Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction

Björn Schelter[a] and Matthias Winterhalder
FDM, Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany and Bernstein Center for Computational Neuroscience Freiburg, University of Freiburg, Germany

Thomas Maiwald
FDM, Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany

Armin Brandt
Epilepsy Center, University Hospital of Freiburg, Breisacher Str. 64, 79106 Freiburg, Germany

Ariane Schad
FDM, Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany

Andreas Schulze-Bonhage
Bernstein Center for Computational Neuroscience Freiburg, University of Freiburg, Germany and Epilepsy Center, University Hospital of Freiburg, Breisacher Str. 64, 79106 Freiburg, Germany

Jens Timmer
FDM, Freiburg Center for Data Analysis and Modeling, University of Freiburg, Eckerstr. 1, 79104 Freiburg, Germany and Bernstein Center for Computational Neuroscience Freiburg, University of Freiburg, Germany

(Received 5 May 2005; accepted 24 October 2005; published online 12 January 2006)

Nonlinear time series analysis techniques have been proposed to detect changes in the electroencephalography dynamics prior to epileptic seizures. Their applicability in practice to predict seizure onsets is hampered by the present lack of generally accepted standards to assess their performance. We propose an analytic approach to judge the prediction performance of multivariate seizure prediction methods. Statistical tests are introduced to assess patient individual results, taking into account that prediction methods are applied to multiple time series and several seizures. Their performance is illustrated utilizing a bivariate seizure prediction method based on synchronization theory. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2137623]

Epileptic seizures present a nonlinear dynamic phenomenon. Seizures are generated by an abnormal synchronization of neurons which is unforeseeable for the patients. A reliable and timely prediction would considerably increase the quality of life of those patients who cannot be treated successfully by common therapeutic strategies. Analysis techniques from the theory of nonlinear dynamics could thus make an important contribution to the medical treatment of a severe neurological disease if a reliable prediction was guaranteed. But the lack of a widely accepted methodology to evaluate and compare seizure prediction algorithms has so far led to contradicting claims about their performance and hampered their practical application. Here, we present a methodology to assess seizure prediction performance and a statistical test to judge its superiority to a random predictor. This methodology is then demonstrated in an application to invasive electroencephalographic recordings using a quantity measuring phase synchronization.

I. INTRODUCTION

Time series analysis techniques originating from the theory of nonlinear dynamics and chaos theory have widely been applied to biomedical time series in recent years.[1–11] The prediction of epileptic seizures by applying time series analysis techniques to electroencephalography (EEG) data obtained from patients with epilepsy is of particular interest.[12–14]

So far, a certain number of epilepsy patients cannot be treated successfully by common therapeutic strategies, so they have to endure unforeseen seizures. A precise prediction at an early stage before seizure onset would offer new therapeutic options such as applying electric stimuli or seizure warning devices.[15–17] Several time series analysis techniques based on the theory of nonlinear dynamics[7–10,18–21] or the

[a]Electronic mail: schelter@fdm.uni-freiburg.de

 © 2006 American Institute of Physics

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

theory of linear stochastic processes[22–27] have been applied to invasive and scalp EEG data. Significant changes in the EEG dynamics in a range between seconds up to hours in advance of seizure onsets have been reported.[13,14] These studies have strengthened the hope that not only *interictal* states between seizures but also specific *preictal* states preceding seizures exist. The existence of preictal periods is the basic requirement for genuine seizure prediction.

When a focal seizure is generated, synchronized epileptic brain activity is initially observed only in a small area of the brain. From this focus, the activity spreads out to other brain areas. Provided that there is any information about an impending seizure contained in the EEG data in advance of the seizure onset, time series analysis techniques are supposed to detect such changes in the data. Visual inspection of the EEG data has not yet led to the detection of any characteristic changes preceding seizure onsets. Recently, it has been proposed that changes in the interactions between different brain areas measured by EEG time series contain information about upcoming seizures. Therefore, analysis techniques suitable to detect interactions in multivariate systems have come into the focus of research. Especially methods originating from synchronization theory promise new opportunities for epileptic seizure prediction.[7–9,21,28]

Application of time-resolved quantities measuring synchronization leads to feature time series. A threshold crossing of such a feature time series or a set of feature time series is often used for raising alarms. For the evaluation of seizure prediction methods, the electrode contacts obtaining highest sensitivity are of particular interest.

A proper assessment of seizure prediction performance motivated by the requirements for a successful therapeutic intervention is achieved by the methodology of the seizure prediction characteristic.[29,30] The seizure prediction characteristic evaluates sensitivity of a prediction method with respect to its specificity and with respect to temporal aspects of a prediction. For a real prediction, there has to be a time period between the prediction and the occurrence of a seizure, the seizure prediction horizon. This time period is necessary, e.g., for drug administration or the application of a seizure warning device. Moreover, a time interval is necessary during which the seizure is predicted to occur, the seizure occurrence period. Otherwise, the necessary duration of the period during which the intervention or warning has to continue is not precisely characterized. The temporal properties of a prediction characterized by these two time intervals must be related to the number of false alarms to ensure specificity of the seizure prediction method. Allowing too many false predictions would lead to impairment due to possible side-effects of interventions or loss of the patients' acceptance of seizure warning devices.

The number of EEG channels utilized by a seizure prediction method ranges from a few to more than a hundred. Using, for instance, symmetric and bivariate synchronization quantities between each pair of electrode contacts leads to an enormous number of possible combinations. Assuming that there is no predictive information in the EEG data and thus in the feature time series, the probability of predicting at least some of the seizure onsets correctly by chance is increasing with increasing number of electrode contacts. To assess the superiority of a prediction method over a random predictor, the same number of electrode contacts has to be considered. There are several possibilities for the choice of a random predictor. We present a completely uninformative random predictor based on alarms following a Poisson process. Utilizing this random predictor, a calculation of a critical sensitivity is possible that enables the decision whether the investigated seizure prediction method is superior to a random prediction. This critical value is adapted to the seizure prediction characteristic and allows to test the statistical significance of seizure prediction performance.

Using invasive EEG data of four representative patients suffering from epilepsy, we demonstrate the performance of a seizure prediction method based on a quantity measuring phase synchronization. This quantity, introduced as mean phase coherence, has been shown to detect changes in the EEG dynamics prior to seizure onsets.[7] Using the concept of the seizure prediction characteristic and the corresponding significance level developed in this article, we show that for two patients the mean phase coherence is superior to a random predictor, while for the remaining two patients the performance compared to a random predictor is questionable. This strengthens the hope that, at least for some patients, a reliable seizure prediction is possible utilizing methods developed within the framework of nonlinear dynamics such as quantities measuring phase synchronization.

The paper is structured as follows. In Sec. II, the concept of phase synchronization is summarized. The application of the mean phase coherence to invasive EEG data is given in Sec. III. The requirements for an appropriate assessment of seizure predictability is presented in Sec. IV. The unspecific random predictor and a test for the statistical significance are introduced in Secs. V and VI, respectively. Finally, the seizure prediction performance of the mean phase coherence is illustrated in Sec. VII utilizing the proposed methodology.

II. PHASE SYNCHRONIZATION OF OSCILLATORY SYSTEMS

Coupled self-sustained oscillators have been observed to be able to synchronize their oscillations. The phenomenon of synchronization has first been described by Huygens.[31] Since then different phenomena have been discovered, ranging from phase synchronization to almost complete synchronization.[31–34] The former cannot be observed for all oscillators in its strict sense, i.e., that there is a defined phase relation without any relation between the amplitudes. Phase synchronization has been observed for weakly coupled chaotic self-sustained oscillators.[35] The theory of phase synchronization has been generalized to stochastic oscillatory systems.[2]

To detect phase synchronization of coupled nonidentical chaotic oscillators with frequencies $\omega_{1,2}=1\pm\Delta\omega$, phase and amplitude of the real-valued signal have to be investigated. To this aim, several procedures have been proposed.[35–38] In the following, the definition based on Gabor's analytic signal representation[39]

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

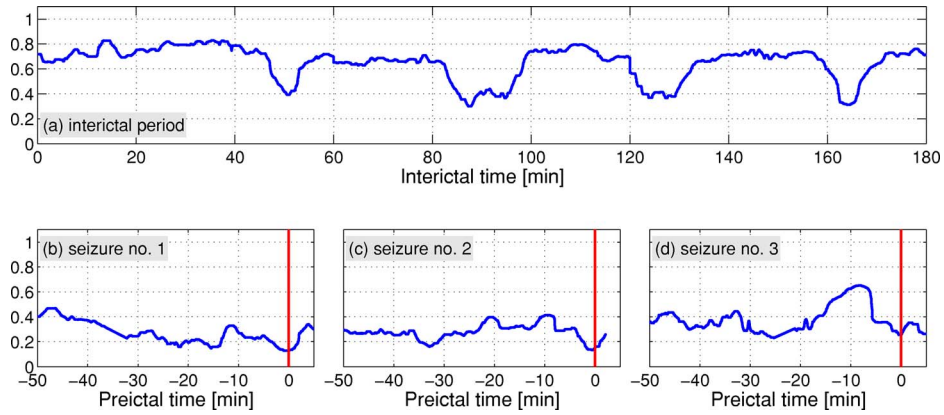

FIG. 1. Time course of the mean phase coherence applied to invasive EEG data recorded for one epilepsy patient. The feature values are calculated using two focal electrode contacts, applying a sliding window technique. A time course of the mean phase coherence during an interictal period of 3 h duration is shown in (a). The temporal evolution of the feature mean phase coherence is shown for three preictal periods of 50 min duration each in (b), (c), and (d). Seizure onsets are marked by the vertical lines at time points zero.

$$\psi(t) = x(t) + i\hat{x}(t), \tag{1}$$

expressed via its polar representation

$$\psi(t) = A(t)e^{i\Phi(t)}, \tag{2}$$

where $A(t)$ denotes amplitude and $\Phi(t)$ phase of the analytic signal, is applied. The imaginary counterpart of the real valued oscillatory signal $x(t)$ can be estimated using the Hilbert transform[40]

$$\hat{x}(s) = \frac{1}{\pi}\text{P.V.}\int x(t)\frac{1}{s-t}dt, \tag{3}$$

where P.V. refers to Cauchy's principal value. This procedure is reasonable for signals with a clearly defined frequency.[41]

Phase synchronization of coupled nonidentical self-sustained chaotic oscillators is realized by an almost constant phase difference guaranteed by the phase locking condition[35]

$$|n\Phi^{(1)}(t) - m\Phi^{(2)}(t)| = |\Phi_{n,m}(t)| < \text{const}, \tag{4}$$

where $\Phi^{(i)}(t)$ denotes the phase of time series i and $\Phi_{n,m}(t)$ the phase difference for given integers n and m. If the phase locking condition is fulfilled for two integers n and m, the processes are referred to as $n:m$ phase locked. In the presence of additional stochastic influence, phase jumps of $\pm 2\pi, \pm 4\pi, \ldots$ occur. Thus, applied to stochastic processes, the distribution of the phase differences

$$\Psi_{n,m}(t) = (n\Phi^{(1)}(t) - m\Phi^{(2)}(t)) \bmod 2\pi, \quad n,m \in \mathbb{Z}, \tag{5}$$

is significantly different from a uniform distribution.[2] The deviation of this distribution from a uniform one can be quantified by

$$R_{n,m}^2 = \langle \cos \Psi_{n,m}(t)\rangle^2 + \langle \sin \Psi_{n,m}(t)\rangle^2, \tag{6}$$

taking values close to zero if there is no deviation and values close to one for preferred values of the phase difference, respectively.[3,7]

III. APPLICATION TO INVASIVE EEG RECORDINGS

Applying the phase synchronization measure *mean phase coherence* $R := R_{1,1}$ [cf. Eq. (6)] to invasive EEG recordings of epilepsy patients, a preictal decrease in synchronization has been observed.[7–9] This decrease of the mean phase coherence is investigated in the following.

A. Invasive EEG recordings

Invasive EEG recordings of four epilepsy patients have been analyzed, which were recorded during presurgical monitoring at the Epilepsy Center of the University Hospital of Freiburg, Germany. The EEG data were sampled at 256 Hz. To prevent aliasing and to eliminate possible line noise and low frequency components, the EEG data were preprocessed by a 50 Hz notch filter and a band pass filter between 0.5 Hz and 120 Hz.

In order to limit the probability of random unspecific threshold crossings related to the number of comparisons between electrode contacts, six electrode contacts have been analyzed. Three contacts were early involved in ictal activity (focal electrode contacts), whereas the remaining three contacts were not involved in ictal activity or only involved late during seizure spread (extra-focal electrode contacts). For all patients, 24 h interictal data and five seizures, each including a preseizure period of 50 min duration, have been examined.

B. Phase synchronization: interictal versus preictal

The mean phase coherence has been applied to the EEG data using a sliding window technique. To avoid the detection of very short-lived changes in the temporal evolution of the feature, a median filter has been applied.

An exemplary time course of the mean phase coherence calculated using two focal electrode contacts is shown in Fig. 1(a) for an interictal period of 3 h duration. Most of the time, the feature fluctuates between 0.6 and 0.8, with distinct drops at approximately 50, 80, 120, and 160 min. In Figs. 1(b)–1(d), the temporal evolution of the feature mean phase coherence is shown for three preictal periods of 50 min duration each. Seizure onsets are marked by the vertical lines at time points zero. A decrease of the mean phase coherence is observed during the preictal periods.

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

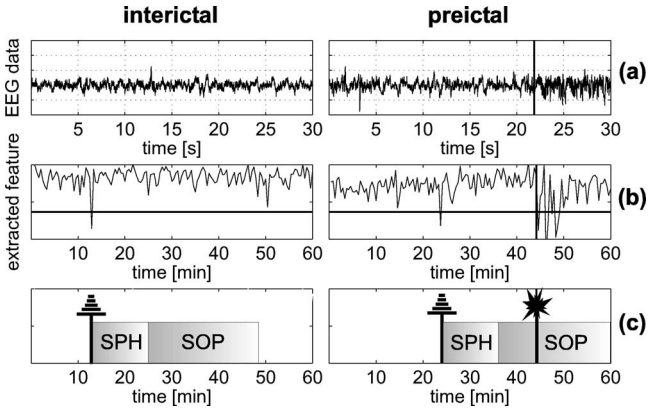

FIG. 2. Basic operation of a prediction method during an interictal and a preictal period. Seizure onset is marked by vertical lines. (a) Examples of EEG recordings and (b) exemplary time course of a feature extracted by a seizure prediction algorithm. The solid, horizontal line indicates the threshold for raising alarms. Alarm events and two consecutive time intervals characterizing a prediction, the seizure prediction horizon SPH and seizure occurrence period SOP are illustrated in (c). Note the different time scales for the EEG data and the feature time series.

The example demonstrates that during preictal periods, the time course of mean phase coherence is different compared to interictal periods far away from a seizure. It remains to be investigated whether these changes in the temporal evolution of the mean phase coherence can be utilized to enable seizure prediction with the accuracy required for the application in therapeutic devices.

IV. ASSESSMENT OF SEIZURE PREDICTABILITY

In the following, a general methodology is presented to evaluate the performance of time series analysis techniques with respect to seizure prediction. In general, potential candidates for seizure prediction work in a similar way (cf. Fig. 2). Time-resolved time series analysis techniques are applied to EEG data recorded from different electrode contacts [Fig. 2(a)]. Values of the time-resolved techniques lead to feature time series [Fig. 2(b)]. The mean phase coherence applied to the EEG data using a sliding window technique is an example for a feature time series.

A threshold crossing of the feature is used to raise an alarm in order to predict an impending seizure [Figs. 2(b) and 2(c)]. The optimal threshold is determined retrospectively as explained below. The number of seizures preceded by a threshold crossing divided by the total number of seizures investigated yields the sensitivity of the prediction method. This calculation can be performed patient individually or averaged over groups of patients.

For a proper prediction, the timing between an alarm before a seizure is the deciding factor. First, a time interval after an alarm, the seizure prediction horizon SPH [Fig. 2(c)], is required to successfully apply therapeutic interventions or seizure warning devices. Within this time interval, drugs or other treatment strategies can be administered or the patient takes care of behavioral adjustments.

Second, subsequent to the prediction horizon, the seizure should start. Due to stochasticity of the recorded time series, it is extremely rare that the SPH is immediately followed by a seizure onset. To avoid a too long warning period, in which the patient may sustain severe stress or physiological disadvantages caused by potential side effects of drugs or long-term electric stimulation of focal brain structures, a second time interval has to be fixed. In this time interval, introduced as seizure occurrence period SOP, the seizure has to start if the prediction is to be classified as correct [Fig. 2(c)].

Both time intervals have to be taken into consideration to decide whether or not a seizure is predicted correctly. If there is no seizure onset during a seizure occurrence period, the alarm has to be classified as a false prediction. During interictal periods, i.e., periods far away from any seizure, all alarms lead to false predictions. This case is illustrated in Fig. 2 in the left-hand column. The number of false predictions has to be controlled to assess prediction performance.[42] Too many false predictions cannot be accepted, e.g., for a seizure warning device due to the loss of the patients' acceptance of the device. For automatic therapeutic devices, the number of false predictions is limited by the number of unnecessary but acceptable interventions. An appropriate criterion for the maximum number of false predictions is the maximum false prediction rate ($\mathrm{FPR}_{\max}$) which limits the tolerated number of false predictions in a given time period.

For a given data set, the false prediction rate as well as the number of correct predictions is determined by the seizure prediction horizon, the seizure occurrence period, and the chosen threshold. If, for instance, the threshold in the example of Fig. 2 would be decreased, the false alarm could be avoided (left-hand column) but only at the expense of losing a correctly predicted seizure (right-hand column).

Therefore, the following procedure is proposed: First the maximum false prediction rate should be fixed together with the seizure prediction horizon and the seizure occurrence period. This leads to a value for the threshold optimized patient individually using the time course of the feature during interictal periods. Afterwards, the sensitivity of the prediction method is estimated as the fraction of preictal periods during which the threshold is crossed within an SOP.

Sensitivity depending on the set of chosen prediction parameters, i.e., the maximum false prediction rate $\mathrm{FPR}_{\max}$, the seizure prediction horizon SPH, and the seizure occurrence period SOP, yields the seizure prediction characteristic[29,30]

$$S(\mathrm{FPR}_{\max}, \mathrm{SOP}, \mathrm{SPH}). \tag{7}$$

Its functional relationship is not known in advance, but can be estimated on the basis of several values for SPH, SOP, and $\mathrm{FPR}_{\max}$. The seizure prediction characteristic provides an efficient way to obtain information about the sensitivity of seizure prediction methods and thus a comparison between different prediction methods. By means of the seizure prediction characteristic, it has been shown that sensitivities for the prediction of seizures are in general lower than sensitivities calculated without taking into account the aforementioned temporal aspects of a prediction.[10,22,23,29,43]

A problem that has so far remained unsolved is a rigorous way to decide the statistical significance of the seizure prediction characteristic or the significance of a difference between seizure prediction characteristics. This problem is

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

addressed in the following for multivariate seizure prediction methods. Basically, any prediction method should be superior to an unspecific random predictor that makes no use of any information contained in the EEG data.

V. THE RANDOM PREDICTOR

The interval between two consecutive alarms of the unspecific random predictor used is exponentially distributed resulting in a Poisson process. The probability of raising an alarm in a given time interval of width h is

$$\text{prob}\{\text{alarm in } [t,t+h] | \text{given the history}\} = \mu h + o(h), \tag{8}$$

where μ is the intensity of the Poisson process.[44] In the case of a sampled process, h is set to the width of one sampling bin. The terms of smaller order $o(h)$ are neglected in the following and μh is abbreviated to P_{Poisson}. The number of false predictions (FP) fixes the parameter

$$P_{\text{Poisson}} = \frac{\text{FP}}{N}, \tag{9}$$

which is the probability of raising an alarm at any single sampling point of the time series with N samples. The probability to raise an alarm in a period of duration equal to SOP is

$$\text{prob}_{\Delta t|\Delta t=\text{SOP}} = 1 - (1 - P_{\text{Poisson}})^{N(\Delta t)}|_{\Delta t=\text{SOP}}, \tag{10}$$

with $N(\Delta t)$ being the number of bins in the period Δt. This probability can be approximated by[30]

$$P = \text{prob}_{\Delta t|\Delta t=\text{SOP}} \approx 1 - e^{-\text{FPR}_{\max}\cdot\text{SOP}} \approx \text{FPR}_{\max}\cdot\text{SOP} \tag{11}$$

if the product of $\text{FPR}_{\max}\cdot\text{SOP}$ is considerably smaller than one. Only the duration of SOP is contained in equation for the probability P [Eq. (11)]. The seizure prediction horizon, while being essential for the concept of prediction, has no influence on the sensitivity of the random predictor.

VI. TESTING STATISTICAL SIGNIFICANCE OF PREDICTION PERFORMANCE

To decide about the statistical significance of sensitivity values, tests based on constructing seizure time surrogates or measure profile surrogates have recently been introduced.[45,46]

Here, we introduce an alternative approach, based on the comparison of sensitivities of prediction methods with the performance of an unspecific random prediction, utilizing no information contained in EEG recordings. The exact distribution of the test statistic is derived. The probability of at least one alarm raised by an unspecific random prediction during a seizure occurrence period SOP has been introduced in the previous section with $P = 1 - e^{-\text{FPR}_{\max}\cdot\text{SOP}} \approx \text{FPR}_{\max}\cdot\text{SOP}$ for a given value of the maximum false prediction rate $\text{FPR}_{\max}$ [cf. Eq. (11)]. A critical value to test whether the sensitivity $S(\text{FPR}_{\max}, \text{SOP}, \text{SPH})$ of a prediction method is higher than the sensitivity of an unspecific random prediction should take into account that usually several electrode contacts are analyzed. Furthermore, more than one seizure is usually investigated, which is necessary to determine sensitivity reliably. The probability of predicting k out of K present seizures by chance follows a binomial distribution with probability P. Critical values for the sensitivity of prediction methods including both extensions are derived analytically in the following.

A binomial distribution with probability P from Eq. (11) accounts for K independent seizures investigated. The probability of predicting at least k of K seizures by an unspecific random prediction is given by

$$P_{\text{binom}}\{k;K;P\} = \sum_{j\geq k} \binom{K}{j} P^j (1-P)^{K-j}. \tag{12}$$

For a rather high value $P=0.5$, for instance, the probability of predicting at least one of five seizures just by chance is 97%, whereas the probability of predicting all five seizures by chance is approximately 3%.

For a prediction method the features extracted from EEG electrode contacts rather than the EEG data themselves are of particular interest. As the unspecific random prediction utilizes no EEG data, only the number of independent extracted features influences the critical value. For example, if n features are extracted from n independent electrode contacts, corresponding to an univariate feature extraction preserving independence, the dimension d of the extracted feature vector is equal to the number of investigated electrode contacts n. In general, the maximum number of independent features for a r variate, symmetric feature extraction using n electrode contacts is given by

$$d_{\max,r}(n) = \binom{n}{r}. \tag{13}$$

The effective value $d_{\text{eff}}(n)$ of d might be smaller than $d_{\max,r}$, as signals from neighboring electrode contacts may be correlated, leading to a lower effective dimension of the extracted feature vector. However, the assumption of independent extracted features has been empirically verified, for example, by means of cross-correlation analysis for two measures of synchronization.[46]

The dimension d of the extracted feature vector leads to a correction of the probability, taken into account by

$$P_{\text{binom},d}\{k;K;P\} = 1 - \left(\sum_{j<k} \binom{K}{j} P^j (1-P)^{K-j} \right)^d. \tag{14}$$

This is the probability of predicting at least k of K seizures by means of at least one of d independent features correctly and is used in the following as a test statistic.

For a given significance level α, the sensitivity of an unspecific random prediction, based on d independent extracted features and K seizures, is given by the critical value $\sigma_{\text{rand},d} = \max_k(P_{\text{binom},d}\{k;K;P\} > \alpha) \times 100\%$. As the underlying dependence structure of the features is unknown, upper and lower critical values are calculated. The lower critical value

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

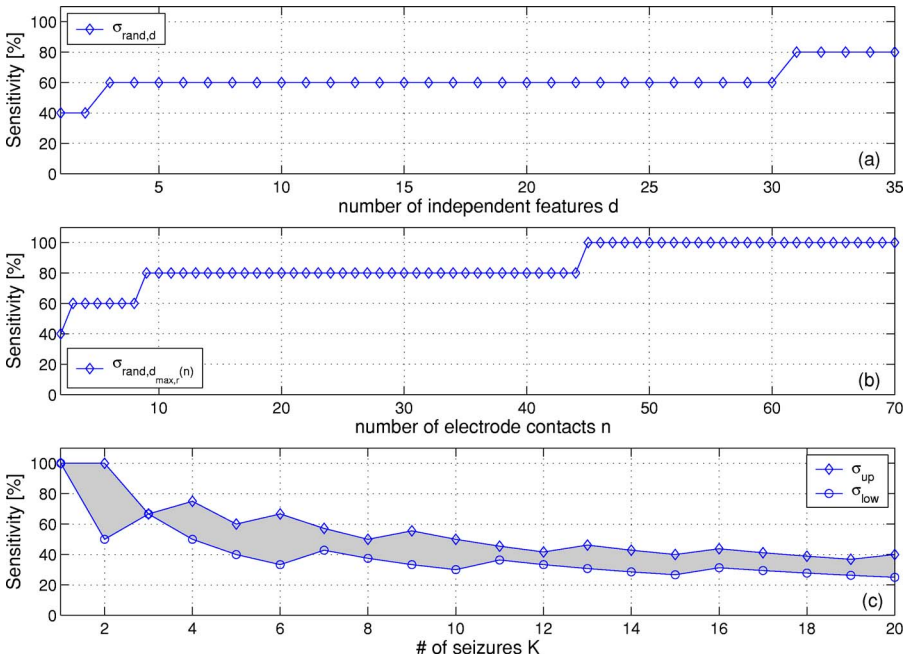

FIG. 3. Dependence of $\sigma_{\mathrm{rand},d}$ on the feature vector dimension d for $K=5$ seizures (a) and dependence of $\sigma_{\mathrm{rand},d_{\mathrm{max},2}(n)}$ on the number of investigated electrode contacts n for $K=5$ seizures (b). Lower critical value σ_{low} and upper critical value σ_{up} of a random prediction as a function of the number of seizures K for a maximum number of independent features $d_{\mathrm{max},2}=15$ (c). Local, positive slopes of σ_{low} and σ_{up}, respectively, in (c) are caused by the discretization.

$$\sigma_{\mathrm{low}} = \sigma_{\mathrm{rand},1} = \max_k (P_{\mathrm{binom},1}\{k;K;P\} > \alpha) \times 100\% \tag{15}$$

is given for a one-dimensional feature vector ($d=1$). For the upper critical value, complete independence of the features is assumed, leading to

$$\sigma_{\mathrm{up}} = \sigma_{\mathrm{rand},d_{\mathrm{max},r}} = \max_k (P_{\mathrm{binom},d_{\mathrm{max},r}}\{k;K;P\} > \alpha) \times 100\,\% \,. \tag{16}$$

In Fig. 3 sensitivity values of the unspecific random prediction $\sigma_{\mathrm{rand},d}$ depending on the number of independent features d [Fig. 3(a)], $\sigma_{\mathrm{rand},d_{\mathrm{max},2}(n)}$ depending on the number of investigated electrode contacts n [Fig. 3(b)] as well as dependence of σ_{up} and σ_{low} on the number of seizures K [Fig. 3(c)] are shown based on a 5% significance level ($\alpha=0.05$). The number of seizures has been fixed to $K=5$ in Figs. 3(a) and 3(b). In Fig. 3(c) the number of electrode contacts ($n=6$) and a bivariate feature extraction ($r=2$) leads to a maximum feature dimension $d_{\mathrm{max},2}=15$. The seizure occurrence period has been chosen to be half an hour and the maximum false prediction rate has been set to 0.3 false predictions per hour (FP/h), corresponding to a maximum number of eight false predictions per day.

When $K=5$ seizures were investigated and $d=35$, sensitivity of the investigated prediction method must obtain a value of at least 81% to achieve a significant result, i.e., all seizures must be predicted correctly in this example [Fig. 3(a)]. The upper critical value for the unspecific random prediction increases with the number of independent features. This demonstrates the problem that high values of sensitivity are possible even for an unspecific random prediction when using high dimensional feature vectors. For example, assuming independence of the extracted feature time series, $d_{\mathrm{max},2}(n)$ increases rapidly with increasing n. Thus, analyzing more than eight electrode contacts for $K=5$ seizures leads to a critical sensitivity value of 80% (b). Even worse, it is impossible to obtain any significant result if more than 44 electrode contacts were analyzed. The critical sensitivity value of the random predictor achieves 100% in these cases. High critical values are also obtained if only very few seizure are analyzed [Fig. 3(c)]. Increasing the number of investigated seizures leads to a considerably decrease in the critical value of the random prediction. For the choice of d, $\mathrm{FPR}_{\mathrm{max}}$, and SOP in this example, this decrease becomes weaker if more than ten seizures were investigated.

When information about the interdependencies between the features is not available, every prediction method should achieve sensitivity values exceeding at least σ_{low} to be considered useful. If sensitivity exceeds σ_{up}, its superior performance compared to an unspecific random prediction is indicated. Values of the seizure prediction characteristic that fall into the range between σ_{low} and σ_{up} cannot be regarded as definitely superior to a random prediction. Comparing the functional relationships shown in Fig. 3, the dependence of the critical value on d [Fig. 3(a)] is rather flat compared to the dependence of the critical value on the number of seizures [Fig. 3(c)].

In summary, if the performance of any seizure prediction method is compared to the random predictor, the results suggest that one should, first, try to obtain long data sets with several seizures. Second, the number of investigated electrode contacts n and thus the number of independent features should be as small as possible. This procedure guarantees that a superiority compared to a random predictor is in general verifiable for an appropriately chosen value of P [Eq. (11)].

Since the proposed significance level is a point-wise level, it is typically exceeded for a certain number of values for $\mathrm{FPR}_{\mathrm{max}}$, SPH, or SOP, even when the obtained sensitivity is not significantly higher than the random predictor. Single crossings of the critical values should therefore not be regarded as statistically significant. We would like to point out

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

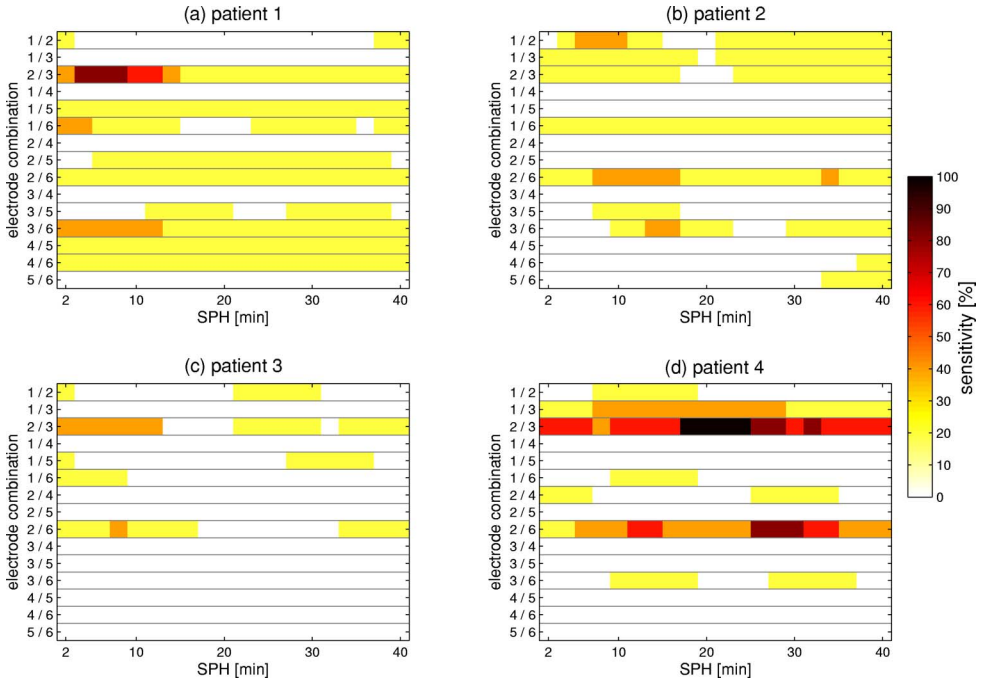

FIG. 4. Seizure prediction characteristic of the mean phase coherence for patients (a) 1, (b) 2, (c) 3, and (d) 4, and all 15 possible combinations between the six electrode contacts. Sensitivity is given as a function of the seizure prediction horizon SPH. The maximum false prediction rate has been set to $FPR_{max}=0.15$ FP/h and the seizure occurrence period to SOP=10 min.

that it is rather common in other applications to use point-wise levels.

VII. SEIZURE PREDICTION PERFORMANCE OF THE MEAN PHASE COHERENCE

For the mean phase coherence, values of the seizure prediction characteristic are shown in Fig. 4 for all four patients under investigation and for all 15 possible combinations between the six electrode contacts. The seizure prediction characteristic is given as a function of the seizure prediction horizon, evaluated in a range between 2 and 40 min. For this analysis, the maximum false prediction rate has been set to $FPR_{max}=0.15$ FP/h, corresponding to a maximum number of false predictions of 3.6 per day. The seizure occurrence period has been set to 10 min.

For all patients investigated, no seizure prediction is possible for several electrode contact combinations. But for patient 1 and for the combination between the focal contact nos. 2 and 3 [Fig. 4(a)], a sensitivity of up to 80% is obtained. For patients 2 and 3, a maximum sensitivity of 40% is achieved [Figs. 4(b) and 4(c)], whereas patient 4 shows the highest performance with 100% sensitivity between electrode contact nos. 2 and 3 [Fig. 4(d)]. For patient 4 a sensitivity of 80% is also possible for one combination between a focal and a extra-focal contact, i.e., contact nos. 2 and 6 [Fig. 4(d)].

To decide on the statistical significance of sensitivity values, the proposed comparison with the performance of an unspecific random prediction has been performed. Values of the seizure prediction characteristic are shown for the mean phase coherence and for all four patients in Figs. 5–7. The corresponding range of an unspecific random prediction, limited by the upper and lower critical values σ_{up} and σ_{low}, is given by the gray areas. In contrast to Fig. 4, only the sensitivity value of the electrode combination with highest sensitivity value is shown.

In Fig. 5, sensitivities are shown as function of the seizure prediction horizon SPH. A maximum false prediction rate of $FPR_{max}=0.15$ FP/h and a seizure occurrence period of 10 min has been chosen for this analysis. For patient 1, significant prediction sensitivity is observed with respect to the upper critical value for small seizure prediction horizons [Fig. 5(a)]. In contrast, for patient 2 and 3 sensitivity values are not significant with respect to the upper critical value [Figs. 5(b) and 5(c)]. In those cases, where sensitivity values are below the upper critical value, the possibility that the prediction of seizures is achieved just by chance cannot be excluded. Patient 4 shows an outstanding performance with only one sensitivity value at SPH=8 min that is not above the upper critical value [Fig. 5(d)]. Sensitivities up to 100% are observed in the range between SPH=18 min and SPH =24 min.

To illustrate the dependence on the uncertainty in the prediction times, sensitivities are shown as function of the seizure occurrence period SOP in Fig. 6. For this analysis, the maximum false prediction rate has been set to FPR_{max} =0.15 FP/h and the seizure prediction horizon to 10 min. For patient 1, sensitivity exceeds the upper critical values considerably for a wide range of SOP$>$2 min [Fig. 6(a)]. For patient 2, sensitivity exceeds the upper critical value only for SOP=6 min [Fig. 6(b)]. Since a point-wise significance level is used, this result cannot be assumed to be statistically significant. The seizure prediction characteristic shows a slightly better performance for patient, which however cannot be considered superior to the upper critical sensitivity value for approximately one half of the SOP values [Fig. 6(c)]. For patient 4, sensitivity exceeds the upper critical values considerably again for a wide range of SOP $>$8 min [Fig. 6(d)]. Sensitivity values of 100% are achieved, demonstrating the outstanding prediction performance observed for this patient.

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

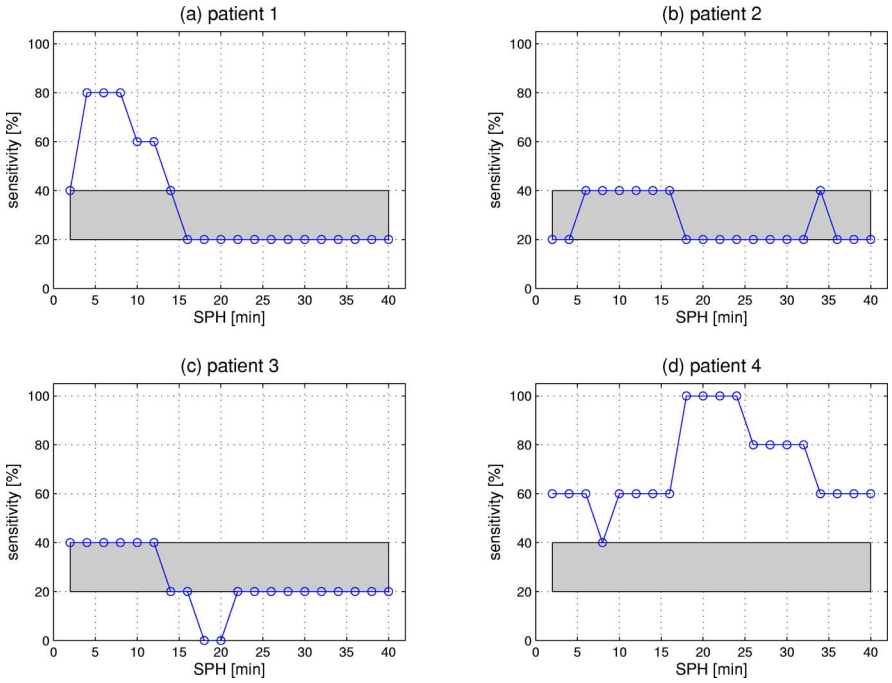

FIG. 5. Sensitivity of the mean phase coherence depending on SPH for the four patients (FPR_{max}=0.15 FP/h and SOP=10 min). The gray areas mark the corresponding ranges of the unspecific random prediction between the lower and upper critical values. (a) For patient 1 a significant seizure prediction performance is observed in the range between 4 and 12 min. (b), (c) The prediction performance for patients 2 and 3 is not significant compared to the random predictor. In contrast to patient 1, patient 4 shows a high overall performance with sensitivity up to 100% for SPH between 18 and 24 min (d).

The seizure occurrence period has been set to half an hour and the seizure prediction horizon to 10 min for the investigation of sensitivity depending on the maximum false prediction rate FPR_{max} (Fig. 7). The prediction performance depending on the maximum false prediction rate achieves significant results in a range from FPR_{max}=4 FP/d to FPR_{max}=0.8 FP/h for patient 1 [Fig. 7(a)]. Sensitivities up to 100% are possible. For patient 2, a significant prediction is only observed with respect to the lower critical value and for a few values of the maximum false prediction rate [Fig. 7(b)]. Patient 3 shows again a superior performance compared to patient 2 [Fig. 7(c)]. However, sensitivity values of 100% are only observed scarcely and only for rather high values of the maximum false prediction rate. For patient 4, an outstanding performance is again observed [Fig. 7(d)]. Values of the seizure prediction characteristic are higher than σ_{up} for maximum false prediction rates lower than 0.8 FP/h, as the upper critical sensitivity value reaches 100% for $FPR_{max} > 0.8$ FP/h. All five seizures are predicted correctly for the patient even for low maximum false prediction rates.

The analysis emphasizes the importance of a significance level to judge the performance of seizure prediction methods. For example, the prediction performance for patient 2 achieves values of 60% for reasonable values of SPH, SOP, and FPR_{max}. Values of 60% seem to be promising at first glance. But the critical value also obtains values of 60%

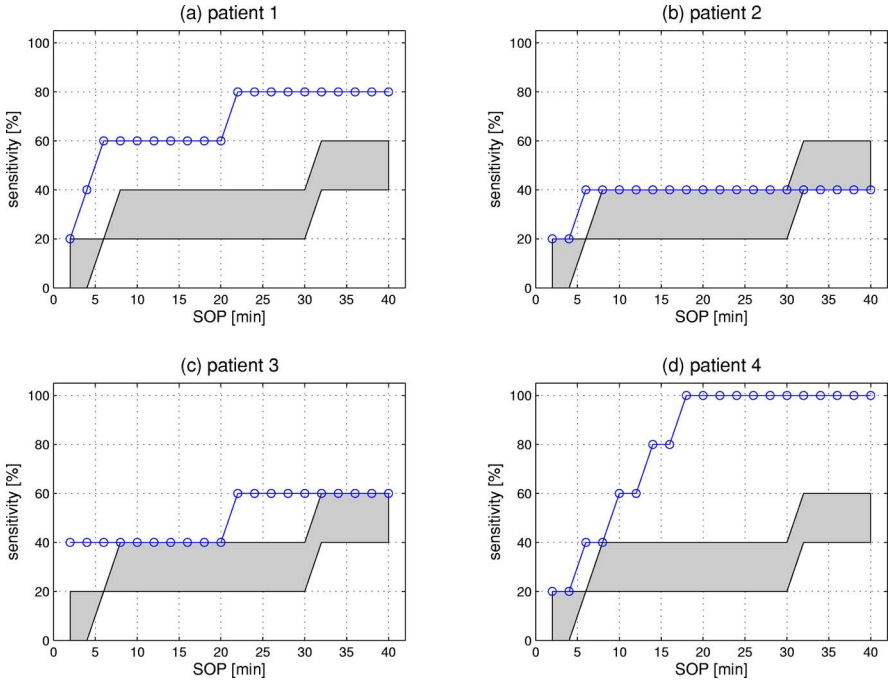

FIG. 6. Sensitivity of the mean phase coherence depending on SOP for the four patients (FPR_{max}=0.15 FP/h and SPH=10 min). The gray areas mark the corresponding ranges of the unspecific random prediction between the lower and upper critical values. Again patients 1 and 4 are characterized by the best prediction performance (a), (d). For almost all SOP values the prediction performance is superior to the random predictor. Patient 2 shows hardly any significant prediction performance compared to the upper critical sensitivity value (b), while for patient 3 results superior to the upper critical sensitivity values are found for distinct SOP values (c).

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

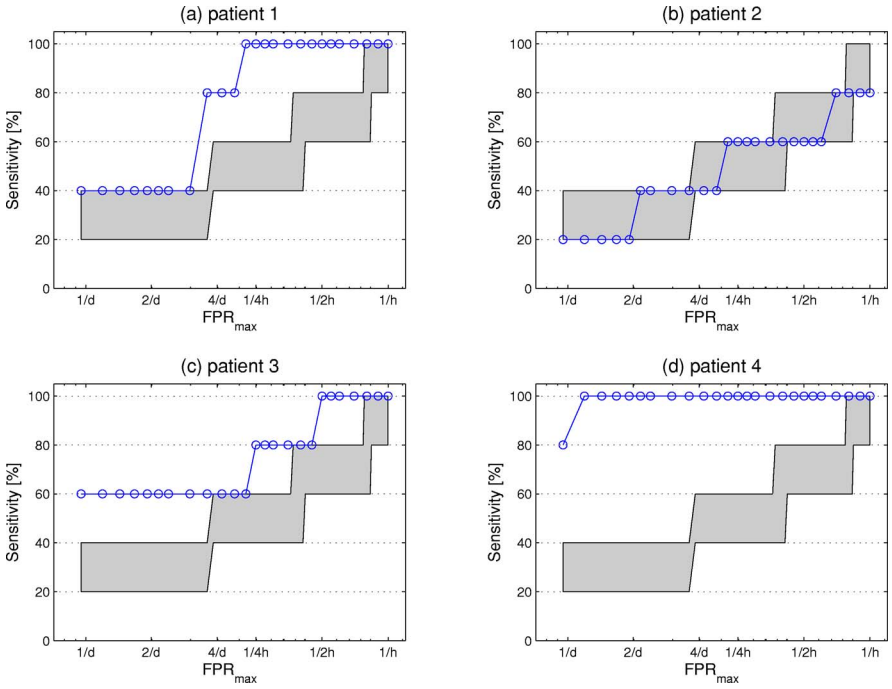

FIG. 7. Sensitivity of the mean phase coherence depending on FPR_{max} for the four patients (SOP=30 min and SPH=10 min). The gray areas mark the corresponding ranges of the unspecific random prediction between the lower and upper critical values. (a) For patient 1 superior performance is detected over a rather broad range of FPR_{max} values, while for patient 2 there is not one significant results compared to the upper critical value (b). Again results for patient 3 are superior to the results obtained for patient 2 (c). An outstanding performance is again observed for patient 4, where the sensitivity values are almost always 100% even for low false prediction rates (d).

for these parameter choices indicating that it is a least doubtful whether or not this performance is superior to a random predictor.

VIII. NONLINEAR DYNAMICS AND SEIZURE PREDICTION

Methods from nonlinear dynamics could contribute to the development of new therapeutic strategies for epilepsy. Nevertheless, it is still under debate whether time series analysis techniques based on the theory of nonlinear dynamics show a predictive power with respect to epileptic seizure prediction.[28,47–49] For instance, by means of simulations and controlled tests, it has been investigated whether prerequisites for applying the concept of Lyapunov exponents to EEG data are given; with the result that application of Lyapunov exponents for seizure prediction is questionable.[28] In another study, it has been shown that basic conditions for phase synchronization analysis are not guaranteed when applied to EEG data.[50]

From a theoretical point of view, it is necessary that basic conditions of time series analysis techniques are fulfilled, if these methods were applied to empirical time series. Such prerequisites are, for instance, stationarity or ergodicity of the underlying processes. Nevertheless, from a practical point of view, time series analysis techniques have to fulfill different requirements to be applicable in therapeutic devices for epilepsy patients. The major aim is to predict seizures with high sensitivity and specificity. For this purpose, the violation of strict prerequisites of the time series technique applied can be accepted.

Taking this as a basis we have presented a methodological framework to assess and statistically validate seizure prediction performances of time series analysis techniques originating from nonlinear dynamics. The seizure prediction characteristic is motivated by the basic requirements for a successful application of prediction methods in seizure warning or therapeutic devices. In combination with its analytically derived critical values, a statistical assessment of patient individual seizure prediction performance is possible. The proposed methodology complements recently introduced strategies to judge the significance of prediction methods by applying seizure time and measure profile surrogates.[45,46] Our approach is derived on the basis of theoretical considerations adjusted to the seizure prediction characteristic which has been proposed to estimate sensitivities of seizure prediction methods while taking seizure prediction horizons, seizure occurrence periods, and false prediction rates into account.

In summary, misinterpretation of the prediction performance of patient individual optimized seizure prediction methods is prevented by the critical sensitivity values derived in this article. For the assessment of prediction performance for an individual patient, we suggest σ_{low} as a critical value that has to be exceeded by any prediction method to be considered useful. Moreover, a sensitivity exceeding σ_{up} can be regarded as a prediction performance that is reliably better than an unspecific random prediction.

ACKNOWLEDGMENTS

This work was supported by the German Science Foundation (DFG Grant No. Ti315/2-1) and the German Federal Ministry of Education and Research (BMBF Grant No. 01GQ0420).

[1] S. Boccaletti, J. Kurths, G. Osipov, D. Valladares, and C. Zhou, Phys. Rep. **366**, 1 (2002).
[2] P. Tass, M. G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H. J. Freund, Phys. Rev. Lett. **81**, 3291 (1998).
[3] M. Rosenblum, A. Pikovsky, J. Kurths, C. Schäfer, and P. A. Tass, in *Handbook of Biological Physics*, Neuro-informatics Vol. 4, edited by F. Moss and S. Gielen (Elsevier, Amsterdam, 2001), pp. 279–321.
[4] M. G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, and R.

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

Mrowka, Phys. Rev. E **65**, 041909 (2002).
[5]F. Varela, J. Lachaux, E. Rodriguez, and J. Martinerie, Nat. Rev. Neurosci. **2**, 229 (2001).
[6]P. A. Tass, T. Fieseler, J. Dammers, K. Dolan, P. Morosan, M. Majtanik, F. Boers, A. Muren, K. Zilles, and G. R. Fink, Phys. Rev. Lett. **90**, 088101 (2003).
[7]F. Mormann, K. Lehnertz, P. David, and C. Elger, Physica D **144**, 358 (2000).
[8]F. Mormann, R. G. Andrzejak, T. Kreuz, C. Rieke, P. David, C. E. Elger, and K. Lehnertz, Phys. Rev. E **67**, 021912 (2003).
[9]F. Mormann, T. Kreuz, R. Andrzejak, P. David, K. Lehnertz, and C. Elger, Epilepsy Res. **53**, 173 (2003).
[10]K. Lehnertz and C. E. Elger, Phys. Rev. Lett. **80**, 5019 (1998).
[11]J. Kurths, A. Voss, A. Witt, P. Saparin, H. Kleiner, and N. Wessel, Chaos **5**, 88 (1995).
[12]J. S. Ebersole, Clin. Neurophysiol. **116**, 489 (2005).
[13]B. Litt and K. Lehnertz, Curr. Opin. Neurol. **15**, 173 (2002).
[14]B. Litt and J. Echauz, Lancet Neurol. **1**, 22 (2002).
[15]J. Milton and P. Jung, in *Epilepsy as a Dynamic Disease*, edited by J. Milton and P. Jung (Springer, New York, 2003), pp. 341–352.
[16]M. Nicolelis, Nature (London) **409**, 403 (2001).
[17]B. Gluckman, H. Nguyen, S. Weinstein, and S. Schiff, J. Neurosci. **21**, 590 (2001).
[18]L. Iasemidis, J. Sackellares, H. Zaveri, and W. Williams, Brain Topogr **2**, 187 (1990).
[19]L. Iasemidis, D.-S. Shiau, P. Pardalos, W. Chaovalitwongse, K. Narayanan, A. Prasad, K. Tsakalis, P. Carney, and J. Sackellares, Clin. Neurophysiol. **116**, 532 (2005).
[20]F. Mormann, T. Kreuz, C. Rieke, R. Andrzejak, A. Kraskov, P. David, C. Elger, and K. Lehnertz, Clin. Neurophysiol. **116**, 569 (2005).
[21]M. Le van Quyen, J. Soss, V. Navarro, R. Robertson, M. Chavez, M. Baulac, and J. Martinerie, Clin. Neurophysiol. **116**, 559 (2005).
[22]M. Le van Quyen, J. Martinerie, M. Baulac, and F. Varela, NeuroReport **10**, 2149 (1999).
[23]B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, M. Dichter *et al.*, Neuron **30**, 51 (2001).
[24]V. Navarro, J. Martinerie, M. Le van Quyen, S. Clemenceau, C. Adam, M. Baulac, and F. Varela, Brain **125**, 640 (2002).
[25]M. Le van Quyen, J. Martinerie, V. Navarro, P. Boon, M. D'Have, C. Adam, B. Renault, F. Varela, and M. Baulac, Lancet **357**, 183 (2001).
[26]M. Le van Quyen, J. Martinerie, V. Navarro, M. Baulac, and F. Varela, Clin. Neurophysiol. **18**, 191 (2001).
[27]M. A. F. Harrison, M. G. Frei, and I. Osorio, Clin. Neurophysiol. **116**, 527 (2005).
[28]Y.-C. Lai, M. A. F. Harrison, M. G. Frei, and I. Osorio, Chaos **14**, 630 (2004).
[29]T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. Voss, A. Schulze-Bonhage, and J. Timmer, Physica D **194**, 357 (2004).
[30]M. Winterhalder, T. Maiwald, H. U. Voss, R. Aschenbrenner-Scheibe, J. Timmer, and A. Schulze-Bonhage, Epilepsy Behav. **4**, 318 (2003).
[31]A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization—A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001).
[32]L. Pecora and T. Carroll, Phys. Rev. Lett. **64**, 821 (1990).
[33]L. Pecora, T. Carroll, G. Johnson, D. Mar, and J. Heagy, Chaos **7**, 520 (1997).
[34]M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, Phys. Rev. Lett. **78**, 4193 (1997).
[35]M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, Phys. Rev. Lett. **76**, 1804 (1996).
[36]J.-P. Lachaux, E. Rodriguez, M. Le van Quyen, A. Lutz, J. Martinerie, and F. Varela, Int. J. Bifurcation Chaos Appl. Sci. Eng. **10**, 2429 (2000).
[37]D. J. DeShazer, R. Breban, E. Ott, and R. Roy, Phys. Rev. Lett. **87**, 044101 (2001).
[38]A. G. Rossberg, K. Bartholome, and J. Timmer, Phys. Rev. E **69**, 016216 (2004).
[39]D. Gabor, J. IEE London **93**, 429 (1946).
[40]A. Oppenheim and R. Schafer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1975).
[41]B. Boashash, Proc. IEEE **80**, 520 (1992).
[42]I. Osorio, M. Frei, and S. Wilkinson, Epilepsia **39**, 615 (1998).
[43]R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. Voss, J. Timmer, and A. Schulze-Bonhage, Brain **126**, 2616 (2003).
[44]D. Cox, *Renewal Theory* (Chapman and Hall, London, 1962).
[45]R. G. Andrzejak, F. Mormann, T. Kreuz, C. Rieke, C. E. Elger, and K. Lehnertz, Phys. Rev. E **67**, 010901(R) (2003).
[46]T. Kreuz, R. G. Andrzejak, F. Mormann, A. Kraskov, H. Stögbauer, C. E. Elger, K. Lehnertz, and P. Grassberger, Phys. Rev. E **69**, 061915 (2004).
[47]Y.-C. Lai, M. A. F. Harrison, M. G. Frei, and I. Osorio, Phys. Rev. Lett. **91**, 068102 (2003).
[48]Y.-C. Lai, M. A. F. Harrison, M. G. Frei, and I. Osorio, Phys. Rev. Lett. **94**, 019802 (2005).
[49]L. D. Iasemidis, K. Tsakalis, J. C. Sackellares, and P. M. Pardalos, Phys. Rev. Lett. **94**, 019801 (2005).
[50]M. Winterhalder, B. Schelter, J. Kurths, A. Schulze-Bonhage, and J. Timmer, Phys. Lett. A (submitted).

Downloaded 26 Jan 2006 to 132.230.70.81. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp