}^{N-k} 1 \right) = 2/(N^2 + N - 2WN - W - W^2).$$

For the correlation dimension computation, the quantity of interest is the scaling of the correlation sum with respect to ϵ, thus Eq. (6) typically is repeated for geometrically spaced values of ϵ. Since a point pair within a hypersphere of radius ϵ_d is also within hyperspheres of radius ϵ_{d+n}, where $\epsilon_{d+n} > \epsilon_d$, the correlation sum can be written as the sum of the fraction of point pairs between consecutive discrete values of ϵ_d.[30,31] This can be expressed as

$$\Delta C(N,d) = F \sum_{k=W}^{N-1} \sum_{i=1}^{N-k} [\Theta(-\epsilon_{d-1} + \|\mathbf{x}_j - \mathbf{x}_i\|) \times \Theta(\epsilon_d - \|\mathbf{x}_j - \mathbf{x}_i\|)],$$

$$C(N,d) = \sum_{i=1}^{d} \Delta C(N,i). \tag{7}$$

If we choose $\epsilon_d = \epsilon_0 10^{d/r}$, then from $\Theta(\epsilon_{d-1} - \|\mathbf{x}_j - \mathbf{x}_i\|) = 1$, we find the indices of the corresponding $\Delta C(N,d)$ bin: $d = \mathrm{int}(r \log_{10}\|\mathbf{x}_j - \mathbf{x}_i\| - r \log_{10}\epsilon_0)$.

A number of additional improvements can be made to the basic correlation sum algorithm computed from time-delay vectors.[31] For time-delay embedding, any given coordinate difference may appear in up to m different positions in point difference vectors. These redundant difference computations can be eliminated by computing all the coordinate differences, storing them in an array, and performing a table lookup to compute the interpoint differences. Additional speed improvement can be made by computing the maximal

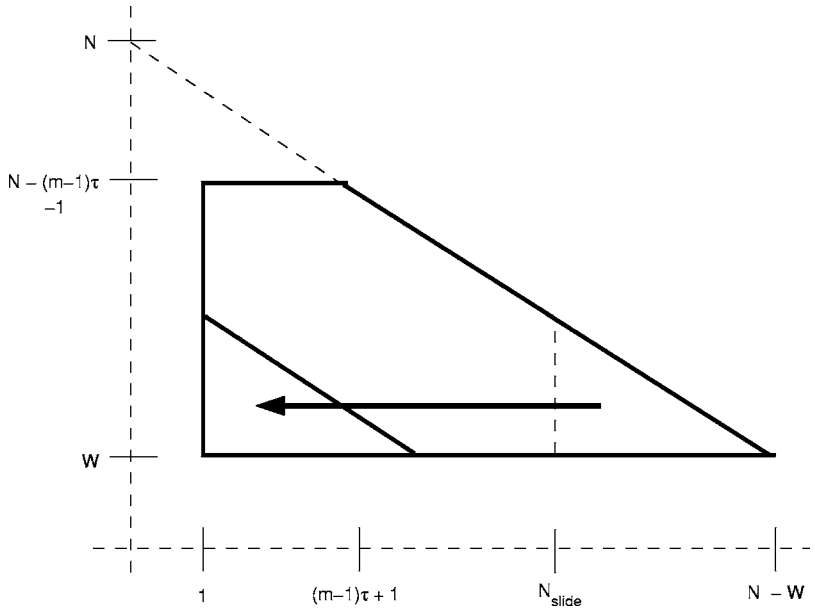

FIG. 1. A schematic of the coordinate difference array is shown. Coordinate differences from the ith window in the triangle at the right are copied into their corresponding positions into the triangle on the left in the $(i+1)$th window. The remainder of the array is filled with the new coordinate differences from the $(i+1)$th window. N is the number of data points per window, W is the Theiler correction, N_{slide} is the number of points between the right-hand indices of consecutive windows, m is the embedding dimension, and τ is the delay time.

coordinate distance for each interpoint difference rather than the Euclidean distance. This is equivalent to changing the geometry of hyperspheres to hypercubes. Switching to the maximal norm also enables us to switch the order of operations of the logarithm and the maximum operators, since the logarithm preserves the sequence order of a set of numbers. This means that the logarithm can be performed at the time of the coordinate differences computation.

Application of this measure to prediction requires that it is applied to a finite window $X_i(j), j=1,\ldots,N$ of time-series data, and then reapplied to following windows $X_k(j), j=1,\ldots,N$ of the same size, whose starting point is offset from the first $X_k(1) > X_i(1)$. To improve temporal resolution in this scheme of sliding windows, there may be overlap between the windows, i.e., $X_k(j) - X_i(j) < N$. In this case, some of the coordinate differences will be in common between adjacent windows. The limits of the coordinate difference array are depicted by the outer boundary of the shape shown in Fig. 1, and the total number of coordinate differences in coordinate difference array D is

$$N_D = 1/2(N-W)^2 - 1/2(m-1)^2\tau^2 + 1/2(N-W-(m-1)\tau).$$

We let $X_k(j) - X_i(j) = N_{\text{slide}}$, so that the points in the coordinate difference array that are in common between the X_kth and X_{k+1}th windows are given by

$$\{D(i,j), i = N_{\text{slide}}+1, \ldots, N-W, j = W, \ldots, N-i\},$$

which totals

$$N_{D\text{slide}} = 1/2(N-W-N_{\text{slide}})(N-W-N_{\text{slide}}-1)$$

points. In the X_{k+1}th window, the indices of the values correspond to $\{D(i,j), i=1,\ldots,N-W-N_{\text{slide}}, j=W,\ldots,i\}$. These values are copied into their new positions in the array, and only the remaining coordinate differences are recomputed.

In any nonstationary time series, such as ECoG, various time-frequency-energy properties of the signal may vary from window to window. We attempt to remove these nonstationary effects prior to computation of the correlation integral-based measures, since these effects are more easily and accurately quantifiable by other measures. One of these properties is the amount of autocorrelation in the signal, but this effect can be compensated for through the Theiler correction. Another simple effect is the amplitude of the data. For a nonperiodic or quasiperiodic signal, there is no standard definition of the signal's amplitude, so here we use rank order statistics to quantify the spread in the data,

$$A = (x_{p_1} - x_{p_2})/2, \tag{8}$$

where x is a window of the scalar time series, and p_1 and p_2 are percentile values of the ranked time series. If the original, unnormalized coordinate difference array is defined by $\{D = r\log_{10}\|x_i - x_{i+k}\|; i = W, N-(m-1)\tau-1; j=1, N-i\}$, the amplitude-normalized coordinate difference array is $D = r\log_{10}\|(x_i - x_{i+k})/A\| + K$, where K is a constant identical for all data windows that shifts the coordinate differences back to the approximate dynamical range of the original values, so that they remain integers. Here, we choose $K = r10\log_{10}2$ for data with 10 bits of precision. Coordinate differences in common between two adjacent windows due to window overlap can be easily renormalized $D_{j+1} = D_j + \log_{10}A_j - \log_{10}A_{j+1}$.

For on-line amplitude normalization based on order statistics, each window is subdivided into sections equal in size to the amount of overlap between windows with an additional section equivalent to the remaining amount for cases where the length of the window is not an integer multiple of the overlap between windows. These subdivisions are stored in a two-dimensional buffer that is circular in one dimension, individually sorted using a quick sort, and then merged together with a merge sort. Then, order-statistics can be determined by the position in the ordered list. In this method, at most only two subdivisions need to be resorted in every window, resulting in a great reduction in the number of operations performed.

III. DATA

Twenty ECoG data sets (anonymized data from a previous HSC-approved study) were selected from consecutive admissions of subjects with pharmaco-resistant seizures who underwent evaluation for epilepsy surgery at the University of Kansas Comprehensive Epilepsy Center. The data were recorded via depth electrodes (Ad-Tech) implanted stereotaxically into the amygdalo-hippocampal region. Correctness of placement was assessed with MRI. The signal was sampled at a rate of 240 Hz, amplified, passed through a 0.5–70 Hz analog bandpass filter, and digitized to 10 bits precision using commercially available devices (Nicolet, Madison WI). The recordings were deemed of good technical quality and suitable for analysis. Data sets were required to be at least 75 h and no more than 125 h long for each subject, and to contain a minimum of five seizures. For subjects with data longer than 125 h, the first 125 h segment of good quality data containing five typical seizures was chosen. The total number of hours of ECoG analyzed was 2347, with a median of 124 h (mean 117 h) per subject.

Seizure scoring was performed using a validated, generic, automated detection method[32,33] and confirmed through visual review by an epileptologist (I. Osorio), who determined false and true positives, as well as visually scoring electrographic onset times. Because the degree of seizure continuum varies, the discriminating statistic in this automated seizure detection algorithm (SDA) may oscillate above and below the preset, generic detection threshold. This can result in multiple, closely spaced detections that could be counted as separate seizures. To eliminate this complicating factor, detections were clustered together using a temporal criterion of 60 s, which was based on the distributions of durations of seizure discontinuities.[33]

After clustering, there were a median of 18 seizure detections (mean 48 detections) per subject. For purposes of parameter training and more detailed analysis of the behavior of the methods, a training data set was chosen consisting of 10 five hour segments of data. Each segment contains at least one typical seizure chosen at random from the time series.

IV. COMPUTATIONAL ISSUES

A. Amplitude normalization

We have previously observed[18,19] that even simple changes in the signal amplitude can cause large changes in the correlation integral or dimension. For instance, consider two windows of ECoG recorded from the same subject, one during the interseizure state with amplitude A_1, and the other during a seizure with amplitude A_2, where $A_2 > A_1$. The correlation integral depends on the number of point pairs separated by at most a distance ϵ. Selecting one arbitrary point $\mathbf{x}_i$, we draw a sphere of radius ϵ around the point and count the number of point pairs enclosed by the sphere. In the interseizure window, the points are clustered into a far tighter group than in the seizure window, and the number of points within the *fixed* sphere of radius ϵ is higher than that generated by the seizure data. This leads to a decrease in the correlation integral during the seizure relative to the interseizure state. If we make the simplistic assumption that the interseizure and seizure data are uniformly distributed in A_1 and A_2, respectively, the probability of a point pair being inside a ball of radius ϵ is proportional to $(\epsilon/A_1)^m$ during the interseizure period, and to $(\epsilon/A_2)^m$ during the seizure period. The relative decrease in the correlation integral then would be on the order of $(A_1/A_2)^m$. While a realistic ECoG signal is not likely to follow a uniform distribution, we intuit that the correlation integral is sensitive to changes in the signal amplitude.

In general, defining the amplitude for a random signal is not as simple as it is in the case of a periodic signal or even a quasiperiodic one. Here, we take the notion that the amplitude can be approximated in terms of rank order statistics according to Eq. (8). One possible definition of amplitude is to use the maximal and minimal values of the windowed scalar time series. However, ECoG often has interseizure spikes. With this definition, any data segment containing one of these spikes would be compressed unreasonably. A more reasonable definition may be to choose an amplitude that represents the typical spread of the data in the window, such as for interdecile ranges (for $p_1=90\%$ and $p_2=10\%$) or interquartile ranges (for $p_1=75\%$ and $p_2=25\%$). In Fig. 2, correlation integral curves are shown for 50 representative seizure (black) and nonseizure (grey) windows. The windows are 15 s in length, and were chosen from the 10 five hour files (one per subject) that composed the training data set. Figure 2(a) shows the curves for the unnormalized signal, while (b)–(d) show the curves for different window-by-window amplitude normalization schemes. In Fig. 2(a), we see that the seizure curves tend to cluster toward larger length scales, while nonseizure curves tend to be at smaller scales. Normalizing the window to the extrema values, as in Fig. 2(b), reduces the spread in the curve locations without changing the shape of the curves. Normalizing to the [10, 90]th and [25, 75]th percentiles further clusters the curves [Figs. 2(c) and 2(d)], though the clustering appears tighter with the [10, 90] normalization scheme [Fig. 2(c)], which appears to be the optimal scheme among these four.

Using the [10, 90] amplitude definition, we now examine the correlation integral time series for correlations with the amplitude of the signal. For each 5 hour file, the correlation integral is computed for embedding dimensions $m=1,\ldots,25$, and $\epsilon=10^{d/20}$, $d=1,\ldots,80$ in 15 s windows with a 10 s overlap between windows. Delay time $\tau=1/12$ s and $W=1$ (no Theiler correction) were kept fixed. The amplitude A is computed for the same 15 s windows. The slope (p) of $\log_{10} C(m,d)$ vs $\log_{10}(1/A)$ is computed using a linear least-squares fit with d fixed, and Pearson's r values is computed to characterize the degree of correlation. For a typical subject, the $C(m=25,d)$ curves for all 1795 windows in the training segment are shown overlaid in Fig. 3(a). In Fig. 3(b), the slopes of the best-fitting line (p) of $\log_{10} C(25,d)$ vs $\log_{10} 1/A$ are plotted versus the length scale d, and the corresponding correlation coefficients are shown in Fig. 3(c). We see that the strongest dependence on amplitude occurs around $d \approx 39$, which corresponds to the higher length scales over which $\log_{10} C(m,d)$ varies. For the same subject, Figs. 4(a) and 4(b) show the slope of the least-squares fit line and Pearson's r values, respectively, varying both m and d. We see that the slope increases with embedding dimension, es-

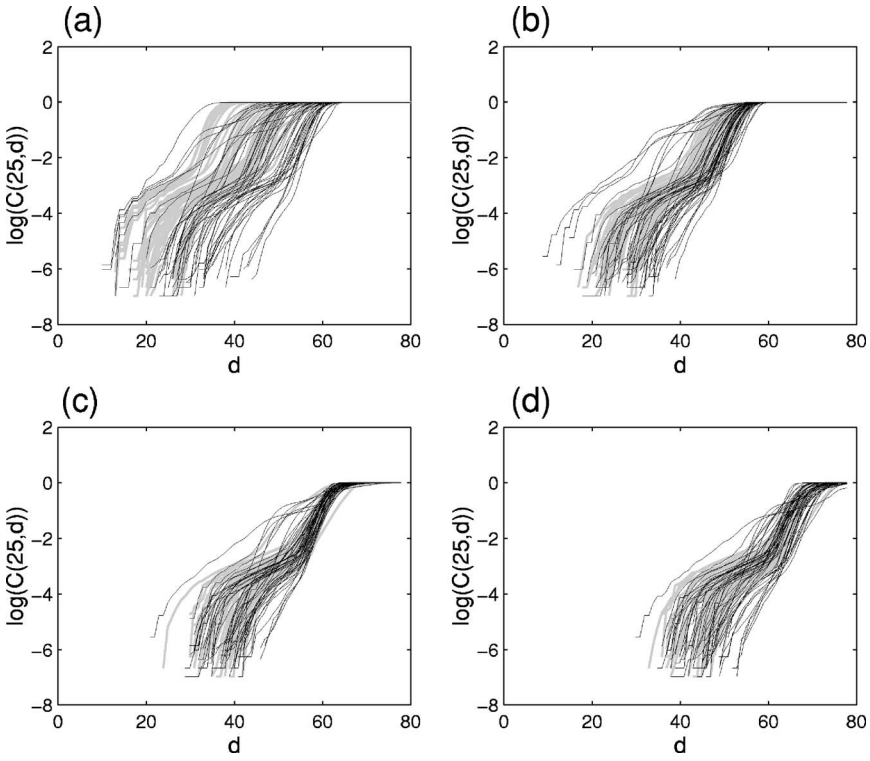

FIG. 2. (a) Correlation integral curves $[\log_{10} C(m=25,d)$, where d is a length scale and m is the embedding dimension] for unnormalized seizure (black) and nonseizure (grey) windows; (b) As in (a), except data are preprocessed by normalizing to each window's extrema; (c) As in (a), except data are normalized with $p_1=90\%$ and $p_2=10\%$; (d) As in (a), except data are normalized with $p_1=75\%$ and $p_2=25\%$.

pecially in the region between $d \approx 30$ and 45. Thus, the larger length scales are most consistent with the stochastic model, in that (1) the value of $\log_{10} C(m,d)$ is highly sensitive to the amplitude of the signal and (2) this sensitivity is dependent on the embedding dimension. Intuitively, this makes sense since our amplitude definition is based on the large absolute deviations of the data rather than the structure at small length scales.

B. Autocorrelation

The correlation dimension D_2 is usually estimated by examining the slope within a linear scaling region of the plot of $\log C_N(m,\epsilon)$ vs $\log \epsilon$ for a series of increasing values of m. For $m < D_2$, the dimension of the reconstructed phase space is not high enough to resolve the structure of the dynamical state and, hence, the slope approximates the embedding dimension. As m increases, the resolution of the dynamical state in the reconstructed phase space improves. For a low-dimensional dynamical system, the slope (S) in the plot of $\log C_N(m,\epsilon)$ vs $\log \epsilon$ increases with m until it reaches a plateau; its value at the plateau is then taken as the estimate of D_2.[24,34] For stochastic dynamics, the slope increases with m, never reaching a plateau.

Once the $\log C$ vs d curves have been normalized window-by-window for amplitude, we compute the slopes of any linear scaling regions present in the data. Figure 5(a) shows the curves for $m=1,\ldots,35$ for a typical window of interseizure data. In the region $\log C=[-3.85,-2.50]$, the slopes of the curves reach a plateau as the value of m is increased. In Figs. 5(b)–5(d), values of the slopes S in this scaling region are plotted for each window for $m=[5,15,25]$, respectively. A seizure (shown by the arrows) occurs at $t=12\,675$ s.

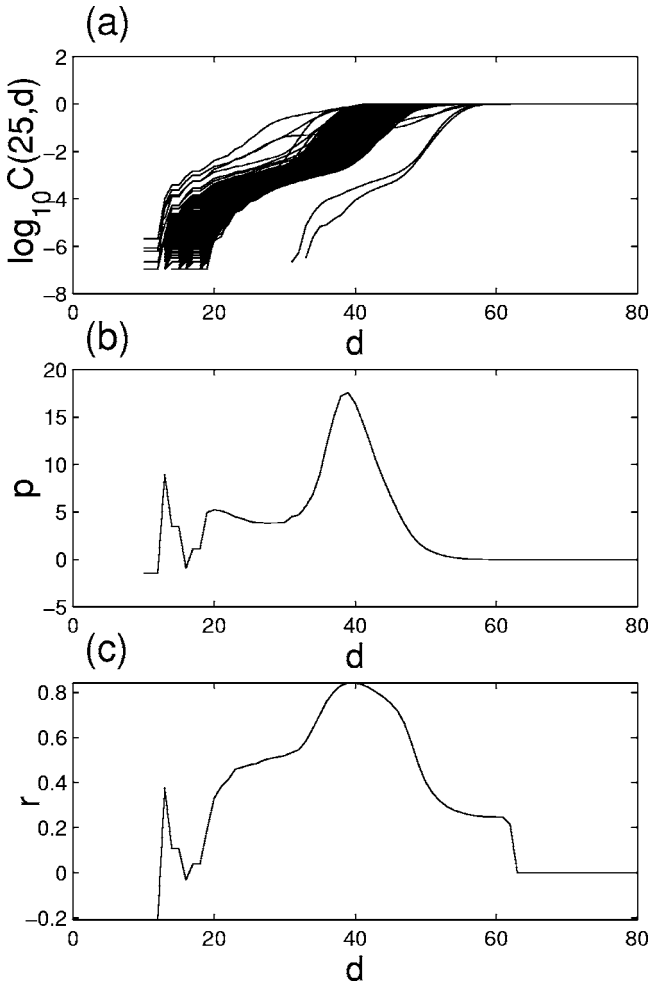

FIG. 3. (a) The logarithm of the correlation integral $[\log_{10} C(m=25,d)]$ vs length scale parameter d. (b) Slopes of the least-squares fit line (p) of $\log_{10} C(m,d)$ vs $\log_{10} 1/A$, plotted vs d. (c) Pearson's r value for the least-squares fit in (b).

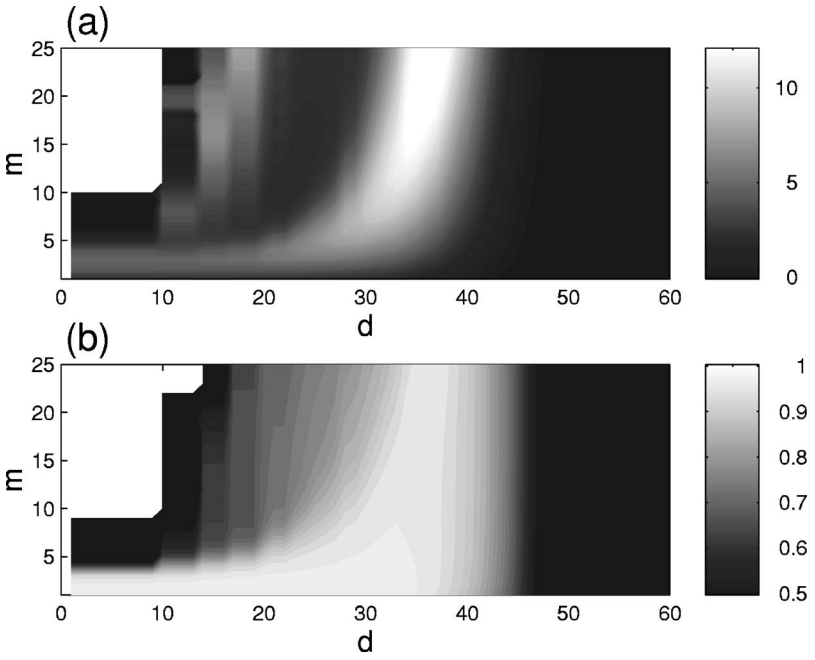

FIG. 4. (a) Slope of best fit line of $\log_{10} C(m,d)$ vs $\log_{10}(1/A)$ as a function of m and d. (b) Pearson's r for the best fit line in (a). The strongest dependence on the amplitude occurs in the region $d \approx 30-45$ for larger m. A is the amplitude of the signal in each window, C is the correlation integral, m is the embedding dimension, and d is a length scale parameter.

Theiler points out[29] that for a finite, autocorrelated data set, the plot of $C(m,\epsilon)$ on a logarithmic scale can exhibit approximately linear regions with distinct slopes, such as those depicted in Fig. 5(a). For a data set $\{x_j\}$, $j=1,\ldots,N$ consisting of autocorrelated Gaussian noise, the autocorrelation α is defined by $\langle x_j x_{j+k}\rangle/\sigma^2=\alpha^k$. We estimate α through the following average: $\alpha=(1/M)\Sigma_{k=1}^{M}|\langle x_j x_{j+k}\rangle/\langle x_j^2\rangle|^{1/k}$, with $m=6$. If this time series has N points with autocorrelation $0<\alpha<1$, Theiler argues that if N is large enough, or if α is small enough (near zero), the effect of autocorrelation is negligible. However, if N is not sufficiently large (as in the case of ECoG analysis where a temporally moving window is slid through the time series) and/or if α is not close to zero, the effect of autocorrelation becomes noticeable, leading to an anomalous scaling region in the plot of $\log C(m,\epsilon)$. The slope of the plot in the anomalous scaling region is not dependent on the embedding dimension, and as such it does not reflect the stochastic nature of the underlying process.

We see from Fig. 6 that the slope of the $\log C(m,\epsilon)$ curves in the anomalous scaling region (a) are highly specific to the detection of the seizure, indicated by an arrow. The decay of the autocorrelation envelope, α (b) also detects the seizure, though is somewhat less specific. But, the number of upcrossings of zero, (c) is highly specific. For a stationary Gaussian process, the expected number of zero upcrossings is related to the autocorrelation through Rice's formula,[35–37]

$$N_{\mathrm{up}} = \frac{1}{2\pi}\left[-\frac{\alpha''(0)}{\alpha(0)}\right]^{1/2}, \tag{9}$$

where $\alpha(0)$ is the autocorrelation at zero lag, and $\alpha''(0)$ is its second derivative. Qualitatively, the seizure discriminating ability of the anomalous slope appears similar to that of the number of zero upcrossings, though for the latter, there is more variance in the measure between seizures, probably related to the variability of α.

The anomalous scaling region can be eliminated by implementing the Theiler correction,[29] which omits point pairs from the correlation sum that are temporally separated by less than some time T. We choose T to be the first zero of the autocorrelation function. Now, we compute T on this basis, window-by-window, to remove the autocorrelative effect. However, when the Theiler correction is performed in this manner, no linear scaling region is apparent up to embedding dimension $m=25$, as shown in Fig. 7 for a nonseizure window. Additionally, the seizure discriminating ability of this window-by-window Theiler corrected method was either greatly reduced or nonexistent for all of the training data.

C. Filtering

To attempt to gain more understanding of the interplay between the correlation integral and the power spectrum of the underlying input data, various filters are applied to the data as a preprocessing step. First, the "seizure component" of the signal is extracted using a 22-coefficient DAUB4 Level 3 wavelet filter, which functions as a bandpass filter roughly in the 10–40 Hz range. This matches the approximate frequency band in which seizure activity typically occurs.[32] Then, the correlation integral iscomputed using the previously described method on the filtered output. Figure 8(a) shows $\log C(m,d=15)$ for $m=1,\ldots,25$. For $m\geqslant 5$, the seizure at $t=1800$ s is easily detectable, but no predictive ability is apparent. The correlation integrals remain distinctly different from each other throughout the 30 min following the seizure.

In order to determine whether the correlation integral might be sensitive to possible preseizure spectral changes outside the seizure activity band, the data are instead filtered to select only the nonseizure component of the signal using a band stop filter that removes the band from 10 to 40 Hz and

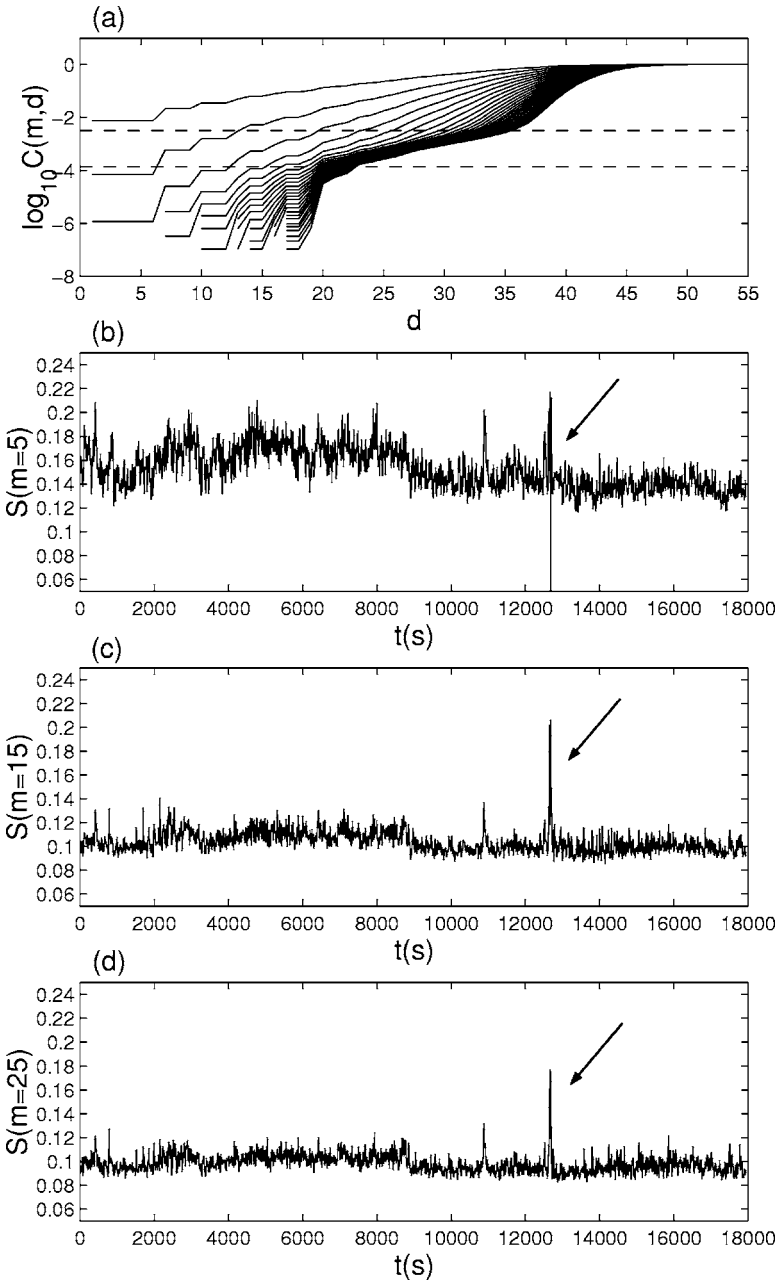

FIG. 5. (a) Anomalous slope (S) of $\log C(m,d)$ vs length scale parameter d for embedding dimension $m=1,\ldots,35$ for a typical window for interseizure data. There is a linear scaling region between the dashed lines. (b) For embedding dimension $m=5$, the slope in the linear scaling region for a 5 h segment of data with a seizure at $t=12\,675$. (c) As in (b), but with $m=15$. (d) As in (b), but with $m=25$.

frequencies greater than 55 Hz (in order to remove 60 Hz line noise). Figure 8(b) shows the correlation integrals at the same parameter settings on the "nonseizure component" of the signal. Sensitivity to seizure is reduced in the nonseizure component. For this subject and at $m=1-14$, there are statistically significant changes within 5 min prior to seizure onset, as measured by a Kolmogorov–Smirnov test[48] to a 95% confidence level. These changes are not generally present at these parameter values for other patients, though some other patients show potential predictive ability in this time frame for other parameter values. There are no parameter regions that consistently show potential predictive ability across multiple patients.

Bandpass filters of width 5 Hz are also applied to the data in order to examine whether discriminating ability is limited to any certain frequency bands. One hour segments of data, each containing at least one seizure, are bandpass filtered into 11 different bands: [0.01, 5], [5, 10],…,[50, 55], and the correlation integral was computed on the resulting 11 filtered data sets. To quantify the ability of the correlation integral to detect seizures in the different bands, a ratio (R) is computed, which is the maximum value of the measure in seizure divided by the mean value of the measure during the nonseizure period. In Fig. 9, we plot R versus the frequency band and see that the ratio maximizes in bands 25–30 and 30–35 Hz. However, the rhythmic seizure patterns occur for this data segment from 12 to 17 Hz, from visual inspection of this seizure's ECoG. Thus, perhaps $C(m,d)$, at least with the parameter settings used in this run, are not even optimal for detection of this seizure activity since the measure does not exhibit seizure discrimination in the dominant seizure frequencies. The segment is too short to accurately quantify the correlation integral's predictive ability in this case, due in

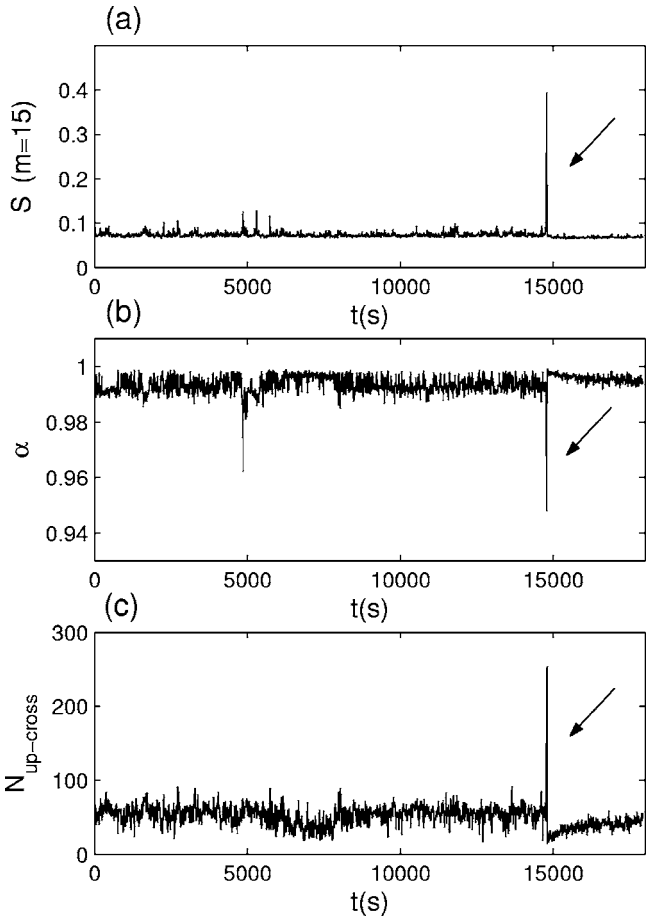

FIG. 6. For a segment of ECoG containing a single seizure (indicated by the arrow), (a) Slope in the anomalous scaling region on $\log C(m,d)$ vs d. (b) Decay in autocorrelation α for the same time series as in (a). (c) Number of zero upcrossings (N_{upcross}) for the same time series as in (a).

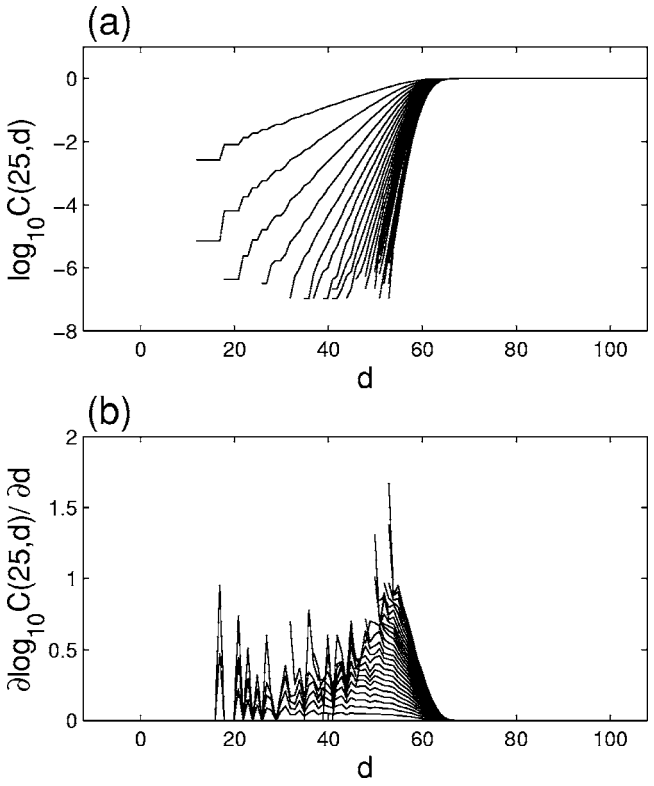

FIG. 7. (a) For $m=1,\ldots,25$, the $\log_{10} C$ vs d curves for a Theiler corrected window. (b) The derivatives of the curves in (a). C is the correlation integral, m is the embedding dimension, and d is a length scale parameter.

part to the fact that clinical seizures are followed by a characteristically different postseizure period that can last 30 min or longer. So, it is not possible to say, for instance, whether activity prior to the second seizure is truly predictive or if it merely represents the postseizure activity gradually decaying to baseline.

D. Embeddings

In addition to time-delay embedding, we have investigated alternate embedding schemes in order to determine whether particular schemes result in improved performance and whether specific embedding schemes differ in the response to time-frequency changes from that of the delay embedded signal. We use the following additional embedding schemes:

(1) Spatial embedding defined by

$$\mathbf{x}(t)=\{x_1(t),x_2(t),\ldots,x_m(t)\}, \tag{10}$$

where m is the number of simultaneous spatially distributed dynamical variables (the number of channels of ECoG).

(2) Spatiotemporal embedding defined by

$$\mathbf{x}(t)=\{x_1(t),x_1(t+\tau),\ldots,x_1(t+(q-1)\tau)\ ,$$

$$x_2(t),x_2(t+\tau),\ldots,x_2(t+(q-1)\tau),$$

$$\vdots$$

$$x_l(t),x_l(t+\tau),\ldots,x_l(t+(q-1)\tau)\}, \tag{11}$$

where $m=ql$ is the embedding dimension, l is the number of spatially distributed time series, and τ is a delay time. The parameters m and τ are selected based on the same considerations as in delay coordinate embedding.

(3) Threshold-crossing interspike intervals (TCII) embedding defined by[38,39]

$$\mathbf{T}_n=\{T_n,T_{n+1},\ldots,T_m\},$$

where T_n is the time between the nth and $(n+1)$th upcrossings of a threshold, x_{th}. This recent embedding strategy has been developed to deal with "spiky" data, such as that recorded from the epileptic brain, and is less sensitive to fluctuations of the amplitude and to noise than other embedding methods.

Delay coordinate and threshold-crossing embeddings are single-channel methods, while spatial and spatiotemporal embeddings require multiple channels. Performing spatial embedding over multiple channels rather than the traditional delay-embedding may be particularly useful to characterize the spatial evolution of seizures. Likewise, a threshold-

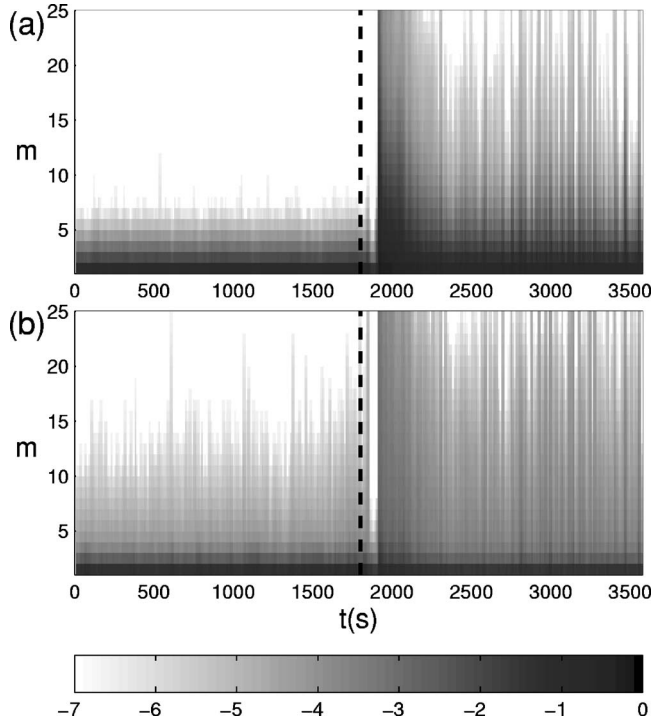

FIG. 8. (a) For data from subject 4 that is preprocessed with the level-3 DAUB4 wavelet filter in order to extract the seizure component, the time series of $\log C(m,d=16)$, computed without Theiler correction or amplitude normalization, and with delay time $\tau=20$. (b) As in (a), except data are preprocessed with a bandpass filter that extracts the [0 10] and [40 55] frequency band, eliminating the seizure component of the signal. Though the seizure at $t=1800$ s is detected by both the seizure and nonseizure components of the signal, no changes are apparent in the measure prior to the electrographic onset. In both plots, C is the correlation integral, m is the embedding dimension, and d is a length scale parameter.

crossing embedding scheme may be sensitive to seizures with preseizure changes such as increases or decreases in "spike" frequency.

Of the embeddings studied, spatiotemporal and threshold-crossing are found to have barriers to practical utility: for spatiotemporal the computational requirements (computations with five channels were taking running 1/6 real time on a 1.9 GHz PC, or on the order of 2.5 days for a 5 h input data file), and for threshold-crossing embedding the severe reduction of the number of computed points, resulting in exceptionally high noise level.

Figure 10 shows the correlation integral computation for the four different embedding schemes, as well as the discriminating statistic from a validated automated seizure detection method.[32] There are seizures at approximately 11,440 and 15,592 s. Parameter values which showed the greatest sensitivity to seizures are used. Except for the threshold-crossing embedding, all embedding methods are found to specifically detect the seizures. In the case of the threshold-crossing embedding [Fig. 10(c)], some sensitivity for the seizures is found, but the method is largely unable to distinguish between seizure and postseizure activity.

Conceptually, it might be thought that the coordinate differences in the spatially embedded correlation integral would reflect the degree of similarity between channels. To examine this further, we compute an index of phase synchronization[40] that examines the interdependence of weakly interacting signals. In this method, we first generate proper rotations by bandpass filtering the windowed signals to a band of $[10,20]$ Hz using a FIR filter, and then compute the analytic signal[41] by using the Hilbert transform. Then, the difference in the complex phases, which are confined in $[0,2\pi]$, are normalized to the interval $[-1,1]$. The interval is divided into $M=e^{0.626+0.4\ln(N-1)}$ bins, where N is the number of points in a window, since this is the optimum number of bins in a histogram.[42] This yields p_k $(k=1,\ldots,M)$, the probability of observing the phase difference in each bin. By summing up $-p_k \ln p_k$, the Shannon entropy can be computed, from which a synchronization index S can be defined.[40] If there is no synchronization, i.e., the phase differences are uniformly distributed in $[-1,1]$, we have $S=0$. If, on the other hand, there is a perfect phase synchronization, then the phase differences distribution follow a delta function, which gives $S=1$. The variation in S from window-to-window thus quantifies the degree of the phase synchronization between the two ECoG channels as a function of time.

In Fig. 11, we see that the phase synchronization measure is more specific for detection of the seizures, with better specificity and lower signal-to-noise ratio. The spatially embedded correlation integral, on the other hand, does not even detect the first of the two seizures.

E. Surrogate analysis

To test the null hypothesis that the observed dynamics of the ECoG are due to a linear Gaussian process, the method of surrogate data analysis is used. Since we have found that the signal amplitude influences the value of the correlation integrals, we choose the amplitude adjusted Fourier transform surrogate.[43] This surrogate specifically tests the null hypothesis that the observed time series is a monotonic nonlinear

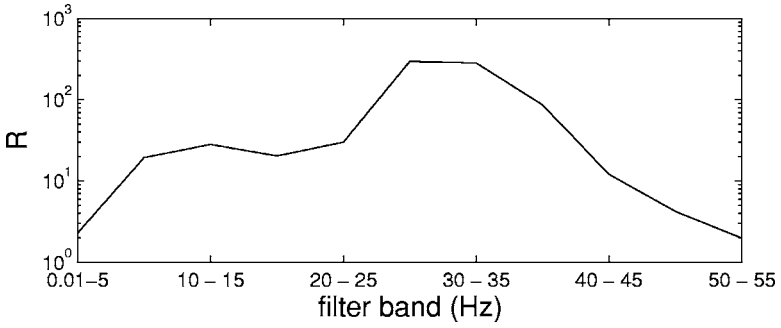

FIG. 9. For subject 3, R is the ratio of the maximum value of $\log C(m=25,d=23)$ in seizure with respect to the mean value between seizures. The data are preprocessed by bandpass filtering to restrict the frequency to bands 5 Hz wide, without amplitude normalization or Theiler correction, and with delay time $\tau=20$. There are two seizures, one at $t\approx250$ s, and one at $t\approx1820$ s. The measure is most sensitive to seizures for 25–35 Hz.

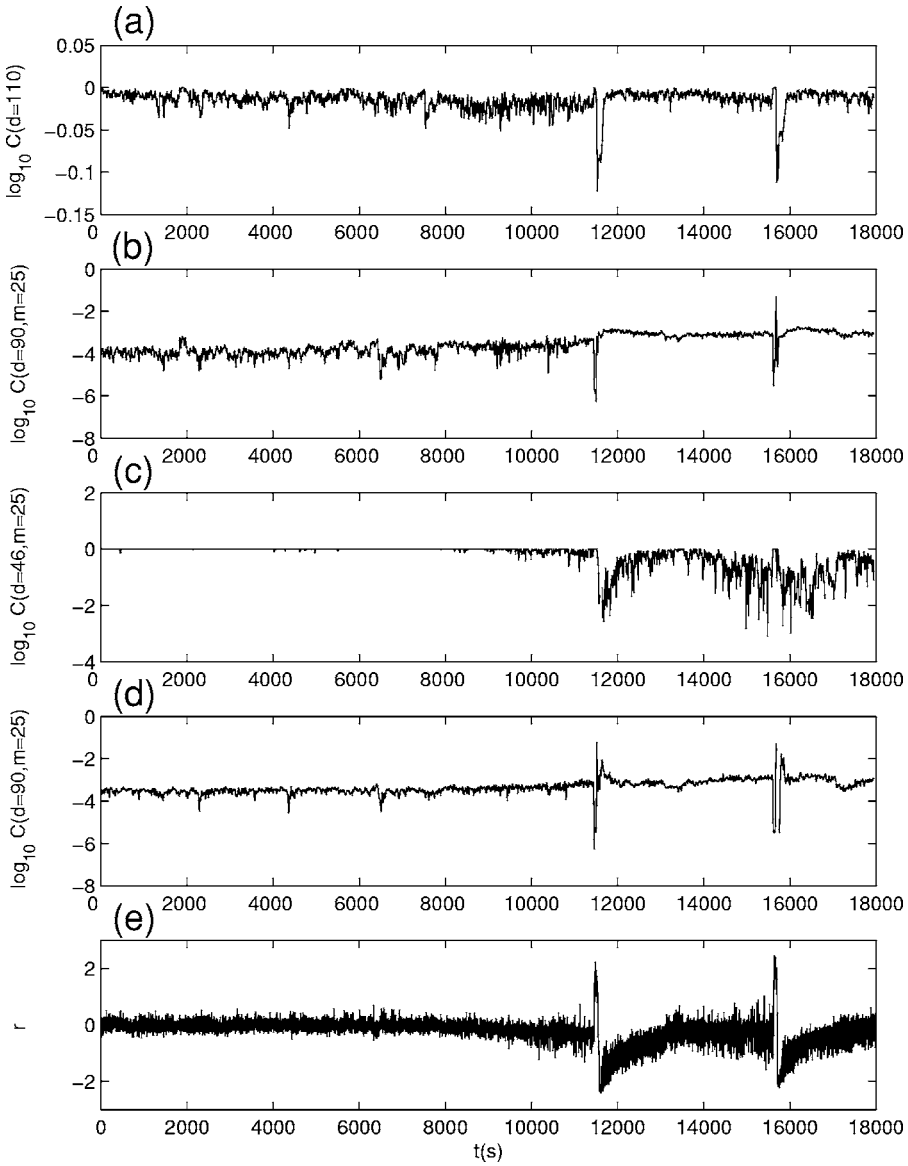

FIG. 10. (a) Log $C(d=110)$ for data spatially embedded across five differential channels; (b) $\log C(m=25,d=90)$ for data spatially embedded across five differential channels and a delay-time embedding dimension of $m=25$; (c) $\log C(d=46,m=25)$ with embedding constructed of $m=25$ successive upcrossings of zero; (d) $\log C(d=90,m=25)$ for delay-time embedding with $m=25$; (e) discriminating measure (r) from an automated seizure detection algorithm (Ref. 32). In all these plots, seizures occur at approximately 11 440 and 15,592 s. C is the correlation integral, m is the embedding dimension, and d is a length scale parameter.

transformation of a linear Gaussian process. Given an original scalar time series x of length N, we first generate a time series y also of length N, consisting of independently generated Gaussian random numbers. Gaussian time series y is then reordered so that the ranks of the elements match those of the original time series x. Thus, if x_i is the pth largest value in x, y is reordered so that y_i is the pth largest value in y. The Fourier transform is performed on y, and the phases are randomized by multiplying each complex amplitude by $e^{i\phi}$, where ϕ is chosen uniformly and randomly from the interval $[0,2\pi]$. Then, the inverse Fourier transform is taken to produce y', and x is reordered so that the ranks of its elements match those of y'. This surrogate time series has the same amplitude distribution as the original time series.

Here, we wish to gain insight into whether the nonlinear or the time-energy-frequency properties of the signal contribute to the predictability or detectability of seizures. To do this, surrogate data is computed window-by-window, and the correlation integrals are computed for each of these windows and compared with the correlation integrals computed on the original data.

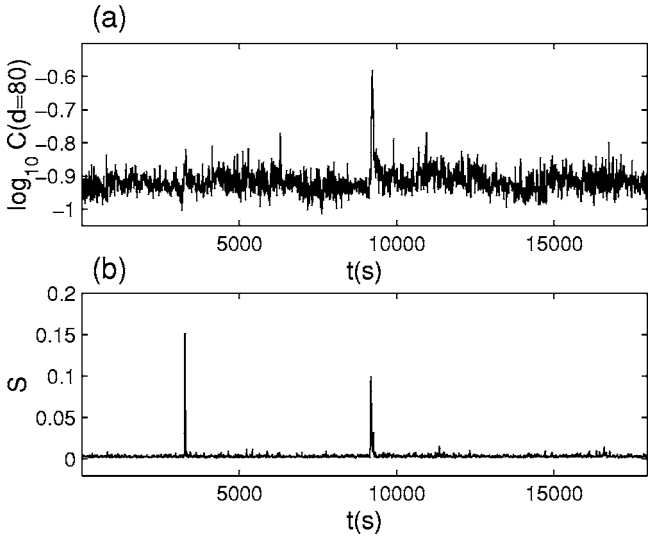

FIG. 11. (a) For a spatially embedded run, $\log_{10} C(d=90)$ vs t, (b) synchronization index S. There are two seizures, one at $t\approx3260$ and $t\approx9150$ s. C is the correlation integral and d is a length scale parameter.

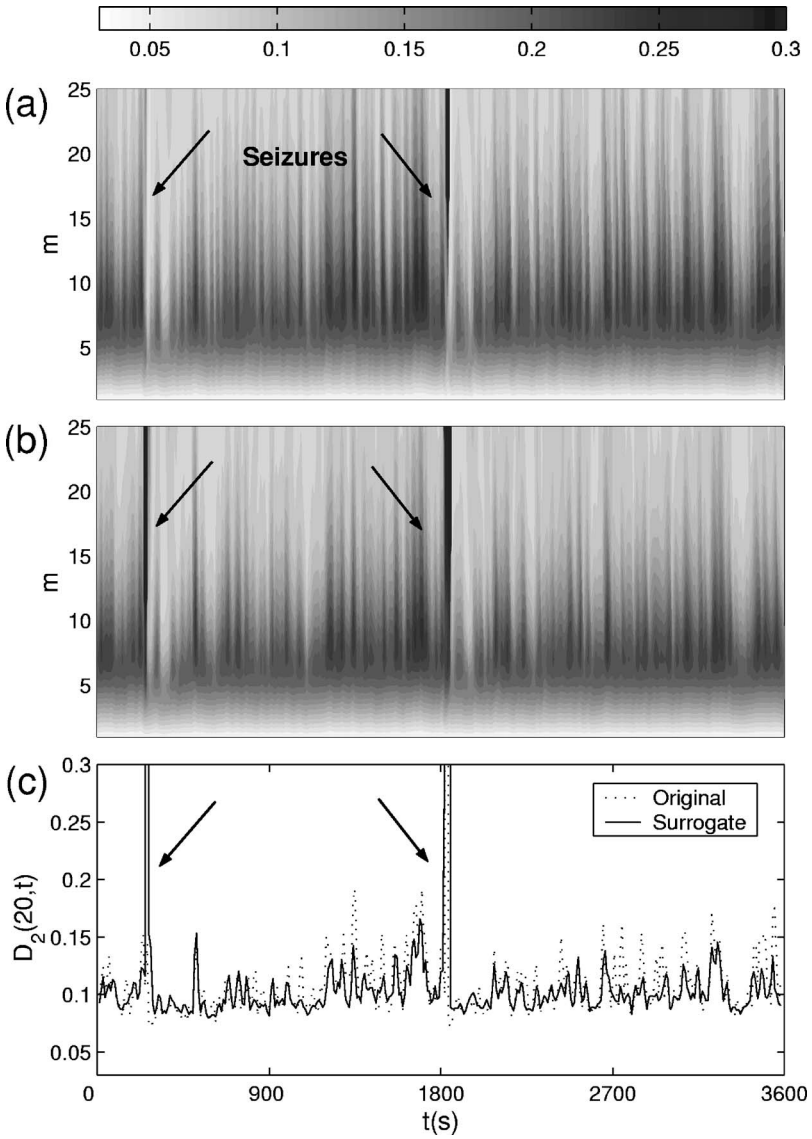

FIG. 12. For a 1 h segment of data with two seizures taken from subject 3, (a) the time series of correlation dimension $D_2(m)$ for the original data, (b) the time series of $D_2(m)$ for the surrogate data, and (c) a single time series at embedding dimension $m=20$ of the original and surrogate data, overlaid on the same axis. Since the seizure detection ability is approximately the same or is actually somewhat improved for the surrogate data vs the original, the implication is that the time-frequency-energy properties of the signal are what detects the seizure, rather than any possible change in the dynamical structure of the signal. Seizures occur at approximately 250 and 1820 s.

Figure 12(a) depicts for subject 3, the time series of the correlation dimension time-locked for a 1 h segment containing two seizures (indicated by arrows). Figure 12(b) shows the same from the surrogate time series of the same data. Surrogate and original D_2 time series are shown overlaid for $m=20$ in Fig. 12(c). The seizure discriminating ability is roughly the same in the surrogate and original time series for one seizure, and is actually improved in the surrogate time series for the other seizure. No predictive ability is apparent in either. Since the seizure detection ability i preserved in the surrogate time series, the implication is that it is the time-frequency-energy properties of the signal that are responsible for the measure's seizure discriminating ability, not the lower-level dynamics of the brain.

V. VALIDATION ON LONG TIME SERIES

In order to validate any seizure detection or prediction abilities of the correlation integral-based measures, we apply the method to the testing data set, consisting of the long time series from 20 subjects described earlier. Based on previous analysis, parameter settings are chosen that eliminate the influence of known time-frequency-energy properties more easily quantifiable through other means, while at the same time providing the maximum seizure-signal content of the ECoG. For these validation runs, the parameters are chosen to be: delay-embedding dimensions $m=[5,20,35]$, $\tau=17$, and Theiler correction $W=30$, based on a sampling rate of $F_s=240$ Hz. An interdecile range amplitude normalization scheme is used, and the length of the sliding windows is chosen to be 15 s, overlapped by 5 s.

The resulting time series is divided into three epochs, time-locked to each seizure onset: 15–30, 45–60, and 75–90 min before seizure onset. These length scales are chosen to be comparable with those cited in the literature as the seizure prediction times for various measures,[44–46] though a few methods using time-frequency-energy techniques have reported longer prediction times. Preseizure epochs are included for each seizure only if the previous seizure has occurred 30 min earlier than the beginning time of the epoch, in order to permit the signal to return to a baseline state after the postseizure period.[47] To quantify whether there is a statistically significant change in these sets of epochs that

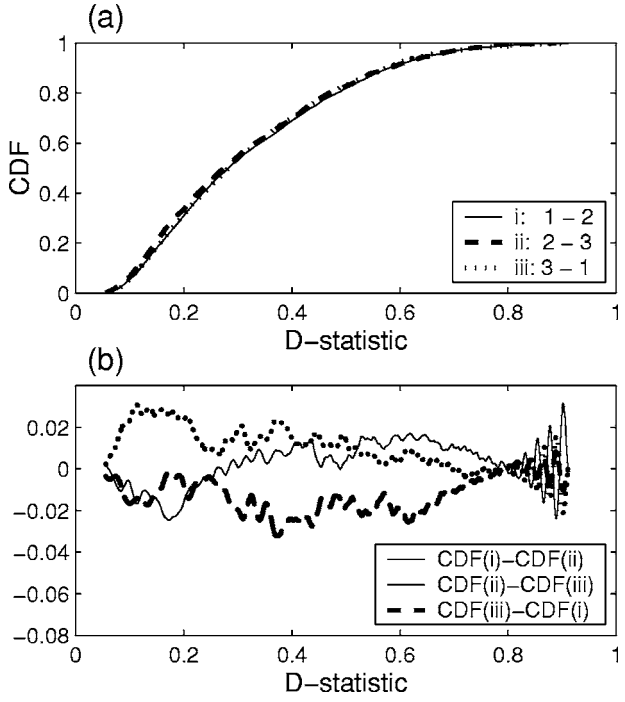

FIG. 13. (a) Cumulative distributions (CDF) of Kolmogorov—Smirnov (KS) *d*-statistic between all pairs of 15 min epochs *between* the following groups: (1) 15–30 min before seizure, (2) 45–60 min before seizure, and (3) 75–90 min before seizure. (b) Differences in the distributions in (a).

would indicate predictive ability between 15 and 90 min before seizure, the Kolmogorov-Smirnov *d*-statistic[48] is computed on every possible pair of epochs between two different sets, e.g., between the set of 15–30 min epochs and the set of 45–60 min epochs.

The Kolmogorov–Smirnov (KS) test is a nonparametric test of whether two distributions are different. The KS statistic is defined as the maximum value of the absolute difference between two cumulative distributions:

$$D = \max_{-\infty<x<\infty} |P_i(x) - P_j(x)|. \tag{12}$$

The cumulative distribution of the resulting *D*-statistic is shown in Fig. 13(a), and the differences in the cumulative distributions are shown in 13(b). Computing the Kolmogrov–Smirnov statistic between these cumulative distributions leads to no significant difference between them to the $p=0.05$ level ($p_{\text{i-ii}} \approx 0.10, p_{\text{ii-iii}} \approx 0.21, p_{\text{iii-i}} \approx 0.52$).

To evaluate whether there is any short-term predictive ability (on a time scale less than 15 min), we overlay correlation integral from preseizure time series 0–15 min before seizure onset. Then, we compute the 0th, 50th, and 90th percentiles of the measure at every time index, as shown in Fig. 14 for subject 2, and Fig. 15 for subject 5. For some subjects, a sudden increase in the value of the measure, especially at lower values of *d*, is seen within 1 min prior to SDA detection, as is the case in Fig. 15(a). In these cases, this is typically due to a late detection of electrographic onset by the automated seizure detection algorithm, as confirmed by visual review. In one case (subject 20), precursor activity in the form of quasiperiodic complexes is detected by the correlation integrals a few minutes before electrographic onset. No statistically significant changes are found in the 50th percentile values between the first and last 7.5 min for any of these overlay plots, as measured by a KS test.

We also quantify the sensitivity of the correlation integral measures to detect seizures. Using the seizure detection scoring provided by the automated seizure detection algo-

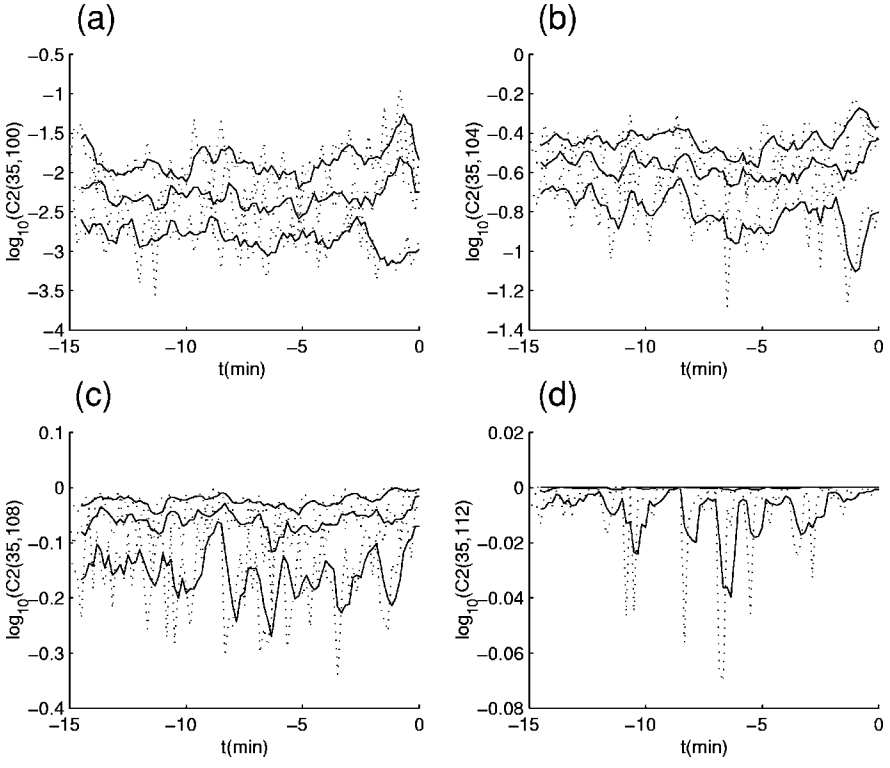

FIG. 14. ECoG segments 0–15 min before seizure are overlaid, and percentage levels computed at each time index are computed. (a) The 0th, 50th, and 100th percentiles of the distribution $\log_{10} C(35,100)$ at every time index for overlaid preseizure epochs for subject 2. The dashed lines are the original percentile levels, and the solid lines are ten point moving averages. (b)–(d) The same as in (a), but for $\log_{10} C(35,104)$, $\log_{10} C(35,108)$, and $\log_{10} C(35,112)$, respectively.

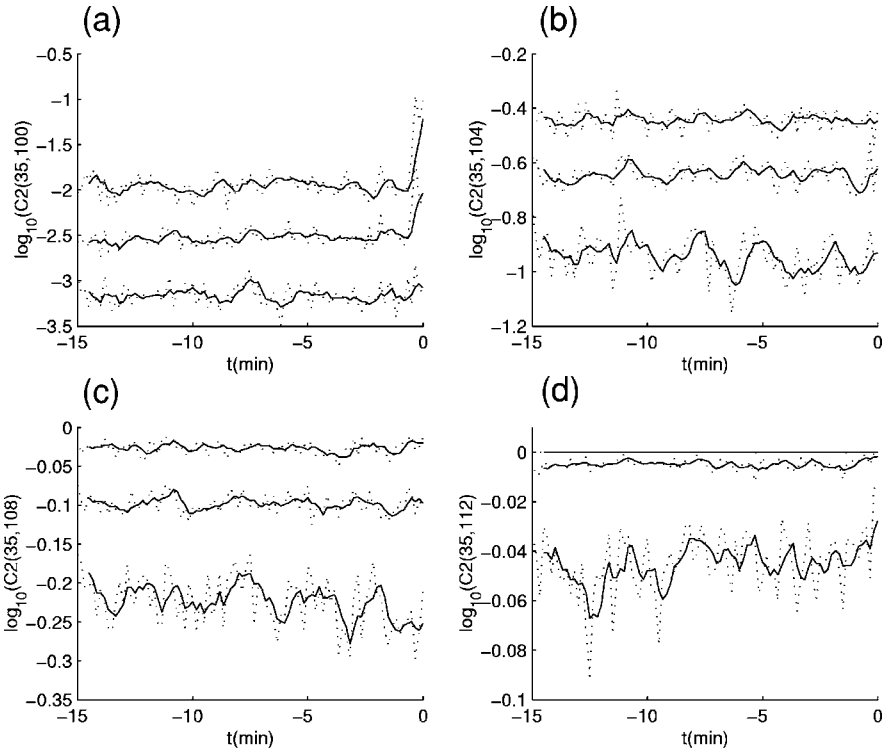

FIG. 15. ECoG segments 0–15 min before seizure are overlaid, and percentage levels at each time index are computed. (a) The 0th, 50th, and 100th percentiles of the distribution $\log_{10} C(35,100)$ at every time index for overlaid preseizure epochs for subject 5. The dashed lines are the original percentile levels, and the solid lines are ten point moving averages. (b)–(d) The same as in (a), but for $\log_{10} C(35,104)$, $\log_{10} C(35,108)$, and $\log_{10} C(35,112)$, respectively.

rithm and confirmed by expert visual analysis, the time series are divided into seizure and nonseizure periods. For correlation values that exceed a specific threshold, values from seizure windows are considered true positive detections, and those from all other windows are considered false positive detections. The number of true or false positives is divided by the number of seizure or nonseizure windows. The threshold can be varied to parametrize a curve (ROC curve) indicating the relationship between true positives and false positives as the threshold changes. The area under the curve is denoted by γ. Figure 16 shows γ for the parameter values in d in the correlation integral that generated the largest γ (solid line), in order to assess the best performance of the measure for each patient. To provide a basis of comparison of the measure's seizure detection ability, we also compute the SDA discriminating statistic[32,33] for the same windows of data as were used for the correlation integral. This is shown as the dashed line in Fig. 16. The correlation integral apparently has no advantage over the SDA in detecting seizures.

VI. DISCUSSION

The current state of the seizure prediction field is that there is little to no consensus on whether seizures are predictable to a degree useful for clinical application, or whether the methods used by any of the groups are reliable. There is also no consensus on the time scale of a distinctive preseizure state though there is speculation that some such state exists.[44–46,49,50]

One class of seizure prediction methods are those based on the correlation integral,[8–10,23] which includes the correlation density, the related dynamical similarity index, and the effective correlation dimension. Retrospective studies with these measures yield seizure prediction times on the order of 10 min.[8–10] Our motivation for this study is to understand precisely what type of signal characteristics these measures are sensitive to, and also to validate the fundamental component of the methods, the correlation integral, on long time series. We previously performed preliminary analyses of short ECoG segments which elucidated the measures' dependence on amplitude and autocorrelation.[18,19] This study is the first to perform a systematic, large-scale study on correlation integral-based measures that not only meticulously examined

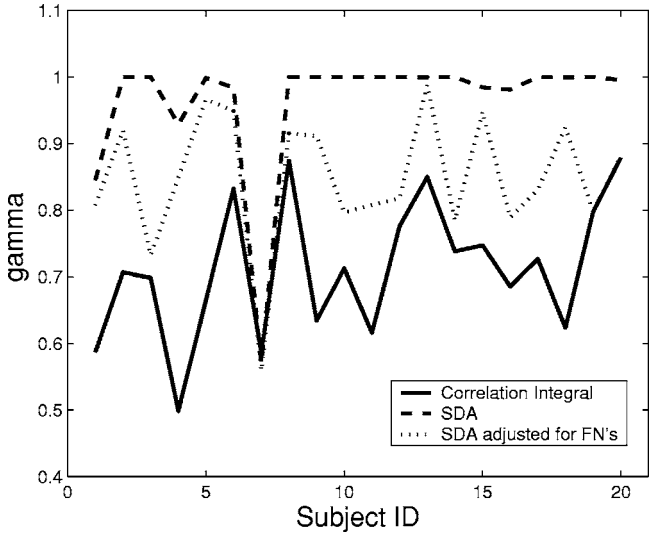

FIG. 16. Maximum area under ROC curve for $\log C(m=25,d)$, $d=100+2i$, $i=1,\ldots,6$ (solid line); area under the ROC curve for the maximum seizure detection algorithm ratio in 15 s windows, overlapped by 10 s (dashed line); area under the ROC curve for the maximum ratio corrected for a worst-case false negative rate (dotted line). C is the correlation integral, m is the embedding dimension, and d is a length scale parameter.

the sensitivity to the individual parameters and method variations, but also tested the method on long time series, including continuous computation on over 2000 h of ECoG data.

We have confirmed in this comprehensive work that seizure detection measures based on the correlation integral are sensitive to time-frequency-energy signal changes such as those quantified by autocorrelation and signal amplitude,[18,19] which we robustly measure using the spread of the windowed data measured with interdecile or interquartile ranges. Window-by-window amplitude normalization and Theiler correction[29] can eliminate these effects but significantly reduce the ability of the measures to discriminate seizures from interseizure periods. This suggests that the seizure detection ability of the correlation-integral based measures is directly linked to the signal's variation in the time-frequency-energy characteristics rather than to any nonlinear dynamics present in the ECoG. This is supported by our results with surrogate data, which illustrate that the seizure detection ability is approximately equivalent for both original and shuffled data.

No significant improvement in seizure detection or prediction ability is seen when the data are preprocessed through filtering to extract frequency bands that may be prominent in seizure or "precursor" activity. Likewise, different embedding schemes do not substantially improve the measures' abilities to detect or predict seizures.

No evidence of prediction is found on a time scale of 15–90 min before seizure onset, and on overlaid time series within a scale of 0–15 min before seizure; the only advanced warning is less than 1 min prior to seizure for some seizures in some subjects. However, advanced warnings can be easily obtained through subject-tuning of an existing automated seizure detection algorithm[32] which is far more computationally efficient than the correlation integral method. Furthermore, the correlation integral method is far less specific in terms of false positives than the discriminating statistic used in the automated seizure detection algorithm, thus the correlation integral probably has little value as a seizure detector, let alone a predictor.

Other measures with claimed predictive ability need to be rigorously tested in a manner similar to what has been done here to examine the validity of prediction claims. We have reported previously preliminary work on the Lyapunov exponents that demonstrates their inability to detect slow parameter drift into a dynamical crisis in a noisy dynamical system (such as the brain[51]), an analogy to a dynamical route to seizure that would justify usage of nonlinear dynamical measures.[52] A recent study which examines the sensitivity and specificity of a large collection of measures used in the literature for seizure prediction found that univariate measures, including many nonlinear dynamical measures like the correlation dimension, were inadequate for clinical application.[49] Those results and ours suggest that seizure prediction is far from being a solved challenge. Studies such as these need to be expanded, particularly focusing on validating prediction claims with Lyapunov exponents, synchronization measures, and the accumulated energy, which are currently the primary focus of the seizure prediction community.[44,50,53–55]

ACKNOWLEDGMENTS

This work was supported in part by NIH under Grant No. 1R43NS43100-01, and in part by internal funding at Flint Hills Scientific, L.L.C.

[1] M. R. Guevara, in *Concepts and Techniques in Bioelectric Measurements: Is the Medium Carrying the Message?*, edited by J. Billette and A.-R. LeBlanc (Èditions de l'École Polytechnique, Montréal, 1997).
[2] A. Babloyantz and A. Destexhe, Proc. Natl. Acad. Sci. U.S.A. **83**, 3513 (1986).
[3] P. E. Rapp, T. R. Bashore, J. Martinerie, A. M. Albano, and A. I. Mees, Brain Topogr **2**, 99 (1989).
[4] G. Mayer-Kress and S. P. Layne, Ann. N.Y. Acad. Sci. **504**, 62 (1987).
[5] P. E. Rapp, I. D. Zimmerman, E. P. Vining, N. Cohen, A. M. Albano, and M. A. Jimenezmontano, J. Neurosci. **14**, 4731 (1994).
[6] L. D. Iasemidis, J. C. Sackellares, H. P. Zaveri, and W. J. Williams, Brain Topogr **2**, 187 (1990).
[7] M. L. V. Quyen, J. Martinerie, C. Adam, and F. J. Varela, Phys. Rev. E **56**, 3401 (1997).
[8] K. Lehnertz and C. E. Elger, Phys. Rev. Lett. **80**, 5019 (1998).
[9] C. E. Elger and K. Lehnertz, Eur. J. Neurosci. **10**, 786 (1998).
[10] J. Martinerie, C. Adam, M. L. V. Quyen, M. Baulac, S. Clemenceau, B. Renault, and F. J. Varela, Nat. Med. **4**, 1173 (1998).
[11] S. J. Schiff, Nat. Med. **4**, 1117 (1998).
[12] L. Diambra and C. P. Malta, Phys. Rev. E **59**, 929 (1999).
[13] L. M. Hively, V. A. Protopopescu, and P. C. Gailey, Chaos **10**, 864 (2000).
[14] R. Q. Quiroga, J. Arnhold, K. Lehnertz, and P. Grassberger, Phys. Rev. E **62**, 8380 (2000).
[15] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, Phys. Rev. E **64**, 061907 (2001).
[16] L. D. Iasemidis, P. Pardalos, J. C. Sackellares, and D. S. Shiau, J. Comb. Optim. **5**, 9 (2001).
[17] R. Porcher and G. Thomas, Phys. Rev. E **64**, 010902 (2001).
[18] I. Osorio, M. A. F. Harrison, Y.-C. Lai, and M. G. Frei, J. Clin. Neurophysiol. **18**, 269 (2001).
[19] Y.-C. Lai, I. Osorio, M. A. F. Harrison, and M. G. Frei, Phys. Rev. E **65**, 031921 (2002).
[20] K. K. Jerger, T. I. Netoff, J. T. Francis, T. Sauer, L. Pecora, S. L. Weinstein, and S. J. Schiff, J. Clin. Neurophysiol. **18**, 259 (2001).
[21] R. Ashenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. U. Voss, J. Timmer, and A. Schulze-Bonhag, Brain **126**, 2616 (2003).
[22] P. E. McSharry, L. A. Smith, and L. Tarassenko, IEEE Trans. Biomed. Eng. **50**, 628 (2003).
[23] D. Lerner, Physica D **97**, 563 (1996).
[24] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[25] H. G. E. Hentschel and I. Procaccia, Physica D **8**, 435 (1983).
[26] E. Ott, *Chaos in Dynamical Systems*, 2nd Ed. (Cambridge University Press, New York, 2002).
[27] F. Takens, in *Lecture Notes in Mathematics* (Springer, Berlin, 1981), Vol. 898.
[28] J. Theiler, Phys. Rev. A **41**, 3038 (1990).
[29] J. Theiler, Phys. Rev. A **34**, 2427 (1986).
[30] T. S. Parker and L. O. Chua, *Practical Numerical Algorithms for Chaotic Systems* (Springer, New York, 1989).
[31] G. Widman, K. Lehnertz, P. Jansen, W. Meyer, W. Burr, and C. E. Elger, Physica D **121**, 65 (1998).
[32] I. Osorio, M. G. Frei, and S. B. Wilkinson, Epilepsia **39**, 615 (1998).
[33] I. Osorio, M. G. Frei, J. Giftakis, T. Peters, J. Ingram, M. Turnbull, M. Herzog, M. T. Rise, S. Schaffner, R. A. Wennberg, T. S. Walczak, M. W. Risinger, and C. Ajmone-Marsan, Epilepsia **43**, 1522 (2003).
[34] M. Ding, C. Grebogi, E. Ott, T. Sauer, and J. A. Yorke, Phys. Rev. Lett. **70**, 3872 (1993).
[35] S. O. Rice, in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York, 1954).
[36] J. L. Bendat, *Principles and Applications of Random Noise Theory* (Wiley, New York, 1958).
[37] I. F. Blake and W. C. Lindsey, IEEE Trans. Inf. Theory **19**, 295 (1973).
[38] T. Sauer, Phys. Rev. Lett. **72**, 3811 (1994).
[39] N. B. Janson, A. N. Pavlov, A. Neiman, and V. S. Anishchenko, Phys. Rev. E **58**, R4 (1998).
[40] P. Tass, M. G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, A. J. Volkmann, A. Schnitzler, and H.-J. Freund, Phys. Rev. Lett. **81**, 3291 (1998).

[41]M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, Phys. Rev. Lett. **76**, 1804 (1996).
[42]J. Bhattacharya, Acta Neurobiol. Exp. (Warsz) **61**, 309 (2001).
[43]J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Physica D **58**, 77 (1992).
[44]B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, M. Dichter, and G. Vachtsevanos, Neuron **30**, 51 (2001).
[45]B. Litt and J. Echauz, Lancet Neurology **1**, 22 (2002).
[46]B. Litt and K. Lehnertz, Curr. Opin. Neurol. **15**, 173 (2002).
[47]The discriminating statistic of the SDA typically becomes negative during the postseizure period, which is known to have different dynamics than interseizure periods. We took the length of the postseizure period to be the time from the end of the seizure to the first zero up-crossing of this SDA statistic, and assessed this length for the collection of seizures from all 20 of the subjects. For 94.0% of the seizure events, the discriminating statistic had returned to baseline within 30 min of the seizure.
[48]D. A. Darling, Ann. Math. Stat. **28**, 823 (1957).
[49]K. Lehnertz, F. Mormann, T. Kreuz, R. G. Andrzejak, C. Rieke, P. David, and C. E. Elger, IEEE Eng. Med. Biol. Mag. **22**, 57 (2003).
[50]L. D. Iasemidis, D. S. Shiau, W. Chaovalitwongse, J. C. Sackellares, P. M. Pardalos, J. C. Principe, P. R. Carney, A. Prasad, B. Veeramani, and K. Tsakalis, IEEE Trans. Biomed. Eng. **50**, 616 (2003).
[51]J. A. White, J. T. Rubinstein, and A. R. Kay, Trends Neurosci. **23**, 131 (2000).
[52]Y.-C. Lai, M. A. F. Harrison, M. G. Frei, and I. Osorio, Phys. Rev. Lett. **91**, 068102 (2003).
[53]L. D. Iasemidis, P. M. Pardalos, D. S. Shiau, W. Chaovalitwongse, K. Narayanan, S. Kumar, P. R. Carney, and J. C. Sackellares, Optimization Methods Software **18**, 81 (2003).
[54]F. Mormann, R. G. Andrzejak, T. Kreuz, C. Rieke, P. David, C. E. Elger, and K. Lehnertz, Phys. Rev. E **67**, 021912 (2003).
[55]F. Mormann, T. Kreuz, R. G. Andrzejak, P. David, K. Lehnertz, and C. E. Elger, Epilepsy Res. **53**, 173 (2003).